┃ GAFA Geometric And Functional Analysis

ANDERSON LOCALIZATION FOR TWO INTERACTING QUASIPERIODIC PARTICLES

JEAN BOURGAIN · ILYA KACHKOVSKIY

Abstract. We consider a system of two discrete quasiperiodic 1D particles as an operator on $\ell^2(\mathbb{Z}^2)$ and establish Anderson localization at large disorder, assuming the potential has no cosine-type symmetries. In the presence of symmetries, we show localization outside of a neighborhood of finitely many energies. One can also add a deterministic background potential of low complexity, which includes periodic backgrounds and finite range interaction potentials. Such background potentials can only take finitely many values, and the excluded energies in the symmetric case are associated to those values.

1 Introduction

In this paper, we study Anderson localization for the following family of Schrödinger operators on $\ell^2(\mathbb{Z}^2)$:

$$H(\theta_1, \theta_2) = \Delta + \lambda(v(n_1\omega + \theta_1) + v(n_2\omega + \theta_2)) + U(n_1, n_2). \qquad (1.1)$$

Here Δ is the discrete Laplacian, $v \in C^\omega(\mathbb{T})$ is a real analytic function (identified with a 1-periodic analytic function on $\mathbb{R}$), ω is an irrational number, and $\theta = (\theta_1, \theta_2) \in \mathbb{T}^2$ is the quasiperiodic phase. The case $U = 0$ corresponds to a direct sum of two 1D quasiperiodic operators, which can be treated as a system of two non-interacting particles on $\mathbb{Z}$. The function U in (1.1) is a (deterministic) interaction potential. A typical example of U would be a finite-range interaction, that is, $U(n_1, n_2) = f(n_1 - n_2)$, where f is a function on $\mathbb{Z}$ supported on a finite set. In general, we consider general potentials U of low complexity (see Sect. 2.1). In particular, they will always take finitely many values.

Our main motivation for studying the models (1.1) comes from the localization results for interacting particles in random environments, see [AW09, KN13, CS14]. In one way or another, random analogues of operators (1.1) demonstrate Anderson localization at large disorder λ. Another phenomenon is the following decoupling result: suppose the single particle is localized for some λ. Then, for sufficiently small finite range interaction U, where the smallness depends on λ, the multi-particle system will also be localized. Unfortunately, despite some progress obtained in the present paper, a similar question in the quasiperiodic case still remains open, and our current results are obtained using 2D methods.

The second source of motivation, which seems to be closer to applications, comes from the numerical work [FS15], see also references therein. They authors considered the almost Mathieu case $v(\theta) = \cos(2\pi\theta)$ and conjectured that there are some regimes where the addition of the interaction potential can generate some delocalized states (FIKS, that is, freed by interaction kinetic states), based on numerical evidence. In fact, their work goes as far as proposing real-life experiments with cold atoms, that can possibly confirm this prediction.

Assuming that U has low complexity in the sense of Sect. 2.1, localization results of the present paper depend on whether or not the quasiperiodic potential v has cosine-type symmetries. Generic analytic potentials v have no symmetries, and in this case we are able to obtain a perturbative localization result for large λ, assuming that a small positive measure set set of frequencies is removed, see Theorem 2.2. This includes the regime of strong interaction, as long as its strength is $O(|\lambda|)$, see (2.3). One can consider this result as an evidence of the fact that, in the asymmetric case, interaction of low complexity cannot break down the localization (however, the initial regime may need stronger disorder than just for single-particle localization).

In the case of potentials with symmetries, such as $v(\theta) = \cos(2\pi\theta)$, we still obtain localization at large disorder, however, we have to remove a part of spectrum of size $o(|\lambda|)$ around the (finitely many) values of U from consideration, see Theorem 2.3. Thus, possible FIKS are restricted to neighborhoods of finitely many energies, which comprise a relatively small part of the spectrum of the operator (1.1).

While we do not obtain any delocalization results in the present paper, the almost Mathieu version of (1.1) is studied in [BJK] (currently in preparation), in the regime of strong Hubbard-type interaction: that is,

$$v(\theta) = 2\cos(2\pi\theta), \quad U(n_1, n_2) = u\delta_{n_1,n_2}, \quad \lambda \text{ is fixed}, \quad u \gg \lambda.$$

In this case, the operator (1.1) has some spectrum in the region $[u-4-4\lambda, u+4+4\lambda]$, which is separated from the "bulk" spectrum. One can show that, for fixed phase difference $\theta_1 - \theta_2$ and large λ, the spectrum in the interval $[u - 4 - 4\lambda, u + 4 + 4\lambda]$ is purely point, for a full measure set of frequencies ω. On the other hand, for fixed $u \gg \lambda$ and some subsequent choice of $\theta_1 - \theta_2 \approx \pi$ (depending on u), the operator (1.1) has some non-trivial absolutely continuous spectrum in that interval. Thus, in the symmetric case, delocalization can indeed happen in some of the regions excluded in Theorem 2.3 away from the spectrum of the non-interacting operator (the latter is essential for the analysis in [BJK]). In the case where excluded energies belong to the non-interacting spectrum, the question of localization/delocalization remains largely open.

2 Summary of Main Results

We start from the description of the classes of potentials v and U that can be considered.

2.1 Low complexity interaction potentials. Let $T_{(n_1,n_2)}$ be the translation operator on $\ell^2(\mathbb{Z}^2)$:

$$(T_{n_1,n_2}\psi)(m_1, m_2) = \psi(m_1 - n_1, m_2 - n_2).$$

DEFINITION 2.1. *Let $U\colon \mathbb{Z}^2 \to \mathbb{R}$. For each $(n_1, n_2) \in \mathbb{Z}^2$, consider the function $T_{n_1,n_2}U$ restricted to $[0, N-1]^2$. We say that U has low complexity if the number of different functions among such restricted translations of U admits a power bound in N:*

$$\# \left\{ \mathbf{1}_{[0,N-1]^2} \cdot (T_{n_1,n_2}U) : n_1, n_2 \in \mathbb{Z} \right\} \leqslant N^{C_{\mathrm{int}}}, \quad (n_1, n_2) \in \mathbb{Z}^2, \quad N \geqslant 2, \qquad (2.1)$$

for some constant $C_{\mathrm{int}} > 0$.

In other words, U has low complexity if one can get at most $N^{C_{\mathrm{int}}}$ possible configurations by restricting U to a box of size N. This condition implies that U can only take finitely many values:

$$U(n_1, n_2) \in \{U_1, \ldots, U_{N_{\mathrm{int}}}\}, \qquad (2.2)$$

where say, $N_{\mathrm{int}} \leqslant 4^{C_{\mathrm{int}}}$. The class of possible U contains all periodic potentials on $\mathbb{Z}^2$, all finite range translationally invariant interaction potentials, and, in fact, some interesting examples such as Fibonacci-type potentials (see [DGY] and references therein).

We will also assume that U is not very large compared to the disorder:

$$|U(n_1, n_2)| \leqslant m_{\mathrm{int}}|\lambda|, \quad \text{for some } m_{\mathrm{int}} > 0. \qquad (2.3)$$

The possibility of considering background potentials of low complexity can already be traced back to the methods of [BGS02], although it was not stated there. While some low complexity potentials (such as Fibonacci hamiltonians) are known to cause singular continuous spectra by themselves even at small coupling, our results show that these effects are dominated by localization caused by quasiperiodic potentials at large disorder.

2.2 Symmetric and asymmetric single-particle potentials. The results also depend on whether or not the potential v has certain symmetries. In the latter, we will always assume that $v \in C^\omega(\mathbb{T}; \mathbb{R})$ and will identify $C^\omega(\mathbb{T}, \mathbb{R})$ with the space of all 1-periodic real analytic functions on $\mathbb{R}$. In addition, we will assume, without loss of generality, that $\int_{\mathbb{T}} v(\theta)\, d\theta = 0$ and that 1 is the smallest period of v. We call v *symmetric* if at least one of the following conditions holds:

(1) $v(\theta_{\mathrm{sym}} + \theta) = -v(\theta_{\mathrm{sym}} - \theta)$ for some $\theta_{\mathrm{sym}} \in \mathbb{T}$ and all θ.
(2) $v(\theta + 1/2) = -v(\theta)$ for all θ.

This symmetries will be called Type I and Type II, respectively. If none of them holds, we will call v *asymmetric*. In case of Type I symmetry, by shifting the function, *we can assume without loss of generality that $\theta_{\mathrm{sym}} = 0$.* For example, the (shifted) almost Mathieu potential $\cos(2\pi\theta - \pi/2) = \sin 2\pi\theta$ satisfies both symmetries, the function $\sin 2\pi\theta + \sin 4\pi\theta$ has only Type I, and $\cos 2\pi\theta + \sin 6\pi\theta$ has only Type II symmetry.

2.3 Main results. The following two theorems are main results of the present paper.

Theorem 2.2. *Suppose that v is asymmetric. For any $\varepsilon_{\text{freq}} > 0$ there exists $\lambda_0 = \lambda_0(v, \varepsilon_{\text{freq}}, m_{\text{int}}, C_{\text{int}})$ such that the following is true: for every $\lambda \geqslant \lambda_0$, every $(\theta_1, \theta_2) \in \mathbb{T}^2$, and every interaction potential U satisfying the assumptions of Sect. 2.1, there is a set*

$$\Omega(U, \lambda, \theta_1, \theta_2) \subset \mathbb{T}, \quad |\mathbb{T} \backslash \Omega(U, \lambda, \theta_1, \theta_2)| < \varepsilon_{\text{freq}},$$

with the operator $H(\theta_1, \theta_2)$ satisfying complete Anderson localization for all $\omega \in \Omega(U, \lambda, \theta_1, \theta_2)$ and for all possible translations of U.

Theorem 2.3. *Fix a background potential U of low complexity. Suppose v is symmetric and admits a bounded analytic extension into $|\operatorname{Im} z| \leqslant 20$. For any $\varepsilon_{\text{freq}} > 0$ there exist $\lambda_0 = \lambda_0(v, \varepsilon_{\text{freq}}, m_{\text{int}}, C_{\text{int}})$ and $\mu = \mu(v, \varepsilon_{\text{freq}}, m_{\text{int}}, C_{\text{int}}) > 0$ such that, for every $\lambda \geqslant \lambda_0$, every $(\theta_1, \theta_2) \in \mathbb{T}^2$, and every interaction potential U satisfying the assumptions of Sect. 2.1, there exists a set*

$$\Omega(U, \lambda, \theta_1, \theta_2) \subset \mathbb{T}, \quad |\mathbb{T} \backslash \Omega(U, \lambda, \theta_1, \theta_2)| < \varepsilon,$$

with the operator $H(\theta_1, \theta_2)$ satisfying Anderson localization in the region of energies E

$$|E - U_j| \geqslant \frac{\lambda}{e^{(\log \lambda)^\mu}}, \quad j = 1, \ldots, N_{\text{int}}, \tag{2.4}$$

for all $\omega \in \Omega(U, \lambda, \theta_1, \theta_2)$ and for all possible translations of U.

REMARK 2.4. (1) The part of the spectrum removed by (2.4) is contained in N_{int} intervals of size $o(|\lambda|)$.

(2) The condition on separability of v is irrelevant in Theorem 2.2. The potential $v(\theta_1) + v(\theta_2)$ can be replaced by an analytic function $w(\cdot, \cdot) \in C^\omega(\mathbb{T}^2)$ of two variables that is not constant on any straight line segment, with the same proof.

(3) An analogue of Theorem 2.3 can also be obtained for non-separable case, assuming $|E - (\lambda w_i + U_j)| \geqslant \lambda \varepsilon$ for every w_i such that $w(\theta_1, \theta_2) \equiv w_i$ on some straight line segment, and every value U_j of U, also with the same proof.

(4) As discussed above, the inclusion of a background potential U of low complexity could have been done already in [BGS02], as well as in the other papers that establish perturbative results by semi-algebraic techniques and do not involve Lyapunov exponents/cocycles (for example, in [BJ]).

(5) Suppose that v satisfies Type II symmetry, $U = 0$, $\theta_1 = \theta_2 = 1/4$. Then one can easily check that $\psi(n_1, n_2) = (-1)^{n_1} \delta_{n_1 n_2}$ solves the eigenvalue equation $H(1/4, 1/4)\psi = 0$. While this does not contradict purely point spectrum, all known proofs of Anderson localization show that any solution of the eigenvalue equation decays exponentially, which does not allow the existence of states like ψ. This example suggests that some stronger versions of localization can break

down at zero energy, but only in symmetric cases (because otherwise Theorem 2.2 holds). Possible scenarios of delocalization at zero energy in different models are described in [JSS03, ES17].

(6) The condition in Theorem 2.3 of analyticity of v in the strip of size 20 is technical and can possibly be removed with some extra work.

(7) The operator family $H(\theta_1, \theta_2)$ is not ergodic because U is not assumed to have any translational invariance. However, one can prove localization simultaneously for all translations of U.

(8) Our results are perturbative, in the sense that one always has to remove a positive measure set of frequencies. However, our requirements on the frequency are more explicit. The bound on λ_0 can be expressed, in principle, through the Diophantine constant C_{dio} of ω (see (5.1)). The parameter $\varepsilon_{\mathrm{freq}}$ in Theorems 2.2 and 2.3 is, essentially, the measure of frequencies for which (5.1) does not hold with this C_{dio}. Afterwards, as usually happens in localization proofs, one has to remove an extra set of frequencies of measure zero, for which we do not have any arithmetic description, and which depends on θ_1, θ_2, λ, and other parameters.

(9) Theorems 2.2 and 2.3 are formulated for the case of a single phase (θ_1, θ_2). However, one can extend them for a full measure set of phases, see Remark 9.1. We do not believe this argument is new, however, in the case of perturbative results, it has not been explicitly stated in the literature.

(10) In the case when $\|U\|_\infty$ is bounded by a constant independent of λ, the removed energy intervals in Theorem 2.3 become one neighborhood of zero energy of the same size.

2.4 Structure of the paper. The main lines of argument are parallel to the only known techniques of establishing multi-dimensional Anderson localization [BGS02, Bou07]. The proof consists of two main steps: a large deviation theorem for Green's function at fixed energy (Sect. 3–6), and elimination of energy (Sect. 7–9).

In the large deviation result, we need to obtain stronger inductive assumptions to carry over from a scale to the next scale. That is, we require that the large deviation set has small sections by lines in all directions, rather than in the coordinate directions only. In Sect. 3, we establish these bounds at the initial scale. In the case of no symmetries, the initial scale bounds at all energies follow from Proposition 3.2 and are essentially known. In the case when symmetries are present, similar argument would immediately work for $|E - U_j| \geqslant \varepsilon |\lambda|$. To get better initial scale bounds in the wider region (2.4), one has to apply 1D large deviations more carefully, as done in Sect. 3. Section 4 contains preliminaries from theory of semi-algebraic sets (some proofs are provided in the "Appendix"). In order to pass to the next scale, one needs an arithmetic bound on the number of bad boxes (cf. [BGS02, Section 3]). In our proof, a stronger inductive assumption is carried over to the next scale, but there is less freedom in removal of frequencies, as the frequency vector is always of the form (ω, ω). Still, a relatively simple argument in Sect. 5 shows that one can get

a sub-linear bound on the number of bad sites assuming a Diophantine condition on ω. In Sect. 6, we provide the inductive argument for obtaining Green's function estimates at fixed energy. After the preparations from previous sections, the proof goes along the lines of [BGS02]. In particular, our Proposition 6.4 is an analogue of [BGS02, Lemma 4.4]. For the convenience of the reader, we include the proof based on Cartan's lemma in the "Appendix". Our Corollary 6.5 is the main result of that section and is similar to [BGS02, Proposition 4.6]; however, the control of the constants is somewhat more delicate, and we provide the argument in the main text.

The elimination of energy requires more work, as the steep planes argument used in [BGS02], [Bou07] significantly relies on the fact that the dimension of the lattice is the same as the dimension of the set of frequency vectors, which is not the case in our situation. However, our large deviation theorem is stronger, and we take advantage of that. In Sect. 7, we use more elaborate semi-algebraic arguments and bounds on Kakeya maximal functions, in order to estimate the number of directions in which the large deviation set can contain a long line segment, see Lemmas 7.1 and 7.2. In Sect. 8, we show that double resonances can be avoided by removing a small frequency set, if one combines the results of Sect. 7 with some careful choice of scales. Section 9 completes the proof of localization, which at this point becomes fairly standard. In Remark 9.1, we also explain how to obtain results for sets of phases of full measure. While this argument is not new, it has not been explicitly stated in previous works.

3   Symmetries of the Potential and the Initial Scale

Everywhere in this section, we will assume

(1) $v \in C^\omega(\mathbb{T}; \mathbb{R})$.
(2) 1 is the smallest period of v (in particular, $v \neq \mathrm{const}$).
(3) $\int_{\mathbb{T}} v(\theta)\, d\theta = 0$.
(4) $|v(\theta)| \leqslant 1$ for all $\theta \in \mathbb{T}$.

The results will heavily rely on the structure of level sets of the function $v(\theta_1)+v(\theta_2)$. We will have to avoid the situation when these sets contain straight line segments.

LEMMA 3.1. *The level set $v(\theta_1) + v(\theta_2) = E$ contains a straight line segment if and only if $E = 0$ and v is symmetric. In the case of Type I symmetry $v(\theta_{\mathrm{sym}} + \theta) = -v(\theta_{\mathrm{sym}} - \theta)$ this segment has the equation $\theta_{\mathrm{sym}} - \theta_1 = \theta_{\mathrm{sym}} + \theta_2$, and in case of type II symmetry $v(\theta + 1/2) = -v(\theta)$ this segment is $\theta_1 = \frac{1}{2} + \theta_2$ (both equalities are modulo $\mathbb{Z}$).*

Proof. Suppose that $v(\theta) + v(a\theta + b) \equiv E$. Since 1 is the smallest period of v, we have $a = \pm 1$, and hence, from comparing the mean values, we have $E = 0$. Suppose that $a = 1$. Then $v(\theta) = -v(\theta + b) = v(\theta + 2b)$, and hence $b = 1/2$. If $a = -1$,

then $v(\theta) = -v(b - \theta)$. Then the function $v_1(\theta) = v(b/2 + \theta)$ is odd, and has same symmetry for $a = 1$ if and only if v has it. $\qquad \square$

In the rest of this section, we obtain Green's function estimates at the initial scale for H. We will need to use some large deviation theorems for analytic functions. The case of asymmetric v is essentially known.

