

QUANTUM DIFFUSION AND EIGENFUNCTION DELOCALIZATION IN A RANDOM BAND MATRIX MODEL

LÁSZLÓ ERDŐS^{1*}, ANTTI KNOWLES^{2†}

*Institute of Mathematics, University of Munich,
Theresienstr. 39, D-80333 Munich, Germany
lerdos@math.lmu.de*¹

*Department of Mathematics, Harvard University
Cambridge MA 02138, USA
knowles@math.harvard.edu*²

20 September, 2010

We consider Hermitian and symmetric random band matrices H in $d \geq 1$ dimensions. The matrix elements H_{xy} , indexed by $x, y \in \Lambda \subset \mathbb{Z}^d$, are independent, uniformly distributed random variables if $|x - y|$ is less than the band width W , and zero otherwise. We prove that the time evolution of a quantum particle subject to the Hamiltonian H is diffusive on time scales $t \ll W^{d/3}$. We also show that the localization length of the eigenvectors of H is larger than a factor $W^{d/6}$ times the band width. All results are uniform in the size $|\Lambda|$ of the matrix.

AMS Subject Classification: 15B52, 82B44, 82C44

Keywords: Random band matrix, Anderson model, localization length,

1. INTRODUCTION

The general formulation of the universality conjecture for disordered systems states that there are two distinctive regimes depending on the energy and the disorder strength. In the strong disorder regime, the

*Partially supported by SFB-TR 12 Grant of the German Research Council.

†Partially supported by U.S. National Science Foundation Grant DMS 08-04279.

eigenfunctions are localized and the local spectral statistics are Poisson. In the weak disorder regime, the eigenfunctions are delocalized and the local statistics coincide with those of a Gaussian matrix ensemble.

Random band matrices are natural intermediate models to study eigenvalue statistics and quantum propagation in disordered systems as they interpolate between Wigner matrices and random Schrödinger operators. Wigner matrix ensembles represent mean-field models without spatial structure, where the quantum transition rates between any two sites are i.i.d. random variables with zero expectation. In the celebrated Anderson model [5], only a random on-site potential V is present in addition to a short range deterministic hopping (Laplacian) on a graph that is typically a regular box in \mathbb{Z}^d .

For the Anderson model, a fundamental open question is to establish the metal-insulator transition, i.e. to show that in $d \geq 3$ dimensions the eigenfunctions of $-\Delta + \lambda V$ are delocalized for small disorder λ . The localization regime at large disorder or near the spectral edges has been well understood by Fröhlich and Spencer with the multiscale technique [29, 30], and later by Aizenman and Molchanov by the fractional moment method [3]; many other works have since contributed to this field. In particular, it has been established that the local eigenvalue statistics are Poisson [38] and that the eigenfunctions are exponentially localized with an upper bound on the localization length that diverges as the energy parameter approaches the presumed phase transition point [15, 43].

The progress in the delocalization regime has been much slower. For the Bethe lattice, corresponding to the infinite-dimensional case, delocalization has been established in [4, 27, 35]. In finite dimensions only partial results are available. The existence of an absolutely continuous spectrum (i.e. extended states) has been shown for a rapidly decaying potential, corresponding to a scattering regime [8, 10, 39]. Diffusion has been established for a heavy quantum particle immersed in a phonon field in $d \geq 4$ dimensions [28]. For the original Anderson Hamiltonian with a small coupling constant λ , the eigenfunctions have a localization length of at least λ^{-2} (see [9]). The time and space scale λ^{-2} corresponds to the kinetic regime where the quantum evolution can be modelled by a linear Boltzmann equation [24, 45]. Beyond this time scale the dynamics is diffusive. This has been established in the scaling limit $\lambda \rightarrow 0$ up to time scales $t \sim \lambda^{-2-\kappa}$ with an explicit $\kappa > 0$ in [18–20]. There are no rigorous results on the local spectral statistics of the Anderson model, but it is conjectured – and supported by numerous arguments in the physics literature, especially by supersymmetric methods (see [14]) – that the local correlation function of the eigenvalues of the finite volume Anderson model follows the GOE statistics in the thermodynamic limit.

Due to their mean-field character, Wigner matrices are simpler to study than the Anderson model and they are always in the delocalization regime. The complete delocalization of the eigenvectors was proved in [21]. The local spectral statistics in the bulk are universal, i.e. they follow the statistics of the corresponding Gaussian ensemble (GOE, GUE, GSE), depending on the symmetry type of the matrix (see [37] for explicit formulas). For an arbitrary single entry distribution, bulk universality has been proved recently in [17, 22, 23] for all symmetry classes. A different proof was given in [46] for the Hermitian case.

Random band matrices $H = \{H_{xy}\}_{x,y \in \Gamma}$ represent systems on a large finite graph Γ with a metric. The matrix elements between two sites, x and y , are independent random variables with a variance $\sigma_{xy}^2 := \mathbb{E}|H_{xy}|^2$ depending on the distance between the two sites. The variance typically decays with the distance on a characteristic length scale W , called the *band width* of H . This terminology comes from the simplest one-dimensional model where the graph is a path on N vertices, labelled by $\Gamma = \{1, 2, \dots, N\}$, and the matrix elements H_{xy} vanish if $|x - y| \geq W$. If $W = N$ and all variances are equal, we recover the usual Wigner matrix. The case $W = O(1)$ is a one-dimensional Anderson-type model with random hoppings at bounded range. Higher-dimensional models are obtained if the graph Γ is a box in \mathbb{Z}^d . For more general random band matrices and for a systematic presentation, see [44].

Since the one-dimensional Anderson-type models are always in the localization regime, varying the band width W offers a possibility to test the localization-delocalization transition between an Anderson-type

model and the Wigner ensemble. Numerical simulations and theoretical arguments based on supersymmetric methods [31] suggest that the local eigenvalue statistics change from Poisson, for $W \ll N^{1/2}$, to GOE (or GUE), for $W \gg N^{1/2}$. The eigenvectors are expected to have a localization length ℓ of order W^2 . In particular the eigenvectors are fully delocalized for $W \gg N^{1/2}$. In two dimensions the localization length is expected to be exponentially large in W ; see [1]. In accordance with the extended states conjecture for the Anderson model, the localization length is expected to be macroscopic, $\ell \sim N$, independently of the band width in $d \geq 3$ dimensions.

Extending the techniques of the rigorous proofs for Anderson localization, Schenker has recently proved the upper bound $\ell \leq W^8$ for the localization length in $d = 1$ dimensions [40]. In this paper we prove a counterpart of this result from the side of delocalization. More precisely, we show a lower bound $\ell \geq W^{1+d/6}$ for the eigenvectors of d -dimensional band matrices with uniformly distributed entries. We remark that the lower bound $\ell \geq W$ was proved recently in [25] for very general band matrices.

On the spectral side, we mention that, apart from the semicircle law (see [2, 25, 33] for $d = 1$ and [11] for $d = 3$), the question of bulk universality of *local* spectral statistics for band matrices is mathematically open even for $d = 1$. In the spirit of the general conjecture, one expects GUE/GOE statistics in the bulk for the delocalization regime, $W \gg N^{1/2}$. The GUE/GOE statistics have recently been established [25] for a class of generalized Wigner matrices, where the variances of different matrix elements are not necessarily identical, but are of comparable size, i.e. $\mathbb{E}|H_{xy}|^2 \sim \mathbb{E}|H_{x'y'}|^2$; in particular, the band width is still macroscopic ($W \sim N$).

Supersymmetric methods offer a very attractive approach to study the delocalization transition in band matrices but the rigorous control of the functional integrals away from the saddle points is difficult and it has been performed only for the density of states [11]. Effective models that emerge near the saddle points can be more accessible to rigorous mathematics. Recently Disertori, Spencer and Zirnbauer studied a related statistical mechanics model that is expected to reflect the Anderson localization and delocalization transition for real symmetric band matrices. They proved a quasi-diffusive estimate for the two-point correlation functions in a three dimensional supersymmetric hyperbolic nonlinear sigma model at low temperatures [13]. Localization was also established in the same model at high temperatures [12].

We also mention that band matrices are not the only possible interpolating models to mimic the metal-insulator transition. Other examples include the Anderson model with a spatially decaying potential [8, 34] and a quasi one-dimensional model with a weak on-site potential for which a transition in the sense of local spectral statistics has been established in [6, 47].

A natural approach to study the delocalization regime is to show that the quantum time evolution is diffusive on large scales. We normalize the matrix entries so that the rate of quantum jumps is of order one. The typical distance of a single jump is the band width W . If the jumps were independent, the typical distance travelled in time t would be $W\sqrt{t}$. Using the argument of [9], we show that a typical localization length ℓ is incompatible with a diffusion on spatial scales larger than ℓ . Thus we obtain $\ell \geq W\sqrt{t}$, provided that the diffusion approximation can be justified up to time t .

The main result of this paper is that the quantum dynamics of the d -dimensional band matrix is given by a superposition of heat kernels up to time scales $t \ll W^{d/3}$. Although diffusion is expected to hold up to time $t \sim W^2$ for $d = 1$ and up to any time for $d \geq 3$ (assuming the thermodynamic limit has been taken), our method can follow the quantum dynamics only up to $t \ll W^{d/3}$. The threshold exponent $d/3$ originates in technical estimates on certain Feynman graphs; going beyond the exponent $d/3$ would require a further resummation of certain four-legged subdiagrams (see Section 11).

Finally, we remark that our method also yields a bound on the largest eigenvalue of a band matrix; see Theorem 3.4 in the forthcoming paper [16] for details.

Acknowledgements. The problem of diffusion for random band matrices originated from several discussions with H.T. Yau and J. Yin. The authors are especially grateful to J. Yin for various insights and for pointing out an improvement in the counting of the skeleton diagrams.

2. THE SETUP

Let the dimension $d \geq 1$ be fixed and consider the d -dimensional lattice \mathbb{Z}^d equipped with the Euclidean norm $|\cdot|_{\mathbb{Z}^d}$ (any other norm would also do). We index points of \mathbb{Z}^d with x, y, z, \dots . Let $W > 1$ denote a large parameter (the *band width*) and define

$$M \equiv M(W) := |\{x \in \mathbb{Z}^d : 1 \leq |x|_{\mathbb{Z}^d} \leq W\}|,$$

the number of points at distance at most W from the origin. In the following we tacitly make use of the obvious relation $M \sim CW^d$. For notational convenience, we use both W and M in the following.

In order to avoid dealing with the infinite lattice directly, we restrict the problem to a finite periodic lattice Λ_N of linear size N . More precisely, for $N \in \mathbb{N}$, we set

$$\Lambda_N := \{-[N/2], \dots, N-1-[N/2]\}^d \subset \mathbb{Z}^d,$$

a cube with side length N centred around the origin. Here $[\cdot]$ denotes integer part. We regard Λ_N as periodic, i.e. we equip it with periodic addition and the periodic distance

$$|x| := \inf\{|x + N\nu|_{\mathbb{Z}^d} : \nu \in \mathbb{Z}^d\}.$$

Unless otherwise stated, all summations \sum_x are understood to mean $\sum_{x \in \Lambda_N}$.

We consider random matrices $H^\omega \equiv H$ whose entries H_{xy} are indexed by $x, y \in \Lambda_N$. Here ω denotes the running element in probability space. The large parameter of the model is the band width W . We shall always assume that $N \geq WM^{1/6}$. Under this condition all our results hold uniformly in N .

We assume that H is either Hermitian or symmetric. The entries H_{xy} satisfying $1 \leq |x - y| \leq W$ are i.i.d. (with the obvious restriction that $H_{yx} = \overline{H_{xy}}$). In the Hermitian case they are uniformly distributed on a circle of appropriate radius in the complex plane,

$$H_{xy} \sim \frac{1}{\sqrt{M-1}} \text{Unif}(\mathbb{S}^1), \quad 1 \leq |x - y| \leq W. \quad (2.1a)$$

In the symmetric case they are Bernoulli random variables,

$$\mathbb{P}\left(H_{xy} = \frac{1}{\sqrt{M-1}}\right) = \mathbb{P}\left(H_{xy} = \frac{-1}{\sqrt{M-1}}\right) = \frac{1}{2}, \quad 1 \leq |x - y| \leq W. \quad (2.1b)$$

If $|x - y| \notin [1, W]$ then $H_{xy} = 0$. An important consequence of our assumptions (2.1a) and (2.1b) is

$$|H_{xy}|^2 = \frac{1}{M-1} \mathbf{1}(1 \leq |x - y| \leq W). \quad (2.2)$$

We remark that the assumption that the matrix entries have the special form (2.1a) or (2.1b) is not necessary for our results to hold. We make it here because it greatly simplifies our proof. The reason for

this is that, as observed by Feldheim and Sodin [26, 42], the condition (2.2) allows one to obtain a simple algebraic expression for the nonbacktracking powers of H ; see Lemma 5.2.

In the forthcoming paper [16] we extend our results to random matrix ensembles in which the matrix elements H_{xy} are allowed to have a general distribution (and thus in particular a genuinely random absolute value); moreover their variances $\mathbb{E}|H_{xy}|^2$ are given by a general profile on the scale W in $x - y$ (as opposed to the step function profile in (2.2)). Under these assumptions, the algebraic identity of Lemma 5.2 is no longer exact, and needs to be amended with additional random terms. The resulting graphical expansion is considerably more involved than in the case (2.2), and its control requires essential new ideas. However, the fundamental mechanism underlying quantum diffusion for band matrices is already apparent in the special case (2.2) discussed in this paper.

Let $\alpha \in \mathfrak{A} := \{1, \dots, |\Lambda_N|\}$ index the orthonormal basis $\{\psi_\alpha^\omega\}_{\alpha \in \mathfrak{A}}$ of eigenvectors of the matrix H^ω , i.e. $H^\omega \psi_\alpha^\omega = \lambda_\alpha^\omega \psi_\alpha^\omega$, where $\lambda_\alpha^\omega \in \mathbb{R}$. The normalization of the matrix elements is chosen in such a way that the typical eigenvalue of the matrix is of order one:

$$\frac{1}{|\mathfrak{A}|} \sum_{\alpha} \mathbb{E} \lambda_{\alpha}^2 = \frac{1}{|\mathfrak{A}|} \mathbb{E} \operatorname{Tr} H^2 = \frac{1}{|\mathfrak{A}|} \mathbb{E} \sum_{x,y} |H_{xy}|^2 = \frac{M}{M-1}.$$

3. SCALING AND RESULTS

The central quantity of our analysis is

$$\varrho(t, x) := \mathbb{E} |\langle \delta_x, e^{-itH/2} \delta_0 \rangle|^2,$$

where $\delta_x \in \ell^2(\Lambda_N)$ denotes the standard basis vector, defined by $(\delta_x)_y = \delta_{xy}$. The factor $1/2$ is a convenient normalization since, by a standard result of random matrix theory, the spectrum of $H/2$ is asymptotically equal to the unit interval $[-1, 1]$. The function $\varrho(t, x)$ describes the ensemble average of the quantum transition probability of a particle starting from position 0 ending up at position x after time t . Note that $\sum_x \varrho(t, x) = 1$ for any $t \in \mathbb{R}$. Heuristically, the particle performs a series of random jumps of size W . The typical number of jumps in time $t = O(1)$ is of order one. Indeed, by first order perturbation theory, the small-times probability distribution for $1 \leq |x| \leq W$ is given by

$$\varrho(t, x) \sim \mathbb{E} |\langle \delta_x, (\mathbb{1} - itH/2) \delta_0 \rangle|^2 = \frac{t^2}{4} \mathbb{E} |H_{x0}|^2 = \frac{t^2}{4} \frac{1}{M-1},$$

up to higher order terms in t . Thus $\sum_{x \neq 0} \varrho(t, x)$ is an $O(1)$ quantity, separated away from zero, indicating that the distance from the origin is of $O(W)$ for times $t \sim O(1)$.

In time t the particle performs $O(t)$ jumps of size $O(W)$. We expect that the jumps are approximately independent and the trajectory is a random walk consisting of $O(t)$ steps with size $O(W)$ each. Thus, the typical distance from the origin is of order $t^{1/2}W$. We rescale time and space $(t, x) \mapsto (T, X)$ so as to make the macroscopic quantities T and X of order one, i.e. we set

$$t = \eta T, \quad x = \eta^{1/2} W X,$$

where W and η are two large parameters. Ideally, one would like to study the long time limit $\eta \rightarrow \infty$ for a fixed W . In this case, however, we know that the dynamics cannot be diffusive for $d = 1$. Indeed, as explained in the introduction, it is expected that the motion cannot be diffusive for distances larger than

W^2 ; this has in fact been proved [40] for distances larger than W^8 . Thus we have to consider a scaling limit where η and W are related and they tend simultaneously to infinity. To that end we choose an exponent $\kappa > 0$ and set $\eta \equiv \eta(W) := W^{d\kappa}$.

Our first main result establishes that $\varrho(t, x)$ behaves diffusively up to time scales $t = O(W^{d\kappa})$ if $\kappa < 1/3$.

THEOREM 3.1 (QUANTUM DIFFUSION). *Let $0 < \kappa < 1/3$ be fixed. Then for any $T_0 > 0$ and any continuous bounded function $\varphi \in C_b(\mathbb{R}^d)$ we have*

$$\lim_{W \rightarrow \infty} \sum_{x \in \Lambda_N} \varrho(W^{d\kappa}T, x) \varphi\left(\frac{x}{W^{1+d\kappa/2}}\right) = \int_{\mathbb{R}^d} dX L(T, X) \varphi(X), \quad (3.1)$$

uniformly in $N \geq W^{1+d/6}$ and $0 \leq T \leq T_0$. Here

$$L(T, X) := \int_0^1 d\lambda \frac{4}{\pi} \frac{\lambda^2}{\sqrt{1-\lambda^2}} G(\lambda T, X),$$

and G is the heat kernel

$$G(T, X) := \left(\frac{d+2}{2\pi T}\right)^{d/2} e^{-\frac{d+2}{2T} |X|^2}, \quad (3.2)$$

REMARK 3.2. The factor $d+2$ arises from a random walk in d dimensions with steps in the unit ball. If B is a random variable uniformly distributed in the d -dimensional unit ball, the covariance matrix of B is $(d+2)^{-1}\mathbf{1}$.

This result can be interpreted as follows. The limiting dynamics at macroscopic time T is not given by a single heat kernel, but by a weighted superposition of heat kernels at times λT , for $0 \leq \lambda \leq 1$. The factor λ expresses a delay arising from backtracking paths, in which the quantum particle “wastes time” by retracing its steps. If the particle is not backtracking, it is moving according to diffusive dynamics. The backtracking paths correspond to two-legged subdiagrams, and have the interpretation of a self-energy renormalization in the language of diagrammatic perturbation theory. Thus, out of the total macroscopic time T during which the particle moves, a fraction λ of T is spent moving diffusively, and a fraction $(1-\lambda)$ of T backtracking. Theorem 3.1 gives an explicit expression for the probability density $f(\lambda) = \frac{4}{\pi} \frac{\lambda^2}{\sqrt{1-\lambda^2}} \mathbf{1}(0 \leq \lambda \leq 1)$ for the particle to move during a fraction λ of T .

Our proof precisely exhibits this phenomenon. As explained in Section 4, the proof is based on an expansion of the quantum time evolution in terms of *nonbacktracking paths*. At time $t = W^{d\kappa}T$, this expansion yields a weighted superposition of paths of lengths $n = 1, \dots, [t]$ (higher values of n are strongly suppressed). Here n is the number of nonbacktracking steps, i.e. the number of steps that contribute to the effective motion of the particle. The difference $[t] - n$ is the number of steps that the particle spends backtracking. Our expansion (or, more precisely, its leading order *ladder terms*) shows that the weight of a path of n nonbacktracking steps is given by $|\alpha_n(t)|^2$, where $\alpha_n(t)$ is the Chebyshev transform of the propagator $e^{-it\xi}$ in ξ ; see (5.3). The probability density f arises from this microscopic picture by setting $n = [t]$. Then we have, as proved in Proposition 8.5 below, $t|\alpha_{[t]}(t)|^2 \rightarrow f(\lambda)$ weakly as $t \rightarrow \infty$.

