LIFSHITZ TAILS FOR 2-DIMENSIONAL RANDOM SCHRÖDINGER OPERATORS

FRÉDÉRIC KLOPP AND THOMAS WOLFF

ABSTRACT. The purpose of this paper is to prove that, for a 2-dimensional continuous Anderson model, at an open band edge, the density of states exhibits a Lifshitz behavior. The Lifshitz exponent need not be $d/2 = 1$ but is determined by the behaviour of the Floquet eigenvalues of a well chosen background periodic Schrödinger operator.

RÉSUMÉ. Le but de cet article est de démontrer qu'en dimension 2, la densité d'états de toute perturbation aléatoire du type d'Anderson d'un opérateur de Schrödinger périodique satisfait à des asymptotiques de Lifshitz au bord d'un gap ouvert. L'exposant de Lifshitz n'est pas forcément égal à $d/2 = 1$ mais est déterminé par le comportement des valeurs propres de Floquet en bord de bande associé à un opérateur de Schrödinger périodique sous-jacent bien choisi.

0. INTRODUCTION

The purpose of this paper is to prove that, for a 2-dimensional continuous Anderson model, at every band edge, the density of states exhibits a Lifshitz behavior.

This behavior was first discovered by I.M. Lifshitz (see [26]) in the early 60's. I.M. Lifshitz produced a heuristic showing that, at the bottom of the spectrum of a random Schrödinger operator, the density of states decays exponentially fast i.e. if $N(E)$ is the density of states of the random operator and 0 is the bottom of the spectrum, one has

$$\log|\log N(E)| \underset{E\to 0^+}{\sim} -\frac{d}{2}\log E. \tag{0.1}$$

Mathematical work on Lifshitz tails only appeared much later. The behavior near the bottom of the spectrum was studied first. The results were not obtained following the original Lifshitz heuristic but using Wiener integrals and the Donsker-Varadhan technique to estimate the Laplace transform of N (see [31, 32]); the results obtained were more precise than (0.1). This point of view is directly related to the study of Brownian motion among Poissonian obstacles (see [43]). Proofs of Lifshitz tails at the bottom of the spectrum for the continuous Anderson model following the original heuristic of Lifshitz were developed in the 80's (see [14, 15, 29, 42]).

Physically Lifshitz tails were also predicted at all fluctuational edges of the spectrum (see [27]). For the continuous Anderson model, this means all finite edges of the spectrum. The first mathematical result in this direction was obtained in dimension 1 in [30]. In dimension larger than 1, it was proved in [23] that (0.1) holds if and only if the density of states of a well chosen underlying periodic Schrödinger operator,say N_p satisfies

$$\log N_p(E) \underset{E\to 0^+}{\sim} \frac{d}{2}\log E. \tag{0.2}$$

1991 *Mathematics Subject Classification.* 82B44, 42B99.

Key words and phrases. Random Schrödinger operators, Lifshitz tails, random Fourier series.

F. K. acknowledges support from Caltech where this work was initiated, as well as from U.C.L.A and from U.C. Berkeley where it was partially done. F.K. also wants to thanks C. Sabbah for illuminating discussions on analytic vector bundles.

The problem was then to understand what happens in the general case i.e. in the case when (0.2) is not necessarily satisfied; this is the question we will answer in dimension $d = 2$.

0.1. **The main results.** Let (e_1, e_2) be a basis of $\mathbb{R}^2$ and $\Gamma = \mathbb{Z}e_1 \oplus \mathbb{Z}e_2$ be an integer lattice in $\mathbb{R}^2$; call its fundamental cell $C_0 = \{x_1e_1 + x_2e_2;\ \forall 1 \leq j \leq 2,\ -1/2 < x_1, x_2 \leq 1/2\}$. Let W be a Γ-periodic potential in $L^2_{\text{loc}}(\mathbb{R}^2)$ and $H = -\Delta + W$ be a periodic Schrödinger operator acting on $L^2(\mathbb{R}^2)$. We will consider a random Schrödinger of the form

$$H_\omega = H + V_\omega = H + \sum_{\gamma\in\Gamma} \omega_\gamma V_\gamma, \tag{0.3}$$

where $V_\gamma(\cdot) = V(\cdot - \gamma)$ and we assume that
(H.1)

1. $0 \not\equiv V$ is continuous such that $\sum_{\gamma\in\Gamma} |V(\cdot - \gamma)|$ be bounded.
2. V is non-negative.
3. $(\omega_\gamma)_{\gamma\in\Gamma}$ is a collection of independent identically distributed bounded random variables.
4. The common support of the random variables $(\omega_\gamma)_{\gamma\in\Gamma}$ is connected.

Let us recall a few well known facts from the theory of random Schrödinger operators (see e.g. [33, 12, 3]). For all ω, H_ω is essentially self-adjoint on $\mathcal{C}_0^\infty(\mathbb{R}^2)$, and we denote by H_ω its self-adjoint extension. There exists Σ, a closed subset of $\mathbb{R}$ such that, for almost all ω, the spectrum of H_ω coincides with Σ. This set is called the *the (almost sure) spectrum* of H_ω. Our assumptions on V_ω and H imply that Σ is lower semi-bounded.

It follows from the Bethe-Sommerfeld conjecture (see e.g. [41, 40, 10, 7]) and the standard characterization of Σ using admissible periodic operators (see e.g.[13, 33]) that Σ is a finite union of intervals. The connected components of $\mathbb{R} \setminus \Sigma$ will be the gaps of Σ. The edges on the gaps are the points in $\Sigma \cap \overline{\mathbb{R} \setminus \Sigma}$.

One defines the integrated density of states of H_ω in the following way: let $H_\omega^{l,D}$ to be the Dirichlet restriction of H_ω to the cube C_l centered in 0 of side length l. $H_\omega^{l,D}$ has only discrete spectrum. For $E \in \mathbb{R}$, we define

$$N_\omega^l(E) = \frac{1}{\text{Vol}(C_l)} \sharp\{\text{eigenvalues of } H_\omega^{l,D} \leq E\} \tag{0.4}$$

where $\sharp\mathcal{E}$ denotes the number of points of the set $\mathcal{E}$ and $\text{Vol}(C_l)$, the volume of C_l.

Then, ω-almost surely, $N_\omega^l(E)$ has a limit when $l \to +\infty$. This limit is independent of the realization of ω. It is the *integrated density of states* or *IDS* of H_ω. We will denote it by $N(E)$. Physically, it can be interpreted as the number of states per unit of volume below energy E for a system governed by H_ω. The function N is a non random, non decreasing and right continuous function. Its set of increasing points is Σ.

0.1.1. *The density of states decays exponentially at band edges.* We prove

Theorem 0.1. *Assume (H.1) is satisfied. For any edge of a gap of Σ, say E_0, one has*

$$\limsup_{\substack{E\to E_0 \\ E\in\Sigma}} \frac{\log|\log|N(E) - N(E_0)||}{\log|E - E_0|} < 0. \tag{0.5}$$

Theorem 0.1 says that the integrated density of states decays exponentially fast at any band edge of Σ. This can be used to prove band-edge localization for H_ω (see section 0.3).

The idea of the proof of Theorem 0.1 goes roughly as follows. First, following the ideas developed in [23], we show that the study of the density of states of H_ω near E_0 can be reduced to the study of the density of states of a discrete random operator; this is the purpose of section 1. This discrete Hamiltonian is the sum of a kinetic energy part and a potential energy part. The kinetic energy part is made of the Floquet eigenvalues giving rise to the band starting at E_0 for

a well chosen background periodic operator (see section 0.2). The potential energy part is the random potential projected onto the space of states generating the band starting at E_0. This discrete Hamiltonian will be called the reduced operator.
In section 2, we then show that, with good enough precision, we can estimate the integrated density of states of the reduced operator between $E_0-\varepsilon$ and $E_0+\varepsilon$ by the probability that the reduced operator (restricted to some large cube) has an eigenvalue in the interval $[E_0-\varepsilon, E_0+\varepsilon]$. At last, we estimate these probabilities. Carrying out this estimation is essentially a problem in random Fourier series, which we discuss in more detail later in the introduction.

0.1.2. *The Lifshitz exponent.* Under more restrictive assumptions on the random operator, we are able to compute the rate of exponential decay of the integrated density of states.

We consider a random Schrödinger operator for the form (0.3) satisfying assumption (H.1')

1. $0\not\equiv V$ is continuous function with compact support.
2. V is non-negative.
3. $(\omega_\gamma)_{\gamma\in\Gamma}$ is a collection of independent identically distributed random variables.
4. The support of these random variables is the interval $[\omega^-,\omega^+]$.
5. The common distribution of the $(\omega_\gamma)_{\gamma\in\Gamma}$ satisfies

$$\lim_{\varepsilon\to 0}\frac{\log|\log\mathbb{P}\{|\omega_0-\omega^\pm|\le\varepsilon\}|}{-\log\varepsilon}=0$$

To state the additional assumptions, we will need to introduce some notations. The Lifshitz exponent is computed using a well chosen background periodic operator that can but need not be H_0. This background operator depends on the spectral edge E_0 one is considering. More precisely it depends on whether E_0 is the upper (resp. lower) edge (i.e. maximum (resp. minimum)) of a gap. Recall that the essential infimum (resp. supremum) of the random variables $(\omega_\gamma)_\gamma$ is ω_+ (resp. ω_-). Let H^+ (resp. H^-) be the periodic Schrödinger operator defined by $H^-=H+\omega_-\sum_{\gamma\in\Gamma}V_\gamma$ (resp. $H^+=H+\omega_+\sum_{\gamma\in\Gamma}V_\gamma$). Then $H_\omega=H^-+\sum_{\gamma\in\Gamma}\omega_\gamma^+V_\gamma$ (resp. $H_\omega=H^++\sum_{\gamma\in\Gamma}\omega_\gamma^-V_\gamma$) where $\omega_\gamma^+=\omega_\gamma-\omega_-$ (resp. $\omega_\gamma^-=\omega_\gamma-\omega_+$). The random variables $(\omega_\gamma^+)_\gamma$ (resp. $(\omega_\gamma^-)_\gamma$) are i.i.d. and non-negative (resp. non-positive). Using the results of [13], we see that

- E_0 is the upper edge of a gap of Σ if and only if E_0 is the upper edge of a gap of the spectrum of H^-.
- E_0 is the lower edge of a gap of Σ if and only if E_0 is the lower edge of a gap of the spectrum of H^+.

Let us for a while assume that E_0 is the upper edge of a gap of Σ. Then the correct background operator will be H^-.
We recall some facts on the Floquet spectrum of H^- (see e.g. [37, 25, 39]). The Floquet spectrum of H^- is the spectrum of the differential operator H^- acting on $L^2_{\text{loc}}(\mathbb{R}^2)$ with quasi-periodic boundary conditions. For $\theta\in\mathbb{T}^2=\mathbb{R}^2/(\Gamma^*)$ (here Γ^* is the dual lattice of Γ i.e. for $\gamma\in\Gamma$ and $\gamma^*\in\Gamma^*$, one has $\gamma\gamma^*\in 2\pi\mathbb{Z}$), consider the following eigenvalue problem on $L^2_{\text{loc}}(\mathbb{R}^2)$

$$\begin{cases} H^-\varphi=E\varphi \\ \varphi(x+\gamma)=e^{i\gamma\theta}\varphi(x),\ \forall x\in\mathbb{R}^2,\ \forall\gamma\in\Gamma. \end{cases} \tag{0.6}$$

As H^- is elliptic, one knows that the eigenvalues of (0.6) are discrete; when repeated according to multiplicity, we denote them by $E_0(\theta)\le E_1(\theta)\le\cdots\le E_n(\theta)\le\dots$. They are called the *Floquet eigenvalues.* These functions are Lipschitz in the variable θ; they are even analytic in θ when simple. Moreover, Weyl's law tells us that

$$E_n(\theta)\to+\infty \text{ as } n\to+\infty \text{ (uniformly in } \theta). \tag{0.7}$$

The spectrum of H^- is given by $\sigma(H^-) = \cup_{n\geq 0} E_n(\mathbb{T}^2)$. Let E be in the spectrum of H^-. E will be *simple* if the set $\{p \in \mathbb{N};\ \exists\theta \in \mathbb{T}^2,\ E_p(\theta) = E\}$ is reduced to a single integer. Note that, by (0.7), this set of integers is always finite. An edge of a gap of Σ will be called *simple* if it is the upper edge (resp. the lower edge) of a gap of Σ and it is simple for H^- (resp. H^+). By [22], we know that for a generic periodic potential, edges of gaps are simple.

We prove

Theorem 0.2. *Assume (H.1') is satisfied. Let E_0 be an edge of a gap of Σ that is simple. Then, there exists $0 < \alpha \leq +\infty$ such that*

$$\lim_{\substack{E\to E_0\\ E\in\Sigma}} \frac{\log|\log|N(E) - N(E_0)||}{\log|E - E_0|} = -\alpha. \tag{0.8}$$

The exponential decay on the integrated density of states at the edges of the spectrum Σ is known as "Lifshitz tails" for the IDS. The exponent α is the *Lifshitz exponent.* The computation of α is not obvious; section 0.2 is devoted to its description.

Let us now make a few remarks about assumption (H.1). First, the continuity assumption on V is not really needed; one needs some regularity for V to ensure that the random operators are well defined, and one needs the fact that, on some open set, V does not vanish.
Second, assumptions (H.1.4) and (H.1'.4) about the connectedness of the essential support of the random variables $(\omega_\gamma)_\gamma$ can be replaced by the weaker assumption (H.1.4') described in section 1 (see also remark 1.1). This assumption roughly states that the gap below E_0 should stay open while the perturbation is switched on. These remarks also hold for assumption (H.1) and Theorem 0.1.
Third, the compactness of the support of the potential V is not really needed; one only needs that the potential decay sufficiently fast at ∞. However, it is known that the decay of V, when slow enough, changes the value of the Lifshitz exponent (see remark 1.4 and e.g. [12, 15, 23, 20]).

0.2. **How to compute the Lifshitz exponent.** We now assume that (H.1') is satisfied. For the sake of definiteness, let us assume that E_0 is the upper edge of a gap of Σ. Assume E_0 is simple. Let $E_{n_0}(\cdot)$ be the unique Floquet eigenvalue taking the value E_0. Define the set $\mathcal{S} = \{\theta \in \mathbb{T}^2;\ E_{n_0}(\theta) = E_0\}$. Then, there exists $\eta > 0$ such that

- For $n < n_0$, for all $\theta \in \mathbb{T}^2$, $E_n(\theta) < E_0 - \eta$.
- For $n > n_0$, for all $\theta \in \mathbb{T}^2$, $E_n(\theta) > E_0 + \eta$.

This implies that, for θ in some neighborhood of $\mathcal{S}$, the Floquet eigenvalue $E_{n_0}(\theta)$ is simple, hence, that in the same neighborhood, the function $\theta \mapsto E_{n_0}(\theta)$ is real analytic. The Lifshitz exponent will depend on the behavior of E_{n_0} near $\mathcal{S}$.

To describe it precisely, we need to introduce some objects from analytic geometry (see section 3). If $\mathcal{E}$ is a set contained in the closed first quadrant in $\mathbb{R}^2$ then its *exterior convex hull* is the convex hull of the union of the rectangles $R_{xy} = [x, \infty) \times [y, \infty)$, where the union is taken over all $(x, y) \in \mathcal{E}$.
Pick $\theta_0 \in \mathcal{S}$ and consider the Newton diagram of E_{n_0} at θ_0, i.e.,

1. Express E_{n_0} as a Taylor series at θ_0, $E_{n_0}(\theta^1, \theta^2) = \sum_{ij} a_{ij}(\theta^1 - \theta_0^1)^i(\theta^2 - \theta_0^2)^j$, $\theta = (\theta^1, \theta^2)$.
2. Form the exterior convex hull of the points (i, j) with $a_{ij} \neq 0$. This is a convex polygon, called the *Newton polygon.*
3. The boundary of the polygon is the *Newton diagram.*

The Newton decay exponent is then defined as follows. The Newton diagram consists of certain line segments. Extend each to a complete line and intersect it with the diagonal line $\theta^1 = \theta^2$. This gives a collection of points (a_k, a_k), one for each boundary segment. Take the reciprocal of the largest a_k and call this number $\tilde{\alpha}(E_{n_0}, \theta_0)$. Define $\alpha(E_{n_0}, \theta_0) = \min\{\tilde{\alpha}(E_{n_0} \circ T_0, \theta_0) : T_0(\cdot) = \theta_0 + T(\cdot - \theta_0),\ T \in SL(2, \mathbb{R})\}$.

Similarly, define $\alpha(E_{n_0},\theta)$ if θ is any other point in $\mathcal{S}$, the zero set of E_{n_0}. Then, one defines the Lifshitz exponent α as

$$\alpha = \min_{\theta\in\mathcal{S}} \alpha(E_{n_0},\theta). \tag{0.9}$$

The Lifshitz exponent α is positive as $\theta \mapsto \alpha(E_{n_0},\theta)$ is a positive, lower semi-continuous function and $\mathcal{S}$ is compact.

A similar Newton decay exponent arises in problems concerning estimation of oscillatory integrals - see [45] and [35] - and we show here that it also controls the behavior of certain random Fourier series, in a way which to some extent is suggested by the results of [45] and [35]. We will now try to explain our estimate on random Fourier series in a nontechnical way; the precise result is Theorem 3.1 stated in section 3 below.

Let f be a nonnegative real analytic function on T^2, let $\delta > 0$ and $M < \infty$ be parameters, let S be a random subset of the square $[-M,M]\times[-M,M]$ with (say) density $\frac{1}{2}$ and let $P(M,\delta)$ be the probability that there is a function $u \in L^2(T^2)$ such that

$$\int_{T^2} f(\theta)|u(\theta)|^2 < \delta\|u\|_2^2$$

and

$$\mathrm{supp}(\hat{u}) \subset S$$

Thus the Fourier coefficients of u are to be supported on the random set, and u is to be mostly supported near the zero set of f. We will study the behavior of the probability $P(M(\delta),\delta)$ as $\delta \to 0$, letting $M = M(\delta)$ go to ∞ in an appropriate manner. It turns out that $P(M(\delta),\delta)$ behaves roughly like $e^{-\delta^{-\alpha}}$.

One can apply this bound taking $f(\theta)$ to be roughly the lowest Floquet eigenvalue to complete the proofs of Theorems 0.1 and 0.2, in a manner analogous to the one dimensional case treated in [19]. See the discussion in section 0.1.1.

To conclude this subsection, let us compare N to N_p, the density of states of the underlying periodic operator H. Up to a constant factor, $N_p(E_0+\varepsilon) - N_p(E_0)$ is the volume of $\{\theta \in \mathbb{T}^2;\ E_{n_0}(\theta) \le E_0+\varepsilon\}$. From previously known results (see e.g. [12, 23]), one may have conjectured that, at band edges N behaves roughly like e^{-1/N_p} (if one takes the double logarithm of both sides). We see here that this will not be true in general if the set $\mathcal{S}$ has some curvature. For example, if we assume that $E_{n_0}(\theta) - E_0$ vanishes to constant order q on a curve, and if the maximum order of vanishing of the curvature of the curve is m (thus non vanishing curvature means $m = 0$) then $\alpha = \frac{m+3}{q(m+2)}$ and the volume will have the exponent $\frac{1}{q}$.

0.3. **Application to localization.** We apply our results to study band edge localization for random perturbations of periodic operators. Consider H_ω a random Schrödinger operator of the form (0.3) satisfying assumption (H.1). Let l be an integer and H^θ_{ω,C_l} be the operator H_ω restricted to the box C_l with quasi-periodic boundary condition θ (see (0.6)). Then, one proves the

Theorem 0.3. *Fix $\theta \in \mathbb{R}^d$ and E_0 a band edge of the spectrum of H_ω. Then, for any $\beta > 0$ integer, $\nu > 0$, for $l \in \mathbb{N}$ sufficiently large, one has*

$$\mathbb{P}\left(\left\{ dist(\sigma(H^\theta_{\omega,C_{l^\beta}}),E_0) \le \frac{1}{l}\right\}\right) \le l^{-\nu} \tag{0.10}$$

where C_l is the cube centered in 0 of side length $2l+1$ and H^θ_{ω,C_l} is the operator H_ω restricted to this cube with θ-quasi-periodic boundary conditions i.e. with the boundary conditions $\varphi(x+\gamma) = e^{i\theta\gamma}\varphi(x)$ for any $\gamma \in 2k\Gamma$.

Remark 0.1. This result improves upon the earlier result found in [46] where Theorem 0.3 is essentially proven in the case when V has compact support.

Let us now assume that the random operator satisfies a *Wegner estimate* i.e that for some $\beta > 0$ and $\nu > 0$, for $E \in \mathbb{R}$, for $l \geq 1$ and $0 < \varepsilon < 1$, there exists $C(E) > 0$ such that one has

$$\mathbb{P}\left(\mathrm{dist}(\sigma(H^{\theta}_{\omega,C_l}, E) \leq \varepsilon\right) \leq C(E)\mathrm{Vol}(C_l)^{\beta}\varepsilon^{\nu} \tag{0.11}$$

where the constant $C(E)$ is locally uniform in energy.
To be able to apply the multi-scale analysis developed in [17] (see also [1, 16]) , we assume that, for some $\eta > 8$, one has

$$\sum_{\gamma\in\Gamma} \|V\|_{L^2(\gamma+C_0)}|\gamma|^{\eta} < +\infty. \tag{0.12}$$

Picking a band edge energy E_0, using estimate (0.10) for $\theta = 0$, we obtain the initial estimate to start the multi-scale analysis. This shows that the spectrum of H_ω is exponentially localized in some interval around the energy E_0 i.e. that, in some interval around the energy, the spectrum is purely punctual and the eigenfunctions associated to energies in that interval are exponentially localized. Therefore, we have proved the

Theorem 0.4. *Let H_ω be a random operator described by* (0.3) *such that (H.1) and* (0.12) *(with $\eta > 8$) are satisfied. Assume moreover that the random operator satisfies a Wegner estimate* (0.11)*. Then, for every band edge of Σ, the almost sure spectrum of H_ω, there exists a neighborhood of that band edge where the spectrum of H_ω is purely punctual. Moreover the eigenfunctions associated to energies in that neighborhood are exponentially localized.*

Contents

1. Reduction to a discrete model

We will now show how to reduce the computation of the asymptotics of N near a spectral edge to the computation of the asymptotics of the probability of the existence of extremal eigenvalues for some well chosen discrete operator.

We assume that we are in the setting described in section 0.1. We fix E_0 and assume it is the upper edge of a gap. To simplify the notations, from now on and in the rest of the paper, we will assume that $E_0 = 0$. As it is H^- that will be the crucial background operator, we will call it H; the same way, the random variable $(\omega_\gamma^+)_\gamma$ will just be denoted by $(\omega_\gamma)_\gamma$. Hence, we are considering a random Schrödinger operator of the form

$$H_\omega = H + V_\omega = H + \sum_{\gamma\in\Gamma} \omega_\gamma V_\gamma,$$

where we assume that H is periodic Schrödinger operator,

- $0 \not\equiv V \in \mathcal{C}_{\text{comp}}(\mathbb{R}^2))$,
- V is non-negative, $V_\gamma(\cdot) = V(\cdot - \gamma)$,
- $(\omega_\gamma)_{\gamma\in\Gamma}$ is a collection of independent, identically distributed, bounded, non-negative random variables, the support of which contains 0.

and that, for some $\delta > 0$, we have

$$(-\delta, 0) \cap \sigma(H) = (-\delta, 0) \cap \Sigma = \emptyset, \quad [0, \delta] \cap \sigma(H) = [0, \delta] \cap \Sigma = [0, \delta]. \tag{1.1}$$

For $t \in [0, 1]$, we define the random Schrödinger operator $H_{t,\omega} = H + tV_\omega$ and $\Sigma(t)$ to be its almost sure spectrum. Then, instead of assuming that the random variables $(\omega_\gamma)_{\gamma\in\Gamma}$ have connected support, we assume that
(H.1.4')

- for some $\delta' > 0$, for any $t \in [0, 1]$, $\Sigma(t) \cap (-\delta', 0) = \emptyset$.

Remark 1.1. If H_ω satisfies (H.1.3) and if we assume that (1.1) holds, then (H.1.3') is also satisfied. Indeed, as the support of the $(\omega_\gamma)_\gamma$ is an interval, we know that $\Sigma(t) \subset \Sigma$ for all $t \in [0, 1]$ (see e.g. [13]).

We note that assumption (H.1.3') can always be replaced by a stronger assumption on the size of V_ω of the form

- for some $\delta' > 0$, almost surely $\Pi_- V_\omega \Pi_- \leq (\delta - \delta')\Pi_-$.

Here Π_- is the spectral projection associated to H and the interval $(-\infty, 0)$. This will ensure that the gap below energy 0 will stay open for almost every ω. Notice that, as $\Pi_- H \Pi_-$ is a bounded operator and as V_ω is uniformly relatively H-bounded, $(\Pi_- V_\omega \Pi_-)_\omega$ is a bounded family of operators.

Our reduction will be based on the work done in [23]. Therefore we recall some of the notions used in this paper.

1.1. **Floquet theory, Wannier bases, etc.**

Let us describe the Floquet theory for a periodic operator a little further. Therefore, we will need to introduce some definitions.

1.1.1. *Some definitions.* Let $\mathbb{T} = \mathbb{R}^2/\Gamma$ be the quotient torus generated by Γ. Recall that Γ^* is Γ's dual lattice; let C_0^* be Γ^*'s fundamental cell and $\mathbb{T}^2$ be the quotient torus generated by Γ^*. $\mathbb{T}^2$ and C_0^* are identified in a canonical way. The dual basis of $(e_j)_{1\leq j\leq 2}$ is denoted by $(e_j^*)_{1\leq j\leq 2}$. By means of this basis, Γ^* is identified with $2\pi\mathbb{Z}^2$. For $\theta \in (\mathbb{R}^2)^*$, we write $\theta = (\theta_j)_{1\leq j\leq 2}$ when $\theta = \theta_1 e_1^* + \theta_2 e_2^*$.

For $v \in \mathbb{R}^2$, $\tau_v : L^2(\mathbb{R}^2) \to L^2(\mathbb{R}^2)$ denotes the "translation by v" operator i.e for $\varphi \in L^2(\mathbb{R}^2)$ and $x \in \mathbb{R}^2$, $(\tau_v\varphi)(x) = \varphi(x-v)$.

We will also need to define some functional spaces. Let $\mathcal{D}'_\theta$ be the space of θ-quasiperiodic distribution in $\mathbb{R}^2$ i.e the space of distributions $u \in \mathcal{D}'(\mathbb{R}^2)$ such that, for any $\gamma \in \Gamma$, we have $\tau_\gamma u = e^{-i\gamma\cdot\theta}u$. Here $\theta \in \mathbb{T}^2$. Let $H^k_{\text{loc}}(\mathbb{R}^2)$ be the space of distributions that locally belong to $H^k(\mathbb{R}^2)$ and we define $\mathcal{H}^k_\theta = H^k_{\text{loc}}(\mathbb{R}^2) \cap \mathcal{D}'_\theta$. For $k = 0$, we define $\mathcal{H}_\theta = \mathcal{H}^0_\theta \subset L^2(C_0)$; equipped with the L^2-norm over C_0, it is a Hilbert space. The scalar product will be denoted by $\langle\cdot,\cdot\rangle_\theta$.

At last, define $\mathcal{H} = \{u \in L^2_{\text{loc}}(\mathbb{R}^2) \otimes L^2(\mathbb{T}^2);\ \forall (x,\gamma,\theta) \in \mathbb{R}^2 \times \Gamma \times \mathbb{T}^2,\ u(x+\gamma,\theta) = e^{i\gamma\cdot\theta}u(x,\theta)\}$. Equipped with the norm $\int_{\mathbb{T}^2} \|u(.,\theta)\|^2_{L^2(C_0)} d\theta$, it is Hilbert space. Here and in the rest of the paper, $d\theta$, the Lebesgue measure on $\mathbb{T}^2$, is normalized so that $\int_{\mathbb{T}^2} d\theta = 1$. $\mathcal{H}$ is a fiber bundle over $\mathbb{T}^2$ with fiber $\mathcal{H}_\theta$.

1.1.2. *Floquet theory, etc.* Floquet theory can be constructed as follows (see e.g. [39]). For $\theta \in \mathbb{R}^2$ and $u \in \mathcal{S}(\mathbb{R}^2)$, the Schwartz space of rapidly decreasing functions, we define

$$(Uu)(x,\theta) = \sum_{\gamma\in\Gamma} e^{i\gamma\cdot\theta}u(x-\gamma).$$

U can be extended as an isometry from $L^2(\mathbb{R}^2)$ to $\mathcal{H}$. Its inverse is given by the following formula:

$$\text{for } v \in \mathcal{H}, \quad (U^*v)(x) = \int_{\mathbb{T}^2} v(x,\theta)d\theta.$$

One sees that, for $1 \leq j \leq d$, $[\partial_j, U] = 0$ (here ∂_j denotes the j-th partial derivative). Moreover, as W is periodic, $[W,U] = 0$. Hence, H admits the Floquet decomposition

$$UHU^* = \int_{\mathbb{T}^2}^{\oplus} H_\theta d\theta, \tag{1.2}$$

where H_θ is the differential operator H acting on $\mathcal{H}_\theta$ with domain $\mathcal{H}^2_\theta$. H_θ is self-adjoint and elliptic. The Floquet eigenvalues $E_1(\theta), \cdots, E_n(\theta), \cdots$ are the eigenvalues of H_θ.

Let $\mathcal{E} \subset L^2(\mathbb{R}^2)$ be a closed subspace Γ-translation invariant i.e if $\Pi^{\mathcal{E}}$ is the orthogonal projector on $\mathcal{E}$, for $\gamma \in \Gamma$, one has

$$\Pi^{\mathcal{E}} = \tau_\gamma \Pi^{\mathcal{E}} \tau_\gamma^*.$$

The Floquet decomposition of $\Pi^{\mathcal{E}}$ tells us that

$$\Pi^{\mathcal{E}} = \int_{\mathbb{T}^2}^{\oplus} \Pi^{\mathcal{E}}_\theta d\theta,$$

where $\Pi^{\mathcal{E}}_\theta$ is $\Pi^{\mathcal{E}}$ acting on $\mathcal{H}_\theta$. $\Pi^{\mathcal{E}}_\theta$ is an orthogonal projector acting on $L^2(C_0)$. The family of projectors $(\Pi^{\mathcal{E}}_\theta)_{\theta\in\mathbb{T}^2}$ is continuous in θ. Hence, it is of constant rank. For $\theta \in \mathbb{T}^2$, we can find an orthonormal system of vectors $(u_n(\cdot,\theta))_{n\in N}$ (where $N \subset \mathbb{N}$ is a set of indices independent of θ) that spans the range of $\Pi^{\mathcal{E}}_\theta$.

So, transporting $(u_n(\cdot,\theta))_{n\in N}$ over to $L^2(\mathbb{R}^2)$ using U^*, we see that there exists an orthonormal system of vectors $(\hat{u}_{n,0})_{n\in N}$ such that, if, for $\gamma \in \Gamma$, we set $\hat{u}_{n,\gamma} = \tau_\gamma(\hat{u}_{n,0})$, then, $(\hat{u}_{n,\gamma})_{(n\in N;\ \gamma\in\Gamma)}$ is an orthonormal basis for $\mathcal{E}$. We will call it a *Wannier basis* of $\mathcal{E}$. The vectors $(\hat{u}_{n,0})_{n\in N}$ are the *Wannier generators* for $\mathcal{E}$.

We will say that $\mathcal{E} \subset H^2(\mathbb{R}^2)$, a Γ-translation invariant subspace, is of *finite energy* if $\Pi^{\mathcal{E}} H \Pi^{\mathcal{E}}$ is a bounded operator. If this is the case, $\mathcal{E}$ has only finitely many Wannier generators.

1.2. **The first step of the reduction.** Using the notations of section 0.2, let $(E_n(\theta))_{n\geq 0}$ be the Floquet eigenvalues of H and $(u_n(\cdot,\theta))_{n\geq 0}$ be the associated Floquet eigenvectors.

Let $(E_n(\theta))_{n_0\leq n\leq n_0+p_0-1}$ be the Floquet eigenvalues of H for which there exists $\theta \in \mathbb{T}^2$ such that $E_n(\theta) = 0$. We denote by $\mathcal{E}_0$ the subspace of $L^2(\mathbb{R}^2)$ generated by the vectors $(\hat{u}_{n,\gamma})_{n_0\leq n\leq n_0+p_0-1;\gamma\in\Gamma}$ defined above. Let $\Pi^{\mathcal{E}_0}$ be the orthogonal projector onto $\mathcal{E}_0$.

In Proposition 4.1 of [23], it was proved that the operator $\Pi^{\mathcal{E}_0} H_\omega \Pi^{\mathcal{E}_0}$ is a bounded, ergodic random operator for which one can define an integrated density of states. It will be denoted by N_0 and can be defined by means of the following formula:

$$\int_{\mathbb{R}} \varphi(\lambda) dN_0(\lambda) = \frac{1}{\mathrm{Vol}(C_0)} \sum_{j=n_0}^{n_0+p_0-1} \mathbb{E}\left(\langle u_{j,0}, \varphi\left(\Pi^{\mathcal{E}_0} H_\omega \Pi^{\mathcal{E}_0}\right) u_{j,0}\rangle\right) \text{ for } \varphi \in \mathcal{C}_0^\infty(\mathbb{R}). \tag{1.3}$$

Here $\mathbb{E}(f)$ denotes the expectation of a function f with respect to the random variable $(\omega_\gamma)_{\gamma\in\Gamma}$.

