Acta Math., 200 (2008), 211–277
DOI: 10.1007/s11511-008-0027-2

Quantum diffusion of the random Schrödinger evolution in the scaling limit

by

László Erdős
University of Munich
Munich, Germany

Manfred Salmhofer
Max Planck Institute
Leipzig, Germany

Horng-Tzer Yau
Harvard University
Cambridge, MA, U.S.A.

1. Introduction

The fundamental equations governing the basic laws of physics, the Newton and the Schrödinger equations, are time reversible and have no dissipation. It is remarkable that dissipation is nevertheless ubiquitous in nature, so that almost all macroscopic phenomenological equations are dissipative. The oldest such example is perhaps the equation of heat conductance found by Fourier.

On a microscopic level, Brown observed almost two centuries ago that the motion of a pollen suspended in water was erratic [6], [7]. This led to the kinetic explanation by Einstein in 1905 [15] that such a motion was created by the constant "kicks" on the relatively heavy pollen by the light water molecules. It should be noted that at that time even the atomic-molecular structure of matter was not universally accepted. Einstein's theory was strongly supported by Boltzmann's kinetic theory, which, however, was phenomenological and seriously debated at the time. Finally in 1908 Perrin [43] experimentally verified Einstein's theory and used it, among others, to give a precise estimate on the Avogadro number. These experiments gave the strongest evidence for atoms and molecules at that time.

In Einstein's kinetic theory both the heavy particle (the pollen) and the light particles (the water molecules) obey Newton's law. Therefore, Einstein's kinetic theory in

The first auhor was partially supported by NSF grant DMS-0307295 and MacArthur Fellowship. The third author was partially supported by NSF grant DMS-0200235 and EU-IHP Network "Analysis and Quantum" HPRN-CT-2002-0027.

fact postulated the emergence of the Brownian motion from a classical non-dissipative reversible dynamics. The key difficulty of a mathematically rigorous derivation of Brownian motion from reversible dynamics is similar to the justification of Boltzmann's molecular chaos assumption (Stoßzahlansatz); the dissipative character emerges only in a scaling limit, as the number of degrees of freedom goes to infinity.

The first mathematical definition of the Brownian motion was given in 1923 by Wiener, who constructed the Brownian motion as a scaling limit of random walks. This construction was built upon a stochastic microscopic dynamics which by itself is dissipative. The derivation of the Brownian motion from a time-reversible Hamiltonian system, however, was not seriously considered until more than half a century later. Kesten and Papanicolaou [36] in 1978 proved that the velocity distribution of a particle moving in an environment consisting of random scatterers (i.e., Lorentz gas with random scatterers) converges to a Brownian motion in a weak coupling limit for $d \geqslant 3$. In this model the bath of light particles is replaced with random static impurities. The same result was obtained later in $d=2$ dimensions by Dürr, Goldstein and Lebowitz [14]. In a recent work [39], Komorowski and Ryzhik have controlled the same evolution on a longer time scale and proved the convergence to Brownian motion of the position process as well. Bunimovich and Sinai [10] proved the convergence of the periodic Lorentz gas with a hard core interaction to a Brownian motion in 1980. In this model the only source of randomness is the distribution of the initial condition. Finally, Dürr, Goldstein and Lebowitz [13] proved that the velocity process of a heavy particle in a light ideal gas, which is a model with a dynamical environment, converges to the Ornstein–Uhlenbeck process.

Wiener's construction of Brownian motion is based on a random walk. The random walk could easily be replaced by the Markovian process generated by a linear Boltzmann equation. The linear Boltzmann equation was rigorously derived from the classical Lorentz gas by Gallavotti [31], Spohn [49] and Boldrighini, Bunimovich and Sinai [4]. (The non-linear Boltzmann equation was derived by Lanford [40] for short time.) Although Brownian motion was discovered and theorized in the context of classical dynamics, we shall prove that it also describes the motion of a quantum particle in a random environment.

The *random Schrödinger equation*, or the *quantum Lorentz model*, is given by the evolution equation:

$$i\partial_t \psi_t(x) = H\psi_t(x), \quad H = H_\omega = -\tfrac{1}{2}\Delta_x + \lambda V_\omega(x), \tag{1.1}$$

where λ is the coupling constant and V_ω is the random potential. The first scale with a non-trivial limiting dynamics is the weak coupling limit, $\lambda \to 0$, where the space, time

and the coupling constant are subject to the kinetic scaling:

$$t \longmapsto t\varepsilon^{-1}, \quad x \longmapsto x\varepsilon^{-1}, \quad \lambda = \sqrt{\varepsilon}. \tag{1.2}$$

Under this limit, the appropriately rescaled Wigner distribution (see (2.11)) of the solution to the Schrödinger evolution (1.1) converges weakly to a linear Boltzmann equation. This was first established by Spohn [48] for a random potential with Gaussian distribution and small macroscopic time. This method was extended to study higher-order correlations in [35]. A different method (applicable to the lattice setting and general random potentials, see remarks later on) was developed in [20] where the short time restriction was removed. This method was also extended to the phonon case [16] and Lukkarinen and Spohn [42] have employed a similar technique for studying the energy transport in a harmonic crystal with weakly perturbed random masses.

Since the long time limit of a Boltzmann equation is a heat equation, we shall take a time scale longer (see (2.17)) than in the kinetic scaling limit (1.2). Our aim is to prove that the limiting dynamics of the Schrödinger evolution in a random potential under this scaling is governed by a heat equation. This requires to control the Schrödinger dynamics up to a time scale $\lambda^{-2-\varkappa}$, $\varkappa>0$. Quantum correlations that are small on the kinetic scale and are neglected in the first limit may contribute on the longer time scale. The derivation of the heat equation is thus much more difficult than first deriving the Boltzmann equation from Schrödinger dynamics on the kinetic scale and then showing that the Boltzmann equation converges to a diffusive equation under a different limiting procedure. Notice that the limit in our approach is a long time scaling limit which involves *no semiclassical limit.*

The approach of this paper also applies to lattice models and yields a derivation of Brownian motion from the Anderson model [17], [18] and in fact, we present our main technical steps in a unified framework. The dynamics of the Anderson model was postulated by Anderson [3] to be localized for large coupling constants λ and extended for small coupling constants (away from the band edges and in dimension $d\geqslant 3$). The localization conjecture was first established rigorously by Goldsheid, Molchanov and Pastur [34] in one dimension, by Fröhlich–Spencer [30], and later by Aizenman–Molchanov [1] in several dimensions, and many other works have since contributed to this field. The progress for the extended state conjecture, however, has been limited. It was proved by Klein [38] that all eigenfuctions are extended on the Bethe lattice (see also [2] and [29]). In Euclidean space, Schlag, Shubin and Wolff [46] proved that the eigenfunctions cannot be localized in a region smaller than $\lambda^{-2+\delta}$ for some $\delta>0$ in $d=2$. Chen [11], partly based on the method [20], extended this result to all dimensions $d\geqslant 2$ and $\delta=0$ (with logarithmic corrections). Extended states for the Schrödinger equation with a sufficiently

decaying random potential were proven by Rodnianski and Schlag [44] and Bourgain [5] (see also [12]).

In summary, all known results for the Anderson model (or its modifications) in Euclidean space are in regions where the dynamics have either no effective collision or there are typically only finitely many of them. Under the diffusive scaling of this paper, see (2.17), the number of effective scatterings is a negative fractional power of the scaling parameter. In particular, it goes to infinity in the scaling limit, as should be the case if we aim to obtain a Brownian motion. As in [11], one may derive from our dynamical result that the eigenfunctions cannot be localized in a region smaller than $\lambda^{-2-\varkappa/2}$ in dimension $d\geqslant 3$.

Acknowledgement. The authors are grateful for the financial support and hospitality of the Erwin Schrödinger Institut, Vienna, Max Planck Institut, Leipzig, Stanford University and Harvard University, where part of this work has been done. We thank Jani Lukkarinen for pointing out a simplification of Lemma 6.1 for the continuum model and we also thank the referee for his comments and suggestions.

2. Statement of the main result

2.1. Notation

Let

$$H := -\frac{\hbar^2}{2m}\Delta + \lambda V \tag{2.1}$$

denote a random Schrödinger operator acting on $L^2(\mathbf{R}^d)$, $d\geqslant 3$, with a random potential $V=V_\omega(x)$ and a small positive coupling constant λ. We shall choose the units so that $h=1$ and $m=1$, and thus $\hbar^2/2m=[2(2\pi)^2]^{-1}$. The potential is given by

$$V_\omega(x) := \int_{\mathbf{R}^d} B(x-y)\, d\mu_\omega(y), \tag{2.2}$$

where B is a single site potential profile and μ_ω is a Poisson point process on $\mathbf{R}^d$ with homogeneous unit density and with independent, identically distributed (i.i.d.) random masses. More precisely, for almost all realizations ω, it consists of a countable, locally finite collection of points $\{y_\gamma(\omega)\in\mathbf{R}^d : \gamma=1,2,\dots\}$, and random weights $\{v_\gamma(\omega)\in\mathbf{R} : \gamma=1,2,\dots\}$ such that the random measure is given by

$$\mu_\omega = \sum_{\gamma=1}^{\infty} v_\gamma(\omega)\delta_{y_\gamma(\omega)}, \tag{2.3}$$

where δ_y denotes the Dirac mass at $y \in \mathbf{R}^d$. The Poisson process $\{y_\gamma(\omega)\}$ is independent of the weights $\{v_\gamma(\omega)\}$. The weights are real i.i.d. random variables with distribution $\mathbf{P}_v$ and with moments $m_k := \mathbf{E}_v v_\gamma^k$ satisfying

$$m_2 = 1, \quad m_{2d} < \infty \quad \text{and} \quad m_1 = m_3 = m_5 = 0. \tag{2.4}$$

The expectation with respect to the random process $\{y_\gamma, v_\gamma\}$ is denoted by $\mathbf{E}$.

For the single-site potential, we assume that B is a spherically symmetric Schwartz function with 0 in the support of its Fourier transform, i.e.

$$0 \in \operatorname{supp}(\widehat{B}). \tag{2.5}$$

More precisely, we introduce the norm

$$\|f\|_{m,n} := \sum_{|\alpha| \leqslant n} \|\langle x \rangle^m \partial^\alpha f(x)\|_\infty$$

with $\langle x \rangle := (1+x^2)^{1/2}$ (here α is a multi-index) and we assume that

$$\|B\|_{k,k} < C_k \quad \text{for all } k \in \mathbf{N}. \tag{2.6}$$

Actually, it is sufficient to assume (2.6) for all $k \leqslant k_0(\varkappa)$.

We note that the operator H_ω is not bounded from below due to the possible large concentration of Poisson points in some region. Nevertheless, H_ω is self-adjoint under very general conditions, see [37].

We introduce a few notational conventions. The letters x, y and z will always denote configuration space variables, while p, q, r, u, v and w will be reserved for d-dimensional momentum variables. The norm without indices, $\|\cdot\|$, will always denote the standard $L^2(\mathbf{R}^d)$-norm. The bracket $(\cdot,\cdot)$ denotes the standard scalar product on $L^2(\mathbf{R}^d)$ and $\langle \cdot, \cdot \rangle$ will denote the pairing between the Schwartz space and its dual on the phase space $\mathbf{R}^d \times \mathbf{R}^d$.

Integrals without explicit domains will always denote integration over $\mathbf{R}^d$ with respect to the Lebesgue measure. For any $f \in L^2(\mathbf{R}^d)$, the *Fourier transform* is given by

$$\hat{f}(p) := \int_{\mathbf{R}^d} \exp(-2\pi i p \cdot x) f(x)\, dx, \quad p \in \mathbf{R}^d, \tag{2.7}$$

and the *inverse Fourier transform* is given by

$$g(x) = \int_{\mathbf{R}^d} \hat{g}(p) \exp(2\pi i p \cdot x)\, dp, \quad x \in \mathbf{R}^d.$$

For functions defined on the phase space, $f(x,v)$, the Fourier transform will always be taken only in the space variable, i.e.

$$\hat{f}(\xi,v) := \int_{\mathbf{R}^d} \exp(-2\pi i \xi\cdot x) f(x,v)\,dx, \quad \xi\in\mathbf{R}^d.$$

The Fourier transform of the kinetic energy operator is given by

$$\widehat{\left[-\frac{\hbar^2}{2m}\Delta f\right]}(p) = \widehat{\left[-\frac{1}{2(2\pi)^2}\Delta f\right]}(p) = e(p)\hat{f}(p),$$

where

$$e(p) := \tfrac{1}{2}p^2 \tag{2.8}$$

is the dispersion law. The velocity is given by

$$\frac{1}{2\pi}\nabla e(p) = \frac{1}{2\pi}p.$$

We note that the additional $\hbar^2/m = (2\pi)^{-2}$ factor was erroneously omitted in the definition of the Hamiltonian in [19, (1.2)].

Define the *Wigner transform* of a function $\psi\in L^2(\mathbf{R}^d)$ as

$$W_\psi(x,v) := \int_{\mathbf{R}^d} \exp(2\pi i v\cdot\eta)\overline{\psi\left(x+\frac{\eta}{2}\right)}\psi\left(x-\frac{\eta}{2}\right)\,d\eta. \tag{2.9}$$

The Fourier transform of $W_\psi(x,v)$ in the x variable is therefore

$$\widehat{W}_\psi(\xi,v) = \overline{\widehat{\psi}\left(v-\frac{\xi}{2}\right)}\widehat{\psi}\left(v+\frac{\xi}{2}\right). \tag{2.10}$$

Define the *rescaled Wigner distribution* as

$$W_\psi^\varepsilon(X,V) := \varepsilon^{-d} W_\psi\left(\frac{X}{\varepsilon}, V\right). \tag{2.11}$$

Its Fourier transform in X is given by

$$\widehat{W}_\psi^\varepsilon(\xi,V) = \overline{\widehat{\psi}\left(V-\frac{\varepsilon\xi}{2}\right)}\widehat{\psi}\left(V+\frac{\varepsilon\xi}{2}\right).$$

For any function $h\colon \mathbf{R}^d\to\mathbf{C}$ and energy value $e\geqslant 0$ we introduce the notation

$$[h](e) := \int_{\mathbf{R}^d} h(v)\delta(e-e(v))\,dv := \int_{\Sigma_e} h(q)\frac{d\nu(q)}{|\nabla e(q)|}, \tag{2.12}$$

where $d\nu(q)$ is the restriction of the Lebesgue measure onto the energy surface

$$\Sigma_e := \{q : e(q) = e\},$$

that is the ball of radius $\sqrt{2e}$. More explicitly,

$$[h](e) := (2e)^{d/2-1} \int_{S^{d-1}} h\big(\sqrt{2e}\,\phi\big)\, d\phi.$$

Clearly

$$\int_{\mathbf{R}^d} h(v)\, dv = \int_0^\infty [h](e)\, de. \tag{2.13}$$

The normalization of the measure $[\cdot](e)$ is given by

$$[1](e) := c_{d-1}(2e)^{d/2-1}, \tag{2.14}$$

where c_{d-1} is the volume of the unit sphere S^{d-1}.

2.2. Main theorem

The weak coupling limit is defined by the following scaling:

$$t = \frac{\mathcal{T}}{\varepsilon}, \quad x = \frac{\mathcal{X}}{\varepsilon}, \quad \varepsilon = \lambda^2. \tag{2.15}$$

The Wigner distribution $W^{\varepsilon}_{\psi_{\mathcal{T}/\varepsilon}}(\mathcal{X}, V)$ converges weakly to a function $F_{\mathcal{T}}(\mathcal{X}, V)$ that satisfies the Boltzmann equation

$$\begin{aligned} &\partial_{\mathcal{T}} F_{\mathcal{T}}(\mathcal{X}, V) + V \cdot \nabla_{\mathcal{X}} F_{\mathcal{T}}(\mathcal{X}, V) \\ &\qquad = 2\pi \int_{\mathbf{R}^d} |\widehat{B}(U-V)|^2 \delta(e(U)-e(V))[F_{\mathcal{T}}(\mathcal{X}, U) - F_{\mathcal{T}}(\mathcal{X}, V)]\, dU. \end{aligned} \tag{2.16}$$

Note that the Boltzmann equation can be viewed as the generator of a Markovian semigroup on phase space. The proof of (2.16) for the continuum Gaussian model was given in [20]; the $\mathbf{Z}^d$ lattice case with general i.i.d. random potential was considered in [11]. The derivation of the Boltzmann equation for potential (2.2) follows from these two proofs in a straightforward way.

In this paper we consider the long time scaling

$$t = \lambda^{-\varkappa}(\lambda^{-2}T), \quad x = \lambda^{-\varkappa/2}(\lambda^{-2}X) = \frac{X}{\varepsilon}, \quad \varepsilon = \lambda^{\varkappa/2+2}, \tag{2.17}$$

with some $\varkappa > 0$. This scaling corresponds to the long time limit of the Boltzmann equation with diffusive scaling.

For some energy $e>0$, let

$$L_e f(v) := \int_{\mathbf{R}^d} \sigma(u,v)[f(u)-f(v)]\,du, \quad e(v)=e, \tag{2.18}$$

be the generator of the momentum jump process $v(t)$ on Σ_e with collision kernel

$$\sigma(u,v) := 2\pi|\widehat{B}(u-v)|^2\delta(e(u)-e(v)). \tag{2.19}$$

Notice that we use u, v, etc. to denote the momenta; the corresponding velocities are $u/2\pi$, $v/2\pi$, etc.

A well-known argument shows that $B\not\equiv 0$ and the regularity of B guarantees the following properties. Some details will be given in [19].

LEMMA 2.1. *For each $e>0$ the Markov process $\{v(t)\}_{t\geqslant 0}$ with generator L_e is uniformly exponentially mixing. The unique invariant measure is the uniform distribution, $[\,\cdot\,](e)/[1](e)$, on the energy surface Σ_e.*

Let

$$D_{hj}(e) := \frac{1}{(2\pi)^2}\int_0^\infty \mathcal{E}_e[v^{(h)}(t)v^{(j)}(0)]\,dt, \quad v=(v^{(1)},...,v^{(d)}), \quad h,j=1,2,...,d,$$

be the velocity autocorrelation matrix, where $\mathcal{E}_e$ denotes the expectation with respect to this Markov process in equilibrium. By the spherical symmetry of $\widehat{B}$ and $e(U)$, the autocorrelation matrix is constant times the identity:

$$D_{hj}(e) = D_e\delta_{hj}, \quad D_e := \frac{1}{(2\pi)^2 d}\int_0^\infty \mathcal{E}_e[v(t)\cdot v(0)]\,dt. \tag{2.20}$$

The main result of the paper is the following theorem.

THEOREM 2.2. *Let $d\geqslant 3$ and $\psi_0\in L^2(\mathbf{R}^d)$ be a normalized initial wave function. Let $\psi(t):=\exp(-itH)\psi_0$ solve the Schrödinger equation (1.1). Let $\mathcal{O}(x,v)$ be a Schwartz function on $\mathbf{R}^d\times\mathbf{R}^d$. For almost all energies $e>0$, $[|\widehat{\psi}_0(v)|^2](e)$ is finite and let f be the solution to the heat equation*

$$\partial_T f(T,X,e) = D_e\Delta_X f(T,X,e) \tag{2.21}$$

with the initial condition

$$f(0,X,e) := \delta(X)[|\widehat{\psi}_0(v)|^2](e)$$

for these energies. Then there exist $0<\varkappa_0(d)\leqslant 2$ such that for $0<\varkappa<\varkappa_0(d)$ and for ε and λ related by (2.17), the rescaled Wigner distribution satisfies

$$\lim_{\lambda\to 0}\int_{\mathbf{R}^d}\int_{\mathbf{R}^d}\mathcal{O}(X,v)\mathbf{E}W^{\varepsilon}_{\psi(\lambda^{-\varkappa-2}T)}(X,v)\,dv\,dX = \int_{\mathbf{R}^d}\int_{\mathbf{R}^d}\mathcal{O}(X,v)f(T,X,e(v))\,dv\,dX, \tag{2.22}$$

and the limit is uniform on $T\in[0,T_0]$ with any fixed T_0. In $d=3$ one can choose

$$\varkappa_0(3) = \tfrac{1}{370}.$$

Remark 1. The coefficient $[|\widehat{\psi}_0(v)|^2](e)$ in the initial condition $f(0,X,e)$ is finite for almost all e by using (2.13) for $h=|\widehat{\psi}_0|^2$.

Remark 2. The total cross section of the collision process (2.18),

$$\sigma_0(e) := \int_{\mathbf{R}^d} \sigma(u,v)\,du, \quad e=e(v), \tag{2.23}$$

is a function of $e=e(v)$ only. Assuming $\widehat{B}(0)\neq 0$, we see that $\sigma_0(e)\sim[1](e)$ for small e, and $\sigma_0(e)\sim e^{-1/2}$ for large e. It follows from Lemma 2.1 and from standard probability arguments that the diffusion constant (2.20) scales as $D_e\sim e/\sigma_0(e)$ for small $e\ll 1$ and $D_e\sim e^2/\sigma_0(e)$ for large $e\gg 1$. If $\widehat{B}$ vanishes at 0 (but (2.5) still holds), then the small energy behaviour of $\sigma_0(e)$ and D_e depend on the rate of vanishing of $\widehat{B}$ at 0 in a straighforward way.