Our second main result shows that the eigenvectors of H have a typical localization length larger than $W^{1+d\kappa/2}$, for any $\kappa < 1/3$. For $x \in \Lambda_N$ and $\ell > 0$ we define the characteristic function $P_{x,\ell}$ projecting onto the complement of an ℓ -neighbourhood of x ,

$$P_{x,\ell}(y) := \mathbf{1}(|y - x| \geq \ell).$$

Let $\varepsilon > 0$ and define the random subset $\mathfrak{A}_{\varepsilon, \ell}^\omega \subset \mathfrak{A}$ of eigenvectors through

$$\mathfrak{A}_{\varepsilon, \ell}^\omega := \left\{ \alpha \in \mathfrak{A} : \sum_x |\psi_\alpha^\omega(x)| \|P_{x, \ell} \psi_\alpha^\omega\| < \varepsilon \right\}.$$

The set $\mathfrak{A}_{\varepsilon, \ell}^\omega$ contains, in particular, all eigenvectors that are exponentially localized in balls of radius $O(\ell)$; see Corollary 3.4 below for a more general and precise statement.

THEOREM 3.3 (DELOCALIZATION). *Let $\varepsilon > 0$ and $0 < \kappa < 1/3$. Then*

$$\limsup_{W \rightarrow \infty} \mathbb{E} \frac{|\mathfrak{A}_{\varepsilon, W^{1+d\kappa/2}}|}{|\mathfrak{A}|} \leq 2\sqrt{\varepsilon},$$

uniformly in $N \geq W^{1+d/6}$.

Theorem 3.3 implies that the fraction of eigenvectors subexponentially localized on scales $W^{1+\kappa d/2}$ converges to zero in probability.

COROLLARY 3.4. *For fixed $\gamma > 0$ and $K > 0$ define the random subset of eigenvectors*

$$\mathfrak{B}_\ell^\omega := \left\{ \alpha \in \mathfrak{A} : \exists u \in \Lambda_N : \sum_x |\psi_\alpha^\omega(x)|^2 \exp\left[\frac{|x-u|}{\ell}\right]^\gamma \leq K \right\}. \quad (3.3)$$

Then for $0 < \kappa < 1/3$ we have

$$\lim_{W \rightarrow \infty} \mathbb{E} \frac{|\mathfrak{B}_{W^{1+\kappa d/2}}|}{|\mathfrak{A}|} = 0,$$

uniformly in $N \geq W^{1+d/6}$.

4. MAIN IDEAS OF THE PROOF

We need to compute the expectation of the squared matrix elements of the unitary time evolution $e^{-itH/2}$. A natural starting point is the power series expansion $e^{-itH/2} = \sum_{n \geq 0} (-itH/2)^n / n!$. Unfortunately, the resulting series is *unstable* for $t \rightarrow \infty$, as is manifested by the large cancellations in the sum

$$\mathbb{E} |\langle \delta_x, e^{-itH/2} \delta_y \rangle|^2 = \sum_{n, n'} \frac{i^{n-n'} t^{n+n'}}{2^{n+n'} n! n!} \mathbb{E} H_{xy}^n H_{yx}^{n'}. \quad (4.1)$$

This can be seen as follows. The expectation

$$\mathbb{E} H_{xy}^n H_{yx}^{n'} = \mathbb{E} \sum_{x_1, \dots, x_{n-1}} \sum_{y_1, \dots, y_{n'-1}} H_{xx_1} H_{x_1 x_2} \dots H_{x_{n-1} y} H_{y y_{n'-1}} \dots H_{y_1 x} \quad (4.2)$$

is traditionally represented graphically by drawing the labels $x, x_1, x_2, \dots, y_1, x$ as vertices of a path, and by identifying vertices whose labels are identical. Since the matrix elements are centred (i.e. $\mathbb{E} H_{xy} = 0$ for all x, y), each edge must be traveled at least twice in any path that yields a nonzero contribution to (4.2). It is well known that the leading order contribution to (4.2) is given by the so-called *fully backtracking paths*.

A fully backtracking path is a path generated by successively applying the transformation $a \mapsto aba$ to the trivial path x . A typical fully backtracking path may be thought of as a tree with double edges. It is not hard to see that, after summing over y , each fully backtracking path yields a contribution of order 1 to (4.2). Also, the number of fully backtracking paths is of order $4^{n+n'}$, so that the expectation (4.2) is of order $4^{n+n'}$. In particular, this implies that the main contribution to (4.1) comes from terms satisfying $n+n' \sim t$. Moreover, the series (4.1) is unstable in the sense that the sum of the absolute values of its summands behaves like e^{4t} as $t \rightarrow \infty$.

The large terms in (4.1) systematically cancel each other out similarly to the *two-legged subdiagram renormalization* in perturbative field theory. In perturbative renormalization, these cancellations are exploited by introducing appropriately adjusted fictitious counter-terms. In the current problem, however, we make use of the *Chebyshev transformation*, which removes the contribution of all backtracking paths in one step. The key observation is that, if U_n denotes the n -th Chebyshev polynomial of the second kind, then $U_n(H/2)$ can be expressed in terms of *nonbacktracking paths*. A nonbacktracking path is a path which contains no subpath of the form aba . Thus the strongest instabilities in (4.1) can be removed if $e^{-itH/2}$ is expanded into a series of Chebyshev polynomials. This idea appeared first in [7] and has recently been exploited in [26, 42] to prove, among other things, the edge-universality for band matrices. In [42] it is also stated that the same method can be used to prove delocalization of the edge eigenvectors if $W \geq N^{5/6}$, i.e. to get the bound $\ell \geq W^{6/5}$ on the localization length ℓ . Our estimate gives a slightly weaker bound, $\ell \geq W^{7/6}$, for this special case, but it applies to bulk eigenvectors as well as higher dimensions.

After the Chebyshev transform, we need to compute expectations

$$\mathbb{E} \sum'_{x_1, \dots, x_{n-1}, y_1, \dots, y_{n'-1}} \sum' H_{xx_1} H_{x_1 x_2} \cdots H_{x_{n-1} y} H_{y y_{n'-1}} \cdots H_{y_1 x}, \quad (4.3)$$

where the summations are restricted to nonbacktracking paths. As above, since $\mathbb{E}H_{ab} = 0$ every matrix element must appear at least twice in the non-trivial terms of (4.3). Taking the expectation effectively introduces a pairing, or more generally a lumping, of the factors, which can be conveniently represented by Feynman diagrams. The main contribution comes from the so-called *ladder diagrams*, corresponding to $n = n'$ and $x_i = y_i$. The contribution of these diagrams can be explicitly computed, and showed to behave diffusively. More precisely: Since we express nonbacktracking powers of H as Chebyshev polynomials in $H/2$, the contribution of each graph to the propagator $e^{-itH/2}$ carries a weight equal to the Chebyshev transform $\alpha_n(t)$ of $e^{-it\xi}$ in ξ . We shall show that $\alpha_n(t)$ is given essentially by a Bessel function of the first kind. In order to identify the limiting behaviour of the ladder diagrams, we therefore need to analyse a probability distribution on \mathbb{N} of the form $\{|\alpha_n(t)|^2\}_{n \in \mathbb{N}}$ for large t (Section 8).

The main work consists in proving that the non-ladder diagrams are negligible. Similarly to the basic idea of [18–20], the non-ladder diagrams are classified according to their combinatorial complexity. The large number of complex diagrams is offset by their small value, expressed in terms of powers of W . Conversely, diagrams containing large pieces of ladder subdiagrams have a relatively large contribution but their number is small.

More precisely, focusing only on the pairing diagrams in the Hermitian case, it is easy to see that ladder subdiagrams are marginal for power counting. We define the *skeleton* of a graph by collapsing parallel ladder rungs (called *bridges*) into a single rung. We show that the value of a skeleton diagram is given by a negative power of $M \sim CW^d$ that is proportional to the size of the skeleton diagram. This is how the dimension d enters our estimate. We then sum up all possible ladder subdiagrams corresponding to a given skeleton. Although the ladder subdiagrams do not yield additional W -powers, they represent classical random walks for which dispersive bounds are available, rendering them summable. The restriction $t \ll W^{d/3}$ comes

from summing up the skeleton diagrams. In Section 11 we present a critical skeleton that shows that this restriction is necessary without further resummation or a more refined classification of complex graphs.

5. THE PATH EXPANSION

We start by writing the expansion of $e^{-itH/2}$ in terms of nonbacktracking paths by using the Chebyshev transform.

5.1. The Chebyshev transform of $e^{-it\xi}$. The Chebyshev transform $\alpha_k(t)$ of $e^{-it\xi}$ is defined by

$$e^{-it\xi} = \sum_{k=0}^{\infty} \alpha_k(t) U_k(\xi).$$

Here U_k denotes the Chebyshev polynomial of the second kind, defined through

$$U_k(\cos \theta) = \frac{\sin(k+1)\theta}{\sin \theta} \tag{5.1}$$

for $k = 0, 1, 2, \dots$. The Chebyshev polynomials satisfy the orthogonality relation

$$\frac{2}{\pi} \int_{-1}^1 d\xi \sqrt{1-\xi^2} U_k(\xi) U_l(\xi) = \delta_{kl}.$$

Therefore the coefficients $\alpha_k(t)$ are given by

$$\alpha_k(t) = \frac{2}{\pi} \int_{-1}^1 d\xi \sqrt{1-\xi^2} e^{-it\xi} U_k(\xi). \tag{5.2}$$

The coefficient $\alpha_k(t)$ can be evaluated explicitly using the standard identities (see [32])

$$\begin{aligned} U_k(\xi) &= \frac{\xi T_{k+1}(\xi) - T_{k+2}(\xi)}{1-\xi^2}, & T_{k+2}(\xi) - 2\xi T_{k+1}(\xi) + T_k(\xi) &= 0, \\ \frac{2}{\pi} \int_{-1}^1 d\xi \frac{T_{2l}(\xi) \cos(t\xi)}{\sqrt{1-\xi^2}} &= 2(-1)^l J_{2l}(t), & \frac{2}{\pi} \int_{-1}^1 d\xi \frac{T_{2l+1}(\xi) \sin(t\xi)}{\sqrt{1-\xi^2}} &= 2(-1)^l J_{2l+1}(t), \\ J_k(t) + J_{k+2}(t) &= \frac{2(k+1)}{t} J_{k+1}(t). \end{aligned}$$

Here T_k denotes the Chebyshev polynomial of the first kind and J_k the Bessel function of the first kind; they are defined through

$$T_k(\cos \theta) := \cos(k\theta), \quad J_k(t) := \frac{1}{\pi} \int_0^\pi d\theta \cos(t \sin \theta - t\theta).$$

If $k = 2l$ is even we may therefore compute

$$\begin{aligned} \alpha_{2l}(t) &= \frac{2}{\pi} \int_{-1}^1 d\xi \sqrt{1-\xi^2} \cos(t\xi) U_k(\xi) = \frac{2}{\pi} \int_{-1}^1 d\xi \sqrt{1-\xi^2} \cos(t\xi) \frac{\xi T_{k+1}(\xi) - T_{k+2}(\xi)}{1-\xi^2} \\ &= \frac{2}{\pi} \int_{-1}^1 d\xi \cos(t\xi) \frac{T_k(\xi) - T_{k+2}(\xi)}{2\sqrt{1-\xi^2}} = (-1)^l [J_{2l}(t) + J_{2l+2}(t)] = 2(-1)^l \frac{2l+1}{t} J_{2l+1}(t). \end{aligned}$$

If $k = 2l + 1$ is odd a similar calculation yields

$$\alpha_{2l+1}(t) = -2i(-1)^l \frac{2l+2}{t} J_{2l+2}(t).$$

Thus we have the following result.

LEMMA 5.1. *We have that*

$$e^{-it\xi} = \sum_k \alpha_k(t) U_k(\xi),$$

where

$$\alpha_k(t) = 2(-i)^k \frac{k+1}{t} J_{k+1}(t). \quad (5.3)$$

Also, for all $t \in \mathbb{R}$ we have the identity

$$\sum_{k \geq 0} |\alpha_k(t)|^2 = 1, \quad (5.4)$$

as follows from the orthonormality of the Chebyshev polynomials.

5.2. Expansion in terms of nonbacktracking paths. For $n = 0, 1, 2, \dots$ let $H^{(n)}$ denote the n -th nonbacktracking power of H . It is defined by

$$H_{x_0, x_n}^{(n)} := \sum'_{x_1, \dots, x_{n-1}} H_{x_0 x_1} \cdots H_{x_{n-1} x_n},$$

where \sum' means sum under the restriction $x_i \neq x_{i+2}$ for $i = 0, \dots, n-2$. We call this restriction the *nonbacktracking condition*.

The following key observation is due to Bai and Yin [7].

LEMMA 5.2. *The nonbacktracking powers of H satisfy*

$$H^{(0)} = \mathbb{1}, \quad H^{(1)} = H, \quad H^{(2)} = H^2 - \frac{M}{M-1} \mathbb{1},$$

as well as the recursion relation

$$H^{(n)} = HH^{(n-1)} - H^{(n-2)} \quad (n \geq 3). \quad (5.5)$$

PROOF. For the convenience of the reader we give the simple proof. The cases $n = 0, 1, 2$ are easily checked.

Moreover,

$$\begin{aligned}
(HH^{(n-1)})_{x_0x_n} &= \sum_{x_1, \dots, x_{n-1}} \prod_{i=1}^{n-2} \mathbf{1}(x_i \neq x_{i+2}) H_{x_0x_1} \cdots H_{x_{n-1}x_n} \\
&= \sum_{x_1, \dots, x_{n-1}} \prod_{i=0}^{n-2} \mathbf{1}(x_i \neq x_{i+2}) H_{x_0x_1} \cdots H_{x_{n-1}x_n} \\
&\quad + \sum_{x_1, \dots, x_{n-1}} \mathbf{1}(x_0 = x_2) \prod_{i=1}^{n-2} \mathbf{1}(x_i \neq x_{i+2}) H_{x_0x_1} \cdots H_{x_{n-1}x_n} \\
&= (H^{(n)})_{x_0, x_n} + \sum_{x_3, \dots, x_{n-1}} \mathbf{1}(x_0 \neq x_4) \prod_{i=3}^{n-2} \mathbf{1}(x_i \neq x_{i+2}) H_{x_0x_3} H_{x_3x_4} \cdots H_{x_{n-1}x_n} \\
&\quad \times \sum_{x_1} \mathbf{1}(x_1 \neq x_3) |H_{x_0x_1}|^2 \\
&= (H^{(n)})_{x_0, x_n} + (H^{(n-2)})_{x_0x_n}.
\end{aligned}$$

Notice that in the last step we used (2.2). □

Feldheim and Sodin have observed [26, 42] that (5.5) is reminiscent of the recursion relation for the Chebyshev polynomials of the second kind. Let us abbreviate $\tilde{U}_n(\xi) := U_n(\xi/2)$. Then we have (see e.g. [32])

$$\tilde{U}_0(\xi) = 1, \quad \tilde{U}_1(\xi) = \xi, \quad \tilde{U}_2(\xi) = \xi^2 - 1,$$

and for $n \geq 2$

$$\tilde{U}_n(\xi) = \xi \tilde{U}_{n-1}(\xi) - \tilde{U}_{n-2}(\xi).$$

Comparing this to Lemma 5.2, we get, following [26, 42],

$$H^{(n)} = \tilde{U}_n(H) - \frac{1}{M-1} \tilde{U}_{n-2}(H).$$

Solving for $\tilde{U}_n(H)$ yields

$$\tilde{U}_n(H) = \sum_{k \geq 0} \frac{1}{(M-1)^k} H^{(n-2k)},$$

with the convention that $H^{(n)} = 0$ for $n < 0$. Therefore Lemma 5.1 yields

$$e^{-itH/2} = \sum_{n \geq 0} \alpha_n(t) \tilde{U}_n(H) = \sum_{m \geq 0} H^{(m)} \sum_{k \geq 0} \frac{\alpha_{m+2k}(t)}{(M-1)^k}.$$

We have proved the following result.

LEMMA 5.3. *We have that*

$$e^{-itH/2} = \sum_{m \geq 0} a_m(t) H^{(m)},$$

where

$$a_m(t) := \sum_{k \geq 0} \frac{\alpha_{m+2k}(t)}{(M-1)^k}.$$

6. GRAPHICAL REPRESENTATION

For ease of presentation, we assume throughout the proof of Theorem 3.1 (Sections 6 – 8) that we are in the Hermitian case (2.1a). How to extend our arguments to cover the symmetric case (2.1b) is described in Section 9.

Using Lemma 5.3 we get

$$\varrho(t, x) = \sum_{n, n' \geq 0} a_n(t) \overline{a_{n'}(t)} \mathbb{E} H_{0x}^{(n)} H_{x0}^{(n')}.$$

Expanding in nonbacktracking paths yields a graphical expansion. Let us write $H_{0x}^{(n)} H_{x0}^{(n')}$ as a sum over paths $x_0, x_1, \dots, x_{n+n'-1}, x_0$, where $x_0 = 0$ and $x_n = x$. Such a path is graphically represented as a loop of $n + n'$ vertices belonging to the set $\mathcal{V}_{n, n'} := \{0, \dots, n + n' - 1\}$; see Figure 6.1. Vertices $i \in \mathcal{V}_{n, n'}$ satisfying the nonbacktracking condition (i.e. $x_{i-1} \neq x_{i+1}$) are drawn using black dots; other vertices are drawn using white dots. There are $n + n'$ oriented edges $e_0, \dots, e_{n+n'-1}$ defined by $e_i := (i, i + 1)$ (here, and in the

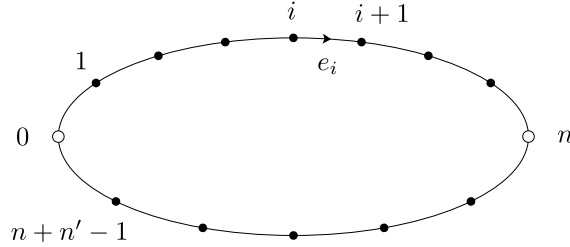


Figure 6.1: The graphical representation of paths of vertices.

following, $\mathcal{V}_{n, n'}$ is taken to be periodic). We denote by $\mathcal{E}_{n, n'} := \{e_0, \dots, e_{n+n'-1}\}$ the set of edges. In Figure 6.1 the edges are oriented clockwise. Each vertex has an outgoing and an incoming edge, and each edge e has an initial vertex $a(e)$ and final vertex $b(e)$. Moreover, we order the edges using their initial vertices.

Each vertex $i \in \mathcal{V}_{n, n'}$ carries a label $x_i \in \Lambda_N$. The labels $\mathbf{x} = (x_0, \dots, x_{n+n'-1})$ are summed over under the restriction $Q_x(\mathbf{x}) = 1$, where

$$Q_x(\mathbf{x}) := \delta_{0x_0} \delta_{x_n x} \prod_{i=0}^{n+n'-1} \mathbf{1}(1 \leq |x_i - x_{i+1}| \leq W) \prod_{i=0}^{n-2} \mathbf{1}(x_i \neq x_{i+2}) \prod_{i=n}^{n+n'-2} \mathbf{1}(x_i \neq x_{i+2}).$$

The two last products implement the nonbacktracking condition. We define the unordered pair of labels corresponding to the edge e through

$$\varrho_{\mathbf{x}}(e) := \{x_{a(e)}, x_{b(e)}\}.$$

Next, to each configuration of labels $\mathbf{x} = (x_0, \dots, x_{n+n'-1})$ we assign a *lumping* $\Gamma = \Gamma(\mathbf{x})$ of the set of edges $\mathcal{E}_{n, n'}$. Here a lumping means a partition of $\mathcal{E}_{n, n'}$ or, equivalently, an equivalence relation on $\mathcal{E}_{n, n'}$. We use the notation $\Gamma = \{\gamma\}_{\gamma \in \Gamma}$, where $\gamma \in \Gamma$ is *lump* of Γ , i.e. an equivalence class. The lumping $\Gamma = \Gamma(\mathbf{x})$ associated with the labels \mathbf{x} is defined according to the rule that e and e' are in the same lump $\gamma \in \Gamma$ if and only if $\varrho_{\mathbf{x}}(e) = \varrho_{\mathbf{x}}(e')$. Let $\tilde{\mathcal{G}}_{n, n'}$ denote the set of lumpings of $\mathcal{E}_{n, n'}$ obtained in this manner. Thus we may write

$$\mathbb{E} H_{0x}^{(n)} H_{x0}^{(n')} = \sum_{\Gamma \in \tilde{\mathcal{G}}_{n, n'}} V_x(\Gamma).$$

Here

$$V_x(\Gamma) = \sum_{\mathbf{x}}^* Q_x(\mathbf{x}) \mathbb{E} H_{x_0 x_1} \cdots H_{x_{n+n'-1} x_0},$$

where the summation is restricted to label configurations yielding the lumping Γ .