The first tool needed for the reduction will be

Theorem 1.1. *Under the assumptions made in section 1, there exists $C > 0$ and $E_0 > 0$ such that, for $0 \leq E \leq E_0$, we have*

$$N_0(E/C) \leq N(E) - N(0^-) \leq N_0(C \cdot E). \tag{1.4}$$

Proof of Theorem 1.1. To prove estimate (1.4), we shall use the same techniques as in [23]. We will therefore recall some of the notions used in that paper; we refer to [23] for more details. The upper bound in (1.4) was already proved in Theorem 4.1 of [23]; for the sake of completeness, we will reprove it here in a slightly different way.
Let $k \in \mathbb{N} \setminus \{0\}$ and define the following periodic Schrödinger operator

$$H_{\omega,k} = H + \sum_{\gamma\in C_k\cap\Gamma} \omega_\gamma \sum_{\beta\in(2k+1)\Gamma} V(x-\gamma-\beta) = H + V_{\omega,k} \tag{1.5}$$

where $C_k = \{x \in \mathbb{R}^2;\text{ for } j = 1,2,\ -k-1/2 < x_j \leq k+1/2\}$.
Then, for any $\omega \in \Omega$ and $k \in \mathbb{N}^*$, $H_{\omega,k}$ is a $(2k+1)\Gamma$-periodic essentially self-adjoint Schrödinger operator. It is an H-bounded perturbation of H with relative bound 0. We notice that the boundedness assumption on the $(\omega_\gamma)_\gamma$ implies that the family $(H_{\omega,k})_{\omega,k}$ is uniformly H-bounded. $H_{\omega,k}$ will also denote its self-adjoint extension. Being a periodic Schrödinger operator, $H_{\omega,k}$ admits an integrated density of states denoted by $n(H_{\omega,k})$ (see e.g. [37, 38, 39]); it is given by

$$n(H_{\omega,k}, E) = \frac{1}{(2\pi(2k+1))^2} \sum_{n\in\mathbb{N}} \int_{\{\theta\in\mathbb{R}^2/((2k+1)\Gamma)^*;\ E_n^{k,\omega}(\theta)\leq E\}} d\theta. \tag{1.6}$$

where the $E_n^{k,\omega}(\theta)$ are the Floquet eigenvalues of $H_{\omega,k}$.
In [23], we proved that, for any $\varphi \in \mathcal{C}_0^\infty(\mathbb{R})$, we have

$$\lim_{k\to+\infty} \mathbb{E}\left(\int_{\mathbb{R}} \varphi(\lambda) dn(H_{\omega,k})\right) = \int_{\mathbb{R}} \varphi(\lambda) dN, \tag{1.7}$$

and, that for $\lambda \in \mathbb{R}$ a continuity point of $N(\lambda)$,

$$\lim_{k\to+\infty} \mathbb{E}\left(n(H_{\omega,k},\lambda)\right) = N(\lambda). \tag{1.8}$$

Hence, to prove estimate (1.4), it is enough to prove that, for some constant $C > 0$, for $E \geq 0$ sufficiently small, almost surely, one has

$$n_0(H_{\omega,k}, E/C) \leq n(H_{\omega,k}, E) - n(H_{\omega,k}, 0^-) \leq n_0(H_{\omega,k}, C \cdot E) \tag{1.9}$$

where $n_0(H_{\omega,k},\cdot)$ is the density of states of the periodic discrete operator $\Pi_0 H_{\omega,k}\Pi_0$. Here, we used the fact that $\mathbb{E}(n_0(H_{\omega,k},E/C))$ converges to N_0 as $k\to+\infty$ (see [23]).

To obtain the estimate (1.9), we will use Floquet theory for the periodic operator $H_{\omega,k}$. By $\mathcal{E}_-$ (resp. $\mathcal{E}_+$), we denote the subspace of $L^2(\mathbb{R}^2)$ generated by the vectors $(\hat u_{n,\gamma})_{0\leq n<n_0;\gamma\in\Gamma}$ (resp. $(\hat u_{n,\gamma})_{n_0+p_0\leq n;\gamma\in\Gamma}$) defined above. Let Π_- (resp. Π_+) be the orthogonal projector onto $\mathcal{E}_-$ (resp. $\mathcal{E}_+$). For $\cdot\in\{-,0,+\}$, we define

$$H_{\omega,k,\cdot}=\Pi_\cdot H_{\omega,k}\Pi_\cdot=\Pi_\cdot H\Pi_\cdot+\Pi_\cdot V_{\omega,k}\Pi_\cdot=H_\cdot+V_{\omega,k,\cdot}$$

Under our assumptions, we know that, for some $a>0$, one has

$$H_{\omega,k,+}\geq H_+\geq a\Pi_+. \tag{1.10}$$

We also define $\Pi^+=\Pi_0+\Pi_+$ and

$$H^+_{\omega,k}=\Pi^+H_{\omega,k}\Pi^+=\Pi^+H\Pi^++\Pi^+V_{\omega,k}\Pi^+=H^++V^+_{\omega,k}.$$

The operators $(H_{\omega,k,\cdot})_{\cdot\in\{-,0,+\}}$ and $H^+_{\omega,k}$ are $(2k+1)\Gamma$-periodic; we can then consider their Floquet decomposition $H_{\omega,k,\cdot}(\theta)$ for $\theta\in\mathbb{T}^2_k=\mathbb{R}^2/((2k+1)\Gamma)^*$. Define

$$\vartheta(H_{\omega,k,\cdot}(\theta),E)=\sharp\{\text{eigenvalues of } H_{\omega,k,\cdot}(\theta) \text{ less than or equal to } E\}.$$

The notation $\vartheta(H_{\omega,k,\cdot}(\theta),E^-)$ will designate the number of eigenvalues of $H_{\omega,k,\cdot}(\theta)$ strictly less than E. Let us stress that, when counting the eigenvalues of $H_{\omega,k,\cdot}(\theta)$, we consider this operator as acting on $\mathcal{E}_\cdot$.

The estimate (1.9) is an immediate consequence of

Lemma 1.1. *There exists $E_0>0$ and $C>0$ such that, for $0\leq E\leq E_0$, for $k\geq1$ and for $\theta\in\mathbb{T}^2_k$, we have*

$$\vartheta\left(H_{\omega,k,0}(\theta),E/C\right)\leq\vartheta(H_{\omega,k}(\theta),E)-\vartheta(H_{\omega,k}(\theta),0^-)\leq\vartheta\left(H_{\omega,k,0}(\theta),C\cdot E\right). \tag{1.11}$$

Indeed, to get estimate (1.9) from Lemma 1.1, one only needs to integrate (1.11) with respect to θ (see [23]).
Lemma 1.1 is an immediate consequence of the following two results

Lemma 1.2. *Let $\mathcal{H}$ be a separable Hilbert space. Fix $\delta>0$. There exists $C>1$ such that, for any A and B self-adjoint operators on the same dense domain $\mathcal{D}\subset\mathcal{H}$ satisfying*

1. *A and B have compact resolvent and are both lower semi-bounded,*
2. *$V=B-A$ is non negative and V satisfies $\|V\|\leq1/\delta$,*
3. *the operator $A_t=A+tV$ has no spectrum in the interval $(-\delta,0)$ for $t\in[0,1]$,*

if χ_+ denotes the spectral projector of A onto the energy interval $[0,+\infty)$ (i.e. $\chi_+=\mathbf{1}_{[0,+\infty)}(A)$), then, for $E\geq0$

$$\vartheta(\chi_+B\chi_+,E/C)\leq\vartheta(B,E)-\vartheta(B,0^-)\leq\vartheta(\chi_+B\chi_+,E). \tag{1.12}$$

Note that when counting the eigenvalues of $\chi_+B\chi_+$, we consider this operator as acting on $\chi_+\mathcal{H}$.

and

Lemma 1.3. *Let $\mathcal{H}$ be a separable Hilbert space. Fix $\delta>0$. There exists $C>1$ such that, for any A and B positive, self-adjoint operators on the same dense domain $\mathcal{D}\subset\mathcal{H}$ satisfying*

1. *A and B have compact resolvent,*
2. *$V=B-A$ is non negative and V satisfies $\|V\|\leq1/\delta$,*

if χ denotes a spectral projector of A (i.e. $[\chi,A]=0$) such that $(1-\chi)A(1-\chi)\geq\delta(1-\chi)$, then, for $0\leq E\leq\delta/2$, one has

$$\vartheta(\chi B\chi,E)\leq\vartheta(B,E)\leq\vartheta(\chi B\chi,C\cdot E). \tag{1.13}$$

Note that when counting the eigenvalues of $\chi B\chi$, we consider this operator as acting on $\chi\mathcal{H}$.

The proof of Lemma 1.1 is completed in a very simple way using Lemma 1.2 and 1.3. By assumption (H.1.3') and using the fact that the random potential $V_{\omega,k}$ is bounded uniformly in ω and k, we can apply Lemma 1.2 to $H_{\omega,k}$ (taking $\chi_+ = \Pi^+$), and then Lemma 1.3 to $H^+_{\omega,k}$ (taking $\chi = \Pi_0$). This yields Lemma 1.1. □

Proof of Lemma 1.2. We will first prove the upper bound. Assumption (1) and (2) in Lemma 1.2 guaranty that A_t has discrete spectrum and that the eigenvalues of A_t are continuous increasing functions of t.
By assumption (3) in Lemma 1.2, we obtain that, for any $t \in [0,1]$, one has

$$\vartheta(A, 0^-) = \vartheta(A_t, 0^-) = \vartheta(B, 0^-). \tag{1.14}$$

Let us denote this number by r. If we denote the eigenvalues of B by $(E_n(B))_{n\geq 1}$ (we order them increasingly and repeat them according to multiplicity), the min-max formula for lower semi-bounded operators (see [37]) tells us that

$$E_n(B) = \sup_{\substack{\mathcal{E}\subset\mathcal{H}\\ \dim(\mathcal{E})=n-1}} \inf_{\substack{\varphi\perp\mathcal{E}\\ \|\varphi\|=1\\ \varphi\in\mathcal{D}(H)}} \langle B\varphi, \varphi\rangle. \tag{1.15}$$

So, if $n \geq r+1$, in the variational formula (1.15), we may choose $\mathcal{E}$ so that it contains $\chi_-\mathcal{H}$ where $\chi_- = 1 - \chi_+$; this gives us that, for $n \geq r+1$,

$$E_n(B) \geq \sup_{\substack{\mathcal{E}\subset\mathcal{H}\\ \dim(\mathcal{E})=n-1-r}} \inf_{\substack{\varphi\perp\mathcal{E}\oplus(\chi_-\mathcal{H})\\ \|\varphi\|=1\\ \varphi\in\mathcal{D}}} \langle B\varphi, \varphi\rangle = E_{n-r}(\chi_+ B\chi_+)$$

where $E_n(\chi_+ B\chi_+)$ denotes the n-th eigenvalue of $\chi_+ B\chi_+$ (acting on $\chi_+\mathcal{H}$). This completes the proof of the upper bound in (1.12).

Let us prove the lower bound of (1.12). First note that equation (1.14) stays valid if we replace A and B by λA and λB where $\lambda > 0$. Let r denote the number of negative eigenvalues of A.
Equation (1.14) implies that, if $\varphi \in \mathcal{D}$ is orthogonal to $\mathbf{1}_{(-\infty,0)}(B)\mathcal{H}$, then $\langle A\varphi, \varphi\rangle \geq 0$. Indeed, if this would not be the case, as $\mathbf{1}_{(-\infty,0)}(B)A\mathbf{1}_{(-\infty,0)}(B) \leq \mathbf{1}_{(-\infty,0)}(B)B\mathbf{1}_{(-\infty,0)}(B) < 0$, one could find a $r+1$-dimensional subspace of $\mathcal{H}$ such that A be negative on this subspace; hence, the number of negative eigenvalues of A would be at least $r+1$.
Therefrom, we deduce that, for $n \geq 1$ and $t \in (0,1]$, one has

$$t \cdot E_{n+r}(B) \leq E_{n+r}(A_t). \tag{1.16}$$

Indeed, if we let $\mathcal{E}$ be the space generated by the $n+r-1$-st eigenvectors of B, the min-max principle tells us that

$$E_n(A_t) \geq \inf_{\substack{\varphi\perp\mathcal{E}\\ \|\varphi\|=1\\ \varphi\in\mathcal{D}}} \langle (A+tV)\varphi, \varphi\rangle. \tag{1.17}$$

As $\mathcal{E}$ contains $\mathbf{1}_{(-\infty,0)}(B)\mathcal{H}$, for $\varphi \in \mathcal{D}$ satisfying $\varphi \perp \mathcal{E}$, we know that $\langle A\varphi, \varphi\rangle \geq 0$, hence, that $\langle A\varphi, \varphi\rangle \geq t\langle A\varphi, \varphi\rangle$. Plugging this into equation (1.17), we obtain

$$E_n(A_t) \geq \inf_{\substack{\varphi\perp\mathcal{E}\\ \|\varphi\|=1\\ \varphi\in\mathcal{D}}} \langle (tA+tV)\varphi, \varphi\rangle = t \cdot E_{n+r}(B).$$

This implies that, for $t \in [0,1]$, one has

$$\vartheta(A_t, E) - \vartheta(A_t, 0^-) \leq \vartheta(tB, E) - \vartheta(tB, 0^-). \tag{1.18}$$

Now, decompose $A + tV - E$ over the orthogonal direct sum $\chi_+\mathcal{H} \overset{\perp}{\oplus} \chi_-\mathcal{H}$, we get the matrix

$$M(t,E) = \begin{pmatrix} \chi_+(A+tV)\chi_+ - E\chi_+ & t\chi_+V\chi_- \\ t\chi_-V\chi_+ & \chi_-(A+tV)\chi_- - E\chi_- \end{pmatrix} \tag{1.19}$$

As A and V satisfy assumptions (2) and (3), for $t \in (0, \delta^2/2]$, one has $\chi_-(A+tV)\chi_- \leq -(\delta/2)\chi_-$. Hence, for $t \in (0, \delta^2/2]$, the lower right entry of the matrix $M(t,E)$ in (1.19) is invertible (as an operator on $\chi_-\mathcal{H}$); so the matrix $M(t,E)$ is invertible if and only if $L(t,E) - E$ is invertible (as an operator on $\chi_+\mathcal{H}$) where

$$L(t,E) = \chi_+(A+tV)\chi_+ + t^2\chi_+V\chi_-(E - \chi_-(A+tV)\chi_-)^{-1}\chi_-V\chi_+.$$

Moreover, $\varphi \in \mathcal{D}$ satisfies $(A + tV - E)\varphi = 0$ if and only if one has

$$\begin{cases} (L(t,E) - E)(\chi_+\varphi) = 0 \\ \chi_-\varphi = \left[\chi_-(E - \chi_-(A+tV)\chi_-)^{-1}\chi_-V\chi_+\right] \cdot (\chi_+\varphi) \end{cases}$$

This proves that, for $t \in (0, \delta^2/2]$, E is a non negative eigenvalue of multiplicity m of $A + tV$ if and only if E is an eigenvalue of multiplicity m of the operator $L(t,E)$ acting on $\chi_+\mathcal{H}$.

Now, we note that $L(t,E)$ is decreasing and real analytic in E. Hence, if $(E_n(E,t))_{n\geq 1}$ are the eigenvalues of $L(t,E)$ (ordered increasingly and repeated according to multiplicity), the functions $E \mapsto E_n(E,t)$ are continuous and decreasing; so the functions $E \mapsto f_n(E,t) = E_n(E,t) - E$ are strictly decreasing and continuous. Hence, each of them vanishes exactly once on the interval $[0, +\infty)$ (as $f_n(0,t) \geq 0$ and $f_n(E,t) \to -\infty$ as $E \to +\infty$). Let $E_n^*(t)$ be the zero of $f_n(E,t)$ in $[0, +\infty)$ i.e

$$E_n(E_n^*(t), t) = E_n^*(t). \tag{1.20}$$

The sequence $(E_n^*(t))_{n\geq 1}$ is increasing; moreover, E is an eigenvalue of multiplicity m of $L(t,E)$ if and only if $E_n(E,t) = E$ for m values of n that is if and only if $\#\{n \geq 1; E_n^*(t) = E\} = m$.

We will now prove that the sequence $(E_n^*(t))_{n\geq 1}$ coincides with the sequence of the non negative eigenvalues of $A + tV$.

Let $(E_n^+(t))_{n\geq 1}$ be the non negative eigenvalues of $A + tV$ (ordered increasingly and repeated according to multiplicity). First, note that, by what has been said above, $\{E_n^*(t); n \geq 1\} = \{E_n^+(t); n \geq 1\}$ and, for any $E \geq 0$, $\#\{n \geq 1; E_n^*(t) = E\} = \#\{n \geq 1; E_n^+(t) = E\}$. This implies that $E_1^*(t) = E_1^+(t)$. Now either $E_1^+(t)$ is a single eigenvalue, then we have $\{E_n^*(t); n \geq 2\} = \{E_n^+(t); n \geq 2\}$ and, for any $E \geq 0$, $\#\{n \geq 2; E_n^*(t) = E\} = \#\{n \geq 2; E_n^+(t) = E\}$. Or $E_1^+(t)$ is a multiple eigenvalue with multiplicity larger than 2 i.e. $E_2^+(t) = E_1^+(t)$. In this case, $E_1^*(t)$ is an eigenvalue of multiplicity at least 2 of $L(t, E_1^*(t))$ i.e. $f_2(E_1^*(t), t) = 0$ so that $E_2^*(t) = E_1(t)$. So, we obtain again that $\{E_n^*(t); n \geq 2\} = \{E_n^+(t); n \geq 2\}$ and , for any $E \geq 0$, $\#\{n \geq 2; E_n^*(t) = E\} = \#\{n \geq 2; E_n^+(t) = E\}$. An induction then immediately gives that, for any $k \geq 1$, $\{E_n^*(t); n \geq k\} = \{E_n^+(t); n \geq k\}$ and , for any $E \geq 0$, $\#\{n \geq k; E_n^*(t) = E\} = \#\{n \geq k; E_n^+(t) = E\}$. Thus, one proves that, for $n \geq 1$, $E_n^*(t) = E_n^+(t)$.

Taking (1.20) into account, we see that the n-th non negative eigenvalue of $A + tV$ satisfies

$$E_n(E_n^+(t), t) = E_n^+(t). \tag{1.21}$$

For $\varphi \in \mathcal{D}$, using assumptions (2) and (3), we compute

$$\langle \varphi, \chi_+V\chi_-(E - \chi_-(A+tV)\chi_-)^{-1}\chi_-V\chi_+\varphi \rangle \leq \frac{1}{\delta^2} \cdot \langle \varphi, \chi_+V\chi_+\varphi \rangle.$$

Plugging this into the definition of $L(t,E)$, and using the fact that $\chi_+A\chi_+ \geq 0$, we obtain that, for $0 < t \leq \delta$ and $E \geq 0$, one has

$$L(t,E) \leq \chi_+(A+2V)\chi_+ \leq 2\chi_+(A+V)\chi_+ = 2\chi_+B\chi_+.$$

So that, if $(E_n(\chi_+B\chi_+))_{n\geq 1}$ are the eigenvalues of $\chi_+B\chi_+$ (ordered increasingly and repeated according to multiplicity), for $0 < t \leq \delta$ and $E \geq 0$, we have

$$E_n(E,t) \leq 2E_n(\chi_+B\chi_+).$$

Combining this with (1.21), for $0 < t \leq \inf(\delta^2/2, \delta)$, one gets

$$E_n^+(t) \leq 2E_n(\chi_+B\chi_+).$$

Hence, for $0 < t \leq \inf(\delta^2/2, \delta)$, one has

$$\vartheta(\chi_+B\chi_+, E/2) \leq \vartheta(A_t, E) - \vartheta(A_t, 0^-).$$

Combining this with (1.18), we complete the proof of the lower bound in Lemma 1.2. $\square$

Proof of Lemma 1.3. Let us start with the proof of the lower bound. For C, a self-adjoint operator with discrete spectrum on a Hilbert space $\mathcal{H}$, one has

$$\vartheta(C,E) = \sup\{\text{rank}(\mathcal{E});\ \mathcal{E} \subset \mathcal{H},\ \text{linear s.t. } \forall \varphi \in \mathcal{E} \subset \mathcal{H},\ \|\varphi\| = 1,\ \langle B\varphi, \varphi\rangle \leq E\}.$$

This immediately implies the lower bound in (1.13).

The upper bound is proved as follows. We decompose $\mathcal{H}$ into $\mathcal{H} = (\chi\mathcal{H}) \overset{\perp}{\oplus} (\chi_+\mathcal{H})$ where $\chi_+ = 1 - \chi$. This induces the following matrix form for $B = A + V$

$$B = \begin{pmatrix} \chi B\chi & \chi V\chi_+ \\ \chi_+V\chi & \chi_+B\chi_+ \end{pmatrix}.$$

As $\chi_+A\chi_+ \geq \delta\cdot\chi_+$ and $V \geq 0$, the upper bound in Lemma 1.3 is then an immediate consequence of

Lemma 1.4. *There exists $C > 0$ depending only on δ such that*

$$\begin{pmatrix} \chi B\chi & \chi V\chi_+ \\ \chi_+V\chi & \chi_+B\chi_+ \end{pmatrix} \geq \frac{1}{C}\begin{pmatrix} \chi B\chi & 0 \\ 0 & \chi_+B\chi_+ \end{pmatrix}.$$

Proof. Lemma 1.4 was proved in [23]; but, for the sake of completeness, we will give this proof here.

First, note that, for $\varphi \in \mathcal{D}$, one has

$$\langle \chi B\chi\varphi, \varphi\rangle + \langle \chi_+B\chi_+\varphi, \varphi\rangle \leq \langle B\varphi, \varphi\rangle + 2\left|\langle \chi_+V\chi\varphi, \varphi\rangle\right|.$$

If we prove that, for some $C > 0$ depending only on δ, one has

$$|\langle \chi_+V\chi\varphi, \varphi\rangle| = |\langle \chi_+B\chi\varphi, \varphi\rangle| \leq C\langle B\varphi, \varphi\rangle, \tag{1.22}$$

we will have completed the proof of Lemma 1.4.

By assumption (2), we know that for some $C > 0$ depending only on δ, we have

- $\chi_+A\chi_+ \geq \delta\chi_+$;
- $\|(\chi_+A\chi_+)^{-1/2}V(\chi_+A\chi_+)^{-1/2}\| \leq C$.

This implies that, for any $\varphi \in (\chi_+\mathcal{H}) \cap \mathcal{D}$, one has

$$|\langle V\chi_+\varphi, \chi_+\varphi\rangle| \leq C\langle A\chi_+\varphi, \chi_+\varphi\rangle \leq C\langle A\varphi, \varphi\rangle.$$

One computes

$$\begin{aligned} 2|\langle \chi_+V\chi\varphi, \varphi\rangle| &\leq 2|\langle V^{1/2}\varphi, \tilde{V}^{1/2}\chi_+\varphi\rangle| + 2\langle V\chi_+\varphi, \chi_+\varphi\rangle \\ &\leq 2\|V^{1/2}\varphi\| \cdot \|V^{1/2}\chi_+\varphi\| + 2\langle \chi_+V\chi_+\varphi, \varphi\rangle \\ &\leq \langle V\varphi, \varphi\rangle + 3\langle \chi_+V\chi_+\varphi, \varphi\rangle \\ &\leq \langle V\varphi, \varphi\rangle + C\langle A\varphi, \varphi\rangle \\ &\leq C\langle B\varphi, \varphi\rangle. \end{aligned}$$

for some $C > 1$ depending only on δ. This completes the proof of Lemma 1.4. $\square$

Remark 1.2. In Lemmas 1.2 and 1.3, the boundedness assumption on V can be replaced by a relative form boundedness assumption.

In Lemma 1.2 and 1.3 we have only studied the reduction in the case of a lower band edge. To deal with an upper band edge, one proves an analogue of Lemma 1.2 namely

Lemma 1.5. *Let $\mathcal{H}$ be a separable Hilbert space. Fix $\delta > 0$. There exists $C > 1$ such that, for any A and B self-adjoint operators on the same dense domain $\mathcal{D} \subset \mathcal{H}$ satisfying*

1. *A and B have compact resolvent and are both lower semi-bounded,*
2. *$V = A - B$ is non negative and V satisfies $\|V\| \leq 1/\delta$,*
3. *the operator $A_t = A - tV$ has no spectrum in the interval $(0, \delta)$ for $t \in [0, 1]$,*

if χ_- denotes the spectral projector of A onto the energy interval $(-\infty, 0]$ (i.e. $\chi_+ = \mathbf{1}_{(-\infty,0]}(A)$), then, for $E \geq 0$,

$$\vartheta(-(\chi_- B\chi_-), E/C) \leq \vartheta(B, 0) - \vartheta(B, -E) \leq \vartheta(-(\chi_- B\chi_-), E). \tag{1.23}$$

Note that when counting the eigenvalues of $-(\chi_- B\chi_-)$, we consider this operator as acting on $\chi_-\mathcal{H}$.

We note here that $-(\chi_- B\chi_-) \geq -(\chi_- A\chi_-) \geq 0$.

Proof of Lemma 1.5. Let us start with the proof of the upper bound. By assumptions (2) and (3) of Lemma 1.5, for $t \in [0, 1]$, we know that

$$\vartheta(B, 0) = \vartheta(A_t, 0) = \vartheta(A, 0). \tag{1.24}$$

Denote this number by r. Pick $E \leq 0$. Then, the proof of the lower bound in Lemma 1.3 yields $\vartheta(\chi_- B\chi_-, E) \leq \vartheta(B, E)$. On the other hand, as V is non negative, we have $\vartheta(\chi_- B\chi_-, E) \leq \vartheta(\chi_- A\chi_-, E)$. Combining this with (1.24) at energy $E = 0$, we obtain $\vartheta(\chi_- B\chi_-, 0) = \vartheta(B, 0)$ and

$$\vartheta(B, 0) - \vartheta(B, E) \leq \vartheta(\chi_- B\chi_-, 0) - \vartheta(\chi_- B\chi_-, E) \tag{1.25}$$

Using the fact that $\chi_-\mathcal{H}$ is of dimension r and that $-\chi_- B\chi_- \geq 0$, we see that the right hand side of (1.25) is nothing but $\vartheta(-\chi_- B\chi_-, -E)$. This proves the upper bound in equation (1.23).

Turn now to the proof of the lower bound. We notice that, if $\varphi \in \mathcal{D}$ is orthogonal to $\mathbf{1}_{(0,+\infty)}(B)\mathcal{H}$, then $\langle A\varphi, \varphi\rangle \leq 0$. Indeed, if this were not the case, as $\langle B\varphi, \varphi\rangle \leq \langle A\varphi, \varphi\rangle$, one would have at most an $r-1$-dimensional subspace of $\mathcal{H}$ such that A be negative on this subspace. This contradicts (1.24).

Therefrom, we deduce that, for $1 \leq n \leq r$ and $t \in (0, 1]$, one has

$$t \cdot E_n(B) \geq E_n(A_t).$$

Indeed, if we let $\mathcal{E}$ be the space generated by the $r - 1$-st eigenvectors of A_t, the min-max principle tells us that

$$E_n(t \cdot B) \geq \inf_{\substack{\varphi \perp (\mathcal{E} \oplus \mathbf{1}_{(0,+\infty)}(B)\mathcal{H}) \\ \|\varphi\|=1 \\ \varphi \in \mathcal{D}}} \langle (tA + tV)\varphi, \varphi\rangle \tag{1.26}$$

as B is non negative on $\mathbf{1}_{(0,+\infty)}(B)\mathcal{H}$. But, for $\varphi \in \mathcal{D}$ satisfying $\varphi \perp \mathcal{E} \oplus \mathbf{1}_{(0,+\infty)}(B)\mathcal{H}$, we know that $\langle A\varphi, \varphi\rangle \leq 0$, hence, that $\langle A\varphi, \varphi\rangle \leq t\langle A\varphi, \varphi\rangle$. Plugging this into equation (1.26), we obtain

$$t \cdot E_n(B) \geq \inf_{\substack{\varphi \perp (\mathcal{E} \oplus \mathbf{1}_{(0,+\infty)}(B)\mathcal{H}) \\ \|\varphi\|=1 \\ \varphi \in \mathcal{D}}} \langle (A - tV)\varphi, \varphi\rangle \geq E_n(A_t).$$

This implies

$$\vartheta(A_t, 0) - \vartheta(A_t, E) \leq \vartheta(tB, 0) - \vartheta(tB, E).$$

To complete the proof of Lemma 1.5 from this step on, one follows the steps of the end of the proof of Lemma 1.2. $\square$

1.3. **The study of N_0.** Equation (1.4) will be the starting point of our analysis. Our aim will now be to compare N_0 with the density of states of a well chosen simple discrete model. Using the Wannier basis for $\mathcal{E}_0$, the operator $\Pi_0 H_\omega \Pi_0$ can be considered as acting on $L^2(\mathbb{T}^2)\otimes\mathbb{C}^{p_0}$ (or on $\ell^2(\Gamma)\otimes\mathbb{C}^{p_0}$ conjugating with the discrete Fourier transform). We will now show that this operator can be efficiently compared to a new operator that is regular i.e. that has smooth kernel when considered acting on $L^2(\mathbb{T}^2)\otimes\mathbb{C}^{p_0}$ (or, in the Fourier picture, that has off-diagonal coefficients that are rapidly decaying when considered acting on $\ell^2(\Gamma)\otimes\mathbb{C}^{p_0}$). Here, "efficiently" is meant in the sense that it does not alter the behavior of the integrated density of states too much. The regularity that we gain through this change will be crucial in the rest of our argument.
Let $\mathcal{M}at_{p_0}(\mathbb{C})$ denote the space of $p_0\times p_0$ complex matrices and $\mathcal{H}erm_{p_0}(\mathbb{C})$, the subspace of hermitian matrices. We prove

Theorem 1.2. *There exists*

1. $\mathcal{U}:\ \Pi_0(L^2(\mathbb{R}^d))\to L^2(\mathbb{T}^2)\otimes\mathbb{C}^{p_0}$ *a unitary equivalence satisfying* $\forall\gamma\in\Gamma$, $U\circ\tau_\gamma=e_\gamma\circ U$ *where* e_γ *denotes the multiplication by the function* $\theta\mapsto e^{i\gamma\cdot\theta}$ *acting on* $L^2(\mathbb{T}^2)\otimes\mathbb{C}^{p_0}$,
2. $h:\ \mathbb{T}^2\to\mathcal{H}erm_{p_0}(\mathbb{C})$, *a smooth function,*
3. $G:\ \mathbb{T}^2\times\mathbb{T}^2\to\mathcal{M}at_{p_0}(\mathbb{C})$ *a smooth function such that, for any* $\theta\in\mathbb{T}^2$, $G(\theta,\theta)$ *is a positive definite matrix,*

and a constant $C>1$ *such that:*

$$\frac{1}{C}h\le\mathcal{U}\Pi_0H_0\Pi_0\mathcal{U}^*\le Ch\quad\text{and}\quad\frac{1}{C}(h+G_\omega)\le\mathcal{U}\Pi_0H_\omega\Pi_0\mathcal{U}^*\le C(h+G_\omega).\tag{1.27}$$

Here, h is considered as a multiplication operator on $L^2(\mathbb{T}^2)\otimes\mathbb{C}^{p_0}$, and G_ω is the operator with the kernel

$$G_\omega(\theta,\theta')=\sum_{\gamma\in\Gamma}\omega_\gamma G(\theta,\theta')e^{i\gamma(\theta-\theta')}.\tag{1.28}$$

Obviously, to estimate the density of states N_0, it is now enough to estimate the density of states, say N^*, of the operator $h_\omega=h+G_\omega$. This will be done in the section 2, 3, 4 and 5. The rest of this section will be devoted to the proof of Theorem 1.2.

1.4. **The proof of Theorem 1.2.** For $\theta\in\mathbb{T}^2$, let $(u_n(\cdot,\theta))_{n_0\le n\le n_0+p_0-1}$ be an orthonormal system of vectors that spans the range of $\Pi_\theta^{\mathcal{E}_0}$ (see section 1.1.2). Then, $\Pi_0H_\omega\Pi_0$ is unitarily equivalent to the operator $\tilde h_\omega$ acting on $L^2(\mathbb{T}^2)\otimes\mathbb{C}^{p_0}$ defined by

$$(\tilde h_\omega\phi)(\theta)=([h+\tilde v_\omega]\phi)(\theta)=h(\theta)\phi(\theta)+\int_{\mathbb{T}^2}\tilde v_\omega(\theta,\theta')\phi(\theta')d\theta',\quad \phi\in L^2(\mathbb{T}^2)\otimes\mathbb{C}^{p_0}\tag{1.29}$$

where

$$h(\theta)=\begin{pmatrix}\langle Hu_{n_0}(\theta),u_{n_0}(\theta)\rangle & \cdots & \langle Hu_{n_0}(\theta),u_{n_0+p_0-1}(\theta)\rangle\\ \vdots & \ddots & \vdots\\ \langle Hu_{n_0+p_0-1}(\theta),u_{n_0}(\theta)\rangle & \cdots & \langle Hu_{n_0-p_0-1}(\theta),u_{n_0+p_0-1}(\theta)\rangle\end{pmatrix}\tag{1.30}$$

where $\langle\cdot,\cdot\rangle$ denotes the scalar product in $L^2(C_0)$ and

$$\tilde v_\omega(\theta,\theta')=\begin{pmatrix}\tilde v_{\omega,0,0}(\theta,\theta') & \cdots & \tilde v_{\omega,0,p_0-1}(\theta,\theta')\\ \vdots & \ddots & \vdots\\ \tilde v_{\omega,p_0-1,0}(\theta,\theta') & \cdots & \tilde v_{\omega,p_0-1,p_0-1}(\theta,\theta')\end{pmatrix}\tag{1.31}$$

where

$$\tilde{v}_{\omega,j,k}(\theta,\theta') = \sum_{\gamma\in\Gamma} \omega_\gamma \int_{\mathbb{R}^2} V(x-\gamma) u_{n_0+i}(x,\theta)\overline{u_{n_0+j}(x,\theta')}dx.$$

The vectors $(u_n)_{n_0\leq n\leq n_0+p_0-1}$ have been continued to $\mathbb{R}^2$ using the quasi-periodic boundary conditions they satisfy. Note that, if we chose the $(u_n)_{n_0\leq n\leq n_0+p_0-1}$ to be the Floquet eigenvectors, we get that $\langle Hu_i(\theta), u_j(\theta)\rangle = \delta_{ij}E_i(\theta)$ i.e. $h(\theta)$ is diagonal.
The operator $\tilde{v}_\omega$ is bounded on $L^2(\mathbb{T}^2)\otimes\mathbb{C}^{p_0}$ as the $(v_j(x,\theta))_{1\leq j\leq p_0}$ generate $\Pi_\theta^{\mathcal{E}_0}\mathcal{H}_\theta$ (on which H_θ is bounded) and as V is relatively bounded with respect to H.

