Vietnam Journal of Mathematics **26**:3 (1998) 189–216

Vietnam Journal
of
MATHEMATICS

Survey

A Short Introduction to the Theory of Random Schrödinger Operators

Frédéric Klopp

Department of Mathematics, Institute of Galilée, University of Paris-Nord
Avenue J.-B. Clément, F-93430 Villetaneuse, France

Received March 16, 1998

Abstract. This paper is an enlarged version of a talk given at the *French-Vietnamese Colloquium of Mathematics* held in Ho Chi Minh City, 3–9 March 1997. It is a quick, non-exhaustive presentation of some points of the theory of random Schrödinger operators. We essentially restrict ourselves to a single model, the continuous Anderson model.

1. Some Introductory Words

As a first approximation, solid state matter may be considered to be perfectly periodic. In the one electron approximation, the Hamiltonian H used to describe this perfect crystal is a periodic Schrödinger operator acting on $L^2(\mathbf{R}^d)$, that is,

$$H = -\Delta + W,$$

where

- $-\Delta = \sum_{j=1}^{d} \frac{\partial^2}{\partial x_j^2}$ is the usual Laplace operator (with domain $H^2(\mathbf{R}^d)$);
- W is a Γ-periodic potential, (i.e., for $\gamma \in \Gamma$, $W(x+\gamma) = W(x)$);
- Γ is a non-degenerate lattice in $\mathbf{R}^d$.

The spectral theory of periodic Schrödinger operators is well known (see [37, 53, 59]). Under mild conditions on the potential W, it can be proved that the spectrum of H is made of bands of purely absolutely continuous spectrum. Intuitively, this characterizes a diffusive behavior of the evolution semi-group associated to H. Physically, this roughly means that, at some energies (i.e., energies inside the spectrum), solid state matter will be a conductor and, at other energies (i.e., energies outside the spectrum), it will be an insulator.

However, in experiments, it was discovered that this diffusive behavior is not always seen where expected. One of the possible explanations is the presence of impurities, that change the scattering process. Of course, it is impossible to know the exact location of the impurities in a given chunk of matter. Nevertheless, one may understand some global features of the impurity distribution (e.g., concentration, homogeneity, etc.). One way to take these global features into account is to study a large number of pieces of the same material. Hence, a natural thing to do is to study a random distribution of impurities [1]. This naturally leads to random Schrödinger operators.

A random Schrödinger operator is a family of Schrödinger operators endowed with some probabilistic structure. As we shall see below in a simple example, if this structure is ergodic, then, at least spectrally, the family of Schrödinger operators behaves like a single operator. Indeed, one can associate a spectrum to the family. In the same way, many spectral objects can be defined for the family of operators (e.g., density of states, etc.). These concepts are then the natural objects to be studied. The aim of this paper is to briefly present some of these notions and some of the most striking results concerning them.

As a conclusion to this introduction, we emphasize the fact that this paper does not aim at giving an exhaustive picture of the state of the art on the topic of random Schrödinger operators. It simply presents a brief introduction[1] to some notions, results and techniques that have proved valuable and useful in the study of random Schrödinger operators. In particular, we will restrict ourselves to a very simple and specific model: the continuous Anderson model (see (3.1)). We will also make the necessary assumptions to have as few technical difficulties as possible. For a more exhaustive treatment of this topic, one may consult, for example, [7, 52] (or [41] for a more physical point of view).

2. Some Basic Facts about Periodic Schrödinger Operators

We will first briefly review some standard facts from the Floquet theory for periodic Schrödinger operators. Basic references where this material may be found are [53, 55, 59].

Let Γ be a non-degenerate lattice in $\mathbf{R}^d$ generated by $(e_i)_{1\leq i\leq d}$, a basis of $\mathbf{R}^d$ (i.e., $\Gamma = \oplus_{j=1}^d \mathbf{Z}e_j$), and let W be a Γ-periodic bounded potential (i.e., a measurable function).[2]

Define $H = -\Delta + W$ to be a periodic Schrödinger operator acting on $L^2(\mathbf{R}^d)$. Then, by [53], H is essentially self-adjoint on $\mathcal{C}_0^\infty(\mathbf{R}^d)$ with domain $H^2(\mathbf{R}^d)$. Its unique self-adjoint extension will also be denoted by H (cf. Appendix).

2.1. The Floquet Decomposition

As W is Γ-periodic, we know that, for any $\gamma \in \Gamma$,

$$\tau_\gamma \circ H \circ \tau_\gamma^* = \tau_\gamma \circ H \circ \tau_{-\gamma} = H.$$

[1] This paper is essentially meant for students.

[2] This assumption as well as many of the following assumptions may be relaxed [52, 53].

For $\theta \in \mathbf{R}^d$ and $u \in S(\mathbf{R}^d)$, the Schwartz space of rapidly decreasing functions, we define

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

U can be extended as a unitary isometry from $L^2(\mathbf{R}^d)$ to $\mathcal{H}$ where

$$\mathcal{H} = \big\{u(x,\theta) \in L^2_{\mathrm{loc}}(\mathbf{R}^d) \otimes L^2(\mathbf{T}^*);\ \forall x \in \mathbf{R}^d,\ \forall\theta \in \mathbf{T}^*,$$
$$\forall\gamma \in \Gamma,\ u(x+\gamma,\theta) = e^{i\gamma.\theta} u(x,\theta)\big\},$$

and $\mathcal{H}$ is equipped with the norm $\dfrac{1}{\mathrm{Vol}(\mathbf{T}^*)}\displaystyle\int_{\mathbf{T}^*} \|u(\cdot,\theta)\|^2_{L^2(C_0)} d\theta$ (here, C_0 is the unit cell of the lattice Γ, i.e., $C_0 = \{x_1e_1 + \cdots + x_de_d,\ -1/2 \le x_1, \ldots, x_d < 1/2\}$). $\mathcal{H}$ is a fiber bundle over the torus $\mathbf{T}^*$, the fibers being the Hilbert spaces $\mathcal{H}_\theta = \{u \in L^2(\mathbf{R}^d);\ \forall x \in \mathbf{R}^d,\ \forall\gamma\in\Gamma,\ u(x+\gamma) = e^{i\gamma.\theta}u(x)\}$ equipped with the scalar product on $L^2(C_0)$.

The inverse of U is given by the formula:

$$\text{for } v \in \mathcal{H},\quad (U^*v)(x) = \frac{1}{\mathrm{Vol}(\mathbf{T}^*)}\int_{\mathbf{T}^*} v(x,\theta)d\theta.$$

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

$$UHU^* = \int^{\oplus} H_\theta d\theta, \tag{2.1}$$

where H_θ is the differential operator H acting on $\mathcal{H}_\theta$ with domain $\mathcal{H}^2_\theta := \mathcal{H}_\theta \cap H^2_{\mathrm{loc}}(\mathbf{R}^d)$. H_θ is self-adjoint.

As H is elliptic, we know that H_θ has a compact resolvent; its spectrum is discrete. Let us denote its eigenvalues (also called Floquet eigenvalues of H) by

$$E_0(\theta) \le E_1(\theta) \le \cdots \le E_n(\theta) \ldots .$$

The functions $(\theta \mapsto E_n(\theta))_{n\in\mathbf{N}}$ are Lipschitz continuous and one has

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

Let σ denote the spectrum of H. It is purely absolutely continuous and given by

$$\sigma = \bigcup_{n\in\mathbf{N}} E_n(\mathbf{T}^*).$$

2.2. The Density of States

One defines $n(E)$, the *integrated density of states* of H, in the following way (cf. [55, 59]) for $E \in \mathbf{R}$,

$$n(E) = \frac{1}{(2\pi)^d}\sum_{n\in\mathbf{N}} \int_{\{\theta\in\mathbf{T}^*;\ E_n(\theta)\le E\}} d\theta. \tag{2.2}$$

It is a continuous, positive, non-decreasing function. It is constant in the gaps of the spectrum of H and its growth points are exactly the points of σ. One can give another formula defining the density of states n. Let $\varphi \in \mathcal{C}_0^\infty(\mathbf{R})$ and let χ_0 be the characteristic function of C_0. Then, the operator $\chi_0\varphi(H)$ is trace-class and we have

$$\int_{\mathbf{R}} \varphi(E)dn(E) = \langle \varphi, dn\rangle = \frac{1}{\mathrm{Vol}(C_0)}\mathrm{tr}(\chi_0\varphi(H)\chi_0). \tag{2.3}$$

3. The Continuous Anderson Model

Let H be a self-adjoint periodic Schrödinger operator as defined in the previous section. And let $(\omega_\gamma)_{\gamma\in\Gamma}$ be a collection of independent identically distributed (i.i.d. for short) real-valued random variables. Furthermore, we will assume they are bounded.

Let $V : \mathbf{R}^d \to \mathbf{R}$ be a L^∞-function such that $\|V\|_{1,\infty} = \sum_{\gamma\in\Gamma} \|V\|_{L^\infty(\tau_\gamma(C_0))}$ is bounded. Then we define

$$H_\omega = H + V_\omega \text{ where } V_\omega(x) = \sum_{\gamma\in\Gamma} \omega_\gamma V(x-\gamma). \tag{3.1}$$

Let us now consider the function $\omega \mapsto H_\omega$. As H_ω is a bounded perturbation of H, it is a well-defined mapping from the probability space on which ω lives into the space of self-adjoint operators with domain $H^2(\mathbf{R}^d)$. This mapping is weakly measurable in ω (see [54]), i.e., for $(\phi, \psi) \in H^2(\mathbf{R}^d)^2$, the mapping $\omega \mapsto \langle \psi, H_\omega\phi\rangle$ is measurable. Indeed, it is an absolutely convergent sum of measurable functions as

$$\langle \psi, H_\omega\phi\rangle = \langle \psi, H\phi\rangle + \langle \psi, V_\omega\phi\rangle = \langle \psi, H\phi\rangle + \sum_{\gamma\in\Gamma} \omega_\gamma \int_{\mathbf{R}^d} V(x-\gamma)\psi(x)\overline{\phi(x)}dx.$$

We now consider the family $(H_\omega)_\omega$ endowed with the natural probability measure defined by the random variables $(\omega_\gamma)_{\gamma\in\mathbf{R}^d}$ (cf. Appendix).

3.1. Ergodicity and the Almost Sure Spectrum

We will see that, due to ergodicity, the family of Schrödinger operators $(H_\omega)_\omega$ may in some respects be considered as a single operator.

3.1.1. Ergodicity

For $\gamma \in \Gamma$, let $\tau_\gamma : L^2(\mathbf{R}^d) \to L^2(\mathbf{R}^d)$ be defined by

$$\forall u \in L^2(\mathbf{R}^d), \quad (\tau_\gamma u)(x) = u(x-\gamma).$$

Then τ_γ is a unitary operator on $L^2(\mathbf{R}^d)$ and $\tau_\gamma^* = \tau_{-\gamma}$. Moreover, as W is Γ-periodic, we have

$$\tau_\gamma H \tau_\gamma^* = H.$$

Let $t_\gamma : \mathbf{R}^\Gamma \to \mathbf{R}^\Gamma$ be defined by

$$\forall \omega \in \mathbf{R}^\Gamma, \quad (t_\gamma\omega)_\beta = \omega_{\beta-\gamma}. \tag{3.2}$$

One computes

$$\tau_\gamma H_\omega \tau_\gamma^* = \tau_\gamma H \tau_\gamma^* + \tau_\gamma V_\omega \tau_\gamma^* = H + V_{t_\gamma \omega}. \tag{3.3}$$

The action of the group $\{t_\gamma;\ \gamma \in \Gamma\}$ on the probability space underlying the random variables $(\omega_\gamma)_{\gamma \in \Gamma}$ is ergodic (cf. Appendix). Therefore, as the $(\tau_\gamma)_\gamma$ are unitary operators, we will say that the operator H_ω is an *ergodic random operator.*

3.1.2. The Almost Sure Spectrum

The first consequence of ergodicity is as follows:

Theorem 3.1. [49] *There exists $\Sigma \subset \mathbf{R}$ a closed set such that, for almost every ω, the spectrum of H_ω is equal to Σ.*

We call Σ the (almost sure) spectrum of H_ω. This result partially justifies the fact that we speak of the family of operators $(H_\omega)_\omega$ as a single operator (at least from a probabilistic point of view).

