

Lecture notes on Quantum Diffusion and Random Matrix Theory

Felipe Hernández

October 30, 2025

Abstract

In joint work with Adam Black and Reuben Drogin [BDH25b; BDH25a], we develop a new approach to understanding the diffusive limit of the random Schrodinger equation based on ideas taken from random matrix theory. These lecture notes present the main ideas from this work in a self-contained and simplified presentation. The lectures were given at the summer school “PDE and Probability” at Sorbonne Université from June 16-20, 2025.

Contents

1	Introduction	2
2	The Kinetic Time Scale	7
2.1	The non-commutative Khintchine inequality	9
2.1.1	Proof of the non-commutative Khintchine inequality	11
2.1.2	Applying non-commutative Khintchine to $T_1(t)$	13
2.2	Bounds on $T_k(t)$ from $T_1(t)$	14
2.3	Bounds for spectral projections	16
2.3.1	Resolvent bounds for Δ and H	18
2.3.2	ℓ^p bounds for eigenfunctions	20
3	The Diffusive Time Scale	21
3.1	The direct perturbative approach	22
3.2	The resolvent of a Wigner matrix	25

3.2.1	More general ensembles	28
3.3	The local law for the Anderson model	29
3.3.1	The self-consistent equation for θ	31
3.3.2	Concentration of the resolvent entries	33
3.3.3	Conclusion of the proof of Theorem 3.3	34
3.3.4	On the assumptions on $\tilde{\theta}$	35
3.4	The T equation	35
3.5	Finishing the proof of Theorem 3.1	39

1 Introduction

The goal of these lectures is to say what we can about the random Schrödinger equation in the weak coupling limit:

$$i\partial_t\psi = \Delta\psi + \lambda V\psi. \tag{1.1}$$

Above, Δ is the Laplacian, $d \geq 2$, λ is a small coupling parameter $\lambda \ll 1$, and V is a random potential. One can consider (1.1) on $\mathbb{R}^d$ or $\mathbb{Z}^d$. On $\mathbb{R}^d$ one can take V to be a stationary Gaussian field (for example), and on $\mathbb{Z}^d$ it is simplest to take V to have independent standard Gaussian entries, and Δ to be the nearest-neighbor Laplacian.

The motivation for studying (1.1) is that it is a model for studying wave transport in random media. Indeed, the Schrodinger equation is the simplest example of a dispersive PDE, and the term λV is the simplest kind of random perturbation that can be made to it. Examples of wave transport in random media arise in telecommunications, geological imaging, and condensed matter physics. This last example is closest to the specific model described above – in [And58], Anderson introduced (1.1) as a model for electron transport in disordered materials. More specifically, Anderson considered a model on the lattice $\mathbb{Z}^d$ where Δ represents a discrete “hopping” term¹ between lattice sites and λV is the potential associated to each site. In this case the Hamiltonian $H = \Delta + \lambda V$ describes the effective energy of a single electron in a disordered material, and thus encodes the electrical properties of this material.

The simplest electrical property is the conductivity, and it is natural to ask whether H describes an insulator or a conductor. A direct way to model this mathematically (as considered already in Anderson’s original work) is to observe the dispersion of an initially localized wavefunction $\psi_0 \in \ell^2(\mathbb{Z}^d)$ – for example one may take $\psi_0 = |0\rangle$, the wavefunction localized at the site at the origin. In this case one is interested in the mean square displacement of the wavefunction at time t , for ψ_t solving (1.1). That is,

¹Anderson actually considered much more general hopping terms which could be nonlocal and nonuniform.

one defines

$$r^2(t) := \sum_{x \in \mathbb{Z}^d} |x|^2 |\psi_t(x)|^2. \quad (1.2)$$

In [And58], Anderson argued that for sufficiently large λ there is some constant $r_{\max}$ depending on the initial data ψ_0 and the potential λV such that $r(t) \leq r_{\max}$ for all t . The physical interpretation of this fact is that sufficiently disordered materials are *insulators*.

That $r(t)$ is bounded at high disorder is derived a consequence of the fact that the operator H has a pure point spectrum of orthonormal eigenfunctions, each exponentially localized to a finite interval. Anderson's original paper inspired an entire subfield of condensed matter physics characterizing the localization of eigenfunctions of random Schrodinger operators. A thorough survey of localization is outside the scope of these lectures (see [AW15; Sto01] for more comprehensive treatments), but we can summarize what is known rigorously as follows:

- In $d = 1$, the operator H has a complete orthonormal basis of localized eigenfunctions, and therefore $r(t) \leq r_{\max} < \infty$ almost surely [GMP77; KS80]. In the weak coupling limit $\lambda \ll 1$, it is the case that $r_{\max} \simeq \lambda^{-2}$ [Pas80; CL90].
- In any $d \geq 1$ and any $\lambda > 0$, there exist localized eigenfunctions of H near the spectral edges [FS83].

A conspicuous gap in the mathematically rigorous theory of localization is what happens to the *bulk* of the spectrum in $d \geq 2$. It is conjectured that in $d \geq 3$ there is a “metal-insulator” transition in the spectrum between localized eigenfunctions and pure point spectrum near the edges and a continuous spectrum consisting of “extended” or delocalized eigenstates in the bulk. In $d = 2$ it is instead conjectured that the bulk consists of localized eigenfunctions with localization length scale on the order $e^{c\lambda^{-2}}$, as predicted in [Abr+79]. These conjectures appear in the list of Simon's problem's [Sim84; Sim00].

In terms of bounds on $r(t)$, the above results and conjectures correspond to the following bounds on $r(t)$:

$$\sup_{t < \infty} r(t) = r_{\max} \simeq \begin{cases} \lambda^{-2}, & d = 1 \\ e^{\lambda^{-2}}, & d = 2 \\ +\infty, & d \geq 3. \end{cases}$$

The conjectured values for $r_{\max}$, the extended states conjecture, and the existence of the mobility edge seem to be far out of the reach of current methods. A more modest goal is to try to characterize $r(t)$ for finite (as opposed to infinite) times. This is the goal of these lecture notes. There are two heuristics that determine the behavior of $r(t)$:

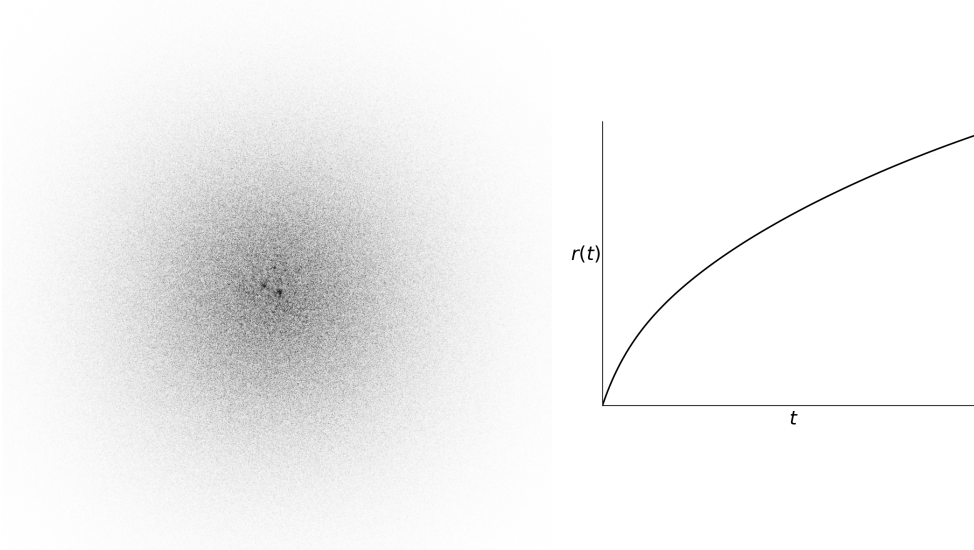


Figure 1: The result of a simulation in $d = 2$, $\lambda = 0.1$ at time $T = 2000$ with initial data $\psi_0 = |0\rangle$. On the left is the square of the wavefunction at time T , and on the right is the plot of $r(t)$ as a function of t . Note that $|\psi_T|^2$ weakly resembles a Gaussian distribution, and that $r(t)$ appears to have square-root growth at long times.

1. For times $t \ll \lambda^{-2}$, the evolution e^{-itH} behaves like the free evolution $e^{-it\Delta}$.
2. The effect of the potential V is to scatter the wavefunction, and each scattering event is independent.

The combination of these heuristics suggests that $\psi_t(x)$ should be some random superposition of random walk paths of $\lambda^2 t$ steps and step size λ^{-2} . Therefore, one expects that $|\psi_t|^2$ resembles a Gaussian and $r(t) \approx \lambda^{-1} t^{1/2}$. See Figure 1 for a numerical simulation.

In these notes we provide a self-contained proof of the following result.

Theorem 1.1 (Simplification of Theorem 1.1 in [BDH25a]). *Let ψ_0 be the Kronecker delta at the origin and ψ_t solve (1.1) on $\mathbb{Z}^d$, $d \geq 2$, and let $\kappa < \frac{1}{10}$. Then for any $\lambda^{-2} \leq T \leq \lambda^{-2-\kappa}$, the bound*

$$\frac{1}{T} \int_0^T \sum_{x \in \mathbb{Z}^d} |x|^2 |\psi_t(x)|^2 dt \geq c \lambda^{-1} T^{1/2}$$

holds with probability at least $1 - C\lambda^{1000}$.

Prior work towards understanding diffusion in the random Schrodinger equation

can be characterized loosely into three threads.² The first thread consists of *direct perturbative* approaches. This can be traced back to the work of van Hove [Van54], who first observed conditions under which perturbations λW to a Hamiltonian H_0 would have “effective strength” λ^2 . In [Spo77] considered the “van Hove limit” $\lambda \rightarrow 0$ and $t \sim \lambda^{-2}$ for the random Schrodinger equation, establishing a kinetic equation for the momentum distribution up to time $c\lambda^{-2}$. This was improved by László Erdős and HT Yau in [EY00], who reached time scales $C\lambda^{-2}$ and established a kinetic equation for the Wigner phase space distribution of ψ_t . The difficulty with the “direct perturbative” approach is that to reach longer time scales one must go up to higher orders in perturbation theory and tam a combinatorial explosion of diagrams. Nevertheless in a technical tour-de-force, Erdős, Salmhofer, and Yau [ESY08; ESY07a; ESY07b] reached diffusive time scales on the order $\lambda^{-2-\varepsilon}$ (with $\varepsilon = \frac{1}{9800}$ on $\mathbb{Z}^3$, and $\varepsilon = \frac{1}{370}$ on $\mathbb{R}^3$). This work was revisited in [Her24] which provided a different perspective on the diagrammatic expansions. These papers only addressed convergence to a diffusive limit *in expectation*, whereas convergence in higher order moments was established by Thomas Chen up to kinetic time in [Che06].

The second thread consists of *harmonic analysis* approaches to the random Schrodinger equation. In [SSW02], Schlag, Shubin, and Wolff proved that in $d \leq 2$, eigenfunctions for the random Schrodinger operator have Fourier transforms localized to a λ^2 -thick annulus around the corresponding level set of the dispersion relation. In dimension 2, this proof used as input the Córdoba L^4 argument used in two-dimensional restriction theory. Later [Bou01] established similar estimates without the use of wavepacket analysis, using instead stronger inputs from probability. Both of these papers pushed the first heuristic, that e^{itH} behaves like $e^{it\Delta}$ up to time λ^{-2} , nearly as far as it could go by using the properties of the Laplacian in a strong way.

The third thread of research is *random matrix* analysis of related models. The Anderson Hamiltonian $H = \Delta + \lambda V$ is itself a random matrix, but the fact that the randomness appears only on the diagonal has thus far precluded the use of traditional ideas from the field. Indeed, the most traditional random matrix models have randomness in *every* entry. Nevertheless, there has been recent and rapid progress on understanding random matrices with fewer and fewer random entries. A simple example is the random band matrix on $\mathbb{Z}_L^d := \mathbb{Z}^d / L\mathbb{Z}^d$, which has Hamiltonian H

$$H_{xy} = \begin{cases} g_{xy}, & |x - y| \leq W \\ 0, & \text{else.} \end{cases}$$

The random entries g_{xy} are independent up to the symmetry constraint $g_{xy} = g_{yx}$. Recent progress on this model uses the framework of computing moments for the resolvent $R(z) = (H - z)^{-1}$ via *self-consistent equations* developed first for Wigner

²Of course, this categorization is incomplete and overly simple. Some very interesting works that the reader may also find inspiring are the papers of Magnen, Poirot, and Rivasseau [MPR97; MPR98; MPR99] and Duerinckx and Shirley [DS21].

matrices [ESY09] (see also [Erd+13a] for a clear exposition) and extended to random band matrices in [Erd+13b]. In a recent series of papers [YYY21; YYY22; YY21], it was shown that in $d \geq 8$ the eigenvectors of H are completely delocalized so long as $W \geq L^\varepsilon$ and L is large enough. Even more recently, this work has been extended to other random matrix models that are closer to the Anderson model [YY25] and, in spectacular recent breakthroughs, even to dimension $d = 2$ [Dub+25b] and $d \geq 3$ [Dub+25a]. Until now, however, these random matrix methods have never been applied directly to the Anderson model.

In these notes we explain how to use ideas from each of these threads of prior work to prove Theorem 1.1. We break up the exposition into two sections. Section 2 covers the required results up the kinetic time scale λ^{-2} which were established in [BDH25b]. Section 3 then explains how to prove results at the diffusive time scale and proves Theorem 1.1. In these lecture notes the goal is to communicate my favorite ideas from the proofs, and to skip details I find less interesting. The reader can of course refer to [BDH25b; BDH25a] for more details and stronger results.

Acknowledgements. I thank the organizers of the “PDE and Probability” summer school at Sorbonne Université for the invitation to deliver these lectures. In particular, I would like to thank Antoine Gloria for his generous hospitality during this visit. Of course the ideas and perspectives contained in these notes are due to many discussions with Adam Black and Reuben Drogin. This work was supported by NSF grant DMS-2303094.

2 The Kinetic Time Scale

In this section we try to squeeze out as much information about e^{itH} as possible just from considerations at short times (short compared to the kinetic timescale λ^{-2}). Previously it was known that one could obtain delocalization lower bounds at scale λ^{-2} [SSW02; Che06]. In this chapter we establish this as a consequence of the fact that $e^{itH} \approx e^{it\Delta}$ for $t \ll \lambda^{-2}$, which itself has much farther reaching consequences. In particular, we are able to prove $\ell^p \rightarrow \ell^q$ bounds for the spectral projections of H which we will use as a priori estimates to prove diffusion.

We can write $e^{itH} - e^{it\Delta}$ using the Duhamel formula as follows:

$$e^{-itH} - e^{-it\Delta} = i\lambda \int_0^t e^{-i(t-s)H} V e^{is\Delta} ds. \quad (2.1)$$

Thinking of V as a bounded potential, one naively has $e^{-itH} - e^{-it\Delta} = O(\lambda t)$. In fact however there is a square-root cancellation in the integral above which leads to the estimate

$$e^{-itH} - e^{-it\Delta} = O(\lambda\sqrt{t}).$$

Such an estimate would imply that the effective strength of the potential is λ^2 rather than λ , in the sense that for times $t \ll \lambda^{-2}$ the potential is not noticed.