Remark 3. The condition (2.5) is not essential, but the theorem needs to be modified if $\widehat{B}$ vanishes on $D(0,\delta)$, a ball of radius $\delta>0$ about the origin. Let $\delta>0$ be the maximal radius so that $D(0,\delta)\cap\mathrm{supp}(\widehat{B})=\varnothing$. In this case the total cross section $\sigma_0(e)$ is zero for all energy values $e\leqslant\delta^2/8$, because the diameter of the energy surface Σ_e is smaller than the minimal range of $\widehat{B}$. Therefore the evolution is ballistic for the part of the initial wave function that is supported on energy shells $e\leqslant\delta^2/8$. For the other part of the wave function the diffusion equation still holds.

Figure 1 below shows the three different scales schematically. On the Schrödinger scale both time and space are of order 1 in atomic units. On the kinetic scale time and space are rescaled by λ^{-2}. The dynamics is given by the Boltzmann equation characterized by finitely many collisions. On the diffusive scale we rescaled the time and space by an additional factor $\lambda^{-\varkappa}$ and $\lambda^{-\varkappa/2}$, respectively. The typical number of collisions is of order $\lambda^2 t\sim\lambda^{-\varkappa}$.

If we assume that the Boltzmann equation holds under all scalings, Theorem 2.2 can be easily understood. From the Boltzmann equation (2.16), the momentum distribution develops according to the Markovian generator L_e. Therefore, the Boltzmann equation (2.16) describes the process that a particle travels with a fixed momentum v up to an exponentially distributed random time with average value $\sigma_0(e(v))^{-1}$, then it changes momentum from v to a new momentum u on the same energy surface Σ_e chosen by the probability distribution $P(u)=\sigma(u,v)/\sigma_0(e(v))$. The different energy sectors do not interact. Clearly, this process then converges to a Brownian motion in configuration space with a diffusion coefficient given by (2.20) and with momentum restricted to a fixed energy shell Σ_e.

Under the assumption that the Boltzmann equation is valid for all time, this argument applies in all dimensions. The random Schrödinger evolution, however, is expected

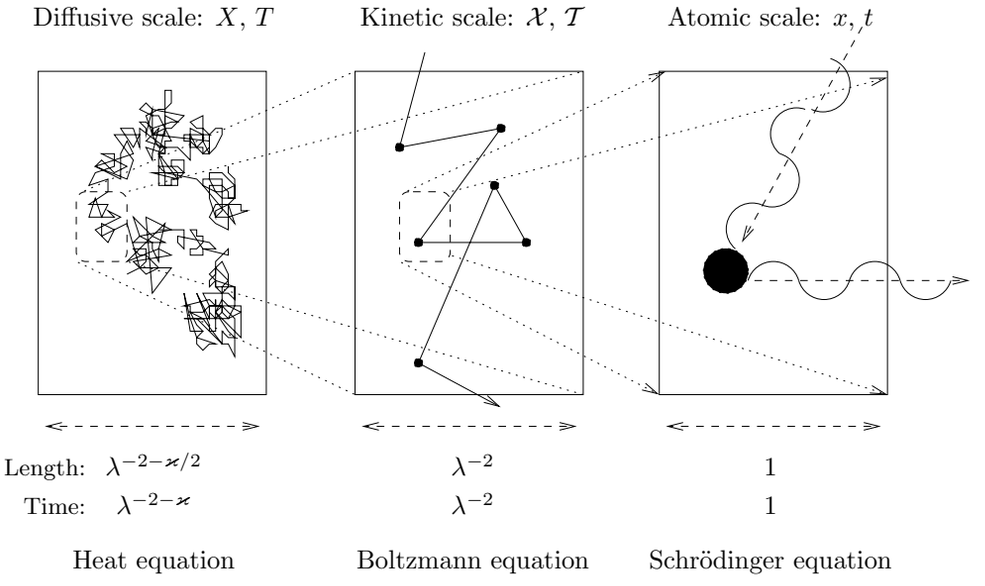

Figure 1. Evolution equations on three scales.

to be localized for $d\leqslant 2$ even for small coupling constants. Therefore, even though the Boltzmann approximation was proved to be valid for $d\geqslant 2$ [20], [11] (it is not valid for $d=1$) in the weak coupling limit, it will not be valid for all time in $d=2$. It is expected that memory effects and quantum correlations eventually dominate the evolution and ruin the Markovian character of the Boltzmann picture. Heuristic ideas show that this transition happens at an exponentially large time (see, e.g., [50]).

The effects of the quantum correlations and memory are not expected to change the Boltzmann picture drastically in $d\geqslant 3$, but one expects corrections to the diffusion equation and a transition between different energy shells for $\varkappa\geqslant 2$ (see [41]).

Although Theorem 2.2 is formulated for the Euclidean space, quantum diffusion on large scales is not restricted to continuum models. In fact, we proved in [18] an analogous result for the Anderson model where the Hamiltonian (2.1) is defined on $\ell^2(\mathbf{Z}^d)$, $d\geqslant 3$. The kinetic energy is the discrete Laplacian and the random potential is given by

$$V_\omega(x) := \sum_{\gamma\in\mathbf{Z}^d} v_\gamma\, \delta(x-\gamma), \quad x\in\mathbf{Z}^d, \tag{2.24}$$

where $\{v_\gamma = v_\gamma(\omega) : \gamma\in\mathbf{Z}^d\}$ are i.i.d. random variables.

The discrete model is technically more involved for two reasons. First, the dispersion relation of the lattice Laplacian lacks convexity which simplifies several estimates in the continuum model. We address this issue in [18].

Second, our choice of the potential in the continuum model contains more randomness than in the discrete one. Comparing (2.2)–(2.3) with (2.24), we note that in the continuum model both the locations of the obstacles and the coupling constants are random variables, while in the Anderson model the locations are deterministic. Formally, the potential (2.24) corresponds to (2.2)–(2.3) with the deterministic choice $y_\gamma = \gamma \in \mathbf{Z}^d$ and with $B(x)$ being the lattice delta function $\delta(x)$. The random choice of the centers in the continuum model simplifies the formula in the calculation of the high moments of the random potential. Nevertheless, we will formulate our estimates in a unified setup that can be applied directly to the Anderson model in [18] as well.

2.3. Strategy of the proof

The above heuristic argument using the Boltzmann equation, besides being misleading for $d=2$, also masks the difficulties in proving Theorem 2.2, namely that one has to follow the full *quantum mechanical* time evolution through infinitely many collisions. The main tool of our proof is to use the Duhamel expansion to decompose the wave function into elementary wave functions characterized by their collision histories. We then apply two strategies to simplify the expansion: (i) renormalization of the propagator, i.e., resumming the two legged subdiagrams; (ii) stopping rules to control recollisions. Apart from these two steps, the bulk of our proof is devoted to giving sharp estimates for a large class of Feynman graphs.

To get an idea, imagine that we expand the solution to the Schrödinger equation by using the Duhamel formula repeatedly. This rewrites the solution into a sum of elementary wave functions, each of which is characterized by a sequence of collisions with random obstacles. When we take the expectation of $\|\psi_t\|^2$ with respect to a Gaussian randomness, we pair the random obstacles by Wick's theorem and obtain a sum of amplitudes of Feynman graphs. In case of a non-Gaussian randomness the higher-order cumulants are also present due to various recollision patterns (their contribution turns out to be negligible, but proving this is rather involved).

If we take only the Laplacian as the free part in the expansion, even the amplitudes of individual graphs diverge in the limit we consider. However, this can be remedied by a simple resummation of all two-legged insertions caused by the lowest order self-energy contribution (it turns out that higher-order corrections to the self-energy do not play a role in the scaling limit we consider). The resummation is performed by choosing an appropriate reference Hamiltonian H_0 for the expansion. After this rearrangement, all graphs have a finite amplitude in our scaling limit, and the ladder graphs give the leading contribution.

However, we have to estimate not only individual graphs but the sum of all graph amplitudes, which requires beating down the factorial growth of the number of graphs. This problem has been addressed in constructive field theory. For field theories with bosons, the graphical expansion to infinite order diverges. Borel summability was proven by cluster expansion and renormalization group methods [8], [9], [23], [24], [26], [33]. In fermionic theories, the anticommutation relations entail cancellations which lead to analyticity in the presence of regulators [21], [25], [32], [45]. Our method to control the combinatorial growth is completely different: it is by very sharp bounds on the individual graphs. We give a classification of arbitrary large graphs, based on counting the number of vertices carrying extra oscillatory effects. The number of these vertices is called the *degree* of the graph and it measures the improvement over the standard power counting. For the ladder graphs, the degree is zero, for the anti-ladder (i.e., complete crossing) graph, it is 2. For general graphs, the degree is roughly the number of vertices after removing all ladder and anti-ladder subgraphs. We thus obtain an extra λ^c factor (for some $c>0$) *per non-(anti)ladder vertex*. This strong improvement is sufficient to beat the growth of the combinatorics in the time scale we consider. To our knowledge, nothing like this has been done in a graphical expansion before. Improved phase space estimates have been used to prove regularity in two-dimensional many-fermion systems, but the improvement exponent was fixed independently of the number of vertices [22], [27], [28].

For a comparison, the unperturbed Green functions in the perturbation expansion for the many-fermion systems and for the random Schrödinger equation are given by

$$\frac{1}{ip_0+p^2-\mu} \quad \text{and} \quad \frac{1}{p^2-\alpha+i\eta}.$$

In the many-fermion case, $\mu>0$ and $p_0\in M_F=\{(2n+1)\pi/\beta : n\in\mathbf{Z}\}$, where $\beta\sim T^{-1}$ is the inverse temperature. In the random Schrödinger case, $\eta\sim t^{-1}$. Their L^2-properties are different:

$$\frac{1}{\beta}\sum_{p_0\in M_F}\int_{\mathbf{R}^d}\frac{dp}{|ip_0+p^2-\mu|^2}\sim|\log\beta|, \quad \int_{\mathbf{R}^d}\frac{dp}{|p^2-\alpha+i\eta|^2}\sim\frac{1}{\eta}.$$

Notice that the divergence is more severe for the random Schrödinger equation case. In the many-fermion case, there is one p_0-summation per line of the graph; in the random Schrödinger case there are just two overall α-integrals for graphs with arbitrarily many lines.

This paper is organized as follows. In §3 we perform the self-energy renormalization, we smooth out the data and restrict the problem to a finite box. The Duhamel expansion is introduced in §4. In §5 we reduce the main theorem to Theorems 5.1, 5.2 and 5.3. The key result is Theorem 5.2, which we prove in the rest of this paper. The other two theorems are more technical and they are proven in the companion paper [19].

The Feynman graphs are introduced in §7. In §8 and §9 we reduce all estimates to Theorem 8.4. This theorem is our main technical bound on Feynman graphs and it is proven in §10.

Since the random potential in our model is given by general i.i.d. random variables, the rule for taking the expectation is different from the case of the Gaussian random field used in [20]. This produces technical difficulties especially for the Anderson model, where, in addition to the usual pairing from the Wick theorem, we have to introduce higher-order partitions of the vertices, called *non-trivial lumps*. Our continuum model avoids this complication due to the additional randomness of the obstacle centers. Nevertheless, we present the general proof here so that the key technical results in this paper could be applied to the Anderson model as well. For readers interested only in the continuum case, we recommend to ignore the non-trivial lumps.

Universal constants and constants that depend only on the dimension d, on the final time T_0 and on the Schwartz norms $\|B\|_{k,k}$ from (2.6) will be denoted by C and their value may vary from line to line.

3. Preparations

3.1. Renormalization

The purpose of this procedure is to include immediate recollisions with the same obstacle into the propagator itself. This is also called the renormalization of "one-particle propagators" or two-legged subdiagrams. Without renormalization, these graphs individually are exponentially large ("divergent"), but their sum is finite. Renormalization removes this instability and the analysis of the resulting Feynman graphs will become simpler.

The self-energy operator is given by the multiplication operator in momentum space

$$\theta(p) := \Theta(e(p)), \quad \Theta(\alpha) := \lim_{\varepsilon \to 0+} \Theta_\varepsilon(\alpha), \quad \Theta_\varepsilon(\alpha) := \Theta_\varepsilon(\alpha, r) \tag{3.1}$$

for any r with $e(r)=\alpha$, where

$$\Theta_\varepsilon(\alpha, r) := \int_{\mathbf{R}^d} \frac{|\widehat{B}(q-r)|^2 \, dq}{\alpha - e(q) + i\varepsilon}. \tag{3.2}$$

Note that by spherical symmetry of B and $e(q)$, $\Theta_\varepsilon(\alpha, r)$ depends only on $|r|$, and note that $|r|=\sqrt{2\alpha}$, therefore $\Theta_\varepsilon(\alpha)$ in (3.1) is well defined. Clearly $\theta(p)$ is spherically symmetric. The existence of the limit and related properties of Θ have been proven in [20] using that $\|\widehat{B}^2\|_{2d,2d} < \infty$. Here we summarize the results.

LEMMA 3.1. *In $d\geqslant 3$ the following hold*:

$$|\Theta_\varepsilon(\alpha,r)-\Theta_\varepsilon(\alpha,r')|\leqslant C\big||r|-|r'|\big| \tag{3.3}$$

(*equation* (3.80) *in* [20]) *and*

$$|\Theta_\varepsilon(\alpha,r)-\Theta_{\varepsilon'}(\alpha',r)|\leqslant C(|\varepsilon-\varepsilon'|+|\alpha-\alpha'|)\varepsilon^{-1/2} \tag{3.4}$$

if $\varepsilon\geqslant\varepsilon'>0$ (*equation* (3.68) *in* [20]). *From this latter estimate, the existence of the limit* $\lim_{\varepsilon\to 0+}\Theta_\varepsilon(\alpha,r)$ *follows. Moreover, Θ is Hölder continuous*:

$$|\Theta(\alpha)-\Theta(\alpha')|\leqslant C|\alpha-\alpha'|^{1/2}. \tag{3.5}$$

Proof. We only have to prove the Hölder continuity. For any ε and for any r and r' with $\alpha=e(r)$ and $\alpha'=e(r')$ we have

$$\begin{aligned}|\Theta(\alpha)-\Theta(\alpha')|&\leqslant \lim_{\varepsilon'\to 0+}|\Theta_{\varepsilon'}(\alpha,r)-\Theta_\varepsilon(\alpha,r)|+|\Theta_\varepsilon(\alpha,r)-\Theta_\varepsilon(\alpha,r')|\\&\quad+|\Theta_\varepsilon(\alpha,r')-\Theta_\varepsilon(\alpha',r')|+\lim_{\varepsilon'\to 0+}|\Theta_\varepsilon(\alpha',r')-\Theta_{\varepsilon'}(\alpha',r')|\\&\leqslant C\big(\varepsilon^{1/2}+\big||r|-|r'|\big|+|\alpha-\alpha'|\varepsilon^{-1/2}\big).\end{aligned}$$

By optimizing ε and using $e(r)=\alpha$, we obtain (3.5). □

We have the following estimate on $\theta(p)$ and similarly on $\Theta(e)$.

LEMMA 3.2. *For any $d\geqslant 3$ there exist universal positive constants c_1, c_2 such that*

$$|\theta(p)|\leqslant\frac{c_2\log\langle p\rangle}{\langle p\rangle},\quad |\Theta(e)|\leqslant\frac{c_2\log\langle e\rangle}{\langle e\rangle^{1/2}}, \tag{3.6}$$

$$\operatorname{Im}\Theta(e)\leqslant -c_1\min\{|e|^{d/2-1},|e|^{-1/2}\},\quad \operatorname{Im}\theta(p)\leqslant -c_1\min\{|p|^{d-2},|p|^{-1}\}. \tag{3.7}$$

Proof. By performing the angular integration, we can write $\Theta_\varepsilon(\alpha,p)$ with $e(p)=\alpha$ as

$$\Theta_\varepsilon(\alpha,p)=\int_0^\infty S(e)\frac{(2e)^{d/2-1}\,de}{\alpha-e+i\varepsilon},\quad \text{with } S(e):=\int_{S^{d-1}}\big|\widehat{B}\big(\sqrt{2e}\,(\phi_r-\phi)\big)\big|^2\,d\phi, \tag{3.8}$$

where ϕ_r is a fixed vector on the unit sphere S^{d-1}. For small values of e,

$$|S(e)|=O(1)\quad\text{and}\quad|\nabla S(e)|=O(e^{-1/2}).$$

For large values of e, using the regularity of $\widehat{B}$,

$$|S(e)|=O(e^{-(d-1)/2})\quad\text{and}\quad|\nabla S(e)|=O(e^{-d/2}).$$

These estimates, in particular, stand behind the proof that $\lim_{\varepsilon\to 0+}\Theta_\varepsilon(\alpha,p)$ is finite, since they guarantee a sufficient decay for large e and a sufficient smoothness around the singularity of the denominator in (3.8). The imaginary part therefore is

$$\operatorname{Im}\theta(p)=\operatorname{Im}\lim_{\varepsilon\to 0+}\Theta_\varepsilon(\alpha,p)=-\pi(2\alpha)^{d/2-1}S(\alpha),$$

which behaves as $\sim -|p|^{d-2}$ for small p and as $\sim -|p|^{-1}$ for large p. The real part of $\Theta_\varepsilon(\alpha,r)$ is bounded for small α. For large α one splits the integration

$$\operatorname{Re}\Theta_\varepsilon(\alpha,r)=\Big(\int_{\alpha-1}^{\alpha+1}+\int_{|\alpha-e|\geqslant 1}\Big)S(e)\frac{(2e)^{d/2-1}\,de}{\alpha-e+i\varepsilon}.$$

After Taylor expanding $(2e)^{d/2-1}S(e)$ around $\alpha\gg 1$, the first term is bounded by

$$\Big|(2\alpha)^{d/2-1}S(\alpha)\int_{\alpha-1}^{\alpha+1}\frac{de}{\alpha-e+i\varepsilon}\Big|+2\sup_{|e-\alpha|\leqslant 1}\Big|\frac{d}{de}[(2e)^{d/2-1}S(e)]\Big|=O(\alpha^{-1/2}),$$

and the second term by

$$\int_{|\alpha-e|\geqslant 1}\frac{de}{|\alpha-e|e^{1/2}}\leqslant\frac{c\log\langle\alpha\rangle}{\langle\alpha\rangle^{1/2}}.$$

If we write $\Theta(e)=\mathcal{R}(e)-i\mathcal{I}(e)$, where $\mathcal{R}(e)$ and $\mathcal{I}(e)$ are real functions, and recall that $\operatorname{Im}(x+i0)^{-1}=-\pi\delta(x)$, we have

$$\mathcal{I}(e)=-\operatorname{Im}\Theta(e)=\pi\int_{\mathbf{R}^d}\delta(e(q)-e)|\widehat{B}(q-r)|^2\,dq \tag{3.9}$$

for any r satisfying $e=e(r)$. □

We rewrite the Hamiltonian as

$$H=H_0+\widetilde{V},$$

where

$$H_0:=\omega(p):=e(p)+\lambda^2\theta(p)\quad\text{and}\quad\widetilde{V}:=\lambda V-\lambda^2\theta(p). \tag{3.10}$$

We note that our renormalization is only an approximation of the standard self-consistent renormalization given by the solution to the equation

$$\omega(p)=e(p)+\lambda^2\lim_{\varepsilon\to 0+}\int_{\mathbf{R}^d}\frac{|\widehat{B}(p-q)|^2\,dq}{\omega(p)-\omega(q)+i\varepsilon}. \tag{3.11}$$

Due to our truncation procedure, the definition (3.1) is sufficient and is more convenient for us. Since $e(p)$ is spherically symmetric, so are $\theta(p)$ and $\omega(p)$.

The following lemma collects some estimates on the renormalized propagators that we shall use to prove Theorem 2.2. The proof is fairly simple and will be given in [19]. We note that formula (2.8) in [19] erroneously contain a factor 2 in the exponent of η, the correct bound is (3.14).

LEMMA 3.3. *Suppose that* $\lambda^2 \geqslant \eta \geqslant \lambda^{2+4\varkappa}$ *with* $\varkappa \leqslant \frac{1}{12}$. *Then we have*

$$\int_{\mathbf{R}^d} \frac{|h(p-q)|\,dp}{|\alpha-\omega(p)+i\eta|} \leqslant \frac{C\|h\|_{2d,0}\,|\log\lambda|\log\langle\alpha\rangle}{\langle\alpha\rangle^{1/2}\langle|q|-\sqrt{2|\alpha|}\,\rangle}, \tag{3.12}$$

and for $0 \leqslant a < 1$,

$$\int_{\mathbf{R}^d} \frac{|h(p-q)|\,dp}{|\alpha-\omega(p)+i\eta|^{2-a}} \leqslant \frac{C_a\|h\|_{2d,0}\,\lambda^{-2(1-a)}}{\langle\alpha\rangle^{a/2}\langle|q|-\sqrt{2|\alpha|}\,\rangle}, \tag{3.13}$$

$$\int_{\mathbf{R}^d} \frac{|h(p-q)|\,dp}{|\alpha-e(p)+i\eta|^{2-a}} \leqslant \frac{C_a\|h\|_{2d,0}\,\eta^{-(1-a)}}{\langle\alpha\rangle^{a/2}\langle|q|-\sqrt{2|\alpha|}\,\rangle}. \tag{3.14}$$

For $a=0$ *and with* $h:=\widehat{B}^2$, *the following more precise estimate holds. There exists a constant* C_0, *depending only on finitely many constants* C_k *from* (2.6), *such that*

$$\int_{\mathbf{R}^d} \frac{\lambda^2|\widehat{B}(p-q)|^2\,dp}{|\alpha-\overline{\omega(p)}-i\eta|^2} \leqslant 1+C_0\lambda^{-12\varkappa}[\lambda+|\alpha-\omega(q)|^{1/2}]. \tag{3.15}$$

3.2. Smoothing the initial data and the potential

In this section we show that it is sufficient to prove the main theorem under the assumptions that $\widehat{\psi}_0(p)$ is a bounded, smooth, compactly supported function, and $\widehat{B}(p)$ is supported on $\{p:|p|\leqslant\lambda^{-\delta}\}$ for any fixed $\delta>0$.