We will not give a rigorous proof of this classical statement (see [7, 11, 49, 52] and references therein). We will heuristically explain it. It is well known that a random variable that is invariant under the action of an ergodic group is almost surely constant (see the appendix and, e.g., [19]). Hence, Eq. (3.3) tells us that the spectrum of H_ω is invariant under the action of $\{t_\gamma;\ \gamma \in \Gamma\}$. So this spectrum should be almost surely constant. Of course to make this rigorous, one has to understand what is meant by the measurability of the spectrum with respect to ω. Therefore, one prefers to work with the spectral resolution of H_ω which is easier to handle.

Theorem 3.1 can be made more precise. For a self-adjoint operator A, we denote its pure point spectrum by $\sigma_{pp}(A)$ (the closure of the set of eigenvalues of A), its absolutely continuous spectrum by $\sigma_{ac}(A)$ and its singular continuous spectrum by $\sigma_{sc}(A)$ (cf. Appendix). Then, we have the following:

Theorem 3.2. [26, 38, 51] *There exist three closed subsets Σ_{ac}, Σ_{sc}, Σ_{pp} of $\mathbf{R}$, such that, for almost every ω,*

$$\sigma_{ac}(H_\omega) = \Sigma_{ac},\quad \sigma_{sc}(H_\omega) = \Sigma_{sc},\quad \sigma_{pp}(H_\omega) = \overline{\{\textit{eigenvalues of } H_\omega\}} = \Sigma_{pp}. \tag{3.4}$$

We see that the spectral types of the operators of the family H_ω are also almost surely constant. Let us note that in the case of the pure point spectrum (i.e., the eigenvalues), the set of eigenvalues itself is not almost surely constant; only its closure is almost surely constant [52].

One of the main questions of the theory is then to determine the previously defined sets Σ, Σ_{ac}, Σ_{sc} and Σ_{pp}.

3.2. The Density of States

Let us define the density of states. Physically, what we would like to do is count the number of states in some energy range, say, below energy E. The problem will be that, typically, we will be in the presence of a continuum of states (as in the periodic case). So the obvious counting function will not make any sense. Therefore, the natural progression is to count the number of states per unit of volume.

To do this "per unit volume counting", let $\Lambda \subset \mathbf{R}^d$ be a large bounded cube. Consider the restriction of H_ω to Λ with Dirichlet boundary conditions imposed on $\partial\Lambda$. Let us momentarily forget about the random nature of H_ω and denote this restriction by H_Λ. As H_ω is a rather ideal perturbation of the Laplacian, the spectrum of H_Λ will be discrete. So, given an interval $I \subset \mathbf{R}$, we may count the number of eigenvalues of H_Λ in I. Let us say that there are $N_\Lambda(I)$ such eigenvalues. The quantity we are interested in is $\nu_\Lambda(I) = N_\Lambda(I)/|\Lambda|$, the number of states per unit of volume (here, $|\Lambda|$ denotes the volume of Λ). For the Laplace operator restricted to Λ (with Dirichlet boundary conditions), we know that the number of eigenvalues in I grows with λ proportionately to $|\Lambda|$. If H_ω is not a nasty perturbation of $-\Delta$, we may expect $N_\Lambda(I)$ to behave in the same way as the counting function for the Dirichlet Laplacian. So $\nu_\Lambda(I)$ seems to be a reasonable quantity to consider. In particular, we may hope that $\nu_\Lambda(I)$ converges when Λ tends to $\mathbf{R}^d$. In this case, we shall call this limit the integrated density of states $\nu(I)$ of H_ω.

Although we did not do so, for the periodic Schrödinger operator H, the integrated density of states n (defined in (2.2)) may be obtained by the procedure described above.

The procedure described in the above paragraph works well and is somewhat classical (see [7, 25, 52]). Here, we will develop a different approach based on periodic approximations. The main advantage of this approximation is the speed of convergence towards its limit. Essentially, the speed of convergence is exponential in the size of the side length of the cube on which one approximates. More details about this approach may be found in [35, 36].

Pick $n \in \mathbf{N}^*$ and a realization of the random variables $(\omega_\gamma)_{\gamma\in\Gamma}$. Consider the following $(2n+1)\Gamma$-periodic Schrödinger operator:

$$H_{\omega,n} = H + \sum_{\gamma\in\Gamma/(2n+1)\Gamma} \omega_\gamma \Big(\sum_{\beta\in(2n+1)\Gamma} \tau_{\gamma+\beta} V \Big). \tag{3.5}$$

Here, we identify $\Gamma/(2n+1)\Gamma$ with the cube of center 0 and side length $2n$ in Γ, i.e., $\big\{\gamma = \sum_{i=1}^d \gamma_i e_i;\ -n \le \gamma_i \le n,\ 1 \le i \le d\big\}$. Note that only finitely many of random variables come into the definition of $H_{\omega,n}$.

In Subsec. 2.2, we defined the integrated density of states of a periodic Schrödinger operator. We will denote the integrated density of states of $H_{\omega,n}$ by $N_{\omega,n}(E)$. Let $dN_{\omega,n}(E)$ denote its distributional derivative; it is a positive measure. Then we have the following:

Theorem 3.3. *For almost every $\omega \in \Omega$, $dN_{\omega,n}(E)$ converges $*$-weakly to dN, a non-random positive measure, i.e.,*

$$\forall \varphi \in \mathcal{C}_0^\infty(\mathbf{R}),\quad \int_{\mathbf{R}} \varphi dN_{\omega,n} \underset{n\to+\infty}{\longrightarrow} \int_{\mathbf{R}} \varphi dN.$$

N is called the integrated density of states of H_ω.

Proof. Pick $\varphi \in \mathcal{C}_0^\infty(\mathbf{R})$. By formula (2.3), we know that

$$\langle \varphi, dN_{\omega,n}\rangle = \frac{1}{\mathrm{Vol}(C_n)} \mathrm{tr}(\chi_{C_n}\varphi(H_{\omega,n})\chi_{C_n}) \tag{3.6}$$

where $C_n = (2n+1)C_0$ is the fundamental domain of $(2n+1)\Gamma$. As the trace is cyclic and $\chi_{C_n} = \sum_{\gamma \in C_n \cap \Gamma} \tau_\gamma \chi_{C_0}$, we obtain

$$\langle \varphi, dN_{\omega,n} \rangle = \frac{1}{\mathrm{Vol}(C_0)\#(C_n \cap \Gamma)} \sum_{\gamma \in C_n \cap \Gamma} \mathrm{tr}(\chi_{C_0} \tau_\gamma^* \varphi(H_{\omega,n}) \tau_\gamma \chi_{C_0}) \tag{3.7}$$

where $\#(E)$ denotes the number of points of E. For $n \geq 1$ and $\gamma \in C_n \cap \Gamma$, define

$$a_{\gamma,n}(\omega) = \mathrm{tr}(\chi_{C_0} \tau_\gamma^* \varphi(H_{\omega,n}) \tau_\gamma \chi_{C_0}).$$

Obviously, $a_{\gamma,n}(\omega)$ are random variables. Let $m \geq n$ and pick $\gamma \in C_n \cap \Gamma$. Then, using the Helffer-Sjöstrand formula (see Appendix), we compute

$$\begin{aligned} a_{\gamma,n}(\omega) - a_{\gamma,m}(\omega) &= \mathrm{tr}(\chi_{C_0+\gamma}(\varphi(H_{\omega,n}) - \varphi(H_{\omega,m}))\chi_{C_0+\gamma}) \\ &= \frac{i}{2\pi} \int_C \frac{\partial \tilde{\varphi}}{\partial \bar{z}}(z) \cdot \Delta_{\omega,n,m}(z) dz \wedge d\bar{z}, \end{aligned}$$

where $\tilde{\varphi}$ is an almost analytic extension of φ and

$$\begin{aligned} \Delta_{\omega,n,m}(z) = \mathrm{tr}(\chi_{C_0+\gamma}((\lambda_0 + H_{\omega,n})^{-q}(z - H_{\omega,n})^{-1} \\ - (\lambda_0 + H_{\omega,m})^{-q}(z - H_{\omega,m})^{-1})\chi_{C_0+\gamma}). \end{aligned}$$

Here, λ_0 is real number chosen such that it is a uniform lower bound for all the operators $H_{\omega,n}$; q is chosen larger than $d/2$ so that $\Delta_{\omega,n,m}(z)$ is a trace-class.

Let us now estimate the trace class norm of the operator $\Delta_{\omega,n,m}(z)$. We have

$$\|\Delta_{\omega,n,m}(z)\|_{\mathrm{tr}} \leq A + B \tag{3.8}$$

where

$$\begin{aligned} A &= \|\chi_{C_0+\gamma}((z - H_{\omega,n})^{-1} - (z - H_{\omega,m})^{-1})(\lambda_0 + H_{\omega,m})^{-q} \chi_{C_0+\gamma}\|_{\mathrm{tr}} \\ &= \|\chi_{C_0+\gamma}(z - H_{\omega,n})^{-1} \Delta^V_{\omega,n,m}(z - H_{\omega,m})^{-1}(\lambda_0 + H_{\omega,m})^{-q} \chi_{C_0+\gamma}\|_{\mathrm{tr}}, \end{aligned}$$

$$\begin{aligned} B &= \|\chi_{C_0+\gamma}(z - H_{\omega,n})^{-1}((\lambda_0 + H_{\omega,n})^{-q} - (\lambda_0 + H_{\omega,m})^{-q})\chi_{C_0+\gamma}\|_{\mathrm{tr}} \\ &= \|\chi_{C_0+\gamma}(z - H_{\omega,n})^{-1} \sum_{l=1}^{q-1} (\lambda_0 + H_{\omega,n})^{l-q} \Delta^V_{\omega,n,m} (\lambda_0 + H_{\omega,m})^{-l} \chi_{C_0+\gamma}\|_{\mathrm{tr}}, \end{aligned}$$

and

$$\Delta^V_{\omega,n,m}(x) = \sum_{\substack{|\alpha|>n \\ \alpha \in \Gamma}} (\omega_\alpha^n - \omega_\alpha^m) V(x - \alpha).$$

Here, $\|\cdot\|_{\mathrm{tr}}$ denotes the trace-class norm, and $\omega_\alpha^{n,m} = \omega_{\tilde{\alpha}}$ where $\alpha = \tilde{\alpha} \bmod ((2n+1)\Gamma)$ or $\bmod ((2m+1)\Gamma)$.

We separately control A and B. For A, as the $(\omega_\gamma)_\gamma$ are bounded, we obtain, for some $M > 0$,

$$\begin{aligned} A &\leq \sum_{\beta\in\Gamma} \|\chi_{C_0+\gamma}(z - H_{\omega,n})^{-1}\Delta^V_{\omega,n,m}\chi_{\beta+C_0}\|_{\mathcal{L}(L^2(\mathbf{R}^d))} \\ &\qquad \times \|\chi_{\beta+C_0}(z - H_{\omega,m})^{-1}(\lambda_0 + H_{\omega,m})^{-q}\chi_{C_0+\gamma}\|_{\mathrm{tr}} \\ &\leq M \sum_{\substack{\beta\in\Gamma \\ |\alpha|>k,\alpha\in\Gamma}} \|\chi_{C_0+\gamma}(z - H_{\omega,n})^{-1}V\chi_{\beta+\alpha+C_0}\|_{\mathcal{L}(L^2(\mathbf{R}^d))} \\ &\qquad \times \|\chi_{\beta+C_0}(z - H_{\omega,m})^{-1}(\lambda_0 + H_{\omega,m})^{-q}\chi_{C_0+\gamma}\|_{\mathrm{tr}}. \end{aligned}$$

By the appendix of [32] (Proposition 4.3) and by the proof of Theorem XIII.96 of [53], we know that, for some $K > 1$, we have, uniformly in ω, for $|\mathrm{Im}\, z| \leq 1$

$$\|\chi_{\beta+C_0}(z - H_{\omega,m})^{-1}(\lambda_0 + H_{\omega,m})^{-q}\chi_{C_0+\gamma}\|_{\mathrm{tr}} \leq \frac{K}{|\mathrm{Im} z|^{2e}} - |\mathrm{Im} z||\beta - \gamma|_\Gamma/K$$

and

$$\|\chi_{C_0+\gamma}(z - H_{\omega,n})^{-1}V\chi_{\beta+\alpha+C_0}\|_{\mathcal{L}(L^2(\mathbf{R}^d))} \leq \frac{K}{|\mathrm{Im} z|^2}\|\tau_{\beta+\alpha-\gamma}V\|_{L^\infty(C_0)}$$

(where τ_γ is the translation by γ).
Hence, we obtain

$$A \leq \sum_{\beta\in\Gamma}\sum_{\substack{|\alpha|>n \\ \alpha\in\Gamma}} \frac{K}{|\mathrm{Im} z|^4}\|\tau_{\beta+\alpha-\gamma}V\|_{L^\infty(C_0)}.e^{-|\mathrm{Im} z||\beta-\gamma|_\Gamma/K}.$$

Doing the same computation for B, we obtain the same estimate for $\Delta_{\omega,n,m}(z)$.