So far I have been imprecise about the precise norm one should use to measure $e^{-itH} - e^{-it\Delta}$. For example, one precise (and correct!) statement one can make is

$$\sup_{\|\psi\|_{\ell^2}=1} \left(\mathbb{E} \|e^{-itH}\psi - e^{-it\Delta}\psi\|_{\ell^2}^{2k} \right)^{\frac{1}{2k}} \lesssim C_k \lambda \sqrt{t}. \quad (2.2)$$

The bound (2.2) is useful in some applications, but it cannot directly be used to say anything about *eigenfunctions* of H , for the simple reason that eigenfunctions are random functions and the statement above is about deterministic ψ . A stronger and more useful bound would swap the expectation and supremum above, and thus be a bound on operator norm as in

$$\left(\mathbb{E} \|e^{-itH} - e^{-it\Delta}\|_{\ell^2 \rightarrow \ell^2}^{2k} \right)^{\frac{1}{2k}} \lesssim C_k \lambda \sqrt{t}. \quad (2.3)$$

Such a bound cannot be true on $\mathbb{Z}^d$ for the trivial reason that a random potential resembles any deterministic potential on arbitrarily large sets (although extremely sparsely). There are several ways to remedy this. One is to add a cutoff to the potential so that it has compact support, as done in [BDH25c]. An alternate remedy, which we use in these notes, is to consider a model on a discrete torus $\mathbb{Z}_L^d := \mathbb{Z}^d / L\mathbb{Z}^d$ for a length $L = \lambda^{-10}$ (say), which is much larger than any length scale relevant to the dynamics we consider.

We are now ready to state the main result of this chapter.

Theorem 2.1. *Let $L = \lambda^{-100}$ and $H = \Delta_L + \lambda V$. Then with probability at least $\exp(-cK^2)$, we have that the estimate*

$$\|e^{-itH} - e^{-it\Delta}\|_{\ell^2 \rightarrow \ell^2} \leq K(\log \lambda^{-1})^2 \lambda \sqrt{t}.$$

holds for all $t \in \mathbb{R}$.

What we need in our application is a result on spectral projections. To this end, let $\chi \in C_c^\infty(\mathbb{R})$ be any smooth and compactly supported function. We write $\chi(H)$ to mean the operator satisfying $\chi(H)\psi_E = \chi(E)\psi_E$ for any eigenfunction $H\psi_E = E\psi_E$ (since we are on $\mathbb{Z}_L^d$ anyway, H is a finite dimensional matrix). Then we define $\chi_{\delta,E}(x) := \chi((x - E)/\delta)$. Then we have the following result.

Corollary 2.2. *Let $L = \lambda^{-100}$ and $H = \Delta_L + \lambda V$. Then with probability at least $\exp(-cK^2)$ we have that*

$$\|\chi_{\delta,E}(H) - \chi_{\delta,E}(\Delta_L)\|_{\ell^2 \rightarrow \ell^2} \leq K(\log \lambda^{-1})^2 \lambda \delta^{-1/2}.$$

The above result allows us to compare the spectral statistics of Δ_L with that of H down to intervals of width λ^2 , and is the crucial ingredient in our approach to proving quantum diffusion.

Indeed, we use Corollary 2.2 to establish the following less obvious consequence of Theorem 2.1:

Theorem 2.3. *Let $\varepsilon > 0$ be a small number and $N > 0$ be a big number. Suppose $E \in [-2d, 2d]$ is not a critical value of the dispersion relation ω . Then resolvent for $H = \Delta_L + \lambda V$ satisfies, for any $1 \leq p \leq \frac{6}{5}$ and $6 \leq q \leq \infty$, the estimate*

$$\|R(E + i\eta)\|_{p \rightarrow q} \leq C_{\varepsilon,N} \lambda^{-\varepsilon} (\lambda^2 \eta^{-1} + 1)$$

with probability at least $1 - C_{\varepsilon,N} \lambda^N$.

To work with the above theorem more concisely we introduce the following notation. We say that B stochastically dominates A , written $A \prec B$, if for any $\varepsilon, N > 0$ there is a $C_{\varepsilon,N}$ such that

$$A \leq C_{\varepsilon,N} \lambda^{-\varepsilon} B$$

holds with probability at least $1 - C_{\varepsilon,N} \lambda^N$. We will derive Theorem 2.3 from Corollary 2.2 later in Section 2.3. The bulk of this chapter is dedicated to the proof of Theorem 2.1, which we sketch below.

To prove Theorem 2.1 we iterate (2.1) to write out $e^{-itH} - e^{-it\Delta}$ as a Dyson series,

$$e^{-itH} - e^{-it\Delta} = \sum_{j=1}^{\infty} e^{-it\Delta} (i\lambda)^j T_j(t), \quad (2.4)$$

where

$$T_1(t) := \int_0^t e^{-is\Delta} V e^{is\Delta} ds$$

and for $j \geq 1$ we have

$$T_j(t) := \int_{0 \leq s_1 \leq \dots \leq s_j \leq t} V(s_j) V(s_{j-1}) \cdots V(s_1) d\vec{s},$$

with $V(s) := e^{-is\Delta} V e^{is\Delta}$.

A simple idea that would work to bound $\mathbb{E} \|T_j\|_{op}^k$ would be to use the moment method, using

$$\mathbb{E} \|T_j\|_{op}^{2k} \leq \mathbb{E} \|T_j^* T_j\|_{op}^k \leq \mathbb{E} \text{tr}(T_j^* T_j)^k.$$

The expectation on the right could be computed using a Wick expansion, involving the introduction of a combinatorial explosion of terms. Such an analysis was indeed done in [Her24], but this is cumbersome and it is difficult to extract reasonable quantitative bounds.

To avoid any diagrammatic expansion we use two ideas. The first is to observe that $T_1(t)$ is a random symmetric matrix that is *linear* in the randomness V . This allows us to import a standard tool from random matrix theory, the non-commutative Khintchine inequality. The noncommutative Khintchine inequality is introduced and proven in Section 2.1. The second idea is a remarkable reduction which allows us to bound the operator norms of T_j in terms of the operator norm of T_1 . For the second trick to work, it is crucial that we work in operator norm since we make use of an approximate structure of T_2 (in particular, we will use the inequality $\|AB\| \leq \|A\| \|B\|$).

2.1 The non-commutative Khintchine inequality

The presentation in this section closely follows Section 3.1 of van Handel's "Structured Random Matrices" [Van17].

The main ingredient we need is a bound on the operator norm of random matrices X of the form

$$X = \sum_{j=1}^s g_j A_j, \tag{2.5}$$

where g_j are independent Gaussian variables and A_j are symmetric random $n \times n$ matrices. Note that any random matrix X with jointly Gaussian entries can be written in this way.

As examples, note that a GOE random matrix (having variance $2N^{-1}$ on the diagonal and variance N^{-1} on the off-diagonal) can be written as

$$X_{\text{GOE}} := \frac{1}{\sqrt{N}} \sum_{i \leq j} g_{ij} E_{ij} \tag{2.6}$$

where

$$E_{ij} := \begin{cases} \sqrt{2} |i\rangle \langle i|, & i = j \\ |i\rangle \langle j| + |j\rangle \langle i|, & i \neq j. \end{cases}$$

Another example is a diagonal random matrix,

$$X_{\text{diag}} := \sum_i g_i |i\rangle \langle i|. \quad (2.7)$$

In both cases, one can verify that $\|X\|_{\text{op}} \lesssim \|\sum_j A_j^2\|^{1/2}$ with high probability, which is a noncommutative version of the square root cancellation expected if A_j were scalars.

Theorem 2.4 (Non-commutative Khintchine inequality). *For matrices X of the form (2.5),*

$$(\mathbb{E} \text{tr}[X^{2p}])^{1/2p} \leq \sqrt{2p-1} \left(\text{tr}[(\mathbb{E} X^2)^p] \right)^{1/2p} \quad (2.8)$$

This result is due to Lust-Piquard and Pisier [Lus86; LP91].

Before we prove Theorem 2.4 we note that it implies a concentration inequality on the operator norm of X , up to a logarithmic loss in the dimension.

Corollary 2.5. *Let X be of the form (2.5) and let n be the dimension of X (so X is an $n \times n$ matrix). Then*

$$\mathbb{E} \|X\|_{\text{op}} \lesssim \sqrt{\log(n)} \left\| \sum_{j=1}^s A_j^2 \right\|_{\text{op}}^{1/2}. \quad (2.9)$$

Moreover, for $\alpha > 4$ we have the bound

$$\mathbb{P}(\|X\|_{\text{op}} \geq \alpha \sqrt{\log(n)} \left\| \sum_{j=1}^s A_j^2 \right\|_{\text{op}}^{1/2}) \leq \exp\left(-\frac{\log 2}{8} \alpha^2 \log n\right). \quad (2.10)$$

To see where the factor of $\sqrt{\log n}$ comes from, observe that with X_{diag} defined as in (2.7),

$$\|X_{\text{diag}}\|_{\text{op}} = \sup_{1 \leq i \leq n} |g_i|.$$

This is on the order $\sqrt{\log n}$ with high probability. Note that in this case $\sum_i A_i^2 = \text{Id}$. On the other hand, for the GOE ensemble (2.6),

$$N^{-1} \sum_{ij} E_{ij}^2 = \frac{N+1}{N} \text{Id},$$

so that the non-commutative Khintchine inequality implies

$$\|X_{\text{GOE}}\|_{\text{op}} \lesssim \sqrt{\log N}.$$

This is lossy, as the truth is that the largest eigenvalue of X_{GOE} is close to 2 with high probability. In general, determining the precise dependence on dimension is an interesting problem, but it is completely irrelevant to our analysis which is insensitive to logarithms.

Exercise. Derive Corollary 2.5 from Lemma 2.4.

2.1.1 Proof of the non-commutative Khintchine inequality

The key point of the proof of Theorem 2.4 is the following bound, which can be interpreted as saying that the worst case is that A_j all commute.

Lemma 2.6. *For any symmetric matrices A and B , and $0 \leq \ell \leq 2p - 2$*

$$\text{tr}[AB^{2p-2-\ell}AB^\ell] \leq \text{tr}[A^2B^{2p-2}]. \quad (2.11)$$

Proof. We can always rotate into a basis so that B is diagonal with entries $B_{jk} = b_j \delta_{jk}$. Then the trace can be estimated using Holder's inequality as follows:

$$\begin{aligned} \text{tr}[AB^{2p-2-\ell}AB^\ell] &= \sum_{j,j'} b_j^\ell b_{j'}^{2p-2-\ell} |a_{j,j'}|^2 \\ &\leq \left(\sum_{j,j'} |b_j|^{2p-2} |a_{j,j'}|^2 \right)^{\ell/(2p-2)} \left(\sum_{j,j'} |b_{j'}|^{2p-2} |a_{j,j'}|^2 \right)^{(2p-2-\ell)/(2p-2)}. \end{aligned} \quad (2.12)$$

Now we recognize each sum in the product above as being equal to $\text{tr}[AB^{2p-2}A] = \text{tr}[A^2B^{2p-2}]$. $\square$

Lemma 2.7 (Matrix Jensen's inequality). *Let A be a symmetric matrix and let $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ be a convex function. Then*

$$\sum_{j=1}^d \varphi(a_{jj}) \leq \text{tr}[\varphi(A)]. \quad (2.13)$$

Proof. Write $A = Q\Lambda Q^*$ where Λ is diagonal, and q_{jk} are the entries of Q . Then

$$a_{jj} = \sum_k \lambda_k |q_{jk}|^2.$$

Then by Jensen's inequality (using that $\sum_k |q_{jk}|^2 = 1$ for all j ,

$$\begin{aligned} \sum_j \varphi(a_{jj}) &= \sum_j \varphi\left(\sum_k \lambda_k |q_{jk}|^2\right) \\ &\leq \sum_j \sum_k \varphi(\lambda_k) |q_{jk}|^2 \\ &= \sum_k \varphi(\lambda_k) = \text{tr}[\varphi(A)]. \end{aligned} \quad (2.14)$$

□

Lemma 2.8 (Matrix Holder's inequality). *For symmetric matrices A and B and $\frac{1}{p} + \frac{1}{p'} = 1$ with $p > 1$,*

$$\mathrm{tr}[AB] \leq \mathrm{tr}[|A|^p]^{1/p} \mathrm{tr}[|B|^{p'}]^{1/p'}. \quad (2.15)$$

Proof. Without loss of generality we can take B to be diagonal. Then

$$\mathrm{tr}[AB] = \sum_j b_{jj} a_{jj} \quad (2.16)$$

The result now follows from Holder's inequality, since $\sum |a_{jj}|^p \leq \mathrm{tr}[|A|^p]$. □

Proof of Theorem 2.4. We use the Gaussian integration by parts formula $\mathbb{E}gf(g) = \mathbb{E}f'(g)$ for Gaussians g to compute

$$\begin{aligned} \mathbb{E}[\mathrm{tr}[X^{2p}]] &= \sum_{j=1}^s \mathbb{E}[g_j \mathrm{tr}[A_j X^{2p-1}]] \\ &= \sum_{\ell=0}^{2p-2} \sum_{j=1}^s \mathbb{E}[\mathrm{tr}[A_j X^{2p-2-\ell} A_j X^\ell]] \\ &\leq (2p-1) \mathbb{E}[\mathrm{tr}[(\sum_{j=1}^s A_j^2) X^{2p-2}]] \\ &\leq (2p-1) \mathrm{tr}[(\mathbb{E}X^2)^p]^{1/p} \mathrm{tr}[\mathbb{E}X^{2p}]^{1-\frac{1}{p}}. \end{aligned} \quad (2.17)$$

In the last step we have used the matrix Holder's inequality and the identity $\sum_{j=1}^s A_j^2 = \mathbb{E}X^2$. The proof now follows from rearranging. □

Let us take a moment to reflect on the proof of the non-commutative Khintchine inequality. The identity

$$\mathbb{E} \mathrm{tr}[X^{2p}] = \sum_{\ell=0}^{2p-2} \sum_j \mathbb{E} \mathrm{tr}[A_j X^{2p-2-\ell} A_j X^\ell]$$

can be interpreted diagrammatically. Indeed, the Wick expansion dictates that the left hand side can be expanded as a sum over the $(2p-1)!!$ perfect matchings of $[2p]$. The right hand side can be thought of as a decomposition of these matchings according to where the first matching got sent. The inequality

$$\mathrm{tr}[A_j X^{2p-2-\ell} A_j X^\ell] \leq \mathrm{tr}[A_j^2 X^{2p-2}],$$

combined with the Holder inequality, allows us to argue that the worst case matching is the one in which each matrix factor gets paired with its neighbor. In particular it allows us to “uncross” all of the diagrams. This inequality is in fact an equality when all of the A_j commute (indeed, in this case $A_j X^a A_j X^b = A_j^2 X^{a+b}$). On the other hand in the GOE example, the E_{ij} matrices are very much non-commutative, so the inequality is very lossy.

2.1.2 Applying non-commutative Khintchine to $T_1(t)$

Now we turn back to the random Schrodinger equation on $\mathbb{Z}_L^d$. In this section we write $\Delta = \Delta_L = \Delta_{\mathbb{Z}_L^d}$.