Next, observe that the expectation of a monomial $\prod_{y,z} (H_{yz})^{\nu_{yz}}$ is nonzero if and only if $\nu_{yz} = \nu_{zy}$ for all y, z (here we only use that the law of the matrix entries is invariant under rotations of the complex plane). In particular, $V_x(\Gamma)$ vanishes if one lump $\gamma \in \Gamma$ is of odd size. Defining the subset $\mathcal{G}_{n,n'} \subset \tilde{\mathcal{G}}_{n,n'}$ of lumpings whose lumps are of even size, we find that

$$\mathbb{E} H_{0x}^{(n)} H_{x0}^{(n')} = \sum_{\Gamma \in \mathcal{G}_{n,n'}} V_x(\Gamma).$$

We summarize the key properties of $\mathcal{G}_{n,n'}$.

LEMMA 6.1. *Let $\Gamma \in \mathcal{G}_{n,n'}$. Then each lump $\gamma \in \Gamma$ is of even size. Moreover, any two edges $e, e' \in \gamma$ in the same lump γ are separated by either at least two edges or a vertex in $\{0, n\}$ (nonbacktracking property).*

Next, we give an explicit expression for $V_x(\Gamma)$. We start by assigning to each lump $\gamma \in \Gamma$ an unordered pair of labels ϱ_γ . Then we pick a partition π_γ of γ into two subsets of equal size. Abbreviate these families as $\varrho_\Gamma = \{\varrho_\gamma\}_{\gamma \in \Gamma}$ and $\pi_\Gamma = \{\pi_\gamma\}_{\gamma \in \Gamma}$. Thus we get

$$V_x(\Gamma) = \sum_{\mathbf{x}} Q_x(\mathbf{x}) \sum_{\varrho_\Gamma} \sum_{\pi_\Gamma} \left(\prod_{\gamma \in \Gamma} \Delta_{\mathbf{x}}(\varrho_\gamma, \pi_\gamma) \right) \left(\prod_{\gamma \neq \gamma'} \mathbf{1}(\varrho_\gamma \neq \varrho_{\gamma'}) \right) \mathbb{E} H_{x_0 x_1} \cdots H_{x_{n+n'-1} x_0}. \quad (6.1)$$

Here, for each $\gamma \in \Gamma$, ϱ_γ ranges over all unordered pairs of labels and π_γ ranges over all partitions of γ into two subsets of equal size; $\Delta_{\mathbf{x}}(\varrho_\gamma, \pi_\gamma)$ is the indicator function of the following event: For all $e \in \gamma$ we have that $\varrho_{\mathbf{x}}(e) = \varrho_\gamma$, and

$$\begin{aligned} e, e' \in \gamma \text{ belong to the same subset of } \pi_\gamma &\implies x_{a(e)} = x_{a(e')}, \quad x_{b(e)} = x_{b(e')} \\ e, e' \in \gamma \text{ belong to different subsets of } \pi_\gamma &\implies x_{a(e)} = x_{b(e')}, \quad x_{b(e)} = x_{a(e')}. \end{aligned}$$

This definition of $\Delta_{\mathbf{x}}(\varrho_\gamma, \pi_\gamma)$ has the following interpretation. All edges in γ (corresponding to matrix elements) have the same unordered pair of labels (and hence represent copies of the same random variable H_{yz} or its complex conjugate). Moreover, each random variable H_{yz} must appear as many times as its complex conjugate; random variables indexed by two edges $e, e' \in \gamma$ are identical if e, e' belong to the same subset of π_γ , and each other's complex conjugates if e, e' belong to different subsets of π_γ .

Note that the expectation in (6.1) is equal to

$$\frac{1}{(M-1)^{\bar{n}}}, \quad (6.2)$$

where $\bar{n} := \frac{n+n'}{2}$. In particular, $V_x(\Gamma) \geq 0$.

An important subset of lumpings of $\mathcal{E}_{n,n'}$ is the set of pairings, $\mathcal{P}_{n,n'} \subset \mathcal{G}_{n,n'}$, which contains all lumpings Γ satisfying $|\gamma| = 2$ for all $\gamma \in \Gamma$. We call two-element lumps $\sigma \in \mathcal{P}_{n,n'}$ *bridges*. Given a pairing $\Gamma \in \mathcal{P}_{n,n'}$, we say that e and e' are *bridged* (in Γ) if there is a $\sigma \in \Gamma$ such that $\sigma = \{e, e'\}$. Bridges are represented graphically by drawing a line, for each $\{e, e'\} \in \Gamma$, from the edge e to e' ; see Figure 6.2. Thus a pairing

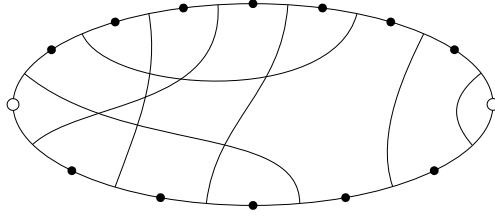


Figure 6.2: A pairing of edges.

$\Gamma \in \mathcal{P}_{n,n'}$ is the edge set of a graph whose vertex set is $\mathcal{E}_{n,n'}$. If Γ is a pairing, each bridge $\sigma \in \Gamma$ has a unique partition π_σ of its edges, so that the expression (6.1) for $V_x(\Gamma)$ may be rewritten in the simpler form

$$V_x(\Gamma) = \sum_{\mathbf{x}} Q_x(\mathbf{x}) \left(\prod_{\{e,e'\} \in \Gamma} \mathbf{1}(x_{a(e)} = x_{b(e')}) \mathbf{1}(x_{b(e)} = x_{a(e')}) \right) \times \left(\prod_{\sigma \neq \sigma'} \prod_{e \in \sigma} \prod_{e' \in \sigma'} \mathbf{1}(\varrho_{\mathbf{x}}(e) \neq \varrho_{\mathbf{x}}(e')) \right) \frac{1}{(M-1)^{\bar{n}}}. \quad (6.3)$$

The main contribution to the expansion is given by the *ladder pairing* $L_n \in \mathcal{P}_{n,n}$. It is defined as

$$L_n := \{ \{e_0, e_{2n-1}\}, \{e_1, e_{2n-2}\}, \dots, \{e_{n-1}, e_n\} \}.$$

The ladder is represented graphically in Figure 6.3.

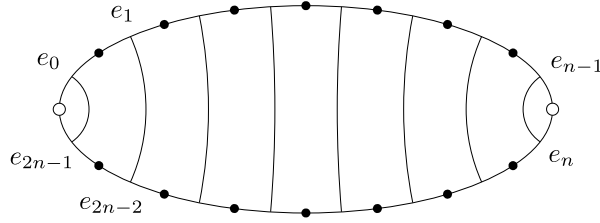


Figure 6.3: The ladder pairing.

7. THE NON-LADDER LUMPINGS

In this section we estimate the contribution of the non-ladder lumpings and show that it vanishes in the limit $W \rightarrow \infty$. Let $\mathcal{G}_{n,n'}^* \subset \mathcal{G}_{n,n'}$ denote the set of non-ladder lumpings, i.e. $\mathcal{G}_{n,n'}^* := \mathcal{G}_{n,n'}$ if $n \neq n'$ and $\mathcal{G}_{n,n}^* := \mathcal{G}_{n,n} \setminus \{L_n\}$. Similarly, let $\mathcal{P}_{n,n'}^* := \mathcal{P}_{n,n'} \cap \mathcal{G}_{n,n'}^*$ denote the set of non-ladder pairings.

We shall prove the following result.

PROPOSITION 7.1. *Let $0 < \kappa < 1/3$ and pick a β satisfying $0 < \beta < 2/3 - 2\kappa$. Then there is a constant C such that*

$$\sum_x \sum_{n, n' \geq 0} |a_n(\eta T) a_{n'}(\eta T)| \sum_{\Gamma \in \mathcal{G}_{n, n'}^*} V_x(\Gamma) \leq \frac{C}{W^{d\beta}},$$

for W larger than some $W_0(T, \kappa)$ and $N \geq W^{1+d/6}$.

The rest of this section is devoted to the proof of Proposition 7.1.

7.1. Controlling the non-pairings. Replacing the expectation in (6.1) with (6.2) we get

$$V_x(\Gamma) = \sum_{\mathbf{x}} Q_x(\mathbf{x}) \sum_{\varrho_\Gamma} \sum_{\pi_\Gamma} \left(\prod_{\gamma \in \Gamma} \Delta_{\mathbf{x}}(\varrho_\gamma, \pi_\gamma) \right) \left(\prod_{\gamma \neq \gamma'} \mathbf{1}(\varrho_\gamma \neq \varrho_{\gamma'}) \right) \frac{1}{(M-1)^{\bar{n}}}.$$

We start by estimating the sum over all lumpings $\Gamma \in \mathcal{G}_{n, n'}^*$ in terms of a sum over all pairings $\Gamma \in \mathcal{P}_{n, n'}^*$. Let us define

$$R_x(\Gamma) := \sum_{\mathbf{x}} Q_x(\mathbf{x}) \sum_{\varrho_\Gamma} \sum_{\pi_\Gamma} \left(\prod_{\gamma \in \Gamma} \Delta_{\mathbf{x}}(\varrho_\gamma, \pi_\gamma) \right) \frac{1}{(M-1)^{\bar{n}}}. \quad (7.1)$$

LEMMA 7.2. *For all $n, n' \in \mathbb{N}$ we have*

$$\sum_{\Gamma \in \mathcal{G}_{n, n'}^*} V_x(\Gamma) \leq \sum_{\Gamma \in \mathcal{P}_{n, n'}^*} R_x(\Gamma).$$

PROOF. Let ϱ_γ and π_γ be given for each $\gamma \in \Gamma$. For each γ , pick any pairing Σ_γ of γ that is compatible with π_γ in the sense that, for each bridge $\sigma \in \Sigma_\gamma$, the two edges of σ belong to different subsets of π_γ . If $n = n'$, we additionally require that not all Σ_γ 's are subsets of the Ladder L_n (such a choice is always possible). Next, set $\varrho_\sigma = \varrho_\gamma$ for all $\sigma \in \Sigma_\gamma$. Note that each bridge σ carries a unique partition π_σ . It is then easy to see that for any pairing Σ_γ as above, we have

$$\Delta_{\mathbf{x}}(\varrho_\gamma, \pi_\gamma) \leq \prod_{\sigma \in \Sigma_\gamma} \Delta_{\mathbf{x}}(\varrho_\sigma, \pi_\sigma).$$

Thus, by partitioning each $\gamma \in \Gamma$ into bridges, we see that each term in $\sum_{\Gamma \in \mathcal{G}_{n, n'}^*} V_x(\Gamma)$ is bounded by a corresponding term in $\sum_{\Gamma \in \mathcal{P}_{n, n'}^*} R_x(\Gamma)$. In fact, there is an overcounting arising from the different ways of partitioning γ into bridges. \square

Because of Lemma 7.2 we may restrict ourselves to pairings. We estimate $\sum_{\Gamma \in \mathcal{P}_{n, n'}^*} R_x(\Gamma)$. If Γ is a pairing we may write, just like (6.3), the expression (7.1) in the simpler form

$$R_x(\Gamma) = \sum_{\mathbf{x}} Q_x(\mathbf{x}) \left(\prod_{\{e, e'\} \in \Gamma} \mathbf{1}(x_{a(e)} = x_{b(e')}) \mathbf{1}(x_{b(e)} = x_{a(e')}) \right) \frac{1}{(M-1)^{\bar{n}}}. \quad (7.2)$$

7.2. Collapsing of parallel bridges. Let us introduce the set $\overline{\mathcal{P}}_{n,n'}^*$, defined as the set of all non-ladder pairings of $\mathcal{E}_{n,n'}$. Clearly, $\mathcal{P}_{n,n'}^*$ is a proper subset of $\overline{\mathcal{P}}_{n,n'}^*$ (due to the nonbacktracking condition of Lemma 6.1 which is imposed on pairings in $\mathcal{P}_{n,n'}^*$).

Let $n, n' \geq 0$ and $\Gamma \in \overline{\mathcal{P}}_{n,n'}^*$. For any i, j , we say that the two bridges $\{e_i, e_j\}$ and $\{e_{i+1}, e_{j-1}\}$ of Γ are *parallel* if $i+1, j \notin \{0, n\}$; see Figure 7.1. Two parallel bridges may be collapsed to obtain a new pairing Γ' of a smaller set of edges, in which the parallel bridges are replaced by a single bridge. More precisely: We obtain $\Gamma' \in \overline{\mathcal{P}}_{m,m'}^*$ from $\Gamma \in \overline{\mathcal{P}}_{n,n'}^*$ by removing the vertices $i+1$ and j , by creating the edges $(i, i+2)$ and $(j-1, j+1)$, and by bridging them. Finally, we rename the vertices using the increasing integers $0, 1, 2, \dots, n+n'-3$; by definition, the new name of the vertex n is m , and m' is defined through $m+m'+2 = n+n'$. The converse operation of collapsing bridges, *expanding bridges*, is self-explanatory.

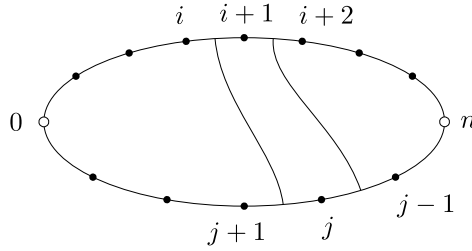


Figure 7.1: Two parallel bridges.

In the next lemma we iterate the above procedure $\Gamma \mapsto \Gamma'$ until all parallel bridges have been collapsed.

LEMMA 7.3. *Let $\Gamma \in \overline{\mathcal{P}}_{n,n'}^*$. Then there exist $m \leq n$, $m' \leq n'$, and a pairing $S(\Gamma) \in \overline{\mathcal{P}}_{m,m'}^*$ containing no parallel bridges, such that Γ may be obtained from $S(\Gamma)$ by successively expanding bridges. This defines $S(\Gamma)$ uniquely.*

PROOF. Successively collapse all parallel bridges in Γ ; see Figure 7.2. The result is clearly independent of the order in which this is done. \square

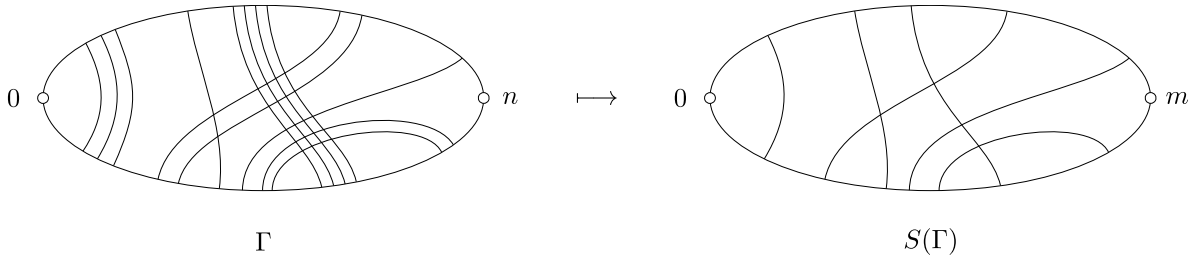


Figure 7.2: Collapsing parallel bridges to obtain the skeleton pairing.

We call the pairing $\Sigma = S(\Gamma)$ the *skeleton* of Γ . The set of skeleton pairings of the edges $\mathcal{E}_{m,m'}$ is denoted by

$$\mathcal{S}_{m,m'}^* := \bigcup_{n,n' \geq 0} \{S(\Gamma) : \Gamma \in \overline{\mathcal{P}}_{n,n'}^*\} \cap \overline{\mathcal{P}}_{m,m'}^*.$$

Note that $\mathcal{S}_{m,m'}^*$ is in general not a subset of $\mathcal{P}_{m,m'}^*$. The following lemma summarizes the key properties of $\mathcal{S}_{m,m'}^*$.

LEMMA 7.4. (i) Each $\Sigma \in \mathcal{S}_{m,m'}^*$ contains no parallel bridges.

(ii) Let $\Sigma \in \mathcal{S}_{m,m'}^*$ and $\sigma = \{e, e'\} \in \Sigma$. Then e, e' are adjacent only if $e \cap e' \in \{0, m\}$.

(iii) If $\bar{m} := \frac{m+m'}{2} = 1$ then $\mathcal{S}_{m,m'}^* = \emptyset$.

PROOF. Statement (i) follows immediately from the definition of $S(\Gamma)$. Statement (ii) is a consequence of the nonbacktracking property of pairings in $\mathcal{P}_{n,n'}^*$, i.e. Lemma 6.1. To see this, let $\Sigma \in \mathcal{S}_{m,m'}^*$ be of the form $\Sigma = S(\Gamma)$ for some $\Gamma \in \mathcal{P}_{n,n'}^*$. If $\Sigma = S(\Gamma)$ contains a bridge $\{e, e'\}$ consisting of two consecutive edges e, e' , then Γ must also contain a bridge $\{f, f'\}$ consisting of two consecutive edges f, f' . If $e \cap e' \notin \{0, m\}$, then $f \cap f' \notin \{0, n\}$, in contradiction to Lemma 6.1. Statement (iii) is an immediate consequence of (ii) and the requirement that $L_1 \notin \mathcal{S}_{1,1}^*$. \square

7.3. Contribution of parallel bridges. For given n and n' , we estimate $\sum_{\Gamma \in \mathcal{P}_{n,n'}^*} R_x(\Gamma)$ by summing over skeleton pairings Σ , followed by summing over all possible ways of expanding the bridges of Σ .

We observe that a pairing $\Gamma \in \mathcal{P}_{n,n'}^*$ is uniquely determined by its skeleton $\Sigma = S(\Gamma) \in \mathcal{S}_{m,m'}^*$ for some positive integers m, m' as well as a family $\ell_\Sigma = \{\ell_\sigma\}_{\sigma \in \Sigma}$ satisfying $|\ell_\Sigma| = \bar{n}$, where ℓ_σ encodes the number of parallel bridges that were collapsed to form the bridge σ . Here $\ell_\sigma \geq 1$ is a positive integer and $|\ell_\Sigma| := \sum_{\sigma \in \Sigma} \ell_\sigma$. Let $G_{\ell_\Sigma}(\Sigma)$ denote the pairing obtained from Σ by expanding the bridge σ into ℓ_σ parallel bridges, for each $\sigma \in \Sigma$. Thus Γ may be recovered from its skeleton through $\Gamma = G_{\ell_\Sigma}(\Sigma)$ for a unique family ℓ_Σ . For given $p \in \mathbb{N}$, the sum over all pairings Γ satisfying $|\Gamma| = p$ therefore becomes

$$\sum_{n+n'=2p} \sum_{\Gamma \in \mathcal{P}_{n,n'}^*} R_x(\Gamma) = \sum_{m+m' \leq 2p} \sum_{\Sigma \in \mathcal{S}_{m,m'}^*} \sum_{\ell_\Sigma: |\ell_\Sigma|=p} R_x(G_{\ell_\Sigma}(\Sigma)). \quad (7.3)$$

Next, we define and estimate the contribution to $R_x(\Gamma)$ of a set of ℓ parallel bridges. Let $\ell \geq 1$, and two labels y, z be given. Then we define

$$D_\ell(y, z) := \sum_{x_0, \dots, x_\ell} \delta_{x_0 y} \delta_{x_\ell z} \prod_{i=0}^{\ell-1} \mathbf{1}(1 \leq |x_i - x_{i+1}| \leq W).$$

Thus, $D_\ell(y, z)$ is equal to the number of paths of length ℓ from y to z , whereby each step takes values in $\{x : 1 \leq |x| \leq W\}$. (We could also have included the nonbacktracking restriction in the definition of D_ℓ , but this is not needed as we only want an upper bound on $R_x(\Sigma)$). Graphically, D_ℓ corresponds to the contribution of ℓ parallel bridges; see Figure 7.3.

We need the following straightforward properties of D_ℓ .