The approximation procedure relies on the following L^2-continuity property of the Wigner transform. If a random wave function is decomposed as $\psi=\psi_1+\psi_2$, then

$$|\mathbf{E}\langle\widehat{\mathcal{O}},\widehat{W}^{\varepsilon}_{\psi}\rangle-\mathbf{E}\langle\widehat{\mathcal{O}},\widehat{W}^{\varepsilon}_{\psi_1}\rangle| \leqslant C\bigg(\int_{\mathbf{R}^d}\sup_v|\widehat{\mathcal{O}}(\xi,v)|\,d\xi\bigg)\sqrt{\mathbf{E}(\|\psi_1\|^2+\|\psi_2\|^2)\cdot\mathbf{E}\|\psi_2\|^2} \tag{3.16}$$

by Schwarz inequality. (Due to a misprint, the $\|\psi_2\|^2$ term was erroneously omitted in §2.1 of our earlier paper [20].)

Approximation of the initial data. Let $\widehat{\psi}_0\in L^2$ and let $\widehat{\psi}_n$ be a sequence of smooth, compactly supported functions with $\|\widehat{\psi}_n-\widehat{\psi}_0\|\to 0$. We decompose $\widehat{\psi}_0=\widehat{\psi}_n+(\widehat{\psi}_0-\widehat{\psi}_n)$. Then

$$\widehat{\psi}(t)=\exp(-itH)\widehat{\psi}_n+\exp(-itH)(\widehat{\psi}_0-\widehat{\psi}_n).$$

Since

$$\|\exp(-itH)(\widehat{\psi}_0-\widehat{\psi}_n)\|=\|\widehat{\psi}_0-\widehat{\psi}_n\|\to 0,$$

as $n\to\infty$, uniformly in t, we see that

$$\lim_{n\to\infty}|\mathbf{E}\langle\widehat{\mathcal{O}},\widehat{W}^{\varepsilon}_{\psi(t)}\rangle-\mathbf{E}\langle\widehat{\mathcal{O}},\widehat{W}^{\varepsilon}_{\psi_n(t)}\rangle|=0$$

uniformly in t (and thus in ε), where $\psi_n(t) := \exp(-itH)\widehat{\psi}_n$ is the time evolution of the approximated initial data. This means that the approximation procedure is continuous on the left-hand side of (2.22).

Similarly, on the right-hand side of (2.22), we can define $f_n(T,X,e)$ to be the solution to (2.21) with initial data $f_n(0,X,e) := \delta(X)[|\widehat{\psi}_n|^2](e)$. Clearly $[|\widehat{\psi}_n|^2](e)$ converges to $[|\widehat{\psi}_0|^2](e)$ in $L^1(de)$. Therefore

$$f_n(T,X,e) \to f(T,X,e) \tag{3.17}$$

in $L^1(dX\,de)$, uniformly in T. The right-hand side of (2.22) is therefore also continuous as $n \to \infty$.

We remark that if $\widehat{\psi}_0$ is smooth, e.g. if $|\nabla_p \widehat{\psi}_0(p)| \leqslant C\langle p\rangle^{-4d}$, then a bounded, smooth and compactly supported approximant, $\widehat{\psi}_n$, can be chosen so that $[|\widehat{\psi}_n|^2](e) \to [|\widehat{\psi}_0|^2](e)$ for every $e>0$, and then the convergence in (3.17) also holds in $L^1(dX)$ for any e. The smoothness of $\widehat{\psi}_0$ is used only at the point when we explicitly compute the main term of the perturbation expansion and identify it with the Boltzmann equation, see [19].

Propagation estimate. To verify that a truncation is allowed for $\widehat{B}$, we first need a crude propagation estimate. Define the following event for any $Z>0$:

$$\Omega_Z := \left\{ \omega : \int_{|y-k|\leqslant 1} d|\mu_\omega|(y) \leqslant Z\langle k\rangle \text{ for all } k \in \mathbf{Z}^d \right\},$$

where $|\mu_\omega|$ denotes the total variation of the (random) measure μ_ω. For any fixed $k \in \mathbf{Z}^d$, let N_k be the number of Poisson points in the ball $\{x : |x-k| \leqslant 1\}$. We compute

$$\mathbf{E}\left\| \int_{|y-k|\leqslant 1} d|\mu_\omega|(y) \right\|^{d+1} \leqslant \mathbf{E} N_k^{d+1} |v|^{d+1} \leqslant C_d,$$

using (2.4) and the fact that N_k is a Poisson random variable with expectation $\mathbf{E}N_k$, which is the volume of the unit ball. By Markov inequality

$$P(\Omega_Z^c) \leqslant \sum_{k\in\mathbf{Z}^d} \frac{C_d}{Z^{d+1}\langle k\rangle^{d+1}} = O(Z^{-d-1}),$$

thus we have

$$\lim_{Z\to\infty} \mathbf{P}(\Omega_Z) = 1. \tag{3.18}$$

We decompose

$$\mathbf{E}\langle \widehat{\mathcal{O}}, \widehat{W}^\varepsilon_{\psi_t}\rangle = \mathbf{E}[\mathbf{1}(\Omega_Z)\langle \widehat{\mathcal{O}}, \widehat{W}^\varepsilon_{\psi_t}\rangle] + \mathbf{E}[\mathbf{1}(\Omega_Z^c)\langle \widehat{\mathcal{O}}, \widehat{W}^\varepsilon_{\psi_t}\rangle],$$

where $\mathbf{1}(\cdot)$ is the characteristic function. On the set Ω_Z^c we use that

$$|\mathbf{E}[\mathbf{1}(\Omega_Z^c)\langle\widehat{\mathcal{O}},\widehat{W}^{\varepsilon}_{\psi_t}\rangle]| \leqslant \left(\int_{\mathbf{R}^d}\sup_v|\widehat{\mathcal{O}}(\xi,v)|\,d\xi\right)\|\psi_t\|^2\mathbf{P}(\Omega_Z^c)\to 0, \tag{3.19}$$

as $Z\to\infty$, uniformly in t (and hence in λ).

For $\omega\in\Omega_Z$ we have $|V_\omega(x)|\leqslant CZ\langle x\rangle$, using the decay properties of B. Computing the time derivative of the mean square displacement, we obtain $\partial_t(\psi_t,x^2\psi_t)=i(\psi_t,[H,x^2]\psi_t)$. Using that $[H,x^2]=-(\nabla\cdot x+x\cdot\nabla)$ and a Schwarz estimate, we have

$$|\partial_t(\psi_t,x^2\psi_t)|\leqslant C(\psi_t,x^2\psi_t)^{1/2}[E+\lambda(\psi_t,|V_\omega|,\psi_t)]^{1/2} \tag{3.20}$$

with $E:=(\psi_t,H\psi_t)=(\psi_0,H\psi_0)$, by energy conservation. We estimate

$$(\psi,|V_\omega|\psi)\leqslant CZ+CZ(\psi,x^2\psi)^{1/2},$$

in particular the energy E is bounded (depending on ψ_0 and Z). From (3.20) we thus have

$$(\psi_t,x^2\psi_t)\leqslant c_1(Z,\psi_0)t^4+c_2(Z,\psi_0) \tag{3.21}$$

on Ω_Z with some constants $c_j(Z,\psi_0)$, $j=1,2$.

Approximation of the potential. We define the truncation of B in Fourier space as $\widehat{B}^\delta(p):=\varphi(\lambda^\delta\langle p\rangle)\widehat{B}(p)$, where $\varphi\colon\mathbf{R}_+\to[0,1]$ is a fixed smooth cutoff function with $\varphi(a)\equiv1$ for $a\leqslant\frac12$ and $\varphi(a)\equiv0$ for $a\geqslant1$. In position space, we have, for any $M\in\mathbf{N}$,

$$|B(x)-B^\delta(x)|\leqslant\langle x\rangle^{-2d}\int_{\mathbf{R}^d}|\langle\nabla_p\rangle^{2d}[\widehat{B}(p)[1-\varphi(\lambda^\delta\langle p\rangle)]]|\,dp\leqslant C_{\delta,M}\lambda^M\langle x\rangle^{-2d}, \tag{3.22}$$

by using the fact that B is in the Schwartz space (2.6).

Let

$$H^\delta:=-\frac12\Delta_x+\lambda\int_{\mathbf{R}^d}B^\delta(x-y)\,d\mu_\omega(y)$$

be the Hamiltonian with the truncated potential. Let $\psi_t^\delta:=\exp(-itH^\delta)\psi_0$ be the evolution of the wave function under the modified Hamiltonian H^δ. On the set Ω_Z and for $t\ll\lambda^{-4}$,

$$\partial_t\|\psi_t-\psi_t^\delta\|^2=-2\operatorname{Im}(\psi_t^\delta,(H-H^\delta)\psi_t)\leqslant C_\delta Z\lambda^{13}(\psi_t,\langle x^2\rangle\psi_t)^{1/2}\leqslant C(Z,\delta,\psi_0)\lambda^5,$$

by using (3.22) with $M=12$. In particular, ψ_t and ψ_t^δ remain close up to time scale $t\sim\lambda^{-2-\varkappa}$, $\varkappa<2$. This bound, together with the L^2-continuity of the Wigner transform (3.16) guarantees that the truncation of B does not influence the left-hand side of (2.22).

As for the right-hand side of (2.22), notice that the collision kernel, $\sigma(U,V)$, of the momentum jump process (2.18) is restricted to the energy surface $e(V)=e(U)=e$. Therefore U and V are bounded, depending on e, so $\widehat{B}(U-V)=\widehat{B}^{\delta}(U-V)$ for these momenta, if λ is sufficiently small. Thus, the truncation of B does not influence the right-hand side of (2.22).

Armed with these results, we assume for the rest of the paper that $\widehat{\psi}_0(p)$ is smooth, compactly supported, bounded, and that $\widehat{B}(p)$ is supported on $\{p:|p|\leqslant\lambda^{-\delta}\}$ for any fixed $\delta>0$. We thus extend the convention from the end of §2 that general constants denoted by C may depend on the truncated versions of $\widehat{B}$ and $\widehat{\psi}_0$. The same applies to the hidden constants in the $O(\cdot)$ and $o(\cdot)$ notations.

3.3. Restriction to a finite box

We will reduce the problem to a finite box of size L, $L\gg1$, with periodic boundary conditions. In this way, for technical convenience, we avoid the infinite summation in (2.3). Let $\Lambda_L:=\left[-\frac{1}{2}L,\frac{1}{2}L\right]^d\subset\mathbf{R}^d$ be a finite torus and let $\Lambda_L^*:=(\mathbf{Z}/L)^d$ be the dual lattice. We introduce the notation

$$\int_{\Lambda_L^*} f(x)\,dp := \frac{1}{|\Lambda_L^*|}\sum_{p\in\Lambda_L^*} f(p). \tag{3.23}$$

The integrals $\int_{\Lambda_L}$ and $\int_{\Lambda_L^*}$ converge to their infinite volume counterparts as $L\to\infty$. Let $(\cdot,\cdot)_L$ and $\|\cdot\|_L$ denote the scalar product and the norm on $L^2(\Lambda_L)$.

For any $L,M\gg1$ we consider the random Schrödinger operator

$$H'=H'_{L,M}:=-\frac{1}{2}\Delta+\lambda V'_\omega,\quad V'_\omega(x):=\sum_{\gamma=1}^{M} v'_\gamma\,B(x-y'_\gamma)=\int_{\Lambda_L} B(x-y)\,d\mu'_\omega,$$

with periodic boundary conditions on Λ_L and $\mu'_\omega:=\sum_{\gamma=1}^{M} v'_\gamma\delta_{y'_\gamma}$. Here $\{y'_\gamma:\gamma=1,...,M\}$ are i.i.d. random variables uniformly distributed on Λ_L and $\{v'_\gamma:\gamma=1,...,M\}$ are i.i.d. variables distributed according to $\mathbf{P}_v$ and they are independent of the y'_γ. The variable M itself will be random; it is chosen to be an independent Poisson variable with expectation $|\Lambda_L|$. The expectation with respect to the joint measure of $\{M,y'_\gamma,v'_\gamma\}$ is denoted by $\mathbf{E}'$. Sometimes we will use the decomposition

$$\mathbf{E}'=\mathbf{E}_M\mathbf{E}_y^{\otimes M}\mathbf{E}_v^{\otimes M} \tag{3.24}$$

referring to the expectation of M, $\{y_\gamma\}$ and $\{v_\gamma\}$ separately. The parameter L is implicit in these notation. In particular, $\mathbf{E}_y^{\otimes M}$ stands for the normalized integral

$$\frac{1}{|\Lambda_L|^M}\int_{(\Lambda_L)^M} dy_1\,...\,dy_M. \tag{3.25}$$

It is well known that the restriction of the random measure μ_ω (see (2.3)) to the box Λ_L has the same distribution as μ'_ω. In particular, given a realization ω of the infinite volume random measure μ_ω, we can associate with it the number of points in Λ_L ($M=M(\omega)$) and the operator $H'_\omega=H_{L,M(\omega)}$ with random measure μ'_ω. We can thus realize the random operator H'_ω on the same probability space as H_ω. Due to the periodic boundary and the non-trivial support of B, the potentials of H_ω and H'_ω will not be the same on Λ_L, but the difference will be negligible far away from the boundary.

Let χ_L be a smooth cutoff function, supported on Λ_L, with $\chi_L\equiv 1$ on $\Lambda_{L/2}$ and $|\nabla\chi_L|\leqslant CL^{-1}$. Let $\psi_L(t):=\chi_L\exp(-itH)\psi'_0$ and let $\psi'(t):=\exp(-itH')\psi'_0$ be the two dynamics applied to the cutoff initial data $\psi'_0:=\chi_L\psi_0$ supported on Λ_L. We also define the cutoff observable $\mathcal{O}_L:=\chi_L\mathcal{O}$. Clearly

$$\lim_{L\to\infty}\mathbf{E}\langle\widehat{\mathcal{O}}_L,\widehat{W}^\varepsilon_{\psi_L(t)}\rangle_L=\mathbf{E}\langle\widehat{\mathcal{O}},\widehat{W}^\varepsilon_{\psi(t)}\rangle \tag{3.26}$$

for any t. We estimate

$$\partial_t\|\psi_L(t)-\psi'_L(t)\|_L^2\leqslant C\|(H-H')\psi_L(t)\|_L^2+C\|[H,\chi_L]\psi_L(t)\|_L^2. \tag{3.27}$$

The second term is bounded by $CL^{-1}\|\nabla\psi_L(t)\|$ and on Ω_Z it can be estimated by the total energy as in (3.20). With a propagation estimate similar to (3.21) but applied to the evolution $\exp(-itH)\psi'_0$, we easily obtain that the right-hand side of (3.27) vanishes as $L\to\infty$ for any t. On the complement set, $\omega\in\Omega_Z^c$, we use the uniform bound (3.19) and finally let $Z\to\infty$. In summary, we have shown the following result.

LEMMA 3.4. *Let $\psi'(t):=\exp(-itH'_{L,M})\psi'_0$, where M is a Poisson random variable with mean $|\Lambda_L|$. Then*

$$\limsup_{L\to\infty}|\mathbf{E}\langle\widehat{\mathcal{O}},\widehat{W}^\varepsilon_{\psi(t)}\rangle-\mathbf{E}'\langle\widehat{\mathcal{O}}_L,\widehat{W}^\varepsilon_{\psi'(t)}\rangle_L|=0$$

whenever $\int_{\mathbf{R}^d}\sup_v|\widehat{\mathcal{O}}(\xi,v)|\,d\xi<\infty$.

4. The Duhamel expansion

We expand the unitary kernel of $H=H_0+\widetilde{V}$ (see (3.10)) by the Duhamel formula. Due to the restriction to Λ_L, we really work with $H'_{L,M}=H'_0+\widetilde{V}'$, where the renormalized free evolution, H'_0, is given by $\omega(p)$ in Fourier space and $\widetilde{V}'=\lambda V'-\lambda^2\theta(p)$, $p\in\Lambda_L^*$. The prime indicates the restriction to Λ_L and the dependence on L and M. In this section we work on Λ_L but we will mostly omit the primes in the notation.

For any fixed integer $N\geqslant 1$,

$$\psi_t := \exp(-itH)\psi_0 = \sum_{n=0}^{N-1} \psi_n(t) + \Psi_N(t), \tag{4.1}$$

with

$$\psi_n(t) := (-i)^n \int_0^t \exp(-is_{n+1}H_0)\widetilde{V} \exp(-is_nH_0)\widetilde{V} \dots \widetilde{V} \exp(-is_1H_0)\psi_0 \, [ds_j]_{j=1}^{n+1} \tag{4.2}$$

being the fully expanded terms and

$$\Psi_N(t) := -i \int_0^t \exp(-i(t-s)H)\widetilde{V}\psi_{N-1}(s)\, ds \tag{4.3}$$

being the non-fully expanded or error term. We used the shorthand notation

$$\int_0^t [ds_j]_{j=1}^n := \int_0^t \dots \int_0^t \delta\Big(t - \sum_{j=1}^n s_j\Big)\, ds_1 \dots ds_n.$$

Since each potential $\widetilde{V}$ in (4.2) and (4.3) is a summation itself,

$$\widetilde{V} = -\lambda^2\theta(p) + \sum_{\gamma=1}^{M} V_\gamma, \quad V_\gamma(x) := v_\gamma B(x - y_\gamma),$$

both of these terms in (4.2) and (4.3) are actually big summations over so-called elementary wave functions, which are characterized by their collision history, i.e. by a sequence of obstacles and, occasionally, by $\theta(p)$. Denote by $\widetilde{\Gamma}_n$, $n\leqslant\infty$, the set of sequences

$$\widetilde{\gamma} = (\widetilde{\gamma}_1, \widetilde{\gamma}_2, \dots, \widetilde{\gamma}_n), \quad \widetilde{\gamma}_j \in \{1, 2, \dots, M\} \cup \{\vartheta\} \tag{4.4}$$

and by $W_{\widetilde{\gamma}}$ the associated potential

$$W_{\widetilde{\gamma}} := \begin{cases} \lambda V_{\widetilde{\gamma}}, & \text{if } \widetilde{\gamma} \in \{1, \dots, M\}, \\ -\lambda^2\theta(p), & \text{if } \widetilde{\gamma} = \vartheta. \end{cases}$$

The tilde refers to the fact that the additional $\{\vartheta\}$ symbol is also allowed. An element $\widetilde{\gamma}\in\{1, \dots, M\}\cup\{\vartheta\}$ is identified with the potential $W_{\widetilde{\gamma}}$ and it is called *potential label* if $\widetilde{\gamma}\in\{1, \dots, M\}$, otherwise it is a *$\vartheta$-label*. A potential label carries a factor λ, a ϑ-label carries λ^2.

For any $\widetilde{\gamma}\in\widetilde{\Gamma}_n$ we define the following fully expanded wave function with truncation

$$\psi_{*t,\widetilde{\gamma}} := (-i)^{n-1} \int_0^t W_{\widetilde{\gamma}_n} \exp(-is_nH_0) W_{\widetilde{\gamma}_{n-1}} \dots W_{\widetilde{\gamma}_1} \exp(-is_1H_0)\psi_0 \, [ds_j]_{j=1}^n, \tag{4.5}$$

and without truncation

$$\psi_{t,\widetilde{\gamma}} := (-i)^n \int_0^t \exp(-is_{n+1}H_0) W_{\widetilde{\gamma}_n} \exp(-is_nH_0) W_{\widetilde{\gamma}_{n-1}} \dots W_{\widetilde{\gamma}_1} \exp(-is_1H_0)\psi_0 \, [ds_j]_{j=1}^{n+1}. \tag{4.6}$$

In the notation, the star $(*)$ will always refer to truncated functions. Note that

$$\psi_{t,\widetilde{\gamma}} = -i \int_0^t \exp(-i(t-s)H_0)\psi_{*s,\widetilde{\gamma}}\, ds.$$

Each term (4.6) along the expansion procedure is characterized by its order n and by a sequence $\widetilde{\gamma}\in\widetilde{\Gamma}_n$. We now identify the main terms.

Denote by $\Gamma_k^{nr}\subset\widetilde{\Gamma}_k$ the set of *non-repetitive* sequences that contain only potential labels, i.e.

$$\Gamma_k^{nr} := \{\gamma = (\gamma_1, ..., \gamma_k) : \gamma_j \in \{1, ..., M\}, \text{ with } \gamma_i \neq \gamma_j \text{ if } i \neq j\}.$$

Let

$$\psi_{t,k}^{nr} := \sum_{\gamma\in\Gamma_k^{nr}} \psi_{t,\gamma}$$

denote the corresponding elementary wave functions.