The unitary equivalence $\mathcal{T}^* :\ L^2(\mathbb{T}^2)\otimes\mathbb{C}^{p_0} \to \Pi_0(L^2(\mathbb{R}^d))$ such that $\Pi_0 H_\omega \Pi_0 = \mathcal{T}^*\tilde{h}_\omega\mathcal{T}$ is given by

$$(\mathcal{T}^*v)(x,\theta) = \sum_{1\leq j\leq p_0} v_j u_{n_0+j}(x,\theta), \quad v=(v_j)_{1\leq j\leq p_0}\in L^2(\mathbb{T}^2)\otimes\mathbb{C}^{p_0}.$$

As the functions $(u_{n_0+j})_{0\leq j\leq p_0-1}$ satisfy the quasi-periodic boundary conditions (0.6), we see that $\mathcal{T}$ satisfies: $\forall\gamma\in\Gamma$, $\mathcal{T}\circ\tau_\gamma = e_\gamma\circ\mathcal{T}$.

The orthonormal system $(u_n)_{n_0\leq n\leq n_0+p_0-1}$ can be chosen arbitrarily. The main drawback of choosing the Floquet eigenvectors is that the functions $h(\theta)$ and $\tilde{v}_\omega(\theta,\theta')$ will not in general be smooth. The crucial remark is that, as we are interested in the integrated density of states of $\tilde{h}_\omega$ near 0 and as h and $\tilde{v}_\omega$ are non negative, the behavior of $h(\theta)$ near the points θ where $h(\theta)$ does not vanish (i.e. where its determinant does not vanish) will not be very important. So that we only need the functions $h(\theta)$ and $\tilde{v}_\omega(\theta,\theta')$ to be smooth near the points where the determinant of $h(\theta)$ vanishes i.e. near the set $\mathcal{S}=\{\theta\in\mathbb{T}^2;\ \exists 1\leq p\leq p_0,\ E_{n_0+p-1}(\theta)=0\}$.

The structure of $\mathcal{S}$ will be important. As the Floquet eigenvalues are continuous (see Lemma 7.1), $\mathcal{S}$ is compact. We will use

Lemma 1.6. *There exists $(U_k^p)_{1\leq p\leq p_0;1\leq k\leq k_p}$, a finite collection of connected open sets covering $\mathcal{S}$ such that, for all $1\leq p\leq p_0$ and $1\leq k\leq k_p$,*

1. *For some $\varepsilon_{p,k}>0$, the number of Floquet eigenvalues $(E_n)_{n_0\leq n\leq n_0+p_0-1}$ such that $0\leq E_n(\theta)\leq\varepsilon_{k,p}$ is constant and equal to p i.e. the spectral space associated to H_θ and the energy interval $[0,\epsilon_{p,k}]$ has constant dimension p.*
2. *For $\theta\in U_k^p$, there exists a system of vectors $(v_j^{p,k}(\cdot,\theta))_{1\leq j\leq p}$ in $\mathcal{H}_\theta$ such that*
   - *In U_k^p, for $1\leq j\leq p$, the mapping $\theta\in\mathbb{T}^2\mapsto v_j^{p,k}(\cdot,\theta)\in\mathcal{H}_\theta^2$ is smooth,*
   - *For $\theta\in U_k^p$, $span(v_j^{p,k}(\cdot,\theta))_{1\leq j\leq p}$ $=span(u_{n_0+j}(\cdot,\theta))_{1\leq j\leq p}$ i.e. the system of vectors $(v_j^{p,k}(\cdot,\theta))_{1\leq j\leq p}$ spans the spectral space associated to H_θ and to the energy interval $[0,\epsilon_{p,k}]$,*
   - *For $\theta\in U_k^p$, the system $(v_j^{p,k}(\cdot,\theta))_{1\leq j\leq p}$ is orthonormal,*
3. *If $p\leq q$, then, for $\theta\in U_k^p\cap U_l^q$, one has $(v_j^{p,k}(\cdot,\theta))_{1\leq j\leq p} = (v_j^{q,l}(\cdot,\theta))_{1\leq j\leq p}$.*

Moreover, if $p_0=1$ then one can pick $k_{p_0}=1$ and $(v_1^{p_0,k_{p_0}}(\cdot,\theta))$ to be a normalized vector associated to the eigenvalue $E_{n_0}(\theta)$ of H_θ for $\theta\in U_{k_{p_0}}^{p_0}$.

Remark 1.3. The $\mathcal{H}_\theta$-valued function $\theta\mapsto v_j(\cdot,\theta)$ is said to be smooth (resp. analytic) if and only if the $\mathcal{H}_{\theta_0}$-valued function $\theta\mapsto(T_{\theta\to\theta_0}v_j)(\cdot,\theta):=e^{ix(\theta_0-\theta)}v_j(\cdot,\theta)$ is smooth (resp. analytic). The mapping $T_{\theta\to\theta_0}:\ \mathcal{H}_\theta\to\mathcal{H}_{\theta_0}$ is an isometry.

We postpone the proof of this result to the end of this section. Let us now use it to complete our reduction. For $1\leq p\leq p_0$, let $U^p=\cup_k U_k^p$, and, define the family of vectors $(v_j(\cdot,\theta))_{1\leq j\leq p_0}$ in the following inductive way

1. For $\theta \in U^{p_0}$, let
$$v_j(\cdot,\theta) = v_j^{p_0,k}(\cdot,\theta) \text{ for } 1 \leq j \leq p_0.$$
2. For $1 \leq p \leq p_0 - 1$ and $\theta \in U^p \setminus \cup_{q>p} U^q$, let
$$v_j(\cdot,\theta) = \begin{cases} v_j^{p,k}(\cdot,\theta) \text{ for } 1 \leq j \leq p, \\ u_{n_0+j}(\cdot,\theta) \text{ for } p+1 \leq j \leq p_0. \end{cases}$$
3. For $\theta \notin \cup_{1\leq p\leq p_0} U^q$, let
$$v_j(\cdot,\theta) = u_{n_0+j}(\cdot,\theta) \text{ for } 1 \leq j \leq p_0.$$

Then, by point (2) of Lemma 1.6, for all $\theta \in \mathbb{T}^2$, $(v_j(\cdot,\theta))_{1\leq j\leq p_0}$ is an orthonormal basis of $\Pi_\theta^{\mathcal{E}_0}\mathcal{H}_\omega$. Moreover, by points (2) and (3) of Lemma 1.6, the family of vectors $(v_j(\cdot,\theta))_{1\leq j\leq p}$ is smooth for $\theta \in U^p$.

We write $\Pi_0 H_\omega \Pi_0$ in the basis $(v_j(\cdot,\theta))_{1\leq j\leq p_0}$ to obtain an operator described by equations (1.29), (1.30) and (1.31) where the vectors $(u_j(\cdot,\theta))_{1\leq j\leq p_0}$ are replaced with the vectors $(v_j(\cdot,\theta))_{1\leq j\leq p_0}$. One gets that, there exists $\eta > 0$ such that h has the following form:
1. For $\theta \in U^{p_0}$, $h(\theta)$ is smooth.
2. For $1 \leq p \leq p_0 - 1$ and $\theta \in U^p \setminus \cup_{q>p} U^q$,
$$h(\theta) = \begin{pmatrix} h^p(\theta) & 0 \\ 0 & h'_p(\theta) \end{pmatrix}$$
where the $p \times p$-matrix h^p is smooth and , the $(p_0 - p) \times (p_0 - p)$-matrix h'_p is the diagonal matrix with entries $(E_{n_0+j})_{p+1\leq j\leq p_0-1}$; h'_p is larger than η (for some $\eta > 0$).
3. For $\theta \notin \cup_{1\leq p\leq p_0} U^q$, $h(\theta)$ the diagonal matrix with entries $(E_{n_0+j})_{0\leq j\leq p_0-1}$; it is larger than η.

Now, by points (2) and (3) of Lemma 1.6, if $p \leq q$ and $\theta \in U^p \cap U^q$, then, one has
$$h^q(\theta) = \begin{pmatrix} h^p(\theta) & 0 \\ 0 & h_{pq}(\theta) \end{pmatrix}$$
for some square matrix h_{pq}. Indeed,
- in the case where $p = q$, it is obvious;
- in the case where $p < q$: pick $i \leq p < j \leq q$ and compute

$$\begin{aligned} \langle H_\theta v_j(\cdot,\theta), v_i(\cdot,\theta)\rangle &= \langle H_\theta v_j^{q,k}(\cdot,\theta), v_i^{q,k}(\cdot,\theta)\rangle \quad &&\text{for some } k \\ &= \langle v_j^{q,k}(\cdot,\theta), H_\theta v_i^{q,k}(\cdot,\theta)\rangle \quad &&\text{as } H_\theta \text{ is self-adjoint} \\ &= \langle v_j^{q,k}(\cdot,\theta), H_\theta v_i^{p,k}(\cdot,\theta)\rangle \quad &&\text{using point (3) of Lemma 1.6} \\ &= \sum_{l=1}^{p} a_l \langle v_j^{q,k}(\cdot,\theta), v_i^{p,k}(\cdot,\theta)\rangle = 0 \quad &&\text{using points (2) and (3) of Lemma 1.6} \end{aligned}$$

Let $(\chi_p)_{1\leq p\leq p_0}$ be a smooth partition of unity relative to $(U^p)_{1\leq p\leq p_0}$ and $\mathcal{S}$ i.e.
- $0 \leq \chi_p$, $\sum_p \chi_p \leq 1$ and $\operatorname{supp} \chi_p \subset U^p$,
- $\sum_p \chi_p = 1$ in some open neighborhood of $\mathcal{S}$.

Let $\chi_0 = 1 - \sum_p \chi_p$. Let I_k denote the $k \times k$ identity matrix. Then, there exists $C > 1$ such that, for $\theta \in \mathbb{T}^2$, one has
$$\frac{1}{C} h_0(\theta) \leq h(\theta) \leq C h_0(\theta) \text{ where } h_0(\theta) = \sum_{p=1}^{p_0} \chi_p(\theta) \begin{pmatrix} h^p(\theta) & 0 \\ 0 & I_{p_0-p} \end{pmatrix} + \chi_0(\theta). \tag{1.32}$$

The matrix h_0 is smooth. It will be the background Hamiltonian (i.e. kinetic energy) for our reduced model. Because of the definition of the functions $(\chi_p)_{1\leq p\leq p_0}$, one knows that, for some $C>1$, for any point θ_0 such that the determinant of $h_0(\theta_0)$ vanishes, there exists a neighborhood of θ_0 and an index $1\leq \tilde p_0\leq p_0$ such that, for $\theta\in\theta_0$, one has

$$\frac{1}{C}\begin{pmatrix} h^p(\theta) & 0\\ 0 & I_{p_0-p}\end{pmatrix}\leq h_0(\theta)\leq C\begin{pmatrix} h^p(\theta) & 0\\ 0 & I_{p_0-p}\end{pmatrix}. \tag{1.33}$$

Let us now turn to the the reduction of the potential. Recall that the operator $\tilde v_\omega$ is defined by (1.31) where the vectors $(u_j(\cdot,\theta))_{1\leq j\leq p_0}$ are replaced with the vectors $(v_j(\cdot,\theta))_{1\leq j\leq p_0}$. We will modify $\tilde v_\omega$ to get a smooth kernel. The idea is that one needs only to modify the kernel of $\tilde v_\omega$ near the points where $h_0(\theta)$ is positive. Indeed,on the points where $h_0(\theta)$ vanishes, $\tilde v_\omega$ already is regular, so we do not need to modify it. But, elsewhere, $h_0(\theta)$ is positive, hence, these points do not contribute to the integrated density of states near 0.
For $1\leq i\leq p_0$, pick χ'_i, a smooth function such that $\operatorname{supp}\chi'_i\subset\cup_{p\geq i}U^p$, $0\leq\chi'_i\leq 1$ and $\chi'_i\equiv 1$ on $\cup_{p\geq i}\operatorname{supp}\chi_p$. Pick $\phi=(\phi_j)_{1\leq j\leq p_0}\in L^2(\mathbb{T}^2)\otimes\mathbb{C}^{p_0}$ and compute

$$\begin{aligned}
\langle\tilde v_\omega\phi,\phi\rangle &= \sum_{\gamma\in\Gamma}\omega_\gamma\int_{\mathbb{R}^2}V(x)\left|\sum_{j=1}^{p_0}\int_{\mathbb{T}^2}v_j(x,\theta)e^{i\gamma\theta}\phi_j(\theta)d\theta\right|^2dx\\
&=\sum_{\gamma\in\Gamma}\omega_\gamma\int_{\mathbb{R}^2}V(x)\left|\sum_{j=1}^{p_0}\int_{\mathbb{T}^2}(\chi'_j(\theta)+[1-\chi'_j(\theta)])v_j(x,\theta)e^{i\gamma\theta}\phi_j(\theta)d\theta\right|^2dx\\
&=\sum_{\gamma\in\Gamma}\omega_\gamma\int_{\mathbb{R}^2}V(x)\left[\left|\sum_{1\leq j\leq p_0}\Phi_j(\gamma,x)\right|^2+\left|\sum_{1\leq j\leq p_0}\Phi'_j(\gamma,x)\right|^2+\right.\\
&\qquad\left.+2\mathrm{Re}\left(\sum_{1\leq j\leq p_0}\Phi_j(\gamma,x)\cdot\overline{\sum_{1\leq j\leq p_0}\Phi'_j(\gamma,x)}\right)\right]dx
\end{aligned}$$

where

$$\begin{aligned}
\Phi_j(\gamma,x)&=\sum_{1\leq j\leq p_0}\int_{\mathbb{T}^2}\chi'_j(\theta)v_j(x,\theta)e^{i\gamma\theta}\phi_j(\theta)d\theta,\\
\Phi'_j(\gamma,x)&=\sum_{1\leq j\leq p_0}\int_{\mathbb{T}^2}(1-\chi'_j(\theta))v_j(x,\theta)e^{i\gamma\theta}\phi_j(\theta)d\theta
\end{aligned}$$

Pick $0<\alpha<1$. Then, one has

$$\begin{aligned}
&\left|\langle\tilde v_\omega\phi,\phi\rangle-\sum_{\gamma\in\Gamma}\omega_\gamma\int_{\mathbb{R}^2}V(x)\left[\left|\sum_{1\leq j\leq p_0}\Phi_j(\gamma,x)\right|^2+\left|\sum_{1\leq j\leq p_0}\Phi'_j(\gamma,x)\right|^2\right]dx\right|\leq\\
&\qquad\leq\sum_{\gamma\in\Gamma}\omega_\gamma\int_{\mathbb{R}^2}V(x)\left[\alpha\left|\sum_{1\leq j\leq p_0}\Phi_j(\gamma,x)\right|^2+\frac{1}{\alpha}\left|\sum_{1\leq j\leq p_0}\Phi'_j(\gamma,x)\right|^2\right]dx.
\end{aligned}$$

As the random variables $(\omega_\gamma)_\gamma$ are bounded, we obtain

$$(1-\alpha)\sum_{\gamma\in\Gamma}\omega_\gamma\int_{\mathbb{R}^2}V(x)\left|\sum_{1\le j\le p_0}\Phi_j(\gamma,x)\right|^2dx$$

$$\le\langle\tilde v_\omega\phi,\phi\rangle_{L^2(\mathbb{T}^2)\otimes\mathbb{C}^{p_0}}+C\left(\frac{1}{\alpha}-1\right)\sum_{\gamma\in\Gamma}\int_{\mathbb{R}^2}V(x)\left|\sum_{1\le j\le p_0}\Phi_j'(\gamma,x)\right|^2dx \quad (1.34)$$

and

$$\langle\tilde v_\omega\phi,\phi\rangle_{L^2(\mathbb{T}^2)\otimes\mathbb{C}^{p_0}}\le(1+\alpha)\sum_{\gamma\in\Gamma}\omega_\gamma\int_{\mathbb{R}^2}V(x)\left|\sum_{1\le j\le p_0}\Phi_j(\gamma,x)\right|^2dx+$$

$$+C\left(1+\frac{1}{\alpha}\right)\sum_{\gamma\in\Gamma}\int_{\mathbb{R}^2}V(x)\left|\sum_{1\le j\le p_0}\Phi_j'(\gamma,x)\right|^2dx \quad (1.35)$$

Using Parseval's formula, we compute

$$\sum_{\gamma\in\Gamma}\int_{\mathbb{R}^2}V(x)\left|\sum_{1\le j\le p_0}\Phi_j'(\gamma,x)\right|^2dx=\int_{\mathbb{R}^2}V(x)\int_{\mathbb{T}^2}\left|\sum_{1\le j\le p_0}v_j(x,\theta)(1-\chi_j'(\theta))\phi_j(\theta)\right|^2d\theta dx \quad (1.36)$$

Let $(\chi_c'\phi)(\theta)$ be the vector $((1-\chi_j'(\theta))\phi_j(\theta))_{1\le j\le p_0}$. As $(1-\chi_j')\chi_i=0$ for $i\ge j$, we have that

$$\chi_p(\theta)\begin{pmatrix}h^p(\theta)&0\\0&I_{p_0-p}\end{pmatrix}(\chi_c'\phi)(\theta)=\chi_p(\theta)\begin{pmatrix}0&0\\0&I_{p_0-p}\end{pmatrix}(\chi_c'\phi)(\theta)=$$

$$=\chi_p(\theta)\begin{pmatrix}I_p&0\\0&I_{p_0-p}\end{pmatrix}(\chi_c'\phi)(\theta)=\chi_p(\theta)(\chi_c'\phi)(\theta).$$

Hence

$$h_0(\theta)(\chi_c'\phi)(\theta)=\sum_{p=1}^{p_0}\chi_p(\theta)\begin{pmatrix}h^p(\theta)&0\\0&I_{p_0-p}\end{pmatrix}(\chi_c'\phi)(\theta)+\chi_0(\theta)(\chi_c'\phi)(\theta)=(\chi_c'\phi)(\theta).$$

Plugging this into (1.36), we get that, for some $C>0$,

$$\begin{aligned}\sum_{\gamma\in\Gamma}\int_{\mathbb{R}^2}V(x)\left|\sum_{1\le j\le p_0}\Phi_j'(\gamma,x)\right|^2dx&\le C\langle h_0(\chi_c'\phi),(\chi_c'\phi)\rangle_{L^2(\mathbb{T}^2)\otimes\mathbb{C}^{p_0}}\\&\le C\langle h_0\phi,\phi\rangle_{L^2(\mathbb{T}^2)\otimes\mathbb{C}^{p_0}}.\end{aligned} \quad (1.37)$$

Now, we need to build a replacement for $\tilde v_\omega$ near the points where its kernel is not regular. Therefore, pick a sequence of positive exponentially decaying real numbers $(a_\gamma)_{\gamma\in\Gamma}$ and set

$$g_0(\theta,\theta')=g_0(\theta-\theta')=\sum_{\gamma\in\Gamma}a_\gamma e^{i\gamma(\theta-\theta')}.$$

The operator with kernel g_0 that we also denote by g_0 is compact. By the same calculation as above, for some $C>0$, we have

$$\sum_{\gamma\in\Gamma}\omega_\gamma\sum_{1\le j\le p_0}\int_{\mathbb{T}^2\times\mathbb{T}^2}g_0(\theta,\theta')e^{i\gamma(\theta-\theta')}(1-\chi_j'(\theta))\phi_j(\theta)\overline{(1-\chi_j'(\theta'))\phi_j(\theta')}d\theta d\theta'\le$$

$$\le C\langle h_0\phi,\phi\rangle_{L^2(\mathbb{T}^2)\otimes\mathbb{C}^{p_0}}. \quad (1.38)$$

Now, adding (1.38) to the left hand side of (1.34) and to the right hand side of (1.35), and using (1.36) and (1.37), we obtain that, there exists $C > 1$ such that, for $\phi \in L^2(\mathbb{T}^2) \otimes \mathbb{C}^{p_0}$, one has

$$\frac{1}{C}\langle h_\omega \phi, \phi\rangle_{L^2(\mathbb{T}^2)\otimes\mathbb{C}^{p_0}} \leq \langle \tilde{h}_\omega \phi, \phi\rangle_{L^2(\mathbb{T}^2)\otimes\mathbb{C}^{p_0}} \leq C\langle h_\omega \phi, \phi\rangle_{L^2(\mathbb{T}^2)\otimes\mathbb{C}^{p_0}} \tag{1.39}$$

where h_ω is defined by $h_\omega = h_0 + v_\omega$ and

$$v_\omega(\theta,\theta') = \begin{pmatrix} v_{\omega,1,1}(\theta,\theta') & \cdots & v_{\omega,1,p_0}(\theta,\theta') \\ \vdots & \ddots & \vdots \\ v_{\omega,p_0,1}(\theta,\theta') & \cdots & v_{\omega,p_0,p_0}(\theta,\theta') \end{pmatrix}$$

where

$$v_{\omega,i,j}(\theta,\theta') = \sum_{\gamma\in\Gamma} \omega_\gamma \left(\left[\int_{\mathbb{R}^2} V(x) v_i(x,\theta)\overline{v_j(x,\theta')}dx \right] e^{i\gamma(\theta-\theta')}\chi'_i(\theta)\chi'_j(\theta') + \right.$$
$$\left. + g_0(\theta,\theta')e^{i\gamma(\theta-\theta')}(1-\chi'_i(\theta))(1-\chi'_j(\theta'))\delta_{ij} \right).$$

We see that we can write

$$v_\omega(\theta,\theta') = \sum_{\gamma\in\Gamma} \omega_\gamma G(\theta,\theta') e^{i\gamma(\theta-\theta')} \tag{1.40}$$

where the matrix $G(\theta,\theta')$ has entries

$$g_{i,j}(\theta,\theta') = \int_{\mathbb{R}^2} V(x) v_i(x,\theta)\overline{v_j(x,\theta')}dx \chi'_i(\theta)\chi'_j(\theta') + g_0(\theta,\theta')(1-\chi'_i(\theta))(1-\chi'_j(\theta'))\delta_{ij}$$

We have

Lemma 1.7. *The kernel $G(\theta,\theta')$ is smooth. Moreover, for any $\theta \in \mathbb{T}^2$, $G(\theta,\theta)$ is a positive definite matrix.*

Proof. The regularity of $G(\theta,\theta')$ will follow from the regularity of

$$a_{i,j} : \ (\theta,\theta') \mapsto \int_{\mathbb{R}^2} V(x) v_i(x,\theta)\overline{v_j(x,\theta')}dx$$

in the set $\mathrm{supp}(\chi'_i)\times\mathrm{supp}(\chi'_j)$. The smoothness of $a_{i,j}$ comes from the smoothness of the mapping $\theta \in \mathbb{T}^2 \mapsto v_i(x,\theta) \in \mathcal{H}^2_\theta$ on the support of χ'_i and the fact that V has compact support.

Remark 1.4. We see here that we can relax the assumption on the compact support of V and still get the same result. We also notice that, because of the quasi-periodicity of the $(v_i(\cdot,\theta))_i$, the rate of decay of V determines the regularity of the kernel G. This regularity will then play a role in the determination of the Lifshitz exponent (see section 3). Moreover it is known that, if the potential decays sufficiently slowly, the Lifshitz tail regime is changed (see [12, 20]).

Let us now turn to the proof of the positive definiteness of G. Fix $\theta \in \mathbb{T}^2$. Pick $a = (a_j)_{1\leq j\leq p_0}$ a vector in $\mathbb{C}^{p_0}$. One has

$$\langle G(\theta,\theta)a,a\rangle = \int_{\mathbb{R}^2} V(x) \left| \sum_{i=1}^{p_0} a_i v_i(x,\theta)\chi'_i(\theta) \right|^2 dx + g_0(0,0)\sum_{i=1}^{p_0}(1-\chi_i(\theta))^2|a_i|^2 \geq 0. \tag{1.41}$$

So the matrix $G(\theta,\theta)$ is non negative as V is non negative.
Assume $\langle G(\theta,\theta)a,a\rangle = 0$. By (1.41), if $\chi'_j(\theta) \neq 1$, then $a_j = 0$. So that $\langle G(\theta,\theta)a,a\rangle = 0$ becomes

$$\int_{\mathbb{R}^2} V(x) \left| \sum_{i=1}^{p_0} a_i v_i(x,\theta) \right|^2 dx = 0.$$

As V is continuous and not vanishing identically, there exists O, an open subset of $\mathbb{T}^2$ on which V does not vanish; this implies that $\sum_{i=1}^{p_0} a_i v_i(x,\theta) = 0$ identically on O.
The system $(v_i(\cdot,\theta))_i$ is an orthonormal basis of $\Pi_\theta^{\mathcal{E}_0}\mathcal{H}_\theta$; let $Q(\theta)$ be the unitary matrix transforming this basis into the the basis $(u_i(\cdot,\theta))_i$ of Floquet eigenvectors. Then, we get that,

$$\sum_{i=1}^{p_0}(Q(\theta)a)_i u_i(x,\theta) = 0,\ x \in O \tag{1.42}$$

The function $(u_i)_i$ are eigenfunctions of the periodic Schrödinger operator H_0. Applying H_0 to (1.42) repeatedly, we get, for $n \in \mathbb{N}$,

$$\sum_{i=1}^{p_0}(Q(\theta)a)_i (E_i(\theta))^n u_i(x,\theta) = 0 \quad \text{ if } u_i(\cdot,\theta) \text{ is associated to the eigenvalue } E_i(\theta).$$

This in turn implies that, for $E \in \mathbb{R}$, one has

$$\sum_{i;\ E_i=E}(Q(\theta)a)_i u_i(x,\theta) = 0,\ x \in O.$$

Now, the Unique Continuation Principle (see e.g. [9, 47]) implies that $x \mapsto \sum_{i;\ E_i=E}(Q(\theta)a)_i u_i(x,\theta)$ vanishes identically as it is an eigenfunction of H_0. The functions $(u_i)_i$ are linearly independent. Hence, $Q(\theta)a = 0$. As $Q(\theta)$ is unitary, this implies $a = 0$ and completes the proof of Lemma 1.7. □

1.5. **The proof of Lemma 1.6.** Recall that $\mathcal{S}$ is the set of θ in $\mathbb{T}^2$ where at least one of the Floquet eigenvalues of H vanishes. Recall that $\mathcal{S}$ is compact. For $\theta \in \mathcal{S}$, let $1 \le p(\theta) \le p_0$ be the dimension of the kernel of H_θ i.e. $E_{n_0}(\theta) = \cdots = E_{n_0+p(\theta)-1}(\theta) = 0$ and $0 < E_{n_0+p(\theta)}(\theta) \le \cdots \le E_{n_0+p_0-1}(\theta)$. We write

$$\mathcal{S} = \bigcup_{p=1}^{p_0} \mathcal{S}_p \text{ where } \mathcal{S}_p = \{\theta \in \mathbb{T}^2;\ p(\theta) = p\}$$

We prove

Lemma 1.8. *The set $\mathcal{S}$ admits the following finite decomposition*

$$\mathcal{S} = \bigcup_{p=1}^{p_0}\bigcup_{j=1}^{j_p} \mathcal{A}_p^j \tag{1.43}$$

where the union is disjoint and, for $1 \le p \le p_0$, one has

1. *For $1 \le j \le j_p$, $\mathcal{A}_p^j \subset \mathcal{S}_p$.*
2. *The set $\mathcal{A}_p^j$ either is reduced to a single point or is a connected real-analytic curve (i.e. a connected real analytic manifold of dimension 1).*
3. *If $(j,p) \neq (j',p')$ and $\mathcal{A}_p^j \cap \overline{\mathcal{A}_{p'}^{j'}} \neq \emptyset$, then*
   - *$p' \le p$,*
   - *$\mathcal{A}_p^j \subset \overline{\mathcal{A}_{p'}^{j'}}$,*
   - *$\mathcal{A}_p^j$ is reduced to a single point,*
   - *$\mathcal{A}_{p'}^{j'}$ is a real-analytic curve.*
4. *Assume $\mathcal{A}_p^j = \{\theta_0\}$. Then, either θ_0 is isolated in $\mathcal{S}$ or, for some $\varepsilon_0 > 0$ sufficiently small, one has $\mathcal{S} \cap \dot{\overline{D}}(\theta_0,\varepsilon_0) = \bigcup_{(p',j')\in E} \mathcal{A}_{p'}^{j'} \cap \dot{\overline{D}}(\theta_0,\varepsilon_0)$ where E is a non empty, finite set of indices such that, for $(p',j') \in E$,*

- *$p' \leq p$;*
- *the set $\mathcal{A}_{p'}^{j'}$ is a real analytic curve.*

Here $\dot{\overline{D}}(\theta,\varepsilon) = \overline{D}(\theta,\varepsilon) \setminus \{\theta\}$ and $\overline{D}(\theta_0,\varepsilon) = \{|\theta-\theta_0| \leq \varepsilon\}$.

Proof. The set $\mathcal{S}$ is a real analytic subset of $\mathbb{T}^2$. Indeed, for $\theta_0 \in \mathcal{S}$, pick $\mathcal{C}_0$, a small circle around 0 in the complex plane of energies that does not enclose any eigenvalue of H_{θ_0} different from 0 and consider the Riesz projector

$$\Pi_{\theta,\theta_0} = \int_{\mathcal{C}_0} (z - H_\theta)^{-1} dz. \tag{1.44}$$

For θ close to θ_0, this projector is analytic and of finite, constant rank; moreover, θ belongs to $\mathcal{S}$ if and only if $\Pi_{\theta,\theta_0} H_\theta \Pi_{\theta,\theta_0}$ has a non empty kernel i.e. $\det(\Pi_{\theta,\theta_0} H_\theta \Pi_{\theta,\theta_0}) = 0$. This proves that $\mathcal{S}$ is a real analytic subset of $\mathbb{T}^2$ (see [28]). This immediately implies that $\mathcal{S}$ is a semi-analytic subset of $\mathbb{T}^2$ (see [2]).

Consider the set $\mathcal{S}_p^+ = \cup_{q \geq p} \mathcal{S}_q$, the set where the multiplicity of the eigenvalue 0 is at least p. $\mathcal{S}_p^+$ is a real analytic subset of $\mathbb{T}^2$. Indeed, pick $\theta_0 \in \mathcal{S}_p^+$ and let q be the rank of Π_{θ_0,θ_0}. Then, $q \geq p$. Moreover, there exists $\mathcal{V}$, a neighborhood of θ_0 such that, $\theta \in \mathcal{V} \cap \mathcal{S}_p^+$ if and only the first p eigenvalues of $\Pi_{\theta,\theta_0} H_\theta \Pi_{\theta,\theta_0}$ are equal to 0. This is equivalent to saying that the elementary symmetric functions $(s_i(\theta))_{q-p \leq j \leq q}$ vanish. The functions $(s_i(\theta))_{1 \leq j \leq q}$, being the coefficients of the characteristic polynomial of $\Pi_{\theta,\theta_0} H_\theta \Pi_{\theta,\theta_0}$, are real analytic in θ. Hence, the sets $\mathcal{S}_p^+$ are real analytic subsets of $\mathbb{T}^2$.

Corollary 2.11 in [2] then states the existence of a locally finite (hence finite) semi-analytic stratification of $\mathcal{S}$ compatible with the sets $(\mathcal{S}_p^+)_p$. That is:

- $\mathcal{S}$ is the disjoint union of the $(\tilde{A}_k)_{1 \leq k \leq K}$ such that
  - Each $\tilde{A}_k$ is a connected semi-analytic subset and an analytic sub-manifold of $\mathbb{T}^2$.
  - If $k \neq j$ and $\tilde{A}_k \cap \overline{\tilde{A}_j}$ then $\tilde{A}_k \subset \overline{\tilde{A}_j}$ and $\dim \tilde{A}_k < \dim \tilde{A}_j$.
- Each $\mathcal{S}_p^+$ is a finite union of certain $\tilde{A}_k$.

As $\mathcal{S}$ has a non empty interior (otherwise 0 would be an eigenvalue of the periodic operator H), we know that the sets $\tilde{A}_k$ have dimension less than 1. Let us now proceed by induction to reindex the sets $\tilde{A}_k$ to obtain Lemma 1.8.
Pick $p = p_0$. Then $\mathcal{S}_{p_0} = \mathcal{S}_{p_0}^+$. Hence, $\mathcal{S}_{p_0}$ is the disjoint union of certain sets $\tilde{A}_k$; renumber these sets to obtain the sets $(A_{p_0}^j)_{1 \leq j \leq j_{p_0}}$.
Pick $p < p_0$ and assume that one has constructed the sets $(A_l^j)_{p+1 \leq l \leq p_0; 1 \leq j \leq j_k}$ that are two by two disjoint and chosen so that, for $p+1 \leq l \leq p_0$, one has $\mathcal{S}_l = \bigcup_{j=1}^{j_l} \mathcal{A}_j^l$. Then, as $\mathcal{S}_p^+ = \cup_{p+1 \leq l \leq p_0} \mathcal{S}_{p_0} \cup \mathcal{S}_p$ and the union is disjoint, if $\tilde{A}_k \subset \mathcal{S}_p^+$ and $\tilde{A}_k \notin \{A_l^j;\ p+1 \leq l \leq p_0;\ 1 \leq j \leq j_k\}$, then $\tilde{A}_k \subset \mathcal{S}_p$. As $\mathcal{S}_p^+$ is the union of certain $\tilde{A}_k$, we get that $\mathcal{S}_p$ also is the union of certain $\tilde{A}_k$ taken outside of $\{A_l^j;\ p+1 \leq l \leq p_0;\ 1 \leq j \leq j_k\}$. we renumber these sets to obtain the sets $(A_p^j)_{1 \leq j \leq j_p}$.
Points (1)-(4) of Lemma 1.8 are immediate consequences of the definition of a semi-analytic stratification (see [2]). This completes the proof of Lemma 1.8. □

Remark 1.5. As $\mathcal{S}$ is closed, point (3) implies that if $\mathcal{A}_p^j$ is an analytic curve then, $\overline{\mathcal{S}_p^j}$, its closure, admits the decomposition $\overline{\mathcal{S}_p^j} = \mathcal{A}_p^j \cup \mathcal{A}_{p_1}^{j_1} \cup \mathcal{A}_{p_2}^{j_2}$ where $\mathcal{A}_{p_{1,2}}^{j_{1,2}}$ are reduced to a single point or empty.

We will now prove Lemma 1.6. This will be done in three steps; first, we will state Lemma 1.9 that shows the existence of the functions (v_j) locally near a point of $\mathcal{S}$. Then, in Lemma 1.10, we show how to extend this construction along an analytic curve contained in $\mathcal{S}$

where the multiplicity of the eigenvalue 0 is constant. At last, in Lemma 1.11, we deal with the points of $\mathcal{S}$ where the multiplicity of the eigenvalue 0 changes.