Recall that we need a bound on the first collision operator,

$$T_1(t) = \int_0^t e^{is\Delta} V e^{-is\Delta} ds, \quad (2.18)$$

in operator norm. Note that $T_1(t)$ can be written in the form

$$T_1(t) = \sum_{j \in \mathbb{Z}_L^d} g_j A_j(t),$$

where

$$A_j(t) = \int_0^t e^{is\Delta} |j\rangle \langle j| e^{-is\Delta} ds,$$

and $|j\rangle \langle j|$ is the rank-one projection onto the site at $j \in \mathbb{Z}_L^d$.

We have set things up so that we can directly apply the non-commutative Khintchine inequality directly to $T_1(t)$:

Lemma 2.9. *With $L = \lambda^{-100}$, operator $T_1(t)$ satisfies the estimates*

$$\mathbb{E}(\|T_1(t)\|_{op}) \lesssim \begin{cases} (\log \lambda^{-1}) \sqrt{t \log t}, & d = 2 \\ (\log \lambda^{-1}) \sqrt{t}, & d \geq 3. \end{cases} \quad (2.19)$$

for $t \ll R$.

We have made two simplifications in the statement of the above result. The first is that we did not state the tail bounds for $\|T_1(t)\|$, but this follows from just applying the correct version of the non-commutative Khintchine inequality (with the tail bound, not just the point estimate). The second is that we have obtained a bound only for a single (arbitrary) time t . It is possible to improve this to a bound on all $t \ll \lambda^{-10}$ by a union bound argument.

Now we prove Lemma 2.9 by a more or less direct computation.

Proof. By the non-commutative Khintchine inequality, it suffices to bound the operator norm of

$$B(t) := \sum_j A_j(t)^2 = \sum_j \int_0^t \int_0^t e^{is'\Delta} |j\rangle \langle j| e^{i(s-s')\Delta} |j\rangle \langle j| e^{-is\Delta} ds ds'. \quad (2.20)$$

Using that $\sum_j |j\rangle \langle j| = \text{Id}$ and that $\langle j| e^{i\sigma\Delta} |j\rangle = \langle 0| e^{i\sigma\Delta} |0\rangle = f_d(t)$, we can write

$$B(t) = \int_0^t \int_0^t f_d(s-s') e^{i(s'-s)\Delta} ds ds',$$

so that

$$\|B(t)\|_{op} \leq \int_0^t \int_0^t |f_d(s-s')| ds ds' \leq t \int_{-t}^t |f_d(\sigma)| d\sigma.$$

A Fourier-analysis computation shows that $|\langle 0| e^{i\tau\Delta_{\mathbb{Z}^d}} |0\rangle| \lesssim (1+|\tau|)^{-d/2}$. On $d=2$ this is borderline integrable and we pick up a factor of $\log t$, and otherwise this is bounded directly by t . The proof is now completed with an application of NCK. $\square$

Exercise. Show that $\|T_1(t+s)\|_{op} \leq \|T_1(t)\|_{op} + \|T_1(s)\|_{op}$. Deduce that for $T \geq 1$,

$$\sup_{t \in [0, T]} \|T_1(t)\|_{op} \leq \sup_{t \in [0, 1]} \|T_1(t)\|_{op} + \sum_{j=1}^{\lceil \log_2 T \rceil} \|T_1(2^j)\|_{op}.$$

Using this and a union bound one may turn an estimate for $\|T_1(t)\|_{op}$ at a single time into an estimate for $\sup_{t \in [0, T]} \|T_1(t)\|_{op}$.

2.2 Bounds on $T_k(t)$ from $T_1(t)$

Now we use the bounds on T_1 to prove bounds on the remaining terms in the Duhamel expansion. By iterating the Duhamel formula we obtain the identity

$$e^{-itH} = e^{-itH_0} + \sum_{j=1}^k (i\beta)^j T_j(t) + i\beta \int_0^t e^{-i(t-s)H} V T_k(s) ds, \quad (2.21)$$

where T_j is the operator defined by $T_0(t) = e^{-itH_0}$ and recursively using

$$T_j(t) = \int_0^t e^{-i(t-s)H_0} V T_{j-1}(s) ds.$$

For simplicity we will see how to deal with $T_2(t)$, which is given by

$$T_2(t) = \int_{0 \leq s_1 \leq s_2 \leq t} V(s_2) V(s_1) ds_1 ds_2.$$

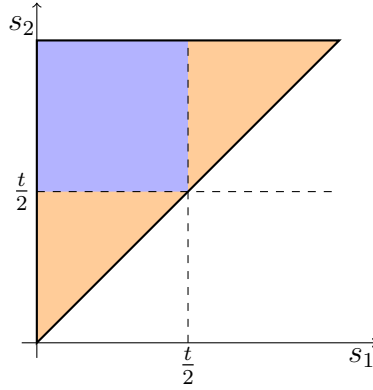


Figure 2: A schematic for the decomposition of the operator $T_2(t)$. The region in the blue square corresponds to a term of the form $T_1(t/2)^2$.

We can think of $T_2(t)$ as the integral over a triangular region in $\mathbb{R}^2$ of the operators $V(s_2)V(s_1)$, as in Figure 2.

We can subdivide this triangle into a square of side length $t/2$ and two smaller triangles. The integral over the square corresponds to the operator

$$\begin{aligned} \int_{t/2}^t \int_0^{t/2} V(s_2)V(s_1) ds_1 ds_2 &= \left(\int_{t/2}^t V(s_2) ds_2 \right) \left(\int_0^{t/2} V(s_1) ds_1 \right) \\ &= e^{-it\Delta/2} T_1(t/2) e^{it\Delta/2} T_1(t/2). \end{aligned}$$

The triangle in the lower left integrates to the operator $T_2(t/2)$ and the triangle in the upper right integrates to $e^{-it\Delta/2} T_2(t/2) e^{it\Delta/2}$. Since $e^{it\Delta}$ is unitary we have the following bound in operator norm:

$$\|T_2(t)\|_{op} \leq \|T_1(t/2)\|_{op}^2 + 2\|T_2(t/2)\|_{op}.$$

Iterating this bound yields

$$\|T_2(t)\|_{op} \leq 2t \sup_{s \in [0,1]} \|T_2(s)\|_{op} + \sum_{j=1}^{\lfloor \log_2 t \rfloor} 2^{j-1} \|T_1(t/2^j)\|_{op}^2.$$

Each of the terms in the sum is of order $\lesssim t$, so that we conclude

$$\|T_2(t)\|_{op} \lesssim t,$$

where $\lesssim$ is hiding a logarithmic factor in t .

In the exercises below we show how to extend this to higher order terms T_j .

Exercise. Prove the decomposition formula

$$T_k(s+t) = \sum_{j=0}^k T_j(s)T_{k-j}(t). \quad (2.22)$$

Exercise. Using the decomposition formula above, observe the bound

$$\|T_k(t)\|_{op} \leq \sum_{\substack{\vec{m} \in \mathbb{N}_{\geq 0}^k \\ \sum m_j = k}} \prod_{j=1}^k \|T_{m_j}(t/k)\|_{op}. \quad (2.23)$$

Assume that

$$\|T_j(t)\|_{op} \leq (C_j A^2(t \log t))^{k/2}$$

for $j < k$, and let $\bar{C} := \max_{j < k} C_j$. Using the above bound, prove that

$$\|T_k(t)\|_{op} \leq (4k^{-1/2} A \bar{C} (t \log t)^{1/2})^k.$$

Exercise. Using the above exercise, conclude that

$$\|T_k(t)\|_{op} \leq (Ck^{-1} A^2(t \log t))^{k/2}$$

with probability at least $\exp(-cA^2)$.

By summing the above estimate for $T_k(t)$, we obtain a result for e^{-itH} .

Theorem 2.10. *The family of estimates*

$$\|T_k(t)\|_{op} \leq (C\alpha k^{-1} \log(R)(t \log t))^{k/2} \quad (2.24)$$

holds with probability at least $\exp(-c\alpha^2)$. In particular, for $t \ll (\log \beta^{-1})^{-1} \beta^{-2}$, it follows that

$$\|e^{-itH} - e^{-itH_0}\|_{op} \lesssim \beta |\log \beta| \sqrt{t}. \quad (2.25)$$

2.3 Bounds for spectral projections

To conclude the chapter we show how to derive Theorem 2.3 from Corollary 2.2. First, we list some easy consequences of Corollary 2.2 (which we leave to the reader to state precisely and prove if desired):

- On $\mathbb{Z}^d / L\mathbb{Z}^d$ with $L = \lambda^{-100}$, the spectrum of H is very likely to be contained in $[-2d - \varepsilon, 2d + \varepsilon]$ for $\varepsilon = (\log \lambda^{-1})^C \lambda^2$.
- Any eigenfunction $H\psi = E\psi$ has its Fourier transform localized (in an ℓ^2 sense) to the set $\{|\omega(\xi) - E| \lesssim \lambda^2\}$.

- As a consequence of the above item, if E is not a critical value of ω and $H\psi = E\psi$ then ψ cannot have too much of its ℓ^2 mass in any ball of radius $r \ll (\log \lambda^{-1})^{-C} \lambda^{-2}$.

What we need to reach diffusive timescales (and length scales) are $\ell^p \rightarrow \ell^q$ mapping properties of the resolvent. At the moment what we have are estimates of the form

$$\|\Pi_{\delta,E}(H) - \Pi_{\delta,E}(\Delta)\|_{2 \rightarrow 2} \lesssim \lambda \delta^{-1/2}.$$

It may be surprising at first that we can say anything about $\|R(z)\|_{p \rightarrow q}$ using only this bound. The first observation is that we can transfer bounds when $p = 2$ or $q = 2$. In fact, we have the following result:

Lemma 2.11. *For $z = E + i\eta$ with $\eta \gtrsim \lambda^{-2}$,*

$$\|R^{1/2}\|_{2 \rightarrow X} \prec \|R_0^{1/2}\|_{2 \rightarrow X}$$

Proof sketch. We write

$$R^{1/2} = R_0^{1/2}(\Delta - z)^{1/2}R^{1/2}.$$

So then it suffices to prove that

$$\|(\Delta - z)^{1/2}R^{1/2}\|_{2 \rightarrow 2} \prec 1.$$

We write out a partition of unity

$$1 = \sum_j \chi_j$$

where χ_0 is a smooth cutoff to the interval $[E - \eta, E + \eta]$, and for $j \geq 1$ χ_j is supported on $|t - E| \sim 2^j \eta$. Write $\Pi_j^0 := \chi_j(\Delta)$ and $\Pi_j := \chi_j(H)$. Then

$$(\Delta - z)^{1/2}R^{1/2} = \sum_{k,\ell} (\Delta - z)^{1/2}\Pi_k^0 \Pi_\ell R^{1/2}.$$

The diagonal terms contribute a logarithmic factor:

$$\left\| \sum_k (\Delta - z)^{1/2}\Pi_k^0 \Pi_k R^{1/2} \right\| \lesssim \sum_k 2^{k/2} \eta^{1/2} \eta^{-1/2} 2^{-k/2} \lesssim \sum_k 1 \lesssim \log \eta^{-1}.$$

For the off-diagonal terms we can use Corollary 2.2. For example, for the terms $k > \ell$ we write

$$\Pi_k^0 \Pi_\ell = \Pi_k \Pi_\ell + (\Pi_k^0 - \Pi_k) \Pi_\ell = (\Pi_k^0 - \Pi_k) \Pi_\ell,$$

so that

$$\|\Pi_k^0 \Pi_\ell\|_{2 \rightarrow 2} \leq \|\Pi_k^0 - \Pi_k\| \prec \lambda 2^{-k/2} \eta^{-1/2}.$$

Therefore,

$$\|(\Delta - z)^{1/2}\Pi_0^0 \Pi_\ell R^{1/2}\|_{2 \rightarrow 2} \lesssim (2^{k/2} \eta^{1/2})(\lambda 2^{-k/2} \eta^{-1/2})(2^{-\ell/2} \eta^{-1/2}) \lesssim \lambda 2^{-\ell/2} \eta^{-1/2},$$

which is a convergent sum in ℓ . □

Now we explain how to use this. Of course the above lemma also implies

$$\|R^{1/2}\|_{X \rightarrow 2} \prec \|R_0^{1/2}\|_{X \rightarrow 2}$$

by duality. Therefore for any $p \leq 2 \leq q$ and $R = R(E + i\eta)$ and $\eta \gtrsim \lambda^2$,

$$\|R\|_{p \rightarrow q} \prec \|R_0^{1/2}\|_{p \rightarrow 2} \|R_0^{1/2}\|_{2 \rightarrow q}.$$

Moreover, if $z_0 = E + i\lambda^2$ and $z = E + i\eta$ for $\eta \ll \lambda^2$, we can write

$$R(z) = R^{1/2}(z_0)(R^{-1/2}(z_0)R(z)R^{-1/2}(z_0))R^{1/2}(z_0). \quad (2.26)$$

The operator in the brackets in the middle satisfies

$$\|R^{-1/2}(z_0)R(z)R^{-1/2}(z_0)\|_{2 \rightarrow 2} \leq \lambda^2 \eta^{-1},$$

so that for any $\eta > 0$ we have the estimate

$$\|R(z)\|_{p \rightarrow q} \prec (\lambda^2 \eta^{-1} + 1) \|R_0(z_0)\|_{p \rightarrow 2} \|R_0(z_0)\|_{2 \rightarrow q}.$$

2.3.1 Resolvent bounds for Δ and H

At this point it is clear that we need estimates for $\|R_0^{1/2}(E + i\eta)\|_{2 \rightarrow q}$. We can first decompose $R_0^{1/2}$ into spectral projections,

$$R_0^{1/2} = \sum_j R_0^{1/2} \chi_j(\Delta),$$

where χ_j on a window of width $2^j \delta$. If we had an estimate of the form

$$\|\chi_j(\Delta)\|_{2 \rightarrow q} \lesssim 2^{-j/2} \eta^{1/2},$$

then by summing over j it would follow that

$$R_0^{1/2} \lesssim \log \eta^{-1}.$$

Now $\chi_j(\Delta)$ can be explicitly written as a Fourier multiplier:

$$\widehat{\chi_j(\Delta)f(\xi)} = \chi_j(\omega(\xi)) \widehat{f(\xi)}.$$

For χ_j we have the following estimate, which is essentially equivalent to the Tomas-Stein restriction estimate [Tom75], and we repeat the proof below.

Lemma 2.12 (Tomas-Stein estimate). *Let χ be a bounded function supported in an interval $I = [E_0 - \varepsilon, E_0 + \varepsilon]$ not containing any critical values of ω . Then on $d \geq 2$ we have the estimate*

$$\|\chi(\Delta)\|_{2 \rightarrow 6} \lesssim \varepsilon^{1/2} \log \varepsilon^{-1}.$$

As a consequence of this lemma and the “transfer” argument, we have established Theorem 2.3, which we restate below.