The typical number of collisions up to time t is of order $\lambda^2 t$. To allow for some room, we set

$$K := \lfloor \lambda^{-\delta}(\lambda^2 t)\rfloor, \tag{4.7}$$

where $\lfloor\cdot\rfloor$ denotes the integer part and $\delta=\delta(\varkappa)>0$ is a small positive number to be fixed later on. K will serve as an upper threshold for the number of collisions in the expansion.

5. Proof of the main theorem

The proof is divided into three theorems. The first one states that all terms other than $\psi_{t,k}^{nr}$, $0\leqslant k<K$, are negligible. For the precise statement we use the previous notation, in particular we recall that the prime indicates the dependence on L and M.

THEOREM 5.1. (L^2-estimate of the error terms) *Let $t=O(\lambda^{-2-\varkappa})$ and K be given by (4.7). If $\varkappa<\varkappa_0(d)$ and δ is sufficiently small (depending only on $\varkappa$), then*

$$\lim_{\lambda\to 0}\lim_{L\to\infty} \mathbf{E}' \left\| \psi_t' - \sum_{k=0}^{K-1} \psi_{t,k}'^{\,nr} \right\|_L^2 = 0.$$

In $d=3$ dimensions, one can choose $\varkappa_0(3)=\frac{1}{370}$.

The second key theorem gives an explicit formula for the main terms, $\psi'^{nr}_{t,k}$. It really identifies the so-called ladder diagram as the only contributing term. We introduce the notation

$$R_\eta(\alpha,v) := \frac{1}{\alpha-\omega(v)+i\eta}$$

for the renormalized propagator.

THEOREM 5.2. (Only the ladder diagram contributes) *Assume that* $\varkappa<2/(33d+36)$ *and let* $\varepsilon=\lambda^{2+\varkappa/2}$, $t=O(\lambda^{-2-\varkappa})$ *and* K *be given by* (4.7). *For a sufficiently small positive* δ, *for* $\eta=\lambda^{2+\varkappa}$ *and for any* $1\leqslant k<K$, *we have*

$$\lim_{L\to\infty} \mathbf{E}'\|\psi'^{nr}_{t,k}\|_L^2 = V_\lambda(t,k)+O(\lambda^{1/3-(6+11d/2)\varkappa-O(\delta)}), \tag{5.1}$$

$$\lim_{L\to\infty} \langle \widehat{\mathcal{O}}_L, \mathbf{E}'\widehat{W}^\varepsilon_{\psi'^{nr}_{t,k}}\rangle_L = W_\lambda(t,k,\mathcal{O})+O(\lambda^{1/3-(6+11d/2)\varkappa-O(\delta)}) \tag{5.2}$$

as $\lambda\ll 1$. *Here*

$$V_\lambda(t,k) := \frac{\lambda^{2k}\exp(2t\eta)}{(2\pi)^2}\int_{\mathbf{R}}\int_{\mathbf{R}}\exp(i(\alpha-\beta)t)\int_{(\mathbf{R}^d)^{k+1}}|\widehat{\psi}_0(p_1)|^2 \times\prod_{j=1}^{k+1}\overline{R_\eta(\alpha,p_j)}R_\eta(\beta,p_j)\prod_{j=1}^{k}|\widehat{B}(p_{j+1}-p_j)|^2\,dp_1\ldots dp_{k+1}\,d\alpha\,d\beta \tag{5.3}$$

$$W_\lambda(t,k,\mathcal{O}) := \frac{\lambda^{2k}\exp(2t\eta)}{(2\pi)^2}\int_{\mathbf{R}}\int_{\mathbf{R}}\exp(i(\alpha-\beta)t)\int_{\mathbf{R}^d}\int_{(\mathbf{R}^d)^{k+1}}\widehat{\mathcal{O}}(\xi,v_{k+1})\overline{\widehat{W}^\varepsilon_{\psi_0}(\xi,v_1)} \times\prod_{j=1}^{k+1}\overline{R_\eta\Big(\alpha,v_j+\frac{\varepsilon\xi}{2}\Big)}R_\eta\Big(\beta,v_j-\frac{\varepsilon\xi}{2}\Big)\prod_{j=1}^{k}|\widehat{B}(v_j-v_{j+1})|^2\,dv_1\ldots dv_{k+1}\,d\xi\,d\alpha\,d\beta. \tag{5.4}$$

We adopt the notation $O(\delta)$ in the exponent of λ. This always means (constant)δ with universal, explicitly computable positive constants that depend on $\varkappa$ and that can be easily computed from the proof.

The formula (5.3) is the value of the so-called *ladder Feynman graph* in the diagrammatic expansion of $\mathbf{E}'\|\psi'^{nr}_{t,k}\|^2$. We will see, in Proposition 7.2, that this expansion generates $k!B_k$ terms, where B_k is the number of partitions of a set with k elements (note that B_k is almost of order $k!$). Theorem 5.2 states that only one diagram is relevant; the contribution of all the other Feynman graphs is negligible even after summation. The extension of (5.1) to the Wigner transform (5.2) is straightforward. Theorem 5.2 is the most important step in the proof of the main theorem.

The third theorem identifies the limit of $\sum_{k=0}^{K-1}W_\lambda(t,k,\mathcal{O})$ as $\lambda\to 0$ with the solution to the heat equation. We note that the definition (5.4) does not apply literally to the free evolution term $k=0$; this term is defined separately:

$$W_\lambda(t,k=0,\mathcal{O}) := \int_{\mathbf{R}^d}\int_{\mathbf{R}^d}\exp(it\varepsilon v\cdot\xi)\exp(2t\lambda^2\operatorname{Im}\theta(v))\widehat{\mathcal{O}}(\xi,v)\overline{\widehat{W}_0(\varepsilon\xi,v)}\,dv\,d\xi. \tag{5.5}$$

THEOREM 5.3. (The ladder diagram converges to the heat equation) *Under the conditions of Theorem* 5.2 *and setting* $t=\lambda^{-2-\varkappa}T$, *we have*

$$\lim_{\lambda\to 0}\sum_{k=0}^{K-1} W_\lambda(t,k,\mathcal{O}) = \int_{\mathbf{R}^d}\int_{\mathbf{R}^d} \mathcal{O}(X,v) f(T,X,e(v))\,dv\,dX, \tag{5.6}$$

where f *is the solution to the heat equation* (2.21).

Proof of Theorem 2.2 *using Theorems* 5.1, 5.2 *and* 5.3. We compute the expectation of the rescaled Wigner transform, $\mathbf{E}W_t^\varepsilon = \mathbf{E}W_{\psi_t}^\varepsilon$, tested against a Schwartz function

$$\int_{\mathbf{R}^d}\int_{\mathbf{R}^d} \mathcal{O}(X,v)\mathbf{E}W_t^\varepsilon(X,v)\,dv\,d\xi = \int_{\mathbf{R}^d}\int_{\mathbf{R}^d} \widehat{\mathcal{O}}(\xi,v)\mathbf{E}\overline{\widehat{W}_t^\varepsilon(\xi,v)}\,dv\,d\xi = \langle \mathcal{O}, \mathbf{E}W_t^\varepsilon\rangle.$$

Combining Lemma 3.4, Theorem 5.1 and the finite box version of the L^2-continuity of the Wigner transform (3.16), it is sufficient to compute the Wigner transform of $\psi'(t,K) := \sum_{k=0}^{K-1} \psi'^{nr}_{t,k}$. The Wigner transform $W_{\psi'(t,K)}$ is quadratic in ψ', so it contains a double sum over k and k':

$$\widehat{W}_{\psi'(t,K)}(\xi,v) = \sum_{k,k'=0}^{K-1} \overline{\widehat{\psi}'^{nr}_{t,k}\left(v-\frac{\xi}{2}\right)}\,\widehat{\psi}'^{nr}_{t,k'}\left(v+\frac{\xi}{2}\right).$$

The potential labels are not repeated within $\bar\psi$ and ψ. Moreover, the expectation of a single potential in (4.6) is zero. Thus the potential labels in the ψ and $\bar\psi$ must pair, in particular taking expectation reduces this double sum to a single sum over k:

$$\mathbf{E}'\widehat{W}_{\psi'(t,K)} = \sum_{k=0}^{K-1} \mathbf{E}'\widehat{W}_{\psi'^{nr}_{t,k}}.$$

By using (5.2) and (5.6) together with $K=O(\lambda^{-\varkappa-\delta})$, we obtain Theorem 2.2. □

The main result of the present paper is the proof of Theorem 5.2. The proofs of Theorems 5.1 and 5.3 will be given in the companion paper [19]. For the reader's convenience, we summarize below the key ideas of the proof of Theorem 5.1 from [19].

The Duhamel expansion allows for the flexibility that at every new term of the expansion we perform the separation into elementary waves, $\psi_{*s,\widetilde\gamma}$, and we can decide whether we want to stop (keeping the full propagator as in (4.3)) or to continue to expand that term further. This decision will depend on the collision history, $\widetilde\gamma$. In particular, not every error term will be expanded up to the same order N, in some cases we may decide to stop the expansion earlier.

To estimate a non-fully expanded term, we will use the unitarity of the full evolution,

$$\left\| -i \int_0^t \exp(-i(t-s)H)\psi_{*s,\widetilde{\gamma}}\, ds \right\|^2 \leqslant t \int_0^t \|\psi_{*s,\widetilde{\gamma}}\|^2\, ds. \tag{5.7}$$

Typically we lose a factor of t by using this estimate, since the oscillatory character of the time integration is lost. We can use this crude estimate only if the fully expanded term, $\|\psi_{*s,\widetilde{\gamma}}\|^2$, is small, i.e. if $\widetilde{\gamma}$ represents an atypical collision sequence. Once $\widetilde{\gamma}$ is "sufficiently" atypical, we stop the expansion for that elementary wave function to reduce the number and the complexity of the expanded terms.

There are basically two patterns how a collision history can become atypical; either the total number of collisions exceeds the typical number of collisions, $O(\lambda^2 t)$, or there is a recollision. This explains why only the non-repetition terms $\psi_{t,k}^{nr}$ with $k<K$ contribute to the main term.

A recollision is typically penalized by a factor λ^2 in the weak coupling environment. This is, however, not the case for the immediate repetition of a potential label, $\widetilde{\gamma}_j=\widetilde{\gamma}_{j+1}\in\{1,...,M\}$. The renormalization (3.10) compensates for these terms. Up to the highest order, the contribution of a sequence with an immediate repetition cancels that of the same sequence where the repetition is replaced by a θ-label. Technically, all these estimates have to be combined with the key method of the present paper (proof of Theorem 5.2) to show that the sum of all $k!B_k$ repetition diagrams is sufficiently small to compensate for the unitarity estimate (5.7).

The result of the current paper (Theorem 5.2) has been fundamentally used in [19]. While that paper was already in print, we have improved the possible range of $\varkappa$ and the exponent in the error bounds in Theorem 5.2 (compare with [19, Theorem 2.3]) and, therefore, several exponents in [19] can be improved. While these improvements are minor (and the exponents used in [19] are still correct), we list them in the following for the convenience of the readers. Instead of $\varkappa<2/(34d+39)$ required in [19, Theorem 2.3], the upper bound $\varkappa<2/(33d+36)$ is sufficient. The exponent $\frac{1}{3}-\left(\frac{17}{3}d+\frac{13}{2}\right)\varkappa-O(\delta)$ appearing in [19, Theorem 2.3] has been improved to $\frac{1}{3}-\left(6+\frac{11}{2}d\right)\varkappa-O(\delta)$. Using the improved exponent in (9.4) instead of $\frac{1}{3}-\left(\frac{17}{3}d+\frac{3}{2}\right)\varkappa-O(\delta)$ in [19, (5.18)], the exponent in [19, Proposition 4.6, (4.37)] can be improved to $\frac{1}{3}-\left(7+\frac{11}{2}d\right)\varkappa-O(\delta)$. The better estimates on terms (I) and (II) at the end of §4.5 and §4.6 of [19] will lead to a somewhat better threshold $\varkappa_0(d)$ for $\varkappa$. More precisely, the upper bound (4.31) of [19] is changed to $\varkappa<(2q-48)/((33d+36)q+108+72d)$ and (4.32) of [19] is changed to $\varkappa<(2q-72)/((33d+36)q+108+72d)$ after correcting a typo $-(16+2d)\varkappa$ to $-(16+12d)\varkappa$ in the exponent of the previous line. In [19, §4.6], the bound (4.40) is changed to $\varkappa<2/(18q+33d+96)$ and (4.41) is unchanged. These bounds yield an explicit $\varkappa_0(d)>0$ depending on the dimension, so that [19, Theorem 4.1] holds for $\varkappa<\varkappa_0(d)$ (the explicit

upper bound in Theorem 4.1 was a typo, it should have been $\varkappa<\varkappa_0(d)$). For $d=3$, explicitly $\varkappa_0(3)>\frac{1}{370}$.

6. Pairing potential labels

The wave function

$$\psi'^{\,nr}_{t,k}=(-i)^k\sum_{\gamma\in\Gamma^{nr}_k}\int_0^t \exp(-is_{k+1}H'_0)V'_{\gamma_k}\exp(-is_kH'_0)\,...\,V'_{\gamma_1}\exp(-is_1H'_0)\psi'_0\,[ds_j]_{j=1}^{k+1}$$

contains k potential terms with different potential labels. Every term in

$$\mathbf{E}'\|\psi'^{\,nr}_{t,k}\|_L^2=\sum_{\gamma,\gamma'}\mathbf{E}'\langle\psi_{t,\gamma},\psi_{t,\gamma'}\rangle$$

has $2k$ potential terms, and their expectation is

$$\mathbf{E}'\,\overline{V'_{\gamma_1}V'_{\gamma_2}\,...\,V'_{\gamma_k}}V'_{\gamma'_1}V'_{\gamma'_2}\,...\,V'_{\gamma'_k}. \tag{6.1}$$

Since there is no repetition within γ and γ', and $\mathbf{E}'V'_\gamma=0$, the expectation in (6.1) is non-zero only if there is a complete pairing between γ and γ'. Such pairings correspond to permutations on $I_k=\{1,2,...,k\}$. We denote by $\mathfrak{S}_k$ the set of all permutations on k elements.

We recall the K-identity from [20, Lemma 3.1] (with a corrected $(2\pi)^{-1}$ factor):

$$\int_0^t\Big(\prod_{j=1}^{k+1}\exp(-is_j\omega(p_j))\Big)\,[ds_j]_{j=1}^{k+1}=\frac{i\exp(\eta t)}{2\pi}\int_{\mathbf{R}}\exp(-i\alpha t)\Big(\prod_{j=1}^{k+1}\frac{1}{\alpha-\omega(p_j)+i\eta}\Big)\,d\alpha \tag{6.2}$$

for any $\eta>0$. Therefore, we have

$$\begin{aligned}\mathbf{E}'\|\psi'^{\,nr}_{t,k}\|_L^2=\frac{\lambda^{2k}\exp(2t\eta)}{(2\pi)^2}&\sum_{\sigma\in\mathfrak{S}_k}\int_{(\Lambda_L^*)^{k+1}}\int_{(\Lambda_L^*)^{k+1}}\delta(p_{k+1}-\tilde p_{k+1})\\&\times\mathbf{E}'\sum_{\substack{\gamma_1,...,\gamma_k=1\\ \gamma_i\neq\gamma_j}}^{M}\prod_{j=1}^{k}\overline{\widehat V_{\gamma_j}(p_{j+1}-p_j)}\widehat V_{\gamma_j}(\tilde p_{\sigma(j)+1}-\tilde p_{\sigma(j)})\overline{\widehat{\psi'_0}(p_1)}\widehat{\psi'_0}(\tilde p_1)\\&\times\int_{\mathbf{R}}\int_{\mathbf{R}}\exp(i(\alpha-\beta)t)\Big(\prod_{j=1}^{k+1}\frac{1}{\alpha-\overline{\omega(p_j)}-i\eta}\cdot\frac{1}{\beta-\omega(\tilde p_j)+i\eta}\Big)\,d\alpha\,d\beta\,d\mathbf{p}\,d\tilde{\mathbf{p}},\end{aligned} \tag{6.3}$$

where the summation runs over all ordered k-tuples $(\gamma_1, ..., \gamma_k)$ of $\{1, 2, ..., M\}$ with distinct elements. We compute the expectation, using $m_2=1$ from (2.4) and the factorization of $\mathbf{E}'$ from (3.24):

$$\mathbf{E}\prod_{j=1}^{k}\overline{\widehat{V}_{\gamma_j}(p_{j+1}-p_j)}\widehat{V}_{\gamma_j}(\tilde{p}_{\sigma(j)+1}-\tilde{p}_{\sigma(j)}) = P(\sigma,\mathbf{p},\tilde{\mathbf{p}})\overline{\mathcal{B}(\mathbf{p})}\mathcal{B}(\tilde{\mathbf{p}}), \tag{6.4}$$

with

$$\mathcal{B}(\mathbf{p}) := \prod_{j=1}^{k}\widehat{B}(p_{j+1}-p_j) \tag{6.5}$$

and

$$P(\sigma,\mathbf{p},\tilde{\mathbf{p}}) := \mathbf{E}_M\mathbf{E}_y^{\otimes M}\sum_{\substack{\gamma_1,...,\gamma_k=1\\ \gamma_i\neq\gamma_j}}^{M}\prod_{j=1}^{k}\exp[2\pi i y_{\gamma_j}(p_{j+1}-p_j-(\tilde{p}_{\sigma(j)+1}-\tilde{p}_{\sigma(j)}))]. \tag{6.6}$$

We obtain from (6.3) that

$$\begin{aligned}\mathbf{E}'\|\psi_{t,k}'^{\,nr}\|_L^2 = \lambda^{2k}\sum_{\sigma\in\mathfrak{S}_k}\int_{(\Lambda_L^*)^{k+1}}\int_{(\Lambda_L^*)^{k+1}}\delta(p_{k+1}-\tilde{p}_{k+1})\\ \times P(\sigma,\mathbf{p},\tilde{\mathbf{p}})M^\circ(k,\mathbf{p},\tilde{\mathbf{p}})\overline{\widehat{\psi}_0'(p_1)}\widehat{\psi}_0'(\tilde{p}_1)\,d\mathbf{p}\,d\tilde{\mathbf{p}},\end{aligned} \tag{6.7}$$

with

$$M^\circ(k,\mathbf{p},\tilde{\mathbf{p}}) := \frac{\exp(2t\eta)}{(2\pi)^2}\int_{\mathbf{R}}\int_{\mathbf{R}}\exp(i(\alpha-\beta)t)\Bigg(\prod_{j=1}^{k+1}\frac{\overline{\widehat{B}(p_{j+1}-p_j)}}{\alpha-\overline{\omega(p_j)}-i\eta}\frac{\widehat{B}(\tilde{p}_{j+1}-\tilde{p}_j)}{\beta-\omega(\tilde{p}_j)+i\eta}\Bigg)\,d\alpha\,d\beta \tag{6.8}$$

(with the convention that for $j=k+1$ we set the superfluous term $\widehat{B}(p_{j+1}-p_j):=1$).

The expectation value in (6.6) can be easily computed to yield a product of delta functions since the variables y_{γ_j} are independent. The constraint $\gamma_i\neq\gamma_j$ induces only a trivial combinatorial factor that becomes irrelevant in the $L\to\infty$ limit.

If the obstacle centers were deterministic, i.e., $y_{\gamma_j}=\gamma_j$, then the constraint $\gamma_i\neq\gamma_j$ has more serious consequences. This is the case for the lattice Anderson model, where the summation in (6.6) extends to all $\gamma_j\in\mathbf{Z}^d$ with $\gamma_i\neq\gamma_j$ and the momentum variables are on the dual torus, $\mathbf{T}^d:=\left[-\frac{1}{2},\frac{1}{2}\right]$. Due to the constraint $\gamma_i\neq\gamma_j$, the formula (6.6) is not a simple product of delta functions and we have to use a connected graph expansion that is well known from statistical physics.

Let $\mathcal{A}_n$ be the set of partitions of $I_k:=\{1,2,...,k\}$, i.e. $\mathbf{A}=\{A_\mu:\mu\in I(\mathbf{A})\}\in\mathcal{A}_k$ if $\bigcup_{\mu\in I(\mathbf{A})}A_\mu=I_k$ and the elements of $\mathbf{A}$ are disjoint and non-empty. The sets in the

partition are labelled by the index set $I(\mathbf{A})$ and let $m(\mathbf{A})=|I(\mathbf{A})|$ denote the number of elements in $\mathbf{A}$. The elements of the partition $\mathbf{A}$ will be called *lumps*. A lump is *trivial* if it has only one element. The trivial partition, where every lump is trivial, is denoted by $\mathbf{A}_0$.

LEMMA 6.1. (i) (Continuum model) *For any fixed L, k and M, with $k\leqslant M$, and any fixed momenta $q_j\in\Lambda_L^*$,*

$$\mathbf{E}_y^{\otimes M}\sum_{\substack{\gamma_1,\dots,\gamma_k=1\\ \gamma_i\neq\gamma_j}}^{M}\prod_{j=1}^{k}\exp(2\pi i q_j y_{\gamma_j})=\frac{M!}{|\Lambda_L|^k(M-k)!}\sum_{\mathbf{A}\in\mathcal{A}_k}\prod_{\nu\in I(\mathbf{A})}c(|A_\nu|)\delta\Big(\sum_{l\in A_\nu}q_l\Big), \quad (6.9)$$

with $c(1)=1$ and $c(n)=0$ for any $n\geqslant 2$.