PROPOSITION 3.2. *Suppose $w \in C^\omega(\mathbb{T}^2)$ is non-constant on any line segment in $\mathbb{T}^2$. Then there are positive constants C, α depending only on v such that, for any unit line segment $L \subset \mathbb{R}^2$, any $E \in \mathbb{R}$, and any $\delta > 0$, we have*

$$|\{(\theta_1, \theta_2) \in L \colon |w(\theta_1, \theta_2) - E| \leqslant \delta\}|_1 \leqslant C\delta^\alpha, \tag{3.1}$$

where $|\cdot|_1$ denotes the one-dimensional Lebesgue measure. The constant C can be chosen uniformly in E on any compact interval.

Proof. Define an analytic function $f \in C^\omega((-1,1)^3)$ by

$$f(\theta_1, \theta_2, \eta) = w(\theta_1 \cos 2\pi\eta + \theta_2 \sin 2\pi\eta, -\theta_1 \sin 2\pi\eta + \theta_2 \cos 2\pi\eta).$$

The function $f(\cdot, \theta_2, \eta)$ is non-constant in θ_1 at any fixed θ_2, η. Moreover, by the choice of θ_2 and η, one can parametrize any line segment with θ_1 being the natural length parameter. Then, one can refer to the beginning of Section 4 of [BGS02] and [GS01, Lemma 11.4]. $\qquad \square$

REMARK 3.3. If v is asymmetric, then $w(\theta_1, \theta_2) = v(\theta_1) + v(\theta_2)$ satisfies the assumptions of Proposition 3.1.

We now consider the symmetric separable case. Any line segment in $L \subset \mathbb{T}^2$ can be parametrized either by θ_1 or θ_2. Since v is 1-periodic, the restriction of $v(\theta_1) + v(\theta_2)$ onto L can be completely described by the function

$$\theta \mapsto v(\theta) + v(a\theta + b), \quad -1 \leqslant a \leqslant 1, \ 0 \leqslant b \leqslant 1. \tag{3.2}$$

Let

$$g(v, a, b) = \max_{-1/2 \leqslant \theta \leqslant 1/2} |v'(\theta) + av'(a\theta + b)|.$$

LEMMA 3.4. *Suppose that v is symmetric. Let also $-1 \leqslant a \leqslant 1, 0 \leqslant b < 1$. Then, for some $C(v) > 0$, we have*

(1) *If v has only Type I symmetry with $\theta_{\mathrm{sym}} = 0$, then*

$$g(v, a, b) \geqslant C(v)(|a + 1| + b(1 - b)). \tag{3.3}$$

(2) *If v has only Type II symmetry, then*

$$g(v, a, b) \geqslant C(v)(|a - 1| + |b - 1/2|). \tag{3.4}$$

(3) *If v has both Type I (with $\theta_{\mathrm{sym}} = 0$) and Type II symmetries, then*

$$g(v, a, b) \geqslant C(v)(|a^2 - 1| + b(1 - b)|b - 1/2|). \tag{3.5}$$

Proof. The inequality

$$g(v, a, b) \geqslant C_1 \min\{|a - 1|, |a + 1|\} \tag{3.6}$$

is clearly satisfied in all three cases with $C_1 = C_1(v) = \max_{\theta \in \mathbb{T}} |v'(\theta)|$ (since the first term in the derivative of (3.2) attains its maximum for some $\theta \in [0, 1]$). Let us also note that, in all three cases, $g(v, a, b)$ is Lipschitz in a uniformly in b:

$$|g(v, a_1, b) - g(v, a_2, b)| \leqslant C_{\mathrm{Lip}}|a_1 - a_2|, \quad C_{\mathrm{Lip}} = C_{\mathrm{Lip}}(v) > 0. \tag{3.7}$$

We now address each of the three cases separately.

Case 1. Due to continuity of g and absence of Type II symmetry, for any $\varepsilon > 0$, we have

$$g(v, a, b) \geqslant C(\varepsilon) > 0, \quad \text{for} \quad -1 + \varepsilon \leqslant a \leqslant 1, \ 0 \leqslant b \leqslant 1.$$

Hence, (3.6) can be replaced by $g(v, a, b) \geqslant C_1|a+1|$, which implies (3.3) for $C_{\mathrm{Lip}}|a + 1| \geqslant \frac{1}{2}b(1 - b)$. Suppose now that $C_{\mathrm{Lip}}|a + 1| < \frac{1}{2}b(1 - b)$. Then, (3.7) implies that it would be sufficient to establish (3.3) for $a = -1$, that is, to estimate $g(v, -1, b)$ from below. Let

$$v(\theta) = \sum_{n \in \mathbb{Z}} c_n e^{2\pi i n \theta}.$$

Then

$$g(v, -1, b)^2 \geqslant \int_0^1 |v'(\theta) - v'(b - \theta)|^2 \, d\theta = \sum_{n \in \mathbb{Z}} 4n^2 |c_n|^2 \sin^2(\pi n b).$$

Since 1 is the smallest period of v, there is a finite index set $I \subset \mathbb{Z}$ with $\gcd(I) = 1$ and $c_n \neq 0$ for $n \in I$, which implies

$$g(v, -1, b)^2 \geqslant \sum_{n \in I} 4n^2 |c_n|^2 \sin^2(\pi n b) \geqslant Cb^2(b - 1)^2,$$

since the last expression can only vanish for $b = 0$ or $b = 1$, and in both cases admits a quadratic lower bound.

Case 2. Similarly to Case 1, (3.6) can be replaced by $g(v, a, b) \geqslant C_1|a - 1|$, which implies (3.4) for $C_{\mathrm{Lip}}|a + 1| \geqslant \frac{1}{2}|b - 1/2|$. Hence, we can assume that the opposite inequality holds, which allows to consider $a = 1$ and obtain a similar Fourier estimate:

$$g(v, 1, b)^2 \geqslant \sum_{n \in \mathbb{Z}} 4n^2 |c_n|^2 \cos^2(\pi n b) \geqslant \sum_{n \in I} 4n^2 |c_n|^2 \cos^2(\pi n b),$$

where $I \subset \mathbb{Z}$ is an index set with $\gcd(I) = 1$ and $c_n \neq 0$ for $n \in I$. Type II symmetry implies that all such n must be odd, and hence the right hand side can only vanish for $b = 1/2$, with any of the terms providing a lower bound $C(b - 1/2)^2$.

Case 3. Similarly to Case 1 and Case 2, (3.6) immediately implies (3.5) for $C_{\mathrm{Lip}}|a^2 - 1| \geqslant \frac{1}{4}b(1 - b)|b - 1/2|$. In case of the opposite inequality, one can replace a by 1 or -1 using (3.7) and then apply the same Fourier lower bound from Case 2 or Case 1, respectively. $\qquad\square$

One can also obtain upper bounds, which immediately follow from Lipschitz continuity of v, and combine both results into the following

COROLLARY 3.5. *Under the assumptions of Lemma* 3.4, *we have the following two-sided bounds with $C_-(v), C_+(v) > 0$*

(1) *If v has only Type I symmetry with $\theta_{\mathrm{sym}} = 0$, then*

$$C_-(v)(|a + 1| + b(1 - b)) \leqslant g(v, a, b) \leqslant C_+(v)(|a + 1| + |b(b - 1)|).$$

(2) *If v has only Type II symmetry, then*

$$C_-(v)(|a - 1| + |b - 1/2|) \leqslant g(v, a, b) \leqslant C_+(v)(|a - 1| + |b - 1/2|).$$

(3) *If v has both Type I (with $\theta_{\mathrm{sym}} = 0$) and Type II symmetries, then*

$$C_-(v)(|a^2 - 1| + b(1 - b)|b - 1/2|) \leqslant g(v, a, b) \leqslant C_+(v)(|a^2 - 1| \\ + b(1 - b)|b - 1/2|).$$

Moreover, each upper bound also holds for $|v(\theta) + v(a\theta + b)|$ uniformly in $\theta \in [0, 1]$.

REMARK 3.6. In a separable asymmetric case, similar arguments imply that $g(v, a, b) \geqslant \varepsilon(v) > 0$ uniformly in a, b.

REMARK 3.7. In case of Type I symmetry with $\theta_{\mathrm{sym}} \neq 0$, the second term in Cases 1 and 3 will be different due to the shift of b. As mentioned earlier, we will always assume $\theta_{\mathrm{sym}} = 0$.

PROPOSITION 3.8. *Let f be an analytic function in the disk $|z| \leqslant 2e$, $|f(z)| \leqslant M$ for $|z| \leqslant 2e$, and $f(0) = 1$. Let $D = \{z \colon |z| \leqslant 1, |f(z)| \leqslant \lambda\}$. Then D can be covered by a union of disks of total diameter bounded by $C \exp\{\frac{\log \lambda}{\log M}\}$ (where C is an absolute constant). In particular,*

$$|D \cap [-1, 1]|_1 \leqslant C \exp\left\{2\frac{\log \lambda}{\log M}\right\}.$$

As earlier, $|\cdot|_1$ denotes the 1D Lebesgue measure. For the proof, see Theorem 4 in Section 11.3 of [Lev96].

Theorem 3.9. *Suppose v is symmetric and extends to a bounded analytic function in the strip $|\operatorname{Im} z| \leqslant 20$. Then, for any line segment $L \subset \mathbb{R}^2$ of unit length, and any $E \in [-3,3]\backslash\{0\}$, $\delta > 0$, we have*

$$|\{(\theta_1, \theta_2) \in L \colon |v(\theta_1) + v(\theta_2) - E| \leqslant \delta\}|_1 \leqslant c_1(v)\delta^{-\frac{c_2(v)}{\log|c_3(v)E|}}, \qquad (3.8)$$

where $|\cdot|_1$ denotes the one-dimensional Lebesgue measure and $c_1(v), c_2(v) > 0$, $0 < c_3(v) \leqslant 1/6$ depend only on v.

Proof. We will assume that L is described by a, b as in (3.2), and the points of L are parametrized by θ (which is either θ_1 or θ_2). Without loss of generality, we can also assume that $\delta \leqslant |E|/2$, otherwise the bound can be obtained by choosing a sufficiently large $c_1(v)$. Finally, one only needs to consider a, b satisfying

$$g(v, a, b) \geqslant \frac{C_-(v)}{2C_+(v)}|E|, \qquad (3.9)$$

otherwise the set in the left hand side of (3.8) is empty. The function

$$h(\theta) = v'(\theta) + av'(a\theta + b),$$

satisfies $|h(\theta_0)| \geqslant \frac{C_-(v)}{2C_+(v)}|E|$ for some $\theta_0 \in [-1/2, 1/2]$. Also, h extends to the strip $|\operatorname{Im} z| \leqslant 20$ and satisfies $|h(z)| \leqslant h_{\max}(v)$ in that strip for some $h_{\max}(v) > 0$. We can pick the bound $h_{\max}(v)$ in such a way that $h_{\max}(v) \geqslant |h(\theta_0)|$ uniformly in a, b, in the range considered. Lemma 3.8 applied to the function

$$h_1(\theta) = \frac{h(\theta - \theta_0)}{h(\theta_0)}.$$

with $M = h_{\max}(v)/h(\theta_0)$ and $\lambda = \eta/h(\theta_0)$ yields the following bound

$$|\{\theta \in [-1/2, 1/2] \colon |h(\theta)| < \eta\}| \leqslant C \exp\left\{2\frac{\log \eta - \log|h(\theta_0)|}{\log h_{\max}(v) - \log|h(\theta_0)|}\right\}. \qquad (3.10)$$

We have

$$\exp\left\{2\frac{-\log|h(\theta_0)|}{\log h_{\max}(v) - \log|h(\theta_0)|}\right\} = \exp\left\{2 - \frac{\log h_{\max}(v)}{\log h_{\max}(v) - \log|h(\theta_0)|}\right\} \leqslant C_1(v),$$

and

$$0 < \log h_{\max}(v) - \log|h(\theta_0)| \leqslant \max\{\log 2, -\log C_2(v)|E|\} \leqslant -\log C_3(v)|E|,$$

where $C_3(v) = \min\{C_2(v), 1/6\}$ (since we are only considering $|E| \leqslant 3$). Using the last two estimates, (3.10) becomes

$$|\{\theta \in [-1/2, 1/2] \colon |h(\theta)| < \eta\}| \leqslant C_4(v)\eta^{-\frac{2}{\log C_3(v)|E|}}, \quad |E| \leqslant 3,$$

where the implicit dependence on E of the left hand side is in the choice a, b restricted by (3.9).

To estimate the set in the left hand side of (3.8), note that the number of intervals of monotonicity of the function $\theta \mapsto v(\theta) + v(a\theta + b)$ is bounded by a constant $M(v)$ that depends only of v (note that this fact is not trivial and is shown in [Gab68]). By considering the sets where $|h(\theta)| < \eta$ and $|h(\theta)| \geqslant \eta$, we arrive to

$$|\text{l. h. s. of } (3.8)| \leqslant C_5(v)(\eta^{-1}\delta + \eta^{-\frac{2}{\log C_3(v)|E|}})$$

(where $M(v)$ is absorbed by $C_5(v)$). Balancing the powers leads to

$$|\text{l. h. s. of } (3.8)| \leqslant C_6(v)\delta^{\frac{2}{2-\log C_3(v)|E|}},$$

which implies (3.8). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Following [BGS02], define an *elementary region* $\Lambda \subset \mathbb{Z}^2$ as a difference of a rectangle and its translation over some non-zero lattice vector. Let $\mathcal{ER}(N)$ be the set of all elementary regions of diameter N; the diameter of Λ is denoted by $\text{diam}(\Lambda)$. For an elementary region $\Lambda \in \mathcal{ER}(N)$, define the *Green's function* of the operator $H(\theta_1, \theta_2)$ restricted to Λ in the usual way,

$$G_\Lambda(\theta_1, \theta_2, E) = (H_\Lambda(\theta_1, \theta_2) - E)^{-1} = (\mathbf{1}_\Lambda(H(\theta_1, \theta_2) - E)|_{\text{ran } \mathbf{1}_\Lambda})^{-1},$$

where $\mathbf{1}_\Lambda$ is the indicator function of Λ (that is, the restriction operator). For $\gamma > 0$, define also the set of "good" phases,

$$\mathcal{G}^{\gamma,b}(\Lambda, E) = \{\theta_1, \theta_2 \subset \mathbb{T}^2 : G_\Lambda(\theta_1, \theta_2, E) \text{ satisfies } (3.12),(3.13)\}, \text{ that is,} \quad (3.11)$$

$$\|G_\Lambda(\theta_1, \theta_2, E)\| < \lambda^{-1}e^{\sigma(\Lambda)^b}, \qquad\qquad\qquad\qquad\qquad\qquad\qquad (3.12)$$

$$|G_\Lambda(\theta_1, \theta_2, E)(n_1, n_2)| < e^{-\gamma|n_1 - n_2|} \text{ for all } n_1, n_2 \in \Lambda, |n_1 - n_2| \geqslant \frac{1}{4}\sigma(\Lambda).$$
$$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (3.13)$$

Here the norm in (3.12) can be chosen to be, for example, Hilbert–Schmidt norm (however, the choice of a particular matrix norm does not matter as one can essentially ignore factors of N^C). The complementary set of "bad" phases is

$$\mathcal{B}^{\gamma,b}(\Lambda, E) = \mathbb{T}^2 \backslash \mathcal{G}^{\gamma,b}(\Lambda, E). \qquad\qquad\qquad\qquad\qquad (3.14)$$

Theorem 3.10. *Suppose v satisfies the assumptions of Theorem 3.9. Fix $b, \mu \in (0, 1)$. Then, for all unit line segments $L \subset \mathbb{R}^2$, we have*

$$|\mathcal{B}^{\gamma,b}(\Lambda, E) \cap L|_1 \leqslant \exp\{-\sigma(\Lambda)^{b(1-\mu)}\},$$

uniformly in $\Lambda \in \mathcal{ER}(N)$ with $\gamma = \frac{1}{2}\log\lambda$, assuming

$$\lambda \geqslant \lambda_0(N, b, v, \mu), \quad N \geqslant N_0(b, v, \mu),$$
$$|E - U_j| \geqslant \frac{\lambda}{\exp\{\sigma(\Lambda)^{b\mu}\}}, \quad j = 1, \dots, N_{\text{int}}. \qquad (3.15)$$

Proof. Suppose

$$|v(\theta_1 + n_1\omega) + v(\theta_2 + n_2\omega) - (E - U_j)/\lambda| > \delta,$$
$$\text{for all } (n_1, n_2) \in \Lambda, \ j = 1, \ldots, N_{\text{int}}. \tag{3.16}$$

Then, using the resolvent identity (where Δ_Λ and U_Λ denote the restrictions of the corresponding operators)

$$(H_\Lambda(\theta_1, \theta_2) - E)^{-1} = (\lambda V(\theta_1, \theta_2) - E - \Delta_\Lambda - U_\Lambda)^{-1}$$
$$= \{I - (\lambda V_\Lambda(\theta_1, \theta_2) + U_\Lambda - E)^{-1}\Delta_\Lambda\}^{-1}(\lambda V_\Lambda(\theta_1, \theta_2) + U_\Lambda - E)^{-1},$$

we can see (cf. [BGS02, Lemma 4.1], or by direct expansion of the resolvent)

$$|(H_\Lambda(\theta_1, \theta_2) - E)^{-1}(n_1, n_2)| \leqslant (1 - 16\lambda^{-1}\delta^{-1})^{-1}(16\lambda^{-1}\delta^{-1})^{|n_1 - n_2| + 1}, \tag{3.17}$$

and

$$\|(H_\Lambda(\theta_1, \theta_2) - E)^{-1}\| \leqslant 8\delta^{-1}\lambda^{-1}, \tag{3.18}$$

where we used the fact that $\|\Delta_\Lambda\| \leqslant \|\Delta\| = 4$. Take $\lambda \geqslant e^{\sigma(\Lambda)^b}$, $\delta = \lambda^{-1/2}$, and suppose that (θ_1, θ_2) satisfy (3.16). Since $\lambda^{-1}\delta^{-1} = \lambda^{-1/2}$, (3.17) and (3.18) imply

$$|(H_\Lambda(\theta_1, \theta_2) - E)^{-1}(n_1, n_2)| \leqslant e^{-\frac{1}{2}\log(\lambda)|n_1 - n_2|},$$
$$\|(H_\Lambda(\theta_1, \theta_2) - E)^{-1}\| \leqslant C\lambda^{-1}e^{\frac{1}{2}\sigma(\Lambda)^b},$$

which implies (3.12),(3.13). It remains to estimate the measure of the set of θ for which (3.16) fails. Let $L \subset \mathbb{R}^2$ be a line segment. Apply Theorem 3.9:

$$|\{(\theta_1, \theta_2) \in L \colon (3.16) \text{ fails}\}|_1 \leqslant c_1(v)N_{\text{int}}|\Lambda| \exp\left\{-\frac{c_2(v)\log\delta}{\log|c_3(v)E/\lambda|}\right\}$$

$$\leqslant c_1(v)N_{\text{int}}|\Lambda| \exp\left\{\frac{c_2(v)\log\delta}{\sigma(\Lambda)^{b\mu}}\right\} = c_1(v)N_{\text{int}}|\Lambda| \exp\left\{-\frac{1}{2}c_2(v)\sigma(\Lambda)^{b(1-\mu)}\right\}.$$

The constants $c_1(v)$ and $-\frac{1}{2}c_2(v)$, as well as the factor $N_{\text{int}}|\Lambda|$, can be absorbed into the exponent, if one chooses a slightly smaller μ in the beginning. □

REMARK 3.11. The parameter E in Theorem 3.9 corresponds to $\lambda^{-1}(E - U_j)$ in Theorem 3.10.