Theorem 2.13. *The resolvent for $H = \Delta_L + \lambda V$ satisfies, for any $1 \leq p \leq \frac{6}{5}$, $6 \leq q \leq \infty$,*

$$\|R\|_{p \rightarrow q} \prec \lambda^2 \eta^{-1} + 1.$$

Proof. By a duality argument it suffices to prove that

$$\|\chi(\Delta)^2\|_{\frac{6}{5} \rightarrow 6} \lesssim \varepsilon \log \varepsilon^{-1}$$

We drop the square, just relabeling $\chi \mapsto \chi^2$. We use a decomposition of χ as follows:

$$\chi = \sum_k \phi_k,$$

where ϕ_k is smooth to scale 2^{-k} and supported in an interval of width 2^{-k} around E_0 . There are about $\log \varepsilon^{-1}$ terms in this decomposition, going from unit scale to scale ε . We have $\|\phi_k\|_\infty \lesssim \varepsilon 2^{-k}$, so that

$$\|\phi_k(\Delta)\|_{2 \rightarrow 2} \lesssim 2^k \varepsilon.$$

To prove that $\|\phi_k(\Delta)\|_{\frac{6}{5} \rightarrow 6} \lesssim 1$ it therefore suffices by interpolation to prove that

$$\|\phi_k(\Delta)\|_{1 \rightarrow \infty} \lesssim 2^{-\frac{1}{2}k} \varepsilon.$$

The operator $\phi_k(\Delta)$ is a convolution operator with kernel

$$K_k(x) = \int_{\mathbb{T}^d} e^{i\xi x} \phi_k(\omega(\xi)) \, d\xi.$$

Since the function $\phi_k(\omega(\xi))$ oscillates (and is mean-zero) on scales 2^{-k} , and is also smooth to scale 2^{-k} , we have that $|K_k(x)|$ is quite small unless $|x| \sim 2^k$. For such x , one can decompose the integral above using the coarea formula along the slices $X_E := \{\omega(\xi) = E\}$:

$$K_k(x) = \int_{[E-2^{-k}, E+2^{-k}]} \phi_k(E) \left(\int_{X_E} e^{i\xi x} \frac{1}{|\nabla \omega(\xi)|} \, d\mathcal{H}^{d-1}(\xi) \right) dE.$$

We treat the term $|\nabla \omega|^{-1}$ as just a smooth weight because we assume that E is away from critical values of ω . One can now decompose the surface X_E into coordinate charts and apply a stationary phase estimate. The fact that there is at least one direction of nonvanishing curvature at every point of X_E implies that

$$\left| \int_{X_E} e^{i\xi x} \frac{1}{|\nabla \omega(\xi)|} \, d\mathcal{H}^{d-1}(\xi) \right| \lesssim |x|^{-1/2}.$$

Therefore,

$$|K_k(x)| \lesssim 2^{-k/2} \varepsilon$$

as desired, and this completes the proof. $\square$

As a consequence, we obtain that

$$\|R_0^{1/2}(E + i\eta)\|_{2 \rightarrow 6} \lesssim \log \eta^{-1}.$$

2.3.2 ℓ^p bounds for eigenfunctions

An interesting consequence of the resolvent bounds is a bound on the ℓ^p norm of eigenfunctions of H on $\mathbb{Z}_L^d$. Indeed, let ψ_k satisfy $H\psi_k = E_k\psi_k$ and $\|\psi_k\|_2 = 1$, so that it is an ℓ^2 -normalized eigenfunction of H with energy E_k . Then

$$\psi_k = \lambda R^{1/2}(E_k + i\lambda^2)\psi_k$$

so that

$$\|\psi_k\|_{\ell^6} \leq \lambda \|R^{1/2}(E_k + i\lambda^2)\|_{2 \rightarrow 6} \lesssim \lambda.$$

This implies that on $\mathbb{Z}^2$ most eigenfunctions are of order λ in ℓ^6 norm. In higher dimensions one can obtain bounds for ℓ^p with $p_d > 2$ but $p_d \rightarrow 2$ as $d \rightarrow \infty$.

3 The Diffusive Time Scale

In this chapter of the lecture notes, we analyze timescales $t \gg \lambda^{-2}$ on which the potential now has a strong effect. Before we try to compute, it is instructive to understand what one expects to be true. For this purpose it is helpful to return to the setting of $\mathbb{R}^d$. For simplicity we consider a stationary Gaussian potential. Such a potential is completely characterized by its two-point correlation function,

$$K(x) := \mathbb{E}V(0)V(x).$$

To observe diffusion in the random Schrodinger equation

$$i\partial_t\psi = \Delta\psi + \lambda V\psi,$$

one looks at the evolution of observables $\langle\psi_t, \text{Op}(a_0)\psi_t\rangle$, where $a_0 \in C^\infty(\mathbb{R}_x^d \times \mathbb{R}_\xi^d)$ is a smooth function on phase space and $\text{Op}(a_0) \in \mathcal{B}(L^2(\mathbb{R}^d))$ is the quantization of it (for example, using the Weyl quantization). One can derive heuristically (see for example [Spo77]) the following effective kinetic equation for the phase space density $a_t(x, p)$:

$$\partial_t a_t(x, \xi) + \xi \cdot \nabla_x a_t(x, \xi) = \lambda^2 \int_{\mathbb{R}^d} \delta(|\xi|^2/2 - |\xi'|^2/2) \widehat{K}(\xi - \xi') [a(x, \xi') - a(x, \xi)] d\xi'.$$

The term on the left hand side describes free transport by particles along lines (solving $\dot{x} = \xi$ and $\dot{\xi} = 0$). The right hand side is a collision term that scatters particles at momentum ξ' to momentum ξ (and vice-versa). The λ^2 gives the effective strength of the potential (corresponding to the fact that, up to time λ^{-2} , the potential is not noticed), the $\delta(|\xi|^2/2 - |\xi'|^2/2)$ term enforces conservation of kinetic energy, and $\widehat{K}(\xi - \xi')$ is the scattering kernel corresponding to the potential V . This limit was established first in the spatially homogeneous case and for small kinetic times by Spohn [Spo77], then for the full linear Boltzmann equation and at arbitrary kinetic times in [EY00], and finally for diffusive time scales in [ESY08].

Note that for V smooth, $\widehat{K}$ is localized so that each scattering event can only deflect a particle by a small angle. In contrast, on $\mathbb{Z}^d$ one has $K(x) = \delta_0(x)$ where δ_0 is a Kronecker delta at the origin so that $\widehat{K}(\xi) = 1$. Moreover, $|\xi|^2/2$ is replaced by the dispersion relation $\omega(\xi)$. Therefore, after a scattering event the momentum should be uniformly distributed on a kinetic energy shell $\{\omega(\xi) = E\}$ (weighted by $|\nabla\omega|^{-1}$ to account for the small thickness of the shell). Thus the position of the particle at time T is (classically) described by a random walk of $\sim \lambda^2 T$ steps of size λ^{-2} . Strictly speaking the step distribution depends on the energy shell E , so that $p_t(x) = |e_{0x}^{itH}|^2$ should resemble not a single Gaussian but a mixture of Gaussian distributions with different variances, all on the order $\lambda^{-2}t$. Nevertheless, it is not a bad heuristic to imagine that

$$|e_{0x}^{itH}|^2 \approx (\lambda^{-2}t)^{-d/2} \exp(-\lambda^2|x|^2/t). \quad (3.1)$$

For reasons that we will see later, it is more convenient to work with the resolvent $R(z) = (H - z)^{-1}$ than with the propagator e^{-itH} . We write $z = E + i\eta$, where E selects the energy and $\eta \rightarrow 0$ is a small parameter. As a function of E , $f(x) = \frac{1}{x - E + i\eta}$ is smooth to scale η , so by Corollary 2.2 we have $R(E + i\eta) \approx R_0(E + i\eta)$ for $\eta \gg \lambda^2$. Indeed, $R(z)$ has a nice explicit expression in terms of the propagator:

$$R(z) = \int_0^\infty e^{iEt} e^{-\eta t} e^{itH} dt. \quad (3.2)$$

A heuristic for the entries of $|R(z)_{0x}|^2$ is that the entries of e_{0x}^{itH} are completely decorrelated. Combined with (3.1) we arrive at the guess

$$\begin{aligned} |R(z)_{0x}|^2 &\approx \int_0^\infty e^{-2\eta t} |e_{0x}^{itH}|^2 dt \\ &\approx \int_0^\infty e^{-2\eta t} (\lambda^{-2}t)^{-d/2} \exp(-\lambda^2|x|^2/t) dt. \end{aligned}$$

Taking the limit $\eta \rightarrow 0$ we obtain, for $d \geq 3$, the heuristic

$$|R(z)_{0x}|^2 \approx \lambda^2 |x|^{2-d}. \quad (3.3)$$

Although (3.3) and (3.1) are heuristically related, it is not clear that either implies the other. Nevertheless, we do have by applying Parseval's theorem to the identity (3.2)

$$\eta \int_0^\infty e^{-\eta t} |e_{0x}^{itH}|^2 dt = \int_{-\infty}^\infty |R_{0x}(E + i\eta)|^2 dE.$$

Thus by summing over an annulus at the diffusive scale $\lambda^{-1}\eta^{-1/2}$ we can deduce *time-averaged* information about $r(t)$ from the distribution of the mass in the rows of the resolvent.

In particular, Theorem 1.1 is a consequence of the following estimate.

Theorem 3.1. *For any $\varepsilon, \delta > 0$ there exists c, C such that for $E \in [-2d, 2d]$ with $d(E, \Sigma_d) > \varepsilon$ and $\eta > \lambda^{2.1-\delta}$,*

$$\sum_{|x| > c\lambda\eta^{-1/2}} |R_{0x}(E + i\eta)|^2 \geq c\eta^{-1}$$

holds with probability at least $1 - C\lambda^{100}$.

3.1 The direct perturbative approach

As a first try, let's see how one might go about understanding the resolvent directly from perturbation theory. To this end we use the resolvent identity,

$$(H_0 + W - z)^{-1} = (H_0 - z)^{-1} - (H_0 - z)^{-1}W(H_0 + W - z)^{-1}.$$

This can be checked by multiplying both sides on the right by $(H_0 + W - z)$. Applying this identity with $H_0 = \Delta$ and $W = \lambda V$ we arrive at the identity

$$(H - z)^{-1} = (\Delta - z)^{-1} - (\Delta - z)^{-1}(\lambda V)(H - z)^{-1}.$$

Writing $R(z) = (H - z)^{-1}$ and $R_0(z) = (\Delta - z)^{-1}$ and iterating this formula, we arrive at the *Born series* expansion for $R(z)$,

$$R(z) = R_0(z) + \sum_{j=1}^{\infty} (-1)^j R_0(z) (\lambda V R_0(z))^j.$$

This is algebraically much simpler than the Dyson series expansion we used in the previous chapter, and this is the main reason to use $R(z)$ instead of the propagator e^{itH} .

Note that the j -th term in the expansion above can be rearranged to

$$R_0^{1/2} (\lambda R_0^{1/2} V R_0^{1/2})^j R_0^{1/2}.$$

The term in brackets is a random matrix that is linear in the randomness, and is in a form that is amenable to applying the non-commutative Khintchine inequality. The naive bound for the operator, using $\|R_0^{1/2}\|_{\ell^2 \rightarrow \ell^2} = \eta^{-1/2}$ (remember that $z = E + i\eta$, $\eta > 0$), is $\|R_0^{1/2} V R_0^{1/2}\| \lesssim \eta^{-1}$. However, by applying the non-commutative Khintchine inequality, one can obtain a square-root cancellation for this operator, and instead get

$$\|R_0^{1/2} V R_0^{1/2}\|_{\ell^2 \rightarrow \ell^2} \lesssim \eta^{-1/2}.$$

Therefore, for $\eta \gg \lambda^2$, the term $\lambda R_0^{1/2} V R_0^{1/2}$ is of lower order than the main term $R_0(z)$.

Taking η into the diffusive regime $\eta \ll \lambda^2$ requires *renormalization*. In this context, renormalization is simply the fact that one is not forced to take $H_0 = \Delta$, but could instead take $H_0 = \Delta + \Theta$ for any operator Θ we choose. It turns out that for the Anderson model, we can take $\Theta = \lambda^2 \theta \text{Id}$ for a scalar θ (we will see how one might renormalize more general models soon). The scaling λ^2 is for convenience (it will turn out that θ is an $O(1)$ -sized quantity).

To explain the following calculation it is useful to introduce a diagrammatic notation. In this notation, the resolvent is represented by a solid line,

$$R = \text{———},$$

and we define $M = (\Delta - (z + \lambda^2 \theta))^{-1}$, represented by a dashed line, $M = \text{-----}$. With this choice of renormalization we have the Born series expansion

$$R(z) = M + \sum_{j=1}^{\infty} (-1)^j M ((\lambda V - \lambda^2 \theta) M)^j. \quad (3.4)$$

Diagrammatically, this is represented by

$$\text{————} = \text{-----} + \text{---}\otimes\text{---} + \text{---}\otimes\text{---}\otimes\text{---} + \dots,$$

where we use the circled cross to represent $\lambda V - \lambda^2 \theta$ (the cross represents λV and the circle $\lambda^2 \theta$). Upon taking an expectation the term with a single cross vanishes and the two collisions in the third term are paired by the Wick rule, so the first few terms can be written as

$$\mathbb{E}\text{————} = \text{-----} + \text{---}\circ\text{---} + \text{---}\overset{\frown}{\star}\text{---}\star\text{---} + \dots.$$

The latter two terms can be made to cancel if we choose θ to solve the self-consistent equation

$$\overset{\frown}{\star}\text{---}\star + \circ = 0, \tag{3.5}$$

which can also be written as

$$\lambda^2 \mathbb{E} V R V = \lambda^2 \theta \text{Id},$$

and taking a diagonal entry this produces the following defining equation for θ :

$$\theta = M_{00} = (\Delta - (z + \lambda^2 \theta))_{00}^{-1}.$$

Returning to the diagrammatic interpretation of the self-consistent equation, we may use it to simplify away self-loops in more complicated terms coming from the Wick expansion of higher order terms in the Born series. For example, using (3.5) we have the identity

$$\text{---}\overset{\frown}{\star}\text{---}\star\text{---}\star\text{---}\star\text{---}\star\text{---} + \text{---}\overset{\frown}{\star}\text{---}\circ\text{---}\star\text{---}\star\text{---}\star\text{---} = 0.$$

The terms that do not cancel involve crossings. There are *many* crossing terms, the simplest of which is given by

$$\text{---}\overset{\frown}{\star}\text{---}\star\text{---}\star\text{---}\star\text{---} = \lambda^4 \sum_{x,y} M |x\rangle M_{xy}^2 M_{yx} \langle y| M$$

A “naive” estimate would indicate that this term has size $\lambda^4 \eta^{-2}$, but in fact the crossing induces an additional cancellation such that the term actually has size $\lambda^4 \eta^{-1}$ (although this is delicate). In [ESY07a; ESY07b] used combinatorial arguments and estimates from harmonic analysis to show that the sum of all crossing terms has size on the order $\lambda^{2+\varepsilon} \eta^{-1}$. In fact, they handled second moments of the terms in the *Dyson* series which is even more complicated, but their methods should of course apply to the simpler problem of computing the expectation of the resolvent.

3.2 The resolvent of a Wigner matrix

We are not going to try to directly estimate the crossing terms. Instead we take inspiration from random matrix theory, where crossing terms also appear but elegant methods have been developed to handle them. In particular, we use the framework of *self-consistent equations* which was developed to understand the local spacing of eigenvalues in Wigner matrices [ESY09]. An excellent exposition of this approach is provided in [Erd+13a]. We will present a suboptimal result which nevertheless suffices for our purposes.