LEMMA 7.5. Let $\ell \in \mathbb{N}$. Then for each y we have

$$\sum_z D_\ell(y, z) = M^\ell.$$

Moreover, for each y and z we have

$$D_\ell(y, z) \leq M^{\ell-1}.$$

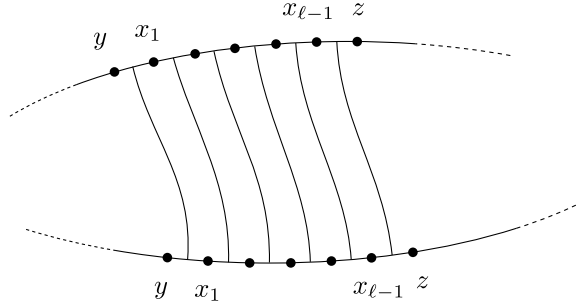


Figure 7.3: Summing up ℓ parallel bridges.

as well as

$$D_\ell(y, z) \leq \frac{C}{\ell^{d/2}} M^{\ell-1} + \frac{C}{N^d} M^\ell.$$

for some constant C .

PROOF. The first two statements are obvious. The last follows from a standard local central limit theorem; see for instance the proof in [42]. \square

7.4. Orbits of vertices. Fix $\Gamma \in \mathcal{P}_{n,n'}^*$. We observe that the product in (7.2) may be interpreted as an indicator function that fixes labels along paths of vertices. To this end, we define a map $\tau \equiv \tau_\Gamma$ on the vertex set $\mathcal{V}_{n,n'}$. Start with a vertex $i \in \mathcal{V}_{n,n'}$. Let e be the outgoing edge of i (i.e. $e = (i, i+1)$), and e' the edge bridged by Γ to e . Then we define τi as the final vertex of e' (i.e. $e' = (\tau i - 1, \tau i)$). Thus the product in (7.2) may be rewritten as

$$\prod_{\{e, e'\} \in \Gamma} \mathbf{1}(x_{a(e)} = x_{b(e')}) \mathbf{1}(x_{b(e)} = x_{a(e')}) = \prod_{i \in \mathcal{V}_{n,n'}} \delta_{x_i x_{\tau i}}.$$

Starting from any vertex $i \in \mathcal{V}_{n,n'}$ we construct a path $(i, \tau i, \tau^2 i, \dots)$. In this fashion the set of vertices is partitioned into orbits of τ ; see Figure 7.4. Let $[i] \subset \mathcal{V}_{n,n'}$ denote the orbit of the vertex $i \in \mathcal{V}_{n,n'}$.

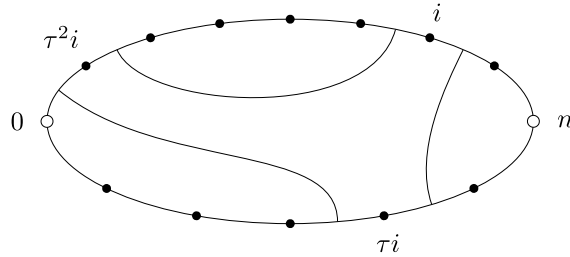


Figure 7.4: Construction of the orbit $[i]$ of the vertex i .

Next, let $\Sigma = S(\Gamma) \in \mathcal{S}_{m,m'}^*$ be the skeleton pairing of Γ , and let the family ℓ_Σ be defined through $\Gamma = G_{\ell_\Sigma}(\Sigma)$. The map $\tau \equiv \tau_\Sigma$ on the skeleton pairing Σ is defined exactly as for Γ above. In order to sum

over all labels $\mathbf{x} = (x_0, \dots, x_{n+n'-1})$ in the expression for $R_x(G_{\ell_\Sigma}(\Sigma))$, we split the set of labels \mathbf{x} into two parts: labels of vertices between two parallel bridges, and labels associated with vertices of Σ . In order to make this precise, we need the following definitions.

Let $Z(\Sigma)$ be the set of orbits of Σ . It contains the distinguished orbits $[0]$ and $[m]$, which receive the labels 0 and x respectively. (Note that we may have $[0] = [m]$, in which case x must be 0 .) We assign a label y_ζ to each orbit $\zeta \in Z(\Sigma)$, and define the family $\mathbf{y}_\Sigma := \{y_\zeta\}_{\zeta \in Z(\Sigma)}$. Each bridge $\sigma \in \Sigma$ ‘‘sits between two orbits’’ $\zeta_1(\sigma)$ and $\zeta_2(\sigma)$. More precisely, let $e = (i, i+1) \in \sigma$ be the smaller edge of σ . Then we set $\zeta_1(\sigma) := [i]$ and $\zeta_2(\sigma) := [i+1]$. (Note that using the larger edge of σ in this definition would simply exchange $\zeta_1(\sigma)$ and $\zeta_2(\sigma)$; this is of no consequence for the following.)

LEMMA 7.6. *For given $\Sigma \in \mathcal{S}_{m,m'}^*$, ℓ_Σ , and $\Gamma = G_{\ell_\Sigma}(\Sigma) \in \mathcal{P}_{n,n'}^*$ we have*

$$R_x(\Gamma) \leq \frac{1}{(M-1)^{\bar{n}}} \sum_{\mathbf{y}_\Sigma} \mathbf{1}(0 = y_{[0]}) \mathbf{1}(x = y_{[m]}) \prod_{\sigma \in \Sigma} D_{\ell_\sigma}(y_{\zeta_1(\sigma)}, y_{\zeta_2(\sigma)}). \quad (7.4)$$

PROOF. The left-hand side of (7.4) is given by the expression (7.2). The summation over all x_i 's between parallel bridges of Γ is contained in the factors D_{ℓ_σ} , and the summation over all the remaining x_i 's is replaced by the sum over \mathbf{y}_Σ . We relaxed the nonbacktracking condition in $Q_x(\mathbf{x})$ to obtain an upper bound. \square

Next, let $Z^*(\Sigma) := Z(\Sigma) \setminus \{[0]\}$ and define $L(\Sigma) := |Z^*(\Sigma)|$. The set $Z^*(\Sigma)$ is the set of orbits whose label is summed over in $\sum_x R_x(\Gamma)$. The following lemma gives an upper bound on $L(\Sigma)$. It states, roughly, that the number of orbits (or free labels) is bounded by $2\bar{m}/3$; we refer to it as the 2/3 rule. Compare this bound with the trivial bound $L(\Sigma) \leq \bar{m}$, which would be sharp if Σ were allowed to have parallel bridges.

LEMMA 7.7 (THE 2/3 RULE). *Let $\Sigma \in \mathcal{S}_{m,m'}^*$. Then $L(\Sigma) \leq \frac{2\bar{m}}{3} + \frac{1}{3}$.*

PROOF. Let $Z'(\Sigma) := Z(\Sigma) \setminus \{[0], [m]\}$. We show that every orbit $\zeta \in Z'(\Sigma)$ consists of at least 3 vertices. Let $i \in \mathcal{V}_{m,m'}$ belong to $\zeta \in Z'(\Sigma)$. Then, by Lemma 7.4 (ii), we have that $\tau i \neq i$. By assumption, $\tau i \notin \{0, m\}$. Hence $\tau^2 i \neq i$, for otherwise Σ would have two parallel bridges, in contradiction to Lemma 7.4 (i). Therefore the orbit of τ contains at least 3 vertices. Note that there are orbits containing exactly 3 vertices, as depicted in Figure 7.4.

The total number of vertices of Σ not including the vertices 0 and m is $2\bar{m} - 2$, so that we get

$$3|Z'(\Sigma)| \leq 2\bar{m} - 2.$$

The claim follows from the bound $|Z^*(\Sigma)| \leq |Z'(\Sigma)| + 1$. \square

7.5. Bound on $R_x(\Gamma)$. As in the previous subsection, we fix $\Gamma \in \mathcal{P}_{n,n'}^*$, $\Sigma = S(\Gamma) \in \mathcal{S}_{m,m'}^*$, and ℓ_Σ satisfying $\Gamma = G_{\ell_\Sigma}(\Sigma)$.

We start by observing that the product in (7.4) may be rewritten in terms of a multigraph $\Pi(\Sigma)$ on the vertex set $Z(\Sigma)$. Each factor $D_{\ell_\sigma}(y_{\zeta_1(\sigma)}, y_{\zeta_2(\sigma)})$ yields an edge connecting the orbits ζ_1 and ζ_2 . In other words, there is a one-to-one map, which we denote by ϕ , between bridges of Σ and edges of $\Pi(\Sigma)$; each bridge $\sigma \in \Sigma$ gives rise to an edge $\phi(\sigma)$ of $\Pi(\Sigma)$ connecting $\zeta_1(\sigma)$ and $\zeta_2(\sigma)$. See Figure 7.5 for an example of such a multigraph.

LEMMA 7.8. *There is a subset of bridges $\Sigma_T \subset \Sigma$ of size $|\Sigma_T| = L(\Sigma)$, such that, in the subgraph of $\Pi(\Sigma)$ with the edge set $\phi(\Sigma_T)$, each orbit $\zeta \in Z^*(\Sigma)$ is connected to $[0]$.*

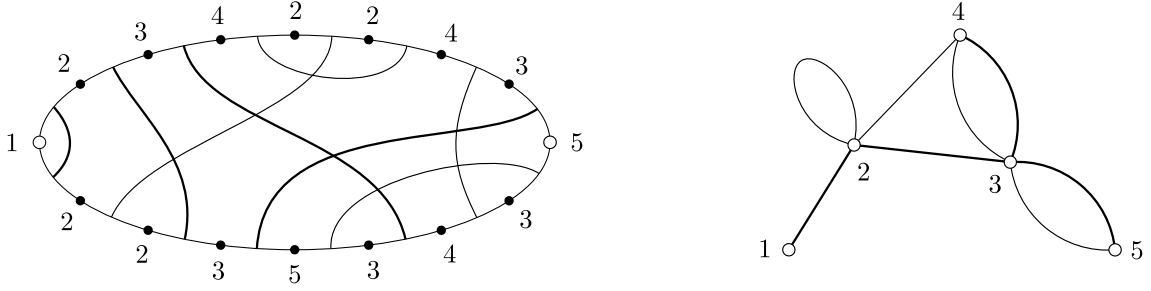


Figure 7.5: Left: a skeleton pairing Σ giving rise to 5 orbits indexed by $Z(\Sigma) = \{1, \dots, 5\}$; the bridges in Σ_T , for one possible choice of Σ_T , are drawn using thick lines. Right: the corresponding multigraph on the vertex set $Z(\Sigma)$; the edges in $\phi(\Sigma_T)$ are drawn using thick lines.

PROOF. Starting from $\zeta_0 = [0]$, we construct a sequence of orbits $\zeta_0, \zeta_1, \dots, \zeta_{L(\Sigma)}$, and a sequence of bridges $\sigma_1, \dots, \sigma_{L(\Sigma)}$, with the property that for all $k = 1, \dots, L(\Sigma)$ there is a $k' < k$ such that ζ_k and $\zeta_{k'}$ are connected by $\phi(\sigma_k)$.

Assume that $\zeta_0, \dots, \zeta_{k-1}$ have already been constructed. Let i be the smallest vertex of $\mathcal{V}_{m, m'} \setminus (\zeta_0 \cup \dots \cup \zeta_{k-1})$. Then we set $\zeta_k = [i]$. By construction, the vertex $i-1$ belongs to an orbit $\zeta_{k'}$ for some $k' < k$. Set σ_k to be the bridge containing $\{i-1, i\}$. Hence, by definition of $\Pi(\Sigma)$, we see that ζ_k and $\zeta_{k'}$ are connected by $\phi(\sigma_k)$.

The set Σ_T is given by $\{\sigma_1, \dots, \sigma_{L(\Sigma)}\}$. □

Because $|\Sigma_T| = L(\Sigma)$, the subgraph of $\Pi(\Sigma)$ with the edge set $\phi(\Sigma_T)$ is a tree that connects all orbits in $Z^*(\Sigma)$ to $[0]$. Let us call this tree $\mathcal{T}(\Sigma)$. Its root is $[0]$.

Next, we observe that

$$|\Sigma \setminus \Sigma_T| \geq 1. \quad (7.5)$$

Indeed, using Lemma 7.7 and $\bar{m} \geq 2$ we find

$$|\Sigma \setminus \Sigma_T| = \bar{m} - L(\Sigma) \geq \frac{\bar{m}}{3} - \frac{1}{3} \geq \frac{1}{3}. \quad (7.6)$$

We now estimate (7.4) as follows. Each factor indexed by $\sigma \in \Sigma \setminus \Sigma_T$ is estimated by $\sup_{y, z} D_{\ell_\sigma}(y, z)$. As it turns out, we need to exploit the heat kernel decay for at least one bridge in $\Sigma \setminus \Sigma_T$. Pick a bridge $\bar{\sigma} \in \Sigma \setminus \Sigma_T$ (By (7.5) there is such a bridge). Using Lemma 7.5, we estimate

$$\sup_{y, z} D_{\ell_\sigma}(y, z) \leq M^{\ell_\sigma - 1} \quad \text{if } \sigma \in \Sigma \setminus (\Sigma_T \cup \{\bar{\sigma}\}), \quad (7.7a)$$

$$\sup_{y, z} D_{\ell_\sigma}(y, z) \leq \frac{C}{\ell_\sigma^{d/2}} M^{\ell_\sigma - 1} + \frac{C}{N^d} M^{\ell_\sigma} \quad \text{if } \sigma = \bar{\sigma}. \quad (7.7b)$$

Since $N \geq WM^{1/6}$ and $M \sim CW^d$ we find

$$\frac{C}{\ell_\sigma^{d/2}} M^{\ell_\sigma - 1} + \frac{C}{N^d} M^{\ell_\sigma} \leq CM^{\ell_\sigma - 1} \left(\frac{1}{\ell_\sigma^{1/2}} + \frac{1}{M^{1/6}} \right)$$

where we replaced d with 1 to obtain an upper bound. Thus we get

$$\sum_x R_x(\Gamma) \leq \frac{C}{(M-1)^{\bar{n}}} \left(\frac{1}{\ell_{\bar{\sigma}}^{1/2}} + \frac{1}{M^{1/6}} \right) \prod_{\sigma \in \Sigma \setminus \Sigma_T} M^{\ell_{\sigma}-1} \sum_{\mathbf{y}_{\Sigma}} \mathbf{1}(0 = y_{[0]}) \prod_{\sigma \in \Sigma_T} D_{\ell_{\sigma}}(y_{\zeta_1(\sigma)}, y_{\zeta_2(\sigma)}).$$

We perform the summation over \mathbf{y}_{Σ} by starting at the leaves of $\mathcal{T}(\Sigma)$ and moving towards the root $[0]$. Each vertex ζ of $\mathcal{T}(\Sigma)$ carries a label y_{ζ} . Let us choose a leaf ζ of $\mathcal{T}(\Sigma)$, and denote by ζ' the parent of ζ in $\mathcal{T}(\Sigma)$. Let $\sigma \in \Sigma$ be the (unique) bridge such that $\phi(\sigma)$ connects ζ and ζ' . Then summation over y_{ζ} yields the factor

$$\sum_{y_{\zeta}} D_{\ell_{\sigma}}(y_{\zeta}, y_{\zeta'}) = M^{\ell_{\sigma}}, \quad (7.8)$$

by Lemma 7.5. Continuing in this manner until we reach the root, we find

$$\begin{aligned} \sum_x R_x(\Gamma) &\leq \frac{C}{(M-1)^{\bar{n}}} \left(\frac{1}{\ell_{\bar{\sigma}}^{1/2}} + \frac{1}{M^{1/6}} \right) \prod_{\sigma \in \Sigma \setminus \Sigma_T} M^{\ell_{\sigma}-1} \prod_{\sigma \in \Sigma_T} M^{\ell_{\sigma}} \\ &= C \left(\frac{M}{M-1} \right)^{\bar{n}} \left(\frac{1}{\ell_{\bar{\sigma}}^{1/2}} + \frac{1}{M^{1/6}} \right) \frac{1}{M^{|\Sigma \setminus \Sigma_T|}}. \end{aligned}$$

Now (7.6) implies

$$|\Sigma \setminus \Sigma_T| \geq \frac{\bar{m}}{3} - \frac{1}{3},$$

so that

$$\sum_x R_x(\Gamma) \leq C \left(\frac{M}{M-1} \right)^{\bar{n}} \left(\frac{1}{\ell_{\bar{\sigma}}^{1/2}} + \frac{1}{M^{1/6}} \right) \frac{M^{1/3}}{M^{\bar{m}/3}}. \quad (7.9)$$

Notice that (7.9) results from an ℓ^1 - ℓ^{∞} -summation procedure, where the ℓ^1 -bound (7.8) was used for propagators associated with bridges in Σ_T , and the ℓ^{∞} -bound (7.7) for propagators associated with bridges in $\Sigma \setminus \Sigma_T$. The bound (7.7a) is a simple power counting bound; the bound (7.7b), improved by the heat kernel decay, is used only for one bridge. Note that in the original setup (2.1) each row and column of H contains M nonzero entries H_{xy} , whose positions are determined by the condition $1 \leq |x - y| \leq W$. If we removed this last condition and only required that each row and column contain M nonzero entries in arbitrary locations off the diagonal, then all bounds relying solely on power counting would remain valid. In particular, (7.9) would be valid without the factor $\ell_{\bar{\sigma}}^{-1/2} + M^{-1/6}$, which results from the heat kernel decay associated with the special band structure.

7.6. Sum over pairings. We may now estimate $\sum_{n+n'=2p} \sum_{\Gamma \in \mathcal{P}_{n,n}^*} \sum_x R_x(\Gamma)$ for fixed p . Let first $p, m, m' \geq 0$ and $\Sigma \in \mathcal{S}_{m,m'}^*$. Then (7.9) yields

$$\sum_{\ell_{\Sigma}: |\ell_{\Sigma}|=p} \sum_x R_x(G_{\ell_{\Sigma}}(\Sigma)) \leq C \left(\frac{M}{M-1} \right)^p \frac{M^{1/3}}{M^{\bar{m}/3}} \sum_{\ell_{\Sigma}: |\ell_{\Sigma}|=p} \left(\frac{1}{\ell_{\bar{\sigma}}^{1/2}} + \frac{1}{M^{1/6}} \right).$$

The sum on the right-hand side is equal to

$$\begin{aligned}
\sum_{\ell_1 + \dots + \ell_{\bar{m}} = p} \left(\frac{1}{\ell_1^{1/2}} + \frac{1}{M^{1/6}} \right) &= \sum_{\ell_1=1}^{p-\bar{m}+1} \left(\frac{1}{\ell_1^{1/2}} + \frac{1}{M^{1/6}} \right) \sum_{\ell_2 + \dots + \ell_{\bar{m}} = p - \ell_1} 1 \\
&\leq \sum_{\ell_1=1}^p \left(\frac{1}{\ell_1^{1/2}} + \frac{1}{M^{1/6}} \right) \binom{p - \ell_1 - 1}{\bar{m} - 2} \\
&\leq C \left(\frac{1}{p^{1/2}} + \frac{1}{M^{1/6}} \right) \frac{p^{\bar{m}-1}}{(\bar{m} - 2)!}.
\end{aligned}$$

Next, we note that

$$|\mathcal{S}_{m,m'}^*| \leq (2\bar{m} - 1)(2\bar{m} - 3) \cdots 3 \cdot 1 \leq 2^{\bar{m}} \bar{m}!.$$

This expresses the fact that the first edge of Σ can be bridged with at most $(2\bar{m} - 1)$ edges, the next remaining edge with at most $(2\bar{m} - 3)$ edges, and so on. Therefore (7.3) and Lemma 7.4 (iii) yield

$$\begin{aligned}
\sum_{n+n'=2p} \sum_{\Gamma \in \mathcal{P}_{n,n'}^*} \sum_x R_x(\Gamma) &\leq C \sum_{4 \leq m+m' \leq 2p} 2^{\bar{m}} \bar{m}! \left(\frac{M}{M-1} \right)^p \frac{M^{1/3}}{M^{\bar{m}/3}} \frac{p^{\bar{m}-1}}{(\bar{m} - 2)!} \left(\frac{1}{p^{1/2}} + \frac{1}{M^{1/6}} \right) \\
&\leq C \frac{M^{1/3}}{p} \left(\frac{1}{p^{1/2}} + \frac{1}{M^{1/6}} \right) \left(\frac{M}{M-1} \right)^p \sum_{4 \leq m+m' \leq 2p} \bar{m}^2 2^{\bar{m}} \left(\frac{p}{M^{1/3}} \right)^{\bar{m}} \\
&\leq \frac{M^{1/3}}{p} \left(\frac{1}{p^{1/2}} + \frac{1}{M^{1/6}} \right) \left(\frac{M}{M-1} \right)^p \sum_{\bar{m}=2}^p \left(\frac{Cp}{M^{1/3}} \right)^{\bar{m}}.
\end{aligned}$$

Thus, Lemma 7.2 yields

$$\sum_{n+n'=2p} h_{n,n'} \leq \frac{M^{1/3}}{p} \left(\frac{1}{p^{1/2}} + \frac{1}{M^{1/6}} \right) \left(\frac{M}{M-1} \right)^p \sum_{r=2}^p \left(\frac{Cp}{M^{1/3}} \right)^r, \quad (7.10)$$

where we abbreviated

$$h_{n,n'} := \sum_{\Gamma \in \mathcal{G}_{n,n'}^*} \sum_x V_x(\Gamma).$$

7.7. Conclusion of the proof. In this subsection we complete the proof of Proposition 7.1 by showing that the error

$$E_W := \sum_{n,n'} |a_n(\eta T) a_{n'}(\eta T)| h_{n,n'} \quad (7.11)$$

satisfies $E_W = o(1)$ as $W \rightarrow \infty$, uniformly in $N \geq W^{1+d/6}$.