(ii) (Lattice model) *For any fixed k and momenta $q_j\in\mathbf{T}^d$,*

$$\sum_{\substack{\gamma_1,\dots,\gamma_k\in\mathbf{Z}^d\\ \gamma_i\neq\gamma_j}}\prod_{j=1}^{k}\exp(2\pi i q_j\gamma_j)=\sum_{\mathbf{A}\in\mathcal{A}_k}\prod_{\nu\in I(\mathbf{A})}c(|A_\nu|)\delta\Big(\sum_{l\in A_\nu}q_l\Big), \quad (6.10)$$

with

$$c(n):=\sum_{\substack{\Gamma\subset K_n\\ \Gamma\text{ connected}}}(-1)^{|\Gamma|},$$

where K_n denotes the complete graph on n vertices and $|\Gamma|$ denotes the number of edges in the subgraph Γ. The following estimate holds for $n\geqslant 2$:

$$|c(n)|\leqslant n^{n-2}. \quad (6.11)$$

Remark 1. Recall that M is a Poisson random variable with expectation $|\Lambda_L|$. Therefore, apart from the prefactor that converges to 1 almost surely as $L\to\infty$, the right-hand side of (6.9) is simply $\prod_{j=1}^k\delta(q_j)$. With an obvious choice of $c(n)$, we write it in the same form as (6.10). In this way, the continuum and lattice models can be treated simultaneously. The explicit form of $c(n)$ will not be needed. The arguments in the sequel will use only the bound (6.11) that is valid for both choices of $c(n)$.

Remark 2. Analogous formulas hold if the natural index set $I_k=\{1,2,...,k\}$ is replaced by an arbitrary finite set S. In this case, the summation on the right-hand side of (6.9)–(6.10) is over all partitions of S. The set of these partitions is denoted by $\mathcal{A}(S)$.

Proof of Lemma 6.1. Part (i) is straightforward from the definition of $\mathbf{E}_y$.

For part (ii) we use the connected graph expansion

$$\prod_{i\neq j=1}^{k}(1-\delta_{\gamma_i,\gamma_j})=\sum_{\mathbf{A}\in\mathcal{A}_k}\prod_{\nu\in I(\mathbf{A})}\delta_c(A_\nu),$$

where

$$\delta_c(A)=c(|A|)\prod_{l,l'\in A}\delta_{\gamma_l,\gamma_{l'}}$$

is the Ursell coefficient of the hard-core lattice gas (see, e.g., [47]). Therefore,

$$\begin{aligned}&\sum_{\substack{\gamma_1,\dots,\gamma_k\in\mathbf{Z}^d\\ \gamma_i\neq\gamma_j}}\prod_{j=1}^{k}\exp(2\pi i q_j\gamma_j)\\ &\quad=\sum_{\gamma_1,\dots,\gamma_k\in\mathbf{Z}^d}\sum_{\mathbf{A}\in\mathcal{A}_k}\prod_{\nu\in I(\mathbf{A})}\biggl(\exp\biggl(2\pi i\sum_{l\in A_\nu}q_l\gamma_l\biggr)c(|A_\nu|)\prod_{l,l'\in A_\nu}\delta_{\gamma_l,\gamma_{l'}}\biggr)\\ &\quad=\sum_{\mathbf{A}\in\mathcal{A}_k}\prod_{\nu\in I(\mathbf{A})}c(|A_\nu|)\delta\biggl(\sum_{l\in A_\nu}q_l\biggr).\end{aligned}$$

□

We will use the identity (6.9) to express P in (6.6) as a linear combination of products of delta functions of the momenta and insert it into (6.7). After the limit $L\to\infty$, each term in the summation $\sum_\sigma\sum_{\mathbf{A}}$ will be expressed by a Feynman graph. The precise definitions will be given in the next section.

7. Graphical representation

Traditionally, the Feynman graphs consist of interaction vertices and particle lines among them. In case of Gaussian random potentials, the interaction vertices are paired according to the Wick theorem [20]. For non-Gaussian randomness, the non-vanishing higher-order cumulants correspond to joining several vertices [11]. In our case, the appearance of the non-trivial subsets are due to selecting the non-repetition sequences. This requires us to define Feynman graphs in a more general setup than usual. In this section we introduce the necessary graphical representation in full generality and we will define the value of a Feynman graph, $V^\circ(\mathbf{A},\sigma)$, with permutation σ and partition $\mathbf{A}$ in (7.21). The final result of this section is given in Proposition 7.2 at the end.

7.1. Circle graphs and their values

We start with an oriented circle graph with two distinguished vertices, denoted by 0 and 0^*. The number of vertices is N. The vertex set is $\mathcal{V}$ and the set of oriented edges

is $\mathcal{L}(\mathcal{V})$. For $v\in\mathcal{V}$ we use the notation $v-1$ and $v+1$ for the vertex right before and after v in the circular ordering. We also let $e_{v-}=(v-1,v)$ and $e_{v+}=(v,v+1)$ be the edges right before and after the vertex v, respectively. In particular $e_{(v+1)-}=e_{v+}$. For each $e\in\mathcal{L}(\mathcal{V})$, we introduce a momentum w_e and a real number α_e associated with this edge. The collection of all momenta is denoted by $\mathbf{w}=\{w_e:e\in\mathcal{L}(\mathcal{V})\}$ and $d\mathbf{w}=\bigotimes_e dw_e$ is the Lebesgue measure. We sometimes use the notation $v\sim e$ to indicate that an edge e is adjacent to a vertex v.

Let $\mathbf{P}=\{P_\mu:\mu\in I\}$ be a partition of the set $\mathcal{V}\setminus\{0,0^*\}$,

$$\mathcal{V}\setminus\{0,0^*\}=\bigcup_{\mu\in I}P_\mu$$

(with all P_μ non-empty and pairwise disjoint), where $I=I(\mathbf{P})$ is the index set to label the sets in the partition. Let $m(\mathbf{P}):=|I(\mathbf{P})|$. The sets P_μ are called $\mathbf{P}$-*lumps* or just *lumps*. If two elements $v,v'\in\mathcal{V}\setminus\{0,0^*\}$ belong to the same lump within a partition $\mathbf{P}$, we denote it by $v\equiv v'$ (mod $\mathbf{P}$). We assign a variable $u_\mu\in\mathbf{R}^d$, $\mu\in I(\mathbf{P})$, to each lump. We call them *auxiliary momenta*; they will be needed for a technical reason. We always assume that the auxiliary momenta add up to zero:

$$\sum_{\mu\in I(\mathbf{P})}u_\mu=0. \tag{7.1}$$

The set of all partitions of the vertex set $\mathcal{V}\setminus\{0,0^*\}$ is denoted by $\mathcal{P}_\mathcal{V}$. For any $P\subset\mathcal{V}$, we let

$$L_+(P):=\{(v,v+1)\in\mathcal{L}(\mathcal{V}):v\in P\text{ and }v+1\notin P\}$$

denote the set of edges that go out of P, with respect to the orientation of the circle graph, and similarly $L_-(P)$ denote the set of edges that go into P. We set

$$L(P):=L_+(P)\cup L_-(P).$$

For any $\xi\in\mathbf{R}^d$ we define the following product of delta functions

$$\Delta(\mathbf{P},\mathbf{w},\mathbf{u}):=\delta\Big(\xi+\sum_{e\in L_\pm(\{0^*\})}\pm w_e\Big)\prod_{\mu\in I(\mathbf{P})}\delta\Big(\sum_{e\in L_\pm(P_\mu)}\pm w_e-u_\mu\Big), \tag{7.2}$$

where $\mathbf{u}:=\{u_\mu:\mu\in I(\mathbf{P})\}\in\mathbf{R}^d$ is a set of auxiliary momenta. The sign $\pm$ indicates that the momentum w_e is added or subtracted depending whether the edge e is outgoing or incoming, respectively. The function $\Delta(\mathbf{P},\mathbf{w},\mathbf{u})=\Delta_\xi(\mathbf{P},\mathbf{w},\mathbf{u})$ depends on ξ, but we will mostly omit this fact from the notation. All estimates will be uniform in ξ.

Summing up all arguments of these delta functions and using (7.1), we see that these delta functions force the two momenta corresponding to the two edges adjacent to 0 to differ by ξ: $w_e - w_{e'} = \xi$ for $e \in L_+(\{0\})$ and $e' \in L_-(\{0\})$.

As a motivation for these definitions, we mention that the lumps naturally arise from the connected graph formula (Lemma 6.1). According to this formula, the Kirchoff law must be satisfied for all lumps, i.e. the incoming and outgoing momenta must sum up to zero. This fact would be described by the delta functions (7.2) with all $u_\mu = 0$. In certain recollision terms, however, the non-repetition condition leading to Lemma 6.1 is not fully satisfied and the Kirchoff law breaks down for a few lumps. The non-trivial auxiliary momenta will bookkeep this deviation from the Kirchoff law (see [19] for more details). Finally, the shift by ξ at the vertex 0^* in (7.2) will be used when computing the Wigner transform in the Fourier representation (2.10).

For each subset $\mathcal{G} \subset \mathcal{V} \setminus \{0, 0^*\}$, we define

$$\mathcal{N}_{\mathcal{G}}(\mathbf{w}) := \prod_{e \sim 0} |\widehat{\psi}_0(w_e)| \prod_{v \in (\mathcal{V} \setminus \{0,0^*\}) \setminus \mathcal{G}} |\widehat{B}(w_{e_{v-}} - w_{e_{v+}})| \prod_{v \in \mathcal{G}} \langle w_{e_{v-}} - w_{e_{v+}} \rangle^{-2d}. \tag{7.3}$$

In our application, the subset $\mathcal{G}$ collects those vertices, where the original potential decay $|\widehat{B}(w_{e_{v-}} - w_{e_{v+}})|$ could not be explicitly kept along the estimates and this will happen only at a few places; the size of $\mathcal{G}$ will be at most 8. For the purpose of this paper, i.e. for the proof of Theorem 5.2, we will need only $\mathcal{G} = \varnothing$, but for the analysis of the repetition terms in [19] we need the more general definition.

Due to the support properties of $\widehat{B}$ and $\widehat{\psi}_0$, we will see that all intermediate momenta w_e satisfy $|w_e| \leqslant N\lambda^{-\delta}$. The maximal number of vertices in our graphs will be

$$2K + 2 = O(\lambda^{-\varkappa - \delta}),$$

therefore all intermediate momenta will be smaller than $\zeta := \lambda^{-\varkappa - 3\delta}$. This justifies the definition of the restricted Lebesgue measures

$$d\mu(w) := \mathbf{1}(|w| \leqslant \zeta)\, dw, \quad \zeta := \lambda^{-\varkappa - 3\delta}, \quad d\mu(\mathbf{w}) := \bigotimes_e d\mu(w_e). \tag{7.4}$$

Moreover, each auxiliary momentum u_μ will always be a sum (or difference) of different w_e momenta (see (7.2)), therefore each of them always satisfies $|u_\mu| \leqslant O(\lambda^{-2\varkappa - 4\delta})$. We will often take the supremum of all possible auxiliary momenta and $\sup_{\mathbf{u}}$ is always considered subject to this bound.

With this notation, we define, for any $\mathbf{P} \in \mathcal{P}_{\mathcal{V}}$ and $g = 0, 1, 2, \dots$, the *E-value* of the partition:

$$E_g(\mathbf{P}, \mathbf{u}, \boldsymbol{\alpha}) := \lambda^{N-2} \sup_{\mathcal{G}: |\mathcal{G}| \leqslant g} \int \prod_{e \in \mathcal{L}(\mathcal{V})} \frac{1}{|\alpha_e - \omega(w_e) + i\eta|}\, \Delta(\mathbf{P}, \mathbf{w}, \mathbf{u}) \mathcal{N}_{\mathcal{G}}(\mathbf{w})\, d\mu(\mathbf{w}). \tag{7.5}$$

The prefactor λ^{N-2} is due to the fact that in the applications all but the two distinguished vertices, $\{0,0^*\}$, will carry a factor λ. The E-value depends also on the parameters λ and η, but we will not specify them in the notation. In the applications, the regularization η will be mostly chosen as $\eta=\lambda^{2+\varkappa}$.

We will also need a slight modification of these definitions, indicated by a lower star in the notation:

$$E_{*g}(\mathbf{P},\mathbf{u},\boldsymbol{\alpha}) := \lambda^{N-2} \sup_{\mathcal{G}:|\mathcal{G}|\leqslant g} \int \prod_{\substack{e\in\mathcal{L}(\mathcal{V})\\ e\notin L(\{0^*\})}} \frac{1}{|\alpha_e-\omega(w_e)+i\eta|}\, \Delta(\mathbf{P},\mathbf{w},\mathbf{u})\mathcal{N}_{\mathcal{G}}(\mathbf{w})\, d\mu(\mathbf{w}). \quad (7.6)$$

The only difference is that the denominators carrying the momenta associated with edges that are adjacent to 0^* are not present in E_{*g}. We call E_{*g} the *truncation* of E_g. We will see that Feynman diagrams arising from the perturbation expansion can naturally be estimated by quantities of the form (7.5) or (7.6).

7.2. Feynman graphs

We apply this general setup to the following situation that we will call *Feynman graph*. Every quantity in our perturbation expansion will be expressed by values of Feynman graphs that are defined below.

For experts we mention that our Feynman graphs differ from those that one typically obtains after averaging over a Gaussian disorder. In the latter, potential lines never appear as external lines but only as pairing lines, and one can identify vertices connected by pairing lines so that the graph becomes 4-valent. In our case, the graph is still trivalent and has external potential lines, with a corresponding dependence on momentum variables $\mathbf{u}$. Also, averaging the disorder will not simply pair up lines but can also join more than two potential lines, which correspond to the higher moments.

Consider the cyclically ordered set $\mathcal{V}_{n,n'}:=\{0,1,2,...,n,0^*,\tilde{n}',\widetilde{n'-1},...,\tilde{1}\}$ and view this as the vertex set of an oriented circle graph on $N=n+n'+2$ vertices. We set

$$I_n := \{1,2,...n\} \quad \text{and} \quad \tilde{I}_{n'} := \{\tilde{1},\tilde{2},...,\tilde{n}'\},$$

so that $\mathcal{V}_{n,n'}=I_n\cup\tilde{I}_{n'}\cup\{0,0^*\}$.

The set of edges $\mathcal{L}(\mathcal{V}_{n,n'})$ is partitioned into $\mathcal{L}(\mathcal{V}_{n,n'})=\mathcal{L}\cup\widetilde{\mathcal{L}}$, where $\mathcal{L}$ contains the edges between $I_n\cup\{0,0^*\}$ and $\widetilde{\mathcal{L}}$ contains the edges between $\tilde{I}_{n'}\cup\{0,0^*\}$.

Let $\mathcal{P}_{n,n'}$ be the set of all partitions $\mathbf{P}$ on the set $I_n\cup\tilde{I}_{n'}$. The lumps of a partition containing only one vertex will be called *single lumps*. The vertices 0 and 0^* are not part of the partitions, hence they will not be considered as single lumps. Let $G=G(\mathbf{P})$

be the set of edges that go into a single lump and let $g(\mathbf{P}):=|G(\mathbf{P})|$ be its cardinality. In case $n=n'$, we will use the shorter notation $\mathcal{V}_n=\mathcal{V}_{n,n}$, $\mathcal{P}_n=\mathcal{P}_{n,n}$, etc. The Feynman graphs arising from the non-repetition terms will always have $n=n'$ and no single lumps, $g(\mathbf{P})=0$, but the more general definition will be needed for the repetition terms in [19]. We remark that even in [19] we will always have

$$|n-n'|\leqslant g(\mathbf{P})\leqslant 4, \quad n,n'\leqslant K. \tag{7.7}$$

We also introduce a function Q that will represent the momentum dependence of the observable. In our estimates, we will always bound Q in supremum norm; no decay or smoothness will be necessary. We will need extra conditions on the observable only to evaluate the ladder in the proof of Theorem 5.3 (see [19] for details). Since Q will always appear linearly in our formulae, we can assume, for convenience, that $\|Q\|_\infty\leqslant 1$. General Q can be accommodated by a multiplicative factor $\|Q\|_\infty$ in the final estimate but it will not be carried along the proofs.

We define the following function to collect all potential terms:

$$\begin{aligned}\mathcal{M}(\mathbf{w}):=\prod_{e\in\mathcal{L}\cap G}[-\overline{\theta(w_e)}]\prod_{e\in\widetilde{\mathcal{L}}\cap G}[-\theta(w_e)]\prod_{\substack{e\in\mathcal{L}\setminus G\\ e\not\sim 0^*}}\overline{\widehat{B}(w_e-w_{e+1})}\prod_{\substack{e\in\widetilde{\mathcal{L}}\setminus G\\ e\not\sim 0}}\widehat{B}(w_e-w_{e+1})\\ \times\overline{\widehat{\psi}(w_{e_{0+}})}\widehat{\psi}(w_{e_{0-}})Q\big[\tfrac{1}{2}(w_{e_{0^*-}}+w_{e_{0^*+}})\big],\end{aligned} \tag{7.8}$$

with $\mathbf{w}:=\{w_e:e\in\mathcal{L}\cup\widetilde{\mathcal{L}}\}$, and recalling that for any $e\in\mathcal{L}(\mathcal{V})$ the edge $e+1$ denotes the edge succeeding e in the circular ordering.

The delta function $\Delta(\mathbf{P},\mathbf{w},\mathbf{u}\equiv 0)$ ensures that the two momenta adjacent to each single lump coincide. This holds even for $\xi\neq 0$, recall that 0 and 0^* are not considered as lumps. Therefore the distribution $\mathcal{M}(\mathbf{w})\Delta(\mathbf{P},\mathbf{w},\mathbf{u}\equiv 0)$ is supported on the regime with $|w_e|\leqslant\zeta$ for all momenta w_e, due to the support properties of $\widehat{\psi}_0$ and $\widehat{B}$, and to the control on the number of terms, $n,n'\leqslant K$ (see §7.1). In particular

$$\mathcal{M}(\mathbf{w})\Delta(\mathbf{P},\mathbf{w},\mathbf{u}\equiv 0)\,d\mathbf{w}=\mathcal{M}(\mathbf{w})\Delta(\mathbf{P},\mathbf{w},\mathbf{u}\equiv 0)\,d\mu(\mathbf{w}). \tag{7.9}$$

Using the boundedness of θ and $|\widehat{\psi}_0(w)|\leqslant C\langle w\rangle^{-10d}$, we easily obtain that

$$|\mathcal{M}(\mathbf{w})|\Delta(\mathbf{P},\mathbf{w},\mathbf{u}\equiv 0)\,d\mathbf{w}\leqslant C^{g(\mathbf{P})}\mathcal{N}_{\mathcal{G}}(\mathbf{w})\Delta(\mathbf{P},\mathbf{w},\mathbf{u}\equiv 0)\,d\mu(\mathbf{w}), \tag{7.10}$$

where $\mathcal{G}$ is the set of single lumps and $g(\mathbf{P})=|\mathcal{G}|$, since the delta function also guarantees that there is no additional decay at the vertices $v\in\mathcal{G}$ in $\mathcal{N}_{\mathcal{G}}(\mathbf{w})$ (the last product in (7.3) is a constant).

Let $\alpha,\beta\in\mathbf{R}$, $\mathbf{P}\in\mathcal{P}_{n,n'}$ and

$$V(\mathbf{P},\alpha,\beta) := \lambda^{n+n'+g(\mathbf{P})} \int \Big(\prod_{e\in\mathcal{L}} \frac{1}{\alpha-\overline{\omega(w_e)}-i\eta} \Big) \Big(\prod_{e\in\widetilde{\mathcal{L}}} \frac{1}{\beta-\omega(w_e)+i\eta} \Big) \times \Delta(\mathbf{P},\mathbf{w},\mathbf{u}\equiv 0)\mathcal{M}(\mathbf{w})\, d\mathbf{w}. \tag{7.11}$$

Due to (7.9), the integration measure could be changed to $d\mu(\mathbf{w})$. The truncated version, $V_*(\mathbf{P},\alpha,\beta)$, is defined analogously but the α and β denominators that correspond to $e\in L(\{0^*\})$ are removed.