Hence, integrating these estimates, as $\partial_{\overline{z}}\tilde{\varphi}$ vanishes to infinite order on the real axis (see (6.2)), and using (3.8) and (6.3), we obtain

$$\begin{aligned} &\mathrm{Vol}(C_0)\Big|\mathbf{E}\Big(\int_{\mathbf{R}} \varphi(\lambda)dN_{\omega,k}\Big) - \langle\varphi, dN\rangle\Big| \leq \mathbf{E}\Big(\frac{1}{\pi}\int_{\mathbf{C}} \Big|\frac{\partial\tilde{\varphi}}{\partial\overline{z}}(z)\Big| \|\Delta_{\omega,n,m}(z)\|_{\mathrm{tr}}dxdy\Big) \\ &\leq \frac{1}{\pi}\int_{\mathbf{C}} \frac{C}{|\mathrm{Im} z|^4}\Big|\frac{\partial\tilde{\varphi}}{\partial\overline{z}}(z)\Big| \Big(\sum_{\substack{|\alpha|>n,\alpha\in\Gamma \\ \beta\in\Gamma}} \|\tau_{\beta+\alpha-\gamma}V\|_{L^\infty(C_0)}e^{-|\mathrm{Im} z||\beta-\gamma|_\Gamma/C}\Big)dxdy. \end{aligned} \tag{3.9}$$

Now, as $\sum_{\gamma\in\Gamma} \|\tau_\gamma V\|_{L^\infty(C_0)} < +\infty$, we apply Lebesgue's dominated convergence theorem to the right-hand side of (3.9) so that it tends to 0 when n tends to $+\infty$. Hence, the sequence $a_{\gamma,n}(\omega)$ is a Cauchy sequence and converges. We call $a_\gamma(\omega)$ its limit.

We may also use the same computation as above to estimate $a_{\gamma,n}(\omega) - a_{0,n}(t_\gamma(\omega))$. This shows that $a_{\gamma,n}(\omega) - a_{0,n}(t_\gamma(\omega)) \to 0$ as $n \to +\infty$. Hence, the random variable a_γ satisfies, for any $\omega \in \Omega$,

$$a_\gamma(t_\beta(\omega)) = a_{\gamma+\beta}(\omega). \tag{3.10}$$

Moreover, (3.9) is sufficient to show that

$$\frac{1}{\#(C_n \cap \Gamma)}\sum_{\gamma\in C_n\cap\Gamma} |a_{\gamma,n}(\omega) - a_\gamma(\omega)| \to 0$$

when $n \to +\infty$. So, by (3.7), $\langle \varphi, dN_{\omega,n} \rangle$ admits a limit when $n \to \infty$ if and only if $S_n(\omega)$ admits a limit, where

$$S_n(\omega) = \frac{1}{\#(C_n \cap \Gamma)} \sum_{\gamma \in C_n \cap \Gamma} a_\gamma(\omega).$$

Now, $(a_\gamma)_{\gamma \in \Gamma}$, being the images of a_0 under the action of the group of shifts $(t_\gamma)_{\gamma \in \Gamma}$ and the action of this group being ergodic on the considered probability space, the celebrated Birkhoff-Khinchine Theorem (see Appendix, [19, 25]) tells us that $S_n(\omega)$ has a limit and that this limit is almost surely independent of ω.

To end this proof, we just have to check that the limit defines a positive distribution on $\mathbf{R}$. We leave this easy task to the reader. ∎

For $E \in \mathbf{R}$, let us define $N(E) = \int_{-\infty}^{E} dN$. N is the integrated density of states of H_ω. The *-weak convergence proved in Theorem 3.3 implies that, for any $E \in \mathbf{R}$, a continuity point of N, we have ω almost surely, $N_{\omega,n}(E) \to N(E)$ as $n \to \infty$.

As a consequence of Theorem 2.2, we obtain a nice closed formula for the density of states. Pick $\varphi \in \mathcal{C}_0^\infty(\mathbf{R})$. As the density of states of H_ω is non-random, by (3.6), we know that is the limit of $\mathbf{E}(\langle \varphi, dN_{\omega,n} \rangle)$ as $n \to +\infty$ (where $\mathbf{E}(X)$ denotes the expectation of the random variable X). Using the stationarity of the $(\omega_\gamma)_{\gamma \in \Gamma}$, we obtain

$$\begin{aligned} \mathbf{E}(\langle \varphi, dN_{\omega,n} \rangle) &= \frac{1}{\mathrm{Vol}(C_n)} \mathrm{tr}(\chi_{C_n} \mathbf{E}(\varphi(H_{\omega,n})) \chi_{C_n}) \\ &= \frac{1}{\mathrm{Vol}(C_0)} \mathrm{tr}(\chi_{C_0} \mathbf{E}(\varphi(H_{\omega,n})) \chi_{C_0}). \end{aligned}$$

Hence, taking the limit as $n \to +\infty$, using the same arguments as in the proof of Theorem 2.2, we obtain

$$\langle \varphi, dN \rangle = \frac{1}{\mathrm{Vol}(C_0)} \mathbf{E}(\mathrm{tr}(\chi_{C_0} \varphi(H_\omega) \chi_{C_0})),$$

hence,

$$dN = \frac{1}{\mathrm{Vol}(C_0)} \mathbf{E}(\mathrm{tr}(\chi_{C_0} E_\omega \chi_{C_0})), \tag{3.11}$$

where E_ω is the spectral resolution of H_ω.

4. The Spectrum: Its Location

Now, the first question one needs to answer about the spectrum of H_ω (i.e., the almost sure spectrum of H_ω) is the question of its location. Therefore, we will use the following theorem:

Theorem 4.1. [2, 51] *Let σ be the spectrum of H_ω and dN its density of states. Then we have*

$$\mathrm{supp}(dN) = \Sigma.$$

Proof. By formula (3.11), one obviously has

$$\operatorname{supp}(dN) \subset \Sigma.$$

Conversely, fix $E \in \mathbf{R}$, $\varepsilon > 0$ and assume $\int_{E-\varepsilon}^{E+\varepsilon} dN = 0$, that is,

$$\mathbf{E}(\operatorname{tr}(\chi_{C_0} E_\omega([E-\varepsilon,\ E+\varepsilon])\chi_{C_0})) = 0.$$

Hence, using the stationarity of H_ω, for any $n > 0$, we have

$$\mathbf{E}\Big(\sum_{|\gamma|\leq n} \operatorname{tr}(\chi_{C_\gamma} E_\omega([E-\varepsilon,\ E+\varepsilon])\chi_{C_\gamma})\Big) = 0.$$

Then, using the positivity of the trace, we have, for almost every ω, for any $\gamma \in \Gamma$,

$$\operatorname{tr}(\chi_{C_\gamma} E_\omega([E-\varepsilon,\ E+\varepsilon])\chi_{C_\gamma}) = 0.$$

Hence, for almost every ω, for any $\gamma \in \Gamma$, we have

$$\|\chi_{C_\gamma} E_\omega([E-\varepsilon,\ E+\varepsilon])\| = 0.$$

So, for almost every ω, $E_\omega([E-\varepsilon,\ E+\varepsilon]) = 0$, thus, $E \notin \Sigma$. This proves

$$\Sigma \subset \operatorname{supp}(dN). \qquad \blacksquare$$

From Theorems 3.3 and 4.1, we obtain a characterization of Σ using periodic approximations of H_ω, namely, the following:

Theorem 4.2. [26] *Σ is the closure of the union of the spectra of the operators $H_{\omega,n}$ (defined by (3.5)) for $n \in \mathbf{N}$ and $\omega \in \operatorname{supp}\mathbf{P}$, the support of the probability measure (i.e., the infinite product of the supports of the individual ω_γ).*

Proof. Denote the spectrum of $H_{\omega,n}$ by $\sigma_{\omega,n}$. By formula (2.3), we see that, for periodic operators, the spectrum is the support of the density of states. Then, by Theorem 3.3, we have

$$\Sigma \subset \overline{\bigcup_{\omega,n\in\mathbf{N}} \sigma_{\omega,n}} = \Sigma'$$

as

$$\bigcup_{\omega} \sigma_{\omega,n} \subset \bigcup_{\omega} \sigma_{\omega,2n}. \tag{4.1}$$

Conversely, fix ω^0 and n_0 as in Theorem 4.2 and take $E \in \sigma_{\omega^0,n_0}$. Then, for any $\varepsilon > 0$, there exists $\varphi_\varepsilon \in L^2(\mathbf{R}^d)$ such that $\|\varphi_\varepsilon\| = 1$ and $\|(H_{\omega^0,n_0} - E)\varphi_\varepsilon\| \leq \varepsilon$. Pick $m \in \mathbf{N}$. Define the event

$$\Omega_{m,\varepsilon} = \left\{\omega;\ \forall |\gamma| \leq 2mn_0 + m + n_0,\ |\omega_\gamma - \omega^0_{[\gamma]}| \leq \frac{\varepsilon}{\|V\|^2_{1,\infty}}\right\}$$

where $[\gamma] \equiv \gamma \bmod ((2n+1)\Gamma)$. Then, for any $m \in \mathbf{N}$ and $\varepsilon > 0$, if ω^0 is in the support of $\mathbf{P}$, then $\mathbf{P}(\Omega_{m,\varepsilon}) > 0$. Moreover, for $\omega \in \Omega_{m,\varepsilon}$, we compute

$$\begin{aligned}\|(H_\omega - H_{\omega^0,n_0})\varphi_\varepsilon\|^2 &\leq \sum_{\gamma\in\Gamma} |\omega_\gamma - \omega_\gamma^0| \|V_\gamma\|_\infty^2 \|\varphi_\varepsilon\|^2 \\ &\leq \sum_{|\gamma|\leq 2mn_0+m+n_0} \frac{\varepsilon}{\|V\|_{1,\infty}} \|V_\gamma\|_\infty^2 \sum_{|\gamma|>2mn_0+m+n_0} C\|V_\gamma\|_\infty^2 \\ &\leq 2\varepsilon\end{aligned}$$

for m large enough as $\sum_{\gamma\in\Gamma} \|\tau_\gamma V\|_\infty < +\infty$.

Hence, for m large enough and $\omega \in \Omega_{m,\varepsilon}$, we have

$$\|(H_\omega - E)\varphi_\varepsilon\| \leq 3\varepsilon.$$

This implies $\sigma(H_\omega) \cap [E - 3\varepsilon, E + 3\varepsilon] \neq \emptyset$. Then, as $\sigma(H_\omega)$ is almost surely constant and as $\mathbf{P}(\Omega_{m,\varepsilon}) > 0$, we know that, ω almost surely, $\sigma(H_\omega) \cap [E - 3\varepsilon, E + 3\varepsilon] = \Sigma \cap [E - 3\varepsilon, E + 3\varepsilon] \neq \emptyset$. This holds for any $\varepsilon > 0$. So $E \in \Sigma$ as Σ is closed. ∎

One can also give characterizations of Σ in terms of other approximations of H_ω. The details of these constructions may be found in [26, 52].

5. Localization

We will now discuss the question of the spectral type of the random operator H_ω, that is, the question of the location of the sets Σ_{ac}, Σ_{sc} and Σ_{pp}. The results that we will describe can be summed up as follows: Under the adequate conditions on the random variables $(\omega_\gamma)_{\gamma\in\Gamma}$, at the edges of the spectrum of H_ω, the spectrum is a pure point spectrum and the associated eigenfunctions associated to this pure point spectrum are exponentially decaying. Another regime where we know that this localization phenomenon occurs is the regime of large randomness, i.e., if the fluctuation of V_ω is very large.