We prove

Lemma 1.9. *Fix $\theta_0 \in \mathcal{S}$ and recall that $p(\theta_0)$ is the dimension of the kernel of H_{θ_0}. There exists V_{θ_0}, an open neighborhood of θ_0 (in $\mathbb{T}^2$) and $\varepsilon_0 > 0$ such that*

1. *For $\theta \in V_{\theta_0}$, the number of Floquet eigenvalues $(E_n)_{n_0 \leq n \leq n_0+p_0-1}$ such that $0 \leq E_n(\theta) \leq \varepsilon_0$ is constant and equal to $p(\theta_0)$ i.e. the spectral space associated to H_θ and the energy interval $[0, \epsilon_0]$ has constant dimension $p(\theta_0)$.*
2. *There exists $p(\theta_0)$ real analytic functions $\theta \in V_{\theta_0} \mapsto v_j(\cdot, \theta) \in \mathcal{H}_\theta$ forming an orthonormal system, such that, for $\theta \in V_{\theta_0}$, one has*

$$span((v_j(\cdot,\theta))_{1 \leq n \leq p(\theta_0)}) = span((u_j(\cdot,\theta))_{n_0 \leq n \leq n_0+p(\theta_0)-1}) \text{ in } \mathcal{H}_\theta.$$

Proof. Pick $\theta_0 \in \mathcal{S}$. Then, $E_{n_0}(\theta_0) = \cdots = E_{n_0+p(\theta_0)-1}(\theta_0) = 0$ and $0 < E_{n_0+p(\theta_0)}(\theta_0) \leq \cdots \leq E_{n_0+p_0-1}(\theta_0)$. Point (1) of Lemma 1.9 is then an immediate consequence of the continuity of the Floquet eigenvalues.

For θ close to θ_0, consider the Riesz projector defined in (1.44) where $\mathcal{C}_0$ is a closed path around 0 that does not enclose any other eigenvalue of H_{θ_0}.
It is well known (see e.g. [39]) that, for $\theta \in V_{\theta_0}$, a sufficiently small neighborhood of θ_0, this integral defines the orthonormal projector onto the vector space spanned by $(u_j(\cdot,\theta))_{n_0 \leq j \leq n_0+p(\theta_0)-1}$ (recall that $(u_j(\cdot,\theta))_{n_0 \leq j \leq n_0+p_0-1}$ denote the eigenvectors associated to the eigenvalues $E_{n_0}(\theta) \leq \cdots \leq E_{n_0+p_0-1}(\theta)$); moreover, this projector is analytic in θ. For $1 \leq j \leq p(\theta_0)$ and $\theta \in V_{\theta_0}$, set $\tilde{v}_j(\cdot,\theta) = \Pi_{\theta,\theta_0}(T_{\theta_0 \to \theta} u_{j+n_0}(\cdot,\theta_0))$. Using the Gram-Schmidt algorithm, we orthonormalize the system $(\tilde{v}_j(\cdot,\theta))_{1 \leq j \leq p(\theta_0)}$ into the system $(v_j(\cdot,\theta))_{1 \leq j \leq p(\theta_0)}$. The system $(u_j(\cdot,\theta_0))_{n_0 \leq j \leq n_0+p(\theta_0)-1}$ is orthonormal; hence, reducing V_{θ_0} somewhat if necessary, we see that, for $1 \leq j \leq p(\theta_0)$, the function $\theta \mapsto v_j(\cdot,\theta)$ is real analytic in V_{θ_0}. □

Let $\mathcal{P}$ be the 0-dimensional strata of $\mathcal{S}$ i.e. $\mathcal{P}$ is the union of the sets $\mathcal{A}_p^j$ consisting of a single point. Then, we prove

Lemma 1.10. *Let $\mathcal{A}_p^j$ be one of the analytic curves contained in the decomposition of $\mathcal{S}$ given in Lemma 1.8. There exists $\varepsilon_0 > 0$ such that, for $0 < \varepsilon < \varepsilon_0$, then*

- *$\mathcal{B}_p^j = \mathcal{A}_p^j \setminus (\mathcal{P} + \overline{D(0,\varepsilon)})$ is connected.*
- *there exists $\mathcal{V}$, a neighborhood of $\mathcal{B}_p^j$ and p real analytic functions $\theta \in \mathcal{V} \mapsto v_j(\cdot,\theta) \in \mathcal{H}_\theta$ forming an orthonormal system, such that, for $\theta \in \mathcal{V}$, one has*

$$span((v_j(\cdot,\theta))_{1 \leq n \leq p}) = span((u_j(\cdot,\theta))_{n_0 \leq n \leq n_0+p-1}) \text{ in } \mathcal{H}_\theta.$$

Proof. The first point in Lemma 1.10 is an immediate consequence of Lemma 6.3 in [2].
Fix ε positive, sufficiently small. As the dimension of the kernel of H_θ, $\ker(H_\theta)$, is constant along $\mathcal{B}_p^j$ is constant, $(\ker(H_\theta))_\theta$ is a real analytic vector bundle over $\mathcal{B}_p^j$. Let $\mathcal{F}$ denote this vector bundle. The curve $\mathcal{B}_p^j$ is diffeomorphic to either the unit circle or a segment. The structure group of the vector bundle $\mathcal{F}$ is $\mathrm{Gl}(n,\mathbb{C})$; as it is connected, the vector bundle $\mathcal{F}$ is topologically trivial (see [8]) i.e. there exists $\theta \in \mathcal{B}_p^j \mapsto v_j(\cdot,\theta) \in \mathcal{H}_\theta$, p continuous sections of $\mathcal{F}$ forming an orthonormal system for every $\theta \in \mathcal{B}_p^j$.

The vector bundle $\mathcal{F}$ can also be defined by the projection on the spectral space associated to the spectrum of H_θ in a small complex disk center at 0. Hence, by standard analytic perturbation theory, $\mathcal{F}$ can be extended to a small complex tubular neighborhood of $\mathcal{B}_p^j$ as an analytic vector bundle $\mathcal{F}_e$. Then, $\mathcal{F}_e$ is also topologically trivial (see [8]).

To prove Lemma 1.10, it is sufficient to prove that $\mathcal{F}_e$ is analytically trivial. That this is indeed the case follows from Grauert's results (see [5, 11]) and from the fact that, both, a complex neighborhood of $\mathbb{T}^2$ and a complex neighborhood of $\mathcal{B}_p^j$ are Stein manifolds. Hence, we get that there exists $\mathcal{V}$, a neighborhood of $\mathcal{B}_p^j$ and p real analytic sections of $\mathcal{F}_e$,$\theta \in \mathcal{V} \mapsto v_j(\cdot,\theta) \in \mathcal{H}_\theta$, forming an orthonormal system. By the definition of $\mathcal{F}_e$, we have completed the proof Lemma 1.10. $\square$

At last, we prove

Lemma 1.11. *Pick $\theta_0 \in \mathcal{P}$ and let $q_0 = p(\theta_0)$. Then, there exists $\varepsilon_0 > 0$ such that for $0 < \varepsilon < \varepsilon_0$, there exists U, an open neighborhood of $\mathcal{S}$ in $D(\theta_0, \varepsilon)$, and $\theta \in U \mapsto (v_j(\cdot,\theta) \in \mathcal{H}_\theta)_{1\leq j\leq q_0}$ a smooth orthonormal system of functions such that*

- *in U, one has*
$$span((v_j(\cdot,\theta))_{1\leq n\leq q_0}) = span((u_j(\cdot,\theta))_{n_0\leq n\leq n_0+q_0-1}) \text{ in } \mathcal{H}_\theta,$$
- *if $\mathcal{S}\cap(D(\theta_0,\varepsilon)\setminus D(\theta_0,\varepsilon/2)) = \bigcup_{1\leq k\leq k_0} \mathcal{A}_{p_k}^{j_k}\cap(D(\theta_0,\varepsilon)\setminus D(\theta_0,\varepsilon/2))$, then, for every $1 \leq k \leq k_0$, there exists $V_k \subset U$, a neighborhood of $\mathcal{A}_{p_k}^{j_k} \cap (D(\theta_0,\varepsilon) \setminus D(\theta_0,\varepsilon/2))$ such that, for $\theta \in V_k$, the orthonormal system $(v_j(\cdot,\theta) \in \mathcal{H}_\theta)_{1\leq j\leq p_k}$ satisfies*
$$span((v_j(\cdot,\theta))_{1\leq n\leq p_k}) = span((u_j(\cdot,\theta))_{n_0\leq n\leq n_0+p_k-1}) \text{ in } \mathcal{H}_\theta,$$

Proof. If the multiplicity of the eigenvalue 0 does not change in a neighborhood of θ_0 or if θ_0 is an isolated point in $\mathcal{S}$, there is nothing to be done as Lemma 1.11 is an immediate consequence of Lemma 1.9.

Let θ_0 be such that, in a neighborhood of θ_0, the multiplicity of the eigenvalue 0 changes. By Lemma 1.8, there exists $\varepsilon_0 > 0$ and finitely many disjoint open analytic curves $(\mathcal{A}_{p_k}^{j_k})_{1\leq k\leq k_0}$ such that :

- $\mathcal{S} \cap \dot{D}(\theta_0,\varepsilon_0) = \bigcup_{1\leq k\leq k_0} \mathcal{A}_{p_k}^{j_k}$,
- for $1 \leq k \leq k_0$, $\overline{\mathcal{A}_{p_k}^{j_k}} \cap D(\theta_0,\varepsilon_0) = \mathcal{A}_{p_k}^{j_k} \cap D(\theta_0,\varepsilon_0) \cup \{\theta_0\}$.

Reducing ε_0 somewhat, Lemma 1.9 gives us q_0 real analytic functions $\theta \in D(\theta_0,\varepsilon_0) \mapsto \tilde{w}_j(\cdot,\theta) \in \mathcal{H}_\theta$ forming an orthonormal system, such that, for $\theta \in D(\theta_0,\varepsilon_0)$, one has
$$\text{span}((\tilde{w}_j(\cdot,\theta))_{1\leq n\leq q_0}) = \text{span}((u_j(\cdot,\theta))_{n_0\leq n\leq n_0+q_0-1}) \text{ in } \mathcal{H}_\theta.$$

Let $\tilde{U} = D(0,\varepsilon_0/2)$. We can also find $\delta_1 > \delta_0 > 0$ $\varepsilon_0 > 0$ such that,

1. for $\theta \in D(\theta_0,\varepsilon_0)$, one has $E_{n_0}(\theta) \leq \cdots \leq E_{n_0+q_0}(\theta) \leq \delta_1$;
2. for $1 \leq k \leq k_0$, there exists V_k, a neighborhood of $\mathcal{A}_{p_k}^{j_k}$ in $D(\theta_0,\varepsilon_0) \setminus \overline{D(\theta_0,\varepsilon_0/3)}$ such that, for $\theta \in V_k$, one has $\delta_0 \leq E_{n_0+p_k+1}(\theta) \leq \cdots \leq E_{n_0+q_0}(\theta)$
3. in each V_k, there exists p real analytic functions $\theta \in V_k \mapsto w_j(\cdot,\theta) \in \mathcal{H}_\theta$ forming an orthonormal system, such that, for $\theta \in V_k$, one has
$$\text{span}((w_j(\cdot,\theta))_{1\leq n\leq p_k}) = \text{span}((u_j(\cdot,\theta))_{n_0\leq n\leq n_0+p_k-1}) \text{ in } \mathcal{H}_\theta.$$
4. taking V_k thin enough, we get $\mathcal{S} \cap V_k = \gamma_k$ where γ_k is a real analytic curve and V_k is a tubular neighborhood of γ_k in $\mathbb{T}^2$.

Now, using the same arguments as in the proof of Lemma 1.9, we can find $q_0 - p_k$ real analytic functions $\theta \in V_k \mapsto w_j(\cdot,\theta) \in \mathcal{H}_\theta$ forming an orthonormal system, such that, for $\theta \in V_k$, one has
$$\text{span}((w_j(\cdot,\theta))_{p_k+1\leq n\leq q_0}) = \text{span}((u_j(\cdot,\theta))_{n_0+p_k\leq n\leq n_0+q_0-1}) \text{ in } \mathcal{H}_\theta.$$

This is done just by projection onto the spectral space associated to the eigenvalues in the interval $[\delta_0,\delta_1]$ (using (1) and (2)).

So, for $\theta \in V_k$, we have two orthonormal systems $(w_j(\cdot,\theta))_{1\leq j\leq q_0}$ and $(\tilde{w}_j(\cdot,\theta))_{1\leq j\leq q_0}$ generating the same subspace that are real analytic in θ. Hence, there exists a $q_0 \times q_0$ real analytic unitary matrix $M(\theta)$ defined in V_k, such that
$$(w_j(\cdot,\theta))_{1\leq j\leq q_0} = M(\theta)(\tilde{w}_j(\cdot,\theta))_{1\leq j\leq q_0}.$$

Let us now glue together the sets of vectors constructed above. Pick $\theta_* \in \gamma_k \subset V_k$. Fix $\varepsilon > 0$ small. As γ_k is real analytic and V_k is a tubular neighborhood of γ_k, we can parameterize V_k by the curvilinear abscissa on γ_k and another analytic coordinate. Assume that the curvilinear abscissa of θ_* is t_*. Pick χ a smooth function such that $\chi \equiv 0$ on $[\delta/2, +\infty)$, $\chi \equiv 1$ on $(-\infty, -\delta/2]$ and $0 \leq \chi \leq 1$ ($\delta > 0$ small). Construct the unitary matrix $\tilde{M}(\theta)$ in the following way. As $M_* = M(\gamma_k(t_*))$ is unitary, there exists L, a self-adjoint matrix such that $M^* = e^{iL}$. Define $T(\theta) = M(\theta)M_*^{-1}$. $T(\theta)$ is unitary and in a neighborhood of $\gamma_k(t_*)$, $T(\theta) = I + o(1)$. Reducing δ and V somewhat, one defines the unitary matrix

$$\tilde{M}(\theta) = [T(\theta)]^{\chi(\theta)} e^{i\chi(\theta)L}.$$

We note that $\tilde{M}(\theta) = Id$ if $\chi(\theta) = 0$, and $\tilde{M}(\theta) = M(\theta)$ if $\chi(\theta) = 1$.

For $\theta \in V_k$, set $(v_j(\cdot, \theta))_j = \tilde{M}(\theta)(\tilde{w}_j(\cdot, \theta))_j$; the matrix $\tilde{M}(\theta)$ being unitary, $(v_j(\cdot, \theta))_j$ is orthonormal and spans the same vector space as $(w_j(\cdot, \theta))_j$. Let $U = \tilde{U} \cup (\cup_k V_k)$. For $\theta \in U$, we define

- $v_j(\cdot, \theta) = w_j(\cdot, \theta)$ if $1 \leq j \leq p_k$ and $\theta \in V_k$,
- $v_j(\cdot, \theta) = \tilde{w}_j(\cdot, \theta)$ if $p_k + 1 \leq j \leq q_0$ and $\theta \in V_k$,
- $v_j(\cdot, \theta) = \tilde{w}_j(\cdot, \theta)$ if $1 \leq j \leq q_0$ and $\theta \in \tilde{U}$.

This definition is consistent as, for $\theta \in \tilde{U} \cap V_k$ and $1 \leq j \leq p_k$, one has $\tilde{w}_j(\cdot, \theta) = w_j(\cdot, \theta)$. This completes the proof of Lemma 1.11. □

Let us now complete the proof of Lemma 1.6. Therefore, we use the stratification of $\mathcal{S}$ given by Lemma 1.8. For every point θ_0 in $\mathcal{P}$, construct the neighborhood U and the vectors $(v_j)_j$ using Lemma 1.11. For every analytic curve $\mathcal{A}_p^j$, construct the neighborhood W_k and the vectors $(\tilde{v}_j)_j$ using Lemma 1.10 so that $U \cap W_k \neq \emptyset$ if $\mathcal{A}_p^j$ is a curve starting at θ_0. Reducing the neighborhood W_k somewhat, in $U \cap W_k$, we can smoothly glue together the vectors $(v_j)_j$ and $\tilde{v}_j)_j$ in the same way as in the proof of Lemma 1.11. Then, if we take the open sets (U_k^p) to be all the $(W_k)_k$ and U, and the functions $(v_j^{p,k})$ to bet the functions (v_j), one checks that points 1., 2. and 3. of Lemma 1.6 are satisfied.

In the case $p_0 = 1$, as the multiplicity of the lowest eigenvalue stays constant near $\mathcal{S}$, it suffices to use Lemma 1.10 to construct the vector announced in Lemma 1.6. This completes the proof of Lemma 1.6. □

We will end this section with a result that will be of use in section 3 namely

Lemma 1.12. *The set $\mathcal{S}$ does not contain any segment of strait line with irrational slope.*

Proof. The proof is done in two steps. First, $\mathcal{S}$ cannot contain a whole strait line with irrational slope. Indeed, if it did, this line being dense in $\mathbb{T}^2$, the eigenvalue $E_{n_0}(\theta)$, being continuous, would be constant on $\mathbb{T}^2$; hence, the periodic operator H would admit an eigenvalue : this is known not to be possible (see e.g. [37, 39]).

Let us now prove that, if $\mathcal{S}$ contains a segment of a strait line, it must contain the whole line. As $\mathcal{S}$ is closed, we may assume that it contains the segment $S_{\theta_0,\theta_1} = \{t\theta_0 + (1-t)\theta_1,\ t \in [0, 1]\}$. As $\mathcal{S}$ is a real analytic set, near θ_0, the segment S_{θ_0,θ_1} is a part of a real analytic curve $\mathcal{C}$ extending beyond θ_0 (see e.g. [28]). The unique continuation principle then implies that $\mathcal{C}$ is contained in the line passing through θ_0 and θ_1; hence, $\mathcal{S}$ contains the segment $\{t\theta_0 + (1-t)\theta_1,\ t \in [-\varepsilon, 1+\varepsilon]\}$ for some $\varepsilon > 0$. Considering the maximal segment of the line passing through θ_0 and θ_1 contained in $\mathcal{S}$, we see that this segment has to be the whole line. This completes the proof of Lemma 1.12. □

2. Approximating the density of states by a probability

In this section, we will show that, at energy E, the density of states of an operator of the form $h_\omega = h + G_\omega$ (as defined in Theorem 1.2) is well approximated by the probability to find an eigenvalue less than E for h_ω (restricted to some well chosen cube). Rather than considering h_ω, we will consider its conjugate under the discrete Fourier transform; we denote it by $\hat{h}_\omega$; it is acting on $\ell^2(\mathbb{Z}^2)\otimes\mathbb{C}^{p_0}$. As the kernel of h_ω is smooth, the coefficients of $\hat{h}_\omega$ (these are $p_0\times p_0$ matrices) satisfy that, for any $n\in\mathbb{N}$, there exists $C_n>0$ such that

$$\left\|\hat{h}_\omega(\gamma,\beta)\right\|\leq C_n(1+|\gamma-\beta|)^{-n},\quad \gamma,\beta\in\Gamma. \tag{2.1}$$

As functions of the random variable ω_α, they also satisfy that, for any $n\in\mathbb{N}$, there exists $C_n>0$ such that

$$\left\|\partial_{\omega_\alpha}\hat{h}_\omega(\gamma,\beta)\right\|\leq C_n(1+|\gamma-\alpha|)^{-n}(1+|\alpha-\beta|)^{-n}\quad \alpha,\gamma,\beta\in\Gamma. \tag{2.2}$$

Pick $(\delta_i)_{1\leq i\leq p_0}$ be the canonical basis of $\mathbb{C}^{p_0}$ and $(\delta_\gamma)_{\gamma\in\Gamma}$ be the canonical basis of $\ell^2(\mathbb{Z}^2)$. If $Z\subset\mathbb{Z}^2$, define Π_Z to be the projector on the subspace of $\ell^2(\mathbb{Z}^2)\otimes\mathbb{C}^{p_0}$ generated by $(\delta_\gamma\times\delta_i)_{(\gamma,i)\in Z\times\{1,\dots,p_0\}}$.
Let $L\in\mathbb{N}$ and $\Lambda_L\in\mathbb{Z}^2$ be the cube centered at 0 of side length $2L$. We set $\Pi_L=\Pi_{\Lambda_L}$. At last, we define

$$\hat{h}_{\omega,L}=\Pi_L\hat{h}_\omega\Pi_L.$$

Our aim in this section is to estimate $N^*(E)$, the integrated density of states of $\hat{h}_\omega$ in terms of certain probabilities. We will prove the

Theorem 2.1. *Fix $0<\varepsilon<1$ and $I\subset\mathbb{R}$ a compact interval.*

1. *There exists $C>0$ such that, for any $n\in\mathbb{N}$, there exists $C_n>0$ such that, for $L\geq 1$ and $E\in I$, one has*

$$N^*(E)\leq p_0\cdot\mathbb{P}(\{\hat{h}_{\omega,L}\text{ admits an eigenvalue }\leq E+C_nL^{-n})+CL^{-(1-\varepsilon)}. \tag{2.3}$$

2. *There exists $C>0$ such that, for $l\geq 1$ and $E\in I$, one has*

$$\frac{1}{Cl^2}\mathbb{P}(\{\hat{h}_{\omega,l}\text{ admits an eigenvalue }\leq E)\leq N^*(E). \tag{2.4}$$

Proof. Define

$$\vartheta^*_L(E)=\#\{\text{eigenvalues of }\hat{h}_{\omega,L}\leq E\}.$$

Then, we know that, for E a continuity point of N^*, ω-almost surely, one has

$$N^*(E)=\lim_{L\to+\infty}\mathbb{E}\left(\frac{\vartheta^*_L(E)}{\#\Lambda_L}\right)=\lim_{L\to+\infty}N^*_L(E).$$

We prove the

Proposition 2.1. *Pick $\varepsilon>0$ and $I\subset\mathbb{R}$ a compact interval. There exists $C>0$ such that, for any $n\in\mathbb{N}$, there exists $C_n>0$ such that, for $L\geq 1$ and $E\in I$, one has*

$$N^*_L(E-C_nL^{-n})-CL^{-(1-\varepsilon)}\leq N^*_{2L}(E)\leq N^*_L(E+C_nL^{-n})+CL^{-(1-\varepsilon)}. \tag{2.5}$$

This implies the

Corollary 2.1. *Pick $\varepsilon>0$ and $I\subset\mathbb{R}$ a compact interval. There exists $C>0$ such that, for any $n\in\mathbb{N}$, there exists $C_n>0$ such that, for $L\geq 1$ and $E\in I$, a continuity point of N^*, one has*

$$N^*_L(E-C_nL^{-n})-L^{-(1-\varepsilon)}\leq N^*(E)\leq N^*_L(E+C_nL^{-n})+L^{-(1-\varepsilon)}. \tag{2.6}$$

The corollary is proved inductively using Proposition 2.1.

We postpone the proof of Proposition 2.1 and use Corollary 2.1 to complete the proof of Theorem 2.1. Let us start with the upper bound (2.3). As we know that

$$\vartheta_L^*(E + C_n L^{-n}) \leq p_0 \cdot \#\Lambda_L,$$

we obtain that

$$N_L^*(E + C_n L^{-n}) \leq p_0 \cdot \mathbb{P}(\{\hat{h}_{\omega,L} \text{ admits an eigenvalue } \leq E + C_n L^{-n}).$$

Taking (2.6) into account, we complete the proof of (2.3).

We now turn to the proof of (2.4). Pick $L \geq 2l \geq 2$. Decompose $\Lambda_L = \cup_{i\in J}(\gamma_i + \Lambda_l) \cup \Lambda_r$ such that $(\gamma_i + \Lambda_{2l}) \cap (\gamma_j + \Lambda_{2l}) = \emptyset$ if $i \neq j$ and such that $(\gamma_i + \Lambda_{2l}) \cap \Lambda_r = \emptyset$ for all i. We can pick the $(\gamma_i)_{i\in J}$ such that $\#\Lambda_r \leq CL \cdot l$. The set of indices J can be chosen so that $C \cdot \#J \cdot l^2 \geq (L-l)^2$ for some $C > 0$ (independent of L and l). Define

$$\hat{h}_{\omega,l,\gamma_i} = \Pi_{\gamma_i+\Lambda_l} \hat{h}_\omega \Pi_{\gamma_i+\Lambda_l}$$

Then, as the $(\hat{h}_{\omega,i})_{i\in J}$ are acting on orthogonal subspaces, one has

$$\begin{aligned}\vartheta_L^*(E) &\geq \#\{i; \hat{h}_{\omega,l,\gamma_i} \text{ has an eigenvalue } \leq E\} \\ &= \sum_{i\in J} \mathbf{1}_{\{\hat{h}_{\omega,l,\gamma_i} \text{ has an eigenvalue } \leq E\}}.\end{aligned}$$

Hence, using the stationarity of the process $(\omega_\gamma)_\gamma$, we obtain

$$\begin{aligned}\mathbb{E}(\vartheta_L^*(E)) &\geq \sum_{i\in J} \mathbb{P}(\{\hat{h}_{\omega,l,\gamma_i} \text{ has an eigenvalue } \leq E\}) \\ &= \#J \cdot \mathbb{P}(\{\hat{h}_{\omega,l} \text{ has an eigenvalue } \leq E\})\end{aligned} \tag{2.7}$$

Taking into account the lower bound on $\#J$, dividing (2.7) by $\#\Lambda_L$ and letting $L \to +\infty$, we obtain (2.4). Thus, we complete the proof of Theorem 2.1. □

Proof of Proposition 2.1. Decompose $\Lambda_{2L} = \cup_{i=1,\dots,4}(\gamma_i + \Lambda_L)$ where $(\gamma_i + \Lambda_L) \cap (\gamma_j + \Lambda_L) = \emptyset$ if $i \neq j$. We write

$$\hat{h}_{\omega,2L} = \sum_{i=1}^{4} \Pi_{\gamma_i+\Lambda_L} \hat{h}_\omega \Pi_{\gamma_i+\Lambda_L} + \sum_{i\neq j} \Pi_{\gamma_i+\Lambda_L} \hat{h}_\omega \Pi_{\gamma_j+\Lambda_L}.$$

Pick $\varepsilon > 0$ and let $K_{i,\varepsilon} \subset \gamma_i + \Lambda_L$ be such that, for $j \neq i$, $\text{dist}((\gamma_i + \Lambda_L) \setminus K_{i,\varepsilon}, \gamma_j + \Lambda_L) \geq L^\varepsilon$. One can pick $K_{i,\varepsilon}$ of size $L^{1+\varepsilon}$. Then, for $i \neq j$, we write

$$\Pi_{\gamma_i+\Lambda_L} \hat{h}_\omega \Pi_{\gamma_j+\Lambda_L} = \Pi_{(\gamma_i+\Lambda_L)\setminus K_{i,\varepsilon}} \hat{h}_\omega \Pi_{\gamma_j+\Lambda_L} + \Pi_{K_{i,\varepsilon}} \hat{h}_\omega \Pi_{\gamma_j+\Lambda_L}$$

The estimate (2.1) and the definition of $K_{i,\varepsilon}$ tell us that, for any $n \in \mathbb{N}$, there exists $C_n > 0$ such that, one has

$$\|\Pi_{(\gamma_i+\Lambda_L)\setminus K_{i,\varepsilon}} \hat{h}_\omega \Pi_{\gamma_j+\Lambda_L}\| \leq C_n L^{-n}.$$

On the other hand, the rank of $\Pi_{K_{i,\varepsilon}} \hat{h}_\omega \Pi_{\gamma_j+\Lambda_L}$ is at most $L^{1+\varepsilon}$. Hence, as the operators $(\Pi_{\gamma_i+\Lambda_L} \hat{h}_\omega \Pi_{\gamma_i+\Lambda_L})_{i=1,2,3,4}$ are acting on two-by-two orthogonal spaces, we obtain that

$$\begin{aligned}\sum_{i=1}^{4} \#\{\text{eigenvalues of } \Pi_{\gamma_i+\Lambda_L} \hat{h}_\omega \Pi_{\gamma_i+\Lambda_L} \leq E - C_n L^{-n}\} - 4L^{1+\varepsilon} \leq \vartheta_{2L}^*(E) \\ \vartheta_{2L}^*(E) \leq \sum_{i=1}^{4} \#\{\text{eigenvalues of } \Pi_{\gamma_i+\Lambda_L} \hat{h}_\omega \Pi_{\gamma_i+\Lambda_L} \leq E + C_n L^{-n}\} + 4L^{1+\varepsilon}.\end{aligned} \tag{2.8}$$

Here we used the following fact: for A and B finite rank operators, one has

$$|\#\{\text{eigenvalues of } A \leq E\} - \#\{\text{eigenvalues of } B \leq E\}| \leq \operatorname{rank}(A - B).$$

As the random variables $(\omega_\gamma)_\gamma$ are stationary, we know that, for $1 \leq i \leq 4$,

$$\mathbb{E}\left(\#\{\text{eigenvalues of } \Pi_{\gamma_i+\Lambda_L}\hat{h}_\omega\Pi_{\gamma_i+\Lambda_L} \leq E\}\right) = \mathbb{E}(\vartheta_L^*(E)).$$

Dividing (2.8) by $\#\Lambda_{2L}$ and taking the expectation, we obtain (2.5). □

3. The probability estimates

We now formulate the probability estimates which were discussed briefly in the introduction. Let G be a smooth function on $\mathbb{T}^2 \times \mathbb{T}^2$ taking values in the $d \times d$ complex matrices, and assume that G is positive on the diagonal:

$$G(\theta, \theta) \text{ is positive definite for each } \theta \in \mathbb{T}^2.$$

If u and v are L^2 functions on $\mathbb{T}^2$ taking values in $\mathbb{C}^d$, then we define $B(u, v) : \mathbb{Z}^2 \to \mathbb{C}$ via

$$B(u, v)(n) = \int_{\mathbb{T}^2\times\mathbb{T}^2} \langle G(\theta, \theta')u(\theta), v(\theta')\rangle e^{in\cdot(\theta-\theta')} d\theta d\theta'.$$

Let $f : \mathbb{T}^2 \to \mathbb{R}$ be a nonnegative smooth function on $\mathbb{T}^2$, vanishing at some point, not identically zero, and satisfying the following condition:

f1. If a is any point such that $f(a) = 0$, then there is a real analytic function F defined in a neighborhood of a and such that $C^{-1}f(x) \leq F(x) \leq Cf(x)$.

It is useful in connection with the proofs of Theorems 0.1 and 0.2 to require property f1 instead of real analyticity of f. We make two further remarks about the condition f1.

First, if f satisfies f1, and if f vanishes on a segment of a (connected) real analytic curve γ then f must vanish identically on γ by a standard analytic continuation argument making use of the functions F. In particular, f cannot vanish on a segment of a line with irrational slope, since such a line is dense in the torus.

Second, the Newton diagram at a of a function F as in f1. is independent of the particular choice of F. Thus the definition of $\alpha(f)$ extends immediately to functions f satisfying f1.

Let $\{\omega_n\}$ be iid random variables satisfying the following conditions:

$$\text{The (essential) infimum of } \omega_n \text{ is } -1, \text{ and } \omega_n \text{ is not identically } -1. \tag{3.1}$$

$$\omega_n \text{ is bounded.} \tag{3.2}$$

$$\lim_{\epsilon\to 0} \frac{\log|\log \operatorname{Prob}(\omega_n < -1+\epsilon)|}{-\log\epsilon} = 0. \tag{3.3}$$

Let $\mathcal{P}(M)$ be the $\mathbb{C}^d$ valued functions on $\mathbb{T}^2$ whose Fourier transforms (these are functions on $\mathbb{Z}^2$) are supported in the square Q_M centered at the origin with side M, and define

$$H(\omega, u) = \int_{\mathbb{T}^2} f(\theta)|u(\theta)|^2 + \sum_n \frac{1}{2}(1+\omega_n)B(u, u)(n)$$

$$h(M, \omega) = \inf_{\substack{u\in\mathcal{P}(M)\\ \|u\|_2=1}} H(\omega, u)$$

$$P(M, \delta) = \operatorname{Prob}(h(M, \omega) < \delta)$$

Theorem 3.1. *Assume that f has property f1 and G is C^∞, and let $\alpha = \alpha(f)$. Assume that the $\{\omega_n\}$ satisfy (3.1), (3.2) and (3.3). If $\eta_0 > 0$, then the following is true. For each δ let $M_\delta \in [\delta^{-\alpha-\eta_0}, e^{\delta^{-\alpha+\eta_0}}]$. Define $P(\delta) = P(M_\delta, \delta)$. Then*

$$\lim_{\delta\to 0} \frac{\log|\log P(\delta)|}{-\log\delta} = \alpha(f).$$

We note that the formulation discussed in the introduction involving functions whose Fourier transforms are supported on a random set is the special case where $G = 1$ and the $\{\omega_n\}$ are Bernoulli. The proof for this case is basically no easier than the proof in the general case, but the general case is needed for the proof of Theorems 0.1 and 0.2.

For Theorem 0.1 we only need the upper bound on the probability in Theorem 3.1, and this is also valid under rather weaker hypotheses: (3.3) is not needed, and the C^∞ regularity of G can be relaxed. Thus we also give the following variant:

Theorem 3.2. *Assume that f has property f1 and that G satisfies*

$$\sum_{\substack{k\in\mathbb{Z}^2\\ \ell\in\mathbb{Z}^2}} (1+|k|+|\ell|)^\sigma |\hat{G}(k,\ell)| < \infty \tag{3.4}$$

for some $\sigma > 0$. Assume that the $\{\omega_n\}$ are iid's satisfying (3.1) and (3.2). If $\eta_0 > 0$, then the following is true. For each δ let $M_\delta \le e^{\delta^{-\alpha+\eta_0}}$. Define $P(\delta) = P(M_\delta, \delta)$. Then

$$\liminf_{\delta\to 0} \frac{\log|\log P(\delta)|}{-\log\delta} \ge \alpha(f).$$

This section and the next two sections are the proof of Theorem 3.1 (and 3.2). The proof will use various arguments from elementary probability, some estimates of Fourier transforms (Bernstein's inequality and the Stein-Tomas theorem), and the following Lemmas 3.1 and 3.2; these are properties of a real analytic function near its zero set, and are related to considerations in [35] and [45]. We fix f as in Theorem 3.1 and let $\alpha = \alpha(f)$. We always let

$$\Omega(\delta) = \{x : |f(x)| < \delta\} \tag{3.5}$$

Lemmas 3.1 and 3.2 are valid without the positivity assumption on f. The terminology *w-neighborhood of γ* means $\{x \in \mathbb{T}^2 :\ \mathrm{dist}(x,\gamma) < w\}$.