We set $Y:=\lambda^{-100}$ and define

$$V_{(*)}(\mathbf{P}) := \frac{\exp(2t\eta)}{(2\pi)^2} \int_{-Y}^{Y}\int_{-Y}^{Y} \exp(it(\alpha-\beta))V_{(*)}(\mathbf{P},\alpha,\beta)\, d\alpha\, d\beta \tag{7.12}$$

and

$$E_{(*)g}(\mathbf{P},\mathbf{u}) := \frac{\exp(2t\eta)}{(2\pi)^2} \int_{-Y}^{Y}\int_{-Y}^{Y} E_{(*)g}(\mathbf{P},\mathbf{u},\boldsymbol{\alpha})\, d\alpha\, d\beta, \tag{7.13}$$

where $\boldsymbol{\alpha}$ in $E_{(*)g}(\mathbf{P},\mathbf{u},\boldsymbol{\alpha})$ is defined as $\alpha_e=\alpha$ for $e\in\mathcal{L}$ and $\alpha_e:=\beta$ for $e\in\widetilde{\mathcal{L}}$. The notation $(*)$ indicates the same formulas with and without truncation. We will call these numbers the *V-value* and the *E-value* of the partition $\mathbf{P}$, or sometimes, of the corresponding Feynman graph. Strictly speaking, the V-value and the E-value depend on ξ through $\Delta=\Delta_\xi$. When this dependence is important, we will make it explicit in the notation, e.g. $V=V_\xi$. The V-value depends on the choice of Q as well. When necessary, the notation $V_\xi(\mathbf{P};Q)$ will indicate this fact.

Clearly, by using (7.10),

$$|V_{(*)}(\mathbf{P})| \leqslant (C\lambda)^{g(\mathbf{P})} E_{(*)g}(\mathbf{P},\mathbf{u}\equiv 0), \tag{7.14}$$

with $g=g(\mathbf{P})$. We will use the notation $E_{(*)g}(\mathbf{P}):=E_{(*)g}(\mathbf{P},\mathbf{u}\equiv 0)$.

As we will see in (6.2), for the graphical representation of the Duhamel expansion we will really need

$$V^\circ_{(*)}(\mathbf{P}) := \frac{\exp(2t\eta)}{(2\pi)^2} \int_{\mathbf{R}}\int_{\mathbf{R}} \exp(it(\alpha-\beta))V_{(*)}(\mathbf{P},\alpha,\beta)\, d\alpha\, d\beta, \tag{7.15}$$

that is, a version of $V_{(*)}(\mathbf{P})$ with unrestricted $d\alpha\, d\beta$ integrations. (The circle superscript in V° will refer to the unrestricted version of V.) However, the difference between the restricted and unrestricted V-values are negligible even when we sum them up over all partitions.

LEMMA 7.1. *Assuming* (7.7) *and that* $\eta \geqslant \lambda^{2+4\varkappa}$, *we have*

$$\sum_{\mathbf{P}\in\mathcal{P}_{n,n'}} |V_{(*)}(\mathbf{P}) - V^{\circ}_{(*)}(\mathbf{P})| = O(\lambda^{5(n+n')}). \tag{7.16}$$

The same result holds if $V_{(*)}(\mathbf{P})$ *were defined by restricting the* α,β*-integral to any domain that contains* $[-Y,Y]\times[-Y,Y]$.

Proof. First we consider the case $n, n' \geqslant 1$. To estimate the difference, we consider the integration domain where either $|\alpha| \geqslant Y$ or $|\beta| \geqslant Y$. We assume, for definiteness, that $|\alpha| \geqslant Y$, and we estimate all α-denominators trivially,

$$\frac{1}{|\alpha - \overline{\omega(w_e)} - i\eta|} \leqslant \frac{C}{\langle\alpha\rangle},$$

by using the fact that $|\omega(w_e)| \leqslant \frac{1}{2}w_e^2 + O(\lambda^2) \leqslant \frac{1}{2}Y + O(\lambda^2)$ on the support of $d\mu(w_e)$. Then we estimate all but the last β-denominators in (7.11) trivially by $\eta^{-1}\langle\beta/Y\rangle^{-1}$. Thus, all w_e integrations are trivial except the last one, where we use (3.12). Due to the bounds $|\widehat{\psi}(w)|, |\widehat{B}(w)| \leqslant C\langle w\rangle^{-10d}$, one easily obtains that

$$|V(\mathbf{P},\alpha,\beta)| \leqslant \frac{(C\lambda)^{n+n'+g(\mathbf{P})}|\log\lambda|\log\langle\beta\rangle}{\langle\alpha\rangle^{n+1}\eta^{n'}\langle\beta/Y\rangle^{n'}\langle\beta\rangle^{1/2}}.$$

Therefore, we have

$$\int_{\mathbf{R}}\int_{\{\alpha:|\alpha|\geqslant Y\}} |V(\mathbf{P},\alpha,\beta)|\,d\alpha\,d\beta \leqslant \frac{(C\lambda)^{n+n'+g(\mathbf{P})}|\log\lambda|}{\eta^{n'}Y^{n-1/2-2\delta}} = O(\lambda^{6(n+n')}),$$

using (7.7). A similar bound holds for the truncated values, V_*. Thus

$$|V_{(*)}(\mathbf{P}) - V^{\circ}_{(*)}(\mathbf{P})| = O(\lambda^{6(n+n')}). \tag{7.17}$$

Since the total number of partitions, $|\mathcal{P}_{n,n'}|$, is smaller than $(n+n')^{n+n'}$, and in our applications $n, n' \leqslant K \ll \lambda^{-\varkappa-2\delta}$, we see that the restriction of the α,β-integral to any domain that contains $[-Y,Y]\times[-Y,Y]$ has a negligible effect of order $O(\lambda^{5(n+n')})$, even after summing up all V-values.

If the condition $n, n' \geqslant 1$ is not satisfied, say there is only one α-denominator ($n=0$), then we will not introduce the auxiliary variable α as in (6.2), because the $\int d\alpha$ integral would be logarithmically divergent after taking the absolute value. In this case, we use the definition

$$V^{\circ}(\mathbf{P}) := \frac{\exp(t\eta)}{2\pi}\int_{\mathbf{R}} \exp(-it\beta)V^{\circ}(\mathbf{P},\beta)\,d\beta,$$

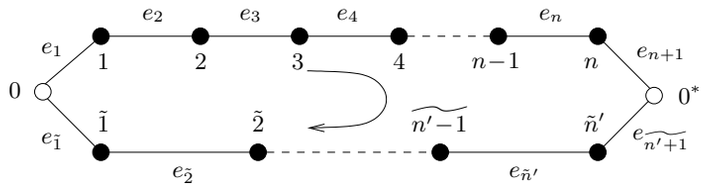

Figure 2. Vertex and edge labels.

with

$$V^\circ(\mathbf{P},\beta) := \lambda^{n+n'+g(\mathbf{P})} \int \Bigg(\prod_{e\in\widetilde{\mathcal{L}}} \frac{1}{\beta-\omega(w_e)+i\eta} \Bigg) \exp(it\overline{\omega(w)}) \\ \times \Delta(\mathbf{P},\mathbf{w},\mathbf{u}\equiv 0)\mathcal{M}(\mathbf{w}) \Bigg(\prod_{e\in\mathcal{L}\cup\widetilde{\mathcal{L}}} dw_e \Bigg),$$

directly, instead of V° given in (7.15). Similar modifications hold for the other cases ($n'=0$, $n\geqslant 1$ and $n'=n=0$) as well. In particular, in our expansion (including all the cases in [19]) only a few such graphs may appear due to the condition $|n-n'|\leqslant 2$. The estimates leading to (7.16) in these cases are similar but much easier than in the general case and they are left to the reader (the same estimates were covered in [20] as well, without the renormalization of the dispersion relation). □

Sometimes we will use a numerical labelling of the edges (see Figure 2). In this case, we label the edge $(j-1,j)$ by e_j, and the edge $(\tilde{j},\widetilde{j-1})$ by $e_{\tilde{j}}$. At the special vertices 0 and 0^*, we denote the edges in the following way: $e_{n+1}:=(n,0^*)$, $e_{\widetilde{n'+1}}:=(0^*,\tilde{n}')$, $e_1=(0,1)$ and $e_{\tilde{1}}:=(\tilde{1},0)$. Therefore, the edge set $\mathcal{L}=\mathcal{L}(\mathcal{V}_{n,n'})$ is identified with the index set $I_{n+1}\cup\tilde{I}_{n'+1}$, and we set

$$p_j := w_{e_j} \quad \text{and} \quad \tilde{p}_j := w_{e_{\tilde{j}}}. \tag{7.18}$$

This notation will sometimes be used in parallel: the $\mathbf{p}$, $\tilde{\mathbf{p}}$ notation is preferred when distinction is needed between momenta on $\mathcal{L}$ and $\widetilde{\mathcal{L}}$ edges, whereas the $\mathbf{w}$ notation is used when no such distinction is necessary. Note that we always have

$$p_1-\tilde{p}_1=\xi. \tag{7.19}$$

7.3. Non-repetition Feynman graphs

A partition $\mathbf{P}\in\mathcal{P}_n$ of $I_n\cup\tilde{I}_n$ is called *even* if for any $P_\mu\in\mathbf{P}$ we have $|P_\mu\cap I_n|=|P_\mu\cap\tilde{I}_n|$. In particular, in an even partition there are no single lumps, i.e. $G(\mathbf{P})=\varnothing$.

Let $\mathfrak{S}_n$ be the set of permutations on I_n and let id be the identity permutation. Note that $\mathbf{A}\in\mathcal{A}_n$ and $\sigma\in\mathfrak{S}_n$ uniquely determine an even partition in $\mathbf{P}(\mathbf{A},\sigma)\in\mathcal{P}_n$, by $I(\mathbf{P}):=I(\mathbf{A})$ and $P_\mu:=A_\mu\cup\sigma(A_\mu)$.

Conversely, given an even partition $\mathbf{P}\in\mathcal{P}_n$, we can define its projection onto I_n, $\mathbf{A}:=\pi(\mathbf{P})\in\mathcal{A}_n$, by $I(\mathbf{A}):=I(\mathbf{P})$ and $A_\mu:=P_\mu\cap I_n$. We let

$$\mathfrak{S}_n(\mathbf{P}):=\{\sigma\in\mathfrak{S}_n:\mathbf{P}(\pi(\mathbf{P}),\sigma)=\mathbf{P}\}$$

be the set of permutations that are *compatible* with a given even partition $\mathbf{P}$. In other words, $\sigma\in\mathfrak{S}_n(\mathbf{P})$ if for each $i\in I_n$ the pair $(i,\sigma(i))$ belongs to the same $\mathbf{P}$-lump. Clearly

$$|\mathfrak{S}_n(\mathbf{P})|=\prod_{\mu\in I(\mathbf{P})}\Big(\frac{|P_\mu|}{2}\Big)!=\prod_{\mu\in I(\pi(\mathbf{A}))}|A_\mu|!\,. \tag{7.20}$$

We will use the notation

$$V_{(*)}(\mathbf{A},\sigma,Q):=V_{(*)}(\mathbf{P}(\mathbf{A},\sigma);Q), \tag{7.21}$$

and similarly for $E_{(*)g}$ and $V^\circ_{(*)}$. In the proofs, Q will be omitted. We also introduce

$$c(\mathbf{A}):=\prod_{\nu\in I(\mathbf{A})}c(|A_\nu|). \tag{7.22}$$

With this notation, we can state the representation of the non-repetition terms as a summation over Feynman diagrams.

PROPOSITION 7.2. *With $Q\equiv1$ and $\xi=0$ we have*

$$\lim_{L\to\infty}\mathbf{E}'\|\psi'^{\,nr}_{t,k}\|_L^2=\sum_{\sigma\in\mathfrak{S}_k}\sum_{\mathbf{A}\in\mathcal{A}_k}c(\mathbf{A})V^\circ_{\xi=0}(\mathbf{A},\sigma,Q\equiv1), \tag{7.23}$$

and with $Q_\xi(v):=\widehat{\mathcal{O}}(\xi,v)$ we have

$$\lim_{L\to\infty}\mathbf{E}'\langle\widehat{\mathcal{O}}_L,\widehat{W}^\varepsilon_{\psi'^{\,nr}_{t,k}}\rangle_L=\sum_{\sigma\in\mathfrak{S}_k}\sum_{\mathbf{A}\in\mathcal{A}_k}c(\mathbf{A})\int_{\mathbf{R}^d}V^\circ_{\varepsilon\xi}(\mathbf{A},\sigma,Q_\xi)\,d\xi. \tag{7.24}$$

Proof. We insert (6.6) and (6.9) into (6.7) and we take the limit as $L\to\infty$. We use the fact that

$$\mathbf{E}_M\bigg[\frac{M!}{|\Lambda_L|^k(M-k)!}\bigg]\to1$$

for any fixed k. We also replace every Riemann sum (3.23) with integrals and we use that $\widehat{\psi}'_0\to\widehat{\psi}_0$. By recalling (7.19) and by choosing $Q\equiv1$ in the definition (7.8), we obtain (7.23). The proof of (7.24) is identical. □

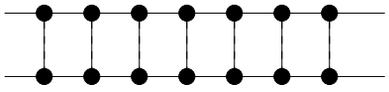

Figure 3. Ladder graph.

Proof of Theorem 5.2. We will only prove (5.1); the proof of (5.2) is analogous. Starting with (7.23), we notice that the graph with the trivial partition $\mathbf{A}_0$ and with the identity permutation on I_k gives the main term in Theorem 5.2, since

$$V_\lambda(t,k) = V^\circ_{\xi=0}(\mathbf{A}_0, \mathrm{id}).$$

This graph is called the *ladder graph* (Figure 3). All other graphs will be negligible.

We first replace V° by V; the error is negligible by Lemma 7.1. In §8 we then estimate $V(\mathbf{A},\sigma)$ for the trivial partition $\mathbf{A}=\mathbf{A}_0$, where every lump has one element. The result is Proposition 8.6. In this case we set $V(\sigma):=V(\mathbf{A}_0,\sigma)$. In §9 we treat the general case $\mathbf{A}\neq\mathbf{A}_0$. The final result of this section is Proposition 9.2. The proof of both propositions rely on Theorem 8.4, which is the core of our method. Its proof is given separately in §10. Finally, the proof of Theorem 5.2 follows from Proposition 9.2, together with (7.23), (7.16) and (6.11). □

We remark that the E- and V-values of the partitions depend on the parameters λ, t, ξ, ζ, k and g; a fact that is not explicitly included in the notation. In §§8–10 we will always assume the following relations:

$$\eta=\lambda^{2+\varkappa},\ \ t=\lambda^{-2-\varkappa}T,\ \ T\in[0,T_0],\ \ K=\lfloor\lambda^{-\delta}(\lambda^2 t)\rfloor,\ k<K,\ \ \zeta=\lambda^{-\varkappa-3\delta},\ \ g\leqslant 8, \tag{7.25}$$

for a sufficiently small $\delta>0$ that is independent of λ but depends on $\varkappa$. All estimates will be uniform in ξ and in $T\in[0,T_0]$. We mention that for the proof of Theorem 5.2 we need only $g=0$, but the more general case is used in [19].

8. Estimates on Feynman graphs without non-trivial lumps

We use the letters p_j and $\tilde{p}_j$, $j\in I_{k+1}$, for the momenta variables (see the convention at the end of §7.2) and $I(\mathbf{A}_0)=I_k$ for the index set of the trivial partition. In the following sections we always assume that $Q\equiv 1$.

We introduce the restricted version of M° (see (6.8)) as

$$M(k,\mathbf{p},\tilde{\mathbf{p}}) := \frac{\exp(2t\eta)}{(2\pi)^2}\int_{-Y}^{Y}\int_{-Y}^{Y}\exp(i(\alpha-\beta)t)\Bigg(\prod_{j=1}^{k+1}\frac{\overline{\widehat{B}(p_{j+1}-p_j)}}{\alpha-\overline{\omega(p_j)}-i\eta}\,\frac{\widehat{B}(\tilde{p}_{j+1}-\tilde{p}_j)}{\beta-\omega(\tilde{p}_j)+i\eta}\Bigg)\,d\alpha\,d\beta, \tag{8.1}$$

and we also define the trivial estimate of M as

$$N(k,\mathbf{p},\tilde{\mathbf{p}}) := \frac{\exp(2t\eta)}{(2\pi)^2}\int_{-Y}^{Y}\int_{-Y}^{Y}\Bigg(\prod_{j=1}^{k+1}\frac{|\widehat{B}(p_{j+1}-p_j)|}{|\alpha-\overline{\omega(p_j)}-i\eta|}\,\frac{|\widehat{B}(\tilde{p}_{j+1}-\tilde{p}_j)|}{|\beta-\omega(\tilde{p}_j)+i\eta|}\Bigg)\,d\alpha\,d\beta. \tag{8.2}$$

The truncated versions of these quantities, denoted by $M_*(k,\mathbf{p},\tilde{\mathbf{p}})$ and $N_*(k,\mathbf{p},\tilde{\mathbf{p}})$, are defined by removing the $(k+1)$-th α and β denominators from the definitions (8.1) and (8.2) but keeping all numerators and all other denominators.

From the definitions (7.12), (7.21) and $V(\sigma)=V(\mathbf{A}_0,\sigma)=V(\mathbf{A}_0,\sigma,Q\equiv 1)$, we obtain

$$V_{(*)}(\sigma) = \lambda^{2k}\int_{(\mathbf{R}^d)^{k+1}}\int_{(\mathbf{R}^d)^{k+1}} M_{(*)}(k,\mathbf{p},\tilde{\mathbf{p}})\Delta_\xi(\sigma,\mathbf{p},\tilde{\mathbf{p}},\mathbf{u}\equiv 0)\overline{\widehat{\psi}_0(p_1)}\widehat{\psi}_0(\tilde{p}_1)\,d\mathbf{p}\,d\tilde{\mathbf{p}}, \tag{8.3}$$

$$E_{(*)}(\sigma,\mathbf{u}) = \lambda^{2k}\int_{(\mathbf{R}^d)^{k+1}}\int_{(\mathbf{R}^d)^{k+1}} N_{(*)}(k,\mathbf{p},\tilde{\mathbf{p}})\Delta_\xi(\sigma,\mathbf{p},\tilde{\mathbf{p}},\mathbf{u})\overline{\widehat{\psi}_0(p_1)}\widehat{\psi}_0(\tilde{p}_1)\,d\mathbf{p}\,d\tilde{\mathbf{p}}, \tag{8.4}$$

with

$$\Delta_\xi(\sigma,\mathbf{p},\tilde{\mathbf{p}},\mathbf{u}) := \delta(\tilde{p}_{k+1}-p_{k+1}+\xi)\prod_{l=1}^{k}\delta(p_{l+1}-p_l-(\tilde{p}_{\sigma(l)+1}-\tilde{p}_{\sigma(l)})-u_l).$$

Clearly $|V_{(*)}(\sigma)|\leqslant E_{(*)g=0}(\sigma,\mathbf{u}\equiv 0)$ for any ξ (see (7.14)).

We introduce a convenient notation. For any $(k+1)\times(k+1)$ matrix M and for any vector of momenta $\mathbf{p}=(p_1,...,p_{k+1})$, we let $M\mathbf{p}$ denote the following $(k+1)$-vector of momenta

$$M\mathbf{p} := \Bigg(\sum_{j=1}^{k+1}M_{1j}p_j, \sum_{j=1}^{k+1}M_{2j}p_j, ...\Bigg). \tag{8.5}$$

Furthermore, we introduce the vector $\mathbf{v}=(v_1,...,v_{k+1})$ as

$$v_l := \xi+u_1+u_2+...+u_{l-1} \quad \text{for all } l=1,2,...,k+1. \tag{8.6}$$

Note that $v_{k+1}=\xi$ by (7.1).

Given a permutation $\sigma\in\mathfrak{S}_k$, we define a $(k+1)\times(k+1)$ matrix $M=M(\sigma)$ as follows

$$M_{hj}(\sigma) := \begin{cases} 1, & \text{if } \tilde{\sigma}(j-1)<h\leqslant\tilde{\sigma}(j),\\ -1, & \text{if } \tilde{\sigma}(j)<h\leqslant\tilde{\sigma}(j-1),\\ 0, & \text{otherwise,}\end{cases} \tag{8.7}$$

where, by definition, $\tilde{\sigma}$ is the *extension* of σ to a permutation of $\{0,1,...,k+1\}$ by $\tilde{\sigma}(0):=0$ and $\tilde{\sigma}(k+1):=k+1$. In particular $[M\mathbf{p}]_1=p_1$ and $[M\mathbf{p}]_{k+1}=p_{k+1}$.

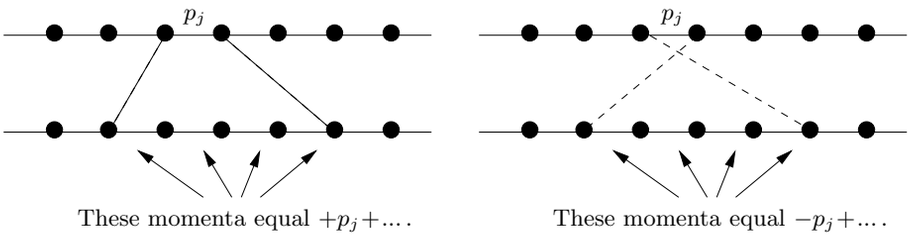

Figure 4. Domain of dependencies of the momenta.

It is easy to check that

$$\Delta_\xi(\sigma,\mathbf{p},\tilde{\mathbf{p}},\mathbf{u})=\prod_{j=1}^{k+1}\delta(\tilde{p}_j-[M\mathbf{p}]_j+[M\mathbf{v}]_j), \tag{8.8}$$

in other words, the matrix M encodes the dependence of the $\tilde{p}$-momenta on the p-momenta and the v-momenta. This rule is transparent in the graphical representation of the Feynman graph: the momentum p_j appears in those $\tilde{p}_h$'s which fall into its "domain of dependence", i.e. the section between the image of the two endpoints of p_j, and the sign depends on the ordering of these images (Figure 4).