So we see that little is known about the nature of the spectrum of H_ω and one of the main open questions is the existence of spectral types other than pure point and the transition between the different existing spectral types.

Localization at spectral edges was first conjectured by physicists in the late 1950s and the beginning of the 1960s (see [1, 40]). Although the models they considered were not exactly the one considered here (they were either discrete or of another type of randomness), the heuristic picture they developed may also be applied in the present case.

Let us now briefly explain the physical heuristics suggesting location at band edges. For the sake of simplicity, we will assume $H_0 = -\Delta$ and V_ω is non-negative. Let us assume $\Sigma = [0, +\infty)$. Let $E > 0$ be small. Lifshits' argument (see [40]) is that, in a finite box, a state of energy less than E has to have a kinetic energy less than E. Hence, the state must spread out in space. On the other hand, if the state is spread out in space, then its expectation on the random potential is approximately the empirical average of a

large number of random variables. The number of random variables that one averages is roughly the size of the support of the state. By the large deviation theory, the probability that the empirical average is small (i.e., greatly differs from the mathematical expectation of random variables) is very small. One computes that it is exponentially small in the volume of the support of the eigenstate. If one optimizes over the possible sizes of this support, one obtains this volume to be roughly of size $E^{-d/2}$. So there is only a very small probability that two states of energy less than E live in the same box of size E^l (for some large l). Hence, there will be almost no tunneling between different states as these will be living too far apart. Hence, the picture of the spectrum one obtains in this way is the following: pure point spectrum associated to eigenfunctions that live far away from each other. This heuristic will find its mathematical realization in the Lifshits tail asymptotic for the density of states described below (see Subsec. 5.1) and in the localization shown to occur in this regime (see Subsec. 5.3).

In the case of large randomness, the other heuristic due to P. W. Anderson goes as follows: In the case of large fluctuations of V_ω, the probability that V_ω (in some finite box) assumes a certain value is very small; hence, the probability that H_ω (in some finite box) has given value E as an eigenvalue is very small. So it is very unlikely that two close values of energy are eigenvalues for H_ω (in the same box of fixed size). Hence, again there should be almost no tunneling and, when one goes over from a finite box to the whole space, energy values that were eigenvalues should remain eigenvalues. The general estimates of the probability that a given energy is an eigenvalue for H_ω restricted to some box are called Wegner estimates (see Subsec. 5.2). To obtain information on the operator in the whole space from information on the same operator in finite boxes, one uses a multiscale analysis (essentially due to Fröhlich and Spencer [18]) which will be briefly described (see Subsec. 5.3).

In the last ten years, localization has been extensively studied mathematically for a large number of different models and regimes. We will not review all this material here. This is far beyond the scope of this elementary introduction.

5.1. The Band-edge Asymptotics for the Density of States

Assume $H = -\Delta$, V_ω is non-negative and $\Sigma = [0, +\infty)$. Then the heuristics described above lead Lifshits to conjecture the following behavior for the asymptotics of the tail of the density of states of $H_\omega = H + V_\omega$ (see (3.1)):

$$N(E) - N(0) \underset{\substack{E \to 0 \\ E > 0}}{\sim} C_1 e^{-C_2 E^{-d/2}}. \tag{5.1}$$

Mathematical proofs of a slightly less precise version (one obtains the asymptotic behavior of $\log(N(E) - N(0))$) of (5.1) have been first given by Pastur [47, 48, 50] (see also [52] for more references) for models different from ours. To prove their asymptotic formula, these authors did not follow the original Lifshits heuristic; they used the Feynman-Kac formula (or Wiener integrals) to estimate the large t behavior of the Laplace transform of the density of states (see also [61, 62]).

The first proofs of (5.1) following the ideas of Lifshits are due to Kirsch and Martinelli [27]. The versions of (5.1) proved in those cases are very rough as one only obtains results on the $\log|\log(N(E) - N(0))|$. Lifshits tails are proved to occur at the bottom of the spectrum of an operator of type (3.1) when $W = 0$. Using the same kind of techniques,

the behavior of the integrated density of states at the bottom of the spectrum has been studied quite extensively for a variety of different continuous and discrete models [28, 45, 57]. In the continuous case of dimension one, and for some models in the discrete case, Lifshits tails have also been proved at internal bands of spectrum [46, 58].

More recently, Klopp recovers many of the previously known results as particular cases [35]. Although the heuristics used to obtain these new results are those of Lifshits, the mathematical technique is quite different from the one used in the papers cited above. We will now describe these results.

Let $\sigma = \sigma(H)$ be the spectrum of H and assume σ has a gap below energy 0 of length at least δ, that is,

- for some $a > 0$ and $\delta > 0$, $\sigma \cap [0, a) = [0, a)$ and $\sigma \cap [-\delta, 0) = \emptyset$.

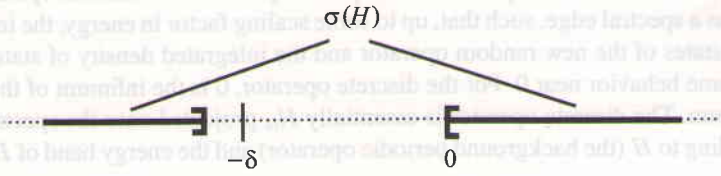

Fig. 1. A part of the spectrum of H.

We will also need some additional assumptions on the random potential V_ω, namely,

- V is a non-negative, bounded and compactly supported potential;
- $(\omega_\gamma)_{\gamma\in\Gamma}$ is a collection of independent identically distributed random variables that have $[0, \omega^+]$ as their common support ($\omega^+ > 0$);
- $\limsup_{\varepsilon\to 0^+} \frac{\log|\log \mathbf{P}(\{\omega_0 \le \varepsilon\})|}{|\log \varepsilon|} = 0$.

Let $\Sigma = \sigma(H_\omega)$ be the almost sure spectrum of H_ω. By Theorem 4.2, we know that 0 belongs to Σ. We assume Σ has a gap below 0, that is,

- for some $\delta' > 0$, $\Sigma \cap [-\delta', 0) = \emptyset$.

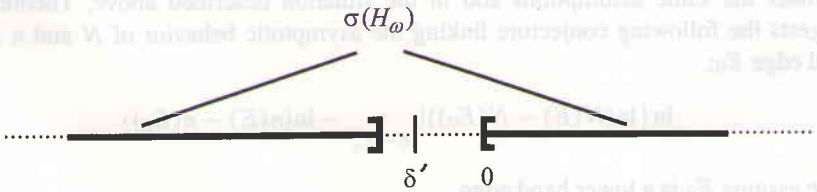

Fig. 2. A part of the spectrum of H_ω.

Let $n(E)$ be the integrated density of states of H and $N(E)$ the integrated density of states of H_ω. Then, we obtain

Theorem 5.1. [35] *Under the assumptions given above, we have*

- *0 is a continuity point for $N(E)$;*

- *under no further assumption on n, one has*

$$\liminf_{E\to 0^+} \frac{\log|\log(N(E)-N(0))|}{\log E} \geq -\frac{d}{2};$$

- *one has the equivalence*

$$\lim_{E\to 0^+} \frac{\log|\log(N(E)-N(0))|}{\log E} = -\frac{d}{2} \Longleftrightarrow \lim_{E\to 0^+} \frac{\log(n(E)-n(0))}{\log E} = \frac{d}{2}.$$

Remark 5.1. We will not give a proof of this result here as it can be found in [35]. Here, we will just say a few words about it. There are two key ingredients to this proof. First, we shall do a cut-off in energy for the density of states. More precisely, this means that we shall construct a random operator (more precisely, a discrete random operator) that also has 0 as a spectral edge, such that, up to some scaling factor in energy, the integrated density of states of the new random operator and the integrated density of states of H_ω have the same behavior near 0. For the discrete operator, 0 is the infimum of the almost sure spectrum. The discrete operator is essentially H_ω projected onto the spectral space corresponding to H (the background periodic operator) and the energy band of H staring at 0.

The second idea is that, if n satisfies (5.2), then the discrete operator is essentially a direct sum of discrete Anderson models for which the behavior of the Lifshits tail is already known [44, 57].

The result described in Theorem 5.1 is valid in a much more general setting. Let us also say that this result is valid for upper edges of the spectrum if one properly symmetrizes the assumptions on $(\omega_\gamma)_\gamma$ and H.

It is well known [37] that at the bottom of the spectrum of H, its integrated density of states $n(E)$ satisfies

$$\lim_{E\to 0^+} \frac{\log(n(E)-n(0))}{\log E} = \frac{d}{2}. \tag{5.2}$$

This enables us to recover the results of [27], [28] and [45]. In dimension one, (5.2) is well known at any edge of the spectrum of a periodic Schrödinger operator [17]. Then, Theorem 5.1 gives the results of [46].

Under the same assumptions and in the situation described above, Theorem 5.1 suggests the following conjecture linking the asymptotic behavior of N and n near a band edge E_0:

$$\ln|\ln(N(E)-N(E_0))| \underset{E\to E_0^+}{\sim} -\ln(n(E)-n(E_0))$$

if we assume E_0 is a lower band edge.

Such a formula was proved in general for one-dimensional discrete random Schrödinger operators [36]. We expect it to be correct in a more general setting [30].

5.2. The Wegner Estimate

Let $n \in \mathbf{N}^*$ and consider the operator $H^n_{\omega\restriction D}$ defined by

$$H^n_{\omega\restriction D} = (H_\omega)_{\restriction C_n},$$

i.e., the operator H_ω^n restricted to the cube $C_n = (2n+1)C_0$ with Dirichlet boundary conditions [54]. This operator is self-adjoint on $L^2(C_n)$ with domain $H_0^2(C_n)$. It has a compact resolvent, hence, its spectrum is discrete.

A so-called *Wegner estimate* is an estimate on

$$N_n((E_0-\varepsilon,\ E_0+\varepsilon]) = \mathbf{E}(\#\{E;\ E \text{ is an eigenvalue of } H^n_{\omega\restriction D} \text{ and } E \in (E_0-\varepsilon,\ E_0+\varepsilon]\})$$

for n large and ε small (E_0 is some fixed energy). As we already know, when renormalized by the volume of C_n, this quantity converges to the integrated density of states at $E_0+\varepsilon$ minus the integrated density of states at $E_0-\varepsilon$. So one expects that it grows with C_n and decreases with ε (if the density of states is assumed to be regular). The main problem is then to obtain an estimate that controls the respective increase in n and decrease in ε.

For the sake of simplicity, let us assume the potential V has compact support in some cube centered at 0 of length side less than 1. So

$$H^n_{\omega\restriction D} = (-\Delta)_{\restriction D} + (W + V_\omega)\mathbf{1}_{C_n}.$$

where $\mathbf{1}_{C_n}$ is the characteristic function of the cube C_n.

$H^n_{\omega\restriction D}$ only depends on a finite-dimensional space of random parameters. Let us denote these finitely many parameters by $\omega = (\omega_\gamma)_{\gamma\in\Gamma/(2n+1)\Gamma}$. Let $(E_j(\omega,n))_{j=1}^N$ be the eigenvalues of $H^n_{\omega\restriction D}$ that lie in $[E_0-\varepsilon,\ E_0+\varepsilon]$ for some possible choice of ω, that is, for some ω in the support of the probability measure defined by the initial random variables (here, $\varepsilon_0 > 0$ is fixed). The existence of the density of states guarantees that their number N is of order of the volume of C_n.

We can write

$$N_n((E_0-\varepsilon,\ E_0+\varepsilon]) = \mathbf{E}\Big(\sum_{1\le j\le N} \mathbf{1}_{(E_0-\varepsilon,\ E_0+\varepsilon]}(E_j(\omega,n))\Big)$$
$$\le \sum_{1\le j\le N} \mathbf{P}\{E_j(\omega,n) \in (E_0-\varepsilon,\ E_0+\varepsilon]\},$$

(here, $\mathbf{P}\{E\}$ is the probability of the event E).