4 Some Preliminaries

Convention regarding the constants. For simplicity of the language, we will often use the following construction: the constant C in N^C will always mean some absolute constant, and the constant does not have to be the same in all claims. For example, a claim "Suppose $A(N^C)$. Then $B(N^C)$" will mean the following: for every C_1 there exists C_2 such that $A(N^{C_1})$ implies $B(N^{C_2})$. Usually, the exact dependence of C_2 on C_1 will not be important.

We formulate two standard covering lemmas [BGS02, Lemma 2.2, Lemma 2.4] with slight changes in notation that does not affect the proofs.

PROPOSITION 4.1. *Suppose $\Lambda \subset \mathbb{Z}^2$ is an arbitrary set with the following property: for every $m \in \mathbb{Z}^2$, there is a subset $W(m) \subset \Lambda$ with $m \in W(m)$, $\operatorname{diam} W(m) \leqslant N$, and the Green's function $G_{W(m)}(E)$ satisfying for some $t, N, A > 0$*

$$\|G_{W(m)}(E)\| < A,$$
$$|G_{W(m)}(E)(m,n)| \leqslant e^{-tN}, \quad \text{for all} \quad n \in \partial_* W(m),$$

where $\partial_* W(m)$ is the internal boundary of $W(m)$ relative to Λ, that is,

$$\partial_* W(m) = \{n \in W(m) \colon \text{ there exists } k \in \Lambda \backslash W(m) \text{ with } |k-n| = 1\}.$$

Then, assuming $4N^2 e^{-tN} \leqslant \frac{1}{2}$, we have

$$\|G_\Lambda(E)\| \leqslant 2N^2 A.$$

PROPOSITION 4.2. *Suppose M, N are positive integers such that for some $0 < \tau < 1$*

$$N^\tau \leqslant M \leqslant 2N^\tau.$$

Let $\Lambda_0 \in \mathcal{ER}(N)$ be an elementary region of size N with the property that for all $\Lambda \subset \Lambda_0$, $\Lambda \in \mathcal{ER}(L)$ with $M \leqslant L \leqslant N$, the Green's function $G_\Lambda(E)$ satisfies

$$\|G_\Lambda(E)\| \leqslant e^{L^b} \tag{4.1}$$

for some fixed $0 < b < 1$. We say that $\Lambda \in \mathcal{ER}(L)$, $\Lambda \subset \Lambda_0$ is good if, in addition to (4.1), we have the off-diagonal decay, that is,

$$|G_\Lambda(m,n)| \leqslant e^{-\gamma|m-n|} \quad \text{for all } m, n \in \Lambda, \ |m-n| > \frac{1}{4}L,$$

where $\gamma > 0$ is fixed. Otherwise, Λ is called bad. Assume that any family of pairwise disjoint bad M'-regions in Λ_0 with $M + 1 \leqslant M' \leqslant 2M + 1$, has at most N^b regions in it. Then, under these assumptions, we have

$$|G_{\Lambda_0}(m,n)| \leqslant e^{-\gamma'|m-n|} \quad \text{for all } m, n \in \Lambda_0, \ |m-n| > \frac{1}{4}N,$$

where $\gamma' = \gamma - N^{-\delta}$ and $\delta = \delta(b, \tau) > 0$, provided $N \geqslant N_0(b, \tau, \gamma)$.

REMARK 4.3. While the result of Proposition 4.2 is proved for a fixed τ, one can choose δ and N_0 uniformly to serve an interval $\tau \in [\tau_0, \tau_1] \subset (0, 1)$ (it is important that the endpoints are separated from 0 and 1).

4.1 Facts from real algebraic geometry. Similarly to all previously known higher-dimensional localization results, we will actively use real semi-algebraic sets. We assume that the reader is familiar with [BGS02, Section 7] or [Bou05, Chapter 9], which contains the definitions and relevant references. Below, we summarize some facts from real algebraic geometry that will be used during later constructions (starting from the definition).

(sa1) A set $S \subset \mathbb{R}^d$ is called (closed) *semi-algebraic* if it is a finite union of sets, each of which is defined by finitely many polynomial inequalities or equalities of the form $Q \geqslant 0$, $Q \leqslant 0$, or $Q = 0$. We say that $\deg S \leqslant sd$, if the set S can be described using s inequalities of the above type, with polynomials Q of degree $\leqslant d$ (and $\deg S$ is defined as the smallest possible value of sd among all such representations). By a slight abuse of notation, bounds of the form $A \leqslant (\deg S)^C$ would actually mean "there exists an absolute constant C such that the quantity is bounded by $A \leqslant (2 + \deg S)^{C}$", in order to avoid considering separate case $\deg S = 1$ or $\deg S = 0$ (we will only be interested in such bounds for $\deg S \gg 1$).

(sa2) If $S_1, S_2 \subset \mathbb{R}^d$ are semi-algebraic, then $S_1 \cup S_2$, $S_1 \cap S_2$ are semi-algebraic, each of degree $\leqslant (\deg S_1 + 1)(\deg S_2 + 1)$.

(sa3) *Tarski–Seidenberg principle*: let $S \subset \mathbb{R}^{d+1}$ be semi-algebraic, and $p \colon \mathbb{R}^{d+1} \to \mathbb{R}^d$ be the standard projection. Then $p(S)$ is semi-algebraic, and $\deg p(S) \leqslant (\deg S)^{C(d)}$. In general, $p(S)$ may not be closed, but it will always be closed if S is compact. We will always apply this property to compact semi-algebraic sets.

(sa4) Let $S \subset [0,1]^2$ be a semi-algebraic set of zero Lebesgue measure. For any $\varepsilon > 0$, the ε-neighborhood of S can be covered by $(\deg S)^{C(d)}\varepsilon^{-1}$ balls of radius ε with centers on S.

(sa5) Let $S \subset [0,1]^d$ be semi-algebraic. Then ∂S is also semi-algebraic, $\deg(\partial S) \leqslant (\deg S)^{C(d)}$. Moreover, ∂S is a union of $\leqslant (\deg S)^{C(d)}$ semi-algebraic sets of dimensions $\leqslant d - 1$, whose Hausdorff measures of the corresponding dimensions are bounded by $(\deg S)^{C(d)}$. One can also obtain stronger statements from Proposition 10.3.

(sa6) A semi-algebraic set $S \subset [0,1]^2$ of Lebesgue measure $\geqslant \varepsilon$ always contains a ball of radius $(\deg S)^{-C}\varepsilon$. This property can also be applied on 2D surfaces. The following example will be important. Suppose $\mathcal{C} \subset [0,1]^2$ is a piecewise algebraic curve of length L with $\leqslant N$ smooth pieces of degree $\leqslant B$ each. Let $I \subset [0,1]$ be an interval of length L. Suppose, $\mathcal{C} \times I \subset [0,1]^3$ is a union of K semi-algebraic subsets of degree $\leqslant B^C$. Then there is a curve segment $\mathcal{C}' \subset \mathcal{C}$ and a sub-interval $J \subset I$ such that $\mathcal{C}' \times J$ is contained in one of the K above-mentioned subsets, where $|\mathcal{C}'|, |J| \geqslant L^2 B^{-C} K^{-1}$.

(sa7) A semi-algebraic set $S \subset \mathbb{R}^d$ has $\leqslant (\deg S)^{C(d)}$ connected components, each of which is a semi-algebraic set of degree $\leqslant (\deg S)^{C(d)}$. If $d = 1$, then S is a union of $\leqslant (\deg S)^{C(d)}$ closed line segments.

(sa8) Instead of $\mathbb{R}^d$, one can consider semi-algebraic subsets of $\mathbb{T}^d$ or $\mathbb{S}^d$, using algebraic local charts (the coordinates induced from $\mathbb{R}^d$ on the torus, or stereographic projection on the sphere), with obvious modifications for previous properties. As an example: let $\mathcal{A} \subset \mathbb{R}^{d+1}$ be semi-algebraic. Fix $\eta > 0$, and let $\Xi \subset \mathbb{S}^d$ be the set of directions in which $\mathcal{A}$ contains a line segment of length $\geqslant \eta$. Then Ξ is semi-algebraic, and $\deg \Xi \leqslant (\deg \mathcal{A})^{C(d)}$. Similar constructions will be used during the course of the energy elimination part. In particular, an analogue of (sa6) holds for $\mathcal{C} \subset \mathbb{S}^2$.

(sa9) Let $S \subset [0,1]^d$ be a (topologically) connected semi-algebraic subset. Then S is algebraically path connected. That is, for every $p, q \in S$ there exists a curve $C \subset S$ connecting p and q such that C is a union of $(\deg S)^{C(d)}$ smooth algebraic pieces of degrees $\leqslant (\deg S)^{C(d)}$ and of total length $\leqslant (\deg S)^{C(d)}$ (uniformly in p, q).

(sa10) Let $S \subset [0,1]^d$ be a semialgebraic subset, $\operatorname{diam}(S) \leqslant 1$. During some proofs, we will use the following dyadic layer expansion of S. Fix $0 < \varepsilon < 1$. Define

$$S(k) = \{x \in S \colon \operatorname{dist}(x, \partial S) = 2^{-k}\}, \quad k = 1, 2, 3, \ldots, \lceil \log_2(\varepsilon^{-1}) \rceil.$$

Then each $S(k)$ is a semi-algebraic subset (possibly empty),

$$\deg S(k) \leqslant (\deg S)^C, \quad \dim S_k \leqslant d - 1,$$

and the following relations between neighborhoods are true (here B_ε denotes the closed ε-ball about the origin, so that $S + B_\varepsilon$ is the ε-neighborhood of S):

$$\bigcup_k (S(k) + B_{2^{-(k+1)}}) \subset S \subset \left(\bigcup_k (S(k) + B_{2^{-(k+1)}}) \right) \cup (\partial S + B_\varepsilon).$$

Properties (sa1)–(sa3), (sa4), (sa7) are well known, see references in [BGS02, Section 7]. The remaining properties essentially follow from the quantitative triangulation theorem (Proposition 10.3). We include proofs of (sa5), (sa6), (sa8), and (sa9) in the "Appendix". Property (sa10) follows from the fact that $\operatorname{dist}(x, \partial S)$ is a semi-algebraic function of x, and the level sets of the distance function are always of dimension $\leqslant d - 1$.

The following lemma is a modification of the cylindrical decomposition for semialgebraic sets, see, for example, [BCR98, Section 5.1], with added condition regarding convexity.

LEMMA 4.4. Let $\mathcal{A} \subset [0,1]^2$ be a closed semialgebraic subset of degree B. Then $\mathcal{A}$ can be decomposed into a (not necessarily disjoint) union of sets of the following form:

$$\{(\theta_1, \theta_2) \colon \theta_1 \in I, \; \Theta_-(\theta_1) \leqslant \theta_2 \leqslant \Theta_+(\theta_1)\}$$

where $I \subset [0,1]$ is an interval, $\Theta_\pm \colon I \to [0,1]$ is a continuous algebraic function on I of degree $\leqslant B^C$, smooth on $I \backslash \partial I$, and either linear, or strictly convex, or strictly concave.

Proof. Using quantitative triangulation theorem (Proposition 10.3), one can assume without loss of generality that $\mathcal{A}$ is an algebraic diffeomorphic image of a simplex. Call a point $a \in \partial\mathcal{A}$ singular, if one of the following is true:

- ∂A is not C^∞-smooth at a.
- The tangent vector to ∂A at a is parallel to one of the coordinate axes, and a is an isolated point of $\partial\mathcal{A}$ with this property.
- ∂A is C^∞-smooth at a, but changes convexity at that point.

The set of singular points is a finite semi-algebraic set of degree $\leqslant B^C$, and hence contains at most B^C points. Consider the grid formed by horizontal and vertical lines drawn at each singular point. This grid will split $[0,1]^2$ into $\leqslant B^C$ rectangles. Without loss of generality, we can restrict ourselves to one of these rectangles R. Since $\mathcal{A} \cap R$ has $\leqslant B^C$ connected components, we can further restrict to a single connected component $\mathcal{A}_1$ of $\mathcal{A} \cap R$. Clearly, the projection of $\mathcal{A}_1$ onto the θ_1 axis is a closed interval I. For each $\theta_1 \in I$ the set $\{\theta_2 \colon (\theta_1, \theta_2) \in \mathcal{A}_1\}$ is also a closed interval, and hence $\mathcal{A}_1$ is the space between two graphs of functions. These two functions must be continuous, but do not have to be smooth due to presence of ∂R. However, one can split I into $\leqslant B^C$ intervals of smoothness, in which case they will be of constant convexity on each new interval. $\square$

5 Arithmetic Conditions on Frequency and Thin Semi-algebraic Sets

For a closed set $\mathcal{A} \subset [0,1]^2$, denote by $\eta(\mathcal{A})$ the length of the longest line segment contained in $\mathcal{A}$ (which exists due to compactness).

Theorem 5.1. *Let $\omega \in [0,1)$ satisfy the following Diophantine condition:*

$$\|k\omega\| \geqslant C_{\mathrm{dio}}|k|^{-1-\delta_{\mathrm{dio}}}, \quad 1 \leqslant |k| \leqslant N, \quad C_{\mathrm{dio}}, \delta_{\mathrm{dio}} > 0. \tag{5.1}$$

Let $\mathcal{A} \subset [0,1]^2$ be a semi-algebraic set with

$$\eta(\mathcal{A}) < \min_{1 \leqslant |k| \leqslant 2N} \|k\omega\|. \tag{5.2}$$

Then

$$\#\{(k_1, k_2) \in \mathbb{Z}^2 \colon (\{k_1\omega\}, \{k_2\omega\}) \in \mathcal{A}, |k_1|, |k_2| \leqslant N\} \leqslant (\deg \mathcal{A})^C C_\omega' N^{3/4+3\delta_{\mathrm{dio}}}. \tag{5.3}$$

Proof. We will refer to the set $\{(\{k_1\omega\}, \{k_2\omega\}), |k_1|, |k_2| \leqslant N\}$ as "N-lattice points", and the intersection with $\mathcal{A}$ as "N-lattice points on $\mathcal{A}$", or "lattice points on $\mathcal{A}$". Since we can essentially ignore a factor $(\deg \mathcal{A})^C$, we may, by splitting $\mathcal{A}$ into $\leqslant (\deg \mathcal{A})^C$ components and possibly switching the variables, make the following reductions:

(1) $\mathcal{A}$ *is a single piece described in Lemma* 4.4. *The functions* $\Theta_\pm$ *satisfy* $|\Theta_+(\theta_1) - \Theta_-(\theta_1)| \leqslant \eta(\mathcal{A})$ *for all* $\theta_1 \in I$, *and the slope angle of the tangent vector to the graph of each function does not change more than by* $\pi/8$ *on* I. All these assumptions can be obtained from Proposition 4.4 by introducing a $(\deg A)^C$ factor, using the assumptions on $\mathcal{A}$.

(2) *Both functions* Θ_+ *and* Θ_- *are strictly convex on* I. Let $R = I \times I'$ be the smallest rectangle that contains the graphs of $\Theta_\pm$. If the functions Θ_+ and Θ_- have opposite convexity or one of them is linear, then one can easily check that $\mathcal{A} \cap R$ will contain a line segment of length at least $\frac{1}{3}\operatorname{diam}(R)$, and hence $\operatorname{diam} R \leqslant 3\eta(\mathcal{A})$. In particular, R contains at most 9 lattice points, and hence one can absorb the contribution from all such pieces or $\mathcal{A}$ into the factor $(\deg \mathcal{A})^C C'_{\mathrm{dio}}$. The remaining pieces are strictly convex or concave, and can be assumed to be convex without loss of generality.