Lemma 3.1. *Assume that $f : \mathbb{T}^2 \to \mathbb{R}$ has property f1. Let A be a large fixed constant. Then, if κ is a small positive constant, $\Omega(A\delta)$ has a covering as follows:*

$$\Omega(A\delta) \subset \cup_{i\le C\log\frac{1}{\delta}} R_i$$

where each R_i is either type I, type II or type III, and the type I, II, and III R_i's have the following properties.

1. *Each type I R_i is a rectangle with sides $\tau_i \times w_i$ where we have set $w_i = C\frac{\delta^\alpha}{\tau_i}$, and where $\tau_i \in [\delta^{\frac{\alpha}{2}}, \delta^\kappa]$*
2. *Each type II R_i is the w_i-neighborhood of a real analytic curve segment γ_i with length τ_i, curvature comparable to k_i, and j-th derivative of curvature (with respect to arc length) bounded by $C_j k_i \tau_i^{-j}$ for each j, where we have set $w_i = C(\delta^\alpha k_i^{\frac{1}{2}})^{\frac{2}{3}}$, and where the parameters w_i, τ_i, k_i satisfy the following conditions: $w_i \gtrsim \delta^{\alpha-\kappa}$, $w_i \le \tau_i$ and $k_i \lesssim \tau_i^{-1}$.*
3. *There are only a fixed finite number of type III R_i's. Each type III R_i is the w_i-neighborhood of a line on which f vanishes identically, where we have set $w_i = C\delta^{\alpha-\kappa}$.*

We will refer to the R_i's in Lemma 3.1 as *components*. Lemma 3.1 will be used in the proof of Theorem 3.2. For the converse bound in Theorem 3.1 we need the following, which is easy and probably well known.

Lemma 3.2. *If $f : \mathbb{T}^2 \to \mathbb{R}$ has property f1 then Ω_δ contains a rectangle with area $\gtrsim \delta^\alpha$.*

In the rest of this section we reduce Theorem 3.1 to Lemmas 3.1 and 3.2 and a more explicit estimate of random Fourier series, Lemma 3.6. Section 4 is the proof of Lemma 3.6, and section 5 is the proof of Lemmas 3.1 and 3.2.

We first make a simple reduction: it suffices to prove Theorems 3.1 and 3.2 assuming the following additional condition on the random variables ω_n:

$$\omega_n \text{ has mean zero.} \tag{3.6}$$

Namely, suppose that the ω_n have a nonzero mean m. Necessarily $-1 < m < C$, where C is the constant implicit in (3.2), and we used (3.1). For any u we have

$$\begin{aligned} H(\omega, u) &= \int_{\mathbb{T}^2} f(\theta)|u(\theta)|^2 + \sum_n \frac{1}{2}(1 + m + \omega_n - m)B(u,u)(n) \\ &= (1+m)\left(\int_{\mathbb{T}^2} \frac{f(\theta)}{1+m}|u(\theta)|^2 + \sum_n \frac{1}{2}(1 + \frac{\omega_n - m}{1+m})B(u,u)(n)\right) \end{aligned}$$

This is $1+m$ times the Hamiltonian associated in the same way to the random variables $\frac{\omega_n - m}{1+m}$, which do have mean zero, and the function $\frac{f(\theta)}{1+m}$. It follows that without loss of generality we can assume (3.6), and we will do so throughout the proof.

Let $\delta > 0$ be the parameter in Theorem 3.1. Let $\epsilon > 0$ be such that

$$\epsilon < \delta^\tau \tag{3.7}$$

where τ is a small fixed constant.

Let B be an integer with $B \approx \epsilon^{-1}$, and let Λ be the sublattice $B\mathbb{Z}^2$. Let $Q(0)$ be the square centered at the origin with side B and for each $\lambda \in \Lambda$ let $Q(\lambda) = Q(0) + \lambda$. Let $\mathcal{Q}$ be the set of squares $\{Q(\lambda) : \lambda \in \Lambda\}$. Let R be large, but small compared with B (in the sequel R will be set equal to $\log \frac{1}{\delta}$) If $\lambda, \mu \in \Lambda$ then we define

$$U_{\lambda\mu} = \{n \in \mathbb{Z}^2 : |n - \lambda| \le RB \text{ and } |n - \mu| \le RB\} \tag{3.8}$$

We also fix a function $\phi = \phi_\epsilon$ with the following properties.

$$\text{supp}\phi \subset D(0, \epsilon) \tag{3.9}$$

$$\sum_{\lambda \in \Lambda} \hat{\phi}(k - \lambda) = 1 \tag{3.10}$$

and ϕ is even, nonnegative and is smooth with bounds

$$|D^\alpha \phi| \lesssim \epsilon^{-2-|\alpha|} \tag{3.11}$$

This is possible since by the Poisson summation formula (3.10) is equivalent with $\phi(0) = B^2$ and $\phi(2\pi\frac{i}{B}, 2\pi\frac{j}{B}) = 0$ if i and j are integers, $(i,j) \ne (0,0)$ mod Λ. so in fact any bump function ϕ with $\phi(0) = B^2$ and with support in $D(0, C^{-1}\epsilon)$ and the natural derivative bounds will work. We may further assume that

$$\phi = \psi * \psi \tag{3.12}$$

where ψ is even and nonnegative and satisfies (3.9) and (3.11), since we can start from a suitable ψ and define $\phi = \psi * \psi$. It follows using (3.9) and (3.11) that

$$|\hat{\phi}(n)| \lesssim (1 + \epsilon|n|)^{-1000} \tag{3.13}$$

and similarly with ψ. In particular, (3.13) implies that

$$\sum_{\lambda\in\Lambda} |\hat{\phi}(k-\lambda)| \leq C \tag{3.14}$$

uniformly in $k \in \mathbb{Z}^2$, and that

$$\sum_{\substack{\lambda\in\Lambda \\ |k-\lambda|\geq RB}} |\hat{\phi}(k-\lambda)| \lesssim R^{-998} \tag{3.15}$$

uniformly in k.
If $\lambda \in \Lambda$ we set $\phi_\lambda(x) = e^{2\pi i\lambda\cdot x}\phi(x)$ and $\psi_\lambda(x) = e^{2\pi i\lambda\cdot x}\psi(x)$; thus

$$\widehat{\phi_\lambda}(k) = \phi(k-\lambda).$$

Lemma 3.3. *Let f be as in Theorem 3.1 and assume that G satisfies (3.4). Assume that $\{\omega_n\}$ is a bounded numerical sequence, $\|\omega\|_{\ell^\infty} \leq C_1$. Let A be a large enough constant (depending on C_1) and assume that δ is sufficiently small.*
Suppose there are functions $u_1, u_2 \in L^2(\mathbb{T}^2, \mathbb{C}^d)$ such that

$$\int_{\mathbb{T}^2} f(\theta)|u_i(\theta)|^2 d\theta < \delta\|u_i(\theta)\|_2^2 \tag{3.16}$$

$$\left|\sum_{n\in Q_M} \omega_n B(u_1,u_2)(n)\right| \geq C_1^{-1}\|u_1\|_2\|u_2\|_2 \tag{3.17}$$

Then, there are functions $v_1, v_2 \in L^2(\mathbb{T}^2, \mathbb{C}^d)$ such that, for $i = 1$ and 2, v_i is supported in a single component R_i and

$$\left|\sum_{n\in Q_M} \omega_n B(v_1,v_2)(n)\right| \geq C^{-1}(\log\frac{1}{\delta})^{-1}\|v_1\|_2\|v_2\|_2 \tag{3.18}$$

Proof. Note to begin with that

$$\sum_{n\in\mathbb{Z}^2} |B(u,v)(n)| \lesssim \|u\|_2\|v\|_2 \tag{3.19}$$

It suffices to prove (3.19) in the $d = 1$ case; otherwise we consider an appropriate matrix coefficient and sum over the matrix coefficients at the end. If $n = 1$ and if $G(\theta,\theta') = a(\theta)b(\theta')$ with $a, b \in L^\infty(\mathbb{T}^2)$ then (3.19) follows immediately from the Plancherel theorem. The general case therefore follows by expanding G in a Fourier series and using the hypothesis (3.4).
It suffices to prove the lemma when $\|u_1\|_2 = \|u_2\|_2 = 1$. With u_1 and u_2 so normalized, let $w_i = \chi_{\Omega(A\delta)}u_i$. Then, we have

$$\|u_i - w_i\|_2^2 \leq A^{-1} \tag{3.20}$$

by (3.16). Accordingly, from (3.19) and (3.20),

$$\sum_{n\in Q_M} |B(w_1,w_2)(n) - B(u_1,u_2)(n)| \lesssim A^{-\frac{1}{2}}.$$

It follows that

$$\left|\sum_{n\in Q_M} \omega_n B(w_1,w_2)(n)\right| \geq \frac{1}{2C_1}\|w_1\|_2\|w_2\|_2$$

provided if A is large enough. Now decompose $w_i = \sum_\alpha v_i^\alpha$ where each v_i^α is supported on a single component. The v_i^α may be taken to have disjoint support and so that there are at most $\mathcal{O}(\log \frac{1}{\delta})$ summands α. Accordingly

$$\begin{aligned}
\sum_{\alpha,\beta} \left| \sum_{n\in Q_M} \omega_n B(v_1^\alpha, v_2^\beta)(n) \right| &\geq |\sum_{n\in Q_M} \omega_n B(w_1,w_2)(n)| \\
&\geq \frac{1}{2C_1}\|w_1\|_2\|w_2\|_2 \\
&= \frac{1}{2C_1}\left(\sum_{\alpha,\beta}\|v_1^\alpha\|_2^2\|v_2^\beta\|_2^2\right)^{\frac{1}{2}} \\
&\gtrsim (\log\frac{1}{\delta})^{-1}\sum_{\alpha,\beta}\|v_1^\alpha\|_2\|v_2^\beta\|_2.
\end{aligned}$$

We conclude that if we set $v_1 = v_1^\alpha$ and $v_2 = v_2^\beta$ for appropriate α and β then (3.18) will hold. □

Lemma 3.4. *Let ϵ satisfy (3.7) and let $R = \log\frac{1}{\delta}$. Let v_1 and v_2 satisfy (3.18). Then, for some $\lambda, \mu \in \Lambda$*

$$\left| \sum_{n\in Q_M\cap U_{\lambda_\mu}} \omega_n B(\phi_\lambda * v_1, \phi_\mu * v_2)(n) \right| \geq C^{-1}(\log\frac{1}{\delta})^{-3}\|\psi_\lambda * v_1\|_2\|\psi_\mu * v_2\|_2 \tag{3.21}$$

where C does not depend on the specific choice of ϵ.

Proof. It suffices to prove the lemma when $\|v_1\|_2 = \|v_2\|_2 = 1$.
We may express

$$v_i = \sum_{\lambda\in\Lambda}\phi_\lambda * v_i$$

by (3.10). Therefore, if $n \in Q_M$,

$$\begin{aligned}
B(v_1,v_2)(n) &= \sum_{\lambda,\mu\in\Lambda} B(\phi_\lambda * v_1, \phi_\mu * v_2)(n) \\
&= \sum_{\lambda,\mu\in\Lambda}\sum_{n\in U_{\lambda_\mu}\cap Q_M} B(\phi_\lambda * v_1, \phi_\mu * v_2)(n) + E(n) \qquad (3.22)
\end{aligned}$$

where (with the sums being over λ and μ)

$$|E(n)| \leq \sum_{|\lambda-n|\geq RB}|B(\phi_\lambda * v_1, \phi_\mu * v_2)(n)| + \sum_{|\mu-n|\geq RB}|B(\phi_\lambda * v_1, \phi_\mu * v_2)(n)|. \tag{3.23}$$

We now show that

$$\sum_n |E(n)| \lesssim R^{-998} \tag{3.24}$$

provided that δ is small.

We may assume as above that $d = 1$. Denote the two sums on the right side in (3.23) by $\sum_n I(n)$ and $\sum_n II(n)$ respectively.

Letting $a_{k\ell}$ ($k \in \mathbb{Z}^2, \ell \in \mathbb{Z}^2$) be the Fourier coefficients of G we have

$$B(u,v)(n) = \sum_{k\ell} a_{k,\ell}\hat{u}(n-k)\overline{\hat{v}}(n-\ell) \tag{3.25}$$

Consider $\sum_n I(n)$. Interchanging the order of summation, we can rewrite this as

$$\sum_{\lambda,\mu} \sum_{n:|n-\lambda|\geq RB} |B(\phi_\lambda * v_1, \phi_\mu * v_2)(n)|$$

which can be estimated by

$$\sum_{k,\ell}\sum_{\lambda,\mu}\sum_{n:|n-\lambda|\geq RB} |a_{k\ell}|\,|\hat{\phi}(n-k-\lambda)\widehat{v_1}(n-k)\hat{\phi}(n-\ell-\mu)\widehat{v_2}(n-\ell)| \tag{3.26}$$

Consider first the terms in (3.26) where $|k|+|\ell| \geq \frac{RB}{100}$. The contribution of these terms is

$$\begin{aligned}
&\lesssim (RB)^{-\sigma}\sum_n\sum_{k,\ell}|a_{k\ell}|(1+|k|+|\ell|)^\sigma \sum_{\lambda,\mu}|\hat{\phi}(n-k-\lambda)\widehat{v_1}(n-k)\hat{\phi}(n-\ell-\mu)\widehat{v_2}(n-\ell)| \\
&\lesssim (RB)^{-\sigma}\sum_n\sum_{k,\ell}|a_{k\ell}|(1+|k|+|\ell|)^\sigma|\widehat{v_1}(n-k)\widehat{v_2}(n-\ell)| &(3.27)\\
&\lesssim (RB)^{-\sigma}\sum_{k,\ell}|a_{k\ell}|(1+|k|+|\ell|)^\sigma &(3.28)\\
&\lesssim (RB)^{-\sigma}
\end{aligned}$$

where the estimate (3.27) followed from (3.14) and (3.28) followed from the Cauchy-Schwartz inequality.
We next consider the terms in (3.26) where $|k|+|\ell| \leq \frac{RB}{100}$. If $|k|+|\ell| \leq \frac{RB}{100}$ and if $|n-\lambda| \geq RB$ then $|n-k-\lambda| \geq \frac{RB}{100}$. Accordingly the contribution of these terms is

$$\lesssim \sum_n\sum_{k,\ell}\sum_\mu\sum_{\lambda:|n-k-\lambda|\geq\frac{RB}{100}} |a_{k\ell}|\,|\hat{\phi}(n-k-\lambda)\widehat{v_1}(n-k)\hat{\phi}(n-\ell-\mu)\widehat{v_2}(n-\ell)| \tag{3.29}$$

Using (3.14) and (3.15) we estimate

$$\begin{aligned}
(3.29) &\lesssim R^{-998}\sum_n\sum_{k,\ell}|a_{k\ell}||\widehat{v_1}(n-k)\widehat{v_2}(n-\ell)| &(3.30)\\
&\lesssim R^{-998}.
\end{aligned}$$

Accordingly the expression (3.26) is $\lesssim R^{-998}+(RB)^{-\sigma}$, i.e. is $\lesssim R^{-998}$ if δ is small. The term $\sum_n II(n)$ is symmetric so we have proved (3.24).
Using (3.24) in (3.22), and using also the hypothesis (3.18), we conclude that

$$\begin{aligned}
|\sum_{n\in Q_M}\omega_n \sum_{\lambda,\mu:n\in U_{\lambda_\mu}} B(\phi_\lambda * v_1,\phi_\mu * v_2)(n)| &\geq (C\log\frac{1}{\delta})^{-1} - CR^{-998}\\
&\gtrsim (\log\frac{1}{\delta})^{-1}.
\end{aligned}$$

Equivalently

$$\left|\sum_{\lambda,\mu:U_{\lambda_\mu}\cap Q_M\neq\emptyset}\ \sum_{n\in U_{\lambda_\mu}\cap Q_M}\omega_n B(\phi_\lambda * v_1,\phi_\mu * v_2)(n)\right| \gtrsim (\log\frac{1}{\delta})^{-1}.$$

On the other hand

$$\sum_{\lambda,\mu:U_{\lambda_\mu}\cap Q_M\neq\emptyset}\|\psi_\lambda * v_1\|_2^2+\|\psi_\mu * v_2\|_2^2\lesssim R^2$$

by (3.14) and the fact that for each λ there are $\lesssim R^2$ values of μ with $U_{\lambda\mu}\neq\emptyset$. We may therefore choose λ and μ so that (3.21) holds. □

Lemma 3.5. *Let f, G, $\{\omega_n\}$ be as in Lemma 3.3 and let ϵ satisfy (3.7). If there are functions $u_1,u_2\in L^2(\mathbb{T}^2,\mathbb{C}^d)$ such that (3.16) and (3.17) hold then for some $k,\ell\in Q_{2M}$, some square $Q\in\mathcal{Q}$ which is contained in Q_{2M}, and some subset $U\subset Q$ there are components R_1 and R_2 and scalar valued L^2 functions w_1 and w_2 supported on the ϵ-neighborhoods of R_1 and R_2 respectively, such that*

$$\left|\sum_{n\in U}\omega_n\widehat{\psi * w_1}(n-k)\overline{\widehat{\psi * w_2}(n-\ell)}\right|\geq C^{-1}(\log\frac{1}{\delta})^{-5}\|w_1\|_2\|w_2\|_2. \tag{3.31}$$

Furthermore for each $Q\in\mathcal{Q}$ there are $\lesssim(\log\frac{1}{\delta})^4$ possibilities for the set U.

Proof. Letting w_1 and w_2 be the functions $\psi_\lambda * v_1$ and $\psi_\mu * v_2$ in (3.21) and using (3.12) we see that

$$\left|\sum_{n\in U_{\lambda_\mu}\cap Q_M}\omega_n B(\psi_\lambda * w_1,\psi_\mu * w_2)(n)\right|\gtrsim(\log\frac{1}{\delta})^{-3}\|w_1\|_2\|w_2\|_2$$

with w_1 and w_2 being vector valued L^2 functions each of which is supported in the ϵ-neighborhood of a single component. Passing to an appropriate matrix coefficient we may assume that G, w_1 and w_2 are scalar valued. Since $U_{\lambda\mu}$ intersects $\mathcal{O}((\log\frac{1}{\delta})^2)$ squares $Q\in\mathcal{Q}$, we then have

$$\left|\sum_{n\in U}\omega_n B(\psi_\lambda * w_1,\psi_\mu * w_2)(n)\right|\gtrsim(\log\frac{1}{\delta})^{-5}\|w_1\|_2\|w_2\|_2 \tag{3.32}$$

for some $Q\in\mathcal{Q}$, where $U=U_{\lambda\mu}\cap Q\cap Q_M$. Note that for fixed Q there are $\lesssim(\log\frac{1}{\delta})^4$ pairs (λ,μ) with $U_{\lambda\mu}\cap Q\neq\emptyset$; this gives the final statement of the lemma.
By (3.25) we have

$$B(\psi_\lambda * w_1,\psi_\mu * w_2)(n)=\sum_{ij}a_{ij}\widehat{\psi_\lambda * w_1}(n-i)\overline{\widehat{\psi_\mu * w_2}(n-j)}$$

so the left side of (3.32) is equal to

$$\left|\sum_{ij}a_{ij}\sum_{n\in U}\omega_n\widehat{\psi_\lambda * w_1}(n-i)\overline{\widehat{\psi_\mu * w_2}(n-j)}\right|$$

and is therefore

$$\lesssim\sum_{ij}|a_{ij}|(1+|i|+|j|)^\sigma\left|\sum_{n\in U}(1+|i|+|j|)^{-\sigma}\omega_n\widehat{\psi_\lambda * w_1}(n-i)\overline{\widehat{\psi_\mu * w_2}(n-j)}\right|$$

For each i and j, the sum over n here is $\lesssim(1+|i|+|j|)^{-\sigma}\|w_1\|_2\|w_2\|_2$. Accordingly, by (3.4), if (3.32) is to hold then there must be values of i and j with

$$\left|\sum_{n\in U}\omega_n\widehat{\psi_\lambda * w_1}(n-i)\overline{\widehat{\psi_\mu * w_2}(n-j)}\right|\gtrsim(1+|i|+|j|)^\sigma(\log\frac{1}{\delta})^{-5}\|w_1\|_2\|w_2\|_2 \tag{3.33}$$

This implies $|i|+|j| \lesssim (\log\frac{1}{\delta})^{\frac{5}{\sigma}}$, so $|i|+|j|$ is small compared with M. Of course (3.33) also implies

$$\left|\sum_{n\in U} \omega_n \widehat{\psi_\lambda * w_1}(n-i)\overline{\widehat{\psi_\mu * w_2}(n-j)}\right| \gtrsim (\log\frac{1}{\delta})^{-5}\|w_1\|_2\|w_2\|_2.$$

Replacing w_1 with $e^{2\pi i\lambda\cdot x}w_1$ and w_2 with $e^{2\pi i\mu\cdot x}w_2$ we obtain functions w_1 and w_2 each supported in the ϵ-neighborhood of a single component and with

$$\left|\sum_{n\in U} \omega_n \widehat{\psi * w_1}(n-i-\lambda)\overline{\widehat{\psi * w_2}(n-j-\mu)}\right| \gtrsim (\log\frac{1}{\delta})^{-5}\|w_1\|_2\|w_2\|_2$$

We now define $k = i+\lambda$ and $\ell = j+\mu$; then $k,\ell \in Q_{2M}$ and (3.31) holds. □

Our main probabilistic estimate is the following one to be proved in the next section.

Lemma 3.6. *Let R_1 and R_2 be two components. Let $\epsilon = \max(w_1, w_2)$, where w_i is the "width" of R_i as defined in Lemma 3.1, and let R_i^ϵ be the ϵ-neighborhood of R_i. Let U be a subset of $\mathbb{Z}^2$ which is contained in a square of side $\leq C\epsilon^{-1}$, and let $k,\ell \in \mathbb{Z}^2$. Let $\{\omega_n\}$ be iid's satisfying (3.2), (3.3) and (3.6). Define*

$$P(\rho) = \mathit{Prob}\Big(\exists u_1, u_2 \in L^2, \mathit{supp}(u_i) \subset R_i^\epsilon \\ \text{and } \left|\sum_{n\in U} \omega_n \widehat{\psi * u_1}(n-k)\overline{\widehat{\psi * u_2}(n-\ell)}\right| \geq \rho\|u_1\|_2\|u_2\|_2\Big).$$

Let $\rho = (\log\frac{1}{\delta})^{-N}$ with N fixed. Then, for any fixed $\eta > 0$, there is an estimate (uniform in R_1, R_2, k, ℓ and U)

$$P(\rho) \leq Ce^{-C^{-1}\delta^{-\alpha+\eta}} \tag{3.34}$$

provided κ in Lemma 3.1 has been chosen sufficiently small.

Proofs of Theorems 3.1 and 3.2. We first prove the requisite upper bound on $P(\delta)$, namely that $P(\delta) \leq e^{-\delta^{-\alpha'}}$ for fixed $\alpha' < \alpha$ and small δ.

If $h(M_\delta, \omega) < \delta$ then, there is a function $u \in \mathcal{P}(M)$ with $\|u\|_2 = 1$ such that

$$\int_{\mathbb{T}^2} f(\theta)|u(\theta)|^2 d\theta < \delta \tag{3.35}$$

and furthermore

$$\sum_n (1+\omega_n)B(u,u)(n) \leq \delta. \tag{3.36}$$

We have (assuming that Haar measure $d\theta$ on $\mathbb{T}^2$ has been normalized to have total mass 1)

$$\sum_n B(u,u)(n) = \int_{\mathbb{T}^2} \langle G(\theta,\theta)u(\theta), u(\theta)\rangle d\theta \geq C^{-1}.$$

This and (3.36) imply

$$\left|\sum_n \omega_n B(u,u)(n)\right| \geq (2C)^{-1}$$

which together with (3.35) gives the hypotheses of Lemma 3.3 with $u_1 = u_2 = u$. Accordingly (3.31) must hold for some choice of k, ℓ, $Q \in \mathcal{Q}$ and U and some choice of w_1 and w_2. The number of possible quadruples (k,ℓ,Q,U) is $\lesssim M^2 \cdot M^2 \cdot M^2 \cdot (\log\frac{1}{\delta})^8$, hence $\lesssim M^7$, say.

Therefore by the assumptions on M the number of possibilities for the quadruple is $\lesssim e^{7\delta^{-\alpha+\eta_0}}$, with a *fixed* value of $\eta_0 > 0$.
Accordingly it will suffice to show that for a fixed choice of k, ℓ, λ, μ, Q and U the probability of (3.31) holding for some u and w is less than $e^{-\delta^{-\alpha+\eta}}$ for arbitrary η. However, the latter estimate follows immediately from Lemma (3.6), taking $\rho = C^{-1}(\log \frac{1}{\delta})^{-5}$. This completes the proof of Theorem 3.2.

It remains to prove a corresponding bound in the opposite direction, under the hypotheses of Theorem 3.1.

Lemma 3.7. *Fix $\eta > 0$. Let R be a rectangle in $\mathbb{T}^2$. Then there is a function u with $\|u\|_2 = 1$ such that $\hat{u}$ is supported in the $\delta^{-\eta}$-dilation of the rectangle dual to R and centered at the origin and*

$$\|u\|_{L^2(\mathbb{T}^2 \setminus R)} \leq C_N \delta^{N\eta} \tag{3.37}$$

for all $N < \infty$.

Proof. This is a standard construction so we will omit some details. We can assume that R is centered at the origin. Let $g : \mathbb{R} \to \mathbb{R}$ be in the Schwartz space with L^2 norm 1 and such that $\text{supp}\hat{g}$ is contained in $[-1, 1]$. Given a rectangle $S \subset \mathbb{R}^2$, let h be a suitable rotation of a function of the form $(AB)^{\frac{1}{2}} g(Ax)g(By)$ for suitable A and B. Then h will satisfy $\text{supp}\hat{h} \subset S^*$ and (from the Schwartz space bounds for g)

$$\|h\|_{L^2(\mathbb{R}^2 \setminus TS)} \leq C_N T^{-N} \tag{3.38}$$

for any N.
If R is as indicated then we can lift R up to a rectangle in $\mathbb{R}^2$ also centered at the origin. Let S be the δ^{η}-dilation of the lifted rectangle and construct h as above. Let $u(x) = \sum_{\nu \in \mathbb{Z}^2} u(x - \nu)$ be the periodization of h. The bound (3.38) implies that u decays rapidly outside the projection of S onto $\mathbb{T}^2$, hence (3.37). The Fourier support statement for u follows from the corresponding one for h. $\square$ To finish the proof of Theorem 3.1, first assume we are in the scalar valued case.
Fix a small value of η. Consider a rectangle R as in Lemma 3.2, with δ replaced by $\frac{\delta}{3}$. Let R_1 be the dilation of R around its center by a factor δ^{η}. Lemma 3.7 shows that there is a function u with the following properties: $\hat{u}$ is supported in R_1^*, $\|u\|_2 = 1$ and $\|u\|_{L^2(R^c)} \leq C_{N\eta}\delta^{N\eta}$ for each fixed N. Then $u \in \mathcal{P}(M)$ if η is small, since $M > \delta^{-(\alpha+\eta_0)}$. Also

$$\begin{aligned} \int_{\mathbb{T}^2} f|u|^2 &= \int_R f|u|^2 + \int_{\mathbb{T}^2 \setminus R} f|u|^2 \\ &\leq \frac{\delta}{3} + C_{N\eta}\delta^{N\eta} \end{aligned}$$

for any N. If we take N large depending on η and then let δ be small this is less than $\frac{\delta}{2}$.
On the other hand, consider the event $E = \{\omega : \omega_n \leq -1 + C^{-1}\delta \text{ for all } n \text{ in the double of } R_1^*\}$.
Suppose that $\omega \in E$.

$$\begin{aligned} \sum_n \frac{1}{2}(1 + \omega_n) B(u, u)(n) &= \sum_{k,\ell} a_{k\ell} \sum_n \frac{1}{2}(1 + \omega_n)\hat{u}(n - k)\overline{\hat{u}(n - \ell)} \\ &= \sum_{|k|+|\ell| \geq J} a_{k\ell} \sum_n \frac{1}{2}(1 + \omega_n)\hat{u}(n - k)\overline{\hat{u}(n - \ell)} \\ &+ \sum_{|k|+|\ell| < J} a_{k\ell} \sum_n \frac{1}{2}(1 + \omega_n)\hat{u}(n - k)\overline{\hat{u}(n - \ell)} \end{aligned}$$

where $J = \delta^{-\eta}$.
The first term is bounded by $\sum_{|k|+|\ell|\geq J} |a_{k\ell}| \sum_n |\hat{u}(n-k)||\hat{u}(n-\ell)|$, which is $\leq C_{N\eta}\delta^{N\eta}$ for any N by smoothness of G, and is therefore $< \frac{\delta}{4}$ for small δ. On the other hand, if $|k|+|\ell| < \delta^{-\eta}$ and if $n-k$ is in R_1^* then n is in the double of R_1^*, so $1+\omega_n \in [0, C^{-1}\delta]$. It follows that the second term is $< \frac{\delta}{4}$ if C has been chosen appropriately. So $\sum_n \frac{1}{2}(1+\omega_n)B(u,u)(n) < \frac{\delta}{2}$ and it follows that $h(M,\omega) < \delta$. By (3.3) E has probability $\geq e^{-\delta^{-\eta}|R_1^*|}$ if δ is small. Also $|R_1^*| \approx \delta^{-\alpha-2\eta}$, so

$$\mathrm{Prob}(E) \quad \geq \quad e^{-C\delta^{-\alpha-3\eta}}.$$

Letting $\eta \to 0$ we obtain the necessary bound

$$\limsup_{\delta\to 0} \frac{\log|\log P(\delta)|}{-\log(\delta)} \leq \alpha$$

in the scalar valued case. The vector case follows by taking u to have one coordinate equal to the function constructed above and the others equal to zero. $\square$

4. Probability estimates, part 2

The purpose of this section is to prove Lemma 3.6. We split the argument into two cases handled in subsections 4.1 and 4.2 respectively: both R_1 and R_2 are type III components, or (at least) one of them is type I or II. This dichotomy arises as follows: if R_i is type I or II then one has ℓ^q information on the Fourier transform of a function supported in R_i, and this information can be used together with estimates for sums of independent random variables, since there are good estimates for a sum of independent random variables assuming control on a higher moment. On the other hand, if both R_1 and R_2 are type III, then one does not have a sufficient ℓ^q bound, but the relevant Fourier transforms are roughly constant on families of parallel lines. This makes it possible to use a simple estimate on random matrices (Lemma 4.3 below).

4.1. Flat components.

The following lemma is basically a reformulation of Lemma 3.6 in the case of the type III components.

Lemma 4.1. *Suppose that ℓ_1 and ℓ_2 are lines with rational slopes, and C_0 is any positive constant. Assume that $\rho \geq \epsilon^{\frac{1}{25}}$. Let $\{\omega_n\}$ satisfy (3.2), (3.6). Then, for any cube Q of side $N \leq C_0\epsilon^{-1}$ and any subset $U \subset Q$, the probability that there are functions u_1 and u_2 supported on the the $C_0\epsilon$-neighborhoods of the ℓ_1 and ℓ_2 respectively and with*

$$\left|\sum_{n\in U} \omega_n \widehat{u_1}(n)\overline{\widehat{u_2}(n)}\right| \geq \rho\|u_1\|_2\|u_2\|_2 \tag{4.1}$$

is $\leq Ce^{-C^{-1}\rho^3\epsilon^{-1}}$. Here C depends on C_0 and the lines ℓ_1 and ℓ_2.

Proof of Lemma 3.6 when R_1 and R_2 are type III, assuming Lemma 4.1. Notice first of all that any type III R_i is a $C\delta^{\alpha-\kappa}$-neighborhood of a line with rational slope; see the beginning of section 3. Next note that Lemma 4.1 remains true with (4.1) replaced by

$$\left|\sum_{n\in U} \omega_n \widehat{u_1}(n-k)\overline{\widehat{u_2}(n-\ell)}\right| \geq \rho\|u_1\|_2\|u_2\|_2$$

for (say) any given k and ℓ, since this just corresponds to multiplying u_1 and u_2 by characters. We can further replace u_i with $\psi * u_i$ with $\psi = \psi_\epsilon$ as in section 3, since ψ is supported in $D(0,\epsilon)$.

Thus Lemma 4.1 is true with (4.1) replaced by

$$\left|\sum_{n\in U}\omega_n\widehat{\psi*u_1}(n-k)\overline{\widehat{\psi*u_2}(n-\ell)}\right|\geq\rho\|\psi*u_1\|_2\|\psi*u_2\|_2$$

a fortiori with (4.1) replaced by

$$\left|\sum_{n\in U}\omega_n\widehat{\psi*u_1}(n-k)\overline{\widehat{\psi*u_2}(n-\ell)}\right|\geq\rho\|u_1\|_2\|u_2\|_2.$$

Applying this with $\epsilon=\max(w_1,w_2)\approx\delta^{\alpha-\kappa}$ and with $\rho\approx(\log\frac{1}{\delta})^{-N}$ and then taking κ sufficiently small, we obtain the bound (3.34). □

For the proof of Lemma 4.1 we need two purely probabilistic lemmas. The first is almost trivial; the second (a crude large deviations type estimate for the norm of a random matrix) is a bit less trivial but presumably not new.

Lemma 4.2. *Let $C_1<\infty$, $\rho>0$. Consider a set E which is the disjoint union of m "blocks" E_i, each with cardinality less than C_1r. Assume that $\frac{\log m}{\rho^2 r}$ is sufficiently small. Let t_{ij} be independent random variables with mean zero and $|t_{ij}|\leq C_1$. Define*

$$P(\rho)=Prob\left(\exists\{a_i\}_{i=1}^m:\left|\sum_{i=1}^m\sum_{j\in E_i}t_{ij}a_i\right|\geq\rho r\sum_{i=1}^m|a_i|\right).$$

Then

$$P(\rho)\leq Ce^{-C^{-1}\rho^2r}$$

where C depends on C_1.