We begin by deriving bounds on the coefficients $a_n(t)$.

LEMMA 7.9. (i) We have

$$\sum_{n \geq 0} |a_n(t)|^2 = 1 + O(M^{-1}), \quad (7.12)$$

uniformly in $t \in \mathbb{R}$.

(ii) We have

$$|a_n(t)| \leq C \frac{t^n}{n!}. \quad (7.13)$$

PROOF. We start with (i). Write

$$\sum_{n \geq 0} |a_n(t)|^2 = \sum_{n \geq 0} \sum_{k, k' \geq 0} \frac{\alpha_{n+2k}(t) \overline{\alpha_{n+2k'}(t)}}{(M-1)^{k+k'}}.$$

The term $k = k' = 0$ yields 1 by (5.4). The rest is equal, by (5.4), to

$$\sum_{n \geq 0} \sum_{k+k' > 0} \frac{\alpha_{n+2k}(t) \overline{\alpha_{n+2k'}(t)}}{(M-1)^{k+k'}} \leq \sum_{k+k' > 0} \sum_{n \geq 0} \frac{|\alpha_{n+2k}(t)|^2}{(M-1)^{k+k'}} \leq \sum_{k+k' > 0} \frac{1}{(M-1)^{k+k'}} = O(M^{-1}).$$

In order to prove (ii), we use the integral representation (see [32])

$$J_n(t) = \frac{\left(\frac{t}{2}\right)^n}{\sqrt{\pi} \Gamma\left(n + \frac{1}{2}\right)} \int_{-1}^1 d\lambda e^{it\lambda} (1 - \lambda^2)^{n-\frac{1}{2}}.$$

Therefore

$$|\alpha_n(t)| \leq 2 \frac{n+1}{t} \frac{\left(\frac{t}{2}\right)^{n+1}}{\sqrt{\pi} \Gamma\left(n + \frac{3}{2}\right)} \frac{\pi}{2} \leq \frac{t^n}{n!}. \quad (7.14)$$

Moreover, (5.4) yields

$$|\alpha_n(t)| \leq 1. \quad (7.15)$$

We use the estimate

$$|a_n(t)| \leq \sum_{k \geq 0} \frac{|\alpha_{n+2k}(t)|}{(M-1)^k}.$$

Let us first consider the case $t \leq n$. Then it is easy to see that $\frac{t^{n+2k}}{(n+2k)!} \leq \frac{t^n}{n!}$. Together with (7.14) this yields

$$|a_n(t)| \leq \sum_{k \geq 0} \frac{1}{(M-1)^k} \frac{t^{n+2k}}{(n+2k)!} \leq \frac{t^n}{n!} \sum_{k \geq 0} \frac{1}{(M-1)^k} \leq C \frac{t^n}{n!}.$$

If $t > n$ we have $\frac{t^n}{n!} \geq C$. Thus the bound (7.15) yields

$$|a_n(t)| \leq \sum_{k \geq 0} \frac{C}{(M-1)^k} \leq C \frac{t^n}{n!}. \quad \square$$

Using the new variables $p := \bar{n} = \frac{n+n'}{2}$ and $q := \frac{n-n'}{2}$ we find from the definition (7.11)

$$E_W \leq \sum_{p \geq 0} \sum_{q=-p}^p |a_{p+q}(\eta T) a_{p-q}(\eta T)| h_{p+q, p-q}.$$

Next, we observe that Lemma 7.9 (ii) implies that terms corresponding to $n, n' \gg t = \eta T \sim CM^\kappa T$ are strongly suppressed. Thus we introduce a cutoff at $p = M^\mu$, where $\kappa < \mu < \frac{1}{3}$. Let us first consider the terms $p \leq M^\mu$. We need to estimate

$$\begin{aligned} E_W^{\leq} &:= \sum_{p=0}^{M^\mu} \sum_{q=-p}^p |a_{p+q}(\eta T) a_{p-q}(\eta T)| h_{p+q, p-q} \\ &\leq \left(\sum_{p=0}^{M^\mu} \sum_{q=-p}^p |a_{p+q}(\eta T) a_{p-q}(\eta T)|^2 \right)^{1/2} \left(\sum_{p=0}^{M^\mu} \sum_{q=-p}^p (h_{p+q, p-q})^2 \right)^{1/2} \\ &\leq C \left(\sum_{p=0}^{M^\mu} \sum_{q=-p}^p (h_{p+q, p-q})^2 \right)^{1/2}, \end{aligned}$$

where we used Lemma 7.9 (i). Thus,

$$(E_W^{\leq})^2 \leq C \sum_{p=0}^{M^\mu} \left[\sum_{q=-p}^p h_{p+q, p-q} \right]^2 \leq \sum_{p=2}^{M^\mu} \left[\frac{M^{1/3}}{p} \left(\frac{1}{p^{1/2}} + \frac{1}{M^{1/6}} \right) \left(\frac{M}{M-1} \right)^p \sum_{\bar{m}=2}^p \left(\frac{Cp}{M^{1/3}} \right)^{\bar{m}} \right]^2,$$

by (7.10). For $p \leq M^\mu$ and W large enough, the term in the square brackets is bounded by

$$C \frac{M^{1/3}}{p} \left(\frac{1}{p^{1/2}} + \frac{1}{M^{1/6}} \right) \frac{1}{\left(1 - \frac{1}{M}\right)^p} \left(\frac{p}{M^{1/3}} \right)^2 \leq C \frac{p^{1/2}}{M^{1/3}} \leq \frac{C}{M^{1/6}}.$$

Thus we find $(E_W^{\leq})^2 \leq CM^{\mu-1/3}$.

Let us now consider the case $p > M^\mu$, i.e. estimate

$$E_W^> := \sum_{p > M^\mu} \sum_{q=-p}^p |a_{p+q}(\eta T) a_{p-q}(\eta T)| h_{p+q, p-q}.$$

By (7.13) and the elementary inequality $\frac{p!}{(p-q)!} \leq \frac{(p+q)!}{p!}$ we have

$$|a_{p+q}(t) a_{p-q}(t)| \leq C \frac{t^{2p}}{(p+q)!(p-q)!} \leq C \frac{t^{2p}}{p!p!}.$$

This gives

$$\begin{aligned} E_W^> &\leq C \sum_{p > M^\mu} \frac{(\eta T)^{2p}}{p!p!} \sum_{q=-p}^p h_{p+q, p-q} \\ &\leq C \sum_{p > M^\mu} \frac{(\eta T)^{2p}}{p!p!} \frac{M^{1/3}}{p} \left(\frac{1}{p^{1/2}} + \frac{1}{M^{1/6}} \right) \left(\frac{M}{M-1} \right)^p \sum_{\bar{m}=2}^p \left(\frac{Cp}{M^{1/3}} \right)^{\bar{m}} \end{aligned}$$

by (7.10). Setting $\eta \sim CM^\kappa$ yields

$$\begin{aligned}
E_W^{\geq} &\leq \sum_{p>M^\mu} \frac{(CM^\kappa T)^{2p}}{p!p!} \sum_{\bar{m}=0}^{p-2} \left(\frac{Cp}{M^{1/3}} \right)^{\bar{m}} \\
&\leq \sum_{p>M^\mu} \left(\frac{CM^\kappa T}{p} \right)^{2p} + \sum_{p>M^\mu} \left(\frac{CM^{2\kappa} T^2}{pM^{1/3}} \right)^p \\
&\leq \sum_{p>M^\mu} (CM^{\kappa-\mu} T)^{2p} + \sum_{p>M^\mu} (CM^{2\kappa-1/3-\mu} T^2)^p \\
&\leq (CM^{\kappa-\mu} T)^{2M^\mu} + (CM^{2\kappa-1/3-\mu} T)^{M^\mu}.
\end{aligned}$$

Choosing $\mu = 1/3 - \beta$ (where, we recall, $0 < \beta < 2/3 - 2\kappa$) completes the proof of Proposition 7.1.

8. THE LADDER PAIRINGS

In this section we analyse the contribution of the ladder pairings, $\sum_{n \geq 0} |a_n(\eta T)|^2 V_x(L_n)$, and complete the proof of Theorem 3.1. (Recall that $\eta := W^{d\kappa}$ is the time scale.) Recalling the expression (6.3), and noting that in the case of the ladder the variables x_0, \dots, x_n determine the value of all variables x_0, \dots, x_{2n-1} , we readily find

$$\begin{aligned}
V_x(L_n) &= \frac{1}{(M-1)^n} \sum_{\mathbf{x} \in \Lambda_N^{n+1}} \delta_{0x_0} \delta_{xx_n} \prod_{i=0}^{n-1} \mathbf{1}(1 \leq |x_{i+1} - x_i| \leq W) \\
&\quad \times \prod_{i=0}^{n-2} \mathbf{1}(x_i \neq x_{i+2}) \prod_{0 \leq i < j \leq n-1} \mathbf{1}(\{x_i, x_{i+1}\} \neq \{x_j, x_{j+1}\}). \quad (8.1)
\end{aligned}$$

Throughout this section we assume that $\eta = W^{d\kappa}$ for some $\kappa < 1/3$.

We perform a series of steps to simplify the expression (8.1). In a first step, we get rid of the last product.

LEMMA 8.1. *Under the assumptions of Proposition 7.1 we have*

$$\sum_{n \geq 0} |a_n(\eta T)|^2 V_x(L_n) = \sum_{n \geq 0} |a_n(\eta T)|^2 V_x^1(n) - E_x^1,$$

where

$$V_x^1(n) := \frac{1}{(M-1)^n} \sum_{\mathbf{x} \in \Lambda_N^{n+1}} \delta_{0x_0} \delta_{xx_n} \prod_{i=0}^{n-1} \mathbf{1}(1 \leq |x_{i+1} - x_i| \leq W) \prod_{i=0}^{n-2} \mathbf{1}(x_i \neq x_{i+2})$$

and

$$\sum_x |E_x^1| \leq \frac{C}{W^{d\beta}}.$$

PROOF. For each $\mathbf{x} = (x_0, \dots, x_n) \in \Lambda_N^{n+1}$ we write $1 = \sum_{P: \{0, \dots, n-1\}} \Delta_P(\mathbf{x})$, where the sum ranges over all partitions P of the set $\{0, \dots, n-1\}$, and $\Delta_P(\mathbf{x})$ is the indicator function

$$\Delta_P(\mathbf{x}) := \prod_{0 \leq i < j \leq n-1} \begin{cases} \mathbf{1}(\{x_i, x_{i+1}\} = \{x_j, x_{j+1}\}) & \text{if } i \text{ and } j \text{ belong to the same lump of } P \\ \mathbf{1}(\{x_i, x_{i+1}\} \neq \{x_j, x_{j+1}\}) & \text{if } i \text{ and } j \text{ belong to different lumps of } P. \end{cases}$$

Notice that if $P = P_0 := \{\{0\}, \dots, \{n-1\}\}$ then $\Delta_P(\mathbf{x})$ is the last product of (8.1). Let us define

$$E_x^1(n) := \frac{1}{(M-1)^n} \sum_{\mathbf{x} \in \Lambda_N^{n+1}} \delta_{0x_0} \delta_{xx_n} \prod_{i=0}^{n-1} \mathbf{1}(1 \leq |x_{i+1} - x_i| \leq W) \prod_{i=0}^{n-2} \mathbf{1}(x_i \neq x_{i+2}) \sum_{P \neq P_0} \Delta_P(\mathbf{x}).$$

Thus, by definition, we have

$$E_x^1 = \sum_{n \geq 0} |a_n(\eta T)|^2 E_x^1(n).$$

Next, we estimate $\sum_x |E_x^1|$. We begin by observing that each partition P of $\{0, \dots, n-1\}$ uniquely defines a partition $\Gamma(P) \in \mathcal{G}_{n,n}^*$. Indeed, each lump $p \in P$ gives rise to the lump $\gamma \in \Gamma(P)$ defined by $\gamma = \bigcup_{i \in p} \{e_i, e_{2n-1-i}\}$. In particular, $\Gamma(P) \neq \Gamma(P')$ if $P \neq P'$. We now claim that

$$\frac{1}{(M-1)^n} \sum_{\mathbf{x} \in \Lambda_N^{n+1}} \delta_{0x_0} \delta_{xx_n} \prod_{i=0}^{n-1} \mathbf{1}(1 \leq |x_{i+1} - x_i| \leq W) \prod_{i=0}^{n-2} \mathbf{1}(x_i \neq x_{i+2}) \Delta_P(\mathbf{x}) \leq V_x(\Gamma(P)).$$

This can be directly read off (6.1); there is in fact an overcounting arising from the summation over $\boldsymbol{\pi}_\Gamma$. Thus we find

$$\sum_x |E_x^1| \leq \sum_x \sum_{n \geq 0} |a_n(\eta T)|^2 \sum_{\substack{P: \{0, \dots, n-1\} \\ P \neq P_0}} V_x(\Gamma(P)) \leq \sum_x \sum_{n \geq 0} |a_n(\eta T)|^2 \sum_{\Gamma \in \mathcal{G}_{n,n}^*} V_x(\Gamma).$$

Invoking Proposition 7.1 completes the proof. \square

In a second step, we get rid of the second to last product in (8.1), i.e. the nonbacktracking condition.

LEMMA 8.2. *For any $T \geq 0$ we have*

$$\sum_{n \geq 0} |a_n(\eta T)|^2 V_x^1(n) = \sum_{n \geq 0} |a_n(\eta T)|^2 V_x^2(n) - E_x^2,$$

where

$$V_x^2(n) := \frac{1}{(M-1)^n} \sum_{\mathbf{x} \in \Lambda_N^{n+1}} \delta_{0x_0} \delta_{xx_n} \prod_{i=0}^{n-1} \mathbf{1}(1 \leq |x_{i+1} - x_i| \leq W)$$

and

$$\sum_x |E_x^2| \leq \frac{C}{W^{2d/3}}.$$

PROOF. We find

$$\sum_x |E_x^2| = \sum_{n \geq 0} |a_n(\eta T)|^2 \frac{1}{(M-1)^n} \sum_{\mathbf{x} \in \Lambda_N^{n+1}} \delta_{0x_0} \prod_{i=0}^{n-1} \mathbf{1}(1 \leq |x_{i+1} - x_i| \leq W) \left[1 - \prod_{i=0}^{n-2} \mathbf{1}(x_i \neq x_{i+2}) \right].$$

The expression in the square brackets is equal to

$$1 - \prod_{i=0}^{n-2} (1 - \mathbf{1}(x_i = x_{i+2})) = \sum_{k=1}^{n-2} (-1)^{k+1} \sum_{0 \leq i_1 < \dots < i_k \leq n-2} \prod_{j=1}^k \mathbf{1}(x_{i_j} = x_{i_j+2}).$$

Therefore summing over \mathbf{x} yields

$$\begin{aligned} \sum_x |E_x^2| &\leq \sum_{n \geq 0} |a_n(\eta T)|^2 \frac{1}{(M-1)^n} \sum_{k=1}^{n-2} \binom{n-2}{k} M^{n-k} \\ &= \sum_{n \geq 0} |a_n(\eta T)|^2 \left(\frac{M}{M-1} \right)^n \left[\left(1 + \frac{1}{M} \right)^{n-2} - 1 \right]. \end{aligned}$$

We introduce a cutoff at $n = M^{1/3}$. The part $n \leq M^{1/3}$ is bounded by

$$\sum_{n \geq 0} |a_n(\eta T)|^2 \left(\frac{M}{M-1} \right)^M \left[\left(1 + \frac{1}{M} \right)^{M^{1/3}} - 1 \right] \leq C(e^{M^{-2/3}} - 1) \leq \frac{C}{M^{2/3}},$$

by Lemma 7.9 (i). The part $n > M^{1/3}$ is estimated using Lemma 7.9 (ii), exactly as in the estimate of $E_W^>$ in Section 7.7. \square

We summarize what we have proved so far.

LEMMA 8.3. *Under the assumptions of Proposition 7.1 we have*

$$\sum_{n \geq 0} |a_n(\eta T)|^2 V_x(L_n) = \sum_{n \geq 0} |\alpha_n(\eta T)|^2 P_x(n) + E_x,$$

where

$$P_x(n) := \frac{1}{M^n} \sum_{\mathbf{x} \in \Lambda_N^{n+1}} \delta_{0x_0} \delta_{xx_n} \prod_{i=0}^{n-1} \mathbf{1}(1 \leq |x_{i+1} - x_i| \leq W)$$

and

$$\sum_x |E_x| \leq \frac{C}{W^{d\beta}}.$$

PROOF. The claim follows from Lemmas 8.1 and 8.2, combined with an argument identical to the proof of Lemma 7.9 (i) that allows us to replace $|a_n(t)|^2$ with $|\alpha_n(t)|^2$. We replaced the factor $\frac{1}{(M-1)^n}$ with $\frac{1}{M^n}$ by introducing a cutoff at $n = M^{1/3}$, exactly as in the proof of Lemma 8.2. \square

The expression $P_x(n)$ is the (normalized) number of paths in \mathbb{Z}^d of length n from 0 to any point in the set $x + N\mathbb{Z}^d$, whereby each step takes values in $\{y : 1 \leq |y| \leq W\}$.

In a third step, we use the central limit theorem to replace $P_x(n)$ with a Gaussian. Recall the definition of the heat kernel

$$G(T, X) = \left(\frac{d+2}{2\pi T} \right)^{d/2} e^{-\frac{d+2}{2T} |X|^2}.$$

LEMMA 8.4. *Let $\varphi \in C_b(\mathbb{R}^d)$ and $T \geq 0$. Then we have*

$$\lim_{W \rightarrow \infty} \sum_x P_x([\eta T]) \varphi\left(\frac{x}{W^{1+d\kappa/2}}\right) = \int dX G(T, X) \varphi(X), \quad (8.2)$$

where $[\cdot]$ denotes the integer part.

PROOF. Let $\tilde{P}_x(n)$ denote the normalized number of paths in \mathbb{Z}^d of length n from 0 to x , whereby each step takes values in $\{y : 1 \leq |y| \leq W\}$. Then we have

$$\begin{aligned} \sum_{x \in \Lambda_N} P_x([\eta T]) \varphi\left(\frac{x}{W^{1+d\kappa/2}}\right) &= \sum_{x \in \Lambda_N} \sum_{\nu \in \mathbb{Z}^d} \tilde{P}_{x+\nu N}([\eta T]) \varphi\left(\frac{x}{W^{1+d\kappa/2}}\right) \\ &= \sum_{x \in \mathbb{Z}^d} \tilde{P}_x([\eta T]) \varphi\left(\frac{\pi(x)}{W^{1+d\kappa/2}}\right), \end{aligned}$$

where $\pi(x)$ is defined through $\pi(x) \in \Lambda_N$ and $x - \pi(x) \in N\mathbb{Z}^d$. Define the sequence of i.i.d. random variables A_1, A_2, \dots whose law is $\frac{1}{M} \sum_{a \in \mathbb{Z}^d} \mathbf{1}(1 \leq |a| \leq W) \delta_a$, where δ_a denotes the point measure at a . Then we have

$$\sum_{x \in \mathbb{Z}^d} \tilde{P}_x([\eta T]) \varphi\left(\frac{\pi(x)}{W^{1+d\kappa/2}}\right) = \mathbb{E} \varphi\left(\frac{\pi(A_1 + \dots + A_{[\eta T]})}{W^{1+d\kappa/2}}\right). \quad (8.3)$$

Next, we introduce the partition

$$1 = \mathbf{1}(|A_1 + \dots + A_{[\eta T]}| < N/2) + \mathbf{1}(|A_1 + \dots + A_{[\eta T]}| \geq N/2)$$

in the expectation in (8.3). The second resulting term is bounded by

$$\|\varphi\|_\infty \mathbb{P}(|A_1 + \dots + A_{[\eta T]}| \geq N/2).$$

This vanishes in the limit $W \rightarrow \infty$ by the central limit theorem, since $\frac{N}{W\sqrt{[\eta T]}} \rightarrow \infty$ by assumption.