(3) $\mathcal{A}$ *can be covered by* $(\deg \mathcal{A})^C N^{3/4}$ *balls of raduis* $N^{-3/4}$, *with covering multiplicity* $\leqslant 100$. Follows from (sa4) applied to $\partial\mathcal{A}$, as any cover of $\partial\mathcal{A}$ by balls with centers on $\partial\mathcal{A}$ will also cover $\mathcal{A}$. The multiplicity part follows from Vitali covering lemma (one should first cover $\mathcal{A}$ by balls of size $\frac{1}{3}N^{-3/4}$).

(4) $\mathcal{A}$ *cannot contain four distinct* N-*lattice points* p_1, p_2, p_3, p_4 *satisfying*

$$p_2 - p_1 = p_4 - p_3.$$

Indeed, in that case the convex hull of these points must lie above the graph of Θ_1. Moreover, since tangent vector to the graph of Θ_2 cannot change direction by more that $\pi/8$, the graph will not intersect at least one line segment $p_i p_j$ with $i \neq j$, and hence at least one such segment will be completely contained in $\mathcal{A}$, contradicting (5.2).

We say that $v \in \mathbb{R}^2$ is a *short lattice distance vector* if $v = p_1 - p_2$ for two distinct N-lattice points p_1, p_2, and $|v| \leqslant 2N^{-3/4}$. Clearly, both components of v cannot exceed $2N^{-3/4}$ by absolute value which, due to Diophantine condition, implies that the number of different short lattice distance vectors is bounded by

$$(C_1(\omega)N^{-3/4}N^{1+\delta_{\mathrm{dio}}})^2 = C_1(\omega)^2 N^{1/2 + 2\delta_{\mathrm{dio}}}. \tag{5.4}$$

Let B be one of the covering balls of $\mathcal{A}$ of radius $N^{-3/4}$, obtained in (3), and assume that B contains two N-lattice points $p_1, p_2 \in \mathcal{A}$. Property (4) implies that $\mathcal{A}$ cannot contain two points with difference vector $p_2 - p_1$, unless one of these points coincides with p_1 or p_2. Using the multiplicity property from (3), we obtain that, for each short lattice difference vector v, *there are at most* 200 *balls that contain some pair of lattice points on* $\mathcal{A}$ *with difference vector* v. Since the number of different short lattice distance vectors is bounded by (5.4), we can conclude that the number of covering balls that contain two or more lattice points on $\mathcal{A}$ is bounded by $200 C_1(\omega)^2 N^{1/2 + 2\delta_{\mathrm{dio}}}$.

Let us estimate the number of lattice points on $\mathcal{A}$ in each ball. Since $\mathcal{A}$ intersects each vertical line in a segment of length $\leqslant \eta$, there is at most one N-lattice point

on $\mathcal{A}$ on each vertical line. By counting possible values of the horizontal coordinate, one obtains (similarly to (5.4)) the total bound of $C_1(\omega)N^{1/4+\delta_{\mathrm{dio}}}$ lattice points on $\mathcal{A}$ in each $N^{-3/4}$-ball.

It remains to combine the estimates. We split the balls obtained in (3) into two groups. The number of balls that contain at most one lattice point on $\mathcal{A}$ can be simply bounded by the total number of balls, which is $(\deg\mathcal{A})^C N^{3/4}$. For the balls with two or more points, we estimated the number of such balls by $O(N^{1/2+2\delta_{\mathrm{dio}}})$, and the number of lattice points on $\mathcal{A}$ in each ball by $O(N^{1/4+\delta_{\mathrm{dio}}})$. This gives

$$\#\mathrm{l.\,h.\,s.\ of\ } (5.3) \leqslant (\deg\mathcal{A})^C (200 C_1(\omega)^3 N^{1/4+\delta_{\mathrm{dio}}} N^{1/2+2\delta_{\mathrm{dio}}} + N^{3/4}),$$

which implies the required bound. $\qquad\square$

REMARK 5.2. The constant C'_{dio} can be bounded by, say, $C(1 + C_{\mathrm{dio}}^{-10})$, where C is an absolute constant. One can also replace (5.2) by $\eta(\mathcal{A}) \leqslant N^{-1-\delta_{\mathrm{dio}}-\varepsilon}$ and the right hand side of (5.3) by $(\deg\mathcal{A})^C N^{3/4+3\delta_{\mathrm{dio}}+\varepsilon}$, by introducing an additional requirement $N \geqslant N_0(C_{\mathrm{dio}}, \varepsilon)$.

6 Multi-scale Induction Step and Estimate of the Green's Function

Let H be the operator (1.1) for some fixed background potential U, satisfying the complexity bound (2.1). Recall the definitions of good and bad sets for Green's functions,

$$\mathcal{G}^{\gamma,b}(\Lambda, E) = \{\theta \subset \mathbb{T}^2 : G_\Lambda(\theta_1, \theta_2, E) \text{ satisfies } (3.12),(3.13)\}, \text{ that is,}$$
$$\|G_\Lambda(\theta_1, \theta_2, E)\| \leqslant \lambda^{-1} e^{\sigma(\Lambda)^b}.$$
$$|G_\Lambda(\theta_1, \theta_2, E)(n_1, n_2)| \leqslant e^{-\gamma|n_1-n_2|} \text{ for all } n_1, n_2 \in \Lambda, |n_1-n_2| \geqslant \frac{1}{4}\sigma(\Lambda).$$
$$\mathcal{B}^{\gamma,b}(\Lambda, E) = \mathbb{T}^2 \backslash \mathcal{G}^{\gamma,b}(\Lambda, E).$$

The sets $\mathcal{G}$ and $\mathcal{B}$ depend on U. It will be convenient to introduce the smaller set $\mathcal{G}_U^{\gamma,b}(\Lambda, E)$ as follows: we say that $(\theta_1, \theta_2) \in \mathcal{G}_U^{\gamma,b}(\Lambda, E)$ if $G_\Lambda(\theta_1, \theta_2, E)$ satisfies (3.12), (3.13) *for all possible translations of U* by vectors from $\mathbb{Z}^2$. For $\Lambda \in \mathcal{ER}(M)$, there are at most $M^{C_{\mathrm{int}}}$ translations that give different G_Λ, which allows to retain the possibility of shifting U, while keeping the sets $\mathcal{B}$ exponentially small. Let also

$$\mathcal{B}_U^{\gamma,b}(\Lambda, E) = \mathbb{T}^2 \backslash \mathcal{G}_U^{\gamma,b}(\Lambda, E).$$

We will often use the following elementary consequence of the resolvent identity.

PROPOSITION 6.1. *Let $\Lambda \in \mathcal{ER}(N)$, $H_1 = \Delta + \lambda V_1$, $H_2 = \Delta + \lambda V_2$ are both Schrödinger operators on $\ell^2(\Lambda)$, $\|V_j\|_\infty \leqslant 2$, $0 < b < 1$, $\lambda > 1$, and*

$$\|H_1^{-1}\| < \lambda^{-1} e^{\sigma(\Lambda)^b}, \tag{6.1}$$

$$|H_1^{-1}(n_1, n_2)| < e^{-\gamma|n_1-n_2|} \text{ for all } n_1, n_2 \in \Lambda, |n_1-n_2| \geqslant \frac{1}{4}\sigma(\Lambda). \tag{6.2}$$

Assume also that $\|V_1 - V_2\|_\infty \leqslant e^{-3\gamma_1 N}$, where $\gamma_1 = \max\{\gamma, 1\}$, and $N^b \leqslant \frac{1}{10}\gamma_1 N$. Then

$$\|H_2^{-1}\| < 2\lambda^{-1} e^{\sigma(\Lambda)^b},$$

$$|H_2^{-1}(n_1, n_2)| < 2e^{-\gamma|n_1 - n_2|} \text{ for all } n_1, n_2 \in \Lambda, |n_1 - n_2| \geqslant \frac{1}{4}\sigma(\Lambda).$$

Proof. The resolvent identity implies

$$H_2^{-1} = H_1^{-1} + \lambda H_1^{-1}(V_1 - V_2)H_2^{-1},$$

or

$$H_2^{-1} = H_1^{-1}(1 - \lambda H_1^{-1}(V_1 - V_2))^{-1}.$$

Using Neumann series, we have

$$\|H_2^{-1} - H_1^{-1}\| \leqslant \lambda^{-1} e^{N^b}\left(e^{N^b - 3\gamma_1 N} + e^{2(N^b - 3\gamma_1 N)} + \cdots\right) \leqslant 2\lambda^{-1} e^{-2\gamma_1 N},$$

from which both estimates follow. $\qquad\square$

The following is a refined version of the claims of [BGS02, Remark 3.2, Remark 3.4, Lemma 4.2 and Remark 4.3] which state, essentially, that one can treat sets of parameters $E, \omega, \theta_1, \theta_2$ for which the Green's function $G_\Lambda(E, \theta_1, \theta_2)$ is good/bad, as semi-algebraic sets of degree $\leqslant N^C$, where $\Lambda \in \mathcal{ER}(N)$.

LEMMA 6.2. *Let $\Lambda \in \mathcal{ER}(N)$ be an elementary region. Define*

$$S_\Lambda^{\gamma,b} = \{(E, \omega, \theta_1, \theta_2) \colon G_\Lambda(E, \theta_1, \theta_2) \text{ does not satisfy (3.12) or (3.13)}\}.$$

Then there exists a semialgebraic subset $\mathcal{A}_\Lambda^{\gamma,b} \subset \mathbb{R} \times [0,1]^3$, $\deg \mathcal{A}_\Lambda^{\gamma,b} \leqslant N^C$, such that

$$\mathcal{A}_\Lambda^{\gamma,b} + B_{e^{-10\gamma_1 N}} \subset S_\Lambda^{\gamma,b},$$

and for every $(E, \omega, \theta_1, \theta_2) \in \mathbb{R} \times [0,1] \times [0,1]^2 \backslash \mathcal{A}_\Lambda^{\gamma,b}$ we have

$$\|G_\Lambda(\theta_1, \theta_2, E)\| < 8\lambda^{-1} e^{\sigma(\Lambda)^b}, \tag{6.3}$$

$$|G_\Lambda(\theta_1, \theta_2, E)(n_1, n_2)| < 8e^{-\gamma|n_1 - n_2|} \text{ for all } n_1, n_2 \in \Lambda, |n_1 - n_2| \geqslant \frac{1}{4}\sigma(\Lambda). \tag{6.4}$$

The same is true if one considers the sets in variables $\omega, \theta_1, \theta_2$ with fixed E, or with fixed ω, or both.

Proof. Without loss of generality, one can assume that $(0,0) \in \Lambda$, since the translation of Λ by (k_1, k_2) is equivalent to replacing (θ_1, θ_2) by $(\{\theta_1 + k_1\omega\}, \{\theta_2 + k_2\omega\})$.

We will assume that the potential is of the form

$$v_{n_1, n_2}(\omega, \theta_1, \theta_2) = w(\theta_1 + n_1\omega, \theta_2 + n_2\omega), \quad (n_1, n_2) \in \Lambda,$$

and w is a real analytic function. Let $w_D(\theta_1, \theta_2)$ be a trigonometric polynomial obtained by truncating the Fourier series of w up to order D. Clearly, for $c(v) > 0$, we have

$$\|w_D - w\|_{C^1[0,1]^2} \leqslant e^{-c(w)D}.$$

Proposition 6.1 implies that (3.12) and (3.13) can become worse at most by a factor of 2, if one replaces w by w_D, as long as $D \geqslant C(w, \gamma)N$. We will now replace each trigonometric term of w_D by a polynominal. It is sufficient to consider terms of the form

$$\cos(k\omega + \theta_1) = \cos(k\omega)\cos(\theta_1) - \sin(k\omega)\sin(\theta_1),$$

$|k| \leqslant D$, $|\omega|, |\theta_1| < 1$. Let c_M be the M-th Taylor polynomial for the cosine. Then

$$|c_M(k\omega) - \cos(k\omega)| \lesssim \frac{|k|^M}{M!} \leqslant \frac{D^M}{M!},$$

which can be made small by taking, say, $M \gtrsim D^3$. We thus replace each $v_{n_1,n_2}(\omega, \theta_1, \theta_2)$ by a polynomial $w_{n_1,n_2}(\omega, \theta_1, \theta_2)$ of degree, say, $\leqslant C(w, \gamma)N^{10}$, so that

$$\|v_{n_1,n_2} - w_{n_1,n_2}\|_{C^1([0,1]^3)} \leqslant e^{-100\gamma_1 N}.$$

Now, define $\mathcal{A}_\Lambda$ to be the set of parameters for which (3.12) or (3.13) fail for the operator with potential w_{n_1,n_2}, by at least a factor of 4 in the right hand side. Then, Proposition (6.1) and the fact that v_{n_1,n_2} is very close to w_{n_1,n_2}, implies the claims of the lemma. $\qquad\square$

REMARK 6.3. While the constants in the proof depend on γ, one can easily transfer the dependence on γ into the choice of the initial scale; which, in further arguments, will depend on γ anyway. One can also replace the constant 8 in (6.3), (6.4) by any $C > 1$.

In the following, we will denote by C'_{int}, C'_{dio} some constants that depend, respectively, only on C_{int}, C_{dio}. We will use this notation in a way similar to N^C. Namely, the use of C'_{int} will mean that the claim is true with some constant that depends only on C_{int}. The actual value of that constant may depend on the context.

PROPOSITION 6.4. Fix $\gamma, C_{\mathrm{dio}}, \delta_{\mathrm{dio}}, m_{\mathrm{int}}, C_{\mathrm{int}} > 0$. There exist constants C'_{int} and C'_{dio} such that, if $0 < b < 1$, $0 < \rho < 1$, $C_1 > 0$ satisfy

$$b - 3/4 - 3\delta_{\mathrm{dio}} - 3\rho > 0, \quad C_1 > \frac{C'_{\mathrm{int}} + C'_{\mathrm{dio}} + 1/\rho}{b - 3/4 - 3\delta_{\mathrm{dio}} - 3\rho}, \tag{6.5}$$

then there exists a positive integer $\overline{N}_0 = \overline{N}_0(\gamma, b, \rho, C_1, C_{\mathrm{dio}}, \delta_{\mathrm{dio}}, C_{\mathrm{int}}, m_{\mathrm{int}})$ such that the following is true. Suppose N_0, N_1 are positive integers,

$$\overline{N}_0 \leqslant 100N_0 \leqslant N_1^\rho,$$

and for every $N_0 \leqslant M \leqslant N_1$ and any elementary region $\Lambda \in \mathcal{ER}(M)$ we have

$$\sup_{L \subset \mathbb{R}^2, E \in \mathbb{R}} |\mathcal{B}^{\gamma,b}(\Lambda, E) \cap L|_1 \leqslant \exp(-M^\rho), \tag{6.6}$$

where L runs over all unit line segments. Let N satisfy

$$N_0^{C_1} \leqslant N \leqslant N_1^{\rho C_1},$$

and assume that ω satisfies the finite scale Diophantine condition

$$\|k\omega\| \geqslant C_{\mathrm{dio}}|k|^{-1-\delta_{\mathrm{dio}}}, 1 \leqslant |k| \leqslant 100N. \tag{6.7}$$

Then, for all $\Lambda \subset \mathcal{ER}(N)$ and any unit line segment $L \subset \mathbb{R}^2$ we have (for all possible translations of U)

$$\sup_{E \in \mathbb{R}} |\{(\theta_1, \theta_2) \in L \colon \|G_\Lambda(\theta_1, \theta_2, E)\| > \lambda^{-1} e^{N^b}\}|_1 < e^{-N^{3\rho}}. \tag{6.8}$$

The line of the argument in the proof is very similar to [BGS02, Lemma 4.4]. We include it in the "Appendix" for the convenience of the reader. The following is the main result of this section.

COROLLARY 6.5. *Fix* $C_{\mathrm{dio}}, \delta_{\mathrm{dio}}, b, \rho, C_1, C_{\mathrm{int}}, m_{\mathrm{int}}$ *as in Proposition* 6.4, *with extra assumption* $C_1 > 2/\rho$. *Fix also* $\gamma_0 > 0$. *There exist* $\lambda_0 = \lambda_0(C_{\mathrm{dio}}, \gamma_0, b, \rho, \delta_{\mathrm{dio}}, C_1, C_{\mathrm{int}}, m_{\mathrm{int}})$ *such that, for* $\lambda \geqslant \lambda_0$, *one can find* $\widetilde{N}_0 = \widetilde{N}_0(\lambda, C_{\mathrm{dio}}, \gamma_0, b, \rho, \delta_{\mathrm{dio}}, C_1, C_{\mathrm{int}}, m_{\mathrm{int}})$ *so that the following bound* (6.10) *is true for all* $N \geqslant \widetilde{N}_0$, *all* ω *satisfying the finite scale Diophantine condition*

$$\|k\omega\| \geqslant C_{\mathrm{dio}}|k|^{-1-\delta_{\mathrm{dio}}}, 1 \leqslant |k| \leqslant 100N^2, \tag{6.9}$$

all $\Lambda \in \mathcal{ER}(N)$, *and all unit line segments* L *in* $\mathbb{R}^2$:

$$|\mathcal{B}_U^{\gamma,b}(\Lambda, E) \cap L|_1 \leqslant \exp\{-\sigma(\Lambda)^\rho\}. \tag{6.10}$$

For the asymmetric case, one needs in addition to assume

$$|E - U_j| \geqslant \frac{\lambda}{\exp\{(\log \lambda)^{1/(\varepsilon C_1)}\}}, \quad j = 1, \ldots, N_{\mathrm{int}}. \tag{6.11}$$

where $0 < \varepsilon = \varepsilon(b, m_{\mathrm{int}}, C_{\mathrm{int}}) < 1$ *is a small constant.*

Proof. Fix some $\varepsilon = \varepsilon(b, m_{\mathrm{int}}, C_{\mathrm{int}}) > 0$ (the exact choice will be made later in the proof). There exists a scale $N_0' = N_0'(\gamma_0, b)$ and $\delta = \delta(\gamma_0, b)$, such that Proposition 4.2 is applicable in the range

$$\gamma \in [\gamma_0, 2\gamma_0] = [\gamma_0, \gamma_1], \quad \tau \in [1/4, 3/4] \tag{6.12}$$

for all $N \geqslant N_0'(\gamma_0, b)$ and with $\delta(b, \tau) = \delta$. Choose $\overline{N}_0$ such that Proposition 6.4 is applicable with $\gamma = \gamma_0$ and other parameters introduced earlier, and take $\widetilde{N}_0 \geqslant \max\{100\overline{N}_0, N_0'\}^{1/\varepsilon}$. Suppose also that λ and E are chosen to satisfy (6.6) with

$\gamma = \gamma_1 = 2\gamma_0$ for all "initial scales" $M \in [\widetilde{N}_0^\varepsilon, \widetilde{N}_0^{C_1}]$. We have yet to show that such choice is possible in line with (6.11), which will be discussed in the end of the proof.