The simplest example of a random matrix is the GOE ensemble, having independent (symmetric) Gaussian entries of variance 1 above the diagonal and variance 2 on the diagonal. We can express such matrices as

$$H_{\text{GOE}} = \frac{1}{\sqrt{N}} \sum_{1 \leq i \leq j \leq N} g_{ij} E_{ij}, \quad (3.6)$$

where

$$E_{ij} = \begin{cases} \sqrt{2} e_i \otimes e_i, & i = j, \\ e_i \otimes e_j + e_j \otimes e_i, & i \neq j. \end{cases}$$

For convenience we use the bra-ket notation $|i\rangle \langle j|$ to mean $e_i \otimes e_j$.

The distribution of H_{GOE} is invariant under conjugation by a rotation matrix, so the eigenvectors are uniformly distributed from the unit sphere. What is more interesting is to compute the distribution of eigenvalues of H_{GOE} . Let λ_k be the set of eigenvalues and ψ_k the eigenvectors, so that $H_{\text{GOE}} \psi_k = \lambda_k \psi_k$. We define the empirical measure of eigenvalues μ by

$$\mu = \sum_k \delta_{\lambda_k}.$$

Then we can write

$$R_{\text{GOE}}(z) = (H_{\text{GOE}} - z)^{-1} = \sum_k \frac{1}{\lambda_k - z} \psi_k \otimes \psi_k.$$

Taking a trace, we have

$$\text{tr}[R_{\text{GOE}}(z)] = \sum_k \frac{1}{\lambda_k - z} = \int \frac{1}{t - z} d\mu(t).$$

Writing $z = E + i\eta$ with $\eta > 0$ and taking the imaginary part, we have

$$-\text{Im tr}[R_{\text{GOE}}(z)] = \int \frac{\eta}{(E - t)^2 + \eta^2} d\mu(t).$$

The right hand side is a convolution of μ with the Poisson kernel, which (up to a scaling by π), is an approximate identity. More precisely, $R_{\text{GOE}}(E + i\eta)$ encodes (a smoothed version of) the eigenvalue count in the interval $[E - \eta, E + \eta]$. It is therefore of interest to compute the diagonal elements of $R_{\text{GOE}}(E + i\eta)$ with η as small as possible.

An efficient way to compute R_{GOE} is to use the “cavity method”. That is, one introduces a “cavity” by deleting one site from $[N]$. This corresponds to deleting a row and a column of the matrix, resulting in a sample from the $(N - 1) \times (N - 1)$ GOE ensemble. The Schur complement formula can be used to relate the resolvent from the smaller ensemble to the resolvent of the full ensemble. Moreover, these resolvents can be shown to be closely related (for example, one has the eigenvalue interlacing property). This strategy uses strongly the symmetries of the model and is not well suited for working with the Anderson model (where deleting one site of $\mathbb{Z}^d$ results in a different graph that no longer has translation symmetry).

Fortunately there is an alternative approach that uses only (1) the resolvent identity and (2) Gaussian integration by parts. First we use the resolvent identity using the renormalization $H - z = -(m(z)\text{Id} + z) + (H + m(z)\text{Id})$ (we drop the Id and the dependence on z below):

$$R = -(m + z)^{-1} + (m + z)^{-1}(H + m)R. \quad (3.7)$$

At this point one *could* iterate this formula and obtain an expansion for R_{GOE} , resulting in an expression analogous to (3.4). However there is an alternative approach using the Gaussian integration by parts formula $\mathbb{E}Z_i f(Z) = \mathbb{E}\partial_i f(Z)$. Indeed, we can use the definition of H_{GOE} (3.6) and the resolvent identity to compute

$$\frac{\partial}{\partial g_{ij}} R = -\frac{1}{\sqrt{N}} R E_{ij} R.$$

Thus, taking an expectation on both sides of (3.7) we obtain

$$\begin{aligned} \mathbb{E}R &= -(m + z)^{-1} + (m + z)^{-1}(\mathbb{E}HR + \mathbb{E}mR) \\ &= -(m + z)^{-1} + (m + z)^{-1}(-N^{-1} \sum_{i \leq j} \mathbb{E}E_{ij} R E_{ij} R + \mathbb{E}mR). \end{aligned}$$

To make sense of this we define the superoperator $\mathcal{A}_{\text{GOE}}$

$$\begin{aligned} \mathcal{A}_{\text{GOE}}[B] &:= N^{-1} \sum_{i \leq j} E_{ij} B E_{ij} \\ &= N^{-1} \sum_i 2B_{ii} e_i \otimes e_i + N^{-1} \sum_{i \neq j} B_{ii} e_j \otimes e_j + B_{ij} e_j \otimes e_i \\ &= N^{-1} B^\top + N^{-1} \text{tr}[B] \text{Id}. \end{aligned}$$

Therefore we have

$$\mathbb{E}R = -(m + z)^{-1}(\text{Id} + N^{-1}\mathbb{E}R^\top R + \mathbb{E}(N^{-1}\text{tr}[R] - m)R). \quad (3.8)$$

Using that $\|R\|_{op} \leq \eta^{-1}$, we have that $N^{-1}\mathbb{E}R^\top R$ is bounded by $N^{-1}\eta^{-2}$ in operator norm (which is small so long as we take $\eta \gg N^{-1/2}$). We can make the final term small if we take $m = \mathbb{E}N^{-1}\text{tr} R = \mathbb{E}R_{00}$ and if we can show that $N^{-1}\text{tr} R$ is concentrated around its mean. We will see that this is the case in a moment, but first let us complete the calculation. Taking the diagonal entries of (3.8) above and setting $m = \mathbb{E}R_{00}$, we have that m solves

$$m = -(m + z)^{-1}(1 + \varepsilon)$$

where

$$\varepsilon := N^{-1}\mathbb{E}(R^\top R)_{00} + \mathbb{E}(N^{-1}\text{tr}[R - \mathbb{E}R])R_{00}.$$

Using the quadratic formula, the solution m is given by

$$m = \frac{-z \pm \sqrt{z^2 - 4(1 + \varepsilon)}}{2}.$$

Now if we set $z = E + i\eta$ for small $\eta \ll 1$, we obtain (neglecting η and ε)

$$\text{Im } m \approx \text{Im } \sqrt{E^2/4 - 1} = \begin{cases} \sqrt{1 - (E/2)^2}, & |E| \leq 2 \\ 0, & |E| > 2. \end{cases}$$

This is the semicircular law.

To complete the proof we need to show that $N^{-1}\text{tr}[R]$ is concentrated around its mean. To simplify the calculation we will actually show that R_{00} is concentrated, and this is done by computing the Lipschitz constant of R_{00} as a function of g_{ij} . We compute:

$$\begin{aligned} |\nabla R_{00}|^2 &= \sum_{i \leq j} \left| \frac{\partial}{\partial g_{ij}} R_{00} \right|^2 \\ &= N^{-1} \sum_{i \leq j} |(RE_{ij}R)_{00}|^2 \\ &\lesssim N^{-1} \sum_{i,j} |R_{0i}|^2 |R_{0j}|^2 \\ &= N^{-1} \left(\sum_i |R_{0i}|^2 \right)^2 \\ &= N^{-1} \eta^{-2} (\text{Im } R_{00})^2. \end{aligned}$$

Deterministically we have $\text{Im } R_{00} \leq \eta^{-1}$, so we have

$$|\nabla R_{00}|^2 \lesssim \eta^{-4} N^{-1}.$$

Therefore,

$$\mathbb{P}(|R_{00} - \mathbb{E}R_{00}| \geq K\eta^{-2}N^{-1/2}) \leq \exp(-cK^2).$$

We have therefore proved the following:

Proposition 3.2. *For H_{GOE} as in (3.6) and $z = E + i\eta$, with $E \in (-2, 2)$ we have that*

$$\mathbb{P}(|\operatorname{Im} \operatorname{tr} R(z) - \sqrt{1 - (E/2)^2}| \geq K\eta^{-2}N^{-1/2}) \leq \exp(-cK^2).$$

This kind of result is called a “local law”, as it provides information about the local density of eigenvalues at a microscopic scale. The optimal local law provides information for imaginary parts $\eta \gg N^{-1}$, and thus nearly captures the location of individual eigenvalues. The bound above only provides information down to scale $\eta \gg N^{-1/4}$.

3.2.1 More general ensembles

The nice thing about the derivation of the semicircular law that we just saw is that it is rather robust, not relying on any particular symmetry of the GOE ensemble. Let’s go through the calculation again, this time for a more general Gaussian random matrix of the form

$$H = A_0 + \sum_j g_j A_j.$$

In the case of Wigner random matrices we have $A_0 = 0$ and $A_{ij} = E_{ij}$, and in the case of the Anderson tight-binding model we will take $A_0 = \Delta_{\mathbb{Z}^d}$ and $A_j = |j\rangle\langle j|$. Other lattices can be accommodated for by changing A_0 , and nontrivial covariance structure can be added to the potential by changing A_j .

We renormalize by writing $H - z = (A_0 - (z + M)) + (\sum_j g_j A_j + M)$, and so by the resolvent identity have

$$R = (A_0 - (z + M))^{-1} - (A_0 - (z + M))^{-1} \left(\sum_j g_j A_j + M \right) R.$$

Taking an expectation and performing Gaussian integration by parts we arrive at

$$\mathbb{E}R = (A_0 - (z + M))^{-1} - (A_0 - (z + M))^{-1} (\mathbb{E}(M - \mathcal{A}[R])R),$$

where $\mathcal{A}$ is the superoperator

$$\mathcal{A}[B] := \sum_j A_j B A_j.$$

Setting $M = \mathcal{A}[\mathbb{E}R]$ we arrive at the following approximate self-consistent equation for $\mathbb{E}R$.

$$\mathbb{E}R = (A_0 - (z + \mathcal{A}[\mathbb{E}R]))^{-1} + (A_0 - (z + \mathcal{A}[\mathbb{E}R]))^{-1} (\mathbb{E}\mathcal{A}[R - \mathbb{E}R]R). \quad (3.9)$$

If $\mathcal{A}[R]$ is sufficiently concentrated, we can treat the second term as an error term. We then arrive at an (approximate) self-consistent equation for $\Theta := \mathbb{E}\mathcal{A}[R]$

$$\Theta = \mathcal{A}[(A_0 - (z + \Theta))^{-1}] + \mathfrak{E}, \quad (3.10)$$

where

$$\mathfrak{E} = \mathcal{A}[(A - (z + \Theta))^{-1} \mathbb{E} \mathcal{A}[R - \mathbb{E} R] R].$$

Before we specialize to the Anderson model, we stop to make an observation that applies generally. The formal calculation we started with for the Anderson model suggests that $\mathfrak{E}$ should encode the “crossing” terms, and it is not clear from the expression above what $\mathfrak{E}$ has to do with crossings. To see the connection, we will derive an alternative expression for $\mathbb{E} \mathcal{A}[R - \mathbb{E} R] R$. Note first that, if R' is an independent sample of the resolvent, then

$$\mathbb{E} \mathcal{A}[R - \mathbb{E} R] R' = 0.$$

To make use of this we interpolate between R and R' in the second factor. Define R^q to be the resolvent of H^q , where

$$H^q = A_0 + \sum_j (g_j \sqrt{1-q} + g'_j \sqrt{q}) A_j,$$

and g'_j are independent samples of g_j . Then $R = R^0$ and $R' = R^1$, so using the fundamental theorem of calculus to interpolate as in [BBH23] we have

$$\begin{aligned} \mathbb{E} \mathcal{A}[R - \mathbb{E} R] R &= \mathbb{E} \mathcal{A}[R^0 - \mathbb{E} R^0] R^1 + \int_0^1 \frac{d}{dq} \mathbb{E} \mathcal{A}[R^q - \mathbb{E} R^q] R^1 dq \\ &= \int_0^1 \frac{d}{dq} \mathbb{E} \mathcal{A}[R^q] R^1 dq. \end{aligned}$$

Then we compute, using the resolvent identity and another application of GIBP:

$$\begin{aligned} \frac{d}{dq} \mathbb{E} \mathcal{A}[R^q] R^1 &= \sum_j \frac{1}{2\sqrt{q}} \mathbb{E} g'_j \mathcal{A}[R^q A_j R^q] R^1 - \frac{1}{2\sqrt{1-q}} \mathbb{E} g_j \mathcal{A}[R^q A_j R^q] R^1 \\ &= \sum_j \mathbb{E} \mathcal{A}[R^q A_j R^q] R^1 A_j R^1 \\ &= \sum_{j,k} \mathbb{E} A_k R^q A_j R^q A_k R^1 A_j R^1. \end{aligned}$$

The final expression above is a “crossing” term, because the pairing of the A_j and A_k are crossed. This explains that although $\mathfrak{E}$ encodes the contributions of all of the crossing terms, it can be bounded without reference to any crossings at all.

3.3 The local law for the Anderson model

We are finally ready to start thinking about the Anderson model on $\mathbb{Z}^d$, defined by the Hamiltonian

$$H = \Delta + \lambda \sum_{j \in \mathbb{Z}^d} g_j |j\rangle \langle j|.$$

It is more convenient for us to work on the torus $\mathbb{Z}_L^d := \mathbb{Z}^d / L\mathbb{Z}^d$, with $L = \lceil \lambda^{-100} \rceil$ (as in the previous chapter).

The superoperator $\mathcal{A}$ corresponding to this matrix ensemble is $\lambda^2 \mathcal{D}$, where the diagonal superoperator $\mathcal{D}$ is defined by

$$\mathcal{D}[B] = \sum_i |i\rangle \langle i| B |i\rangle \langle i| = \sum_i B_{ii} |i\rangle \langle i|.$$

Therefore, defining $\tilde{\theta}(z) := \mathbb{E}R_{00}(z)$, the general self-consistent equation (3.9) becomes

$$\mathbb{E}R = (\Delta - (z + \lambda^2 \tilde{\theta}))^{-1} + \lambda^2 (\Delta - (z + \lambda^2 \tilde{\theta}))^{-1} \mathbb{E} \mathcal{D}[R - \mathbb{E}R] R.$$

Taking the diagonal entry of both sides we arrive at the following approximate self-consistent equation for $\tilde{\theta}$ (which is the analogue of (3.10)):

$$\tilde{\theta} = (\Delta - (z + \lambda^2 \tilde{\theta}))_{00}^{-1} + \lambda^2 ((\Delta - (z + \lambda^2 \tilde{\theta}))^{-1} \mathbb{E} \mathcal{D}[R - \mathbb{E}R] R)_{00}. \quad (3.11)$$

As a first step towards proving quantum diffusion we will show that, for $z = E + i\eta$ and $\eta \gg \lambda^{2+\frac{1}{6}}$, $\tilde{\theta}$ is well approximated by the solution to

$$\theta = (\Delta - (z + \lambda^2 \theta))_{00}^{-1}. \quad (3.12)$$

To state our result precisely it is convenient to introduce stochastic domination notation. We say that B stochastically dominates A , written $A \prec B$, if for any $\varepsilon, N > 0$ there exists $C(\varepsilon, N)$ such that

$$A \leq C(\varepsilon, N) \lambda^\varepsilon B$$

holds with probability at least $1 - C(\varepsilon, N) \lambda^N$.