The first term resulting from the partition is

$$\mathbb{E} \varphi\left(\frac{A_1 + \dots + A_{[\eta T]}}{W^{1+d\kappa/2}}\right) \mathbf{1}(|A_1 + \dots + A_{[\eta T]}| < N/2) = \mathbb{E} \varphi\left(\frac{A_1 + \dots + A_{[\eta T]}}{W^{1+d\kappa/2}}\right) + o(1),$$

by the same argument as above. Therefore we get

$$\sum_{x \in \Lambda_N} P_x([\eta T]) \varphi\left(\frac{x}{W^{1+d\kappa/2}}\right) = \mathbb{E} \varphi\left(\frac{B_1 + \dots + B_{[\eta T]}}{\sqrt{[\eta T]}}\right) + o(1),$$

where $B_i := \frac{A_i}{W} \frac{\sqrt{[\eta T]}}{\sqrt{\eta}}$. The covariance matrix of B_i is $\frac{T}{d+2} \mathbf{1} + o(1)$, and the claim (8.2) follows by the central limit theorem. \square

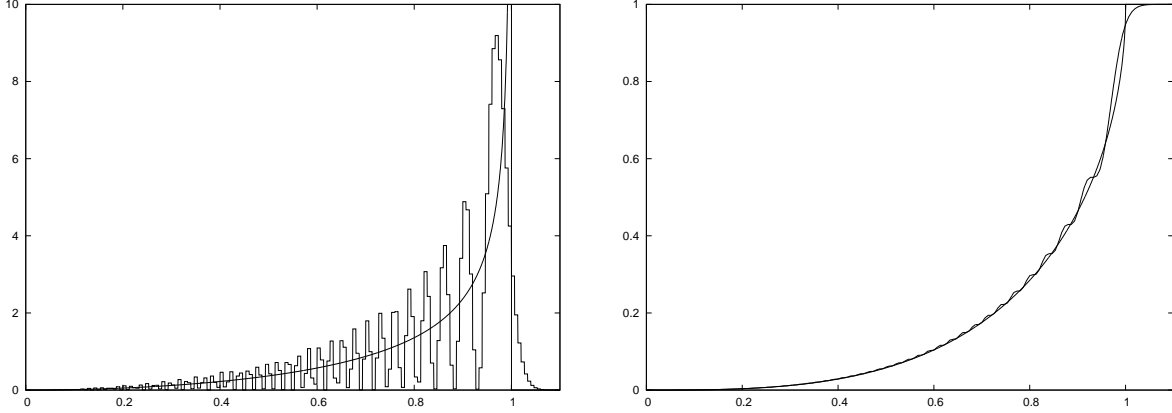


Figure 8.1: The functions $f_t(\lambda)$, $f(\lambda)$ (left) and $F_t(\lambda)$, $F(\lambda)$ (right). Here we chose $t = 150$.

In a fourth and final step, we replace the probability distribution $|\alpha_n(t)|^2$ with its asymptotic distribution. For the following we fix some test function $\varphi \in C_b(\mathbb{R}^d)$. Testing against φ in Lemma 8.3 yields

$$\sum_{n \geq 0} |a_n(\eta T)|^2 \sum_x V_x(L_n) \varphi\left(\frac{x}{W^{1+d\kappa/2}}\right) = \sum_{n \geq 0} |\alpha_n(\eta T)|^2 \sum_x P_x(n) \varphi\left(\frac{x}{W^{1+d\kappa/2}}\right) + O\left(\frac{\|\varphi\|_\infty}{W^{d\beta}}\right). \quad (8.4)$$

While the distribution $|\alpha_n(t)|^2$ has no limit as $t \rightarrow \infty$, it turns out that the rescaled distribution,

$$f_t(\lambda) := t |\alpha_{[t\lambda]}(t)|^2,$$

converges weakly to

$$f(\lambda) := \frac{4}{\pi} \frac{\lambda^2}{\sqrt{1-\lambda^2}} \mathbf{1}(0 \leq \lambda \leq 1).$$

In order to prove this, we consider the integrated distribution

$$F_t(\lambda) := \int_0^\lambda d\xi f_t(\xi).$$

We now show that $F_t(\lambda)$ converges pointwise to $F(\lambda) = \int_0^\lambda f$. See Figure 8.1 for a graph of the functions f_t, f, F_t, F .

PROPOSITION 8.5. *The pointwise limit*

$$F(\lambda) := \lim_{t \rightarrow \infty} F_t(\lambda)$$

exists for all $\lambda \geq 0$ and satisfies

$$F(\lambda) = \int_0^\lambda d\xi \frac{4}{\pi} \frac{\xi^2}{\sqrt{1-\xi^2}} = \frac{2}{\pi} \left(\arcsin \lambda - \lambda \sqrt{1-\lambda^2} \right) \quad (\lambda \in [0, 1]) \quad (8.5a)$$

$$F(\lambda) = 1 \quad (\lambda > 1). \quad (8.5b)$$

PROOF. See Appendix A. □

In order to conclude the proof of Theorem 3.1, we need the following result.

PROPOSITION 8.6. *Let $T \geq 0$. Then*

$$\lim_{W \rightarrow \infty} \sum_{n \geq 0} |a_n(\eta T)|^2 \sum_x V_x(L_n) \varphi\left(\frac{x}{W^{1+d\kappa/2}}\right) = \int_0^\infty d\lambda f(\lambda) \int dX G(\lambda T, X) \varphi(X).$$

Indeed, Theorem 3.1 is an immediate consequence of Propositions 7.1 and 8.6. The rest of this section is devoted to the proof of Proposition 8.6.

We begin by observing that the family of probability measures defined by the densities $\{f_t\}_{t \geq 0}$ is tight, so that we may cut out values of λ in the range $[0, \delta) \cup (1 - \delta, \infty)$.

LEMMA 8.7. *Let $\varepsilon > 0$. Then there is a $\delta > 0$ and a $t_0 \geq 0$ such that*

$$F(\delta) + 1 - F(1 - \delta) \leq \frac{\varepsilon}{\|\varphi\|_\infty}$$

and

$$F_t(\delta) + 1 - F_t(1 - \delta) \leq \frac{\varepsilon}{\|\varphi\|_\infty}$$

for all $t \geq t_0$.

PROOF. By Proposition 8.5 we have that

$$F(\delta) + 1 - F(1 - \delta) \rightarrow 0 \tag{8.6}$$

as $\delta \rightarrow 0$. Choose $\delta > 0$ small enough that the left-hand side of (8.6) is bounded by $\frac{\varepsilon}{2\|\varphi\|_\infty}$. Moreover, Proposition 8.5 also implies that there is a t_0 such that

$$F_t(\delta) + 1 - F_t(1 - \delta) \leq F(\delta) + 1 - F(1 - \delta) + \frac{\varepsilon}{2\|\varphi\|_\infty}$$

for all $t \geq t_0$. □

Now by (8.4), Proposition 8.6 will follow if we can show

$$\sum_{n \geq 0} |\alpha_n(\eta T)|^2 \sum_x P_x(n) \varphi\left(\frac{x}{W^{1+d\kappa/2}}\right) = \int_0^\infty d\lambda f(\lambda) \int dX G(\lambda T, X) \varphi(X) + o(1),$$

i.e.

$$\int_0^\infty d\lambda f_{\eta T}(\lambda) \sum_x P_x([\eta T \lambda]) \varphi\left(\frac{x}{W^{1+d\kappa/2}}\right) = \int_0^\infty d\lambda f(\lambda) \int dX G(\lambda T, X) \varphi(X) + o(1). \tag{8.7}$$

Lemma 8.7 implies that in order to prove (8.7) it suffices to prove

$$\int_\delta^{1-\delta} d\lambda f_{\eta T}(\lambda) \sum_x P_x([\eta T \lambda]) \varphi\left(\frac{x}{W^{1+d\kappa/2}}\right) = \int_\delta^{1-\delta} d\lambda f(\lambda) \int dX G(\lambda T, X) \varphi(X) + o(1), \tag{8.8}$$

for every $\delta > 0$.

Next, note that, by Lemma 8.4, the sum on the left-hand side of (8.8) converges to $\int dX G(\lambda T, X) \varphi(X)$ for each $\lambda \in [\delta, 1 - \delta]$. In order to invoke the dominated convergence theorem, we need an integrable bound on $f_t(\lambda)$.

LEMMA 8.8. *Let $\delta > 0$. Then there is a $C > 0$ such that $f_t(\lambda) \leq C$ for all $\lambda \in [\delta, 1 - \delta]$ and t large enough.*

PROOF. From Lemma 5.1 we get

$$|\alpha_{[t\lambda]}(t)|^2 t \leq C |J_{[t\lambda]+1}(t)|^2 t.$$

We estimate this using the following result due to Krasikov (see [36], Theorem 2). Setting $\mu := (2\nu+1)(2\nu+3)$ and assuming that $\nu > -1/2$ and $t > \sqrt{\mu + \mu^{2/3}}/2$, we have the bound

$$|J_\nu(t)|^2 \leq \frac{4}{\pi} \frac{4t^2 - (2\nu+1)(2\nu+5)}{(4t^2 - \mu)^{3/2} - \mu}.$$

Setting $\nu = [t\lambda] + 1$ yields $|J_{[t\lambda]+1}(t)|^2 \leq \frac{C}{t}$ for $\lambda \in (\delta, 1 - \delta)$ and t large enough. This completes the proof. \square

By Lemmas 8.8 and 8.4, it is enough to prove that

$$\int_{\delta}^{1-\delta} d\lambda f_{\eta T}(\lambda) \int dX G(\lambda T, X) \varphi(X) = \int_{\delta}^{1-\delta} d\lambda f(\lambda) \int dX G(\lambda T, X) \varphi(X) + o(1). \quad (8.9)$$

Let us abbreviate

$$g(\lambda) := \int dX G(\lambda T) \varphi(X).$$

The proof of Proposition 8.6 is therefore completed by the following result.

LEMMA 8.9. *Let $\delta > 0$. Then*

$$\lim_{t \rightarrow \infty} \int_{\delta}^{1-\delta} d\lambda f_t(\lambda) g(\lambda) = \int_{\delta}^{1-\delta} d\lambda f(\lambda) g(\lambda).$$

PROOF. The proof is a simple integration by parts. It is easy to check that on $[\delta, 1 - \delta]$ the function g is smooth and its derivative is bounded. We find

$$\int_{\delta}^{1-\delta} d\lambda f_t(\lambda) g(\lambda) = \int_{\delta}^{1-\delta} d\lambda F_t'(\lambda) g(\lambda) = - \int_{\delta}^{1-\delta} d\lambda F_t(\lambda) g'(\lambda) + F_t(1 - \delta) g(1 - \delta) - F_t(\delta) g(\delta).$$

Proposition 8.5 and dominated convergence yield the claim. \square

9. SYMMETRIC MATRICES

In this section we describe how to extend the argument of Sections 6 – 8 to the symmetric case (2.1b). While in the Hermitian case (2.1a) we had

$$\mathbb{E} H_{xy} H_{yx} = \frac{1}{M-1}, \quad \mathbb{E} H_{xy} H_{xy} = 0,$$

we now have

$$\mathbb{E} H_{xy} H_{yx} = \mathbb{E} H_{xy} H_{xy} = \frac{1}{M-1}. \quad (9.1)$$

Since the distribution of H_{xy} is symmetric, Lemma 6.1 also holds in the symmetric case. However, (9.1) implies that there is no restriction on the order of the labels associated with an edge. Thus we replace (6.1) with

$$V_x(\Gamma) = \sum_{\mathbf{x}} Q_x(\mathbf{x}) \sum_{\varrho_\Gamma} \left(\prod_{\gamma \in \Gamma} \Delta_{\mathbf{x}}(\varrho_\gamma) \right) \left(\prod_{\gamma \neq \gamma'} \mathbf{1}(\varrho_\gamma \neq \varrho_{\gamma'}) \right) \frac{1}{(M-1)^{\bar{n}}}, \quad (9.2)$$

where

$$\Delta_{\mathbf{x}}(\varrho_\gamma) := \prod_{e \in \gamma} \mathbf{1}(\varrho_{\mathbf{x}}(e) = \varrho_\gamma).$$

Next, we define the set $\mathcal{G}_{n,n'}^*$ as the set of lumpings $\mathcal{G}_{n,n'}$ without the complete ladder and the complete antiladder (see its definition below). It is easy to see that the analogue of Lemma 7.2 holds with

$$R_x(\Gamma) := \sum_{\mathbf{x}} Q_x(\mathbf{x}) \sum_{\varrho_\Gamma} \left(\prod_{\gamma \in \Gamma} \Delta_{\mathbf{x}}(\varrho_\gamma) \right) \frac{1}{(M-1)^{\bar{n}}}.$$

It therefore suffices to estimate the contribution of pairings $\Gamma \in \mathcal{P}_{n,n'}^*$. We have that

$$R_x(\Gamma) = \sum_{\mathbf{x}} Q_x(\mathbf{x}) \frac{1}{(M-1)^{\bar{n}}} \times \prod_{\{e,e'\} \in \Gamma} \left(\mathbf{1}(x_{a(e)} = x_{b(e')}) \mathbf{1}(x_{b(e)} = x_{a(e')}) + \mathbf{1}(x_{a(e)} = x_{a(e')}) \mathbf{1}(x_{b(e)} = x_{b(e')}) \right). \quad (9.3)$$

Thus, the graphical representation of pairings has to be modified as follows. Each bridge $\sigma \in \Gamma$ carries a tag, *straight* or *twisted*, which arises from multiplying out the product in (9.3). Twisted bridges are graphically represented with dashed lines.

In order to find a good notion of combinatorial complexity of pairings, we define *antiparallel bridges* as follows. Two bridges $\{e_i, e_j\}$ and $\{e_{i+1}, e_{j+1}\}$ are antiparallel if $i+1, j+1 \notin \{0, n\}$; see Figure 9.1. An *antiladder* is a sequence of bridges such that two consecutive bridges are antiparallel. It is easy to see that,

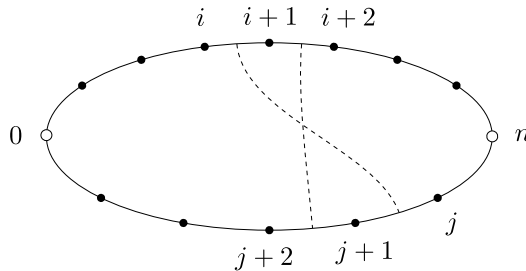


Figure 9.1: Two antiparallel twisted bridges. Compare to Figure 7.1.

in addition to ladders whose rungs are straight bridges, antiladders whose rungs are twisted bridges have a leading order contribution.

The skeleton $\Sigma = S(\Gamma)$ of the pairing Γ is obtained from Γ by the following procedure. A pair of parallel straight bridges is collapsed to form a single straight bridge. A pair of antiparallel twisted bridges is collapsed

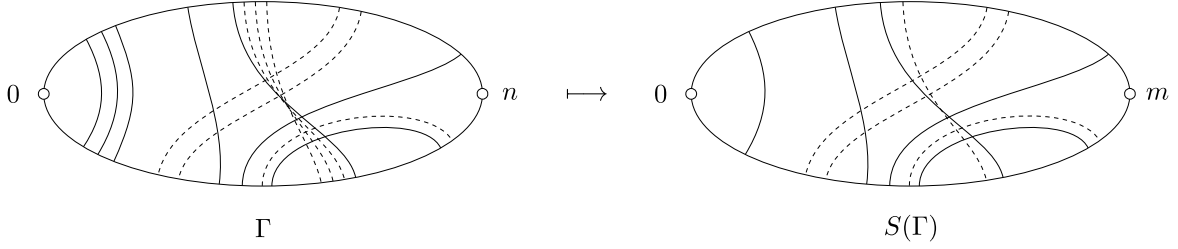


Figure 9.2: The construction of the tagged skeleton graph.

to form a single twisted bridge. This is repeated until no parallel straight bridges or antiparallel twisted bridges remain. The resulting pairing is the skeleton $\Sigma = S(\Gamma)$; see Figure 9.2. Thus we see that Lemma 7.3 holds. Moreover, Lemma 7.4 holds, provided that (i) is replaced with

(i') Each $\Sigma \in \mathcal{S}_{m,m'}^*$ contains no parallel straight bridges and no antiparallel twisted bridges.

Crucially, Lemma 7.7 remains valid for such tagged skeletons. This can be easily seen using the orbit construction of the proof of Lemma 7.7, combined with (i').

Next, we associate a factor $D_\ell(y, z)$ with each bridge $\sigma \in \Sigma$. If σ is straight, this is done exactly as in Section 7.4. If σ is twisted, this association follows immediately from the definition of the antiladder. Thus we find that Lemma 7.6 holds. The rest of the analysis in Section 7 carries over almost verbatim; the only required modification is the summation over 2^m tag configurations of the bridges of Σ . The resulting factor 2^m is immaterial.

Finally, the complete ladder pairing yields (3.1). The complete antiladder is subleading, as its contribution vanishes unless $x = 0$.

10. DELOCALIZATION: PROOFS OF THEOREM 3.3 AND COROLLARY 3.4

In this section we show how to derive Theorem 3.3 from Theorem 3.1, and derive Corollary 3.4 as a consequence.

PROOF OF THEOREM 3.3. We use an argument due to Chen [9] showing that diffusive motion implies delocalization of the vast majority of eigenvectors.

Recall that $P_{x,\ell}(y) := \mathbf{1}(|y - x| \geq \ell)$ is the characteristic function of the complement (in Λ_N) of the ℓ -neighborhood of x . Also, $\mathfrak{A}_{\varepsilon,\ell}^\omega$, defined by

$$\mathfrak{A}_{\varepsilon,\ell}^\omega = \left\{ \alpha \in \mathfrak{A} : \sum_x |\psi_\alpha^\omega(x)| \|P_{x,\ell} \psi_\alpha^\omega\| < \varepsilon \right\},$$

is the set of eigenvectors localized on a scale ℓ up to an error of ε .