Since $C_1 > 2/\rho$, Proposition 6.4 applied with $N_0 = \widetilde{N}_0$, $N_1 = [\widetilde{N}_0^{C_1}]$ thus provides a norm bound for Green's function (6.8) for $N \in [\widetilde{N}_0^{C_1}, \widetilde{N}_0^{2C_1}]$. We now need to extend the off-diagonal decay bound (3.13) from $[\widetilde{N}_0^\varepsilon, \widetilde{N}_0^{C_1}]$ (where we have it due to the assumption), into the new interval $[\widetilde{N}_0^{C_1}, \widetilde{N}_0^{2C_1}]$. Let

$$N \in [\widetilde{N}_0^{C_1}, \widetilde{N}_0^{2C_1}], \quad M_0 = [N^{\varepsilon C_1^{-1}}] \in [\widetilde{N}_0^\varepsilon, \widetilde{N}_0^{2\varepsilon}], \quad M_1 = [N^{1/2}].$$

Both scales M_0 and M_1 satisfy (6.6) with $\gamma = \gamma_1$. We follow the argument in [BGS02, Corollary 4.5] and apply Proposition 4.2 with $M = M_1$ and $N = N$, which agrees with the choice (6.12). We will need to verify the main condition in Proposition 4.2, that is, that the number of bad boxes of sizes in $[M_1/2, 2M_1]$ is bounded by N^b. We will do it through an auxiliary scale M_0. The following argument is very similar to the proof of Proposition 6.4 (see "Appendix"). Let

$$\mathcal{B} = \cup_{M_0+1 \leqslant L \leqslant 2M_0+1} \cup_{\Lambda \in \mathcal{ER}(L), \Lambda \subset [-2M_0, 2M_0]^2} \mathcal{B}_U^{\gamma_1, b}(\Lambda, E). \tag{6.13}$$

In view of Lemma 6.2, $\mathcal{B}$ can be replaced by a semi-algebraic set $\mathcal{A}$, $\deg \mathcal{A} \leqslant M_0^{C'_{\text{int}}}$, satisfying the following: for $(\theta_1, \theta_2) \in [0,1] \backslash \mathcal{A}$, the same conclusion (3.12), (3.13) holds as it was for $(\theta_1, \theta_2) \in [0,1]^2 \backslash \mathcal{B}$, with an extra factor of 8. For any unit line segment $L \subset \mathbb{R}^2$,

$$|\mathcal{A} \cap L|_1 \leqslant M_0^{C'_{\text{int}}} e^{-M_0^\rho}.$$

Suppose that $\widetilde{N}_0$ is chosen to be large enough (depending on ε), in order to have, for all M_0 under consideration,

$$M_0^{C'_{\text{int}}} e^{-M_0^\rho} \leqslant C_{\text{dio}} N^{-1-\delta_{\text{dio}}}. \tag{6.14}$$

Then, one can apply Theorem 5.1 and obtain

$$\#\{n \in [-N, N]^2 \colon \theta + n\omega \in \mathcal{A}\} \leqslant M_0^{C'_{\text{int}}} C'_{\text{dio}} N^{3/4 + 3\delta_{\text{dio}}} \leqslant C'_{\text{dio}} N^{\varepsilon C_1^{-1} + 3/4 + 3\delta_{\text{dio}}}. \tag{6.15}$$

where C'_{int} is a new constant so that the factor $M_0^{C'_{\text{int}}}$ absorbs $(\deg \mathcal{A})^C$ from Theorem 5.1. The assumption (6.5) provides that one can choose $\varepsilon > 0$ in a way that the right hand side of (6.15) is dominated by N^b; one also needs to make another assumption of largeness of the initial scale, of the same type as in Proposition 4.2, which can be taken care of after choosing ε.

The estimate (6.15) provides an upper bound on the number of bad M_0-regions. Similarly to the derivation [BGS02, Equation following (4.36)], one can conclude that most of the M_1-regions $\Lambda_1 \in \mathcal{ER}(M_1)$ satisfy the following:

$$\|G_{\Lambda_1}(\theta_1, \theta_2, E)(m, n)\| \leqslant e^{-\gamma_1 |m-n| - CM_0}, \quad \forall m, n \in \Lambda_1, \quad |m - n| > \frac{1}{4} M_1. \tag{6.16}$$

More precisely, an M_1-region can violate (6.16) only if there is M_0-box with center on that region, violating (3.12) or (3.13). Hence, the number of disjoint M_1-regions violating (6.16), is also bounded by N^b. This verifies the assumptions of Proposition 4.2, which we can now apply and obtain the off-diagonal decay bound (3.13) on the scales $[\widetilde{N}_0^{C_1}, \widetilde{N}_0^{2C_1}]$, with

$$\gamma_2 = \gamma_1 - M_1^{-\delta} \geqslant \gamma_1 - \widetilde{N}_0^{-C_1\delta/2}. \tag{6.17}$$

The process can now be repeated with $\widetilde{N}_1 = \widetilde{N}_0^2$, and $\widetilde{N}_{j+1} = \widetilde{N}_j^2$, as long as γ_j obtained by the analogue of (6.17) stay within the range $[\gamma_0, \gamma_1]$. Due to super-exponential growth of scales $\widetilde{N}_j$, this can be achieved by taking a large initial scale $\widetilde{N}_0$. One can also absorb the extra factor of 8 from Proposition 6.2 into a further decrease of γ_j similar to (6.17), which also will not violate $\gamma_j \geqslant \gamma_0$; note that we only need to do it with respect to (3.13), since (3.12) has extra λ^{-1} in it, which (4.1) does not require.

It remains to explain the choice of the initial scale. Theorem 3.10 needs to be applied with the parameter $b' = b(1-\mu)^{-1}$ with some small μ, so that $b' < 1$. The condition for the required estimate is

$$\lambda \geqslant \lambda_0(\widetilde{N}_0) \approx e^{\sigma(\Lambda)^{b'}} \approx \exp\{\widetilde{N}_0^{b'C_1}\}.$$

One needs to pick $\widetilde{N}_0$ in order to satisfy, say, $\log\lambda_0 > 4\gamma_0$. Then the conclusion of the present Corollary will be true, however, the range of energies would be defined by (3.15), whose denominator does not depend on λ and therefore is not optimal for large λ. One can improve it in the following way: if λ is larger than λ_0 that is required for the scheme to work under all previous assumptions on $\widetilde{N}_0$, one can further increase $\widetilde{N}_0$ as much as Theorem 3.10 permits for that λ; one can check that the resulting range of energies will be of the form (6.11). $\qquad\square$

7 Line Segments in Semi-algebraic Sets

In this section, we consider semi-algebraic subsets of $[0,1]^3$ in the variables $(\omega, \theta_1, \theta_2)$. We will establish several estimates on the amount of long line segments in those sets. Let

$$S \subset [0,1]^3, \quad \deg S \leqslant B, \quad \dim S \leqslant 2 \tag{7.1}$$

be a two-dimensional semi-algebraic subset. Let also

$$S_\varepsilon = S + B_\varepsilon, \quad 0 < \varepsilon < e^{-B^\rho}, \tag{7.2}$$

be the closed ε-neighborhood of S, where $\rho > 0$. The coordinates in $[0,1]^3$ will be denoted by $(\omega, \theta_1, \theta_2)$. Suppose that, for any line segment $L \subset \{\omega\} \times [0,1]^2 \subset [0,1]^3$, perpendicular to the ω axis, we have

$$|L \cap S_{2\varepsilon}|_1 < e^{-B^\rho} \tag{7.3}$$

(note that the condition involves a larger neighborhood $S_{2\varepsilon}$). Let

$$e^{(\log B)^{\beta_1}} \leqslant K \leqslant e^{(\log B)^{\beta_2}}, \quad 1 < \beta_1 < \beta_2.$$

Fix some $\nu > 0$. For $\xi \in \mathbb{S}^2$, consider

$A_\xi^1 = \{(\omega, \theta_1)\colon S_\varepsilon \cap ((\omega, \theta_1, 0) + [-1,1]\xi$ contains an interval of size $K^{-\nu}\} \subset [0,1]^2$.
$A_\xi^2 = \{(\omega, \theta_2)\colon S_\varepsilon \cap ((\omega, 0, \theta_2) + [-1,1]\xi$ contains an interval of size $K^{-\nu}\} \subset [0,1]^2$.

Let us call a direction ξ *singular*, if the set $\mathrm{Proj}_\omega A_\xi^i$ contains an interval of size K^{-1} for $i = 1$ or $i = 2$. For $k = (k_1, k_2) \in [0,K]^2 \cap \mathbb{Z}^2$, define

$$\xi_k = \frac{(1, k_1, k_2)}{\sqrt{1 + k_1^2 + k_2^2}}. \tag{7.4}$$

LEMMA 7.1. *Let S_ε, $S_{2\varepsilon}$, K, B, ν, ρ satisfy the above assumptions. Fix $0 < c_1 < c_2$. There are at most $B^{C(\beta_1, \beta_2, \nu, \rho, c_1, c_2)}$ pairs $k = (k_1, k_2)$ with $c_1 K \leqslant |k| \leqslant c_2 K$ and ξ_k singular.*

Proof. Let

$$\Xi = \{\xi\colon \mathrm{Proj}_\omega A_\xi^1 \text{ contains an interval of size } 1/K\} \subset \mathbb{S}^2.$$

Clearly, Ξ is a semi-algebraic set of degree $\leqslant B^C$ (see (sa8)) and $\{\xi_k\colon k \in \mathcal{K}\} \subset \Xi$. The points ξ_k of the latter set are K^{-2}-separated. Since Ξ has at most B^C connected components, we may assume, by taking large enough $C(\beta_1, \beta_2, \nu, \rho, c_1, c_2)$, that there are at least two singular points ξ_k in the same component (which will, in the end of the proof, lead to a contradiction). By (sa9), there is a curve $\mathcal{C}_0$ in Ξ connecting those two singular points. One can assume that $\mathcal{C}_0$ consists of B^C smooth algebraic pieces of degree B^C, of total diameter $\geqslant K^{-2}$. Hence, one of the pieces is a smooth algebraic curve $\mathcal{C}_1$ in Ξ of diameter $\geqslant B^{-C}K^{-2}$.

Let us make a few reductions with the curve $\mathcal{C}_1$. By definition, each point or $\mathcal{C}_1$ defines a direction ξ such that $\mathrm{Proj}_\omega A_\xi^1$ contains $1/K$-interval. Let us split $[0,1]$ into intervals of size $(2K)^{-1}$. Then, for each $\xi \in \mathcal{C}_1$, the set $\mathrm{Proj}_\omega A_\xi^1$ contains at least one interval of the form $[j/(2K), (j+1)/(2K)]$. Let us split Ξ into $2K$ (possibly overlapping) pieces:

$$\Xi_j = \{\xi \in \Xi\colon \mathrm{Proj}_\omega A_\xi^1 \supset [j/(2K), (j+1)/(2K)]\}, \quad j = 0, 1, \ldots, 2K - 1.$$

Clearly, each Ξ_j is also semi-algebraic of degree B^C. Hence, there is a curve piece

$$\mathcal{C}_2 \subset \mathcal{C}_1, \quad \mathrm{diam}\,\mathcal{C} \geqslant B^{-C}K^{-3}, \quad \mathcal{C}_2 \subset \Xi_j \text{ for some } 0 \leqslant j \leqslant 2K - 1.$$

Without loss of generality, one can assume that $\mathcal{C}_2$ is also smooth and connected (both assumptions lead to a further loss of at most B^C). We thus achieved the following: by sacrificing a factor of $B^C K$ in the size of $\mathcal{C}_1$, the condition "$\mathrm{Proj}_\omega A_\xi^1$

contains some interval of size $1/K$ for each $\xi \in \mathcal{C}_1$" has been replaced by "$\mathrm{Proj}_\omega A_\xi^1$ contains a particular interval of size $1/2K$ for each $\xi \in \mathcal{C}_2$". We will call this construction "freezing" the interval of ω.

By definition of A_ξ^1, for each $\omega \in I$ and $\xi \in \mathcal{C}$, there exists $\theta_1(\xi, \omega)$ such that $(\omega, \theta_1(\xi, \omega)) + [0, 1]\xi) \cap \mathcal{E}$ contains a line segment

$$(\omega, \theta_1(\xi, \omega)) + I'(\xi, \omega)\xi \tag{7.5}$$

of length $K^{-\nu}$. We can repeat the same "freezing" procedure with the interval $I'(\xi, \omega)$. That is, split $[0, 1]$ into $2K^\nu$ segments J_l and define

$$\mathcal{D}_l = \{(\xi, \omega) \in \mathcal{C}_2 \times I : \exists \theta_1 \colon S_\varepsilon \text{ contains (7.5) with } I'(\xi, \omega) \supset J_l\}.$$

Clearly, $\mathcal{C}_2 \times I = \cup_l \mathcal{D}_l$. Note that, while we do not require $\theta_1(\xi, \omega)$ and $I'(\xi, \omega)$ to have any algebraic dependence of (ξ, ω), the existence of such θ_1, I' with $I' \supset J_l$ is a semi-algebraic condition of degree $\leqslant B^C$ for every l, and hence $\mathcal{D}_l$ are semi-algebraic sets of degree $\leqslant B^C$. Since $\mathcal{C}_2 \times I$ has 2D measure $\geqslant K^{-4}N^{-C}$, one can apply (sa6) and conclude that there exists a curve segment $\mathcal{C}_3 \subset \mathcal{C}_2$ and an interval $I_2 \subset I$, both of diameter $\geqslant K^{-4-\nu}B^{-C}$, such that $\mathcal{C}_3 \times I_2$ is completely contained in one of the sets $\mathcal{D}_l$.

We can summarize the previous paragraph in the following claim. There exists a smooth algebraic curve $\mathcal{C}_3 \subset \Xi$, $\mathrm{diam}\,\mathcal{C}_3 \geqslant K^{-5-\nu}$, an interval $I_2 \subset [0, 1]$, $|I_2| \geqslant K^{-5-\nu}$, and an interval $J \subset [0, 1]$, $|J| \geqslant \frac{1}{2}K^{-\nu}$ such that, for $(\omega, \xi) \in I_2 \times \mathcal{C}_3$, we have for some $\theta_1 \in [0, 1]$

$$(\omega, \theta_1, 0) + \xi(J + [0, K^{-7-\nu}]) + B_{\varepsilon/10} \subset S_{3\varepsilon/2}. \tag{7.6}$$

Note that the addition of $[0, K^{-7-\nu}]$ can be achieved by slightly shrinking J; we will need it for future convenience. The addition of an $\varepsilon/10$-ball is possible due to replacement of S_ε by $S_{3\varepsilon/2}$. Also recall that, since $K \gg B^C$, any fixed power of B can be absorbed into an extra factor K^{-1}.