Theorem 3.3. *Fix $\varepsilon, \delta > 0$, and let $E \in [-2d, 2d]$ with $d(E, \Sigma_d) > \varepsilon$. Then for $z = E + i\eta$ with $\eta \gg \lambda^{2+\frac{1}{6}-\delta}$,*

$$|\mathbb{E}R_{00} - \theta| \leq \lambda^{1/2-\delta} (\lambda \eta^{-1})^3$$

and moreover

$$|R_{xy} - \mathbb{E}R_{xy}| \prec \lambda^{1/2} (\lambda \eta^{-1})^2. \quad (3.13)$$

Note that the above theorem does not say anything interesting about the values of the resolvent far away from the diagonal, so it is not yet clear how this is useful for Theorem 3.1.

There are two steps to establishing Theorem 3.3. The first step is to understand the exact self-consistent equation

$$\theta = (\Delta - (z + \lambda^2 \theta))_{00}^{-1}.$$

In particular, one would like to know that a solution θ exists, is unique, and is bounded before one can try to understand what happens to a perturbation.

The second step is to bound the error term in (3.11). The important term is the matrix $\mathcal{D}[R - \mathbb{E}R]$, which is a diagonal matrix with entries $R_{xx} - \mathbb{E}R_{xx}$. Thus we can obtain a bound in operator norm for this matrix by proving a concentration inequality for the entries R_{xx} . This concentration inequality is where we use the $\|R\|_{p \rightarrow q}$ bounds proven in the previous chapter.

3.3.1 The self-consistent equation for θ

First we analyze the equation

$$\theta = (\Delta - (z + \lambda^2 \theta))_{00}^{-1}. \quad (3.14)$$

To this end it is useful to understand the function

$$F(z) := (\Delta - z)_{00}^{-1}.$$

Using the Fourier transform, one can confirm that $F(z)$ is given by

$$F(z) = \int_{\mathbb{T}^d} \frac{d\xi}{\omega(\xi) - z}.$$

This is an analytic function of z , and using the coarea formula we can rewrite the imaginary part as follows:

$$\begin{aligned} \operatorname{Im} F(z) &= \int_{\mathbb{T}^d} \frac{\eta d\xi}{(\omega(\xi) - E)^2 + \eta^2} \\ &= \int_{-2d}^{2d} \frac{\eta}{(E' - E)^2 + \eta^2} \left(\int_{\{\omega(\xi)=E'\}} \frac{1}{|\nabla \omega(\xi)|} d\mathcal{H}^{d-1} \right) dE' \\ &=: \int_{-\infty}^{\infty} \frac{\eta}{(E' - E)^2 + \eta^2} \rho(E') dE', \end{aligned}$$

where we define the density of states ρ

$$\rho(E) := \int_{\{\omega(\xi)=E\}} \frac{1}{|\nabla \omega(\xi)|} d\mathcal{H}^{d-1}. \quad (3.15)$$

Thus,

$$\rho(E) = \lim_{\eta \rightarrow 0} \frac{1}{\pi} \operatorname{Im} F(E + i\eta).$$

Since $F(z)$ is analytic, we have that

$$\lim_{\eta \rightarrow 0} \frac{1}{\pi} F(E + i\eta) = H\rho(E) + i\rho(E), \quad (3.16)$$

where H is the Hilbert transform. We note the following properties of the function ρ which can be verified from the integral formula (3.15):

Lemma 3.4. *The function ρ given by (3.15) is supported in $[-2d, 2d]$. Moreover, $\rho \in C^\infty([-2d, 2d] \setminus \Sigma_d)$, where the Σ_{crit} is the set of critical values $\Sigma_{\text{crit}} := [-2d, 2d] \cap (2d + 4\mathbb{Z})$. Moreover, $\rho > 0$ on $(-2d, 2d)$. Finally, ρ satisfies the bound*

$$\rho(E) \leq C_d(1 + \log(E^{-1})\mathbf{1}_{d=2}).$$

Then using the formula (3.16) and also the fact that $F(E + i\eta)$ is given by a convolution of $F(E + i0^+)$ with the Poisson kernel, we have the following bounds for F :

Lemma 3.5. *For $\eta > 0$, the function F satisfies*

$$|F(E + i\eta)| \lesssim |\log \eta^{-1}|. \quad (3.17)$$

Moreover, on the domain $D_\varepsilon := \{E + i\eta \mid \eta > 0, E \in [-2d, 2d], d(E, \Sigma_d) > \varepsilon\}$, we have

$$\|F\|_{C^1(D_\varepsilon)} \leq C_\varepsilon. \quad (3.18)$$

As a consequence, we can prove a uniqueness result for solutions to (3.11).

Lemma 3.6. *Let $z = E + i\eta$ for $E \in [-2d, 2d]$, $d(E, \Sigma_d) > \varepsilon$, and $\eta > \lambda^{10}$. Suppose that $\theta' = \theta'(z)$ solves*

$$\theta' = F(z + \lambda^2 \theta') + \gamma$$

for some error $|\gamma| < 1$. Then for $\lambda < \lambda(\varepsilon)$, we have the estimate Then

$$|\theta' - F(z)| \lesssim |\gamma|.$$

Proof. First, by (3.17) we have the estimate

$$|\theta'| \lesssim \log \lambda^{-1}.$$

But then, for λ small enough it follows that $B_{\lambda^2 \theta'}(z) \subset D_\varepsilon$ where D_ε is the domain defined in Lemma 3.5. By the C^1 regularity estimate (3.18) it therefore follows that

$$|\theta' - F(z)| \leq |F(z + \lambda^2 \theta') - F(z)| + |\gamma| \lesssim \lambda^2 |\theta'| + |\gamma|.$$

Since $F(z)$ is bounded, we conclude that

$$|\theta' - F(z)| \lesssim \lambda^2 + |\gamma|.$$

□

3.3.2 Concentration of the resolvent entries

In this section we prove a concentration inequality for R_{xy} . First we recall the Gaussian Poincaré inequality.

Lemma 3.7 (Gaussian Poincaré). *For $f \in C^\infty(\mathbb{R}^N)$ and $Z = (Z_1, \dots, Z_N)$ N independent standard Gaussian random variables,*

$$\text{Var } f(Z) \leq \mathbb{E} \sum_j |\partial_{Z_j} f(Z)|^2 =: \mathbb{E} |\nabla f|^2.$$

Applying the Gaussian Poincaré inequality to $|f|^k$ we have the following result for higher order moments:

Lemma 3.8. *For $f \in C^\infty(\mathbb{R}^N)$ with $\mathbb{E} f = 0$, we have for $k \geq 1$ the bound*

$$\mathbb{E} |f|^{2k} \leq (Ck)^{2k} \mathbb{E} |\nabla f|^{2k}.$$

We will apply Lemma 3.8 to the resolvent entries R_{xy} , thought of as functions of the Gaussian potential g_x . Therefore we compute

$$\begin{aligned} |\nabla R_{xy}|^2 &= \sum_w \left| \frac{\partial}{\partial g_w} R_{xy} \right|^2 \\ &= \lambda^2 \sum_w |R_{xw} R_{wy}|^2 \\ &\leq \lambda^2 \left(\sum_w |R_{xw}|^4 \right)^{1/2} \left(\sum_w |R_{yw}|^4 \right)^{1/2} \\ &\leq \lambda^2 \|R\|_{1 \rightarrow 4}^4. \end{aligned} \tag{3.19}$$

By Theorem 2.3, we have $\|R\|_{1 \rightarrow 6} \prec \lambda^2 \eta^{-1}$ (we are taking $\eta \ll \lambda^2$). Therefore (recalling $R = R(z) = R(E + i\eta)$),

$$\begin{aligned} \|R\|_{1 \rightarrow 4} &\leq \|R\|_{1 \rightarrow 2}^{1/4} \|R\|_{1 \rightarrow 6}^{3/4} \\ &\leq \|R^{1/2}\|_{1 \rightarrow 2}^{1/4} \|R^{1/2}\|_{2 \rightarrow 2}^{1/4} \|R\|_{1 \rightarrow 6}^{3/4} \\ &\prec (\lambda \eta^{-1/2})^{1/4} (\eta^{-1/2})^{1/4} (\lambda^2 \eta^{-1})^{3/4} \end{aligned}$$

Combined with Lemma 3.8 and rearranging we conclude the following:

Lemma 3.9. *Let $R = R(z) = R(E + i\eta)$ for a good energy E and $\eta \ll \lambda^2$. Then*

$$|R_{xy} - \mathbb{E} R_{xy}| \prec \lambda^{1/2} (\lambda^2 \eta^{-1})^2$$

3.3.3 Conclusion of the proof of Theorem 3.3

We can now apply Lemma 3.9 to analyze the self-consistent equation (3.11) which we recall

$$\tilde{\theta} = (\Delta - (z + \lambda^2 \tilde{\theta}))_{00}^{-1} + \lambda^2 ((\Delta - (z + \lambda^2 \tilde{\theta}))^{-1} \mathbb{E} \mathcal{D}[R - \mathbb{E}R] R)_{00}.$$

The error term on the right, $\mathfrak{e}$, is bounded as follows:

$$\begin{aligned} |\mathfrak{e}| &= \lambda^2 |\langle 0 | (\Delta - (z + \lambda^2 \tilde{\theta}))^{-1} \mathbb{E} \mathcal{D}[R - \mathbb{E}R] R | 0 \rangle| \\ &\leq \lambda^2 \|(\Delta - (z + \lambda^2 \tilde{\theta}))^{-1}\|_{1 \rightarrow 2} \mathbb{E} \|\mathcal{D}[R - \mathbb{E}R] R\|_{1 \rightarrow 2}. \end{aligned}$$

For the calculation below we will assume that we already know that $\tilde{\theta} = O(1)$ and that $\text{Im } \tilde{\theta} \gtrsim 1$. These bounds are *outputs* of our argument and at first this may appear circular, and this circularity will be addressed shortly. For now we will pretend we have this information already in hand. To bound $\|(H - z)^{-1}\|_{1 \rightarrow 2}$ we use the Ward identity, valid for any symmetric H and $\eta > 0$:

$$\sum_y |(H - z)_{xy}^{-1}|^2 = \eta^{-1} \text{Im}(H - z)_{xx}^{-1}.$$

In particular,

$$\|R\|_{1 \rightarrow 2}^2 = \max_x \sum_y |R_{xy}|^2 \leq \eta^{-1} \max_x |R_{xx}|.$$

Applying this with $H = \Delta$, and assuming that $\text{Im } \tilde{\theta} \gtrsim 1$ we therefore have

$$\|(\Delta - (z + \lambda^2 \tilde{\theta}))^{-1}\|_{1 \rightarrow 2} \lesssim \lambda^{-1}.$$

Therefore

$$\begin{aligned} |\mathfrak{e}| &\lesssim \lambda \left(\mathbb{E} \|\mathcal{D}[R - \mathbb{E}R]\|_{2 \rightarrow 2}^2 \right)^{1/2} \left(\mathbb{E} \|R\|_{1 \rightarrow 2}^2 \right)^{1/2} \\ &\lesssim \lambda^{-\delta} \lambda \left(\mathbb{E} (\sup_x |R_{xx} - \mathbb{E}R_{xx}|^2) \right)^{1/2} (\lambda \eta^{-1}) \\ &\lesssim \lambda^{\frac{1}{2} - \delta} (\lambda^2 \eta^{-1})^3. \end{aligned}$$

In conclusion, we have

$$\tilde{\theta} = (\Delta + (z + \lambda^2 \tilde{\theta}))_{00}^{-1} + O(\lambda^{\frac{1}{2} - \delta} (\lambda^2 \eta^{-1})^3).$$

Theorem 3.3 now follows from Lemma 3.6.

3.3.4 On the assumptions on $\tilde{\theta}$.

What we have actually proved by the above argument is the following lemma.

Lemma 3.10. *Suppose that $\tilde{\theta}(z) := \mathbb{E}R(z)_{00}$ satisfies $|\tilde{\theta}| \leq C$ and $\text{Im } \tilde{\theta} \geq c$. Then for $\theta = \theta(z)$ solving (3.14),*

$$|\tilde{\theta} - \theta| \lesssim \lambda^{\frac{1}{2}-\delta} (\lambda^2 \eta^{-1})^3,$$

For $z = E + i\eta$ with $\eta \gg \lambda^2$, the estimate $|\tilde{\theta} - \theta| \leq \varepsilon$ holds as a consequence of Corollary (2.2). This estimate starts the “bootstrap” – down to $\eta > \lambda^{2+\frac{1}{6}-\delta}$, the bounds in the output of Lemma 3.10 are stronger than the input. At this point a completely correct proof of Theorem 3.3 follows from a continuity argument, using the fact that $\theta(z)$ is a continuous function of z .

3.4 The T equation

Theorem 3.3 specifies the entries of the resolvent R_{xy} up to λ^c absolute error. However, according to the heuristic (3.3) we should have $|R_{xy}|^2 \sim \lambda^2 |x - y|^{2-d}$ for entries $|x - y| \gg \lambda^{-2}$. As a consequence, most of the ℓ^2 mass of the rows of the resolvent is concentrated far from the origin in entries for which the estimate (3.13) has large relative error. In particular, Theorem 3.3 is not immediately helpful to understand sums of the form

$$O[f] := \sum_x f(x) |R_{0x}|^2$$

when f is spread out on the diffusive scale, which is what is required to prove Theorem 3.1.

To address this we need to write down a self-consistent equation for the second moment of R . Notice first that F can be written in the form

$$O[f] = (RFR^*)_{00},$$

where F is the diagonal matrix with entries $f(x)$, that is, $F = \sum_x f(x) |x\rangle \langle x|$. We define

$$\tilde{M} := (\Delta - (z + \lambda^2 \tilde{\theta}))^{-1},$$

and then use the resolvent identity

$$R = \tilde{M} - \tilde{M}(\lambda V + \lambda^2 \tilde{\theta})R$$

to write

$$RFR^* = (\tilde{M} - \tilde{M}(\lambda V + \lambda^2 \tilde{\theta})R)FR^*.$$

Now we take an expectation and perform Gaussian integration by parts:

$$\begin{aligned}\mathbb{E}RFR^* &= \tilde{M}F(\mathbb{E}R^*) + \lambda^2 \tilde{M}\mathbb{E}\mathcal{D}[R - \tilde{\theta}]RFR^* + \lambda^2 \tilde{M}\mathbb{E}\mathcal{D}[RFR^*]R^* \\ &= \tilde{M}F\tilde{M}^* + \lambda^2 \tilde{M}F\tilde{M}^* + \mathfrak{E}_T,\end{aligned}\tag{3.20}$$

where $\mathfrak{E}_T$ is the error:

$$\begin{aligned}\mathfrak{E}_T &= \tilde{M}F(\mathbb{E}R^* - \tilde{M}) + \lambda^2 \tilde{M}\mathbb{E}\mathcal{D}[R - \mathbb{E}R]RFR^* \\ &\quad + \lambda^2 \tilde{M}\mathbb{E}\mathcal{D}[RFR^*](R^* - \tilde{M}^*).\end{aligned}\tag{3.21}$$

The equation (3.20) is admittedly a bit opaque. To interpret it properly, it is helpful to consider only the diagonal entries. That is, we define

$$g(x) := \mathbb{E}(RFR^*)_{xx}.$$

In the case that $F = \sum_x f(x) |x\rangle \langle x|$, the diagonal entries of the first term are given by a convolution,

$$(\tilde{M}F\tilde{M}^*)_{xx} = \sum_y f(y) |\tilde{M}_{xy}|^2 := (\tilde{K} * f)(x),$$

with kernel

$$\tilde{K}(x) := |\tilde{M}_{0x}|^2.$$

Defining the function $\mathfrak{e}_T(x) := (\mathfrak{E}_T)_{xx}$, we therefore obtain the following equation for g in terms of f :

$$g = \tilde{K} * f + \lambda^2 \tilde{K} * g + \mathfrak{e}_T.$$

Define $\mathcal{K}$ to be the operator that is convolution by $\tilde{K}$, so $\mathcal{K}f = \tilde{K} * f$. Then

$$g = (\text{Id} - \lambda^2 \mathcal{K})^{-1} \mathcal{K}f + (\text{Id} - \lambda^2 \mathcal{K})^{-1} \mathfrak{e}_T.$$

To understand what is going on with the main term, note that we can (at least formally) expand $(\text{Id} - \lambda^2 \mathcal{K})^{-1}$ as a Neumann series:

$$(\text{Id} - \lambda^2 \mathcal{K})^{-1} = \text{Id} + \sum_{j=1}^{\infty} (\lambda^2 \mathcal{K})^j.$$

In Section 3.5 we will analyze this series expansion and compare it to a Green's function from a random walk. In the remainder of this section we prove the following result.