So we need to estimate $\mathbf{P}\{E_j(\omega,n) \in (E_0-\varepsilon,\ E_0+\varepsilon]\}$. The mapping $\omega \mapsto E_j(\omega,n)$ realizes a projection from the parameter space onto the real axis, and we would like to measure the size (with respect to the probability measure on the parameter space) of the preimage of some interval. The idea is then to find $\mathcal{V}$ a vector field in the parameters ω such that the eigenvalue $E_j(\omega,n)$ moves when ω moves along the flow of the vector field. The flow of $\mathcal{V}$ foliates the parameter space nicely and the volume we want to measure is that contained in a layer between two leaves (see Fig. 3). This volume will then be of the width of this layer at least when the probability measure has a nice density. So if one is able to do this for all the eigenvalues, one gets an estimate of the form

$$N_n((E_0-\varepsilon,\ E_0+\varepsilon]) \le C\varepsilon \mathrm{Vol}(C_n). \tag{5.3}$$

To be able to do this simultaneously for all eigenvalues, one may choose $\mathcal{V}$ so that H_ω^n derived along $\mathcal{V}$ has nice properties (e.g., positivity). Let us give a simple example. Assume $V \ge \alpha \mathbf{1}_{C_0}$ (for some $\alpha > 0$). If we take $\mathcal{V} = \mathrm{div}_\omega = \sum_{\gamma\in C_n} \frac{\partial}{\partial\omega_\gamma}$, then $\mathcal{V}H_\omega^n \ge \alpha I_{C_n}$. This ensures $\mathcal{V}E_j(\omega,n) \ge \alpha$. Such an estimate immediately gives (5.3).

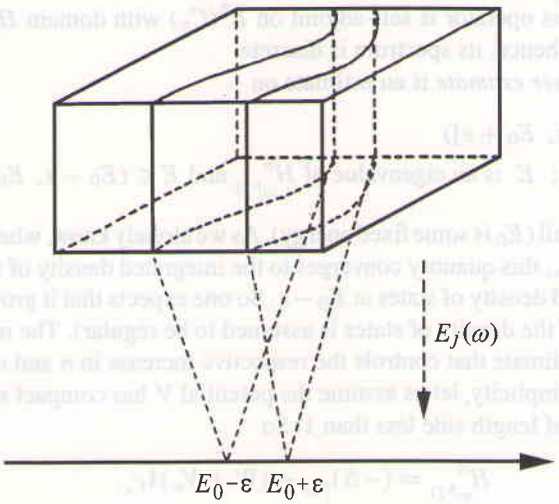

Fig. 3. Foliation and projection of the probability space.

In singular cases like the Bernoulli case, the probability measure is supported by discrete points in the parameter space. To be able to derive (5.3), we need to compute the number of points in the given layer; this is a much harder problem, as it requires knowing more or less exactly the nature of the mappings $(\omega \mapsto E_j(\omega, n))_{1\leq j\leq N}$.

We saw that, to derive (5.3), one needs to assume that the random variables $(\omega_\gamma)_\gamma$ admit a density. In this case, one can define the *randomness* of the system to be the inverse of the supremum norm of this density. The randomness roughly measures the proportionality between the Lebesgue measure and the probability measure in the sense that, for a set of $(\omega_\gamma)_{\gamma\in C_n}$ of a given Lebesgue measure, the larger the randomness, the smaller the probability of this set. From this, one deduces that the constant C in (5.3) can be taken inversely proportional to the randomness of the system.

The right choice of vector field to prove (5.3) is model dependent. For the Anderson model, discrete or continuous, in many cases, one may use the divergence vector field (for example, see [8, 23, 33, 64]). Another useful vector field with respect to this problem is the generator of the dilations $\sum \omega_\gamma \partial\omega_\gamma$; it can be used to obtain a Wegner estimate for the continuous Anderson model without sign assumptions on V [34]. Different types of randomness may require different types of vector fields (e.g., [31, 60]).

For the continuous Anderson model, another more elegant way (but nevertheless based on the philosophy described above) of proving Wegner estimates has been developed in [10].

5.3. Band-edge and Large Randomness Localization

We will now describe the phenomenon of localization. We will first describe the physical heuristics that underlie the understanding of this phenomenon. We will then state some mathematical results. We will not prove any of these localization results as one of the

main technical tools needed to prove localization is the multiscale analysis which is long and complicated.

Physically, localization is roughly the trapping of the electron inside the bulk of solid state matter. The natural interpretation for this is that the wave function of the electron does not spread in space as time evolves. So, if time $t = 0$, the electron has a wave function φ. Then, at time t, one is interested in where the wave function $e^{itH}\varphi$ is localized in space. This is a dynamical property.

Mathematically, this can be interpreted in different ways. The closest to the physical point of view would be to study the time evolution of averages of the form e.g., $\langle x^2 e^{itH}\varphi, e^{itH}\varphi\rangle$ (where φ is some nicely localized function in $L^2(\mathbf{R}^d)$). One would speak of localization if, for some $C > 0$, one has $|\langle x^2 e^{itH}\varphi, e^{itH}\varphi\rangle| \leq C$ for all $t \geq 0$. In this case, one speaks of *dynamical localization*. For more information on dynamical localization, we refer the reader to [5, 15, 16].

There is also a spectral interpretation of localization: it is simply to say that, in some region, the spectrum is purely punctual and dense. This is called *Anderson localization*. One speaks of *exponential localization* when the eigenfunctions associated to eigenvalues in these parts of the spectrum are exponentially decaying at ∞.

The two notions described above are connected. Indeed, dynamical localization (in some energy region) implies Anderson localization (in the same energy region). From a mathematical point of view, more work has been devoted to Anderson localization than to dynamical localization. In the sequel, we will only speak of Anderson localization. Hence, the term "localization" will from now on be used to designate Anderson localization.

5.3.1. *Physical Heuristics*

Localization was first predicted by Anderson [1] for a discrete model (the Anderson model) in the large randomness regime. As we will see below, the main argument that speaks for localization is the absence of resonance, i.e., of tunneling between the different wells of the random potential. Using his work on Lifshits tails, Lifshits [41] gave a different argument to show that localization should occur at the band edges.

One of the main features of quantum systems is the possibility of tunneling and localization is mainly the symptom of the absence of tunneling in a quantum system. Let us explain what tunneling is about. Take two Schrödinger operators $H_1 = -\Delta + W_1$ and $H_2 = -\Delta + W_2$ where W_1 and W_2 are potential wells described in Fig. 4. For the sake of simplicity, let us assume $(W_i)_{i=1,2}$ are compactly supported. Assume the wells are sufficiently deep to bound at least one state (see [54]), i.e., for $i \in \{1, 2\}$, assume that for the well W_i, denote the state by φ_i and the corresponding energy by E_i. Then, by definition, we have $H_i\varphi_i = E_i\varphi_i$ for $i \in \{1, 2\}$. The functions $(\varphi_i)_{i=1,2}$ are localized near the wells $(W_i)_{i=1,2}$. More precisely, one knows that φ_i satisfies an inequality of the type

$$|\varphi_i(x)| \leq Ce^{-|x-x_i|/C}. \tag{5.4}$$

Then the question is: What happens when one puts the two wells together, or more precisely, can one deduce the bound states of the operator $H = -\Delta + W_1(x - x_1) + W_2(x - x_2)$ from the bound states of H_1 and H_2? To obtain this information, one

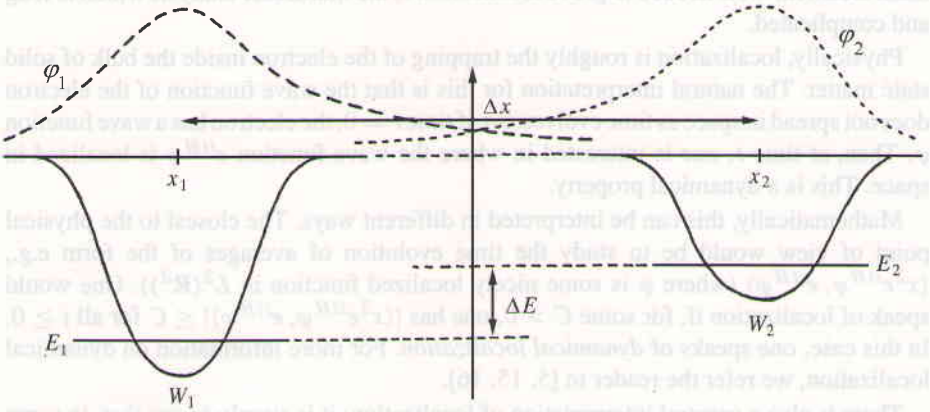

Fig. 4. Two distant wells.

projects H on the vector space spanned by $(\varphi_i)_{i=1,2}$, i.e., one looks at the 2×2 matrix

$$M = \begin{pmatrix} \langle \varphi_1, H\varphi_1 \rangle & \langle \varphi_1, H\varphi_2 \rangle \\ \langle \varphi_2, H\varphi_1 \rangle & \langle \varphi_2, H\varphi_2 \rangle \end{pmatrix}.$$

Because of (5.4), the localization properties of the eigenstates $(\varphi_i)_{i=1,2}$ and as H is a local operator, the off-diagonal terms $\langle \varphi_1, H\varphi_2 \rangle$ and $\langle \varphi_2, H\varphi_1 \rangle = \overline{\langle \varphi_1, H\varphi_2 \rangle}$ are very small, i.e.,

$$|\langle \varphi_1, H\varphi_2 \rangle| \leq Ce^{-|x_1 - x_2|/C} = Ce^{-\Delta x/C}, \tag{5.5}$$

where $\Delta x = |x_2 - x_1|$, i.e., the distance separating the two wells. On the other hand,

$$\langle \varphi_i, H\varphi_i \rangle = \langle \varphi_i, H_i\varphi_i \rangle + \langle \varphi_i, W_j\varphi_i \rangle = E_i + \langle \varphi_i, W_j\varphi_i \rangle,$$

where $\{i, j\} = \{1, 2\}$. Because of (5.4), $\langle \varphi_i, W_j\varphi_i \rangle$ also satisfies an estimate of the type (5.5). So M can be written as

$$M = \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix} + O(e^{-\Delta x/C}) = M_0 + O(e^{-\Delta x/C}).$$

Let us now study the eigenvalues and eigenvectors of M as Δx gets larger. Of course in this limit, the eigenvalues of M tend to E_1 and E_2 as the error tends to 0. To study the eigenvectors, we must distinguish two cases:

(1) If $E_1 \neq E_2$, then one easily computes that the two eigenvectors of M tend to $(1, 0)$ and $(0, 1)$, i.e., the eigenvectors of the matrix M_0. For our quantum system, this means that, when Δx is large, H has two eigenvalues, one close to E_1, the other to E_2. The eigenvectors associated to these two eigenvalues are respectively close to φ_1 and φ_2 (see Fig. 5).

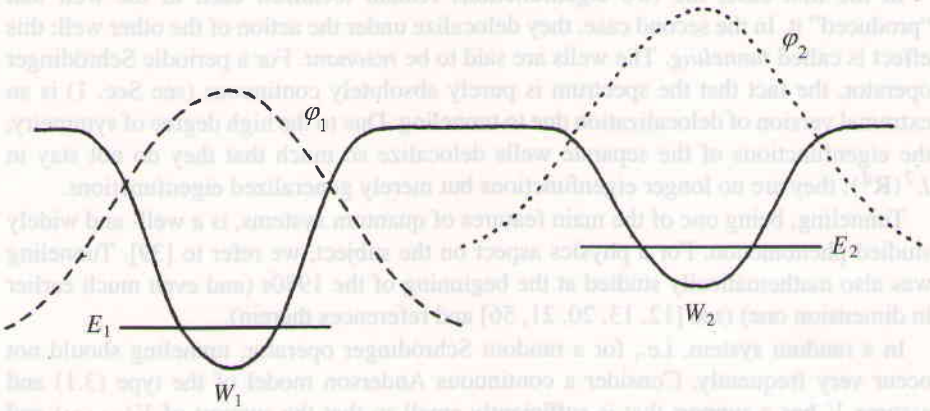

Fig. 5. The asymmetric double well.