Definition 8.1. A matrix M with entries 0, $+1$ or -1 is called *tower matrix* if in each column the non-zero entries are consecutive and identical. The collection of these consecutive 1 or -1 entries are called the *tower* of that column.

By construction, the matrix $M(\sigma)$ is a tower matrix.

PROPOSITION 8.2. *For any permutation $\sigma\in\mathfrak{S}_k$ the matrix $M(\sigma)$ is*
(i) *invertible*;
(ii) *totally unimodular*, *i.e. each subdeterminant is* 0 *or* ±1.

Proof. The invertibility follows from the fact that $\mathbf{p}$ and $\tilde{\mathbf{p}}$ play symmetric roles in (8.8) if $\mathbf{v}\equiv0$ and $\xi=0$, in particular $M(\sigma)^{-1}=M(\sigma^{-1})$. It is easy to prove, by induction on the size of the matrix, that any tower matrix is totally unimodular. □

The following definition is crucial. It establishes the necessary concepts to measure the complexity of a permutation.

Definition 8.3. (Valley, peak, slope and ladder) Given a permutation $\sigma\in\mathfrak{S}_k$ let $\tilde{\sigma}$ be its extension. A point $(j,\sigma(j))$, $j\in I_k=\{1,2,...,k\}$, on the graph of σ is called *peak* if $\sigma(j)<\min\{\tilde{\sigma}(j-1),\tilde{\sigma}(j+1)\}$, it is called *valley* if $\sigma(j)>\max\{\tilde{\sigma}(j-1),\tilde{\sigma}(j+1)\}$. Furthermore, if $\sigma(j)-1\in\{\tilde{\sigma}(j-1),\tilde{\sigma}(j+1)\}$ and $(j,\sigma(j))$ is not a valley, then the point

$(j,\sigma(j))$, $j\in I_k$, is called *ladder*. Finally, a point $(j,\sigma(j))$, $j\in I_k$, on the graph of σ is called *slope* if it is not a peak, a valley or a ladder.

Let $I=\{1,2,...,k+1\}$ denote the set of row indices of M. This set is partitioned into five disjoint subsets, $I=I_p\cup I_v\cup I_l\cup I_s\cup I_{\rm last}$, such that $I_{\rm last}:=\{k+1\}$ is the last index, and $h\in I_p, I_v, I_l, I_s$ if $(\sigma^{-1}(h),h)$ is a peak, valley, ladder and slope, respectively. The cardinalities of these sets are denoted by $p:=|I_p|$, $v:=|I_v|$, $l:=|I_l|$ and $s:=|I_s|$, and, if necessary, we indicate the dependence on σ as $p=p(\sigma)$, etc. We define the *degree* of the permutation σ as

$$\deg(\sigma):=k-l(\sigma). \tag{8.9}$$

A maximal collection of consecutive ladder indices, $h+1,...,h+b\in I_l$, is called a *ladder* of length b. The index h is called the *top index of a ladder*. The *bottom index of a ladder* is defined to be $h+b$ or $h+b+1$, depending on whether

$$|\tilde\sigma^{-1}(h+b+1)-\sigma^{-1}(h+b)|\neq 1 \quad \text{or} \quad |\tilde\sigma^{-1}(h+b+1)-\sigma^{-1}(h+b)|=1,$$

respectively. The set of bottom and top indices are denoted by I_b and I_t. Note that $I_t\subset\{0,1,...,k-1\}$, $I_b\subset\{1,2,...,k+1\}$, and that the top index of a ladder never belongs to I_l and it may be 0. The bottom index is either a ladder, a valley or $k+1$.

Remarks. (i) The terminology of peak, valley, slope and ladder comes from the graph of the permutation $\tilde\sigma$ drawn in a coordinate system where the axis of the dependent variable, $\sigma(j)$, is oriented downward (see Figure 5). It immediately follows from the definition of the extension $\tilde\sigma$ that the number of peaks and valleys are the same, $p(\sigma)=v(\sigma)$.

(ii) The non-zero entries in the matrix $M(\sigma)$ follow the same geometric pattern as the graph: each downward segment of the graph corresponds to a column with a few consecutive 1's, upward segments correspond to columns with (-1)'s. On Figure 5 we also pictured the towers of $M(\sigma)$ drawn inside the graph of σ.

(iii) Because our choice of orientation of the vertical axis follows the convention of labelling rows of a matrix, a peak is a local minimum of $j\mapsto\sigma(j)$. We fix the convention that the notions "higher" or "lower" for objects related to the vertical axis (e.g. row indices) always refer to the graphical picture. In particular the "bottom" or the "lowest element" of a tower is located in the row with the highest index.

Also, a point on the graph of the function $j\mapsto\sigma(j)$ is traditionally denoted by $(j,\sigma(j))$, where the first coordinate j runs on the horizontal axis, while in the labelling of the (h,j)-matrix element M_{hj} of a matrix M the first coordinate h labels rows, i.e. it runs vertically. To avoid confusion, we will always specify whether a double index (h,j) refers to a point on the graph of σ or a matrix element.

(iv) We note that for the special case of the identity permutation $\sigma=\mathrm{id}$ we have $I_p=I_s=I_v=\varnothing$, and $I_l=\{1,2,...,k\}$. In particular, $\deg(\mathrm{id})=0$ and $\deg(\sigma)\geqslant 2$ for any permutation $\sigma\neq\mathrm{id}$.

An example, with $k=8$, is shown in Figure 5. The matrix corresponding to the permutation on this figure is the following (zero entries are left empty):

$$M(\sigma):=\begin{pmatrix} 1 & & & & & & & & \\ & 1 & & & & & & & \\ & & 1 & & & & & & \\ & & 1 & & & -1 & 1 & & \\ & & 1 & & & -1 & & 1 & \\ & & 1 & & -1 & & & 1 & \\ & & 1 & -1 & & & & 1 & \\ & & & & & & & 1 & \\ & & & & & & & & 1 \end{pmatrix}\begin{matrix} 1 & l \\ 2 & l \\ 3 & p \\ 4 & l \\ 5 & s \\ 6 & l \\ 7 & v \\ 8 & s \\ 9 & (\text{last}). \end{matrix} \tag{8.10}$$

The numbers on the right indicate the column indices and the letters show whether it is peak, slope, valley, ladder or last. In this case $I_p=\{3\}$, $I_v=\{7\}$, $I_s=\{5,8\}$, $I_l=\{1,2,4,6\}$, $I_{\mathrm{last}}=\{9\}$, $I_t=\{0,3,5\}$, $I_b=\{2,4,7\}$ and $\deg(\sigma)=4$. There are three ladders, two of them of length 1 and one of length 2.

Now we are ready to estimate $|V(\sigma)|\leqslant E_{g=0}(\sigma,\mathbf{u}\equiv 0)$. The following theorem shows that the degree of the permutation, $\deg(\sigma)$, measures the size of $V(\sigma)$. The proof is the key step in our method and it is given in §10.

THEOREM 8.4. *Assume* (7.25) *with* $\varkappa<2/(6+9d)$ *and let* $\sigma\in\mathfrak{S}_k$. *Then the E-value of the graph of the trivial partition with permutation* σ *is estimated by*

$$\sup_{\mathbf{u}} E_{(*)g}(\sigma,\mathbf{u})\leqslant C(\lambda^{1/3-(1+3d/2)\varkappa-O(\delta)})^{\deg(\sigma)}|\log\lambda|^2 \tag{8.11}$$

if $\lambda\ll 1$.

This theorem is complemented by the following combinatorial lemma.

LEMMA 8.5. *Let* $k\leqslant K=O(\lambda^{-\varkappa-\delta})$, *let* $D\geqslant 0$ *be an integer, and let* $\gamma>\varkappa+\delta$. *Then*

$$\sum_{\substack{\sigma\in\mathfrak{S}_k\\ \deg(\sigma)\geqslant D}} \lambda^{\gamma\deg(\sigma)}\leqslant O(\lambda^{D(\gamma-\varkappa-\delta)}) \tag{8.12}$$

for $\lambda\ll 1$.

Since $\deg(\sigma)\geqslant 2$ if $\sigma\neq\mathrm{id}$, from Theorem 8.4, Lemma 8.5, $g(\mathbf{P})=0$ and the estimate $|V(\sigma)|\leqslant E_{g=0}(\sigma,\mathbf{u}\equiv 0)$, we immediately obtain the following result.

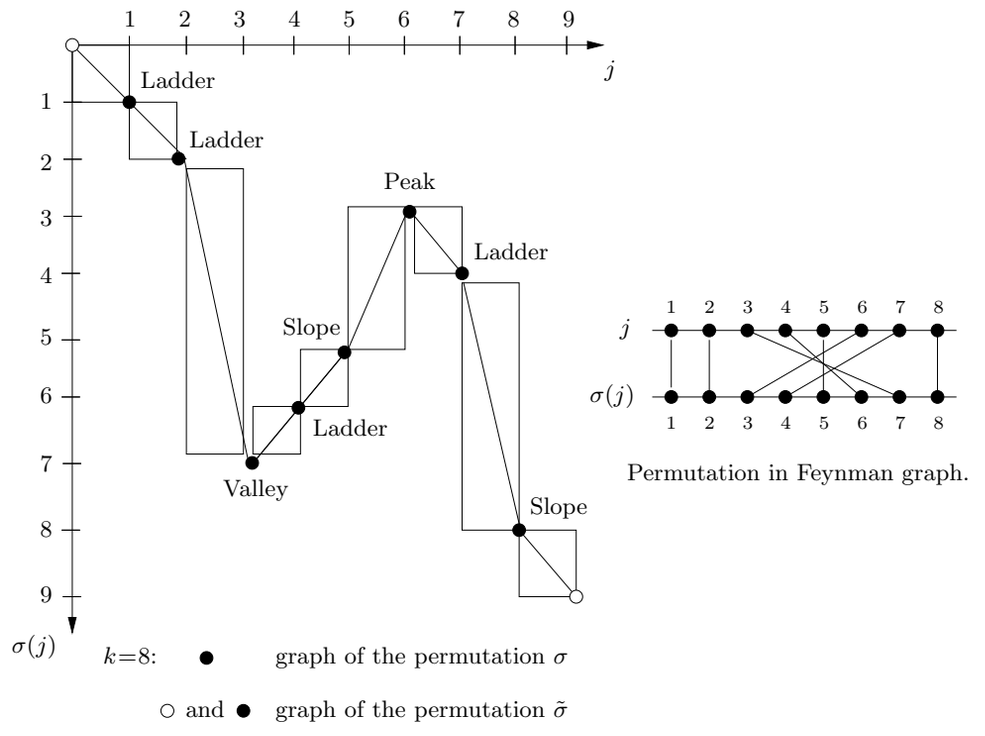

Figure 5. Graph of a permutation with the towers.

PROPOSITION 8.6. *Assuming* (7.25) *with* $\varkappa<2/(12+9d)$, *we have*

$$\sum_{\substack{\sigma\in\mathfrak{S}_k\\ \sigma\neq\mathrm{id}}} |V(\sigma)| \leqslant O(\lambda^{2/3-(4+3d)\varkappa-O(\delta)}) \tag{8.13}$$

for $\lambda\ll 1$.

Proof of Lemma 8.5. Notice that $l(\sigma)=k$ only if $\sigma=\mathrm{id}$, being $l(\sigma)\leqslant k-2$ for all other permutations. We shall prove that, for any l,

$$|\{\sigma\in\mathfrak{S}_k : l(\sigma)=l\}| \leqslant 2(2k)^{k-l}. \tag{8.14}$$

Then (8.12) follows by recalling that $k-l(\sigma)=\deg(\sigma)$, that $k\leqslant K=O(\lambda^{-\varkappa-\delta})$ and by summing up the geometric series

$$\sum_{\substack{\sigma\in\mathfrak{S}_k\\ \deg(\sigma)\geqslant D}} \lambda^{\gamma\deg(\sigma)} \leqslant 2\sum_{m=D}^{k} [2\lambda^{\gamma-\varkappa-\delta}]^m.$$

To prove (8.14), let σ be a permutation with m ladders of size $b_1, ..., b_m$ such that $\sum_{j=1}^m b_j = l$ and $b_j \geqslant 1$. If we remove these ladder indices, we have $k-l$ indices in $\{1, 2, ..., k\} \setminus I_l = \{h_1, h_2, ..., h_{k-l}\}$. The permutation σ induces a unique permutation $\sigma^* \in \mathfrak{S}_{k-l}$ on the indices of this set by $\sigma^*(j) < \sigma^*(j')$ if and only if $\sigma(h_j) < \sigma(h_{j'})$.

Let $I^* := \{0, h_1, h_2, ..., h_{k-l}\}$ and set $\nu := |I^*| = k-l+1$. Clearly, the top of any ladder belongs to I^* and each element of I^* can be the top of at most one ladder. We assign the length b_j of the ladder to its top and, for simplicity, we assign the value zero to any other element of I^*. Thus we obtain numbers $b_1, ..., b_\nu$ with $\sum_{j=1}^\nu b_j = l$ and $b_j \geqslant 0$.

If the permutation σ^* and the numbers $b_1, ..., b_\nu$ are given, then we have 2^ν ways to reconstruct the original permutation σ. To see this, first notice that within I^* the ladder-tops are identified by the condition $b_j > 0$ and the corresponding ladder in σ can emanate either "to the right" or "to the left" down from its top on the graph of σ. Once this choice is made, the permutation σ can be uniquely reconstructed from σ^* by inserting ladders of given length starting from their tops. Therefore, the number of permutations $\sigma \in \mathfrak{S}_k$ with l ladder indices is bounded by

$$2^\nu (k-l)! \left| \left\{ (b_1, b_2, ..., b_\nu) : \sum_{j=1}^{\nu} b_j = l \text{ and } b_j \geqslant 0 \right\} \right| \leqslant 2^\nu (k-l)! \binom{\nu-1+l}{\nu-1} \leqslant 2(2k)^{k-l}.$$

This completes the proof of Lemma 8.5. □

9. Estimates on Feynman graphs with non-trivial lumps

In this section we estimate $V(\mathbf{A}, \sigma)$ for a general partition $\mathbf{A}$. We start with a definition.

Definition 9.1. (i) Let $\mathbf{A} \in \mathcal{A}_k$. Set $a_\nu := |A_\nu|$, $\nu \in I(\mathbf{A})$, to be the size of the νth lump. Let

$$S(\mathbf{A}) := \bigcup_{\substack{\nu \in I(\mathbf{A}) \\ a_\nu \geqslant 2}} A_\nu$$

be the union of the non-trivial lumps. The cardinality of this set, $s(\mathbf{A}) := |S(\mathbf{A})|$, is called the *degree* of the partition $\mathbf{A}$.

(ii) Let $\mathbf{A} \in \mathcal{A}_k$ and $\sigma \in \mathfrak{S}_k$. The number

$$q(\mathbf{A}, \sigma) := \max\{\deg(\sigma), \tfrac{1}{2} s(\mathbf{A})\} \tag{9.1}$$

is called the *joint degree* of the pair $(\sigma, \mathbf{A})$.

The goal is the following generalization of Proposition 8.6 that includes summations over non-trivial lumps.

PROPOSITION 9.2. *We assume* (7.25). *Let* $D\geqslant 0$ *and* $s\geqslant 2$ *be given integers, and let* $q:=\max\{D,\frac{1}{2}s\}$. *For any* $\varkappa<2/(33d+36)$ *we have*

$$\Xi(k,D,s):=\sum_{\substack{\sigma\in\mathfrak{S}_k\\ \deg(\sigma)\geqslant D}}\sum_{\substack{\mathbf{A}\in\mathcal{A}_k\\ s(\mathbf{A})\geqslant s}}\sup_{\mathbf{u},g\leqslant 8}E_{(*)g}(\mathbf{A},\sigma,\mathbf{u})|c(\mathbf{A})|\leqslant C(\lambda^{1/3-(6+11d/2)\varkappa-O(\delta)})^q|\log\lambda|^2. \tag{9.2}$$

We recall that for the continuum model non-trivial lumps do not contribute since $c(\mathbf{A})=0$ unless each A_ν is trivial, $|A_\nu|=1$. Thus (9.2) can be proved directly from Theorem 8.4, Lemma 8.5 and Lemma 9.5 below with a somewhat better exponent. After some preparatory lemmas, we will give a proof of Proposition 9.2 for the general case that is valid for both the discrete and the continuum models.

The following lemma shows that any even partition $\mathbf{P}\in\mathcal{P}_k$ can be generated by a permutation with high degree, depending on the size of non-trivial lumps.

LEMMA 9.3. *For any even partition* $\mathbf{P}\in\mathcal{P}_k$ *there exists a compatible permutation* $\hat\sigma=\hat\sigma(\mathbf{P})\in\mathfrak{S}_k(\mathbf{P})$ *such that*

$$\deg(\hat\sigma)\geqslant\tfrac{1}{2}s(\pi(\mathbf{P})). \tag{9.3}$$

COROLLARY 9.4. *Given* $\sigma\in\mathfrak{S}_k$ *and* $\mathbf{A}\in\mathcal{A}_k$, *we have, for* $\varkappa<2/(6+33d)$,

$$\sup_{\mathbf{u}}E_{(*)g}(\mathbf{A},\sigma,\mathbf{u})\leqslant C|\log\lambda|^2(\lambda^{1/3-(1+11d/2)\varkappa-O(\delta)})^{q(\mathbf{A},\sigma)}. \tag{9.4}$$

Proof. We define a permutation $\sigma^*:=\sigma^*(\mathbf{A},\sigma)$ by

$$\sigma^*:=\begin{cases}\sigma, & \text{if } \deg(\sigma)\geqslant\frac{1}{2}s(\mathbf{A}),\\ \hat\sigma(\mathbf{P}(\mathbf{A},\sigma)), & \text{otherwise.}\end{cases}$$

By Lemma 9.3, we have $\deg(\sigma^*)=q(\mathbf{A},\sigma)$. Clearly $\mathbf{P}(\mathbf{A},\sigma)=\mathbf{P}(\mathbf{A},\sigma^*)$; in particular, $E_{(*)g}(\mathbf{A},\sigma,\mathbf{u})=E_{(*)g}(\mathbf{A},\sigma^*,\mathbf{u})$.

We wish to estimate the value of an arbitrary partition $\mathbf{A}$ by that of the trivial partition $\mathbf{A}_0$. We can artificially break up lumps and use the auxiliary momenta to account for the additional Kirchoff rules. We describe this procedure in full generality for any circle graph. We will call it *Operation I* because further similar operations will be introduced in the companion paper [19].

Operation I. Breaking up lumps.

Consider a circle graph on N vertices (§7.1). Given a partition of the set $\mathcal{V}\setminus\{0,0^*\}$, $\mathbf{P}=\{P_\mu:\mu\in I(\mathbf{P})\}\in\mathcal{P}_\mathcal{V}$, we define a new partition $\mathbf{P}^*$ by breaking up one of the lumps into two smaller non-empty lumps. Let $P_\nu=P_{\nu'}\cup P_{\nu''}$ with $P_{\nu'}\cap P_{\nu''}=\varnothing$ and

$$\mathbf{P}^*=\{P_{\nu'},P_{\nu''},P_\mu:\mu\in I(\mathbf{P})\setminus\{\nu\}\}.$$

In particular $I(\mathbf{P}^*)=I(\mathbf{P})\cup\{\nu',\nu''\}\setminus\{\nu\}$ and $m(\mathbf{P}^*)=m(\mathbf{P})+1$.

LEMMA 9.5. *With the notation above, we have*

$$E_{(*)g}(\mathbf{P},\mathbf{u},\boldsymbol{\alpha}) \leqslant \int_{|r|\leqslant N\zeta} E_{(*)g}(\mathbf{P}^*,\mathbf{u}^*(r,\nu),\boldsymbol{\alpha})\,dr,$$

where the new set of momenta $\mathbf{u}^*=\mathbf{u}^*(r,\nu)$ *is given by* $u^*_\mu := u_\mu$, *for* $\mu\in I(\mathbf{P})\setminus\{\nu\}$, $u^*_{\nu'}=u_\nu - r$ *and* $u^*_{\nu''}=r$. *In our estimates we will always have* $N\leqslant 2K$ *and then*

$$\sup_{\mathbf{u}} E_{(*)g}(\mathbf{P},\mathbf{u},\boldsymbol{\alpha}) \leqslant \Lambda \sup_{\mathbf{u}} E_{(*)g}(\mathbf{P}^*,\mathbf{u},\boldsymbol{\alpha})$$

with $\Lambda := [CK\zeta]^d = O(\lambda^{-2d\varkappa - O(\delta)})$ (*see* (4.7) *and* (7.4)).