Proposition 3.11. *For a bounded diagonal operator $F = \sum_x f(x) |x\rangle \langle x|$ and $\lambda^{2.1-\delta} < \eta < \lambda^{2+\delta}$, we have*

$$|(RFR^*)_{xx} - (\text{Id} - \lambda^2 \mathcal{K})^{-1} f| \prec \lambda^{\frac{1}{2}} (\lambda^2 \eta^{-1})^5 \eta^{-1} \|f\|_{\ell^\infty}.\tag{3.22}$$

Note that, taking $F = \text{Id}$, we have by the Ward identity

$$(RR^*)_{00} = \sum_x |R_{0x}|^2 = \eta^{-1} \text{Im } R_{00},$$

so the error term on the right hand side of (3.22) is small compared to the scale of $(RFR^*)_{xx}$ for a macroscopic-sized observable (meaning, for f having support on a diffusive-sized ball $\lambda^{-1}\eta^{-1/2}$).

The second point we make is that to prove (3.22) it suffices to prove, for every $\delta > 0$, the estimate

$$|\mathbb{E}(RFR^*)_{xx} - (\text{Id} - \lambda^2 \mathcal{K})^{-1} f| \lesssim \lambda^{\frac{1}{2}-\delta} (\lambda^2 \eta^{-1})^5 \eta^{-1} \|f\|_{\ell^\infty}.$$

The concentration inequality then follows from the Gaussian poincare inequality and applications of the a priori $\|R\|_{p \rightarrow q}$ estimates.

The error in the expectation has the form

$$(\text{Id} - \lambda^2 \mathcal{K})^{-1} \mathfrak{e}_T.$$

Proposition 3.11 then follows from the following two estimates.

Lemma 3.12. *The operator $(\text{Id} - \lambda^2 \mathcal{K})^{-1}$ is bounded from $\ell^\infty \rightarrow \ell^\infty$, and in particular*

$$\|(\text{Id} - \lambda^2 \mathcal{K})^{-1}\|_{\infty \rightarrow \infty} \leq C \lambda^2 \eta^{-1}.$$

Lemma 3.13. *The error $\mathfrak{E}_T$ defined in (3.21) satisfies the estimate*

$$\|\mathfrak{E}_T\|_{1 \rightarrow \infty} = \max_{x, y \in \mathbb{Z}_L^d} |(\mathfrak{E}_T)_{xy}| \lesssim [\lambda^{\frac{1}{2}-\delta} (\lambda^2 \eta^{-1})^5] \lambda^{-2} \|F\|_{2 \rightarrow 2}. \quad (3.23)$$

First we prove Lemma 3.12.

Proof of Lemma 3.12. We estimate $\sum_x \tilde{K}(x)$ using the Ward identity and the self-consistent equation (3.14) satisfied by θ :

$$\begin{aligned} \lambda^2 \sum_x \tilde{K}(x) &= \lambda^2 \sum_x |(\Delta - (z + \lambda^2 \tilde{\theta}))_{0x}^{-1}|^2 \\ &= \frac{\lambda^2 \text{Im}(\Delta - (z + \lambda^2 \tilde{\theta}))_{00}^{-1}}{\lambda^2 \text{Im} \tilde{\theta} + \eta} \\ &= \frac{\lambda^2 \text{Im}(\Delta - (z + \lambda^2 \theta))_{00}^{-1} + O(\lambda^{\frac{5}{2}-\delta} (\lambda^2 \eta^{-1})^3)}{\lambda^2 \text{Im} \tilde{\theta} + \eta} \\ &= \frac{\lambda^2 \text{Im} \theta + O(\lambda^{\frac{5}{2}-\delta} (\lambda^2 \eta^{-1})^3)}{\lambda^2 \text{Im} \theta + \eta + O(\lambda^{\frac{5}{2}-\delta} (\lambda^2 \eta^{-1})^3)} \\ &= 1 - (\text{Im} \theta)^{-1} \lambda^{-2} \eta + O(\lambda^{\frac{1}{2}-\delta} (\lambda^2 \eta)^5 \lambda^{-2} \eta) \end{aligned}$$

For $\eta \gg \lambda^{2.1-\delta}$, the last term is small compared to $\lambda^{-2}\eta$. Therefore,

$$\|\lambda^2 \mathcal{K}\|_{\infty \rightarrow \infty} \leq 1 - c\lambda^{-2}\eta.$$

The result now follows for example by summing the Neumann series for $(\text{Id} - \lambda^2 \mathcal{K})^{-1}$. $\square$

The proof of Lemma 3.13 requires only slightly more calculation. First we need the following lemma.

Lemma 3.14. *For $d(E, \Sigma_d) > \varepsilon$, $\eta > \lambda^{2+\frac{1}{6}-\delta}$*

$$\|\tilde{M} - \mathbb{E}R\|_{1 \rightarrow 2} \lesssim \lambda^{-\frac{1}{2}-\delta}(\lambda^2 \eta^{-1})^3. \quad (3.24)$$

Proof. Recall that using the resolvent identity for R and taking expectations we have

$$\mathbb{E}R = \tilde{M} + \lambda^2 \tilde{M} \mathbb{E} \mathcal{D}[R - \mathbb{E}R]R.$$

The result follows upon using Lemma 3.9 and the estimate $\|\tilde{M}\|_{1 \rightarrow 2} \lesssim \lambda^{-1}$,

$$\begin{aligned} \|\mathbb{E}R - \tilde{M}\|_{1 \rightarrow 2} &\leq \lambda^2 \|\tilde{M}\|_{1 \rightarrow 2} \mathbb{E} \|\mathcal{D}[R - \mathbb{E}R]\|_{2 \rightarrow 2} \|R\|_{2 \rightarrow 2} \\ &\lesssim \lambda^{-\frac{1}{2}-\delta}(\lambda^2 \eta^{-1})^3. \end{aligned}$$

$\square$

We can now bound $\mathfrak{E}_T$.

Proof of Lemma 3.13. We write

$$\mathfrak{E}_T =: \text{I} + \text{II} + \text{III} + \text{IV},$$

where

$$\begin{aligned} \text{I} &= \tilde{M}F(\mathbb{E}R^* - \tilde{M}) \\ \text{II} &= \lambda^2 \tilde{M} \mathbb{E} \mathcal{D}[R - \mathbb{E}R]RFR^* \\ \text{III} &= \lambda^2 \tilde{M}(\mathbb{E} \mathcal{D}[RFR^*])\mathbb{E}(R^* - \tilde{M}^*) \\ \text{IV} &= \lambda^2 \tilde{M}(\mathcal{D}[RFR^*] - \mathbb{E} \mathcal{D}[RFR^*])(R^* - \tilde{M}^*). \end{aligned}$$

The first term is bounded using (3.24),

$$\begin{aligned} \|\text{I}\|_{1 \rightarrow \infty} &\leq \|\tilde{M}\|_{1 \rightarrow 2} \|F\|_{2 \rightarrow 2} \|\mathbb{E}R^* - \tilde{M}\|_{1 \rightarrow 2} \\ &\lesssim \lambda^{-\frac{3}{2}-\delta} \|F\|_{2 \rightarrow 2} (\lambda^2 \eta^{-1})^3. \end{aligned}$$

For the second term use Theorem 3.3,

$$\begin{aligned}
\|II\|_{1 \rightarrow \infty} &\leq \lambda^2 \|\tilde{M}\|_{1 \rightarrow 2} \mathbb{E} \|\mathcal{D}[R - \mathbb{E}R]\|_{2 \rightarrow 2} \|R\|_{2 \rightarrow 2} \|F\|_{2 \rightarrow 2} \|R^*\|_{1 \rightarrow 2} \\
&\lesssim \lambda(\lambda^{\frac{1}{2}-\delta})(\lambda^2 \eta^{-1})^2 \eta^{-1} \|F\|_{2 \rightarrow 2} (\lambda \eta^{-1}) \\
&\lesssim \lambda^{-\frac{3}{2}-\delta} (\lambda^2 \eta^{-1})^4 \|F\|_{2 \rightarrow 2}.
\end{aligned}$$

For the third term we use $\|\mathcal{D}[RFR^*]\|_{2 \rightarrow 2} \leq \|R\|_{1 \rightarrow 2}^2 \|F\|_{2 \rightarrow 2}$ as well as (3.24):

$$\begin{aligned}
\|III\|_{1 \rightarrow \infty} &\leq \lambda^2 \|\tilde{M}\|_{1 \rightarrow 2} \mathbb{E} \|\mathcal{D}[RFR^*]\|_{2 \rightarrow 2} \|\mathbb{E}R - \tilde{M}\|_{1 \rightarrow 2} \\
&\leq \lambda^2 \|\tilde{M}\|_{1 \rightarrow 2} \mathbb{E} \|R\|_{1 \rightarrow 2}^2 \|F\|_{2 \rightarrow 2} \|\mathbb{E}R - \tilde{M}\|_{1 \rightarrow 2} \\
&\lesssim \lambda(\lambda \eta^{-1})^2 \|F\|_{2 \rightarrow 2} \lambda^{-\frac{1}{2}-\delta} (\lambda^2 \eta^{-1})^3 \\
&= \lambda^{-\frac{3}{2}-\delta} (\lambda^2 \eta^{-1})^5 \|F\|_{2 \rightarrow 2}.
\end{aligned}$$

For the fourth term we need an estimate for $\|\mathcal{D}[RFR^*] - \mathbb{E}\mathcal{D}[RFR^*]\|$, which is the maximum size of the diagonal fluctuations,

$$\|\mathcal{D}[RFR^*] - \mathbb{E}\mathcal{D}[RFR^*]\| = \sup_x |(RFR^*)_{xx} - \mathbb{E}(RFR^*)_{xx}|.$$

A similar calculation as in (3.19) shows that³

$$\begin{aligned}
|\nabla(RFR^*)_{xx}|^2 &\leq \lambda^2 \|F\|_{2 \rightarrow 2}^2 \|R\|_{1 \rightarrow 4}^2 \|R\|_{1 \rightarrow 2}^2 \|R\|_{2 \rightarrow 4}^2 \\
&\prec \lambda^{-3} (\lambda^2 \eta^{-1})^6
\end{aligned}$$

Therefore,

$$\|\mathcal{D}[RFR^*] - \mathbb{E}\mathcal{D}[RFR^*]\|_{2 \rightarrow 2} \prec \lambda^{-3/2} \|F\|_{2 \rightarrow 2} (\lambda^2 \eta^{-1})^3.$$

Combined with the estimate $\|R\|_{1 \rightarrow 2} \prec \lambda \eta^{-1}$ we have

$$\begin{aligned}
IV &\lesssim \lambda^{-\delta} \lambda^2 \|\tilde{M}\|_{1 \rightarrow 2} (\lambda^{-3/2} (\lambda^2 \eta^{-1})^3) (\lambda \eta^{-1}) \|F\|_{2 \rightarrow 2} \\
&\lesssim \lambda^{-\frac{3}{2}-\delta} (\lambda^2 \eta^{-1})^4 \|F\|_{2 \rightarrow 2}
\end{aligned}$$

which is the same scale as the bound for II.

Rearranging the terms yields the estimate (3.23). □

3.5 Finishing the proof of Theorem 3.1

In this section we complete the proof of Theorem 3.1, which states that the following bound holds with high probability:

$$\sum_{|x| > c\lambda \eta^{-1/2}} |R_{0x}(E + i\eta)|^2 \geq c\eta^{-1}.$$

³Using the bounds for $\|R^{1/2}\|_{1 \rightarrow 2}$, $\|R^{1/2}\|_{2 \rightarrow 2}$, and $\|R^{1/2}\|_{2 \rightarrow 6}$ and interpolating as before. This calculation has been omitted but is largely similar to the calculation for ∇R_{xx}

By the Ward identity, we have that

$$\sum_x |R_{0x}|^2 \gtrsim \eta^{-1},$$

so it suffices to show that for any c_0 we can find c_1 such that

$$\sum_{|x| < c_1 \lambda \eta^{-1/2}} |R_{0x}(E + i\eta)|^2 \leq c_0 \eta^{-1}$$

holds with high probability. By Proposition 3.11 it then suffices to show that there exists c_1 such that

$$((\text{Id} - \lambda^2 \mathcal{K})^{-1} \mathcal{K} \mathbf{1}_{\{|x| < c_1 \lambda \eta^{-1/2}\}})(0) < \frac{1}{2} c_0 \eta^{-1}. \quad (3.25)$$

We need more information about $\mathcal{K}$ to estimate $(\text{Id} - \lambda^2 \mathcal{K})^{-1}$ well enough to compute the main term. In particular, we need the moments of $\tilde{K}$.

Lemma 3.15. *The convolution kernel $\tilde{K}$ satisfies the following estimates:*

$$\lambda^2 \sum \tilde{K}(x) \leq 1 - c \lambda^{-2} \eta \quad (3.26)$$

$$\lambda^2 \sum x_j \tilde{K}(x) = 0 \quad (3.27)$$

$$\lambda^2 \sum |x|^2 \tilde{K}(x) \gtrsim \lambda^{-4} \quad (3.28)$$

$$\lambda^2 \sum |x|^4 \tilde{K}(x) \lesssim \lambda^{-8}. \quad (3.29)$$

Proof. The first bound (3.26) was proven in the calculation above Lemma 3.12. The identity (3.27) follows from the fact that $\tilde{K}$ is symmetric about the origin.

Now we prove (3.28).

$$(\Delta - z)_{0x}^{-1} = \int_{\mathbb{T}^d} e^{i\xi \cdot x} \frac{1}{\omega(\xi) - z} d\xi.$$

For simplicity we set $w = z + \lambda^2 \tilde{\theta}$ below. Then by Plancherel's formula and then the coarea formula we have

$$\begin{aligned} \lambda^2 \sum |x|^2 \tilde{K}(x) &= \lambda^2 \sum |x(\Delta - w)_{0x}^{-1}|^2 \\ &= \lambda^2 \int_{\mathbb{T}^d} \left| \nabla(\omega(\xi) - w)^{-1} \right|^2 d\xi \\ &= \lambda^2 \int_{\mathbb{T}^d} \frac{|\nabla \omega|^2}{|\omega(\xi) - w|^4} d\xi \\ &= \lambda^2 \int \frac{1}{|E' - w|^4} \left(\int_{\{\omega(\xi) = E'\}} |\nabla \omega| d\mathcal{H}^{d-1} \right) dE' \end{aligned}$$

The function in brackets is positive, and $\int |E' - w|^{-4} dE \simeq \lambda^{-6}$, so this completes the proof of (3.28).