(2) If $E_1 = E_2 =: E$, then the situation is more complicated. To simplify the discussion, let us assume the system is symmetric with respect to the plane mediator of x_1 and x_2. In this case, $\langle \varphi_i, W_j \varphi_i \rangle = \langle \varphi_j, W_i \varphi_j \rangle$ for $\{i, j\} = \{1, 2\}$. So we need only to study what happens to the eigenvectors of a matrix of the form

$$J_\varepsilon = \begin{pmatrix} 0 & \varepsilon \\ \varepsilon & 0 \end{pmatrix}$$

when $\varepsilon \to 0$. The fact that the limit matrix has an eigenvalue with multiplicity 2 introduces a degeneracy. One computes that the eigenvectors of J_ε equal $\frac{1}{\sqrt{2}}(1, \varepsilon/|\varepsilon|)$ and $\frac{1}{\sqrt{2}}(1, -\varepsilon/|\varepsilon|)$. Hence, for the quantum system, the eigenvalues are close to E and the eigenvectors to

$$\varphi_\pm = \frac{1}{\sqrt{2}}(\varphi_1 \pm \varepsilon/|\varepsilon| \varphi_2),$$

(see Fig. 6 (we assume $\varepsilon \in \mathbf{R}$)).

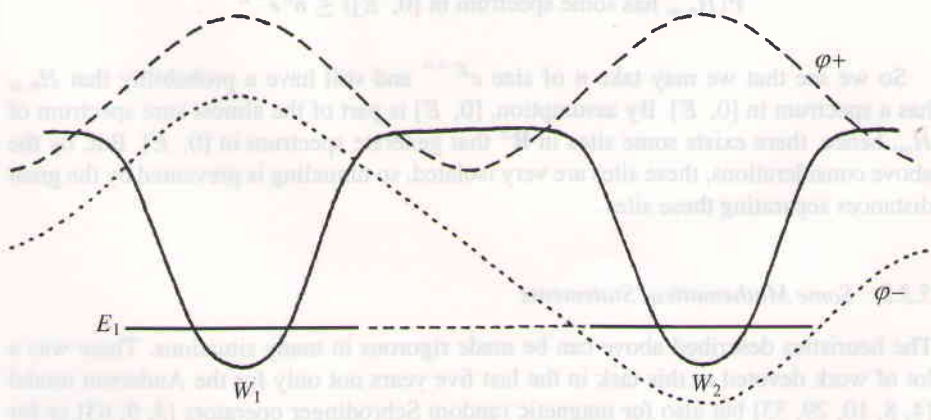

Fig. 6. The symmetric double well.

In the first case, the two eigenfunctions remain localized each in the well that "produced" it. In the second case, they delocalize under the action of the other well: this effect is called *tunneling*. The wells are said to be *resonant*. For a periodic Schrödinger operator, the fact that the spectrum is purely absolutely continuous (see Sec. 1) is an extremal version of delocalization due to tunneling. Due to the high degree of symmetry, the eigenfunctions of the separate wells delocalize so much that they do not stay in $L^2(\mathbf{R}^d)$; they are no longer eigenfunctions but merely generalized eigenfunctions.

Tunneling, being one of the main features of quantum systems, is a well- and widely studied phenomenon. For a physics aspect on the subject, we refer to [39]. Tunneling was also mathematically studied at the beginning of the 1980s (and even much earlier in dimension one) (see [12, 13, 20, 21, 56] and references therein).

In a random system, i.e., for a random Schrödinger operator, tunneling should not occur very frequently. Consider a continuous Anderson model of the type (3.1) and assume V has a support that is sufficiently small so that the support of $V(\cdot - \gamma)$ and $V(\cdot - \beta)$ intersect if and only if $\beta = \gamma$. Then one can consider V_ω as the "union" of infinitely many potential wells $\omega_\gamma V(\cdot - \gamma)$. One may now wonder when tunneling occurs. Of course, as one is dealing with infinitely many potentials, one should not only worry about tunneling between two wells but also between three or four wells, and more generally, between two groups of wells. This is exactly the sense of the Wegner estimate (5.3). Given an energy E_0, the Wegner estimate gives a control on the probability that tunneling occurs at energy E_0 (this in a given fixed cube). Remember that the constant C in (5.3) is inversely proportional to the randomness of the system. So it is immediate that, at large randomness, tunneling should be suppressed, hence, localization should occur. This is Anderson's argument.

Another case where localization should occur is at the band edges. In this case, the mechanism is a little more subtle. As in Subsec. 5.1, let us assume the edge of the spectrum is 0 and that the gap is at the left of 0. For positive energies close to 0, by Lifshits tails (5.1), we know that $N(E) - N(0)$ is very small. Let us now assume that it is well approximated by $\mathbf{E}(N_{\omega,n}(E)) - \mathbf{E}(N_{\omega,n}(0))$ (it actually is the case for $N_{\omega,n}$ defined before Theorem 3.3 (cf. [35])) where n is of size at least a power of E^{-1}. The Bienaymé–Tchebycheff inequality then tells us that,

$$\mathbf{P}(H_{n,\omega} \text{ has some spectrum in } [0,\ E]) \leq n^d e^{-E^{-d/3}}.$$

So we see that we may take n of size $e^{E^{-d/4}}$ and still have a probability that $H_{n,\omega}$ has a spectrum in $[0,\ E]$. By assumption, $[0,\ E]$ is part of the almost sure spectrum of H_ω, hence, there exists some sites in $\mathbf{R}^d$ that generate spectrum in $[0,\ E]$. But, by the above considerations, these sites are very isolated, so tunneling is prevented by the great distances separating these sites.

5.3.2. Some Mathematical Statements

The heuristics described above can be made rigorous in many situations. There was a lot of work devoted to this task in the last five years not only for the Anderson model [4, 8, 10, 29, 33] but also for magnetic random Schrödinger operators [3, 9, 63] or for random translation [31]. We will only describe two rigorous results (taken from [34]) as examples of the two asymptotics we have described above.

We assume V is a real bounded function with compact support. We also assume the i.i.d. random variables $(\omega_\gamma)_{\gamma\in\Gamma}$ have common distribution density g satisfying
(a) $\exists\epsilon_0 > 0$ and $\rho_0 > 0$ such that, $\forall\epsilon \in [0, \epsilon_0]$,

$$\int_{\mathbf{R}} |g((1+\epsilon)t) - g(t)|dt \leq \left(\frac{\epsilon}{\epsilon_0}\right)^{\rho_0}.$$

(b) Let $q_0 = \frac{d}{2}$ if $d \geq 4$ and $q_0 = 2$ if $d \leq 3$. $\exists k > q_0$ such that

$$\int_{\mathbf{R}} |t|^k g(t)dt < +\infty.$$

Remark 5.2. Assumption (a) is a regularity assumption on g that is, for example, satisfied if g is differentiable and $\int |x.g'(x)|dx < +\infty$.

W being defined as in Sec. 1, let H_ω be defined by (3.1). One then proves the following:

Theorem 5.2. [34] *Let H_ω be defined as above and assume Σ, the almost sure spectrum of H_ω, satisfies $\Sigma = \mathbf{R}$. Then, for any $m > 0$, there exists $E_m > 0$ such that, with probability 1,*

- *the spectrum of H_ω in $(-\infty, -E_m]$ is pure point;*
- *if φ is an eigenvector associated with E, an eigenvalue of H_ω in $(-\infty, -E_m]$, then there exists $C_\varphi > 0$ such that, for any $x \in \mathbf{R}^d$,*

$$|\varphi(x)| \leq C_\varphi e^{-m|x|}.$$

Let us comment on this result. In Theorem 5.2, we prove that, in some neighborhood of $-\infty$, the spectrum of H_ω is purely punctual and exponentially localized. We are not exactly in the large randomness limit, but the decay of the density of the random variables at $-\infty$ simulates this large randomness limit. Indeed, negative energies in the spectrum of H_ω can only be created by negative values of the random variables (ω_γ) (as H is lower bounded). So, as g roughly tends to 0 near $-\infty$, the randomness becomes arbitrarily large.

We also notice that, as the single site potential V is compactly supported, the decay rate of the eigenfunctions gets asymptotically faster than any exponential as the energy decreases to $-\infty$.

We take V as above. We assume the i.i.d. random variables $(\omega_\gamma)_{\gamma\in\Gamma}$ have common distribution density g satisfying

- $\exists\epsilon_0 > 0$ and $\rho_0 > 0$ such that, $\forall\epsilon \in [0, \epsilon_0]$,

$$\int_{\mathbf{R}} |g((1+\epsilon)t) - g(t)|dt \leq \left(\frac{\epsilon}{\epsilon_0}\right)^{\rho_0}.$$

- G, the essential support of g, is bounded.

Under these assumptions, we know that Σ, the almost sure spectrum of H_ω, is lower bounded. Define $E_{\inf} = \inf\Sigma$. One proves the following:

Theorem 5.3. [34] *Assume H_ω is defined as above. Assume the above-described assumptions are satisfied. Then there exists $E_0 > E_{\inf}$ and $m > 0$ such that, with probability 1,*

- *the spectrum of $H(t)$ in $[E_{\inf}, E_0]$ is pure point;*
- *if φ is an eigenvector associated with E, an eigenvalue of H_ω in $[E_{\inf}, E_0]$, then there exists $C_\varphi > 0$ such that, for any $x \in \mathbf{R}^d$,*

$$|\varphi(x)| \leq C_\varphi e^{-m|x|}.$$

The above result is not exactly contained in [34]; it is easily obtained from [34] and the results on Lifshitz tails given in Theorem 5.1.

We will not prove any of these results here. The main ingredient needed to mathematically implement the heuristics described above is the multiscale analysis "a la Fröhlich-Spencer" [18]. Many versions of this analysis have been developed [4, 8, 23, 29, 34, 42]; but they are all built around the same idea: to recover the behavior of the resolvent (at some fixed energy or in some energy range) of H_ω restricted to a large cube from the resolvents of H_ω restricted to smaller cubes (covering the large cube). This can be done, for instance, using a resolvent identity of the following form: Let $\Lambda_l \subset \Gamma$ be the cube of center 0 and side length $2l$ (i.e., $\Lambda_l = \left\{ \sum_{1 \leq j \leq d} x_j e_j \in \Gamma;\ -l \leq x_j \leq l \right\}$) and define

$$H_{l,\omega} = H + \sum_{\gamma \in \Delta_l} \omega_\gamma V_\gamma.$$

For $E \notin \sigma(H_{l,\omega})$, we define

$$G_{l,\omega}(E) = (H_{l,\omega} - E)^{-1}.$$

Let $l \leq l'$ and let $\Lambda_l \subset \Lambda_{l'}$ be two cubes of Γ. Then, for $E \notin \sigma(H_{l,\omega}) \cup \sigma(H_{l',\omega})$, we obtain the following resolvent formula:

$$G_{l',\omega}(E) = G_{l,\omega}(E) + G_{l,\omega}(E) V_{l,l'} G_{l',\omega}(E), \tag{5.6}$$

where

$$V_{l,l'} = \sum_{\gamma \in \Lambda_{l'} \setminus \Lambda_l} \omega_\gamma V_\gamma.$$

From (5.6), one sees that one can transfer the decay of the kernel of the resolvent of the small cube to the kernel of the resolvent of the large cube if one knows *a priori* that the resolvent on the large cube is not too big. This is given by the Wegner estimate (5.3). The reader will find the details of the multiscale analysis in the references cited above.

6. Appendix

In the appendix, we shall briefly recall some basic notions from the probability and spectral theory needed for this exposition. For a thorough treatment of this material, we refer to [6, 7, 14, 19, 24, 25, 52, 54].

6.1. Some Notions from the Probability Theory

Let $(\tilde{\Omega}, \mathcal{A}, \tilde{P})$ be a probability space ($\tilde{\Omega}$ is the state space, $\mathcal{A}$ the σ-algebra of events and $\tilde{P}$ the probability measure). $X : \tilde{\Omega} \to \mathbf{R}$ is a *random variable* if it is Borel measurable (where $\mathbf{R}$ is endowed with the σ-algebra of Borel sets denoted by $\boldsymbol{B}(\mathbf{R})$). X induces a probability measure P_X on $(\mathbf{R}, \boldsymbol{B}(\mathbf{R}))$ by $P_X(B) = \tilde{P}(X \in B)$. P_X carries all interesting information on X. P_X is called the *distribution* of X.