For a fixed ξ, the condition (7.6) defines a semi-algebraic set Ω in the (ω, θ_1)-plane, of degree $\leqslant B^C$, whose projection onto the ω axis contains I_2. Since Ω has at most B^C connected components, there is a connected component $\Omega_0 \subset \Omega$ whose projection onto the ω axis contains an interval

$$I_3 \subset I_2, \quad |I_3| \geqslant B^{-C}|I_2| \geqslant K^{-6-\nu}.$$

Hence, using (sa9), one can select a (piecewise algebraic of degree $\leqslant B^C$) curve $\mathcal{C}_\xi \subset \Omega_0$ such that

$$\mathrm{Proj}_\omega \mathcal{C}_\xi = I_3, \quad |I_3| \leqslant \mathrm{diam}\,\mathcal{C}_\xi \leqslant \mathrm{length}(\mathcal{C}_\xi) \leqslant B^C.$$

Since $\mathrm{diam}\,\mathcal{C}_3 \geqslant K^{-5-\nu}$, one can find a large number of separated points on $\mathcal{C}_3$:

$$\xi_s \in \mathcal{C}_3, \quad 0 \leqslant s < K^{10+2\nu}; \quad |\xi_s - \xi_{s'}| \geqslant K^{-20-2\nu}, \quad s \neq s'.$$

For each ξ_s, one can construct a curve $\mathcal{C}_{\xi_s}$ described above. Recall that $|J| \geqslant \frac{1}{2}K^{-\nu}$, $\operatorname{diam}\mathcal{C}_{\xi_s} \geqslant K^{-6-\nu}$, which implies

$$|\mathcal{C}_{\xi_s} + J\xi_s + B_{\varepsilon/200}| \geqslant K^{-8-2\nu}\varepsilon. \tag{7.7}$$

The set in the left hand side is contained in $S_{3\varepsilon/2}$, and $|S_{2\varepsilon}| \leqslant B^C\varepsilon$ (due to (sa4)). Hence, since $K \gg B^C$, the sets in the left hand side of (7.7) must have significant overlap. That is, for some $s \neq s'$, we would have

$$|(\mathcal{C}_{\xi_s} + J\xi_s + B_{\varepsilon/200}) \cap (\mathcal{C}_{\xi_{s'}} + J\xi_{s'} + B_{\varepsilon/200})| \geqslant K^{-20-4\nu}\varepsilon. \tag{7.8}$$

Define the following subset of $\mathcal{C}_{\xi_s} + J\xi_s$:

$$W = \{x \in \mathcal{C}_{\xi_s} + J\xi_s \colon \operatorname{dist}(x, \mathcal{C}_{\xi_{s'}} + J\xi_{s'}) \leqslant \varepsilon/100\}.$$

Clearly, W is a two-dimensional semi-algebraic set of degree $\leqslant B^C$, and

$$(\mathcal{C}_{\xi_s} + J\xi_s + B_{\varepsilon/200}) \cap (\mathcal{C}_{\xi_{s'}} + J\xi_{s'} + B_{\varepsilon/200}) \subset W + B_{\varepsilon/100}.$$

Using (sa4), we have the following bound:

$$|(\mathcal{C}_{\xi_s} + J\xi_s + B_{\varepsilon/200}) \cap (\mathcal{C}_{\xi_{s'}} + J\xi_{s'} + B_{\varepsilon/200})| \leqslant B^C\varepsilon|W|_2,$$

which, combined with (7.8), implies a lower bound on the 2D measure of W in the right hand side, and a consequent 1D bound (since, by Fubini, $\mathcal{C}_\xi$ cannot be longer than B^C):

$$|W|_2 \geqslant K^{-21-4\nu}; \quad |(x + J\xi_s) \cap W|_1 \geqslant K^{-22-4\nu} \text{ for some } x \in \mathcal{C}_{\xi_s}.$$

Since both W and $x + J\xi_s$ are semi-algebraic of degree $\leqslant B^C$, the segment $x + J\xi_s$ contains a smaller segment:

$$x + J_1\xi_s \subset (x + J\xi_s) \cap W, \quad |J_1| \geqslant K^{-23-4\nu}.$$

However, each point $y \in (x + J_1\xi_s)$ is also $\varepsilon/100$-close to $C_{\xi_{x'}} + J\xi_{s'}$. Using (7.6), we can conclude that y will not leave $S_{2\varepsilon}$ after a displacement smaller than $K^{-7-\nu}$ in the direction $\xi_{s'}$. Hence, the point $x \in S_{3\varepsilon/2}$ satisfies

$$x + J_1\xi_s + [0, K^{-7-\nu}]\xi_{s'} \subset S_{2\varepsilon}, \tag{7.9}$$

which is a piece of a plane (convex hull of a triangle). The bound $|\xi_s - \xi_{s'}| \geqslant K^{-20-2\nu}$ implies that the angles of the triangle (7.9) cannot be too small, and hence it must contain a horizontal line segment of length, say, $K^{-100-10\nu} \gg e^{-B^\rho}$. We have obtained a contradiction with the assumption (7.3), which completes the proof for the case of A_ξ^1; the argument for A_ξ^2 is the same. $\qquad\square$

The following lemma establishes a stronger result if one allows for $B^C K$ rather than B^C exceptional directions.

LEMMA 7.2. *Let S_ε and $S_{2\varepsilon}$ satisfy the assumptions from the beginning of the section. Fix $\nu > 0$ and $0 < c_1 < c_2$. There are at most $B^C K$ pairs $k = (k_1, k_2)$, $c_1 K \leqslant |k| \leqslant c_2 K$, such that S_ε intersects some line segment in the direction ξ_k in a set of 1D measure $\geqslant K^{-\nu}$.*

Proof. For $\xi \in \mathbb{S}^2$, denote by

$$M_\varepsilon(\xi) = \max_{x \in \mathbb{R}^3} |S_\varepsilon \cap (x + [0,1]\xi)|_1$$

the measure of the largest 1D section of S_ε in the direction ξ. We first establish that, for all $\delta > 0$,

$$\|M_\varepsilon\|_{L^{10/3}(\mathbb{S}^2)} \leqslant C(\delta) B^C \varepsilon^{1/5 + \delta}. \tag{7.10}$$

Suppose $L \subset \ell \subset S_\varepsilon$ is a subset of a line segment ℓ. Denote by L_ε the round cylinder with axis L and the base being a disc of radius ε:

$$L_\varepsilon = \{x \in \mathbb{R}^3 \colon \operatorname{dist}(x, \ell) \leqslant \varepsilon, \ \operatorname{Proj}_\ell x \in L\}.$$

Clearly, $L_\varepsilon \in S_{2\varepsilon}$. Now let $f_\varepsilon(\xi)$ be the maximal possible volume of cylinders $L_\varepsilon(\xi)$ among all line segments in the direction ξ. With this definition, $f_\varepsilon(\xi) \leqslant f_\varepsilon^*(\xi)$, where $f_\varepsilon^*(\xi)$ is the Kakeya maximal function of the indicator function of $\mathbf{1}_{S_{2\varepsilon}}$, see [Wol95]. Recall also that (sa4) implies $|S_{2\varepsilon}| \leqslant B^C \varepsilon$. The bound (7.10) now follows from the following bound established in [Wol95, Theorem 1]:

$$\|f_\varepsilon^*\|_{L^q(\mathbb{S}^2)} \leqslant C(\delta) \varepsilon^{1 - 3/p - \delta} \|\mathbf{1}_{S_{2\varepsilon}}\|_{L^p(\mathbb{R}^3)}, \quad 1 \leqslant p \leqslant 3, \quad q \leqslant 2p',$$

with $p = 5/2$, $q = 10/3$. As a consequence,

$$\int_{\mathbb{S}^2} M_\varepsilon(\xi)^{10/3} \, d\xi \leqslant C(\delta) B^C \varepsilon^{2/3 - \delta}. \tag{7.11}$$

Let $W_0 = \{\xi \in \mathbb{S}^2 \colon |M_\varepsilon(\xi)| \geqslant K^{-\nu}\}$. (7.11) implies the following bound on the surface measure:

$$|W_0|_2 \leqslant C(\delta) K^{\frac{10}{3}(\nu + 1)} \varepsilon^{2/3 - \delta}, \tag{7.12}$$

Since $S_{2\varepsilon}$ intersects any line segment in at most N^C intervals, $\xi \in W_0$ implies that $S_{2\varepsilon}$ contains a line segment of length $N^{-C} K^{-\nu}$ in the direction ξ. Let $W \supset W_0$ be the set of all directions with the latter property. Clearly, W satisfies (7.12) with extra factor B^C in the right hand side; also, W is semi-algebraic of degree $\leqslant B^C$ (see (sa8)). Due to (7.12), W is contained in $\varepsilon^{1/3 - o(1)}$-neighborhood of its boundary, and, due to (sa6) and with a possible loss of another B^C factor, *can be replaced by the $\varepsilon^{1/4}$-neighborhood of a smooth algebraic curve Γ in $\mathbb{S}^2$, $\deg \Gamma \leqslant B^C$.* All directions ξ_k that we care about, must be in W.

For each ξ in W, denote by $\tilde{\xi}$ the point of intersection of the line $\mathbb{R}\xi$ with the plane $\omega = K^{-1}$ in $\mathbb{R}^3$. This map will transform Γ and W into $\widetilde{\Gamma}, \widetilde{W}$, respectively, and, since $W \subset \Gamma + B_{\varepsilon^{1/4}}$, we have

$$\widetilde{W} \subset \widetilde{\Gamma} + B_{CK\varepsilon^{1/4}}.$$

Moreover, for each

$$\xi_k = \frac{(1, k_1, k_2)}{\sqrt{1 + k_1^2 + k_2^2}}, \quad c_1 K \leqslant |k| \leqslant c_2 K,$$

we have

$$\mathrm{dist}\left(\left(\frac{k_1}{K}, \frac{k_2}{K}\right), \widetilde{\Gamma}\right) \leqslant CK\varepsilon^{1/4}.$$

Note also diam $\widetilde{\Gamma} \leqslant c_2'$ (where c_2' depends on c_2). Using (sa4) and $\varepsilon \ll K^{-4}$, we can cover $\widetilde{W}$ by $B^C K$ (open) squares of size K^{-1}. Since each square contains at most one point $(k_1/K, k_2/K)$, this implies the required bound. $\square$

8 Elimination of Double Resonances

In this section, we will assume that $C_{\mathrm{dio}}, \delta_{\mathrm{dio}}$ are fixed. Their choice will be explained in the beginning of the next section. Denote the finite-scale set of Diophantine frequencies

$$\begin{aligned}
\mathrm{DC}(N) &= \{\omega \in [0,1] \colon \|k\omega\| \geqslant C_{\mathrm{dio}}|k|^{-1-\delta_{\mathrm{dio}}}, \text{ for } 1 \leqslant |k| \leqslant N\} \supset \mathrm{DC}_{C_{\mathrm{dio}},\delta_{\mathrm{dio}}} \\
&= \bigcap_N \mathrm{DC}(N).
\end{aligned}$$

Note that, for finite N, we have dropped the dependence on $C_{\mathrm{dio}}, \delta_{\mathrm{dio}}$ from the notation. In this section and later, we will assume that $\gamma > 10$. The main object of study in this section is the "bad set"

$$\mathcal{E}_N^r = \{(\omega, \theta_1, \theta_2) \colon \exists j \colon (\theta_1, \theta_2) \in \mathcal{B}_U^{\gamma, b}([-N, N]^2, E_j(\omega)), \ \omega \in \mathrm{DC}(N^r)\}, \qquad (8.1)$$

where $E_j(\omega)$ runs over all eigenvalues of $H_{[-M,M]^2}(0,0)$ over all possible translations of U and $N \leqslant M \leqslant N^r$. Note that the set of $E_j(\omega)$ depends on N, ω, and r, and has cardinality $\leqslant N^C$. In this section, the constants may depend on r. Here b, γ are obtained from Corollary 6.5, so that, for $N \geqslant N_0$, $\mathcal{E}_N^r$ satisfies the conclusion of Corollary 6.5: every section of $\mathcal{E}_N^r$ by a unit line segment perpendicular to ω axis has measure $\leqslant e^{-N^\rho}$ (we can slightly decrease ρ to absorb the factor N^C which comes from considering multiple energies $E_j(\omega)$ at the same time).

The set $\mathcal{E}_N^r$ is not semialgebraic. However, as discussed in Lemma 6.2, one can replace it by a smaller set $\mathcal{E}_N^{r,\mathrm{alg}}$, such that, for $(\omega, \theta_1, \theta_2) \notin \mathcal{E}_N^{r,\mathrm{alg}}$, the Green's function $G_{[-N,N]^2}(\theta_1, \theta_2, E)$ satisfies (3.12), (3.13) with a non-essential loss of factor of 10. In other words, one treat the points of the larger set $\mathcal{E}_N^r \setminus \mathcal{E}_N^{r,\mathrm{alg}}$ as "good". Moreover, Lemma 6.2 also implies that we can assume

$$\mathcal{E}_N^{r,\mathrm{alg}} + B_{e^{-10\gamma N}} \subset \mathcal{E}_N^r.$$

In order to meet the assumptions of Lemmas 7.1 and 7.2, we will need some further preparations. Consider the layer expansion of $\mathcal{E}_N^{r,\mathrm{alg}}$ provided by (sa10) with

$\varepsilon = e^{-20\gamma N}$. Each dyadic layer $\mathcal{E}_N^{r,\text{alg}}(k)$ (in the notation of (sa10)) is a $2^{-(k+1)}$-neighborhood of some semialgebraic subset, with the property that the neighborhood of double size is still contained in $\mathcal{E}_N^{r,\text{alg}}$. Moreover, these layers cover $\mathcal{E}_N^{r,\text{alg}}$, except maybe for $e^{-20\gamma N}$-neighborhood of the boundary $\partial \mathcal{E}_N^{r,\text{alg}}$. The total number of layers is N^C, where we are still using the convention that C may depend on r. Note that the points in the neighborhood of the boundary will still satisfy (3.12) and (3.13) with, say, a loss of factor of 20. Hence, we can further decrease $\mathcal{E}_N^{r,\text{alg}}$ and then consider a single dyadic layer $\mathcal{L}$:

$$\mathcal{L} = \mathcal{S} + B_\delta, \quad \mathcal{L} + B_\delta = \mathcal{S} + B_{2\delta} \subset \mathcal{E}_N^{r,\text{alg}}, \quad e^{-20\gamma N} \leqslant \delta \leqslant e^{-N^\rho}, \quad \dim \mathcal{S} \leqslant 2. \tag{8.2}$$

Let us define $\widetilde{\mathcal{B}}_U^{\gamma,b}([-M,M]^2, E)$ to be set of $\theta \in [0,1]^2$ satisfying the same conditions as $\mathcal{B}_U^{\gamma,b}([-M,M]^2, E)$, relaxed by a factor of 100. That is, for all possible translations of U, we assume for $(\omega, \theta_1, \theta_2) \in [0,1]^3 \setminus \widetilde{\mathcal{B}}_U^{\gamma,b}([-M,M]^2, E)$:

$$\|G_{[-M,M]^2}(\theta_1, \theta_2, E)\| \leqslant 100\lambda^{-1} e^{\sigma(\Lambda)^b},$$

$$|G_{[-M,M]^2}(\theta_1, \theta_2, E)(n_1, n_2)| \leqslant 100 e^{-\gamma|n_1 - n_2|} \text{ for all } n_1, n_2 \in \Lambda, \ |n_1 - n_2| \geqslant \frac{1}{4}\sigma(\Lambda).$$

The following theorem is the key step in proving localization. It shows that removal of a small set of frequencies will exclude double resonances.

Theorem 8.1. *Fix* $0 < c_1 < c_2$. *There exists* $N_0 = N_0(r, \lambda, v, \gamma, b, \rho, c_1, c_2, C_{\text{int}}, m_{\text{int}})$ *such that, for*

$$N \geqslant N_0, \quad K = e^{(\log N)^{2/\rho}}, \quad M = [(\log N)^{3/2\rho}], \tag{8.3}$$

there exists

$$\Omega_N^{\text{bad}} \subset \text{DC}(N^r), \quad |\Omega_N| \leqslant e^{-\frac{1}{2}(\log N)^{3/2}}, \tag{8.4}$$

such that, for every $\omega \in \text{DC}(N^r) \setminus \Omega_N^{\text{bad}}$ *and* $c_1 K \leqslant |k| \leqslant c_2 K$, *the following is true. Suppose* $E = E_j(\omega)$ *in the notation of* (8.1). *Then*

$$(\omega, k_1\omega, k_2\omega) \notin \widetilde{\mathcal{B}}_U^{\gamma,b}([-M,M]^2, E) \cap \widetilde{\mathcal{B}}_U^{\gamma,b}([-N,N]^2, E).$$

In other words, either M-*box or* N-*box around the point* (k_1, k_2) *is good at the energy* E.

Theorem 8.1 establishes absence of double resonances. If the energy is close to an eigenvalue of N-box centered at the origin, then, at the distance $\approx K$ from the origin, any point is a center of a good box with respect to that energy. The technical difference is that the size of the good box is variable, although it does not affect the proof too much.

Proof. Ideally, we would like to show that one can remove a small set of frequencies ω such that, for all k_1, k_2 in the considered range, the vector $(\omega, \{k_1\omega\}, \{k_2\omega\})$ avoids the set $\mathcal{E}_N^{r,\mathrm{alg}}$ defined earlier in this section. Using the techniques from Sect. 7, we will be able to establish this for most pairs (k_1, k_2). To deal with the remaining pairs, we would have to consider smaller boxes of size M.

Let $L = \lfloor 3c_2 K + 1 \rfloor$. Let us split $[0,1]$ into L intervals of size L^{-1}, and take $\omega = \omega_0 + \Delta\omega$, $|\Delta\omega| < L^{-1}$, $\omega_0 = l/L$. Each interval contains at most 3 discontinuity points of functions $\{k_1\omega\}$, $\{k_2\omega\}$. Hence, it defines at most three line segments in $[0,1]^3$ in the direction ξ_k, each of the form

$$(\omega_0 + \Delta\omega, k_1\Delta\omega + \theta_1, k_2\Delta\omega + \theta_2), \tag{8.5}$$

parametrized by $\Delta\omega$.

Let $\mathcal{L}$ be defined as in (8.2). Suppose one of the segments (8.5) has large intersection with $\mathcal{L}$ (that is, the intersection contains a segment of size $K^{-\gamma}$). Let us continue the line containing this segment. Since either $|k_1| \geqslant \frac{1}{2}c_1 K$ or $|k_2| \geqslant \frac{1}{2}c_1 K$, the intersection of that line with one of the planes $\theta_1 = 0$ or $\theta_2 = 0$ has coordinates

$$(\omega_0 + O(1)K^{-1}, 0, \theta_2') \quad \text{or} \quad (\omega_0 + O(1)K^{-1}, \theta_1', 0),$$

for some θ_1', θ_2'. Hence, for each ω_0 such that the segment (8.5) has large intersection with $\mathcal{L}$, the semi-algebraic set

$$\mathrm{Proj}_\omega A_{\xi_k}^1 \cup \mathrm{Proj}_\omega A_{\xi_k}^2 \subset [0,1] \tag{8.6}$$

defined in Lemma 7.1, contains a point $\omega_0 + O(1)K^{-1}$. Since the set (8.6) has degree $\leqslant N^C$, there are two possibilities for each pair $k = (k_1, k_2)$:

- There are $\leqslant N^C$ possible values of l with $\omega_0 = l/K$ such that the set (8.6) contains a point $\omega_0 + O(1)K^{-1}$.
- If there are too many possible values of l, then a connected component of one of the sets (8.6) contains two points l_1/K and l_2/K with $l_1 \neq l_2$, and therefore contains an interval of length $\geqslant K^{-1}$. In this case, the pair (k_1, k_2) is singular in the sense of Lemma 7.1 with $S_\varepsilon = \mathcal{L}$, $\varepsilon = \delta$ defined in (8.2).

Using Lemmas 7.1 and 7.2, one can split the pairs (k_1, k_2) and intervals (8.5) into the following groups.

(1) The pairs (k_1, k_2) such that none of the layers $\mathcal{L}$ contain intervals of length $K^{-\nu}$ in the direction ξ_k (that is, there is no layer for which the condition in Lemma 7.2 holds). Each of those pairs and each of the corresponding $3L$ line segments (8.5) can be avoided by removing a set of frequencies of measure $\leqslant N^C K^2 L K^{-\nu-1}$, which can be made small by choosing a large ν. This settles the claim for the majority of pairs (k_1, k_2), except for at most $N^C K$ pairs.