Proof. If any such $\{a_i\}$ exists then there has to be at least one block E_i such that $|\sum_{j\in E_i}t_{ij}|\geq\rho r$. The usual exponential estimate for sums of independent random variables (see e.g. [34] or the proof of Lemma 4.8 below) shows that the probability of the latter for any given i is at most $Ce^{-C^{-1}\rho^2r}$. Since $\log m<<\rho^2r$ the conclusion follows. □

Lemma 4.3. *Let $C_1<\infty$, $\rho>0$. Let T be an $N\times N$ matrix whose entries are independent random variables t_{ij} with mean zero and with $|t_{ij}|\leq C_1$. Then, letting $\|T\|$ be the operator norm, there is an estimate*

$$Prob(\|T\|\geq\rho N)\leq Ce^{-C^{-1}\rho^8N^2}$$

provided N is large compared with $\rho^{-4}\log\frac{1}{\rho}$. Here C depends on C_1.

Proof. Fix i and j with $i\neq j$ and consider the (i,j) entry of TT^*, i.e.

$$m_{ij}=\sum_k t_{ik}t_{jk}.\tag{4.2}$$

The quantity (4.2) is a sum of N independent random variables bounded by C_1^2 and with mean zero, so by the exponential estimate that was used in the preceding proof

$$\mathrm{Prob}(|m_{ij}|\geq\rho^2N)\leq C_2e^{-C_2^{-1}\rho^4N}.\tag{4.3}$$

Now let $E\subset\{1,\dots,N\}\times\{1,\dots,N\}\backslash\text{diagonal}$ be a set with the following property: $(*)$ If $(i,j)\in E$, $(k,\ell)\in E$, then $\{i,j\}\cap\{k,\ell\}=\emptyset$. If E has property $(*)$ then the random variables

$\{m_{ij} : (i,j) \in E\}$ are independent. A set with property $(*)$ has cardinality $\le \frac{N}{2}$, so

$$\begin{aligned} & \mathrm{Prob}(\exists (i_1, j_1), \ldots, (i_{[\rho^4 N]}, j_{[\rho^2 N]}) \in E : |m_{i_\ell j_\ell}| \ge \rho^2 N \ \forall \ell \in \{1, \ldots, [\rho^2 N]\}) \\ \le & \begin{pmatrix} [\frac{N}{2}] \\ [\rho^4 N] \end{pmatrix} (C_2 e^{-C_2^{-1} \rho^4 N})^{[\rho^4 N]} \\ \le & C_3 e^{C_3 N \rho^2 \log \frac{1}{\rho}} e^{-C_3^{-1} \rho^8 N^2} \\ \le & C_4 e^{-C_4^{-1} \rho^8 N^2}. \end{aligned}$$

It is possible to express $\{1, \ldots, N\} \times \{1, \ldots, N\} \backslash$diagonal as the union of (say) $\le 3N$ sets E_k, each of which has property $(*)$. This is entirely elementary; we indicate a way of doing it at the end of the proof. Letting k vary and using the first part of the proof, we have

$$\begin{aligned} & \mathrm{Prob}(\exists k \text{ and } (i_1, j_1), \ldots, (i_{[\rho^4 N]}, j_{[\rho^4 N]}) \in E_k : |m_{i_\ell j_\ell | \ge \rho^2 N \ \forall \ell \in \{1, \ldots, [\rho^4 N]}) \\ \le & 4 N C_4 e^{-C_4^{-1} \rho^8 N^2} \\ \le & C e^{-C^{-1} \rho^8 N^2}. \end{aligned}$$

Let $Y = \{(i,j) : |m_{ij}| \ge \rho^2 N\}$. If the square of the Schmidt norm of TT^* is $\ge 5\rho^4 N^4$, then we have

$$5\rho^4 N^4 \le N^2 |Y| + (\rho^2 N)^2 N^2$$

whence $|Y| \ge 4\rho^4 N^2$, and Y must intersect the complement of the diagonal in cardinality $\ge 3\rho^4 N^2$. Thus for some k the set E_k must contain at least $\rho^4 N$ pairs (i,j) with $|m_{ij}| \ge \rho^2 N$. It follows that the probability that the Schmidt norm of TT^* is $\ge 5^{\frac{1}{2}} \rho^2 N^2$ is $\le C e^{-C^{-1} \rho^8 N^2}$. Hence the same is true with the Schmidt norm of TT^* replaced by the operator norm, hence the probability that the operator norm of T is $\ge 5^{\frac{1}{4}} \rho N$ is $\le C e^{-C^{-1} \rho^8 N^2}$ as claimed.

Here is one way to show that the complement of the diagonal in $\{1, \ldots, N\} \times \{1, \ldots, N\}$ is the union of $\le 3N$ sets each of which has property $(*)$:

1. let $\mathbb{Z}_N$ be the cyclic group with N elements and let $j \in \mathbb{Z}_N$, $j \ne 0$. Then $\mathbb{Z}_N$ is the union of three sets S_1, S_2, S_3 such that $S_i \cap (S_i + j) = \emptyset$ for each $i \in \{1, 2, 3\}$. For this, let S be maximal subject to the condition that $S \cap (S + j) = \emptyset$. The maximality implies $S - j \cup S \cup S + j = \mathbb{Z}_N$, as claimed.
2. For $k \in \{1, \ldots, N-1\}$ let

$$F_k = \{(i, \sigma_k(i) : i \in \{1, \ldots, N\}\}$$

where $\sigma_k(i) = i + k$ if $i \le N - k$ and $\sigma_k(i) = i + k - N$ if $i > N - k$.

Using (1) we can express each F_k as the union of three sets with property $(*)$. Letting k vary we have expressed $\{1, \ldots, N\} \times \{1, \ldots, N\} \backslash$diagonal as the union of $3(N-1)$ such sets. □ We now reduce to a situation where the above lemmas can be applied, by exploiting the fact that if u is supported close to a line then its Fourier transform will be roughly constant on orthogonal lines.

Let p and q be relatively prime integers and let $e = (-q, p)$. Let N be large, and let α be small but such that αN is large. Then we can partition $\mathbb{Z}^2$ as follows:

$$\mathbb{Z}^2 = \cup_i E_i \quad \text{(disjoint union)}$$
$$E_i = \{\nu_i + je : 0 \le j \le \alpha N\}.$$

For this, note that any two of the sets $\{\nu + je : j \in \mathbb{Z}\}$ are identical or disjoint. Now partition $\mathbb{Z}$ in intervals of length αN and form the corresponding partition of each line $\{\nu + je : j \in \mathbb{Z}\}$.

Lemma 4.4. *Assume that $u : \mathbb{T}^2 \to \mathbb{C}$ is supported in the $C_0\epsilon$-neighborhood of the line $\ell = \{x + t(p,q), t \in \mathbb{R}\}$, where $\epsilon \approx N^{-1}$. Let $\{E_i\}$ be the partition of $\mathbb{Z}^2$ corresponding to the given*

p and q.
Then there is a function g constant on each E_i and

$$\|\hat{u}(n)e^{-i\xi\cdot n} - g(n)\|_{\ell^2} \lesssim \alpha\|u\|_2$$

Here $\xi = \xi(\ell)$ depends only on the line and not on u.

Proof. We assume that ℓ passes through the origin; then ξ will be equal to 0. The general case can be easily reduced to this case. We may also clearly assume $\|u\|_2 = 1$.
Let χ be the indicator function of the $C_0\epsilon$-neighborhood of ℓ. Then $\hat{u} = \hat{u} * \hat{\chi}$, and

$$\hat{u}(n+e) - \hat{u}(n) = \widehat{hu}(n)$$

where $h(x) = (e^{-2\pi i x\cdot e} - 1)\chi(x)$ has L^∞ norm $\lesssim \epsilon$. We conclude that

$$\|\hat{u}(\cdot + e) - \hat{u}\|_2^2 \lesssim \epsilon^2 \tag{4.4}$$

Define g by conditional expectation of $\hat{u}$ onto the segments E_i; i.e g is constant on E_i and $\sum_{n\in E_i} g(n) = \sum_{n\in E_i} \hat{u}(n)$ for each i. Then

$$|\hat{u}(n) - g(n)| \lesssim \sum_{|i|\le\alpha N} |\hat{u}(n+(i+1)e) - \hat{u}(n+ie)|$$

whence

$$\begin{aligned}
\sum_n |\hat{u}(n) - g(n)|^2 &\lesssim \sum_n \left(\sum_{|i|\le\alpha N} |\hat{u}(n+(i+1)e) - \hat{u}(n+ie)|\right)^2 \\
&\lesssim \alpha N \sum_n \sum_{|i|\le\alpha N} |\hat{u}(n+(i+1)e) - \hat{u}(n+ie)|^2 \\
&\approx (\alpha N)^2 \sum_n |\hat{u}(n+e) - \hat{u}(n)|^2 \\
&\lesssim \alpha^2
\end{aligned}$$

using (4.4). The lemma follows. □

Lemma 4.5. *Suppose that ℓ_1 and ℓ_2 are lines with the same rational slope (p,q). Let $\{\omega_n\}$, Q and U be as in Lemma 3.1. Then the probability of the event*

There are functions u_1 and u_2 supported on the $C_0\epsilon$-neighborhoods of ℓ_1 and ℓ_2 respectively and with $\left|\sum_{n\in U} \omega_n \widehat{u_1}(n)\overline{\widehat{u_2}(n)}\right| \ge \rho\|u_1\|_2\|u_2\|_2$

is $\le Ce^{-C^{-1}\rho^3 N}$ where C depends on C_0, ℓ_1 and ℓ_2.

Proof. Assume that such functions u_1 and u_2 exist. Let $\{E_i\}$ be the partition as in Lemma 4.5 corresponding to the given p and q, letting $\alpha = C^{-1}\rho$ with C a large constant. Let ξ_1 and ξ_2 correspond to the lines ℓ_1 and ℓ_2 by Lemma 4.4, and let g_1 and g_2 correspond to $\widehat{u_1}$ and $\widehat{u_2}$. Then $\sum_n |\widehat{u_1}(n)\overline{\widehat{u_2}(n)} - g_1(n)\overline{g_2}(n)e^{i(\xi_1-\xi_2)\cdot n}| \lesssim \alpha\|u_1\|_2\|u_2\|_2$. Hence, if $|\sum_{n\in Q}\omega_n\widehat{u_1}(n)\overline{\widehat{u_2}(n)}| \ge \rho\|u_1\|_2\|u\|_2$, then (provided C was chosen large enough) also

$$\left|\sum_{n\in U} \omega_n g_1(n)\overline{g_2}(n)e^{i(\xi_1-\xi_2)\cdot n}\right| \ge \frac{\rho}{2}\|g_1\|_{\ell^2}\|g_2\|_{\ell^2} \tag{4.5}$$

with g_1 and g_2 constant on each E_i.
We now reindex, letting $n \sim (i,j)$ where $n \in E_i$ and $\{(i,j)\}$ for fixed i is an enumeration of E_i. We let E be the union of all "blocks" E_i which intersect Q. If $n \sim (i,j)$ then we let

$$t_{ij} = \omega_n \chi_U(n) e^{i(\xi_1-\xi_2)\cdot n}.$$

There are $\lesssim \rho^{-2}N$ E_i's which intersect Q, and each has cardinality $C^{-1}\rho N$. Accordingly, applying Lemma 4.2 with $E = Q$ and with $m \approx \rho^{-2}N$ and $r = C^{-1}\rho N$, we conclude that the probability that there are numbers a_i with $|\sum_{ij} t_{ij}a_i| \geq \frac{\rho}{2}\cdot C^{-1}\rho N \sum_i |a_i|$ is at most $Ce^{-C^{-1}\rho^3 N}$. It follows that the probability that there is a function $h(n)$ constant on each E_i and with

$$\left|\sum_{n\in U} \omega_n e^{i(\xi_1-\xi_2)\cdot n} h(n)\right| \geq \frac{\rho}{2}\sum_n |h(n)| \tag{4.6}$$

is at most $Ce^{-C^{-1}\rho^3 N}$. On the other hand if (4.5) holds then (4.6) holds with $h(n) = g_1(n)\overline{g_2}(n)$, by the Cauchy-Schwartz inequality. This completes the proof. $\square$

Lemma 4.6. *Suppose that ℓ_1 and ℓ_2 have different slopes; let $\{\omega_n\}$, Q and U be as in Lemma 3.1. Then the probability of the event*

There are functions u_1 and u_2 supported on the $C_0\epsilon$-neighborhoods of ℓ_1 and ℓ_2 respectively

$$\text{and with } \left|\sum_{n\in U} \omega_n \widehat{u_1}(n)\overline{\widehat{u_2}(n)}\right| \geq \rho\|u\|_1\|u\|_2$$

is $\leq Ce^{-C^{-1}\rho^{28}N^2}$ where C depends on C_0, ℓ_1 and ℓ_2.

Proof. Let $\{E_i\}$ be the partition corresponding to the line ℓ_1, and let $\{F_j\}$ be the partition corresponding to the line ℓ_2, where we set $\alpha = C^{-1}\rho$. As in the proof of Lemma 4.5 it suffices (for a certain fixed ξ) to prove the same bound for the probability of the event

There are functions g_1 and g_2 such that g_1 is constant on each E_i,
g_2 is constant on each F_j, and

$$\left|\sum_{n\in U} \omega_n g_1(n)\overline{g_2}(n)e^{i\xi\cdot n}\right| \geq \frac{\rho}{2}\|g_1\|_{\ell^2}\|g_2\|_{\ell^2} \tag{4.7}$$

There are $\lesssim \rho^{-2}N$ E_i's and $\lesssim \rho^{-2}N$ F_j's which intersect Q. Each E_i and F_j intersect at most once, and the (partially defined) map from (i,j) to the intersection point is $1-1$. Define

$$G = \{(i,j):\ E_i \text{ and } F_j \text{ intersect at a point of } U\}$$
$$t_{ij} = \chi_G(i,j)e^{i\xi\cdot n}\omega_n$$

where $n = n(i,j)$ is the intersection point of E_i and F_j. By Lemma 4.3 we have

$$\text{Prob}\left(\exists a,b\in\mathbb{C}^{C\rho^{-2}N}:\left|\sum_{ij} t_{ij}a_i\overline{b_j}\right| \geq \rho^4\cdot\rho^{-2}N\|a\|_{\ell^2}\|b\|_{\ell^2}\right) \leq Ce^{-C^{-1}\rho^{32}(\rho^{-2}N)^2} \tag{4.8}$$

On the other hand, if (4.7) holds for a certain g_1 and g_2 as there, then define $a_i = g_1(n)$ if $n \in E_i$ and $b_j = g_2(n)$ if $n \in F_j$. Since each E_i and F_j has cardinality $\gtrsim \rho N$, we then have

$$\sum_i |a_i|^2 \lesssim \frac{1}{\rho N}\|g_1\|_{\ell^2}^2 \tag{4.9}$$

$$\sum_j |b_j|^2 \lesssim \frac{1}{\rho N}\|g_2\|_{\ell^2}^2 \tag{4.10}$$

Thus

$$\begin{aligned}\left|\sum_{ij} t_{ij} a_i \overline{b_j}\right| &= \left|\sum_{n\in U} \omega_n g_1(n)\overline{g_2}(n) e^{i\xi\cdot n}\right| \\ &\geq \frac{\rho}{2}\|g_1\|_{\ell^2}\|g_2\|_{\ell^2} \\ &\gtrsim \rho^2 N \|a\|_{\ell^2}\|b\|_{\ell^2}.\end{aligned}$$

The first line was just a change in notation, the second line was (4.7) and the last line used (4.9), (4.10). It therefore follows by (4.8) that the probability of the event (4.7) is $\leq Ce^{-C^{-1}\rho^{28}N^2}$, as claimed. □

Proof of Lemma 4.1. If the lines ℓ_1 and ℓ_2 have the same slope we apply Lemma 4.5. If they have different slopes we apply Lemma 4.6 and use that $\rho^{28}N^2 \gtrsim \rho^3 N$. □

4.2. **Nonflat components.** We now consider the remaining case of Lemma 3.6 where either R_1 or R_2 is type I or II.

We need the following entropy type estimate, which will play an analogous role to Lemma 4.4. Here $\psi = \psi_\epsilon$ is as in section 3.

Lemma 4.7. *Fix $\rho > 0$, $\epsilon > 0$. Let γ be a rectifiable curve, let Z be the L^2 functions supported in the ϵ-neighborhood of γ, and let B be the unit ball in Z. Then there is a set of $\lesssim e^{C(\rho^2\epsilon)^{-1}\log\frac{1}{\rho}}$ functions $f_j \in B$, so that if $f \in B$ then $\|\psi * (f - f_j)\|_2 \leq \rho$ for some j.*

Proof. Let $\mathcal{Q}$ be all dyadic squares with side $C_1^{-1}\rho^2\epsilon$ which intersect the ϵ-neighborhood of γ. The cardinality of $\mathcal{Q}$ is $\lesssim (\rho^2\epsilon)^{-1}$.
For each $Q \in \mathcal{Q}$ let c_Q be the center of Q. For each $f \in Z$, let g_f be the conditional expectation of f onto the partition $\mathcal{Q}$, i.e. g_f is constant on each Q and $\int_Q g_f = \int_Q f$. Notice that $\|g_f\|_2 \leq \|f\|_2$. We have

$$\begin{aligned}|\psi * (f - g_f)(x)| &= \left|\sum_Q \int_Q \psi(x-y)(f-g_f)(y)dy\right| \\ &= \left|\sum_Q \int_Q (\psi(x-y) - \psi(x-c_Q)(f-g_f)(y)dy\right| \\ &\leq \int \sum \chi_Q(y)|\psi(x-y) - \psi(x-c_Q)|\,|(f-g_f)(y)|dy. \qquad (4.11)\end{aligned}$$

Using the derivative bounds on ψ, namely that $|\nabla\psi| \lesssim \epsilon^{-1}$, one can easily show that for fixed $y \in Q$ there is an estimate $\int |\psi(x-y) - \psi(x-c_Q)|dx \lesssim \epsilon^{-1}\mathrm{diam}(Q)$. This can be made $< A^{-1}\rho^2$ for any given A by making C_1 large enough. For fixed x the integral $\int \sum_Q \chi_Q(y)|\psi(x-y) - \psi(x-c_Q))|dy$ is bounded essentially by the L^1 norm of ψ, i.e. by a constant. It follows by Schur's test that the L^2 norm of the quantity (4.11) is small compared with ρ if C_1 is large enough.
Thus $\|\psi * (f - g_f)\|_2 \leq \frac{\rho}{3}$ for all $f \in B$ if C_1 has been chosen appropriately. On the other hand, the functions $\{\psi * g_f : f \in B\}$ are a subset of a ball of bounded radius in the d-dimensional Hilbert space $\{\psi * g_f : f \in Z\} \subset L^2$, where $d \lesssim (\rho^2\epsilon)^{-1}$. So there is a subset $\{f_j\} \subset B$ with cardinality $\lesssim e^{Cd\log\frac{1}{\rho}}$ such that if $f \in B$ then $\|\psi * g_f - \psi * g_{f_j}\|_2 \leq \frac{\rho}{3}$ for some j. For this, see for example [36], formula (4.22), the case where (in the notation of [36]) B_1 and B_2 are both

the standard unit ball. It follows that for any $f \in B$ there is j such that

$$\begin{aligned}\|\psi * (f - f_j)\|_2 &\leq \|\psi * (f - g_f)\|_2 + \|\psi * (g_f - g_{f_j})\|_2 + \|\psi * (g_{f_j} - f_j)\|_2 \\ &< \rho.\end{aligned}$$

The lemma is proved. $\square$

We also need the following estimate for sums of independent random variables.

Lemma 4.8. *Let $q \geq 4$. Assume that $\{\omega_n\}$ satisfy (3.6), (3.2). Let $\{a_n\}$ and $\{b_n\}$ be numerical sequences and suppose (for a certain $\epsilon > 0$) that*

$$\|a\|_{\ell^q} \leq \epsilon \|a\|_{\ell^2}. \tag{4.12}$$

Then

$$Prob\left(\left|\sum_j \omega_j a_j b_j\right| \geq \rho \|a\|_2 \|b\|_2\right) \leq C e^{-C_\rho^{-1} \epsilon^{-\frac{2q}{q-2}}} \tag{4.13}$$

where we can take $C_\rho^{-1} = C\rho^{2+\frac{q}{q-2}}$.

Remark 4.1. Observe that no ℓ^q type assumption is made on the $\{b_i\}$. It is important for us that no such assumption is necessary, because we will have a suitable ℓ^q bound for the Fourier transform of only one of the two functions u_1 and u_2 in Lemma 3.6. The "reason" why such an assumption is unnecessary is that if $a \in \ell^q, b \notin \ell^q$, then a and b may be expected to be quite different so that (heuristically speaking; see below for a rigorous argument) $\sum_i |a_i b_i| << \|a\|_{\ell^2} \|b\|_{\ell^2}$.

Proof. Let C_1 be a bound for $|\omega_n|$. We first discuss a basically well known weaker statement which covers the case where both a and b satisfy (4.12). Namely, with $p \geq 2$ suppose that

$$\|c\|_{\ell^1} \leq \beta \tag{4.14}$$

$$\|c\|_{\ell^p} \leq \epsilon \beta. \tag{4.15}$$

Then

$$\text{Prob}(|\sum_j \omega_j c_j| \geq t\beta) \leq C e^{-C^{-1} t^2 \epsilon^{-p'}} \tag{4.16}$$

for an appropriate constant C. Namely, it suffices to prove this when the $\{c_j\}$ are real. Assume that they are and start from the usual subgaussian estimate (the constant depends on the bound in (3.2))

$$\mathbb{E}(e^{\lambda \sum_j \omega_j c_j}) \leq e^{C\lambda^2 \sum_j c_j^2} \tag{4.17}$$

See for example [34]. Let $p = \frac{q}{2}$. We have

$$\sum_j c_j^2 \leq \epsilon^{p'} \beta^2$$

by Hölder's inequality, (4.14) and (4.15). By (4.17)

$$\mathbb{E}(e^{\lambda \sum_j \omega_j c_j}) \leq e^{C\lambda^2 \epsilon^{p'} \beta^2} \tag{4.18}$$

whence

$$\text{Prob}\left(\sum_j \omega_j c_j \geq t\beta\right) \leq e^{-\beta t \lambda + C\lambda^2 \epsilon^{p'} \beta^2}$$

Now take $\lambda = \frac{t}{2C\epsilon^{p'}\beta}$ to obtain

$$\text{Prob}\left(\sum_j \omega_j c_j \geq t\beta\right) \leq e^{-C^{-1}t^2\epsilon^{-p'}}$$

for suitable C, and therefore (4.16) since we can consider also $-c$.

To prove the lemma let $t_\Sigma = \sum_{i\in\Sigma} |a_i b_i|$. Let $X_\lambda = \{i : |a_i| \geq \lambda\}$. Then

$$\begin{aligned}\sum_{i\in X_\lambda} |a_i b_i| &\leq \left(\sum_{i\in X_\lambda} |a_i|^2\right)^{\frac{1}{2}} \|b\|_{\ell^2} \\ &\leq \lambda^{-\frac{q-2}{2}} \left(\sum_{i\in X_\lambda} |a_i|^q\right)^{\frac{1}{2}} \|b\|_{\ell^2} \\ &\leq \lambda^{-\frac{q-2}{2}} \epsilon^{\frac{q}{2}} \|a\|_{\ell^2}^{\frac{q}{2}} \|b\|_{\ell^2}\end{aligned}$$

by (4.12). So if we set $\lambda = C(\frac{\epsilon^q}{\rho^2})^{\frac{1}{q-2}}\|a\|_{\ell^2}$ with C large then

$$t_{X_\lambda} \leq \frac{\rho}{4C_1}\|a\|_{\ell^2}\|b\|_{\ell^2}.$$

Now let $Y_\mu = \{i : |b_i| \geq \mu\}$ and $Z_{\lambda\mu} = Y_\mu \cap X_\lambda^c$. Then

$$\sum_{i\in Z_{\lambda\mu}} |a_i b_i| \leq \lambda \sum_{i\in Y_\mu} |b_i| \leq \frac{\lambda}{\mu}\|b\|_{\ell^2}^2$$

so if we let $\mu = C\rho^{-1}\lambda\frac{\|b\|_{\ell^2}}{\|a\|_{\ell^2}} = C^2\rho^{-\frac{q}{q-2}}\epsilon^{\frac{q}{q-2}}\|b\|_{\ell^2}$ then

$$t_{Z_{\lambda\mu}} \leq \frac{\rho}{4C_1}\|a\|_{\ell^2}\|b\|_{\ell^2}.$$

Let $\Sigma' = Y_\mu^c \cap X_\lambda^c$, $\Sigma'' = Z_{\lambda\mu} \cup X_\lambda$. Then

$$t_{\Sigma''} \leq \frac{\rho}{2C_1}\|a\|_{\ell^2}\|b\|_{\ell^2} \tag{4.19}$$

whereas

$$\sum_{i\in\Sigma'} |b_i|^q \leq \sum_{i\in Y_\mu^c} |b_i|^q \lesssim \mu^{q-2}\|b\|_{\ell^2}^2 \approx \rho^{-q}\epsilon^q\|b\|_{\ell^2}^q \tag{4.20}$$

We conclude

$$\sum_{i\in\Sigma'} |a_i b_i|^{\frac{q}{2}} \leq \rho^{-\frac{q}{2}}\epsilon^q\|a\|_{\ell^2}^{\frac{q}{2}}\|b\|_{\ell^2}^{\frac{q}{2}} \tag{4.21}$$

using Cauchy-Schwartz, (4.12) and (4.20). Now consider the effect of randomization. We have

$$\left|\sum_{i\in\Sigma''} \omega_i a_i b_i\right| \leq \frac{\rho}{2}$$

by (4.19). The contribution from Σ' can be estimated using (4.21) and (4.16) with $p = \frac{q}{2}$. Thus, we take $\beta = \|a\|_{\ell^2}\|b\|_{\ell^2}$, $t = \rho$, and we take ϵ in (4.16) to be approximately $\rho^{-1}\epsilon^2$. We conclude that

$$\begin{aligned}\text{Prob}\left(\left|\sum_j \omega_j a_j b_j\right| \geq \rho\|a\|_{\ell^2}\|b\|_{\ell^2}\right) &\leq \text{Prob}\left(\left|\sum_{j\in\Sigma'} \omega_j a_j b_j\right| \geq \frac{\rho}{2}\|a\|_{\ell^2}\|b\|_{\ell^2}\right)\\ &\leq Ce^{-C^{-1}\frac{\rho^2}{(\rho^{-1}\epsilon^2)^{(\frac{q}{2})'}}}\end{aligned}$$

which is (4.13). □

Lemma 4.9. *Suppose that $u, v \in L^2$, and assume that either*

$$\|\hat{u}\|_{\ell^q} < \epsilon\|u\|_2 \text{ or } \|\hat{v}\|_{\ell^q} < \epsilon\|v\|_2.$$

Then, for any k and ℓ, and any set $U \subset \mathbb{Z}^2$,

$$\textit{Prob}\left(\left|\sum_{n\in U} \omega_n \hat{u}(n-k)\overline{\hat{v}(n-\ell)}\right| \geq \rho\|u\|_2\|v\|_2\right) \leq Ce^{-C^{-1}\rho^{2+\frac{q}{q-2}}\epsilon^{-\frac{2q}{q-2}}}$$

Proof. This follows from Lemma 4.8 and the Plancherel theorem. □

To apply the preceding lemmas we need an ℓ^q bound for the Fourier transform of a function supported on a type I or II R_i, which can be obtained from Bernstein's inequality in the type I case and the Stein-Tomas theorem in the type II case. We first give a minor variant on the Stein-Tomas theorem.

Lemma 4.10. *Fix $q > 6$. Let γ be a piece of curve in $\mathbb{R}^2$ with length $\tau \leq \min(1, k^{-1})$, curvature bounded between $C^{-1}k$ and Ck and such that the ith derivative of the curvature with respect to arc length is bounded by $C_i k\tau^{-i}$ for all i. Then provided $w \leq \tau$ there is a bound*

$$\|\hat{f}\|_q \lesssim w^{\frac{1}{2}}k^{-\frac{1}{q}}\|f\|_2$$

if f is supported in the w-neighborhood of γ.

Proof. We use the argument of P. Tomas [44], and will omit some of the details. Let σ be arclength on γ. The assumptions on γ are sufficient so that we can apply stationary phase to estimate $\hat{\sigma}$; we omit details, just noting that the assumption $\tau \leq k^{-1}$ implies a bound on the number of relevant critical points. We conclude that

$$|\hat{\sigma}(x)| \lesssim \min(\tau, (k|x|)^{-\frac{1}{2}}) \tag{4.22}$$

Let ϕ_j be an appropriate partition of unity such that $\text{supp}\phi_j \subset \{x \in \mathbb{R}^2 : |x| \in (\frac{1}{2}\cdot 2^j, 2\cdot 2^j)\}$ for $j \geq 1$ and ϕ_0 has compact support, and with the natural derivative bounds. If we cut up $\hat{\sigma} = \sum_j K_j$, $K_j = \phi_j\hat{\sigma}$ using this partition of unity then

$$\|K_j * u\|_\infty \lesssim \min(\tau, (k2^j)^{-\frac{1}{2}})\|u\|_1$$

by (4.22) and also

$$\|K_j * u\|_2 \lesssim 2^j\|u\|_2$$

Hence

$$\|K_j * u\|_q \lesssim 2^{\frac{2}{q}j}\min(\tau, (k2^j)^{-\frac{1}{2}})^{1-\frac{2}{q}}\|u\|_{q'}$$

if $q > 2$. For $q > 6$ this is largest when j makes $\tau = (k2^j)^{-\frac{1}{2}}$. We therefore obtain

$$\|\hat{\sigma} * u\|_q \lesssim \tau^{1-\frac{6}{q}}k^{-\frac{2}{q}}\|u\|_{q'}$$

whence the usual P. Tomas TT^* argument shows that

$$\|\widehat{fd\sigma}\|_q \lesssim \tau^{\frac{1}{2}-\frac{3}{q}}k^{-\frac{1}{q}}\|f\|_{L^2(\gamma)}$$

for functions f supported on γ. The factor involving τ is less than 1. Considering the w neighborhood as a union of translates of γ we obtain the lemma. $\square$

Lemma 4.11. *Suppose that u is supported on a type I or type II R_i. Then*

$$\|\hat{u}\|_{\ell^q} \lesssim \delta^{\frac{\alpha}{3}-\eta}\|u\|_2 \tag{4.23}$$

for any $q > 6$ and $\eta > 0$.

Proof. We first prove a corresponding bound on $\mathbb{R}^2$ instead of $\mathbb{T}^2$. Namely, lift the situation up to $\mathbb{R}^2$ by the exponential map; denote a component of the inverse image of the relevant R_i by $\tilde{R}_i$. Let f be a function on $\mathbb{R}^2$ whose support is contained in one of the $\tilde{R}_i$'s. Then, we show that

$$\|\hat{f}\|_q \lesssim C_\eta \delta^{-\eta}\delta^{\frac{\alpha}{3}}\|f\|_2 \tag{4.24}$$

if $q > 6$. Namely, if R_i is type I, then R_i is a rectangle and $|R_i| \lesssim \delta^\alpha$, so Bernstein's inequality gives

$$\|\hat{f}\|_q \lesssim \delta^{\alpha(\frac{1}{2}-\frac{1}{q})}\|f\|_2$$

for any $q > 2$. When $q > 6$ this is better than (4.24). If R_i is type II, then Lemma 4.10 gives

$$\|\hat{f}\|_q \lesssim \frac{w_i^{\frac{1}{2}}}{k_i^{\frac{1}{q}}}\|f\|_2.$$

By definition of w we have then

$$\|\hat{f}\|_q \lesssim \delta^{\frac{\alpha}{3}} k_i^{\frac{1}{6}-\frac{1}{q}}\|f\|_2$$

and this gives (4.24) since k_i is bounded by a power of δ and since it suffices to consider q slightly larger than 6. Thus (4.24) is proved.
To descend to $\mathbb{T}^2$, let u be as in the statement of the lemma and lift u up to $\mathbb{R}^2$ by the exponential map; call the resulting function v. Let ζ be a C_0^∞ function on $\mathbb{R}^2$ such that the $\mathbb{Z}^2$ translations of ζ^2 form a partition of unity. Then u is the periodization of $\zeta^2 v$:

$$\sum_{j\in\mathbb{Z}^2}(\zeta^2 v)(x-j) = u(x)$$

so that

$$\hat{u}(z) = \widehat{\zeta^2 v}(z) \tag{4.25}$$

if $z \in \mathbb{Z}^2$.

On the other hand, for any q

$$\begin{aligned}
\sum_{z\in\mathbb{Z}^2}|\widehat{\zeta^2 v}(z)|^q &= \sum_{z\in\mathbb{Z}^2}|\hat{\zeta} * \widehat{\zeta v}(z)|^q \\
&\lesssim \sum_{z\in\mathbb{Z}^2}\int|\hat{\zeta}(z-x)||\widehat{\zeta v}(x)|^q dx \\
&= \int|\widehat{\zeta v}(x)|^q(\sum_{z\in\mathbb{Z}^2}\int|\hat{\zeta}(z-x)|)dx \\
&\lesssim \int|\widehat{\zeta v}(x)|^q
\end{aligned}$$

by the rapid decay of $\hat{\zeta}$. We now apply (4.24) taking $f = \zeta v$. This shows that

$$\left(\sum_{z\in\mathbb{Z}^2} |\widehat{\zeta^2 v}(z)|^q\right)^{\frac{1}{q}} \lesssim \delta^{-\eta}\delta^{\frac{1}{3}\alpha}\|\zeta v\|_2.$$

Using (4.25) and that $\|\zeta v\|_2 \lesssim \|u\|_2$ we now conclude that $\|\hat{u}\|_{\ell^q} \lesssim \delta^{-\eta}\delta^{\frac{\alpha}{3}}\|u\|_2$ as claimed. $\square$

Proof of Lemma 3.6. Fix $\eta > 0$. We may assume that at least one of the R_i's is type I or II. If both are type I or II then we take R_1 to be the one with the larger "width" w_i. If just one is type I or II then we take R_1 to be this one, and notice that $w_1 \gtrsim w_2$, since the type III components R_i have $w_i \approx \delta^{\alpha-\kappa}$ and the other components have $w_i \gtrsim \delta^{\alpha-\kappa}$. Thus we may assume that R_1 is type I or II and that $w_1 \gtrsim w_2$. Let Z_i be the L^2 functions supported in the ϵ-neighborhood of R_i and let B_i b the unit ball in Z_i. Since $\epsilon \geq w_i$, the functions in Z_i are then also supported on the $C\epsilon$-neighborhood of a rectifiable curve. By Lemma 4.7 (and the lower bound on ρ), we have the following: for each $i \in \{1,2\}$ there is a collection of $\leq e^{-\epsilon^{-(1+\eta)}}$ functions $f_j \in B_i$ with $\|f_j\|_2 \leq 1$, such that if $f \in B_i$ then $\|\psi * (f - f_j)\|_2 \leq C^{-1}\rho$ for some j. Here C is a large constant to be specified later.