Let $(X_i)_{i\in I}$ be a collection of real-valued random variables on $\tilde{\Omega}$. The $(X_i)_{i\in I}$ are *independent* if and only if, for any $n \geq 1$, $B_1, B_2, ..., B_n$ Borel sets and $i_1, i_2, ..., i_n$ in I, we have

$$\tilde{P}(\{X_{i_1} \in B_1, X_{i_2} \in B_2, ..., X_{i_n} \in B_n\}) = \prod_{j=1}^{n} \tilde{P}(X_{i_j} \in B_j). \tag{6.1}$$

The left-hand side of (6.1) is called a *finite-dimensional distribution* of the collection of random variables $(X_i)_{i\in I}$ (even in the case when the $(X_i)_{i\in I}$ are not independent). As in the case of a single random variable, we will only be interested in the finite-dimensional distribution of $(X_i)_{i\in I}$. This enables us to change the underlying probability space (as long as we keep the same distribution). The example we are interested in is the case when $(X_i)_{i\in I}$ are independent and identically distributed (i.e., $\tilde{P}(X_j \in B) = \tilde{P}(X_k \in B)$ for any Borel set B and $(j, k) \in I^2$). Let P_X be the common distribution of $(X_i)_{i\in I}$. Define the probability space $(\Omega, \mathcal{C}, P)$ by

- the state space Ω is $\mathbf{R}^I$;
- the σ-algebra is the σ-algebra generated by the cylinders of Borel sets in $\mathbf{R}$, i.e., it is generated by sets of the form $\prod_{i\in I} B_i$ where, for a finite set of indices in I, B_i is some Borel set in $\mathbf{R}$ and for all other indices, $B_i = \mathbf{R}$;
- P is the product measure $\otimes P_X$.

Any point $\omega \in \Omega$ can be written as $\omega = (\omega_i)_{i\in I}$. For every $i \in I$, $\omega \mapsto \omega_i$ defines a random variable. The collection of random variables $(\omega_i)_{i\in I}$ has the same finite-dimensional distribution as the collection of random variables $(X_i)_{i\in I}$. So one can always replace a collection of i.i.d. random variables by this representation. This is what we did when defining our model (3.1).

6.1.1. Ergodicity and the Birkhoff–Khinchine Theorem

Let $S : \tilde{\Omega} \to \tilde{\Omega}$. S is *measure preserving* if, for any $B \in \mathcal{A}$, $\tilde{P}(S^{-1}B) = \tilde{P}(B)$.

Let us now assume I is a non-degenerate lattice in $\mathbf{R}^d$ that will be denoted by Γ as above. We define $(t_i)_{i\in\Gamma}$, the shifts on Ω by (3.2). These mappings define an Abelian group (under the composition law) isomorphic to the lattice Γ. If the shifts are measure preserving, one says that the random variables $(X_i)_{i\in\Gamma}$ are *stationary*.

Let us now assume the random variables $(X_i)_{i\in\Gamma}$ are stationary. An event $B \in \mathcal{A}$ is invariant under the shift if, for any $i \in \Gamma$, $t_i^{-1}(B) = B$. The shifts $(T_i)_{i\in\Gamma}$ are said to be *ergodic* if any invariant set has measure 0 or 1.

One proves that a collection of i.i.d. random variables indexed by a lattice is ergodic with respect to the shift group (i.e., the shift group is ergodic). Ergodicity is a property of the probability measure $\tilde{P}$.

A well-known consequence of ergodicity is that, if the shifts $(T_i)_{i\in\Gamma}$ are ergodic on $(\tilde{\Omega}, \mathcal{A}, \tilde{P})$, then the random variables invariant under the shifts are almost surely constant.

One of the most important results in ergodic theory is the Birkhoff–Khinchine theorem. It is a version of the strong law of large numbers for an ergodic collection of random variables.

Theorem 6.1. *Let $(X_i)_{i\in\Gamma}$ be an ergodic collection of random variables such that $\mathbf{E}(|X_0|) < +\infty$. Then, $\tilde{P}$ almost surely,*

$$\frac{1}{(2n+1)^d} \sum_{|i|\leq n} X_i \underset{n\to+\infty}{\longrightarrow} \mathbf{E}(X_0).$$

By the above remark, the Birkhoff–Khinchine theorem holds for a collection of i.i.d. random variables (indexed by a lattice).

6.2. Some Notions from the Spectral Theory of Unbounded Linear Operators

Let $\mathcal{H}$ be a Hilbert space and $H : \mathcal{D}(H) \to \mathcal{H}$ an unbounded linear operator with dense domain $\mathcal{D}(H)$. $\overline{H}$ is an *extension* of H if $\mathcal{D}(H) \subset \mathcal{D}(\overline{H})$ and for $f \in \mathcal{D}(H)$, $\overline{H} f = Hf$. One writes $H \subset \overline{H}$.

$z \in \mathbf{C}$ is in the *spectrum* of H if $z - H : \mathcal{D}(H) \to \mathcal{H}$ is not invertible. The spectrum of H is denoted by $\sigma(H)$. The *resolvent set* is the complementary of the spectrum, i.e., the set of points $z \in \mathbf{C}$ where $z - H$ is invertible. The resolvent set is open and the spectrum is closed. For z, a point in the resolvent set, the *resolvent* of H at z is $(z - H)^{-1}$, the inverse of $z - H$. One defines H^*, the adjoint of H, by

$$\mathcal{D}(H^*) = \{f \in \mathcal{H};\ \exists g \in \mathcal{H},\ \forall h \in \mathcal{D}(H),\ \langle Hh, f\rangle = \langle h, g\rangle\}$$
$$H^* f = g.$$

If $\mathcal{D}(H^*)$ is dense, then $\overline{H} = H^{**} = (H^*)^*$. One says that H is *self-adjoint* iff $H^* = H$, i.e., iff $\mathcal{D}(H^*) = \mathcal{D}(H)$ and $H^* f = Hf$ for any $f \in \mathcal{D}(H)$.

An operator H is *symmetric* if $\forall f, g \in \mathcal{D}(H)$, $\langle Hg, f\rangle = \langle g, Hf\rangle$. For H symmetric, we have $H \subset H^{**} \subset H^*$. A symmetric operator is said to be *essentially self-adjoint* if it admits a unique self-adjoint extension. For example, the differential operator $-\Delta$ is essentially self-adjoint on the domain $\mathcal{C}_0^\infty(\mathbf{R}^d)$; its closure is the differential operator $-\Delta$ on the domain $H^2 = \{f \in L^2(\mathbf{R}^d);\ \Delta f \in L^2(\mathbf{R}^d)\}$ (to see this, one can use the Fourier transformation).

An operator H is said to be *semi-bounded below* or *lower bounded* by C iff

$$\forall f \in \mathcal{D}(H),\ \langle Hf, f\rangle \geq C\|f\|^2.$$

If $C = 0$, one says that H is *positive*.

A well-known result by K. Friedrichs states that every non-negative symmetric operator H has at least one non-negative self-adjoint extension. This immediately extends to lower bounded operators.

Let H be a self-adjoint operator with dense domain $\mathcal{D}(H)$ and let B be a symmetric operator whose domain contains $\mathcal{D}(H)$. Then B is said to be *relatively bounded* with respect to H with relative bound $b \geq 0$ if, for some $c \in [0, +\infty)$, we have, $\forall f \in \mathcal{D}(H)$, $\|Bf\| \leq b\|Hf\| + c\|f\|$. Note that a symmetric bounded operator is relatively bounded with respect to any self-adjoint operator.

The Kato–Rellich theorem states that, if H is self-adjoint and B is relative bounded with respect to H with relative bound $b < 1$, then the operator $K := H + B$ is self-adjoint with domain $\mathcal{D}(K) = \mathcal{D}(H)$.

6.2.1. The Spectral Theorem

A function on $\boldsymbol{B}(\mathbf{R})$ with values in the space of orthogonal projections on $\mathcal{H}$ will be called a *resolution of the identity* on $\mathbf{R}$ if and only if

(1) $E(\mathbf{R}) = Id$, the identity of $\mathcal{H}$;
(2) for any f, g in $\mathcal{H}$, $B \in \boldsymbol{B}(\mathbf{R}) \mapsto \langle E(B)f, g\rangle = E_{f,g}(B)$ defines a complex Borel measure on $\mathbf{R}$.

To E, a resolution of the identity on $\mathbf{R}$ and a measurable function $\varphi : \mathbf{R} \to \mathbf{C}$, we may always associate a unique operator $\varphi(E)$ in the following way. Define

$$\mathcal{D}(\varphi(E)) = \Big\{ f \in \mathcal{H};\ \int_{\mathcal{R}} |\varphi(\lambda)|^2 E_{f,f}(d\lambda) < +\infty \Big\}.$$

Then $\mathcal{D}(\varphi(E))$ is dense in $\mathcal{H}$. On this domain, we define $\varphi(E)$ by

$$\forall f \in \mathcal{D}(\varphi(E)),\ \forall g \in \mathcal{H},\ \langle \varphi(E)f, g\rangle = \int_{\mathbf{R}} \varphi(\lambda) E_{f,g}(d\lambda).$$

One writes

$$\varphi(E) = \int_{\mathbf{R}} \varphi(\lambda) dE(\lambda).$$

One notes that $\varphi(E^*) = \overline{\varphi}(E)$. For φ_1, φ_2 measurable, if $\mathcal{D}((\varphi_1\varphi_2)(E)) = \mathcal{D}(\varphi_1(E))$, then we have $\varphi_1(E)\varphi_2(E) = (\varphi_1\varphi_2)(E)$. One has the following spectral theorem:

Theorem 6.2. *H is self-adjoint on $\mathcal{H}$ if and only if there exists a resolution of the identity E on R such that*

$$H = \int_{\mathbf{R}} \lambda dE(\lambda).$$

The resolution of the identity defined by a self-adjoint operator is unique; it is called the *spectral resolution* of H. The spectrum of H is the support of the spectral resolution of H. The spectral resolution of a self-adjoint operator is a projection-valued Borel measure dE so it may be decomposed with respect to the Lebesgue measure. More precisely, if we define

$$\mathcal{H}_\cdot = \{ f \in \mathcal{H};\ E_{f,f} \text{ is } \cdot \}$$

where $\cdot$ stands for "pp = purely punctual", "ac = absolutely continuous" or "sc = singular continuous", then $\mathcal{H}_\cdot$ is a closed subspace of $\mathcal{H}$. One has

$$\mathcal{H} = \mathcal{H}_{\mathrm{pp}} \overset{\perp}{\oplus} \mathcal{H}_{\mathrm{ac}} \overset{\perp}{\oplus} \mathcal{H}_{\mathrm{sc}}.$$

Denote by $\Pi_\cdot$ the orthogonal projection on $\mathcal{H}_\cdot$ and define $E_\cdot = \Pi_\cdot \mathcal{H}_\cdot$. Then E_{pp} is a pure point measure, E_{ac} an absolutely continuous measure and E_{sc} a singular continuous measure. Their respective supports are $\sigma_{\mathrm{pp}}(H)$, the pure point spectrum of H, $\sigma_{\mathrm{ac}}(H)$, the absolutely continuous spectrum and $\sigma_{\mathrm{sc}}(H)$, the singular continuous spectrum of H. Note that, with our definition, $\sigma_{\mathrm{pp}}(H)$ is not the set of eigenvalues of H but the closure of this set. To finish this subsection, we redefine a notation. If E is the spectral resolution of H and $\varphi : \mathbf{R} \to \mathbf{C}$ is a measurable function, then we define $\varphi(H) := \varphi(E)$.

6.2.2. The Helffer–Sjöstrand Formula

Let $\varphi \in \mathcal{C}_0^\infty(\mathbf{R})$. An *almost analytic extension* of φ is a function satisfying

(1) for $z \in \mathbf{R}$, $\tilde{\varphi}(z) = \varphi(z)$;
(2) $\mathrm{supp}(\tilde{\varphi}) \subset \{z \in \mathbf{C};\ |\mathrm{Im}(z)| < 1\}$;
(3) $\tilde{\varphi} \in \mathcal{S}(\{z \in \mathbf{C};\ |\mathrm{Im}(z)| < 1\})$ (here, $\mathcal{S}$ denotes the Schwartz class of rapidly decaying functions);
(4) The family of functions $x \mapsto \frac{\partial \tilde{\varphi}}{\partial \bar{z}}(x+iy) \cdot |y|^{-n}$ (for $0 < |y| < 1$) is bounded in $\mathcal{S}$ for any $n \in \mathbf{N}$.