(2) At most $N^C K$ pairs that satisfy the condition in Lemma 7.2 for some layer $\mathcal{L}$ and are non-singular with respect to the same layer (in the sense of Lemma 7.1). The latter implies that each pair gives only $\leqslant N^C$ values of ω_0, and hence $\leqslant N^C$ problematic intervals (8.5).

(3) At most N^C pairs (k_1, k_2) which are singular with respect to some layer.

To deal with (2) and (3), we will use M-boxes. By perturbation of the diagonal entries,

$$|E_j(\omega) - E_j(\omega_0)| \leqslant 4\lambda\frac{N^r}{L}.$$

(assuming we are taking the eigenvalues in the same order and are considering the same translations of U). If a box of size M satisfies (3.12), (3.13), then both of these properties are preserved (up to a factor of 10) under a perturbation of energy of size $\leqslant e^{-100\gamma M}$, see Proposition 6.1. Hence, in order to "freeze" the dependence on ω and replace $E_j(\omega)$ by $E_j(\omega_0)$, it would be sufficient to impose

$$4\lambda\frac{N^r}{L} \leqslant e^{-100\gamma M}, \tag{8.7}$$

which is satisfied with our choice (8.3) of K, L, M. Suppose now that $\omega = \omega_0 + \Delta\omega$. The condition (8.7) implies that, instead of considering $G_{[-M,M]^2}(\omega, k_1\omega, k_2\omega, E_j(\omega))$, one can consider $G_{[-M,M]^2}(\omega_0, k_1\omega, k_2\omega, E_j(\omega_0))$. This Green's function satisfies the statement of Corollary 6.5, and hence, the bad set for this function can be avoided by removing $CK^{-1}M^C e^{-M^\rho}$ set of ω from each interval $[\omega_0, \omega_0 + 1/L)$. For the singular pairs (3), this gives a total of $CN^C M^C e^{-M^\rho}$ removed frequencies. For non-singular pairs (2), we have $KN^C M^C K^{-1}e^{-M^\rho}$ (there are $N^C K$ pairs remaining, and N^C intervals for each pair). In both cases, the removed set of frequencies has measure bounded by, say, $e^{-\frac{2}{3}M^\rho}$, which is in line with (8.4). $\qquad\square$

REMARK 8.2. Theorem 8.1 is used in the next section to establish localization for zero phase. One can formulate an analogue for arbitrary phase, with the same proof. Fix θ_1', θ_2' and, similarly to (8.1), define

$$\mathcal{E}_N^r(\theta_1', \theta_2') = \{(\omega, \theta_1, \theta_2) \colon \exists j \colon (\theta_1, \theta_2) \in \mathcal{B}_U^{\gamma,b}([-N,N]^2, E_j(\omega, \theta_1', \theta_2')), \omega \in \mathrm{DC}(N^r)\},$$

where $E_j(\omega, \theta_1', \theta_2')$ are eigenvalues of $H_{[-M,M]^2}(\theta_1', \theta_2')$ instead of $H_{[-M,M]^2}(0,0)$. Then the conclusion of Theorem 8.1 holds for $\mathcal{E}_N^r(\theta_1', \theta_2')$, with the same bounds, but Ω_N^{bad} will now depend on (θ_1', θ_2'). For the case of general phases, we will use the notation $\Omega_N^{\mathrm{bad}}(\theta_1', \theta_2')$.

COROLLARY 8.3. *Under the assumptions of Theorem* 8.1, *the set of* $(\theta_1', \theta_2', \omega)$ *satisfying* $\omega \in \Omega_N^{\mathrm{bad}}(\theta_1', \theta_2')$ *has measure* $\leqslant e^{-\frac{1}{4}(\log N)^{3/2}}$. *Moreover, if* $L \subset [0,1]^2$ *is a unit line segment, then the set of* $(\omega, \theta_1', \theta_2')$ *with* $(\theta_1', \theta_2') \in L$ *and* $\omega \in \Omega_N^{\mathrm{bad}}(\theta_1', \theta_2')$, *has 2D measure* $\leqslant e^{-\frac{1}{4}(\log N)^{3/2}}$ *in* $[0,1] \times L$.

Proof. Both follow from the Fubini theorem and Markov's inequality, as the sets under consideration have small sections by any line $(\theta_1', \theta_2') = \mathrm{const}$. One can check that measurability follows from the construction in Theorem 7.1. $\qquad\square$

REMARK 8.4. A short range interaction $U(n_1 - n_2) = f(n_1 - n_2)$ has translation invariance in the diagonal direction. That is, it is natural to consider $\theta = \frac{1}{2}(\theta_1 + \theta_2)$, $\eta = \frac{1}{2}(\theta_1 - \theta_2)$, make θ the ergodic parameter, and η an external parameter. In this case, the conclusion of Corollary 8.3 holds for any fixed η and shows that the set of "bad" pairs (ω, θ) has small measure in $[0,1]^2$.

9 Proof of Localization

Once all ingredients are in place, the proof of localization is fairly standard. Fix some $0 < \delta_{\mathrm{dio}} < 1/24$ and $\gamma > 10$. Let $\varepsilon_{\mathrm{freq}} > 0$. Pick $C_{\mathrm{dio}} = C_{\mathrm{dio}}(\varepsilon_{\mathrm{freq}}) > 0$ such that

$$|\mathrm{DC}_{C_{\mathrm{dio}}, \delta_{\mathrm{dio}}} \cap [0,1]| \geqslant 1 - \varepsilon_{\mathrm{freq}}.$$

Choose other parameters in a way that (6.5) is satisfied and, based on that choice, choose a large λ_0 so that the conclusion of Corollary 6.5 is satisfied. Fix some initial phase (θ_1, θ_2). Consider the sets $\Omega_N^{\mathrm{bad}}(\theta_1, \theta_2)$ defined in Theorem 8.1 and Remark 8.2. Assume that

$$\begin{aligned}
\omega \in \Omega(U, \lambda, \theta_1, \theta_2) \ &:= \mathrm{DC}_{C_{\mathrm{dio}}, \delta_{\mathrm{dio}}} \setminus \limsup_{N \to \infty} \Omega_N^{\mathrm{bad}}(\theta_1, \theta_2) \\
&= \mathrm{DC}_{C_{\mathrm{dio}}, \delta_{\mathrm{dio}}} \setminus \bigcap_{N_0 \geqslant 1} \bigcup_{N \geqslant N_0} \Omega_N^{\mathrm{bad}}(\theta_1, \theta_2).
\end{aligned} \tag{9.1}$$

Choose N sufficiently large enough in order to apply Corollary 6.5, both for the scales M and N from Theorem 8.1.

We will prove Theorems 2.2 and 2.3 at the same time, by showing that the operator $H(\theta_1, \theta_2)$ has Anderson localization, as long as Corollary 6.5 is applicable. In the case of Theorem 2.3, this would mean extra assumption (2.4).

Suppose that ψ is a generalized eigenvector of $H(\theta_1, \theta_2)$, that is, a formal solution of the eigenvalue equation

$$H(\theta_1, \theta_2)\psi = E\psi, \quad \psi(0) = 1, \quad |\psi(n)| \leqslant C_\psi(1 + |n|). \tag{9.2}$$

Our goal is to show that $\psi \in \ell^2(\mathbb{Z}^2)$. At this point, we are allowed to pick large N depending, in particular, on ψ and ω. We will always assume that N and M are large enough, so that $\Omega \notin \Omega_l^{\mathrm{bad}}(\theta_1, \theta_2)$ for all $l \geqslant \min(M, N)$.

From Proposition 5.1, there exists a large constant r_0 (independent of N for N large) such that, among $n \in [-N^{r_0}, N^{r_0}]^2$, there are at most $N^{7r_0/8}$ values of n such that the box $[-N, N] + n$ fails (3.12) or (3.13) with a factor of 100. Hence, there exists $R \in [N^{r_0/2}, N^{r_0}]$ such that the annulus

$$A = [-R - N^{r_0/4}, R + N^{r_0/4}]^2 \setminus [-R + N^{r_0/4}, R - N^{r_0/4}]^2 \tag{9.3}$$

consists of good points (n_1, n_2), that is, the Green's function of any box of size N centered in A satisfies (3.12), (3.13) with a factor of 100, for all possible translations of U.

We will use Poisson's formula. Suppose, ψ satisfies (9.2). If Λ is an elementary region and $m \in \Lambda \backslash \partial \Lambda$, then

$$\psi(m) = \sum_{n \in \Lambda,\, n' \in \mathbb{Z}^2 \backslash \Lambda,\, |n-n'|=1} G_\Lambda(\theta_1, \theta_2, E)(m, n)\psi(n'). \tag{9.4}$$

Suppose $m \in \partial[-R, R]^2$. Apply (9.4) with $\Lambda = m + [-N, N]$:

$$|\psi(m)| \leqslant 100 e^{-\gamma N} \sum_{n' \in \partial(\mathbb{Z}^2 \backslash \Lambda)} |\psi(n')|, \tag{9.5}$$

and keep repeating it for each $\psi(n')$. The size of A guarantees that one can apply (9.5) at least $N^{r_0/5}$ times without leaving the annulus A, which brings an estimate

$$|\psi(m)| \leqslant C_\psi (1000N)^{N^{r_0/5}} e^{-\gamma \cdot N \cdot N^{r_0/5}} (1 + N^{r_0}). \tag{9.6}$$

Clearly, the exponential factor dominates all other factors for large N (depending on ψ). This implies

$$|\psi(m)| \leqslant e^{-\frac{1}{2}N \cdot N^{r_0/5}}, \quad m \in \partial[-R, R]^2, \quad N^{r_0/2} \leqslant R \leqslant N^{r_0}, \quad N \geqslant N_0(C_\psi),$$

and therefore

$$\mathrm{dist}(E, \sigma(H_{[-R,R]^2}(\theta_1, \theta_2))) \leqslant e^{-N^2}. \tag{9.7}$$

Let E_0 be the closest to E eigenvalue of $H_{[-R,R]^2}(\theta_1, \theta_2)$, and suppose $\omega \in \mathrm{DC}(N^2) \backslash \Omega_N(\theta_1, \theta_2)$ (which is valid due to (9.1), maybe after increasing N). Theorem 8.1 implies that, for $c_1 K \leqslant |k| \leqslant c_2 K$, the Green's function $G_\Lambda(\theta_1, \theta_2, E_0)$ satisfies (3.12) and (3.13) with a factor of 100, for $\Lambda = k + [-N, N]^2$ or $\Lambda = k + [-M, M]^2$. Resolvent identity from Proposition 6.1 together with (9.7) implies that (3.13) will also hold at the energy E with a factor of 200, which can be summarized as follows (at this point, we only care about the off-diagonal decay):

$$|G_\Lambda(\theta_1', \theta_2', E)(n_1, n_2)| < 200 e^{-\gamma|n_1 - n_2|} \text{ for all } n_1, n_2 \in \Lambda, |n_1 - n_2| \geqslant \frac{1}{4}\sigma(\Lambda),$$

$$\Lambda = k + [-N, N]^2 \text{ or } \Lambda = k + [-M, M]^2,$$

$$c_1 K \leqslant |k| \leqslant c_2 K, \quad K = e^{(\log N)^{2/\rho}}, \quad M = [(\log N)^{3/2\rho}], \quad 0 < \rho < 1. \tag{9.8}$$

Clearly, the range of k specified in (9.8) covers the whole $\mathbb{Z}^d$ except for finitely many points, as long as $c_2 > c_1 > 0$. We will choose a smaller annulus by taking

$$c_1 < c_1 + \beta < c_1' < c_2' < c_2 - \beta < c_2$$

and will estimate $\psi(k)$, $c_1' K \leqslant |k| \leqslant c_2' K$, in the same way as we obtained (9.6); however, one needs to be more careful about where to stop the expansion, as the box size is now variable. We will repeat applying (9.4) until the total exponential factor gained from iterations of (9.5) exceeds $e^{-\gamma \beta K}$ (which guarantees that we are staying within the annulus). Thus, at the end of the expansion, the "depth" of each

term is determined by the number of N-boxes along the history of applying (9.4) in order to obtain that term. We will provide a coarse bound: for each $0 \leqslant l \leqslant \beta K/N$, we will assume that each term had exactly l N-boxes along the way, and then take the sum over l.

Suppose that $l = 0$, so that each time we have to use an M-box. Then, the total number of steps is $[\beta K/M]$ (rounding errors can be easily absorbed into the final estimate), which leads to the contribution of these terms into $\psi(k)$ estimated by

$$(1000M)^{\beta K/M} \exp\{-\gamma \beta K\} C_\psi (1 + |K|) \leqslant e^{-\frac{1}{2}\gamma \beta K}, \quad \text{assuming} \quad \frac{\log M}{M} \ll \gamma, \quad (9.9)$$

which implies, under the assumptions that all boxes are M-boxes, that

$$|\psi(k)| \leqslant C_\psi \exp\left\{-\frac{1}{2}\gamma \beta c_1' c_1^{-1} |k|\right\}. \quad (9.10)$$

Now, suppose there is only one N-box along the course of applying (9.4), and all other boxes are M-boxes. The total contribution from these configurations has the same exponential factor (by construction), has the factor $(1000M)^{N/M}$ replaced by a factor of $1000N$, and adds an extra combinatorial factor K/M. Thus, (9.9) at $l = 1$ is bounded by (9.9) at $l = 0$, multiplied by the factor

$$1000M^{-1}KN(1000M)^{-N/M} \ll 1.$$

Each additional N-box generates an extra similar (in fact, smaller) factor, which implies that the contribution from the terms with $l > 0$ is dominated by the term with $l = 0$, and (9.10) is true in all cases.

REMARK 9.1. Using Corollary 8.3, one can establish Anderson localization for a full measure set of (θ_1, θ_2), or for a subset of full 1D measure of some line segment L, by considering the sets of $(\theta_1, \theta_2, \omega)$ constructed in Corollary 8.3 instead of sets of ω at fixed (θ_1, θ_2). None of the three cases (fixed phase, full measure subset of $[0, 1]^2$, full measure subset of L) seem to directly imply one another.

One can also approach this argument from the measure-theoretic point of view. We have essentially shown that, for each $(\theta_1, \theta_2) \in [0, 1]^2$, the set of frequencies ω such that the operator $H(\theta_1, \theta_2)$ has Anderson localization for all possible translations of U, has full measure in $\mathrm{DC}_{C_{\mathrm{dio}}, \delta_{\mathrm{dio}}}$, and hence the set of $(\omega, \theta_1, \theta_2)$ with this property has full measure in $\mathrm{DC}_{C_{\mathrm{dio}}, \delta_{\mathrm{dio}}} \times [0, 1]^2$, which allows to apply Fubini's theorem. However, the argument in the previous paragraph is useful in case we do not have an explicit arithmetic condition modulo a zero measure set of frequencies.

10 Appendix

For the convenience of the reader, we include the proof of Proposition 6.4, mostly following [BGS02] with appropriate modifications. As stated in [BGS02], the last steps of the proof can be replaced by an application of the following version of

Cartan's lemma, which became standard in more recent papers. For the proof, see [Bou05, Proposition 14.1], or [Bou07, Lemma A.2]:

PROPOSITION 10.1. *Let $A(\sigma)$ be a self-adjoint $D \times D$ matrix function of a real parameter $\sigma \in [-\delta, \delta]$, satisfying the following conditions:*

(1) *$A(\sigma)$ is real analytic in σ, and admits a holomorphic extension to the strip $(-\delta_1, \delta_1) + i(-\delta_2, \delta_2)$, satisfying in that strip*

$$\|A(z)\| \leqslant B_1.$$

(2) *For each $\sigma \in [-\delta_1, \delta_1]$, there is a subset $\Lambda \subset [1, D]$, such that*

$$|\Lambda| < D_0, \quad \| \left(R_{[1,N]\setminus\Lambda} A(\sigma) R_{[1,N]\setminus\Lambda} \right)^{-1} \| < B_2.$$

(3) *$|\{\sigma \in [-\delta_1, \delta_1] \colon \|A(\sigma)^{-1}\| > B_3\} < 10^{-3}\delta_2(1 + B_1)^{-1}(1 + B_2)^{-1}|.$*

Then, for any $\varkappa < (1 + B_1 + B_2)^{-10D_0}$, we have

$$|\{\sigma \in (-\delta_1/2, \delta_1/2) \colon \|A(\sigma)^{-1}\| > \varkappa^{-1}\}| < \exp \frac{c \log \varkappa}{D_0 \log(D_0 + B_1 + B_2 + B_3)}. \quad (10.1)$$

10.1 Proof of Proposition 6.4 Choose some $M \in [N_0, N_1]$ and define $N = [M_0^{C_1}]$, where C_1 will be a large constant. Let $\Lambda_0 \in \mathcal{ER}(N)$. Consider the partition

$$\Lambda_0 = \cup_\alpha \Lambda_\alpha, \quad \Lambda_\alpha = Q_\alpha \cap \Lambda_0, \quad (10.2)$$

where

$$Q_\alpha \in [-M_0, M_0]^2 + 2M_0\mathbb{Z}^2$$

is a translation of an M_0-cube by an integer vector multiple of $2M_0$, and the union in (10.2) runs over non-empty Λ_α. It is explained in [BGS02] that $\Lambda_\alpha \in \mathcal{ER}(M')$, for some $M_0 \leqslant M' \leqslant 2M_0$, except maybe for at most five values of α.