Consider a fixed choice of L^2 functions $u_1 \in Z_1$ and $u_2 \in Z_2$. The estimate in Lemma 4.11 shows that

$$\|\widehat{u_1}\|_{\ell^q} \lesssim \delta^{-\eta}\delta^{\frac{\alpha}{3}}\|u_1\|_2$$

if $q > 6$. Since ψ is supported in $D(0,\epsilon)$ we then also have

$$\|\widehat{\psi * u_1}\|_{\ell^q} \lesssim \delta^{-\eta}\delta^{\frac{\alpha}{3}}\|\psi * u_1\|_2.$$

Accordingly we can apply Lemma 4.9 with ϵ there equal to $\delta^{-\eta}\delta^{\frac{\alpha}{3}}$. We conclude that

$$\text{Prob}\left(\left|\sum_{n\in U}\omega_n\widehat{\psi * u_1}(n-k)\overline{\widehat{\psi * u_2}(n-\ell)}\right| \geq \rho\|\psi * u_1\|_2\|\psi * u_2\|_2\right) \leq Ce^{-C^{-1}\rho^{2+\frac{q}{q-2}}\delta^{(-\frac{\alpha}{3}+\eta)\frac{2q}{q-2}}}$$

a fortiori

$$\text{Prob}\left(\left|\sum_{n\in U}\omega_n\widehat{\psi * u_1}(n-k)\overline{\widehat{\psi * u_2}(n-\ell)}\right| \geq \rho\|u_1\|_2\|u_2\|_2\right) \leq Ce^{-C^{-1}\rho^{2+\frac{q}{q-2}}\delta^{(-\frac{\alpha}{3}+\eta)\frac{2q}{q-2}}}.$$

The exponent $\frac{\alpha}{3}\frac{2q}{q-2}$ converges to α as $q \to 6$, and the factor involving ρ can be absorbed into the $\delta^{-\eta}$ factor. Thus

$$\text{Prob}\left(\left|\sum_{n\in U}\omega_n\widehat{\psi * u_1}(n-k)\overline{\widehat{\psi * u_2}(n-\ell)}\right| \geq \rho\|u_1\|_2\|u_2\|_2\right) \leq Ce^{-C^{-1}\delta^{-\alpha+\eta}}$$

with $C = C_{\eta\rho}$. This is the requisite bound except that u_1 and u_2 were fixed.

If C has been chosen large enough and if $\|\psi * (u - u_*)\|_2 \leq C^{-1}\rho\|u\|_2$, $\|\psi * (v - v_*)\|_2 \leq C^{-1}\rho\|v\|_2$, and

$$\left|\sum_{n\in U}\omega_n\widehat{\psi * u}(n-k)\overline{\widehat{\psi * v}(n-\ell)}\right| \geq \rho\|u\|_2\|v\|_2$$

then also

$$\left|\sum_{n\in U}\omega_n\widehat{\psi * u_*}(n-k)\overline{\widehat{\psi * v_*}(n-\ell)}\right| \geq \frac{\rho}{2}\|u_*\|_2\|v_*\|_2$$

It follows using the entropy bound discussed at the beginning of the proof that

$$\operatorname{Prob}\left(\exists u_1 \in Z_1, u_2 \in Z_2, \left|\sum_{n\in U} \omega_n \widehat{\psi * u_1}(n-k)\overline{\widehat{\psi * u_2}(n-\ell)}\right| \geq \rho\|u_1\|_2\|u_2\|_2\right)$$
$$\leq e^{C\epsilon^{-1-\eta} - C^{-1}\delta^{-\alpha+\eta}}$$

for any $\eta > 0$. We have $\epsilon \gtrsim \delta^{\alpha-\kappa}$. Taking η small compared with κ it follows that the first term in the exponential can be dropped, and that the bound is of order $e^{-C^{-1}\delta^{-\alpha+\eta}}$ for any $\eta > 0$. This concludes the proof. □

5. Proofs of Lemmas 3.1 and 3.2

As has already been mentioned Lemma 3.2 is easy; it will be proved in subsection 5.2 where we discuss properties of the Newton diagram. Lemma 3.1 will follow from an analysis of the zero set of f using Puiseux expansions, in a similar spirit to [35], although our argument is technically rather different and is easier.

We note to begin with that it suffices to prove the following local version of 3.1: $(*)$ Let f be a real analytic function on a neighborhood of the origin in $\mathbb{R}^2$ and let $\alpha = \alpha(f, 0)$. Then, for some fixed neighborhood N of the origin, and for all small δ, $\Omega(\delta) \cap N$ has a covering as described by Lemma 3.1. In fact, passing from this statement to Lemma 3.1 is a trivial compactness argument. The rest of the section is the proof of the localized version $(*)$ and of Lemma 3.2.

5.1. **Preliminaries on polynomials.**

Let $p(x) = \Pi_i(x - x_i)$ be a polynomial in one variable. For a given i and δ we define $w = w_i(\delta)$ by the relation

$$\Pi_{j=1}^n \max(w, |x_j - x_i|) = \delta \tag{5.1}$$

Lemma 5.1. *$w_i(\delta)$ is uniquely defined and is a strictly increasing function of δ and furthermore has the following doubling type property: if $\delta' > \delta$ then*

$$w(\delta') \leq \frac{\delta'}{\delta} w(\delta) \tag{5.2}$$

$$w(\delta') \geq \left(\frac{\delta'}{\delta}\right)^{\frac{1}{n}} w(\delta). \tag{5.3}$$

Proof. At least one factor in (5.1) is equal to w, namely the factor with $j = i$. This implies that w_i is well defined, and clearly it is strictly increasing. For (5.2), we use

$$\Pi_{j=1}^n \max\left(\frac{\delta'}{\delta} w(\delta), |x_j - x_i|\right) \geq \frac{\delta'}{\delta} \Pi_{j=1}^n \max(w(\delta), |x_j - x_i|) = \delta'$$

by consideration of the $j = i$ factor. So $w(\delta') \leq \frac{\delta'}{\delta} w(\delta)$ as claimed.

Now we prove (5.3):

$$\begin{aligned} \Pi_{j=1}^n \max(w(\delta'), |x_j - x_i|) &\leq \left(\frac{w(\delta')}{w(\delta)}\right)^n \Pi_{j=1}^n \max(w(\delta), |x_j - x_i|) \\ &\leq \left(\frac{w(\delta')}{w(\delta)}\right)^n \delta \end{aligned}$$

which shows that

$$\delta \left(\frac{w(\delta')}{w(\delta)}\right)^n \geq \delta'.$$

This proves (5.3). □

Lemma 5.2. *Let $p(z) = \Pi_{i=1}^n (z - x_i)$ be a monic polynomial. Define w_i via (5.1). Then the set $\{z \in \mathbb{C} : |p(z)| < \delta\}$ is contained in $\cup_i D(x_i, Cw_i)$.*

Proof. Fix a number $C_0 = 4^n$ and subdivide the region $D(x_i, C_0 w_i) \backslash D(x_i, w_i)$ in annuli $4^m w_i \leq |z - x_i| < 4^{m+1} w_i$, $0 \leq m \leq n-1$. One of them must be free of points x_j. Take this one and consider its center circle $c_i = \{|z - x_i| = 2 \cdot 4^{m_i} w_i\}$. If $z \in c_i$ then $|z - x_j| \geq w_i$ for all j. Further, if $|x_j - x_i| \geq 4^{m_i+1} w_i$ then $|z - x_j| \geq |x_j - x_i| - |z - x_i| \geq |x_j - x_i| - \frac{1}{2}|x_j - x_i|$. We conclude that $|p(z)| \gtrsim \Pi_j \max(w_i, |x_j - x_i|) \geq \delta$. Thus $|p| \gtrsim \delta$ on the union of the c_i's. Now consider Ω, the complement of the union of the interiors of the c_i's. This is free of zeroes of p, and its boundary is contained in the union of the c_i's. By the maximum principle $|p| \gtrsim \delta$ on Ω and therefore on the complement of $\cup_i D(x_i, C_0 w_i)$. The lemma follows by replacing δ with a large constant times δ and using the doubling property (5.2). □

5.2. **Preliminaries concerning the Newton diagram.** In this section we assemble some facts about the Newton diagram which are presumably fairly well known, but for which we do not have a convenient reference. In particular we prove Lemma 3.2.
Let $p(x, y) = y^n + \sum_{j=0}^{n-1} a_j(x) y^j$ be a real analytic function which is a Weierstrass polynomial with respect to the y variable. Thus for a certain N there is a factorization

$$p(x, y) = \Pi_{i=1}^n (y - \zeta_i(x))$$

where the "roots" $\zeta_i(x)$ are analytic functions of $x^{\frac{1}{N}}$. It will clearly suffice to prove Lemma 3.1 (more precisely the localized version $(*)$) in the region $x > 0$, and we may therefore always take $x^{\frac{1}{N}}$ to be the principal branch; in particular it is a single valued function, in the region $\{x : re(x) > 0\}$.

We define exponents ν_i and ν_{ij} as follows:

$$\zeta_i(x) = c_i x^{\nu_i} + \text{ higher order,}$$
$$\zeta_i(x) - \zeta_j(x) = c_{ij} x^{\nu_{ij}} + \text{ higher order.}$$

If ζ_i or $\zeta_i - \zeta_j$ vanishes identically we define $\nu_i = \infty$ or $\nu_{ij} = \infty$ respectively.
Thus ν_i and ν_{ij} are (either equal to infinity or) multiples of $\frac{1}{N}$. In addition the Weierstrass polynomial property implies that $\nu_i \geq 1$, since otherwise f would contain a term of total degree less than n.
For each i and j we must have either

$$\nu_i = \nu_j < \nu_{ij} \tag{5.4}$$

or

$$\nu_{ij} = \min(\nu_i, \nu_j) \tag{5.5}$$

In the first case we say that ζ_i and ζ_j are *tangent.*[1]

Lemma 5.3. *For each i with $\nu_i < \infty$,*

$$\sum_j \min(\nu_i, \nu_j) \leq \alpha^{-1}(1 + \nu_i).$$

Furthermore, if the zero root has multiplicity m then $m \leq \alpha^{-1}$.

Proof. A statement equivalent to this is in [35]; see Observation 3 on p. 140 and the preceding discussion. We give the argument for the reader's convenience.

[1]This terminology is convenient, even though it makes a zero root ($\nu_i = \infty$) not be tangent to itself.

An equivalent formulation is as follows: if the roots are ordered so that ν_j is nonincreasing, then for each $k > m$

$$\frac{1+\nu_k}{\sum_{i>k}\nu_i + k\nu_k} \geq \alpha$$

and furthermore $m \leq \alpha^{-1}$.
Consider the points

$$p_k = \left(\sum_{i>k}\nu_i, k\right), \quad k \in \{m, \dots, n\}$$

and the subset $E \subset \{m, \dots, n\}$ consisting of values of k which are either equal to m or n or are such that $\nu_{k+1} < \nu_k$. For $k \in \{m+1, \dots, n\}$ the segment (p_{k-1}, p_k) is a portion of the line

$$\ell_k = \left(\sum_{i>k}\nu_i, k\right) + t(-\nu_k, 1).$$

This implies in the first place that if $\nu_{k+1} = \nu_k$ then the lines ℓ_{k+1} and ℓ_k coincide. It follows readily that $\{p_k\}$ and $\{p_k\}_{k\in E}$ have the same exterior convex hull. In the second place the lines ℓ_k all have negative slopes and the absolute value of the slope is a nondecreasing function of k. It follows that the boundary of the exterior convex hull of the $\{p_k\}$ consists of the segments (p_{k-1}, p_k), $k \in \{m+1, \dots, n\}$ together with the segments $\{0\} \times [n, \infty)$ and $[\sum_{i>m}\nu_i, \infty) \times \{m\}$.
If we expand out the product $\Pi_j(y - \zeta_j(x))$ we see that

1. $p(x, y)$ does not contain any nonzero term of the form $cy^k x^\ell$ with $\ell < \sum_{j>k}\nu_j$.
2. If $k \in E$ then $p(x, y)$ contains a term of the form $cy^k x^{\sum_{j>k}\nu_j}$ with $c \neq 0$.

Statement (2) implies that the $\{p_i\}_{i\in E}$, and therefore also the exterior convex hull of the full set $\{p_i\}$, are contained in the Newton polygon. In fact this exterior convex hull is equal to the Newton polygon. Namely, the Newton polygon is the exterior convex hull of certain lattice points (ℓ, k) with $m \leq k \leq n$. If one of these lattice points were outside the exterior convex hull of the $\{p_i\}$ then necessarily we would have $\ell < \sum_{j>k}\nu_j$, which is impossible.
We may therefore conclude that the lines ℓ_k and the line $(t, m), t \in \mathbb{R}$ are boundary lines of the Newton polygon.
The intersection of ℓ_k with the diagonal is the point (ξ, ξ) where

$$k + t = \xi, \quad \sum_{i>k}\nu_i - t\nu_k = \xi.$$

Hence

$$\xi^{-1} = \frac{1+\nu_k}{\sum_{i>k}\nu_i + k\nu_k}.$$

Furthermore the line $(t, m), t \in \mathbb{R}$ intersects the diagonal at the point (m, m). The lemma follows from these statements and the definition of α. $\square$

Lemma 5.4. *(i) For any i and k with ζ_i not tangent to ζ_k and* $\min(\nu_i, \nu_k) < \infty$

$$\sum_j \min(\nu_{ij}, \nu_{ik}) \leq \alpha^{-1}(1 + \nu_{ik}).$$

(ii) For any i with $\nu_i < \infty$ we have

$$\sum_j \min(\nu_{ij}, \nu_i) \leq \alpha^{-1}(1 + \nu_i).$$

Proof. This will follow from Lemma 5.3 using the definition of tangency (5.4), (5.5). Namely, for (ii) note that $\min(\nu_{ij}, \nu_i) = \min(\nu_j, \nu_i)$, since if ζ_i and ζ_j are tangent then both sides are equal to ν_i.
For (i), note that by (5.4), (5.5)

$$\min(\nu_{ij}, \nu_{ik}) = \min(\nu_{ij}, \nu_i, \nu_k) \leq \min(\nu_j, \nu_i, \nu_k)$$

Substituting this into Lemma 5.3 gives that

$$\sum_j \min(\nu_{ij}, \nu_{ik}) \leq \alpha^{-1}(1 + \min(\nu_i, \nu_k))$$

hence the result by (5.5). $\square$

We fix a particular choice of a root, and from now on we order the roots ζ_j so that our particular root is the first one and ν_{1j} is nonincreasing (this ordering is of course different from the one in the proof of Lemma 5.3). If $\nu_1 < \infty$, then there is a particular value of k_* so that ζ_j is tangent to ζ_1 for $j \leq k_*$ and not for $j > k_*$. Also there is a value of $k_{**} \geq k_*$ such that $\nu_{1j} \geq \nu_1$ for $j \leq k_{**}$ and not for $j > k_{**}$. We have the following:

Lemma 5.5. *If $\nu_1 < \infty$ and $\nu_1 \notin \mathbb{Z}$ then $k_* \leq k_{**} - k_*$.*

Proof. We can express

$$p(z^N, y) = \Pi_j(y - a_j(z))$$

where $a_j(z) = \zeta_j(z^N)$ is an analytic function of z. The assumption that $\nu_1 \notin \mathbb{Z}$ implies that the leading term in the expansion of $a_1(z)$ at the origin is cz^k with k not divisible by N. Likewise if ζ_i is tangent to ζ_1 then the leading term in the expansion of a_i is cz^k.
Let ω be a primitive Nth root of unity. Clearly $p(z^N, y)$ is invariant under the change of variables $(z, y) \to (\omega z, y)$. Suppose $\zeta_i(z)$ is a root which is tangent to ζ_1. Then $a_i(\omega z)$ is again of the form $\zeta_{\ell(i)}(z^N)$, where $\zeta_{\ell(i)}$ is a root with the same leading exponent but a different leading coefficient. This gives us a one to one map from roots tangent to a given one to roots which are not tangent and have the same leading exponent, which proves the lemma. $\square$

Proof of Lemma 3.2. It suffices to prove the corresponding statement in the localized setting $(*)$, i.e, if $\alpha = \alpha(f, 0)$ then there is a rectangle contained in $\Omega(\delta)$ with volume about δ^α.
Choose $T \in SL(2, \mathbb{R})$ so that $\alpha = \tilde{\alpha}(f \circ T, 0)$. Form the Newton polygon for $f \circ T$ and let ℓ be a boundary line intersecting the diagonal at $(\alpha^{-1}, \alpha^{-1})$. Then ℓ is either the horizontal line $\{(t, m) : t \in \mathbb{R}\}$, where m is the multiplicity of the root $y = 0$, or else ℓ has negative slope. Assume at first that ℓ is not horizontal. Let $(0, v)$ and $(u, 0)$ be the intersection points between ℓ and the coordinate axes. We can express α in terms of u and v by

$$\alpha^{-1} = \frac{uv}{u + v} \tag{5.6}$$

If (x, y) is a point of $\mathbb{R}^2$ then let $(X, Y) = (\log \frac{1}{|x|}, \log \frac{1}{|y|})$. If (X, Y) is such that $uX \geq \log(\frac{1}{\delta})$ and $vY \geq \log(\frac{1}{\delta})$, then $(m, n) \cdot (X, Y) \geq \log(\frac{1}{\delta})$ if (m, n) lies on or above the line ℓ. The Newton polygon of $f \circ T$ lies above ℓ, so it follows from the definitions that $|x^m y^n| \leq \delta$ for all (m, n) which appear in the Taylor expansion of $f \circ T$, so $|f \circ T(x, y)| \lesssim \delta$.
We conclude that the set $\{(x, y) : |f \circ T(x, y)| \leq C\delta\}$ contains a rectangle with dimensions $\delta^{\frac{1}{u}} \times \delta^{\frac{1}{v}}$, hence volume $\delta^{\frac{u+v}{uv}}$. This is equal to δ^α by (5.6), so we're done.
On the other hand suppose $\ell = \{(t, m) : t \in \mathbb{R}\}$. Then $\alpha = m^{-1}$ and on the other hand $f \circ T$ is divisible by y^m and is therefore $\leq \delta$ when $|y| \leq \delta^{\frac{1}{m}}$, i.e. on a rectangle with dimensions roughly $1 \times \delta^\alpha$. The proof is complete. $\square$

5.3. **Proof of Lemma 3.1.** Making a preliminary $SL(2,\mathbb{R})$ change of coordinates, we can assume that f is normalized with respect to the y variable and thus is pu where p is a Weierstrass polynomial with respect to the y variable and u is a unit. Factoring out u does not affect the Newton polygon and affects $|f|$ only by a bounded factor, so we can assume that f itself is a Weierstrass polynomial. Now factor into roots,

$$f(x,y) = \Pi_j (y - \zeta_j(x)).$$

We have $\nu_j \geq 1$ for each j by the Weierstrass polynomial property.
For each i define $w = w(i,x)$ by

$$\Pi_{j=1}^n \max(w, |x|^{\nu_{ij}}) = \delta \tag{5.7}$$

Then the set $\{y : |p(x,y)| < \delta\}$ is contained in $\cup_i \{y : |y - \zeta_i(x)| \leq Cw(i,x)\}$. This follows from Lemma 5.2 and the doubling property in Lemma 5.1, since $|\zeta_i(x) - \zeta_j(x)| \approx |x|^{\nu_{ij}}$.
We now fix a root, which we take to be the first one, and order the roots by nonincreasing ν_{1j}, and use the other notation in subsection 5.2. We denote $w(\tau) = w(1,\tau)$. We can make a further $SL(2,\mathbb{R})$ change of coordinates so that the following is true.

f is a Weierstrass polynomial with respect to y, and furthermore either ζ_1 has complex leading coefficient, or else $\nu_1 > 1$. (5.8)

To see this, suppose that initially $\zeta_1(x) = cx+$higher order terms, with c real. Define

$$T = \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix}.$$

Then

$$f \circ T = \Pi_j (y + cx - \zeta_j(x))$$

which is still a Weierstrass polynomial and has $\nu_1 > 1$. We may therefore assume that f itself satisfies (5.8), and we make this assumption for the remainder of the section.

Lemma 5.6. *The parameter w obeys the following estimates when $\tau \leq 1$.*

(i) $w(\tau) \lesssim \delta^{\frac{\alpha}{2}}$ for all τ.

(ii) If $\tau \lesssim \delta^{\frac{\alpha}{\nu_1+1}}$ then $w(\tau) \lesssim \frac{\delta^\alpha}{\tau}$.

(iii) If $\nu_1 > 1$ and if $\tau >> \delta^{\frac{\alpha}{\nu_1+1}}$ then $w(\tau) \lesssim (\delta^\alpha \tau^{\frac{\nu_1-2}{2}})^{\frac{2}{3}}$.

Proof. Let m be the multiplicity of the root $y = 0$.
(i) Consider the product

$$\Pi_{j=1}^n \max(\delta^{\frac{\alpha}{2}}, \tau^{\nu_{1j}}) \tag{5.9}$$

To prove (i) it suffices by Lemma 5.1 to show that (5.9) $\gtrsim \delta$. For this, estimate each factor in (5.9) from below by $\delta^{\frac{\alpha}{2}}$; thus

$$(5.9) \gtrsim \delta^{\frac{n\alpha}{2}}.$$

If $m = n$ then $f(x,y) = cy^n$, $\alpha = \frac{1}{n}$ and it follows that (5.9) $\gtrsim \delta^{\frac{1}{2}} \geq \delta$. If $m < n$ the let ζ_i be a root such that ν_i is as small as possible. Lemma 5.3 then shows that $\alpha \leq \frac{1+\nu_i}{n\nu_i}$. But $\nu_i \geq 1$ since p is a Weierstrass polynomial with respect to y. Thus $\alpha \leq \frac{2}{n}$ and (5.9) $\gtrsim \delta$.

(ii) For clarity we first consider just the case $\nu_1 = \infty$, i.e. the real flat case, although the proof for this case could probably be incorporated into the one for (5.11) below. Then we need to show that $w(\tau) \lesssim \frac{\delta^\alpha}{\tau}$ for all τ. When $\nu_1 = \infty$ we have $\nu_{1j} = \nu_j$, so similarly to (i) we need to show that

$$\Pi_j \max(\frac{\delta^\alpha}{\tau}, \tau^{\nu_j}) \gtrsim \delta \tag{5.10}$$

where we interpret τ^{ν_j} as zero if $\nu_j = \infty$.

Subdivide into the intervals $\tau < \delta^{\frac{\alpha}{\nu_n+1}}$, $\tau \in [\delta^{\frac{\alpha}{\nu_{k+1}+1}}, \delta^{\frac{\alpha}{\nu_k+1}}]$ for $m+1 \leq k \leq n-1$, and $\tau \geq \delta^{\frac{\alpha}{\nu_{m+1}+1}}$. On each of these intervals the left side of (5.10) is proportional to a power of τ and therefore monotone. Since we are working locally it therefore suffices to check (5.10) at the points $\tau = \delta^{\frac{\alpha}{\nu_k+1}}$, $k \geq m+1$, and at $\tau = 1$. If $\tau = \delta^{\frac{\alpha}{\nu_k+1}}$, $k \geq m+1$ then the left side of (5.10) is

$$\Pi_j \max(\delta^{\frac{\alpha\nu_k}{\nu_k+1}}, \delta^{\frac{\alpha\nu_j}{\nu_k+1}})$$

where the second member is again interpreted as zero if $j \leq m$. This is the same as $\delta^{\frac{\alpha \sum_j \min(\nu_j, \nu_k)}{\nu_k+1}}$ so (5.10) follows from Lemma 5.3. On the other hand if $\tau = 1$ then the left side of (5.10) is $\delta^{m\alpha}$, so (5.10) follows from the last statement in Lemma 5.3.

We note next that if $\nu_1 = 1$ then there is nothing to prove; (ii) follows from (i). Now consider the case $\nu_1 \in (1, \infty)$. It suffices as above to show that the product

$$\Pi_{j=1}^n \max(\frac{\delta^\alpha}{\tau}, \tau^{\nu_{1j}}) \tag{5.11}$$

is $\gtrsim \delta$ when $\tau \lesssim \delta^{\frac{\alpha}{\nu_1+1}}$. For this, subdivide into $\tau < \delta^{\frac{\alpha}{\nu_{1\,n}+1}}$, $\tau \in [\delta^{\frac{\alpha}{\nu_{1\,k+1}+1}}, \delta^{\frac{\alpha}{\nu_{1\,k}+1}}]$ for $k > k_*$ and $\tau \in [\delta^{\frac{\alpha}{\nu_{1\,k_*+1}+1}}, \delta^{\frac{\alpha}{\nu_1+1}})$. The latter interval is free of points $\delta^{\frac{\alpha}{\nu_{1\,k}+1}}$, since $\nu_1 \in [\nu_{1\,k_*+1}, \nu_{1k_*}]$. It therefore suffices as with the proof of (5.10) above to check (5.11) at the points $\tau = \delta^{\frac{\alpha}{\nu_{1k}+1}}$ with $k > k_*$ and $\tau = \delta^{\frac{\alpha}{\nu_1+1}}$. If $\tau = \delta^{\frac{\alpha}{\nu_{1k}+1}}$ then

$$(5.11) \quad = \quad \Pi_{j=1}^n \max(\delta^{\alpha\frac{\nu_{1k}}{\nu_{1k}+1}}, \delta^{\alpha\frac{\nu_{1j}}{\nu_{1k}+1}})$$

and this is $\gtrsim \delta$ by Lemma 5.4(i). Likewise if $\tau = \delta^{\frac{\alpha}{\nu_1+1}}$ then

$$(5.11) \quad = \quad \Pi_{j=1}^n \max(\delta^{\alpha\frac{\nu_1}{\nu_1+1}}, \delta^{\alpha\frac{\nu_{1j}}{\nu_1+1}})$$

and this is $\gtrsim \delta$ by Lemma 5.4(ii). This finishes the proof of (ii).

(iii) This is vacuous if $\nu_1 = \infty$, so we may assume $1 < \nu_1 < \infty$. We need to show that the product

$$\Pi_{j=1}^n \max((\delta^\alpha \tau^{\frac{\nu_1-2}{2}})^{\frac{2}{3}}, \tau^{\nu_{1j}}) \tag{5.12}$$

is $\gtrsim \delta$ if $\tau \gtrsim \delta^{\frac{\alpha}{\nu_1+1}}$. We estimate the jth factor in (5.12) below by $(\delta^\alpha \tau^{\frac{\nu_1-2}{2}})^{\frac{2}{3}}$ if ζ_j is tangent to ζ_1 and by $\tau^{\nu_{1j}}$ otherwise. Thus

$$(5.12) \gtrsim \delta^{\frac{2}{3}\alpha k_*} \tau^{k_* \frac{\nu_1-2}{3} + \sum_{j>k_*} \nu_{1j}} \tag{5.13}$$

We claim that the exponent of τ in (5.13) is positive. This is obvious if $\nu_1 \geq 2$. If $\nu_1 < 2$ it follows from Lemma 5.5: $\nu_1 \in (1,2)$ is a noninteger so Lemma 5.5 implies that

$$k_* \frac{\nu_1 - 2}{3} + \sum_{j>k_*} \nu_{1j} \geq k_* \frac{\nu_1 - 2}{3} + k_* \nu_1$$

and the right side is clearly positive.

Accordingly (5.13) is increasing in τ so we need only check that it is $\gtrsim \delta$ when $\tau = \delta^{\frac{\alpha}{\nu_1+1}}$. When

$\tau = \delta^{\frac{\alpha}{\nu_1+1}}$ we have

$$
\begin{aligned}
(5.13) \quad &= \delta^{\frac{2}{3}\alpha k_* + \frac{\alpha}{\nu_1+1}k_*\frac{\nu_1-2}{3}} \cdot \delta^{\frac{\alpha}{\nu_1+1}\sum_{j>k_*}\nu_{1j}} \\
&= \delta^{\frac{\alpha k_* \nu_1}{\nu_1+1}} \cdot \delta^{\frac{\alpha}{\nu_1+1}\sum_{j>k_*}\nu_{1j}} \\
&= \delta^{\frac{\alpha}{\nu_1+1}(k_*\nu_1 + \sum_{j>k_*}\nu_{1j})} \\
&\gtrsim \delta
\end{aligned}
$$

where the last line follows from Lemma 5.4(ii). □

We need one more lemma. Let A_0 be a suitable constant. For given $x \in (0,1]$ let $e(x)$ be a complex unit tangent vector to the graph of ζ_1 in $\mathbb{C}^2$ at the point $(x, \zeta_1(x))$, let $n(x)$ be a complex unit vector normal to $e(x)$, and let $R_0(x)$ be the rectangle

$$\left\{ z \in \mathbb{C}^2 \,|\, |\langle z - (x, \zeta_1(x)), e(x)\rangle| \le A_0 x, |\langle z - (x, \zeta_1(x)), n(x)\rangle| \le A_0 \frac{\delta^\alpha}{x} \right\}.$$

Define $R_1(x)$ like $R_0(x)$ replacing A_0 by a suitable larger constant A_1.

Lemma 5.7. *Assume that $\nu_1 < \infty$ and that τ and x satisfy $\delta^{\frac{\alpha}{2}} \le \tau \le C\delta^{\frac{\alpha}{\nu_1+1}}$, $\frac{\tau}{2} \le x \le 2\tau$. Then $R_0(x) \subset R_1(\tau)$.*

Proof. Note that the "lengths" ($A_0 x$ and $A_0 \tau$) of $R_0(x)$ and $R_0(\tau)$ are comparable, as are their "widths" ($A_0 \frac{\delta^\alpha}{x}$ and $A_0 \frac{\delta^\alpha}{\tau}$). It follows that $R_0(x)$ will be contained in a suitable fixed dilation $R_1(\tau)$ provided that (i) and (ii) below hold.

(i) $(x, \zeta_1(x)) \in R_0(\tau)$.

(ii) $|e(x) - e(\tau)| \lesssim \frac{\delta^\alpha/\tau}{\tau}$.

To verify (ii) observe that $e(x)$ is a normalization of the vector $(1, \zeta_1'(x))$ and that $|\zeta_1'(x) - \zeta_1'(\tau)| \lesssim |\zeta_1'(x)| + |\zeta_1'(\tau)| \lesssim \tau^{\nu_1 - 1}$. Thus

$$
\begin{aligned}
|e(x) - e(\tau)| &\lesssim \tau^{\nu_1-1} \\
&\lesssim \frac{\delta^\alpha/\tau}{\tau}
\end{aligned}
$$

where the last inequality holds since $\tau \lesssim \delta^{\frac{\alpha}{\nu_1+1}}$. This proves (ii), and (i) can be proved in a similar way. □ We fix a sufficiently small positive number κ, and are now ready to define the rectangles R_i.

Ia. The square centered at the origin with side $C_2 \delta^{\frac{\alpha}{2}}$ is a type I R_i, and will be denoted R_0.

In addition, for each choice of root ζ_1 we do the following, after choosing affine coordinates so that (5.8) holds. If $\nu_1 \ne \infty$ then

Ib. For each dyadic $\tau \in [\delta^{\frac{\alpha}{2}}, C_1 \delta^{\frac{\alpha}{\nu_1+1}}]$ we look at the complex rectangle $R_1(\tau)$ in Lemma 5.7. If $R_1(\tau)$ intersects the real plane, then we take the intersection of the double of $R_1(\tau)$ with the real plane. This is contained in a rectangle with dimensions $C\tau \times C\frac{\delta^\alpha}{\tau}$, which we define to be a type I R_i. These R_i's satisfy the conditions given in 1. of Lemma 3.1, provided that $\kappa < \frac{\alpha}{\nu_1+1}$.

II. If ζ_1 has real leading coefficient, then for each dyadic $\tau > C_1 \delta^{\frac{\alpha}{\nu_1+1}}$ we take the curvilinear rectangle

$$\left\{ (x,y) : \frac{\tau}{2} \le x \le 2\tau \text{ and } |y - re\zeta_1(x)| \le C(\delta^\alpha \tau^{\frac{\nu_1-2}{2}})^{\frac{2}{3}} \right\}.$$

These are type II R_i's. Notice that the curve $\{(x,y) : y = re\zeta_1(x)\}$ has curvature about τ^{ν_1-2} when $x \approx \tau$, and that the jth derivative of curvature is $\lesssim \tau^{\nu_1-2-j}$, and therefore behaves as required in 2. of Lemma 3.1. To check the remaining conditions stated there, we must show that

$$\tau^{\nu_1-2} \lesssim \tau^{-1} \text{ (i.e. } k \lesssim \tau^{-1}) \tag{5.14}$$

$$(\delta^\alpha \tau^{\frac{\nu_1-2}{2}})^{\frac{2}{3}} \lesssim \tau \text{ (i.e. } w \lesssim \tau) \tag{5.15}$$

$$(\delta^\alpha \tau^{\frac{\nu_1-2}{2}})^{\frac{2}{3}} \geq \delta^{\alpha-\kappa}. \tag{5.16}$$

Estimate (5.14) follows immediately since $\nu_1 \geq 1$. For (5.15) we estimate

$$\delta^\alpha \leq \tau^{\nu_1+1} \leq \tau^{\frac{5-\nu_1}{2}}$$

using $\nu_1 \geq 1$. This implies (5.15) by algebraic manipulation. Finally (5.16) is equivalent with

$$\tau^{\nu_1-2} \geq \delta^{\alpha-3\kappa}.$$

The latter bound is obvious for small κ if $\nu_1 \leq 2$, and if $\nu_1 > 2$ it follows (for $\kappa < \frac{\alpha}{\nu_1+1}$) since

$$\tau^{\nu_1-2} \geq \delta^{\alpha\frac{\nu_1-2}{\nu_1+1}}.$$

Thus the R_i's just defined have all the properties in 2. of Lemma 3.1.