The proof of (3.29) proceeds along similar lines (one uses Plancherel's theorem to produce a second derivative, otherwise the use of the coarea formula is also similar). $\square$

The main theorem now follows from the following elementary anticoncentration lemma for random walks.

Lemma 3.16. *Let $X \in \mathbb{R}^d$ be a random variable satisfying $\mathbb{E}X = 0$, $\mathbb{E}X_i X_j = \sigma^2 \delta_{ij}$, and $\mathbb{E}|X|^3 \leq C\sigma^3$. Let $Y_N := \sum_{j=1}^N X^{(j)}$ be a sum of N independent copies of X . Then Y_N satisfies, for any $y \in \mathbb{R}^d$,*

$$\mathbb{P}(|Y_N - y| \leq \sigma) \leq CN^{-d/2}.$$

In particular, for $r \geq \sigma$ it follows that

$$\mathbb{P}(|Y_N - y| \leq r) \leq CN^{-d/2}(r/\sigma)^d$$

Proof. We give a proof via characteristic functions. For any random variable Z define its characteristic function $f_Z : \mathbb{R}^d \rightarrow \mathbb{C}$ by $f_Z(\xi) := \mathbb{E}e^{i\langle Z, \xi \rangle}$.

By scaling it suffices to prove the result for $\sigma = 1$. First we estimate f_X near 0. By the moment assumptions on X and a Taylor expansion we can estimate

$$\begin{aligned} |f_X(\xi)| &\leq |\mathbb{E}(1 + i\langle X, \xi \rangle - \frac{1}{2}\langle X, \xi \rangle^2 + C|\langle X, \xi \rangle|^3)| \\ &\leq 1 - \frac{1}{2}|\xi|^2 + C|\xi|^3\mathbb{E}|X|^3 \\ &\leq 1 - \frac{1}{2}|\xi|^2 + C|\xi|^3 \end{aligned}$$

Hence for $c_0 > 0$ small, we have

$$|f_X(\xi)| \leq 1 - \frac{1}{3}|\xi|^2,$$

for $|\xi| \leq c_0$. Now let $\psi : \mathbb{R}^d \rightarrow \mathbb{R}$ satisfy $1_{|x| \leq 1} \leq \psi$ and $|\hat{\phi}| \leq 1_{|\xi| \leq c_0}$, so that by

Plancherel, and the estimate above we get

$$\begin{aligned}
\mathbb{P}(|Y_N - y| \leq 1) &\leq \mathbb{E}\psi(X) \\
&\leq \int_{\mathbb{R}^d} e^{iy \cdot \xi} \hat{\psi}(\xi) f_{Y_N}(\xi) d\xi \\
&\leq C \int_{|\xi| \leq c_0} |f_{Y_N}(\xi)| d\xi \\
&\leq C \int_{|\xi| \leq c_0} |f_X(\xi)|^N \\
&\leq C \int_{|\xi| \leq c_0} (1 - \frac{|\xi|^2}{3})^N d\xi \\
&\leq CN^{-d/2}
\end{aligned}$$

□

We are now ready to establish (3.25), and thereby prove Theorem 1.1. In fact we will prove a somewhat stronger statement that provides bounds for ℓ^2 -type averages of the resolvent on scales much smaller than the diffusive scaling.

Proposition 3.17. *Let $\eta > \lambda^{2.1-\delta}$ and*

$$r > C(\lambda^{1/2}(\lambda^2\eta^{-1})^{5/2})\lambda^{-1}\eta^{-1/2}.$$

Then letting χ_r be the indicator function for the ball of radius r ,

$$(\text{Id} - \lambda^2\mathcal{K})^{-1}(\mathcal{K}\chi_r)(x) \lesssim \lambda^2 r^2.$$

Proof. By Proposition 3.11 it suffices to prove that for any x

$$(\text{Id} - \lambda^2\mathcal{K})^{-1}\mathcal{K}\chi_r(x) \leq \lambda^2 r^2, \tag{3.30}$$

where $\mathcal{K}$ is convolution against the kernel $\tilde{K}_z$

Writing out the Neuman series for $(\text{Id} - \lambda^2\mathcal{K})^{-1}$ we have

$$\begin{aligned}
(\text{Id} - \lambda^2\mathcal{K})^{-1}(\mathcal{K}\chi_r(x_0))(x) &= \lambda^{-2} \sum_{j=1}^{\infty} (\lambda^2 \tilde{K})^{*j} * \chi_r(x) \\
&= \lambda^{-2} \sum_{j=1}^{\infty} (1 - \alpha)^j ((1 - \alpha)^{-1} \tilde{K})^{*j} * \chi_r(x),
\end{aligned}$$

where we set $\alpha = 1 - \lambda^2 \sum_x \tilde{K}(x)$ so that $(1 - \alpha)^{-1} \tilde{K}$ is a probability distribution (note that by Lemma 3.15, $\alpha \simeq \lambda^{-2}\eta$). Letting $Y_k = \sum_{j=1}^k X_j$ be the partial sums of

a random walk with independent steps X_j of distribution $(1 - \alpha)^{-1}\lambda^2\tilde{K}$, we can write the above as

$$\begin{aligned} (\text{Id} - \lambda^2\mathcal{K})^{-1}(\mathcal{K}\chi_r(x_0))(x) &= \lambda^{-2} \sum_{j=1}^{\infty} (1 - \alpha)^j \mathbb{P}(|Y_j - x_0| \leq r) \\ &\lesssim \lambda^{-2} \sum_{j=1}^{\infty} (1 - \alpha)^j \min\{1, r^d (\lambda^{-2}j)^{-d/2}\}, \end{aligned}$$

The estimate (3.30) follows from combining Lemma 3.15 and Lemma 3.16 (with $\sigma = \lambda^{-2}$). For $j \leq \lambda^4 r^2$ we simply bound the terms by 1 and get a contribution of $\lambda^2 r^2$. The terms with $j \gg \lambda^4 r^2$ are bounded similarly because of the summable decay⁴ of $j^{-d/2}$, and we obtain

$$(\text{Id} - \lambda^2\mathcal{K})^{-1}(\mathcal{K}\chi_r(x_0))(x) \lesssim \lambda^2 r^2$$

as desired. □

References

- [Abr+79] Elihu Abrahams, Philip W Anderson, Donald C Licciardello, and Tirupattur V Ramakrishnan. “Scaling theory of localization: Absence of quantum diffusion in two dimensions”. In: *Physical Review Letters* 42.10 (1979), p. 673.
- [And58] Philip W Anderson. “Absence of diffusion in certain random lattices”. In: *Physical review* 109.5 (1958), p. 1492.
- [AW15] Michael Aizenman and Simone Warzel. *Random operators*. Vol. 168. American Mathematical Soc., 2015.
- [BBH23] Afonso S Bandeira, March T Boedihardjo, and Ramon van Handel. “Matrix concentration inequalities and free probability”. In: *Inventiones mathematicae* 234.1 (2023), pp. 419–487.
- [BDH25a] Adam Black, Reuben Drogin, and Felipe Hernández. “Self-consistent equations and quantum diffusion for the Anderson model”. In: *arXiv preprint arXiv:2506.06468* (2025).
- [BDH25b] Adam Black, Reuben Drogin, and Felipe Hernández. “Tail bounds for the Dyson series of random Schrödinger equations”. In: *arXiv preprint arXiv:2502.02566* (2025).

⁴Except of course in $d = 2$, where we lose a logarithm, but anyway this is hidden by the stochastic domination notation.

- [BDH25c] Adam Black, Reuben Drogin, and Felipe Hernández. “Tail bounds for the Dyson series of random Schrödinger equations”. In: *arXiv:2502.02566* (2025).
- [Bou01] Jean Bourgain. “On random Schrödinger operators on $\mathbb{Z}^2$ ”. In: *Discrete and Continuous Dynamical Systems* 8.1 (2001), pp. 1–15.
- [Che06] Thomas Chen. “Convergence in higher mean of a random Schrödinger to a linear Boltzmann evolution”. In: *Communications in mathematical physics* 267.2 (2006), pp. 355–392.
- [CL90] René Carmona and Jean Lacroix. *Spectral Theory of Random Schrödinger Operators*. Probability and Its Applications. Boston: Birkhäuser, 1990.
- [DS21] Mitia Duerinckx and Christopher Shirley. “A new spectral analysis of stationary random Schrödinger operators”. In: *Journal of Mathematical Physics* 62.7 (2021).
- [Dub+25a] Sofia Dubova, Fan Yang, Horng-Tzer Yau, and Jun Yin. “Delocalization of Non-Mean-Field Random Matrices in Dimensions $d \geq 3$ ”. In: *arXiv preprint arXiv:2507.20274* (2025).
- [Dub+25b] Sofia Dubova, Kevin Yang, Horng-Tzer Yau, and Jun Yin. “Delocalization of two-dimensional random band matrices”. In: *arXiv preprint arXiv:2503.07606* (2025).
- [Erd+13a] László Erdős, Antti Knowles, Horng-Tzer Yau, and Jun Yin. “The local semicircle law for a general class of random matrices”. In: *Electron. J. Probab* 18.59 (2013), pp. 1–58.
- [Erd+13b] László Erdős, Antti Knowles, Horng-Tzer Yau, and Jun Yin. “Delocalization and diffusion profile for random band matrices”. In: *Communications in Mathematical Physics* 323 (2013), pp. 367–416.
- [ESY07a] László Erdős, Manfred Salmhofer, and Horng-Tzer Yau. “Quantum diffusion for the Anderson model in the scaling limit”. In: *Annales Henri Poincaré*. Vol. 8. 4. Springer. 2007, pp. 621–685.
- [ESY07b] László Erdős, Manfred Salmhofer, and Horng-Tzer Yau. “Quantum diffusion of the random Schrödinger evolution in the scaling limit II. The recollision diagrams”. In: *Communications in mathematical physics* 271.1 (2007), pp. 1–53.
- [ESY08] László Erdős, Manfred Salmhofer, and Horng-Tzer Yau. “Quantum diffusion of the random Schrödinger evolution in the scaling limit”. In: *Acta Math* 200 (2008), pp. 211–277.
- [ESY09] László Erdős, Benjamin Schlein, and Horng-Tzer Yau. “Semicircle law on short scales and delocalization of eigenvectors for Wigner random matrices”. In: *Annals of Probability* 37.3 (2009).

- [EY00] László Erdős and Horng-Tzer Yau. “Linear Boltzmann equation as the weak coupling limit of a random Schrödinger equation”. In: *Communications on Pure and Applied Mathematics* 53.6 (2000), pp. 667–735.
- [FS83] Jürg Fröhlich and Thomas Spencer. “Absence of diffusion in the Anderson tight binding model for large disorder or low energy”. In: *Communications in Mathematical Physics* 88.2 (1983), pp. 151–184.
- [GMP77] I Ya Gol’dshstein, Stanislav A Molchanov, and Leonid A Pastur. “A pure point spectrum of the stochastic one-dimensional Schrödinger operator”. In: *Functional Analysis and Its Applications* 11.1 (1977), pp. 1–8.
- [Her24] Felipe Hernández. “Quantum diffusion via an approximate semigroup property”. In: *Probability and Mathematical Physics* 5.4 (2024), pp. 1039–1184.
- [KS80] Hervé Kunz and Bernard Souillard. “Sur le spectre des opérateurs aux différences finies aléatoires”. In: *Communications in Mathematical Physics* 78.2 (1980), pp. 201–246.
- [LP91] Françoise Lust-Piquard and Gilles Pisier. “Non commutative Khintchine and Paley inequalities”. In: *Arkiv för matematik* 29 (1991), pp. 241–260.
- [Lus86] Françoise Lust-Piquard. “Inégalités de Khintchine dans C_p ($1 < p < \infty$)”. In: *CR Acad. Sci. Paris Sér. I Math* 303.7 (1986), pp. 289–292.
- [MPR97] J Magnen, G Poirrot, and V Rivasseau. “The Anderson model as a matrix model”. In: *Nuclear Physics B-Proceedings Supplements* 58 (1997), pp. 149–162.
- [MPR98] Jacques Magnen, Gilles Poirrot, and Vincent Rivasseau. “Ward-type identities for the two-dimensional Anderson model at weak disorder”. In: *Journal of statistical physics* 93 (1998), pp. 331–358.
- [MPR99] Jacques Magnen, Gilles Poirrot, and Vincent Rivasseau. “Renormalization group methods and applications: First results for the weakly coupled Anderson model”. In: *Physica A: Statistical Mechanics and its Applications* 263.1-4 (1999), pp. 131–140.
- [Pas80] LA Pastur. “Spectral properties of disordered systems in the one-body approximation”. In: *Communications in mathematical physics* 75.2 (1980), pp. 179–196.
- [Sim00] Barry Simon. “Schrödinger operators in the twenty-first century”. In: *Mathematical physics 2000*. World Scientific, 2000, pp. 283–288.
- [Sim84] Barry Simon. “Fifteen problems in mathematical physics”. In: *Perspectives in mathematics, Birkhäuser, Basel* 423 (1984).

- [Spo77] Herbert Spohn. “Derivation of the transport equation for electrons moving through random impurities”. In: *Journal of Statistical Physics* 17.6 (1977), pp. 385–412.
- [SSW02] Wilhelm Schlag, Carol Shubin, and Thomas Wolff. “Frequency concentration and location lengths for the Anderson model at small disorders”. In: *Journal d’Analyse Mathématique* 88.1 (2002), pp. 173–220.
- [Sto01] Peter Stollmann. *Caught by disorder: bound states in random media*. Vol. 20. Springer Science & Business Media, 2001.
- [Tom75] Peter A Tomas. “A restriction theorem for the Fourier transform”. In: *Bulletin of the American Mathematical Society* 81.2 (1975), pp. 477–478.
- [Van17] Ramon Van Handel. “Structured random matrices”. In: *Convexity and concentration* (2017), pp. 107–156.
- [Van54] Leon Van Hove. “Quantum-mechanical perturbations giving rise to a statistical transport equation”. In: *Physica* 21.1-5 (1954), pp. 517–540.
- [YY21] Fan Yang and Jun Yin. “Random band matrices in the delocalized phase, III: Averaging fluctuations”. In: *Probability Theory and Related Fields* 179 (2021), pp. 451–540.
- [YY25] Fan Yang and Jun Yin. “Delocalization of a general class of random block Schrödinger operators”. In: *arXiv preprint arXiv:2501.08608* (2025).
- [YYY21] Fan Yang, Horng-Tzer Yau, and Jun Yin. “Delocalization and quantum diffusion of random band matrices in high dimensions I: Self-energy renormalization”. In: *arXiv preprint arXiv:2104.12048* (2021).
- [YYY22] Fan Yang, Horng-Tzer Yau, and Jun Yin. “Delocalization and quantum diffusion of random band matrices in high dimensions II: T-expansion”. In: *Communications in Mathematical Physics* 396.2 (2022), pp. 527–622.