By diagonalizing H^ω ,

$$H^\omega = \sum_{\alpha \in \mathfrak{A}} \lambda_\alpha^\omega |\psi_\alpha^\omega\rangle \langle \psi_\alpha^\omega|,$$

we have

$$\begin{aligned} \|P_{x,\ell} e^{-itH^\omega} \delta_x\|^2 &= \left\| \sum_{\alpha \in \mathfrak{A}} P_{x,\ell} e^{-it\lambda_\alpha^\omega} \overline{\psi_\alpha^\omega(x)} \psi_\alpha^\omega \right\|^2 \\ &\leq \left(1 + \frac{1}{\zeta}\right) \left\| \sum_{\alpha \in \mathfrak{A}_{\varepsilon,\ell}^\omega} P_{x,\ell} e^{-it\lambda_\alpha^\omega} \overline{\psi_\alpha^\omega(x)} \psi_\alpha^\omega \right\|^2 + (1 + \zeta) \left\| \sum_{\alpha \in \mathfrak{A} \setminus \mathfrak{A}_{\varepsilon,\ell}^\omega} P_{x,\ell} e^{-it\lambda_\alpha^\omega} \overline{\psi_\alpha^\omega(x)} \psi_\alpha^\omega \right\|^2, \end{aligned}$$

for any $\zeta > 0$. Next, we observe that the norm in the first term may be bounded by 1:

$$\left\| \sum_{\alpha \in \mathfrak{A}_{\varepsilon,\ell}^\omega} P_{x,\ell} e^{-it\lambda_\alpha^\omega} \overline{\psi_\alpha^\omega(x)} \psi_\alpha^\omega \right\|^2 \leq \left\| \sum_{\alpha \in \mathfrak{A}_{\varepsilon,\ell}^\omega} \overline{\psi_\alpha^\omega(x)} \psi_\alpha^\omega \right\|^2 = \sum_{\alpha \in \mathfrak{A}_{\varepsilon,\ell}^\omega} |\psi_\alpha^\omega(x)|^2 \leq \sum_{\alpha \in \mathfrak{A}} |\psi_\alpha^\omega(x)|^2 = 1.$$

Thus we get

$$\begin{aligned} \|P_{x,\ell} e^{-itH^\omega} \delta_x\|^2 &\leq \left(1 + \frac{1}{\zeta}\right) \left\| \sum_{\alpha \in \mathfrak{A}_{\varepsilon,\ell}^\omega} P_{x,\ell} e^{-it\lambda_\alpha^\omega} \overline{\psi_\alpha^\omega(x)} \psi_\alpha^\omega \right\|^2 + (1 + \zeta) \left\| \sum_{\alpha \in \mathfrak{A} \setminus \mathfrak{A}_{\varepsilon,\ell}^\omega} \overline{\psi_\alpha^\omega(x)} \psi_\alpha^\omega \right\|^2 \\ &\leq \left(1 + \frac{1}{\zeta}\right) \sum_{\alpha \in \mathfrak{A}_{\varepsilon,\ell}^\omega} |\psi_\alpha^\omega(x)| \|P_{x,\ell} \psi_\alpha^\omega\| + (1 + \zeta) \sum_{\alpha \in \mathfrak{A} \setminus \mathfrak{A}_{\varepsilon,\ell}^\omega} |\psi_\alpha^\omega(x)|^2. \end{aligned}$$

Averaging over $x \in \Lambda_N$ yields

$$\begin{aligned} \frac{1}{|\mathfrak{A}|} \sum_x \|P_{x,\ell} e^{-itH^\omega} \delta_x\|^2 &\leq \left(1 + \frac{1}{\zeta}\right) \frac{1}{|\mathfrak{A}|} \sum_{\alpha \in \mathfrak{A}_{\varepsilon,\ell}^\omega} \sum_x |\psi_\alpha^\omega(x)| \|P_{x,\ell} \psi_\alpha^\omega\| + (1 + \zeta) \frac{1}{|\mathfrak{A}|} \sum_{\alpha \in \mathfrak{A} \setminus \mathfrak{A}_{\varepsilon,\ell}^\omega} \sum_x |\psi_\alpha^\omega(x)|^2 \\ &\leq \left(1 + \frac{1}{\zeta}\right) \varepsilon + (1 + \zeta) \frac{|\mathfrak{A} \setminus \mathfrak{A}_{\varepsilon,\ell}^\omega|}{|\mathfrak{A}|}, \end{aligned}$$

by definition of $\mathfrak{A}_{\varepsilon,\ell}^\omega$. Therefore

$$\frac{|\mathfrak{A} \setminus \mathfrak{A}_{\varepsilon,\ell}^\omega|}{|\mathfrak{A}|} \geq \frac{1}{1 + \zeta} \frac{1}{|\mathfrak{A}|} \sum_x \|P_{x,\ell} e^{-itH^\omega} \delta_x\|^2 - \frac{\varepsilon}{\zeta}.$$

Taking the expectation yields

$$\mathbb{E} \frac{|\mathfrak{A} \setminus \mathfrak{A}_{\varepsilon,\ell}^\omega|}{|\mathfrak{A}|} \geq \frac{1}{1 + \zeta} \frac{1}{|\mathfrak{A}|} \mathbb{E} \sum_x \|P_{x,\ell} e^{-itH} \delta_x\|^2 - \frac{\varepsilon}{\zeta} = \frac{1}{1 + \zeta} \mathbb{E} \|P_{0,\ell} e^{-itH} \delta_0\|^2 - \frac{\varepsilon}{\zeta}, \quad (10.1)$$

by translation invariance. Note that this estimate holds uniformly in t .

Next, pick a continuous function $\varphi(X)$ that is equal to 0 if $|X| \leq 1$ and 1 if $|X| \geq 2$. Recalling that $\varrho(t, x) = |\langle \delta_x, e^{-itH/2} \delta_0 \rangle|^2$, we find

$$\mathbb{E} \|P_{0,W^{1+d\kappa/2}} e^{-itH/2} \delta_0\|^2 = \sum_x \mathbf{1}(|x| \geq W^{1+d\kappa/2}) \varrho(t, x) \geq \sum_x \varphi\left(\frac{x}{W^{1+d\kappa/2}}\right) \varrho(t, x).$$

Now choose an exponent $\tilde{\kappa}$ satisfying $\kappa < \tilde{\kappa} < 1/3$ and set $t = W^{d\tilde{\kappa}}$. Thus,

$$\mathbb{E} \|P_{0,W^{1+d\kappa/2}} e^{-iW^{d\tilde{\kappa}}H/2} \delta_0\|^2 \geq \sum_x \varphi\left(W^{d/2(\tilde{\kappa}-\kappa)} \frac{x}{W^{1+\tilde{\kappa}d/2}}\right) \varrho(W^{d\tilde{\kappa}}, x).$$

Since we have

$$\lim_{W \rightarrow \infty} \varphi(W^{d/2(\tilde{\kappa}-\kappa)} X) = 1$$

for $X \neq 0$ and $L(1, X)$ is continuous at $X = 0$, a simple limiting argument shows that Theorem 3.1 implies

$$\lim_{W \rightarrow \infty} \sum_x \varphi\left(W^{d/2(\tilde{\kappa}-\kappa)} \frac{x}{W^{1+d\tilde{\kappa}/2}}\right) \varrho(W^{d\tilde{\kappa}}, x) = \int dX L(1, X) = 1.$$

We have hence proved that

$$\lim_{W \rightarrow \infty} \mathbb{E} \|P_{0, W^{1+d\kappa/2}} e^{-iW^{d\tilde{\kappa}} H/2} \delta_0\|^2 = 1,$$

Plugging this into (10.1) yields

$$\liminf_{W \rightarrow \infty} \mathbb{E} \frac{|\mathfrak{A} \setminus \mathfrak{A}_{\varepsilon, W^{1+d\kappa/2}}|}{|\mathfrak{A}|} \geq \frac{1}{1+\zeta} - \frac{\varepsilon}{\zeta}.$$

Setting $\zeta = \sqrt{\varepsilon}$ completes the proof. \square

PROOF OF COROLLARY 3.4. Pick an intermediate exponent $\tilde{\kappa}$ satisfying $\kappa < \tilde{\kappa} < 1/3$ and abbreviate

$$\ell := W^{1+d\kappa/2}, \quad \tilde{\ell} := W^{1+d\tilde{\kappa}/2}.$$

Let $\alpha \in \mathfrak{B}_{\tilde{\ell}}^\omega$ and let $u \in \Lambda_N$ be as in (3.3). Then we find by Cauchy-Schwarz

$$\begin{aligned} \left(\sum_x |\psi_\alpha^\omega(x)| \|P_{x, \tilde{\ell}} \psi_\alpha^\omega\| \right)^2 &\leq \left(\sum_x |\psi_\alpha^\omega(x)|^2 \exp\left[\frac{|x-u|}{\ell}\right]^\gamma \right) \left(\sum_x \exp\left\{-\left[\frac{|x-u|}{\ell}\right]^\gamma\right\} \|P_{x, \tilde{\ell}} \psi_\alpha^\omega\|^2 \right) \\ &\leq K \sum_{|x-y| \geq \tilde{\ell}} \exp\left\{-\left[\frac{|x-u|}{\ell}\right]^\gamma\right\} |\psi_\alpha^\omega(y)|^2 \\ &\leq K e^{-\delta(\tilde{\ell}/\ell)^\gamma} \sum_{|x-y| \geq \tilde{\ell}} \exp\left\{-\left[\frac{|x-u|}{\ell}\right]^\gamma + \delta \left[\frac{|x-y|}{\ell}\right]^\gamma\right\} |\psi_\alpha^\omega(y)|^2, \end{aligned}$$

where $\delta > 0$ is some small constant to be chosen later. Using $(a+b)^\gamma \leq (2a)^\gamma + (2b)^\gamma$ we find

$$\left(\sum_x |\psi_\alpha^\omega(x)| \|P_{x, \tilde{\ell}} \psi_\alpha^\omega\| \right)^2 \leq K e^{-\delta(\tilde{\ell}/\ell)^\gamma} \sum_{x, y} \exp\left\{-\left[\frac{|x-u|}{\ell}\right]^\gamma + \delta \left[2\frac{|x-u|}{\ell}\right]^\gamma + \delta \left[2\frac{|y-u|}{\ell}\right]^\gamma\right\} |\psi_\alpha^\omega(y)|^2.$$

Choosing $\delta < 2^{-\gamma}$ therefore yields

$$\left(\sum_x |\psi_\alpha^\omega(x)| \|P_{x, \tilde{\ell}} \psi_\alpha^\omega\| \right)^2 \leq C \ell^d K^2 e^{-\delta(\tilde{\ell}/\ell)^\gamma} =: \varepsilon_W^2.$$

We have thus proved that $\mathfrak{B}_\ell^\omega \subset \mathfrak{A}_{\varepsilon_W, \tilde{\ell}}^\omega$. Then Corollary 3.4 follows from $\lim_{W \rightarrow \infty} \varepsilon_W = 0$ and Theorem 3.3. \square

11. CRITICAL PAIRINGS

In this section we give an example family of pairings which are critical in the sense that they saturate the 2/3 rule (Lemma 7.7). This implies that extending our results beyond time scales of order $W^{d/3}$ requires either a further resummation of pairings or a more refined classification of graphs in terms of their deviation from the 2/3 rule.

Let $k \geq 1$ and consider the skeleton pairing Σ_k defined in Figure 11.1. It is a critical pairing in the sense that all orbits not containing the vertices $0, m$ consist of 3 vertices.

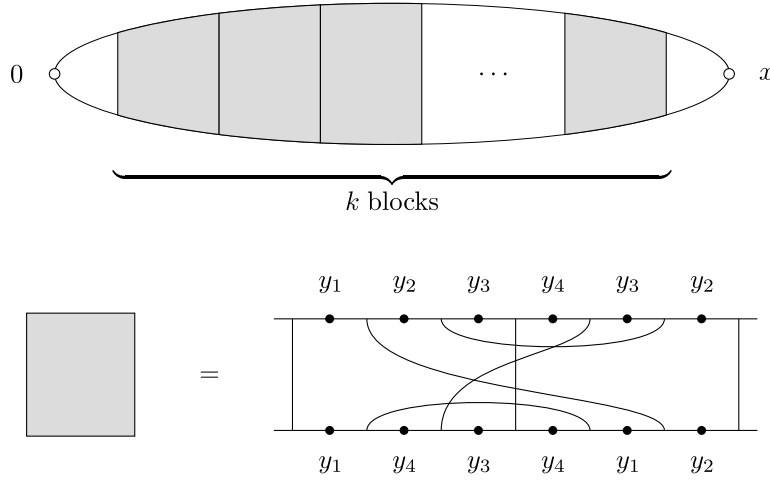


Figure 11.1: A critical skeleton pairing Σ_k . The label of each vertex is indicated next to its vertex.

It is easy to see that for Σ_k we have

$$\bar{m} = 6k + 1, \quad L(\Sigma_k) = 4k + 1.$$

In particular, the 2/3 rule of Lemma 7.7 is saturated. Moreover, if ℓ_{Σ_k} satisfies $\ell_\sigma \geq 2$ for all $\sigma \in \Sigma_k$ then the associated pairing $\Gamma := G_{\ell_{\Sigma_k}}(\Sigma_k)$ has a nonzero contribution $V_x(\Gamma) \sim R_x(\Gamma) \approx M^{-2k}$ (here, and in the following, we ignore any powers of W with exponent of order one). Indeed, it is easy to check that under the condition $\ell_\sigma \geq 2$ for all σ the above Γ satisfies all nonbacktracking conditions. (In fact, it suffices to require that $\ell_{\bar{\sigma}} \geq 2$, where $\bar{\sigma}$ is the bridge drawn as a vertical line in Figure 11.1.)

As shown in the Section 7 (see (7.13)), the coefficients $a_n(t)$ essentially vanish if $n > (1 + o(1))t$. Setting $t = M^\kappa$ thus means restricting the summation to $n, n' \leq M^\kappa$.

Assume, to begin with, that we adopt the strategy of Section 7 in estimating the contribution of each graph, i.e. we use the 2/3 rule for each skeleton pairing and the ℓ^1 - ℓ^∞ -type estimates from Lemma 7.5 on the edges of the associated multigraph. We show that the sum of the contributions of the skeleton pairings Σ_k diverges if $\kappa > 1/3$. Indeed, noting that $n, n' \leq M^\kappa$ implies $\bar{n} \leq M^\kappa$, we find that the contribution of all Σ_k 's is

$$\sum_{k=1}^{p/6} \frac{1}{M^{2k}} \sum_{\ell_1 + \dots + \ell_{6k} = p} 1, \quad (11.1)$$

where $p = M^\kappa$ and the sum over ℓ_i is restricted to $\ell_i \geq 2$ for all i . Here we only sum over the pairing of maximal $\bar{n} = p$ (so as to obtain a lower bound), and set $6k + 1 \approx 6k$. Now (11.1) is equal to

$$\sum_{k=1}^{p/6} \frac{1}{M^{2k}} \binom{p-6k}{6k} \sim \sum_{k=1}^{p/6} \left(\frac{C(M^\kappa - 6k)}{kM^{1/3}} \right)^{6k},$$

which diverges as $W \rightarrow \infty$ if $\kappa > 1/3$. Hence a control of the error term at time scales κ larger than $1/3$ would require further resummation of such critical pairings.

In the estimates of the preceding paragraph we did not make full use of the heat kernel decay associated with each skeleton bridge. For simplicity, the following discussion is restricted to $d = 1$ (it may be easily extended to higher dimensions; in fact some estimates are better in higher dimensions). Using Lemma 7.5, we may improve (11.1) to

$$\sum_{k=1}^{p/6} \frac{1}{M^{2k}} \sum_{\ell_1 + \dots + \ell_{6k} = p} \frac{1}{\sqrt{\ell_1 \dots \ell_{2k}}}; \quad (11.2)$$

this is a simple consequence of the heat kernel bound of Lemma 7.5 and the fact that each six-block of Σ_k contains two bridges in $\Sigma_k \setminus (\Sigma_k)_T$ for which we may apply the bound (7.7b) (in which we drop the unimportant second term for simplicity). Now (11.2) is bounded by

$$\sum_{k=1}^{p/6} \frac{1}{M^{2k}} p^k \binom{p}{4k} \sim \sum_{k=1}^{p/6} \left(\frac{CM^{5\kappa/6}}{k^{2/3}M^{1/3}} \right)^{6k}, \quad (11.3)$$

which is summable for $\kappa < 2/5$. Note, however, that the factor k^{-6k} from (11.1) has been replaced with the larger factor k^{-4k} . Recall that the factor k^{-6k} is used to cancel the combinatorics $\bar{m}! \sim k^{6k}$ arising from the summation over all skeletons. In the present example this small factor is not needed, as the family $\{\Sigma_k\}$ is small. It is clear, however, that a systematic application of this approach requires a more refined classification of skeletons in terms of how much they deviate from the $2/3$ rule. One expects that the number of skeletons saturating the $2/3$ rule is small, and that they are therefore amenable to estimates of type (11.3). Conversely, most of the $\bar{m}!$ skeletons are expected to deviate strongly from the $2/3$ rule, so that their greater number is compensated by their small individual contributions.

Finally, we mention that the upper bound (7.7), used in the ℓ^1 - ℓ^∞ -type estimates above, neglects the spatial decay of the heat kernel, i.e. that

$$D_\ell(x, y) \sim \ell^{-1/2} e^{-(x-y)^2/\ell} \quad (11.4)$$

for $|x - y| \ll N$. Thus a correct *lower* bound on the contribution of each skeleton graph should have taken into account this additional decay as well. A somewhat lengthier calculation shows that with the asymptotics (11.4) the estimate (11.2) may be improved to

$$\sum_{k=1}^{p/6} \frac{1}{M^{2k}} \sum_{\ell_1 + \dots + \ell_{6k} = p} \prod_{j=1}^k \frac{1}{\sqrt{\ell_2^{(j)} \ell_4^{(j)} + \ell_2^{(j)} \ell_6^{(j)} + \ell_3^{(j)} \ell_4^{(j)} + \ell_3^{(j)} \ell_6^{(j)} + \ell_4^{(j)} \ell_5^{(j)} + \ell_4^{(j)} \ell_6^{(j)} + \ell_5^{(j)} \ell_6^{(j)}}},$$

where we abbreviated $\ell_i^{(j)} := \ell_{6(j-1)+i}$. It is not hard to see that the resulting bound is the same as (11.3), with a smaller constant C . In other words, the gain obtained from the spatial decay of the heat kernel is immaterial, and the ℓ^1 - ℓ^∞ -estimates cannot be improved.

In conclusion: Our estimates rely on an indiscriminate application of the 2/3 rule to all skeleton pairings; going beyond time scales of order $W^{d/3}$ would require either (i) a refined classification of the skeleton pairings in terms of how much they deviate from the 2/3 rule, combined with a systematic use of the bound (7.7b) on all bridges in $\Sigma \setminus \Sigma_T$; or (ii) a further resummation of graphs in order to exploit cancellations. The approach (i) can be expected to reach at most times of order $W^{2/5}$ for $d = 1$.

A. PROOF OF PROPOSITION 8.5

Note first that F is monotone nondecreasing and satisfies $0 \leq F(\lambda) \leq 1$, as follows from (5.4). Hence it is enough to prove (8.5a) for $\lambda \in (0, 1)$.

For the following it is convenient to replace F_t with \tilde{F}_t , defined by

$$\tilde{F}_t(\lambda) := \sum_{n=0}^{\lfloor \lambda t \rfloor} |\alpha_n(t)|^2 = F_t\left(\frac{\lfloor t\lambda + 1 \rfloor}{t}\right).$$

By Lemma 8.8 we have $F_t(\lambda) - \tilde{F}_t(\lambda) = o(1)$ as $t \rightarrow \infty$.

Thus let $\lambda \in (0, 1)$ be fixed. From (5.2) we find

$$\tilde{F}_t(\lambda) = \sum_{n=0}^{\lfloor \lambda t \rfloor} \left| \frac{2}{\pi} \int_{-1}^1 d\xi \sqrt{1 - \xi^2} U_n(\xi) e^{-it\xi} \right|^2 = \sum_{n=0}^{\lfloor \lambda t \rfloor} \left| \frac{2}{\pi} \int_0^\pi d\theta \sin \theta \sin[(n+1)\theta] e^{-it \cos \theta} \right|^2,$$

where we used (5.1). Thus,

$$\begin{aligned} \tilde{F}_t(\lambda) &= \frac{1}{\pi^2} \int_0^\pi d\theta \int_0^\pi d\theta' \sin \theta \sin \theta' e^{it(\cos \theta - \cos \theta')} \\ &\quad \times \left[\frac{e^{i(\lfloor \lambda t \rfloor + 1)(\theta + \theta')} - 1}{e^{-i(\theta + \theta')} - 1} + \frac{e^{-i(\lfloor \lambda t \rfloor + 1)(\theta + \theta')} - 1}{e^{i(\theta + \theta')} - 1} - \frac{e^{i(\lfloor \lambda t \rfloor + 1)(\theta - \theta')} - 1}{e^{-i(\theta - \theta')} - 1} - \frac{e^{-i(\lfloor \lambda t \rfloor + 1)(\theta - \theta')} - 1}{e^{i(\theta - \theta')} - 1} \right], \end{aligned} \quad (\text{A.1})$$

We now claim that the limit $t \rightarrow \infty$ of the first two terms of (A.1) vanishes by a stationary phase argument. Let us write the first term of (A.1) as $R_t^1 + R_t^2$, where

$$\begin{aligned} R_t^1 &:= \frac{1}{\pi^2} \int_0^\pi d\theta \int_0^\pi d\theta' \sin \theta \sin \theta' e^{it(\cos \theta - \cos \theta')} \frac{e^{i(\lfloor \lambda t \rfloor + 1)(\theta + \theta')}}{e^{-i(\theta + \theta')} - 1} \\ &= \frac{1}{\pi^2} \int_0^\pi d\theta \int_0^\pi d\theta' \frac{\sin \theta \sin \theta' e^{i(1 - \{\lambda t\})(\theta + \theta')}}{e^{-i(\theta + \theta')} - 1} e^{it(\cos \theta - \cos \theta' + \lambda\theta + \lambda\theta')} \\ &=: \int_0^\pi d\theta \int_0^\pi d\theta' a_t(\theta, \theta') e^{it\phi(\theta, \theta')}, \end{aligned}$$

where $\{\xi\} := \xi - [\xi] \in [0, 1)$. One readily finds the bounds

$$\inf_{\theta, \theta' \in [0, \pi]} |\nabla \phi(\theta, \theta')| \geq \lambda, \quad \sup_{\theta, \theta' \in [0, \pi]} |\nabla^2 \phi(\theta, \theta')| < \infty, \quad \sup_t \sup_{\theta, \theta' \in [0, \pi]} |\nabla a_t(\theta, \theta')| < \infty.$$

A standard stationary phase argument therefore yields $\lim_{t \rightarrow \infty} R_t^1 = 0$.