If $\nu_1 = \infty$, then instead

Ic. For each dyadic $\tau \in [\delta^{\frac{\alpha}{2}}, C_1\delta^\kappa]$ we take the rectangle with length $C\tau$ in the x direction and width $C\frac{\delta^\alpha}{\tau}$ in the y direction. These are type I R_i's, and clearly have the properties in 1. of Lemma 3.1.

III. We also take the rectangle with length 1 in the x direction and width $C\delta^{\alpha-\kappa}$ in the y direction. This is a type III R_i, and clearly has the property in 3. of Lemma 3.1.

Remark 5.1. We note here that, if $\nu_1 < \infty$ and ζ_1 has a real leading coefficient, then $\Omega(\delta)$ contains a set of area at least $\delta^{\frac{\alpha(\nu_1+1)}{\nu_1+2}} >> \delta^\alpha$.

To complete the proof of Lemma 3.1 it now suffices to prove

Lemma 5.8. *T he R_i's cover the intersection of $\Omega(A\delta)$ with a fixed neighborhood of the origin.*

Proof. If $x \leq \delta^{\frac{\alpha}{2}}$ and if $(x,y) \in \Omega(A\delta)$ then Lemma 5.6(i) and the Weierstrass polynomial property imply that $(x,y) \in R_0$.
Next we fix a choice of ζ_1 and a dyadic value of $\tau \geq \delta^{\frac{\alpha}{2}}$. Consider the set

$$\{(x,y) \in \mathbb{R} \times \mathbb{R} : |x-\tau| \leq \frac{\tau}{2} \text{ and } |y - \zeta_1(x)| \leq w(x)\} \tag{5.17}$$

We will show this set is covered by the R_i's associated to ζ_1. In the case $\nu_1 = \infty$ this is clear from the definition and Lemma 5.7. We now consider the case $\nu_1 < \infty$.
Assume first that $\tau \lesssim \delta^{\frac{\alpha}{\nu_1+1}}$. Then (if the constants A_0 and A_1 were chosen appropriately) the set (5.17) is contained in $R_1(\tau)$ by Lemmas 5.6(ii) and 5.7 and is therefore contained in the type I rectangle R_i associated to ζ_1 and τ.

We now assume instead that $\tau >> \delta^{\frac{\alpha}{\nu_1+1}}$ and will show that the set (5.17) is contained in a type II R_i.

Suppose first that ζ_1 has complex leading coefficient. Then, if $|x-\tau| \leq \frac{\tau}{2}$, we have

$$\begin{aligned} \mathrm{dist}(z_1(x), \mathbb{R}) &\approx x^{\nu_1} \\ &>> w(x) \end{aligned}$$

The last inequality is seen as follows: if $x \approx \delta^{\frac{\alpha}{\nu_1+1}}$ then Lemma 5.6(ii) implies that $w(x) \lesssim x^{\nu_1}$. Now use that $w(x)$ is decreasing in x by the definition (5.7).
So the set (5.17) is empty if ζ_1 has complex leading coefficient. On the other hand suppose ζ_1 has real leading coefficient. Then we are assuming that $\nu_1 > 1$. Clearly (5.17) is contained in $\{(x,y) : |x-\tau| \le \frac{\tau}{2}, |y - re\zeta_1(x)| \le w(x)\}$, and this is contained in a type II R_i by Lemma 5.6(iii). This completes the proof of Lemma 5.8 and therefore of Lemma 3.1. □

6. The proof of the main results

6.1. **The proof of Theorem 0.2.** Let $h_\omega = h + G_\omega$ be the discrete random operator defined in Theorem 1.2. By Theorem 1.1 and (1.27), we obtain that, if N^* denotes the density of states of the random operator h_ω then, under the assumptions made in section 1, there exists $C > 0$ and $E_* > 0$ such that, for $0 \le E \le E_*$, we have

$$N^*(E/C) \le N(E) - N(0) \le N^*(C \cdot E). \tag{6.1}$$

Pick $0 < E$ small. Recall the Lifshitz exponent α defined in section 0.2. Fix $\varepsilon > 0$ and pick $L(E) = e^{-E^{-\alpha+\varepsilon}}$. Then, for E sufficiently small, equation (2.3) in Theorem 2.1 and the result of Theorem 3.1 imply

$$N^*(E) \le e^{-(2E)^{-\alpha+\varepsilon}} + CL(E)^{-(1-\varepsilon)}.$$

This and (6.1) imply

$$\limsup_{E\to 0} \frac{\log|\log|N(E) - N(0)|}{\log E} \le -\alpha. \tag{6.2}$$

On the other hand, if we pick $l(E) = E^{-\alpha+\varepsilon}$ and use equation (2.4) in Theorem 2.1 and the result of Theorem 3.1, we get

$$N^*(E) \ge CE^{2\alpha-2\varepsilon} e^{-(E/2)^{-\alpha-\varepsilon}}.$$

This and (6.1) imply

$$\liminf_{E\to 0} \frac{\log|\log|N(E) - N(0)|}{\log E} \ge -\alpha. \tag{6.3}$$

Equations (6.2) and (6.3) imply Theorem 0.2.

6.2. **The proof of Theorem 0.1.** We will first bring ourselves back to the case when V is compactly supported. For the sake of definiteness, assume $E_0 = 0$ is the upper edge of a gap of Σ and assume that the gap contains at least the interval $[-\delta, 0)$ ($\delta > 0$). Pick $R > 0$ large and let χ_R be the characteristic function of the ball of center 0 and radius R. Write $V = V^R + (V - V^R)$ where $V^R = V \cdot \chi_R$. Consider the operator $H^R_\omega = H + V^R_\omega$ (where V^R_ω is constructed in the same way as V_ω using V^R instead of V). Let N_R be the density of states of H^R_ω.
As the random variables $(\omega_\gamma)_{\gamma\in\Gamma}$ are non negative and as V is non negative, one has

$$N(E) \le N_R(E), \quad \forall E \in \mathbb{R}.$$

As the $(\omega_\gamma)_\gamma$ are i.i.d. and bounded, and using assumption (H.1.1), for $p \geq 1$, we compute

$$\begin{aligned}\sup_{(\omega_\gamma)_\gamma} \sup_{\gamma\in\Gamma} \int_{C_\gamma} |V_\omega - V_\omega^R|^p dx &\leq \sup_{(\omega_\gamma)_\gamma} \sup_{\gamma\in\Gamma} \int_{C_0} \left| \sum_{|\beta-\gamma|>R-C} \omega_\beta V(x-\beta+\gamma) \right|^p dx \\ &\leq K \int_{C_0} \left(\sum_{|\beta|>R-C} V(x-\beta) \right)^p dx \\ &\leq K_p \sum_{|\beta|>R-C} \int_{C_0} V(x-\beta) dx.\end{aligned}$$

By assumption (H.1.1), this last term tends to 0 as R tends to $+\infty$. Hence, when $R \to +\infty$, we get that $(1-\Delta)^{-1}(V_\omega - V_\omega^R) \to 0$ in operator norm uniformly in $(\omega_\gamma)_\gamma$ (see [37]). This implies that, for R large enough, for $E \in (-\delta/2, 0)$,

$$N(E) = N_R(E)$$

Hence, we obtain

$$N(E) - N(0) \leq N_R(E) - N_R(0).$$

In the case where V is compactly supported, we follow the same strategy as in the proof of Theorem 0.2 and use Theorem 3.1; to check assumption f1, it will be sufficient to prove

Proposition 6.1. *For any $\theta \in \mathcal{S}$, there exists $C_O > 1$, O a neighborhood of θ in $\mathbb{T}^2$ and $h_O : O \to \mathbb{R}$, a real analytic such that*

$$\frac{1}{C_O} h_O \cdot I_{p_0} \leq h \leq C_O h_O \cdot I_{p_0}. \tag{6.4}$$

Here I_{p_0} is the $p_0 \times p_0$ identity matrix.

Proof of Proposition 6.1. We will use

Lemma 6.1. *Let $\mathcal{U}$ and $\mathcal{T}$ be two unitary equivalences from $\Pi_0(L^2(\mathbb{R}^d))$ to $L^2(\mathbb{T}^2) \otimes \mathbb{C}^{p_0}$ satisfying $\forall \gamma \in \Gamma$, $\mathcal{U} \circ \tau_\gamma = e_\gamma \circ \mathcal{U}$ and $\mathcal{T} \circ \tau_\gamma = e_\gamma \circ \mathcal{T}$. Here e_γ denotes the multiplication by the function $\theta \mapsto e^{i\gamma\cdot\theta}$ acting on $L^2(\mathbb{T}^2) \otimes \mathbb{C}^{p_0}$.*

Then, $\mathcal{T}\mathcal{U}^* : L^2(\mathbb{T}^2) \otimes \mathbb{C}^{p_0} \to L^2(\mathbb{T}^2) \otimes \mathbb{C}^{p_0}$ is the multiplier by a matrix valued function, say $\theta \mapsto P(\theta)$ such that, $P(\theta)$ is unitary almost everywhere.

Proof. $\mathcal{T}\mathcal{U}^*$ is a unitary operator on $^2(\mathbb{T}^2)\otimes\mathbb{C}^{p_0}$. Moreover, it commutes with the multiplication by $\theta \mapsto e^{i\gamma\cdot\theta}$, $\forall\gamma \in \Gamma$. Hence, it is a multiplier by some matrix valued function, say $\theta \mapsto P(\theta)$. The fact that $P(\theta)$ is unitary almost everywhere is immediately deduced from the unitarity of $\mathcal{T}\mathcal{U}^*$. ☐

We now take $\mathcal{U}$ to be the unitary equivalence given by Theorem 1.2 and $\mathcal{T}$ the one constructed using the Floquet eigenvalues (see section 1.4). By Lemma 6.1 and (1.27), there exists $C > 1$ such that, for almost every $\theta \in \mathbb{T}^2$, one has

$$\frac{1}{C} h(\theta) \leq (\mathcal{U}\mathcal{T}^*)(\theta) \begin{pmatrix} E_{n_0}(\theta) & 0 & \cdots & 0 \\ 0 & E_{n_0+1}(\theta) & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & E_{n_0+p_0-1}(\theta) \end{pmatrix} (\mathcal{T}\mathcal{U}^*)(\theta) \leq Ch(\theta).$$

If $\theta_0 \in \mathcal{S}$, by Lemma 1.9, there exists O, a neighborhood of θ_0 such that

- There exists $\eta > 0$ such that, for $\theta \in O$ and $p(\theta_0) + 1 \leq p \leq p_0 - 1$, $E_{n_0+p}(\theta) \geq \eta$.

- For $\theta \in O$, there exists a real analytic matrix $M(\theta)$ such that the matrix
$$\begin{pmatrix} E_{n_0}(\theta) & 0 & \cdots & 0 \\ 0 & E_{n_0+1}(\theta) & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & E_{n_0+p(\theta_0)}(\theta) \end{pmatrix}$$
is unitarily equivalent to $M(\theta)$.

Hence, if h_0 denotes the determinant of the matrix M, we get that, for some constant $\tilde{C}_O > 1$, one has, for $\theta \in O$,
$$\frac{1}{\tilde{C}_0} h_0(\theta) \leq det \left[\begin{pmatrix} E_{n_0}(\theta) & 0 & \cdots & 0 \\ 0 & E_{n_0+1}(\theta) & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & E_{n_0+p_0-1}(\theta) \end{pmatrix} \right] \leq \tilde{C}_0 h_0(\theta)$$
This in turn implies (6.4) for some new constant C_O. □

7. The proof of Theorem 0.3.

As in section 1, we may assume that E_0 and that the random variables have support in $[0, \omega_+]$. To prove Theorem 0.3, we will compare H_ω with a random Schrödinger operator of the same type with a single site potential supported in a fundamental cell of the lattice Γ, and prove a generalized version of Theorem 0.3 for such operators, namely Proposition 7.1.

Let $n \in \mathbb{N} \setminus \{0\}$ and define the periodic potential
$$(V\mathbf{1})_{\omega,n}(x) = \sum_{\gamma \in 2n\Gamma} (V_\omega \mathbf{1}_{C_n})(x - \gamma) \tag{7.1}$$
and the associated periodic Schrödinger operator
$$H_{\omega,n} = H + (V\mathbf{1})_{\omega,n}. \tag{7.2}$$
Here C_n is the cube centered at 0 of side length $2n+1$.

Let us start with the second step. Assume now that V, the single site potential, be supported in $\gamma + C_0$ for some $\gamma \in \Gamma$ where C_0 is the unit cell of Γ. Then, we prove

Proposition 7.1. *For any $E_0 \in \mathbb{R}$, if*
$$N(E_0 + \varepsilon) - N(E_0 - \varepsilon) = O(\varepsilon^\infty) \text{ when } \varepsilon \to 0^+ \tag{7.3}$$
then, for any $\beta > 0$ integer, $\nu > 0$, for $l \in \mathbb{N}$ sufficiently large, one has
$$\mathbb{E}\left(N\left(E_0 + \frac{1}{l}, H_{\omega,l^\beta}\right) - N\left(E_0 - \frac{1}{l}, H_{\omega,l^\beta}\right)\right) \leq l^{-\nu}. \tag{7.4}$$
This statement is proved in section 7.1.

We now come back to the case of a general single site potential V i.e. we make no assumption about its support. Assumption (H.1.3) (or (H.1.3')) imply that there exists $E_0 > 0$ such that, for $t \in [0,1]$ and $0 \leq E \leq E_0$, one has
$$N(E, H_{\omega,n}) - N(0, H_{\omega,n}) \leq N(E, H^t_{\omega,n}) - N(0, H^t_{\omega,n}) \tag{7.5}$$
where $H^t_{\omega,n} = H + t(V\mathbf{1})_{\omega,n}$.

On the other hand, fix W a positive continuous single site potential such that $W(x) \leq V(x)$ on $\mathbb{R}^d$. Let $\tilde{H}^t_{\omega,n} = H + t(W\mathbf{1})_{\omega,n}$ ($(W\mathbf{1})_{\omega,n}$ and W_ω denote the periodic and random potentials obtained by replacing the single site potential V by W). Then, as the $(\omega_\gamma)_\gamma$ are non negative, for $t \geq 0$, one has $\tilde{H}^t_{\omega,n} \leq H^t_{\omega,n}$; hence, $N(E, H^t_{\omega,n}) \leq N(0, \tilde{H}^t_{\omega,n})$. Moreover, for $t > 0$ sufficiently small, one has $N(0, H^t_{\omega,n}) = N(0, \tilde{H}^t_{\omega,n}) = N(0, H_n)$ where H_n is H seen as $(2n+1)\Gamma$-periodic

operator; to prove this use the Floquet reduction of the periodic operators $H^t_{\omega,n}$, $\tilde H^t_{\omega,n}$ and (1.14). Let us now choose W. As V is not vanishing identically, we can find $\gamma\in\Gamma$ and a non-negative, not identically vanishing, continuous function W such that supp$W\subset\gamma+C_0$ and $W(x)\leq V(x)$ on $\mathbb{R}^d$.
Let $\tilde H^t_\omega = H + t\sum_\gamma \omega_\gamma W(\cdot-\gamma)$. Then, 0 is a upper gap edge for the almost sure spectrum of $\tilde H_\omega$ and we can apply Theorem 0.1 at this gap edge. Hence, the density of states of $\tilde H_\omega$ satisfies (7.3) for $E_0=0$; using (7.5) and Proposition 7.1, we immediately get that $N(E,H_{\omega,n})$ satisfies (7.4).

We now need to transform estimate (7.4) into an information on the Floquet eigenvalues of $H_{\omega,n}$. Therefore, we will need the

Lemma 7.1. *There exists $C>0$ such that, for any $n\geq 1$, if $(E_k(\theta))_{k\geq 0}$ denote the Floquet eigenvalues of $H_{\omega,n}$, then, for any $(\theta,\theta')\in([0,2\pi/n]^d)^2$, we have*

$$\left|E_k(\theta)-E_k(\theta')\right|\leq C|\theta-\theta'|\cdot(1+|E_k(\theta)|+|E_k(\theta')|). \tag{7.6}$$

We postpone the proof of this result to the end of the section and complete the proof of Theorem 0.3. Fix $\theta_0\in[0,2\pi l^{-\beta}]^d$. Pick ω such that $\mathrm{dist}(\sigma(H^{\theta_0}_{\omega,C_{l^\beta}},E))\leq 1/l$ i.e. such that $H^{\theta_0}_{\omega,C_{l^\beta}}$ has an eigenvalue in $[E_0-1/l,E_0+1/l]$. Then, for any $\theta\in[0,2\pi l^{-\beta}]^d$, $H^{\theta}_{\omega,C_{l^\beta}}$ has an eigenvalue in $[E_0-l^{-1}-Cl^{-\beta},E_0+l^{-1}+Cl^{-\beta}]=:I_l$. This implies that, for any $\theta\in[0,2\pi l^{-\beta}]^d$, one has

$$\begin{aligned}\mathbb{P}\left(\left\{\mathrm{dist}(\sigma(H^{\theta_0}_{\omega,C_{l^\beta}}),E_0)\leq\frac{1}{l}\right\}\right)&\leq\mathbb{P}\left(\left\{\text{ there exists an eigenvalue of } H^{\theta}_{\omega,C_{l^\beta}} \text{ in } I_l\right\}\right)\\&\leq\mathbb{E}\left(\sharp\left\{\text{eigenvalues of } H^{\theta}_{\omega,C_{l^\beta}} \text{ in } I_l\right\}\right).\end{aligned}$$

Integrating this over $[0,2\pi l^{-\beta}]^d$ in θ, for some $K>0$, we get

$$\begin{aligned}\mathbb{P}&\left(\left\{\mathrm{dist}(\sigma(H^{\theta_0}_{\omega,C_{l^\beta}},E_0)\leq\frac{1}{l}\right\}\right)\mathrm{Vol}([0,2\pi l^{-\beta}]^d)\\&\leq\int_{[0,2\pi l^{-\beta}]^d}\mathbb{E}\left(\sharp\left\{\text{eigenvalues of } H^{\theta}_{\omega,C_{l^\beta}} \text{ in } I_l\right\}\right)d\theta\\&=\mathbb{E}\Big(N(E_0+l^{-1}+Kl^{-\beta},H_{\omega,l^\beta})-N(E_0-l^{-1}-Kl^{-\beta},H_{\omega,l^\beta})\Big)\end{aligned}\tag{7.7}$$

Taking the fact that $N(E,H_{\omega,n})$ satisfies (7.4) into account, (7.7) immediately yields (0.10); thus, we have completed the proof of Theorem 0.3. ☐

7.1. **The proof of Proposition 7.1.** First notice that, as the random variables are i.i.d and as W is supported in $\gamma+C_0$, we may assume that $\gamma=0$. Now, let $N_{\omega,n}$ be the integrated density of states of $H_{\omega,n}$. Then we show

Theorem 7.1. *Pick $\nu>0$ and a compact interval $I\subset\mathbb{R}$. Then, there exists $\varepsilon_0>0$ such that, for $0<\varepsilon<\varepsilon_0$, $E_0\in I$ and $n\geq\varepsilon^{-1-\nu}$, we have*

$$\begin{aligned}\mathbb{E}(N_{\omega,n}(E_0+\varepsilon/2)&-N_{\omega,n}(E_0-\varepsilon/2))-e^{-(\varepsilon/2)^{-\nu/2}}\\&\leq N(E_0+\varepsilon)-N(E_0-\varepsilon)\\&\qquad\leq\mathbb{E}(N_{\omega,n}(E_0+2\varepsilon)-N_{\omega,n}(E_0-2\varepsilon))+e^{-\varepsilon^{-\nu/2}}.\end{aligned}\tag{7.8}$$

Hence, Theorem 7.1 states that the density of states of the random Hamiltonian H_ω is exponentially well approximated by the expectation of the densities of states of the periodic Hamiltonians $H_{\omega,n}$ even for n not very large.
Proof. The proof of Theorem 7.1 relies upon the

Lemma 7.2. *For any $\alpha \in (0,1)$, there exists $C > 1$ and $\rho > 0$ (depending only on d) such that, for any $\varphi \in \mathcal{C}_0^\infty(\mathbb{R})$, for $k \in \mathbb{N}^*$ and $n \in \mathbb{N}^*$, we have*

$$|\mathbb{E}((\varphi, dN_{\omega,n})) - (\varphi, dN)| \leq Cn^{-(1-\alpha)k} \sup_{\substack{x\in\mathbb{R}\\ 0\leq j\leq k+\rho}} \left|(|x|+C)^{\rho+k}\frac{d^j\varphi}{d^jx}(x)\right|. \tag{7.9}$$

Proof. This lemma was proved in [19, 21, 24] in a different setting but essentially the same arguments apply in our case too. The proof essentially relies upon two facts

1. the fact that, for $\varphi \in \mathcal{C}_0^\infty(\mathbb{R})$, one has
$$\mathbb{E}((\varphi, dN_{\omega,n})) = \mathrm{tr}(\mathbf{1}_{C_0}\varphi(H_{\omega,n})\mathbf{1}_{C_0}). \tag{7.10}$$
2. an estimate on the difference between the resolvents of $H_{\omega,n}$ and H_ω at an energy $z \in \mathbb{C}\setminus\mathbb{R}$, namely

$$\begin{aligned}\left\|\mathbf{1}_{C(0,1)}\left((z-H_{\omega,n})^{-1}-(z-H_\omega)^{-1}\right)\mathbf{1}_{C(0,1)}\right\| &= \\ = \left\|\mathbf{1}_{C(0,1)}\left((z-H_{\omega,n})^{-1}((V\mathbf{1})_{\omega,n}-V_\omega)(z-H_\omega)^{-1}\right)\mathbf{1}_{C(0,1)}\right\| &\leq Ce^{-\frac{1+|z|}{C\cdot \mathrm{Im} z}n}. \end{aligned}\tag{7.11}$$

Equality (7.10) results from a direct computation that uses the fact that supp$V \subset C_0$ (see [23]). Estimate (7.11) is a consequence of the fact that, by construction, $((V\mathbf{1})_{\omega,n} - V_\omega)$ vanishes in the cube centered at 0 of side length $2n$ and of the exponential decay of the resolvent kernels $(z-H_\omega)^{-1}$ and $(z-H_{\omega,n})^{-1}$ (this is a consequence of Combes-Thomas estimates (see e.g. [4, 18]). To obtain Lemma 7.2 from (7.11), one then uses a representation formula for $\varphi(H)$ in terms of almost analytic extensions of φ (see e.g. [6, 23]).
To derive Theorem 7.1, in Lemma 7.2, one just picks φ as a Gevrey function with support in $[E_0-2\varepsilon, E_0+2\varepsilon]$ that is identically 1 in $[E_0-\varepsilon/2, E_0+\varepsilon/2]$. One then optimizes the estimate (7.9) in n using the Gevrey estimates on the derivatives of φ (see [19] for more details). $\square$

7.2. **Proof of Lemma 7.1.** Consider the Floquet problem

$$\begin{cases} H_{\omega,n}\varphi = E\varphi \\ \varphi(x+\gamma) = e^{-i\gamma\theta}\varphi(x),\ \forall x \in \mathbb{R}^2,\ \forall \gamma \in \Gamma. \end{cases}$$

Conjugating this Floquet problem with the multiplication by $e^{ix\theta}$, we transform it into the periodic eigenvalue problem

$$\begin{cases} H_{\omega,n}(\theta)\varphi = E\varphi \\ \varphi(x+\gamma) = \varphi(x),\ \forall x \in \mathbb{R}^2,\ \forall \gamma \in \Gamma \end{cases}$$

where

$$H_{\omega,n}(\theta) = (i\nabla+\theta)^2 + (V\mathbf{1})_{\omega,n} = H_{\omega,n} + 2\theta\cdot i\nabla + \theta^2.$$

The domain of $H_{\omega,n}(\theta)$ is $H^2_{\mathrm{loc}}(\mathbb{R}^d)\cap\mathcal{D}'_0$: it does not depend on θ. Hence, for φ in the domain of $H_{\omega,n}(\theta)$, we have

$$\langle H_{\omega,n}(\theta)\varphi,\varphi\rangle = \langle H_{\omega,n}(\theta')\varphi,\varphi\rangle + 2\langle(\theta-\theta')\cdot i\nabla\varphi,\varphi\rangle + (\theta-\theta')\cdot(\theta+\theta')\|\varphi\|^2.$$

Let $\mathcal{V}_k$ be the vector space generate by the $(k-1)$-st eigenfunctions of $H_{\omega,n}(\theta)$. Then, we have

$$E_k(\theta) = \inf_{\substack{\varphi\perp\mathcal{V}_k\\ \|\varphi\|=1}}\langle H_{\omega,n}(\theta)\varphi,\varphi\rangle \leq \inf_{\substack{\varphi\perp\mathcal{V}_k\\ \|\varphi\|=1}}\langle H_{\omega,n}(\theta')\varphi,\varphi\rangle + C|\theta-\theta'|\cdot(1+|E_k(\theta)|). \tag{7.12}$$

as ∇ is relatively bounded with respect to $H_{\omega,n}(\theta)$ with relative bound 0 (this comes from the relative boundedness of $(V\mathbf{1})_{\omega,n}$ with respect to $-\Delta$ with bound 0). The constant C in (7.12) does not depend on ω or on k as the $\ell^1(L^2)$-norm of $(V\mathbf{1})_{\omega,n}$ is bounded uniformly in ω and k. Then, by the min-max formula for eigenvalues (see e.g. [37]), equation (7.12) gives us

$$E_k(\theta) \leq E_k(\theta') + C|\theta - \theta'| \cdot (1 + |E_k(\theta)|).$$

As θ and θ' play symmetric roles, we get (7.6) and complete the proof of Lemma 7.1. $\square$

References

[1] J. M. Barbaroux, J. M. Combes, and P. D. Hislop. Localization near band edges for random Schrödinger operators. *Helv. Phys. Acta*, 70(1-2):16–43, 1997. Papers honouring the 60th birthday of Klaus Hepp and of Walter Hunziker, Part II (Zürich, 1995).

[2] Edward Bierstone and Pierre D. Milman. Semianalytic and subanalytic sets. *Inst. Hautes Études Sci. Publ. Math.*, (67):5–42, 1988.

[3] R. Carmona and J. Lacroix. *Spectral Theory of Random Schrödinger Operators*. Birkhäuser, Basel, 1990.

[4] J. M. Combes and L. Thomas. Asymptotic behavior of eigenfunctions for multi-particle Schrödinger operators. *Communications in Mathematical Physics*, 34:251–270, 1973.

[5] Hans Grauert. Analytische Faserungen über holomorph-vollständigen Räumen. *Math. Ann.*, 135:263–273, 1958.

[6] B. Helffer and J. Sjöstrand. On diamagnetism and the De Haas-Van Alphen effect. *Annales de l'Institut Henri Poincaré, série Physique Théorique*, 52:303–375, 1990.

[7] Bernard Helffer and Abderemane Mohamed. Asymptotic of the density of states for the Schrödinger operator with periodic electric potential. *Duke Math. J.*, 92(1):1–60, 1998.

[8] Dale Husemoller. *Fibre bundles*. McGraw-Hill Book Co., New York, 1966.

[9] D. Jerison and C. Kenig. Unique continuation and absence of positive eigenvalues for Schrödinger operators. *Annals of Mathematics*, 121:463–494, 1984.

[10] Yulia E. Karpeshina. *Perturbation theory for the Schrödinger operator with a periodic potential*. Springer-Verlag, Berlin, 1997.

[11] G. M. Khenkin, editor. *Several complex variables. IV*. Springer-Verlag, Berlin, 1990. Algebraic aspects of complex analysis, A translation of Современные проблемы математики. Фундаментальные направления, Том 10, Akad. Nauk SSSR, Vsesoyuz. Inst. Nauchn. i Tekhn. Inform., Moscow, 1986 [MR 87m:32003], Translation by J. Leiterer and J. Nunemacher, Translation edited by S. G. Gindikin and G. M. Khenkin.

[12] W. Kirsch. Random Schrödinger operators. In A. Jensen H. Holden, editor, *Schrödinger Operators*, number 345 in Lecture Notes in Physics, Berlin, 1989. Springer Verlag. Proceedings, Sonderborg, Denmark 1988.

[13] W. Kirsch and F. Martinelli. On the spectrum of Schrödinger operators with a random potential. *Communications in Mathematical Physics*, 85:329–350, 1982.

[14] W. Kirsch and F. Martinelli. Large deviations and Lifshitz singularities of the integrated density of states of random hamiltonians. *Communications in Mathematical Physics*, 89:27–40, 1983.

[15] W. Kirsch and B. Simon. Lifshitz tails for the Anderson model. *Journal of Statistical Physics*, 38:65–76, 1985.

[16] Werner Kirsch, Peter Stollmann, and Günter Stolz. Anderson localization for random Schrödinger operators with long range interactions. *Comm. Math. Phys.*, 195(3):495–507, 1998.

[17] Werner Kirsch, Peter Stollmann, and Günter Stolz. Localization for random perturbations of periodic Schrödinger operators. *Random Oper. Stochastic Equations*, 6(3):241–268, 1998.

[18] F. Klopp. An asymptotic expansion for the density of states of a random Schrödinger operator with Bernoulli disorder. *Random Operators and Stochastic Equations*, 3(4):315–332, 1995.

[19] F. Klopp. Band edge behaviour for the integrated density of states of random Jacobi matrices in dimension 1. *Journal of Statistical Physics*, 90(3-4):927–947, 1998.

[20] F. Klopp. Erratum to the paper "Internal Lifshits tails for random perturbations of periodic Schrödinger operators". Available on `http://zeus.math.univ-paris13.fr/~klopp/publi.html`, 2000.

[21] F. Klopp. Precise high energy asymptotics for the integrated density of states of an unbounded random Jacobi matrix. *Reviews in Mathematical Physics*, 12(4):575–620, 2000.

[22] F. Klopp and J. Ralston. Endpoints of the spectrum of periodic operators are generically simple. Available on `http://zeus.math.univ-paris13.fr/~klopp/publi.html`, 2000.

[23] Frédéric Klopp. Internal Lifshits tails for random perturbations of periodic Schrödinger operators. *Duke Math. J.*, 98(2):335–396, 1999.

[24] Frédéric Klopp and Leonid Pastur. Lifshitz tails for random Schrödinger operators with negative singular Poisson potential. *Comm. Math. Phys.*, 206(1):57–103, 1999.

[25] P. Kuchment. *Floquet theory for partial differential equations*, volume 60 of *Operator Theory: Advances and Applications*. Birkhäuser, Basel, 1993.

[26] I. M. Lifshitz. Structure of the energy spectrum of impurity bands in disordered solid solutions. *Soviet Physics JETP*, 17:1159–1170, 1963.

[27] I.M. Lifshitz, S.A. Gredeskul, and L.A. Pastur. *Introduction to the theory of disordered systems.* Wiley, New-York, 1988.

[28] S. Łojasiewicz. Sur le problème de la division. *Studia Math.*, 18:87–136, 1959.

[29] G. Mezincescu. Lifshitz singularities for periodic operators plus random potentials. *Journal of Statistical Physics*, 49:1081–1090, 1987.

[30] G. Mezincescu. Internal Lifshits singularities for one dimensional Schrödinger operators. *Communications in Mathematical Physics*, 158:315–325, 1993.

[31] S. Nakao. On the spectral distribution of the Schrödinger operator with random potential. *Japan Journal of Mathematics*, 3:117–139, 1977.

[32] L. Pastur. Behaviour of some Wiener integrals as $t \to +\infty$ and the density of states of the Schrödinger equation with a random potential. *Teor.-Mat.-Fiz*, 32:88–95, 1977. (in russian).

[33] L. Pastur and A. Figotin. *Spectra of Random and Almost-Periodic Operators.* Springer Verlag, Berlin, 1992.

[34] Valentin V. Petrov. *Limit theorems of probability theory.* The Clarendon Press Oxford University Press, New York, 1995. Sequences of independent random variables, Oxford Science Publications.

[35] D. H. Phong and E. M. Stein. The Newton polyhedron and oscillatory integral operators. *Acta Math.*, 179(1):105–152, 1997.

[36] Gilles Pisier. *The volume of convex bodies and Banach space geometry.* Cambridge University Press, Cambridge, 1989.

[37] M. Reed and B. Simon. *Methods of Modern Mathematical Physics, Vol IV: Analysis of Operators.* Academic Press, New-York, 1978.

[38] M.A. Shubin. Spectral theory and index of elliptic operators with almost periodic coefficients. *Russian Mathematical Surveys*, 34:109–157, 1979.

[39] J. Sjöstrand. Microlocal analysis for periodic magnetic Schrödinger equation and related questions. In *Microlocal analysis and applications*, volume 1495 of *Lecture Notes in Mathematics*, Berlin, 1991. Springer Verlag.

[40] M. Skriganov. Geometric and arithmetic methods in the spectral theory of multidimensional periodic operators. *Proceedings of the Steklov Institute of Mathematics*, 171:1–121, 1987.

[41] M. M. Skriganov. Proof of the Bethe-Sommerfeld conjecture in dimension 2. *Dokl. Akad. Nauk SSSR*, 248(1):39–42, 1979.

[42] Peter Stollmann. Lifshitz asymptotics via linear coupling of disorder. *Math. Phys. Anal. Geom.*, 2(3):279–289, 1999.

[43] Alain-Sol Sznitman. *Brownian motion, obstacles and random media.* Springer-Verlag, Berlin, 1998.

[44] Peter A. Tomas. A restriction theorem for the Fourier transform. *Bull. Amer. Math. Soc.*, 81:477–478, 1975.

[45] V. Varchenko. Newton polyhedra and estimations of oscillatory integrals. *Functional Analysis and its Applications*, 8:175–196, 1976.

[46] I. Veselic. Localization for random perturbations of periodic Schrödinger operators with regular floquet eigenvalues. Technical report, Universtät Bochum, 1998.

[47] Thomas H. Wolff. Recent work on sharp estimates in second order elliptic unique continuation problems. In *Fourier analysis and partial differential equations (Miraflores de la Sierra, 1992)*, pages 99–128. CRC, Boca Raton, FL, 1995.

(Frédéric Klopp) Département de Mathématique, Institut Galilée, U.M.R. 7539 C.N.R.S, Université de Paris-Nord, 99 avenue J.-B. Clément, F-93430 Villetaneuse, France

E-mail address: `klopp@math.univ-paris13.fr`

(Thomas Wolff) Department of Mathematics, 253-37 Caltech, Pasadena, CA 91125 USA

E-mail address: `wolff@its.caltech.edu`