Let

$$\mathcal{B} = \cup_{M_0 \leqslant M \leqslant 2M_0} \cup_{\Lambda \in \mathcal{ER}(M), \Lambda \subset [-M,M]^2} \mathcal{B}_U^{\gamma,b}(\Lambda, E). \quad (10.3)$$

In view of Lemma 6.2, one can replace $\mathcal{B}$ by a semi-algebraic set $\mathcal{A}$, $\deg \mathcal{A} \leqslant M_0^{C'_{\mathrm{int}}}$, satisfying the following: for any unit line segment $L \subset \mathbb{R}^2$,

$$|\mathcal{A} \cap L|_1 \leqslant M_0^{C'_{\mathrm{int}}} e^{-M_0^\rho}.$$

The replacement of $\mathcal{B}$ by $\mathcal{A}$ will add a (non-essential for the argument) factor 16 to (10.5). Suppose that $\overline{N}_0$ is chosen to be large enough, in order to have, for all N under consideration,

$$M_0^{C'_{\mathrm{int}}} e^{-M_0^\rho} \leqslant C_{\mathrm{dio}} N^{-1-\delta_{\mathrm{dio}}}. \quad (10.4)$$

Then, one can apply Theorem 5.1 and obtain

$$\#\{n \in [-N, N]^2 \colon \theta + n\omega \in \mathcal{A}\} \leqslant M_0^{C'_{\mathrm{int}}} C'_{\mathrm{dio}} N^{3/4 + 3\delta_{\mathrm{dio}}},$$

where C'_{int} is a new constant so that the factor $M_0^{C'_{\mathrm{int}}}$ absorbs $(\deg \mathcal{A})^C$ from Theorem 5.1. Let us call Λ_α *good* if $\theta + n\omega \notin \mathcal{A}$ for all $n \in \Lambda_\alpha$, and otherwise call Λ_α bad. Let

$$\Lambda_* = \cup_{\Lambda_\alpha \text{ is bad}} \Lambda_\alpha.$$

We have

$$\#\Lambda_* \leqslant M_0^{C'_{\mathrm{int}}} C'_{\mathrm{dio}} N^{3/4 + 3\delta_{\mathrm{dio}}}.$$

(here C'_{int} was increased by 2). Now, one can apply the same derivation as in [BGS02, equation (4.17)] and obtain

$$\|G_{\Lambda_0 \setminus \Lambda_*}(\theta_1, \theta_2, E)\| \leqslant 16\lambda^{-1} M_0^2 e^{(2M_0)^b}, \tag{10.5}$$

assuming $M_0^2 e^{-\gamma M_0/2} \leqslant 1/2$ (which can also be achieved by fixing a large $\overline{N}_0$). The bound (10.5) holds for all $\theta \in [0,1]^2$, with Λ_* depending on θ. The factor 16 comes from the factor 8 from Proposition 6.2.

In order to satisfy the second assumption of Cartan's lemma, we will need to introduce another scale: $M_1 = [(10M_0)^{1/\rho}]$. Suppose, all boxes $[-M_1, M_1]^2 + m$, $m \in \Lambda_0$ are good. Then, Lemma 4.1 implies (after replacing M_1^b by M_1 in the exponent and absorbing the factors):

$$\|G_{\Lambda_0}(\theta_1, \theta_2, E)\| \leqslant \lambda^{-1} e^{M_1}$$

assuming that $\theta \in [0,1]^2 \setminus \Theta$, where, for any unit line segment $L \subset \mathbb{R}^2$, we have

$$|\Theta \cap L|_1 \leqslant e^{-M_1^\rho/2}.$$

The argument behind applying Lemma 4.1 is exactly the same as the one that leads to [BGS02, equation (4.23)]. Note that, in out notation, Θ is a 2D set.

Let $L \subset \mathbb{R}^2$ be a line segment. Introduce new orthonormal coordinates

$$\theta(\eta_1, \eta_2) = (\theta_1(\eta_1, \eta_2), \theta_2(\eta_1, \eta_2)).$$

so that L is parametrized by $(\eta_1, 0)$ as $\eta_1 \in [0,1]$, and consider Green's functions $G_\Lambda(\theta_1(\eta_1, \eta_2), \theta_2(\eta_1, \eta_2), E)$. Denote by Ξ the set of η_1 such that $(\theta_1(\eta_1, 0), \theta_2(\eta_1, 0))$ runs over $\Theta \cap L$. Then $|\Xi| \leqslant e^{-M_1^\rho/2}$. Without loss of generality, one can increase the length of L by 2, so that

$$\|G_{\Lambda_0}(\theta_1(\eta_1, 0), \theta_2(\eta_1, 0), E)\| \leqslant e^{M_1}, \quad \forall \eta_1 \in [-1, 2] \setminus \Xi,$$

with the same bound on $|\Xi|$.

Apply Cartan's lemma (Proposition 10.1) with:

- $A(z) = \frac{1}{\lambda} R_{\Lambda_0}(H(\theta_1(z+1/2, 0), \theta_2(z+1/2, 0)) - E)R_{\Lambda_0}$.
- $[1, D]$ enumerates the points of $[-N, N]^2$.
- $\Lambda = \Lambda_0 \setminus \Lambda_*$.

- $D_0 = M_0^{C'_{\text{int}}} C'_{\text{dio}} (N^{3/4 + 3\delta_{\text{dio}}})$.
- $[-\delta_1, \delta_1] = [-3/2, 3/2]$, $\delta_2 = \delta_2(v)$ is the width of the strip to which v can be analytically extended.
- $B_1 = C(v, m_{\text{int}})$ is the norm bound for $A(z)$ in that strip.
- $B_2 = 2M_0^2 e^{(2M_0)^b}$.
- $B_3 = e^{M_1}$.

In order for the lemma to produce a meaningful bound, we would need to verify Assumption (3), that is, the estimate of the measure of Ξ; we need

$$e^{-5M_0} \approx e^{-M_1^\rho / 2} < \frac{\delta_2(v)}{1000(1 + C(v, m_{\text{int}}))(1 + 2M_0^2 e^{(2M_0)^b})} \approx C_1(v, m_{\text{int}}) M_0^{-2} e^{-(2M_0)^b}.$$

Since $b < 1$, this bound is clearly satisfied at a sufficiently large scale (depending only on v). In order to obtain the final estimate (6.8), we take $\varkappa = e^{-N^b}$. The condition preceding (10.1) becomes

$$e^{-N^b} < (1 + B_1 + B_2)^{-10D_0} = \left\{ 1 + C(v, m_{\text{int}}) + 2M_0^2 e^{(2M_0)^b} \right\}^{-M_0^{C'_{\text{int}}} C'_{\text{dio}} N^{3/4 + 3\delta_{\text{dio}}}}$$

or, after absorbing $1 + C(v, m_{\text{int}})$,

$$N^b > M_0^{C'_{\text{int}}} C'_{\text{dio}} N^{3/4 + 3\delta_{\text{dio}}} \left\{ (2M_0)^b + \log(2M_0^2) \right\}.$$

Recall that $N = [M_0^{C_1}]$. This implies the following condition on C_1 and b:

$$C_1 b > C'_{\text{int}} + C_1(3/4 + 3\delta_{\text{dio}}) + b + C''_{\text{dio}},$$

which can be satisfied as long as $b > 3/4 + 3\delta_{\text{dio}}$ and $\overline{N}_0$ is sufficiently large, depending on C_{dio}.

Finally, we check the conditions under which the right hand side of (10.1) gives at least as good estimate as the one in the right hand side of (6.8): that is, when

$$\frac{-c \log \varkappa}{D_0 \log(D_0 + B_1 + B_2 + B_3)} > N^{3\rho},$$

or

$$cN^b > N^{3\rho} D_0 \log(D_0 + B_1 + B_2 + B_3) \approx N^{3\rho} D_0 M_1.$$

Recalling the definition of M_1 and absorbing the constants, we arrive to the following sufficient condition (note that B_3 dominates D_0, B_1, B_2, assuming again that $\overline{N}_0$ is sufficiently large, depending on C_{dio}):

$$N^b > D_0 M_0^{1/\rho} N^{3\rho},$$

which transforms into the following condition on C_1:

$$C_1 b > C'_{\text{int}} + C_1(3/4 + 3\delta_{\text{dio}} + 3\rho) + 1/\rho + C''_{\text{dio}}.$$

The last condition can be satisfied if

$$b - 3/4 - 3\delta_{\text{dio}} - 3\rho > 0,$$
$$C_1 > \frac{C'_{\text{int}} + C''_{\text{dio}} + 1/\rho}{b - 3/4 - 3\delta_{\text{dio}} - 3\rho}.$$

and then choosing $\overline{N}_0$ in order to satisfy earlier assumptions on a sufficiently large scale. $\square$

REMARK 10.2. The only place where a bound on m_{int} is used, is the Assumption (1) of Cartan's lemma.

10.2   Proofs of some semi-algebraic facts.   We start from the following quantitative triangulation theorem due to Yomdin and Gromov formulated in [Gro87], see also [Yom87, Bur08, PW06]. Let Δ_k be the standard k-dimensional simplex.

PROPOSITION 10.3. *Fix $r \in \mathbb{Z}_+$. Any closed semi-algebraic set $Y \subset [0,1]^d$ can be triangulated into $(\deg Y + 1)^{C(d,r)}$ simplices, where for every closed triangulating k-simplex $\Delta \subset Y$ there exists an algebraic homeomorphism*

$$h_\Delta \colon \Delta_k \to \Delta, \quad \deg h_\Delta \leqslant (\deg Y + 1)^{C(n,r)},$$

such that $h_\Delta|_{\text{int}\,\Delta_k}$ is real analytic with non-vanishing differential, and $\|D_r h_\Delta\| \leqslant 1$.

Proof of (sa5). Apply Proposition 10.3 with $r = 2$. Since the interior point of any d-dimensional simplex is also an interior point of S, the boundary ∂S is contained in the union of all triangulating simplices of dimensions $\leqslant d - 1$. The bound on $\|D_r h_\Delta\|$ will imply the bound on the area of ∂S.

Proof of (sa6). Suppose $|S| \geqslant \varepsilon_1$ and S does not contain a ball of radius δ. Then $S \subset \partial S + B_\delta$. Applying (sa4) with $\varepsilon = \delta$, we obtain an upper bound $|S| \leqslant (\deg S)^{C(d)}\delta$. The contradiction obtained at $\delta < (\deg S)^{C(d)}\varepsilon_1$, proves the first claim.

To prove the second claim, apply Proposition 10.3 to $\mathcal{C}$ and obtain a piece $\mathcal{C}_1 \subset \mathcal{C}$ with the following properties: $\mathcal{C}_1$ is a homeomorphic image of $[0,1]$ under an algebraic map h, which is real analytic with non-zero derivative on $(0,1)$. Moreover, the length of $\mathcal{C}_1$ satisfies $|\mathcal{C}_1| \geqslant B^{-C}L$. With a loss of extra B^C (absorbed into the previous estimate), we can also assume that the direction of tangent vector to $\mathcal{C}_1$ does not change more than by, say, $\pi/100$. Then, one can algebraically re-parametrize $\mathcal{C}_1$ by the distance from the initial point $h(0)$ and obtain an interval I_1 of length $B^{-C}L$ and an algebraic function $h_1 \colon I_1 \to \mathcal{C}_1$ that parametrizes $\mathcal{C}_1$ with $\frac{1}{2} \leqslant |h'_1| \leqslant 2$. With these preparations, one can apply the first claim to $I_1 \times I$: the splitting of $\mathcal{C} \times I$ into K semialgebraic pieces induces a semialgebraic splitting of $I_1 \times I$ into K pieces. At least one of the pieces has area $\geqslant B^{-C}L^2K^{-1}$, and hence contains a ball B of radius $B^{-C}L^2K^{-1}$. By picking a smaller square inside of that ball, one can find intervals I_2, J such that $I_2 \times J \subset B$ and $|I_2|, J| \geqslant B^{-C}L^2K^{-1}$. To conclude, one can pick $\mathcal{C}' = h_1(I_2)$.

Proof of (sa8). The part regarding algebraic local charts is standard, using the fact that the image of a semialgebraic set under an algebraic map is semialgebraic, with the control on the degrees.

To explain the proof for the provided example, consider the set

$$\mathcal{B} = \{(x,y) \in \mathbb{R}^{d+1} \times \mathbb{R}^{d+1} \colon |x-y| \geqslant \eta, \ [x,y] \subset \mathcal{A}\},$$

where $[x,y]$ denotes the line segment connecting x and y. Clearly, $\mathcal{B}$ is semialgebraic of degree $\leqslant (\deg \mathcal{A})^{C(d)}$. Define a function

$$f \colon \mathcal{B} \to \mathbb{S}^d, \quad f(x,y) = \frac{y-x}{|y-x|}.$$

Then f is algebraic, and $\Xi = f(\mathcal{B})$ is a semi-algebraic subset of $\mathbb{S}^d$ of degree $\leqslant (\deg \mathcal{A})^{C(d)}$.

Proof of (sa9). Since S is connected and admits a triangulation, it is also path connected. Let $p, q \in S$, and $f \colon [0,1] \to S$ be a continuous path: $f(0) = p$, $f(1) = q$. Apply Proposition 10.3 to S. For each triangulating simplex Δ, let $[a,b]$ be the smallest interval containing $f^{-1}(\Delta)$. Then, replace the part of the path f between a, b by $h_\Delta(L)$, where L is the line segment in Δ connecting $h_\Delta^{-1}(f(a))$ and $h_\Delta^{-1}(f(b))$. By applying this operation (at most once for each simplex), one gets a path connecting p and q, that visits each simplex only once, and is a straight line segment in the local coordinates on each simplex. Clearly, this new path satisfies all stated properties. See also [BPR06, Section 5.2].

Acknowledgements

References

[AW09]   M. AIZENMAN and S. WARZEL. Localization bounds for multiparticle systems. *Comm. Math. Phys*, (3)290 (2009), 903–934.

[BPR06]   S. BASU, R. POLLACK and ROY M. *Algorithms in Real Algebraic geometry.* Springer, Berlin (2006).

[BCR98]   J. BOCHNAK, M. COSTE and M. ROY. *Real Algebraic Geometry.* Springer, Berlin (1998).

[Bou05]   J. BOURGAIN. *Green's Function Estimates for Lattice Schrödinger Operators and Applications*, Annals of Mathematics Studies, Princeton University Press, Princeton (2005).

[Bou02a] J. BOURGAIN. Estimates on Green's functions, localization and the quantum kicked rotor model. *Ann. Math*, (1)156 (2002), 249–294.

[Bou02b] J. BOURGAIN. On the spectrum of lattice Schrödinger operators with deterministic potential I. *J. Anal. Math.*, 87 (2002), 37–75.

[Bou07]   J. BOURGAIN. Anderson localization for quasi-periodic lattice Schrödinger operators on $\mathbb{Z}^d$, d arbitrary. *Geom. Funct. Anal.*, (3)17 (2007), 682–706.

[BG00]   J. BOURGAIN and M. GOLDSTEIN. On nonperturbative localization with quasi-periodic potential. *Ann. of Math.*, 152 (2000), 835–879.

[BGS02]   J. BOURGAIN, M. GOLDSTEIN and W. SCHLAG. Anderson localization for Schrödinger operators on $\mathbb{Z}^2$ with quasi-periodic potential. *Acta Math.*, 188(1) (2002), 41–86.

[BJ]   J. BOURGAIN and S. JITOMIRSKAYA. Anderson localization for band model. *Geometric Aspects of Functional Analysis. Lecture Notes in Mathematics*, vol 1745, 67–79.

[BJK]   J. BOURGAIN, S. JITOMIRSKAYA and I. KACHKOVSKIY. Localization and delocalization for strongly interacting quasiperiodic particles, in preparation.

[Bur08]   D. BURGET. A proof of Yomdin–Gromov's algebraic lemma. *Israel J. Math.*, 168 (2008), 291–316.

[CS14]   V. CHULAEVSKY and YU. SUHOV. *Multi-Scale Analysis for Random Quantum Systems with Interaction*, Springer, Berlin (2014).

[DGY]   D. DAMANIK, A. GORODETSKI and W. YESSEN. The Fibonacci Hamiltonian. *Inv. Math.*, (3)206, 629–692.

[ES17]   A. ELGART and S. SODIN. The trimmed Anderson model at strong disorder: localisation and its breakup. *J. Spectr. Theory*, 7(1) (2017), 87 – 110.

[GS01]   M. GOLDSTEIN and W. SCHLAG. Hölder continuity of the integrated density of states for quasi-periodic Schrödinger equations and averages of shifts of subharmonic functions. *Ann. Math.*, (1)154 (2001), 155–203.

[FS15]   K. FRAHM and D. SHEPELYANSKY. Freed by interaction kinetic states in the Harper model. *D. Eur. Phys. J. B*, (2015) 88:337.

[Gab68]   A. GABRIELOV. Projections of semi-analytic sets. *Funct. Anal. Appl.*, (4)2 (1968), 282–291.

[Gro87]   M. GROMOV. Entropy, homology and semialgebraic geometry. In: *Séminaire Bourbaki, vol. 1985/86, exp. no. 661. Astérique*, 145 – 146 (1987), 5, 225–240.

[JSS03]   S. JITOMIRSKAYA, H. SCHULTZ-BALDES and G. STOLTZ. Delocalization in random polymer models. *Comm. Math. Phys.*, (1)233 (2003), 27–48.

[KN13]   A. KLEIN and S. NGUYEN. The bootstrap multiscale analysis of the multi-particle Anderson model. *J. Stat. Phys.*, (5)151 (2013), 938–973.

[Lev96]   B. Ya. Levin. *Lectures on entire functions*, American Mathematical Society Translations (1996).

[PW06]   J. Pila and A. Wilkie. The rational points of a definable set. *Duke Math. J.*, (3)133 (2006), 591–616.

[Wol95]   T. Wolff. An improved bound for Kakeya type maximal functions. *Rev. Mat. Iberoam.*, (3)11 (1995), 651–674.

[Yom87]   Y. Yomdin. C^k-resolution of semi-algebraic mappings. *Israel J. Math.*, 57 (1987), 301–317.

J. Bourgain, School of Mathematics, Institute for Advanced Study, Princeton, NJ, 08540, USA   bourgain@math.ias.edu

I. Kachkovskiy, Department of Mathematics, Michigan State University, East Lansing, MI, 48824, USA   ikachkov@msu.edu