Such extensions always exist for $\varphi \in \mathcal{S}(\mathbf{R})$ (see [43]) and one has the following estimates: for $n \geq 0$, $\alpha \geq 0$, $\beta \geq 0$, there exists $C_{n,\alpha,\beta} > 0$ such that

$$\sup_{0<|y|\leq 1} \sup_{x\in\mathbf{R}} \left| x^\alpha \frac{\partial^\beta}{\partial x^\beta} \left(|y|^{-n} \frac{\partial \tilde{\varphi}}{\partial \bar{z}}(x+iy) \right) \right| \leq C_{n,\alpha,\beta} \sup_{\substack{\beta' \leq n+\beta+2 \\ \alpha' \leq \alpha}} \sup_{x\in\mathbf{R}} \left| x^{\alpha'} \frac{\partial^{\beta'} \varphi}{\partial x^{\beta'}}(x) \right|. \quad (6.2)$$

Let H be a lower semi-bounded self-adjoint operator and $1 - \lambda_0$ a lower bound for H. Let $\varphi \in \mathcal{C}_0^\infty(\mathbf{R})$ and $\tilde{\varphi}$ be an almost analytic extension of $(\lambda_0 + x)^q \varphi(x)$ (here, q is any integer). Then, by [22] (see also [32]), we know that the following formula holds:

$$\varphi(H) = \frac{i}{2\pi} \int_{\mathbf{C}} \frac{\partial \tilde{\varphi}}{\partial \bar{z}}(z) \cdot (\lambda_0 + H)^{-q} (z - H)^{-1} dz \wedge d\bar{z}. \quad (6.3)$$

Acknowledgement. The author would like to thank F. Pham for him to participate at the French–Vietnamese Colloquium of Mathematics. Acknowledgement is also given to the Action spécifique de l'Université Paris 13 Vietnam - 0122 for their grant.

References

1. P. W. Anderson, Absence of diffusion in certain random lattices, *Physical Reviews* **109** (1959) 1492–1501.
2. J. Avron and B. Simon, Almost periodic Schrödinger operators, II. The integrated density of states, *Duke Mathematical Journal* **50** (1983) 369–391.
3. J. M. Barbaroux, J. M. Combes, and P. Hislop, Landau Hamiltonians with unbounded random potentials, *Letters in Math. Phys.* **40** (4) (1997) 355–369.
4. J. M. Barbaroux, J. M. Combes, and P. Hislop, Localization near band edges for random Schrödinger operators, *Helvetica Physica Acta* **70**(1&2) (1997) 16–43.
5. S. De Bievre and F. Germinet, Dynamical localization for discrete and continuous random Schrödinger operators, LPTM/97/10, Université Paris VII, 1997 (preprint).
6. R. Carmona, Random Schrödinger operators, in: *Ecole d'été de Probabilité de Saint-Flour*, XIV, Lecture Notes in Mathematics, Vol. 1180, Springer-Verlag, Heidelberg, 1984, pp. 1–121.
7. R. Carmona and J. Lacroix, *Spectral Theory of Random Schrödinger Operators*, Birkhäuser, Basel, 1990.
8. J. M. Combes and P. D. Hislop, Localization for some continuous random Hamiltonians in d-dimensions, *Journal of Functional Analysis* **124** (1994) 149–180.
9. J. M. Combes and P. D. Hislop, Landau Hamiltonians with random potentials: Localization and the density of states, *Commun. Math. Phys.* **177** (1996) 603–630.
10. J. M. Combes, P. D. Hislop, and E. Mourre, Spectral averaging, perturbation of singular spectral and localization, *Trans. AMS* **348** (1996) 4883–4895.

11. H. L. Cycon, R. G. Froese, W. Kirsch, and B. Simon, *Schrödinger Operators*, Springer-Verlag, Berlin, 1987.
12. F. Daumer, Equations de Schrödinger avec champ électrique périodique et champ magnétique constant dans l'approximation de "tight-binding", *Commun. Partial Differential Equations* **18**(5&6) (1993) 1021–1041.
13. F. Daumer, Equations de Schrödinger avec potentiels singuliers et à longue portée dans l'approximation de liaison forte, *Annales de l'Institut Henri Poincaré, série Physique Théorique* **64**(1) (1996) 1–31 (French).
14. B. Davies, *Spectral Theory and Differential Operators*, Cambridge University Press, Cambridge, 1995.
15. R. Del Rio, S. Jitomirskaya, Y. Last, and B. Simon, What is localization?, 1995 (preprint).
16. R. Del Rio, S. Jitomirskaya, Y. Last, and B. Simon, Operators with singular continuous spectrum IV: Hausdorff dimensions, rank one perturbations and localization, *Journal d'Analyse Mathématique* **69** (1996) 153–200.
17. M. Eastham, *The Spectral Theory of Periodic Differential Operators*, Scottish Academic Press, Edinburgh, 1973.
18. J. Fröhlich and T. Spencer, Absence of diffusion in the Anderson tight binding model, *Commun. Mathematical Physics* **88** (1983) 151–184.
19. I. Gikhman and A. Skorokhod, *Introduction to the Theory of Random Processes*, Dover, New York, 1969.
20. B. Helffer and J. Sjöstrand, Multiples wells in the semi-classical limit, I. *Commun. Partial Differential Equations* **9** (1984) 337–408.
21. B. Helffer and J. Sjöstrand, Puits multiples en limite semi-classique, II. Interaction moléculaire, *Annales de l'Institut Henri Poincaré, série Physisque Théorique* **42** (1985) 127–212.
22. B. Helffer and J. Sjöstrand, On diamagnetism and the De Haas-Van Alphen effect, *Annales de l' Institut Henri Poincaré, série Physique Théorique* **52** (1990) 303–375.
23. H. Holden and F. Martinelli, A remark on the absence of diffusion near the bottom of the spectrum for a random Schrödinger operator in $l^2(\mathbf{R}^v)$, *Commun. Math. Phys.* **93** (1984) 197–217.
24. T. Kato, *Perturbation Theory for Linear Operators*, Springer-Verlag, Berlin, 1980.
25. W. Kirsch, Random Schrödinger operators, in: *Schrödinger Operators*, Lecture Notes in Physics, Vol. 345, A. Jensen and H. Holden (Eds.), Springer-Verlag, Berlin, 1989.
26. W. Kirsch and F. Martinelli, On the spectrum of Schrödinger operators with a random potential, *Commun. Math. Phys.* **85** (1982) 329–350.
27. W. Kirsch and F. Martinelli, Large deviations and Lifshits singularities of the integrated density of states of random Hamiltonians, *Commun. Math. Phys.* **89** (1983) 27–40.
28. W. Kirsch and B. Simon, Lifshits tails for the Anderson model, *J. Statistical Physics* **38** (1985) 65–76.
29. W. Kirsch, P. Stollman, and G. Stolz, Localization for random perturbations of periodic Schrödinger operators, Ruhr Universität, Bochum, 1996 (preprint).
30. F. Klopp, Work in progress.
31. F. Klopp, Localization for semi-classical continuous random Schrödinger operators II: the random displacement model, *Helvetica Physica Acta* **66** (1993) 810–841.
32. F. Klopp, An asymptotic expansion for the density of states of a random Schrödinger operators with Bernoulli disorder, *Random Operators and Stochastic Equations* **3**(4) (1995) 315–332.
33. F. Klopp, Localisation pour des opérateurs de Schrödinger aléatoires dans $l^2(\mathbf{R}^d)$: un modèle semi-classique, *Annales de l'Institut Joseph Fourier* **45** (1995) 265–316.
34. F. Klopp, Localization for some continuous random Schrödinger operators, *Commun. Math. Phys.* **167** (1995) 553–570.
35. F. Klopp, Internal Lifshits tails for random perturbations of periodic Schrödinger operators, Université Paris-Nord, Villetaneuse, 1997 (preprint); *Math. Phys. Archive*, Preprint 81–97; *Duke Math. J.* (to appear).
36. F. Klopp, Band edge behaviour for the integrated density of states of random Jacobi matrices in dimension 1, *J. Statistical Physics* (1998) (to appear).
37. P. Kuchment, *Floquet Theory for Partial Differential Equations*, Operator Theory: Advances and Applications, Vol. 60, Birkhäuser, Basel, 1993.

38. H. Kunz and B. Souillard, Sur le spectre des opérateurs aux différences finies aléatoires, *Commun. Math. Phys.* **178** (1980) 201–246.
39. L. Landau and L. Lifshits, *Mécanique Quantique, Théorie Non-relativiste*, Editions MIR, Moscow, 1966.
40. I. M. Lifshits, Structure of the energy spectrum of impurity bands in disordered solid solutions, *Soviet Physics JETP* **17** (1963) 1159–1170.
41. I. M. Lifshits, S. A. Gredeskul, and L. A. Pastur, *Introduction to the Theory of Disordered Systems*, Wiley, New York, 1988.
42. F. Martinelli and E. Scoppola, Introduction to the mathematical theory of Anderson localization, *Rivista Nuovo Cimento* **10** (1997).
43. J. N. Mather, On Nirenberg's proof of Malgrange's preparation theorem, in: *Proceedings of Liverpool Singularities-Symposium I*, Lecture Notes in Mathematics, Vol. 192, Springer-Verlag, Berlin, 1971.
44. G. Mezincescu, Internal Lifshits singularities of disordered finite-difference Schrödinger operators, *Commun. Math. Phys.* **103** (1986) 107–116.
45. G. Mezincescu, Lifshits singularities for periodic operators plus random potentials, *J. Statistical Physics* **49** (1987) 1081–1090.
46. G. Mezincescu, Internal Lifshits singularities for one dimensional Schrödinger operators, *Commun. Math. Phys.* **158** (1993) 315–325.
47. S. Nakao, On the spectral distribution of the Schrödinger operator with random potential, *Japan J. Mathematics* **3** (1977) 117–139.
48. L. Pastur, On the distribution of the eigenvalues of the Schrödinger equation with a random potential, *Functional Analysis and its Applications* **8** (1972) 163–165.
49. L. Pastur, Spectra of random self-adjoint operators, *Russian Mathematical Surveys* **28** (1973).
50. 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. Math. Fiz.* **32** (1977) 88–95 (Russian).
51. L. Pastur, Spectral properties of disordered systems in the one-body approximation, *Commun. Math. Phys.* **75** (1980) 179–186.
52. L. Pastur and A. Figotin, *Spectra of Random and Almost-Periodic Operators*, Springer-Verlag, Berlin, 1992.
53. M. Reed and B. Simon, Methods of modern mathematical physics, in: *Analysis of Operators*, Vol. IV, Academic Press, New York, 1978.
54. M. Reed and B. Simon, Methods of modern athematical physics, in: *Functional Analysis*, Vol. I, Academic Press, New York, 1980.
55. M. A. Shubin, Spectral theory and index of elliptic operators with almost periodic coefficients, *Russian Mathematical Surveys* **34** (1979) 109–157.
56. B. Simon, Semi-classical analysis of low lying eigenvalues I. Non degenerate minima: Asymptotic expansion, *Annales de l'Institut Henri Poincaré, série Physique Théorique* **38** (1983) 295–307.
57. B. Simon, Lifshits tails for the Anderson model, *J. Statistical Physics* **38** (1985) 65–76.
58. B. Simon, Internal Lifshits tails, *J. Statistical Physics* **46** (1987) 911–918.
59. J. Sjöstrand, Microlocal analysis for periodic magnetic Schrödinger equation and related questions, in: *Microlocal Analysis and Applications*, Lecture Notes in Mathematics, Vol. 1495, Springer-Verlag, Berlin, 1991.
60. G. Stolz, Localization for random Schrödinger operator with poisson potential, *Annales de l'Institut Henri Poincaré, série Physique Théorique* **63** (1995) 297–314.
61. A. Sznitman, Lifshitz tails and Wiener sausages, I, *J. Functional Analysis* **94** (1990) 223–246.
62. A. Sznitman, Fluctuations of principal eigenvalues and random scales, ETH, Zürich, *Commun. Math. Phys.* **189**(2) (1997) 337–363.
63. W. M. Wang, Microlocalization, percolation and Anderson localization for the magnetic Schrödinger operator with a random potential, *J. Functional Analysis*, **146**(1) (1997) 1–26.
64. F. Wegner, Bounds on the density of states in disordered systems, *Zeitschrift für Physik B* **44** (1981) 9–15.