Similarly, we find

$$R_t^2 = -\frac{1}{\pi^2} \int_0^\pi d\theta \int_0^\pi d\theta' e^{it(\cos\theta - \cos\theta')} \underbrace{\frac{\sin\theta \sin\theta'}{e^{-i(\theta+\theta')} - 1}}_{=:b(\theta,\theta')}$$

As above, the functions b and ∇b are bounded on $[0, \pi]^2$. The phase $\cos\theta - \cos\theta'$ has four stationary points, $(0, 0)$, $(0, \pi)$, $(\pi, 0)$, (π, π) , all of them nondegenerate. Therefore a standard stationary phase argument implies that $R_t^2 = O(t^{-1/2})$. (Note that the stationary points lie on the boundary of the integration domain. This is not a problem, however, as the usual stationary phase argument may be applied in combination with the identity $\int_0^\infty dx e^{ix^2} = O(t^{-1/2})$.) Similarly, one shows that the second term of (A.1) vanishes as $t \rightarrow \infty$.

Next, as we have just shown, we have

$$\tilde{F}_t(\lambda) = \tilde{F}_t^0(\lambda) + \tilde{F}_t^+(\lambda) + \tilde{F}_t^-(\lambda) + o(1)$$

for $t \rightarrow \infty$, where

$$\begin{aligned} \tilde{F}_t^0(\lambda) &:= \frac{1}{\pi^2} \int_0^\pi d\theta \int_0^\pi d\theta' \sin\theta \sin\theta' e^{it(\cos\theta - \cos\theta')} \left[\frac{1}{e^{-i(\theta-\theta')} - 1} + \frac{1}{e^{i(\theta-\theta')} - 1} \right], \\ \tilde{F}_t^\pm(\lambda) &:= -\frac{1}{\pi^2} \int_0^\pi d\theta \mathcal{P} \int_0^\pi d\theta' \sin\theta \sin\theta' e^{it(\cos\theta - \cos\theta')} \frac{e^{\pm i([\lambda t] + 1)(\theta - \theta')}}{e^{\mp i(\theta - \theta')} - 1}, \end{aligned}$$

where \mathcal{P} denotes principal value. We now show that $\tilde{F}_t^0(\lambda) = o(1)$. Indeed, the expression in square brackets in the definition of $\tilde{F}_t^0(\lambda)$ is equal to -1 . Exactly as above we therefore conclude that $\tilde{F}_t^0(\lambda) = O(t^{-1/2})$.

Next, let us consider $\tilde{F}_t^+(\lambda)$. In a first step, we replace the factor $\frac{1}{e^{-i(\theta-\theta')} - 1}$ with $\frac{1}{-i(\theta-\theta')}$. The error is

$$-\frac{1}{\pi^2} \int_0^\pi d\theta \mathcal{P} \int_0^\pi d\theta' \sin\theta \sin\theta' e^{it(\cos\theta - \cos\theta')} e^{i([\lambda t] + 1)(\theta - \theta')} \left[\frac{1}{e^{-i(\theta-\theta')} - 1} - \frac{1}{-i(\theta-\theta')} \right],$$

which vanishes in the limit $t \rightarrow \infty$ by the above saddle point argument (the expression in the square brackets is analytic, and the phase $\cos\theta - \cos\theta' + \lambda\theta - \lambda\theta'$ has the four nondegenerate saddle points defined by $\sin\theta = \sin\theta' = \lambda$).

In a second step, we choose a scale $2/5 < \varepsilon < 1/2$ and introduce a cutoff in $|\theta - \theta'|$ at $t^{-\varepsilon}$. Thus we have

$$\begin{aligned} \tilde{F}_t^+(\lambda) + o(1) &= \frac{1}{\pi^2} \int_0^\pi d\theta \mathcal{P} \int_0^\pi d\theta' \mathbf{1}(|\theta - \theta'| \leq t^{-\varepsilon}) \sin\theta \sin\theta' e^{it(\cos\theta - \cos\theta')} \frac{e^{i([\lambda t] + 1)(\theta - \theta')}}{i(\theta - \theta')}, \\ &\quad + \frac{1}{\pi^2} \int_0^\pi d\theta \int_0^\pi d\theta' \mathbf{1}(|\theta - \theta'| > t^{-\varepsilon}) \sin\theta \sin\theta' e^{it(\cos\theta - \cos\theta')} \frac{e^{i([\lambda t] + 1)(\theta - \theta')}}{i(\theta - \theta')}, \end{aligned} \quad (\text{A.2})$$

Let us abbreviate $D_t := \{(\theta, \theta') \in [0, \pi]^2 : |\theta - \theta'| > t^{-\varepsilon}\}$. The second term on the right-hand side of (A.2) is equal to

$$\frac{1}{\pi^2} \int_{D_t} d\theta d\theta' e^{it(\cos\theta - \cos\theta' + \lambda\theta - \lambda\theta')} \frac{\sin\theta \sin\theta' e^{i(1 - \{\lambda t\})(\theta - \theta')}}{i(\theta - \theta')} =: \int_{D_t} d\theta d\theta' e^{it\phi(\theta, \theta')} \frac{a_t(\theta, \theta')}{\theta - \theta'}. \quad (\text{A.3})$$

In the domain D_t the phase ϕ has two stationary points defined by $\sin\theta = \sin\theta' = \lambda$ and $\theta \neq \theta'$. For all (θ, θ') not in some fixed neighbourhood of these stationary points and satisfying $|\theta - \theta'| > t^{-\varepsilon}$, we have the bound

$$|\nabla\phi(\theta, \theta)| \geq Ct^{-\varepsilon},$$

for some constant $C > 0$ depending on λ , and large enough t . Thus a standard saddle point analysis shows that (A.3) is of the order $t^{-1/2} + t^{2\varepsilon-1} = o(1)$.

In a third step, we analyse the first term on the right-hand side of (A.2). We introduce the new coordinates

$$u = \frac{\theta + \theta'}{2}, \quad v = \theta - \theta',$$

and write

$$\begin{aligned} \tilde{F}_t^+(\lambda) + o(1) &= \frac{1}{\pi^2} \int_0^\pi d\theta \mathcal{P} \int_0^\pi d\theta' \mathbf{1}(|\theta - \theta'| \leq t^{-\varepsilon}) \sin \theta \sin \theta' e^{it(\cos \theta - \cos \theta')} \frac{e^{i([\lambda t] + 1)(\theta - \theta')}}{i(\theta - \theta')} \\ &= \frac{1}{\pi^2} \int_0^\pi du \mathcal{P} \int_{-a_{t,u}}^{a_{t,u}} dv \sin\left(u + \frac{v}{2}\right) \sin\left(u - \frac{v}{2}\right) e^{i(1 - \{\lambda t\})v} \frac{e^{it(\lambda v - 2 \sin u \sin \frac{v}{2})}}{iv}, \end{aligned}$$

where

$$a_{t,u} := \min\{t^{-\varepsilon}, 2u, 2(\pi - u)\}.$$

Now we replace the factor $e^{it(\lambda v - 2 \sin u \sin \frac{v}{2})}$ with $e^{itv(\lambda - \sin u)}$. The resulting error is

$$R_t := \frac{1}{\pi^2} \int_0^\pi du \mathcal{P} \int_{-a_{t,u}}^{a_{t,u}} dv \sin\left(u + \frac{v}{2}\right) \sin\left(u - \frac{v}{2}\right) e^{i(1 - \{\lambda t\})v} e^{itv(\lambda - \sin u)} \frac{e^{it \sin u (v - 2 \sin \frac{v}{2})} - 1}{iv}. \quad (\text{A.4})$$

It is easy to check that, for $v \in [-a_{t,u}, a_{t,u}]$, we have

$$\left| \frac{e^{it \sin u (v - 2 \sin \frac{v}{2})} - 1}{iv} \right| \leq C t^{1 - \frac{5}{2}\varepsilon} \frac{1}{\sqrt{|v|}}.$$

Therefore

$$|R_t| \leq C \int_0^\pi du \int_0^{2\pi} dv t^{1 - \frac{5}{2}\varepsilon} \frac{1}{\sqrt{|v|}} = o(1).$$

Thus we may write

$$\tilde{F}_t^+(\lambda) + o(1) = \frac{1}{\pi^2} \int_0^\pi du \mathcal{P} \underbrace{\int_{-a_{t,u}}^{a_{t,u}} dv \sin\left(u + \frac{v}{2}\right) \sin\left(u - \frac{v}{2}\right) e^{i(1 - \{\lambda t\})v} \frac{e^{itv(\lambda - \sin u)}}{iv}}_{=: I_t(u)}.$$

In a fourth step, we analyse $I_t(u)$ using contour integration. Abbreviate $b := \lambda - \sin u$. Let us assume that u satisfies $b \neq 0$. Then, setting $z = |b|tv$, we find

$$I_t(u) = \mathcal{P} \int_{-|b|ta_{t,u}}^{|b|ta_{t,u}} dz \sin\left(u + \frac{z}{2|b|t}\right) \sin\left(u - \frac{z}{2|b|t}\right) e^{i(1 - \{\lambda t\})\frac{z}{|b|t}} \frac{e^{iz \operatorname{sgn} b}}{iz}.$$

Let us consider the case $b > 0$. Using the identity

$$\mathcal{P} \frac{1}{v} = i\pi \delta(v) + \frac{1}{v - i0}$$

and Cauchy's theorem, we find

$$I_t(u) = \pi \sin^2(u) + \int_\gamma dz \sin\left(u + \frac{z}{2bt}\right) \sin\left(u - \frac{z}{2bt}\right) e^{i(1 - \{\lambda t\})\frac{z}{bt}} \frac{e^{iz}}{iz},$$

where γ is the arc $\{bta_{t,u}(\cos \varphi, \sin \varphi) : \varphi \in [0, \pi]\}$. The absolute value of the integral is bounded by

$$\int_0^\pi d\varphi e^{a_{t,u} \sin \varphi} e^{-bta_{t,u} \sin \varphi},$$

which is bounded uniformly in t and $b \neq 0$, and vanishes in the limit $t \rightarrow \infty$ for all $b \neq 0$. The case $b < 0$ is treated in the same way. In summary, we have, for each u satisfying $\sin u \neq \lambda$, that

$$|I_t(u)| \leq C, \quad \lim_{t \rightarrow \infty} I_t(u) = \pi \sin^2(u) \operatorname{sgn}(\lambda - \sin u).$$

Hence by dominated convergence we get

$$\lim_{t \rightarrow \infty} \tilde{F}_t^+(\lambda) = \frac{1}{\pi} \int_0^\pi du \sin^2(u) \operatorname{sgn}(\lambda - \sin u).$$

A similar (in fact easier) analysis yields

$$\lim_{t \rightarrow \infty} \tilde{F}_t^-(\lambda) = \frac{1}{\pi} \int_0^\pi du \sin^2(u) \operatorname{sgn}(\lambda + \sin u).$$

Therefore we get

$$\lim_{t \rightarrow \infty} \tilde{F}_t(\lambda) = \frac{2}{\pi} \int_0^\pi du \sin^2(u) \mathbf{1}(\sin u < \lambda) = \frac{4}{\pi} \int_0^\lambda d\xi \frac{\xi^2}{\sqrt{1 - \xi^2}}.$$

This completes the proof of Proposition 8.5.

REFERENCES

- [1] Abrahams, E., Anderson, P.W., Licciardello, D.C., Ramakrishnan, T.V.: *Scaling theory of localization: Absence of quantum diffusion in two dimensions*. Phys. Rev. Lett. **42**, 673–676.
- [2] Anderson, G.; Zeitouni, O.: *A CLT for a band matrix model*. Probab. Theory Related Fields **134** (2006), no. 2, 283–338.
- [3] Aizenman, M. and Molchanov, S.: *Localization at large disorder and at extreme energies: an elementary derivation*, Commun. Math. Phys. **157**, 245–278 (1993).
- [4] Aizenman, M., Sims, R., Warzel, S.: *Absolutely continuous spectra of quantum tree graphs with weak disorder*. Commun. Math. Phys. **264** no. 2, 371–389 (2006).
- [5] Anderson, P.: *Absences of diffusion in certain random lattices*, Phys. Rev. **109**, 1492–1505 (1958).
- [6] Bachmann, S.; De Roeck, W.: *From the Anderson model on a strip to the DMPK equation and random matrix theory*, J. Stat. Phys. **139**, 541–564 (2010).
- [7] Bai, Z.D., Yin, Y.Q.: *Limit of the smallest eigenvalue of a large dimensional sample covariance matrix*. Ann. Probab. **21** (1993), no. 3, 1275–1294.

- [8] Bourgain, J.: *Random lattice Schrödinger operators with decaying potential: some higher dimensional phenomena*. Lecture Notes in Mathematics, Vol. 1807, 70–99 (2003).
- [9] Chen, T.: *Localization lengths and Boltzmann limit for the Anderson model at small disorders in dimension 3*. J. Stat. Phys. **120** (2005), no. 1–2, 279–337.
- [10] Denisov, S.A.: *Absolutely continuous spectrum of multidimensional Schrödinger operator*. Int. Math. Res. Not. **2004** no. 74, 3963–3982.
- [11] Disertori, M., Pinson, H., Spencer, T.: *Density of states for random band matrices*. Commun. Math. Phys. **232**, 83–124 (2002).
- [12] Disertori, M., Spencer, T.: *Anderson localization for a supersymmetric sigma model*. Preprint arXiv:0910.3325.
- [13] Disertori, M., Spencer, T., Zirnbauer, M.: *Quasi-diffusion in a 3D Supersymmetric Hyperbolic Sigma Model*. Preprint arXiv:0901.1652.
- [14] Efetov, K.B.; *Supersymmetry in Disorder and Chaos*, Cambridge University Press, Cambridge, 1997.
- [15] Elgart, A.: *Lifshitz tails and localization in the three-dimensional Anderson model*. Duke Math. J. **146** (2009), no. 2, 331–360.
- [16] Erdős, L., Knowles, A.: *Quantum diffusion and delocalization for band matrices with general distribution*. Preprint arXiv:1005.1838.
- [17] Erdős, L., Péché, G., Ramírez, J., Schlein, B., Yau, H.-T.: *Bulk universality for Wigner matrices*. Comm. Pure Appl. Math. **63** (2010), no. 7, 895–925.
- [18] Erdős, L., Salmhofer, M., Yau, H.-T.: *Quantum diffusion of the random Schrödinger evolution in the scaling limit*. Acta Math. **200**, no. 2, 211–277 (2008).
- [19] Erdős, L., Salmhofer, M., Yau, H.-T.: *Quantum diffusion of the random Schrödinger evolution in the scaling limit II. The recollision diagrams*. Commun. Math. Phys. **271**, 1–53 (2007).
- [20] Erdős, L., Salmhofer, M., Yau, H.-T.: *Quantum diffusion for the Anderson model in scaling limit*. Ann. Inst. H. Poincaré **8** no. 4, 621–685 (2007).
- [21] Erdős, L., Schlein, B., Yau, H.-T.: *Local semicircle law and complete delocalization for Wigner random matrices*. Comm. Math. Phys. **287**, 641–655 (2009).
- [22] Erdős, L., Schlein, B., Yau, H.-T.: *Universality of Random Matrices and Local Relaxation Flow*. Preprint arXiv:0907.5605.
- [23] Erdős, L., Schlein, B., Yau, H.-T., Yin, J.: *The local relaxation flow approach to universality of the local statistics for random matrices*. Preprint arXiv:0911.3687.
- [24] Erdős, L. and Yau, H.-T.: *Linear Boltzmann equation as the weak coupling limit of the random Schrödinger equation*. Commun. Pure Appl. Math. **LIII**, 667–735, (2000).
- [25] Erdős, L., Yau, H.-T., Yin, J.: *Bulk universality for generalized Wigner matrices*. Preprint arXiv:1001.3453.

- [26] Feldheim, O. and Sodin, S.: *A universality result for the smallest eigenvalues of certain sample covariance matrices*. Preprint arXiv:0812.1961.
- [27] Froese, R., Hasler, D., Spitzer, W.: *Transfer matrices, hyperbolic geometry and absolutely continuous spectrum for some discrete Schrödinger operators on graphs*. J. Funct. Anal. **230** no. 1, 184–221 (2006).
- [28] Fröhlich, J. and de Roeck, W.: *Diffusion of a massive quantum particle coupled to a quasi-free thermal medium in dimension $d \geq 4$* . Preprint arXiv:0906.5178.
- [29] Fröhlich, J. and Spencer, T.: *Absence of diffusion in the Anderson tight binding model for large disorder or low energy*, Commun. Math. Phys. **88**, 151–184 (1983).
- [30] Fröhlich, J., Martinelli, F., Scoppola, E., Spencer, T.: *Constructive proof of localization in the Anderson tight binding model*. Commun. Math. Phys. **101** no. 1, 21–46 (1985).
- [31] Fyodorov, Y.V. and Mirlin, A.D.: *Scaling properties of localization in random band matrices: A σ -model approach*. Phys. Rev. Lett. **67** 2405–2409 (1991).
- [32] Gradshteyn, I.S. and Ryzhik, I.M., *Tables of integrals, series, and products*. Academic Press, New York, 2007.
- [33] Guionnet, A.: *Large deviation upper bounds and central limit theorems for band matrices*. Ann. Inst. H. Poincaré Probab. Statist **38** , 341–384 (2002).
- [34] Kirsch, W., Krishna, M., Obermeit, J.: *Anderson model with decaying randomness: Mobility edge*. Math. Z. **235**, 421–433 (2000).
- [35] Klein, A.: *Absolutely continuous spectrum in the Anderson model on the Bethe lattice*. Math. Res. Lett. **1**, 399–407 (1994).
- [36] Krasikov, I., *Uniform bounds for Bessel functions*. J. Appl. Anal. **12**, no. 1, 83–91 (2006).
- [37] Mehta, M.L.: *Random Matrices*. Academic Press, New York, 1991.
- [38] Minami, N.: *Local fluctuation of the spectrum of a multidimensional Anderson tight binding model*. Comm. Math. Phys. **177**, no. 3, 709–725 (1996).
- [39] Rodnianski, I., Schlag, W.: *Classical and quantum scattering for a class of long range random potentials*. Int. Math. Res. Not. **5**, 243–300 (2003).
- [40] Schenker, J.: *Eigenvector localization for random band matrices with power law band width*. Commun. Math. Phys. **290**, 1065–1097 (2009).
- [41] Schlag, W., Shubin, C., Wolff, T.: *Frequency concentration and location lengths for the Anderson model at small disorders*. J. Anal. Math. **88**, 173–220 (2002).
- [42] Sodin, S.: *The spectral edge of some random band matrices*. Preprint arXiv:0906.4047.
- [43] Spencer, T.: *Lifshitz tails and localization*. Preprint (1993).
- [44] Spencer, T.: *Random banded and sparse matrices (Chapter 23)* to appear in “Oxford Handbook of Random Matrix Theory”, edited by G. Akemann, J. Baik and P. Di Francesco.

- [45] Spohn, H.: *Derivation of the transport equation for electrons moving through random impurities*. J. Statist. Phys. **17**, no. 6., 385–412 (1977).
- [46] Tao, T. and Vu, V.: *Random matrices: Universality of the local eigenvalue statistics*. Preprint arXiv:0906.0510.
- [47] Valkó, B., Virág, B.: *Random Schrödinger operators on long boxes, noise explosion and the GOE*. Preprint arXiv:0912.0097.
- [48] Wigner, E.: *Characteristic vectors of bordered matrices with infinite dimensions*. Ann. of Math. **62**, 548–564 (1955).