Proof. The break-up of the lump P_ν corresponds to

$$\delta\Big(\sum_{e\in L_\pm(P_\nu)} \pm w_e - u_\nu\Big) = \int_{\mathbf{R}^d} \delta\Big(\sum_{e\in L_\pm(P_{\nu'})} \pm w_e - u_\nu + r\Big)\delta\Big(\sum_{e\in L_\pm(P_{\nu''})} \pm w_e - r\Big)\,dr. \quad (9.5)$$

Note that $L(P_\nu)\subset L(P_{\nu'})\cup L(P_{\nu''})$ and, for any edge $e\in(L(P_{\nu'})\cup L(P_{\nu''}))\setminus L(P_\nu)$, we inserted an extra $w_e - w_e$ in the left-hand side of (9.5). Note that the property (7.1) on the sum of auxiliary momenta is preserved. The integration in (9.5) can be restricted to $|r|\leqslant N\zeta$, since $|w_e|\leqslant\zeta$ for all e. □

We return to the proof of Corollary 9.4. We apply the break-up operation until all lumps become trivial (singlets). This requires not more than $s(\mathbf{A})$ steps, and, by using Lemma 9.5, we obtain that

$$\sup_{\mathbf{u}} E_{(*)g}(\mathbf{A},\sigma^*,\mathbf{u}) \leqslant \Lambda^{s(\mathbf{A})} \sup_{\mathbf{u}} E_{(*)g}(\mathbf{A}_0,\sigma^*,\mathbf{u}),$$

with $\Lambda := [CK\zeta]^d$. Then (9.4) immediately follows from $s(\mathbf{A})\leqslant 2q(\mathbf{A},\sigma)$ and from Theorem 8.4. This completes the proof of Corollary 9.4. □

Proof of Lemma 9.3. Inequality (9.3) is trivial if the partition has no non-trivial lumps, i.e. $s(\pi(\mathbf{P}))=0$, so we may assume that $s(\pi(\mathbf{p}))\neq 0$. For any $\sigma\in\mathfrak{S}_k$ we define the set of *internal ladder indices* as

$$I^*_l = I^*_l(\sigma) := \{h\in I_l : |\tilde\sigma^{-1}(h-1) - \tilde\sigma^{-1}(h)| = |\tilde\sigma^{-1}(h+1) - \tilde\sigma^{-1}(h)| = 1\},$$

where $\tilde\sigma$ is the extension of σ as before. The indices $h-1$ and $h+1$ are called the *protectors* of the internal ladder index $h\in I^*_l$. They ensure that the index h is neither the bottom nor the top index of the ladder.

We first claim that for any $\sigma\in\mathfrak{S}_k$, $\sigma\neq\mathrm{id}$, we have

$$k - |I^*_l(\sigma)| \leqslant 2\mathrm{deg}(\sigma). \quad (9.6)$$

To see this inequality, we first recall the definition of a ladder and its bottom and top indices from Definition 8.3. Since every ladder has a unique bottom and top index, which are not internal ladder indices, we see that the sets I_l^*, I_b and I_t are disjoint subsets of $\{0,1,...,k+1\}$. Since $I_v, I_l^* \subset \{1,2,...,k\}$, $I_v \cap I_l^* = I_v \cap I_t = \varnothing$ and $k+1 \notin I_t$, we obtain that $|I_l^*|+|I_t|+|I_b|+|I_v \setminus I_b| \leqslant k+1+\gamma$, where γ is the characteristic function of the event $k+1 \in I_b$, i.e. $\gamma := |\{k+1\} \cap I_b|$. Moreover, $|I_t| = |I_b|$, thus

$$|I_l^*|+2|I_b|+|I_v \setminus I_b| \leqslant k+1+\gamma.$$

Notice that $I_l \subset I_l^* \cup (I_b \setminus I_v)$, and $k+1 \notin I_l, I_v$, therefore

$$l \leqslant |I_l^*|+|I_b \setminus I_v|-\gamma = |I_l^*|+|I_b|-|I_v \cap I_b|-\gamma.$$

From the last two inequalities and from $\deg(\sigma) = k-l$, we obtain that

$$k-|I_l^*| \leqslant 2\deg(\sigma)+1-\gamma-|I_v \setminus I_b|-2|I_v \cap I_b| \leqslant 2\deg(\sigma)+1-|I_v|.$$

Since $\sigma \neq \mathrm{id}$, we have $|I_v| \geqslant 1$, which proves (9.6).

Next we will show that there exists a compatible permutation $\hat{\sigma} \in \mathfrak{S}_k(\mathbf{P})$,

$$I_l^*(\hat{\sigma}) \cap \hat{\sigma}(S(\mathbf{A})) = \varnothing, \tag{9.7}$$

where we set $\mathbf{A} := \pi(\mathbf{P})$ for simplicity. Since $S(\mathbf{A}) \neq \varnothing$, we have $I_l^*(\tilde{\sigma}) \neq \{1,2,...,k\}$, i.e. $\hat{\sigma} \neq \mathrm{id}$. Combining (9.6) with (9.7) and with the fact that both I_l^* and $\hat{\sigma}(S(\mathbf{A}))$ are subsets of $\{1,2,...,k\}$, we obtain (9.3).

To construct $\hat{\sigma}$ satisfying (9.7), we apply a greedy algorithm. Since $\mathbf{P}$ is even, $\mathfrak{S}_k(\mathbf{P})$ is non-empty and we pick $\sigma_0 \in \mathfrak{S}_k(\mathbf{P})$. If (9.7) is not satisfied for σ_0, then some internal ladder index h is in the image of a non-trivial lump $A \in \mathbf{A}$: $h = \sigma_0(h')$, $h' \in A$. Let $j' \in A$ be another element of this lump. Flip the permutation σ_0 on these two elements, i.e. define a new permutation σ_1 such that $\sigma_1(h') := \sigma_0(j') = j$, $\sigma_1(j') := \sigma_0(h') = h$ and $\sigma_1(r) := \sigma_0(r)$ for any $r \neq h, j$. Clearly $\sigma_1 \in \mathfrak{S}_k(\mathbf{P})$ and $\sigma_1(s(\mathbf{A})) = \sigma_0(s(\mathbf{A}))$. We claim that

$$|I_l^*(\sigma_1) \cap \sigma_1(S(\mathbf{A}))| < |I_l^*(\sigma_0) \cap \sigma_0(S(\mathbf{A}))|, \tag{9.8}$$

i.e. the total number of internal ladder indices in non-trivial lumps decreased. Continuing this flipping process for σ_1 etc., we obtain a permutation $\hat{\sigma}$ satisfying (9.7).

To see (9.8) we note that, after the flip, the index h is not an internal ladder index anymore. This is clear if $j \neq h-1, h+1$; in this case the points $(\tilde{\sigma}^{-1}(h-1), h-1)$ and $(\tilde{\sigma}^{-1}(h+1), h+1)$ have not changed and they would uniquely fix the location of an internal ladder index in between. The preimage of the index h has moved out from this

position, $\sigma_1^{-1}(h)\neq\sigma_0^{-1}(h)$. The index j however would not become internal ladder, since $\sigma_1^{-1}(j)=h'$ is between $\sigma_1^{-1}(h-1)$ and $\sigma_1^{-1}(h+1)$, but j is not between $h-1$ and $h+1$. It is easy to see that the fixed points $(\tilde\sigma^{-1}(h-1),h-1)$ and $(\tilde\sigma^{-1}(h+1),h+1)$ also prevent any other indices from becoming an internal ladder index after the flip. This could only be possible if due to the new point $(\tilde\sigma_1^{-1}(j),j)=(\tilde\sigma_0^{-1}(h),j)$, one of the neighbors of j, say $j+1$, would become an internal ladder index. It is easy to see that then $j+1$ must be equal to $h-1$, and the other protector of the new internal ladder index $j+1$ must be h. In this case $h-1$ was already an internal ladder index before the flip as well, so no new internal ladder was created.

A similar but simpler argument shows that if $j=h-1$ or $j=h+1$, the number of internal ladder indices also decreases. This completes the proof of Lemma 9.3. □

Proof of Proposition 9.2. Given $\sigma\in\mathfrak{S}_k$ and $\mathbf{A}\in\mathcal{A}_k$, we recall the definition of the permutation $\sigma^*=\sigma^*(\mathbf{A},\sigma)$ defined in the proof of Corollary 9.4. We also note that $s(\mathbf{A})\leqslant 2\deg(\sigma^*)$. Hence

$$\Xi(k,D,s)\leqslant \sum_{\substack{\mathbf{A}\in\mathcal{A}_k\\ s(\mathbf{A})\geqslant s}}|c(\mathbf{A})|\sum_{\substack{\sigma^*\in\mathfrak{S}_k\\ \deg(\sigma^*)\geqslant q}}\sum_{\substack{\sigma\in\mathfrak{S}_k\\ \sigma^*(\mathbf{A},\sigma)=\sigma^*}}\sup_{\mathbf{u}}E_{(*)g}(\mathbf{A},\sigma,\mathbf{u}).$$

Note that $\sigma\in\mathfrak{S}_k(\mathbf{P}(\mathbf{A},\sigma^*))$, so, by (7.20), the summation over σ contributes by a factor of at most $\prod_\nu a_\nu!$ and we obtain that

$$\Xi(k,D,s)\leqslant \sum_{\substack{\sigma^*\in\mathfrak{S}_k\\ \deg(\sigma^*)\geqslant q}}\sum_{\substack{\mathbf{A}\in\mathcal{A}_k\\ 2\leqslant s(\mathbf{A})\leqslant 2\deg(\sigma^*)}}\Big(\prod_{\nu\in I(\mathbf{A})}a_\nu^{a_\nu-2}a_\nu!\Big)\sup_{\mathbf{u}}E_{(*)g}(\mathbf{A},\sigma^*,\mathbf{u}).$$

We also used the estimate (6.11). By using (9.4) and $\deg(\sigma^*)=q(\mathbf{A},\sigma)$, we obtain that

$$\Xi(k,D,s)\leqslant \sum_{\substack{\sigma^*\in\mathfrak{S}_k\\ \deg(\sigma^*)\geqslant q}}(C\lambda^{1/3-(1+11d/2)\varkappa-O(\delta)})^{\deg(\sigma^*)}\sum_{\substack{\mathbf{A}\in\mathcal{A}_k\\ s\leqslant s(\mathbf{A})\leqslant 2\deg(\sigma^*)}}\Big(\prod_{\nu\in I(\mathbf{A})}a_\nu^{a_\nu-2}a_\nu!\Big).$$

We introduce the notation

$$\sum\nolimits^* f(a_\nu):=\sum_{\substack{\nu\in I(\mathbf{A})\\ a_\nu\geqslant 2}}f(a_\nu)\quad\text{and}\quad\prod\nolimits^* f(a_\nu):=\prod_{\substack{\nu\in I(\mathbf{A})\\ a_\nu\geqslant 2}}f(a_\nu).$$

First we fix the sizes of the non-trivial lumps $a_\nu\geqslant 2$. Given these sizes, the number of $\mathbf{A}$ partitions is bounded by

$$\binom{k}{a_1}\binom{k-a_1}{a_2}\binom{k-a_1-a_2}{a_3}\ldots\leqslant\frac{k!}{(k-\sum^* a_\nu)!\prod^* a_\nu!}\leqslant\frac{k^{\sum^* a_\nu}}{\prod^* a_\nu!}.$$

Recalling that $s(\mathbf{A})=\sum^* a_\nu$ and $s(\mathbf{A})\leqslant 2\deg(\sigma^*)$, we have

$$\Xi(k,D,s)\leqslant \sum_{\substack{\sigma^*\in\mathfrak{S}_k\\ \deg(\sigma^*)\geqslant q}} (Ck^2\lambda^\tau)^{\deg(\sigma^*)} \sum_{a_\nu:\sum^* a_\nu\leqslant 2\deg(\sigma^*)} \prod{}^* a_\nu^{a_\nu-2}, \tag{9.9}$$

with $\tau<\frac{1}{3}-\big(1+\frac{11}{2}d\big)\varkappa$. We use the bound $a^{a-2}\leqslant C^{a-1}(a-1)!$. To estimate the summation over the a_ν's, we use the following inequality. For any fixed m and H, we have

$$\sum{}^{\#}\prod{}^*(a_\nu-1)!\leqslant (H-1)!\,, \tag{9.10}$$

where the summation # is over all sequences $(a_1,a_2,...,a_m)$ of positive integers at least 2, whose sum is H. The proof of (9.10) is easily obtained by induction on m from

$$\sum_{a=2}^{H-2}(a-1)!\,(H-a-1)!\leqslant \sum_{a=2}^{H-2}(H-2)!<(H-1)!\,.$$

Summing (9.10) over all $H\leqslant 2\deg(\sigma^*)$ and all $m\leqslant\frac{1}{2}H$, we obtain the bound

$$\sum_{a_\nu:\sum^* a_\nu\leqslant 2\deg(\sigma^*)} \prod{}^* a_\nu^{a_\nu-2}\leqslant 2[2\deg(\sigma^*)]!\leqslant (Ck)^{2\deg(\sigma^*)}$$

for the a_ν summation in (9.9), since $\deg(\sigma^*)\leqslant k+1$ by definition.

In summary, we obtain from (9.9) that

$$\Xi(k,D,s)\leqslant \sum_{\substack{\sigma\in\mathfrak{S}_k\\ \deg(\sigma)\geqslant q}} (Ck^4\lambda^\tau)^{\deg(\sigma)}.$$

Recalling that $k=O(\lambda^{-\varkappa-\delta})$, we can apply Lemma 8.5 with $\gamma=\tau-4(\varkappa+\delta)$ as long as $\gamma>\varkappa+\delta$. For sufficiently small positive δ, this gives the condition $\varkappa<2/(33d+36)$ in Proposition 9.2 and the estimate (9.2). □

10. Proof of Theorem 8.4

For any $(k+1)\times(k+1)$ matrix M we set

$$\begin{aligned} E(M):=\sup_{\tilde{\mathbf{u}}}\lambda^{2k}\int_{-Y}^{Y}\int_{-Y}^{Y}\int_{(\mathbf{R}^d)^{k+1}} &|\mathcal{B}(\mathbf{p})|\,|\mathcal{B}(M\mathbf{p}+\tilde{\mathbf{u}})|\,|\widehat{\psi}(p_1)|\,|\widehat{\psi}(p_1+\tilde{u}_1)|\\ &\times\Bigg(\prod_{j=1}^{k+1}\frac{1}{|\alpha-\overline{\omega(p_j)}-i\eta|}\,\frac{1}{|\beta-\omega([M\mathbf{p}+\tilde{\mathbf{u}}]_j)+i\eta|}\Bigg)\,d\mu(\mathbf{p})\,d\alpha\,d\beta. \end{aligned} \tag{10.1}$$

The key step in the proof of Theorem 8.4 is the following lemma.

LEMMA 10.1. *Assume that the relations* (7.25) *hold, that* $\varkappa<2/(6+9d)$ *and that* $\delta\leqslant\delta(\varkappa)$ *is sufficiently small. Then for any* $\sigma\in\mathfrak{S}_k$ *we have*

$$E(M(\sigma))\leqslant C(\lambda^{1/3-(1+3d/2)\varkappa-O(\delta)})^{\deg(\sigma)}|\log\eta|^2, \tag{10.2}$$

where the matrix $M(\sigma)$ *was defined in* (8.7) *and the degree of the permutation*, $\deg(\sigma)$, *was defined in* (8.9).

From (8.4) and (8.8), clearly

$$\sup_{\mathbf{u}} E_{g=0}(\sigma,\mathbf{u})\leqslant\frac{\exp(2t\eta)}{(2\pi)^2}E(M(\sigma)) \tag{10.3}$$

after integrating out all $\tilde{p}_j$ variables in (8.4), and by using that $\tilde{p}_1=p_1-\xi$. The estimate (10.2) will then complete the proof of Theorem 8.4 for $g=0$.

The proof of Theorem 8.4 for other (but finitely many) values of g follows exactly in the same way. This requires a slight redefinition of $\mathcal{B}(\mathbf{p})$ (see (6.5)) in the definition of $E(M)$ by allowing the factor $\langle p_{j+1}-p_j\rangle^{-2d}$ instead of $\widehat{B}(p_{j+1}-p_j)$ at a few places, exactly as in the definition of $\mathcal{N}_\mathcal{G}$ (see (7.3)). As we will see along the proof of (10.2), this change will require using the less precise bound (3.13) with $a=0$ and $h(p-q)=\langle p-q\rangle^{-2d}$, instead of the more accurate estimate (3.15) at most g times. Each time we lose a constant factor compared with the proof for $g=0$. Since $g\leqslant 8$, this results only in a constant factor. Finally, the proof for the truncated E-values requires us to define a truncated version of $E(M)$, where the last product in (10.1) runs only up to $j=k$, i.e. the last α and β denominators are not present. It will be clear from the proof of Lemma 10.1 that the same bound holds for the truncated version of $E(M)$ as well. This completes the proof of Theorem 8.4.

10.1. Pedagogical detour

The size of the multiple integral in (10.1) heavily depends on the structure of $M=M(\sigma)$. Before we go into the algorithm to evaluate this multiple integral, we present two calculations, that introduce the techniques that we are going to use in the actual proof. The second calculation also provides the bound (10.2), hence (8.11), for the case of the trivial permutation, $\deg(\sigma)=0$.

10.1.1. Method I. Pointwise bound

The most straightforward bound on (10.1) estimates all but one of the β-denominators by L^∞-norm:

$$\sup_{\beta,\mathbf{p},\tilde{\mathbf{u}},j}\frac{1}{|\beta-\omega([M\mathbf{p}+\tilde{\mathbf{u}}]_j)+i\eta|}\leqslant\frac{1}{\eta}. \tag{10.4}$$

It would be possible to estimate this denominator by $C\lambda^{-2}$ apart from in a neighborhood of zero using (3.7), and treat the $|p|{\sim}0$ regime separately. Here we choose the simplest argument and we do not optimize for the best possible exponent $\varkappa$.

We use (10.4) k times to obtain

$$E(M)\leqslant\left(\frac{C\lambda^2}{\eta}\right)^k\sup_{\tilde{\mathbf{u}}}\int_{\mathbf{R}^d}\int_{-Y}^{Y}\int_{-Y}^{Y}\int_{\mathbf{R}^d}\dots\int_{\mathbf{R}^d}\biggl(\prod_{j=1}^{k}\frac{|\widehat{B}(p_{j+1}-p_j)|}{|\alpha-\overline{\omega(p_j)}-i\eta|}\biggr)\,dp_1\dots dp_k \\ \times\frac{d\alpha\,d\beta}{|\alpha-\overline{\omega(p_{k+1})}-i\eta|\,|\beta-\omega(p_{k+1}+\tilde{u}_{k+1})+i\eta|}\,\frac{d\mu(p_{k+1})}{\langle p_{k+1}+\tilde{u}_{k+1}\rangle^{2d}}.$$

Here we have used the following bound for any $\mathbf{q}=(q_1,\dots,q_{k+1})$:

$$|\mathcal{B}(\mathbf{q})|\,|\widehat{\psi}_0(q_1)|\leqslant C^k\langle q_{k+1}\rangle^{-2d},\tag{10.5}$$

to obtain the decay in $\tilde{p}_{k+1}=p_{k+1}+\tilde{u}_{k+1}$. We integrate out $p_1,p_2,\dots,p_k$ using (3.12), then we perform the $d\alpha$ and $d\beta$ integrals and finally dp_{k+1}, to obtain

$$E(M)\leqslant(C|\log\lambda|)^{k+2}(\lambda^2\eta^{-1})^k.\tag{10.6}$$

The estimate (10.6) is off by a factor $(\lambda^2\eta^{-1})^k=(\lambda^{-\varkappa})^k$, because we did not use the stronger estimate mentioned after (10.4). We also collected many logarithmic factors and the constant is not optimal. We note that in the typical term $k{\sim}\lambda^2 t{\sim}\lambda^{-\varkappa}{\gg}1$, so even an error C^k may not be affordable. To improve this estimate, for a typical matrix M, we will not use the pointwise bound (10.4) for all β-denominators. We will carefully select those β-denominators whose singularities cannot overlap with other singularities, hence they can be integrated out at a $|\log\eta|$ expense instead of η^{-1}.

Before we explain this algorithm, we show another method to estimate $E(M)$. It practically estimates $E(M)$ by $E(I)$, i.e. by the ladder graph, that can be computed more precisely. The same calculation will be important when evaluating embedded ladder graphs.

10.1.2. Method II. Successive integration scheme for ladder graphs

We separate all but one α- and β-denominators by the Schwarz inequality. We obtain

$$E(M)\leqslant\lambda^{2k}\sup_{\tilde{\mathbf{u}}}\int_{(\mathbf{R}^d)^{k+1}}\int_{-Y}^{Y}\int_{-Y}^{Y}\biggl[|\widehat{\psi}_0(p_1)|^2\prod_{j=2}^{k+1}\frac{|\widehat{B}(p_j-p_{j-1})|^2}{|\alpha-\overline{\omega(p_j)}-i\eta|^2} \\ +|\widehat{\psi}_0(q_1)|^2\prod_{j=2}^{k+1}\frac{|\widehat{B}(q_j-q_{j-1})|^2}{|\beta-\omega(q_j)+i\eta|^2}\biggr]\frac{d\alpha\,d\beta}{|\alpha-\overline{\omega(p_1)}-i\eta|\,|\beta-\omega(p_1+\tilde{u}_1)+i\eta|}\,d\mu(\mathbf{p}),$$

with the shorthand notation $\mathbf{q}:=M\mathbf{p}+\tilde{\mathbf{u}}$. Because M is an invertible matrix with determinant ±1 (Proposition 8.2), the contributions of the two terms in the square bracket are identical up to exchange of α and β. To estimate the first term, we use iteratively (3.15) and (3.13) (with $a=\frac{1}{2}$) to integrate out $p_{k+1}, p_k, ..., p_2$ (in this order):

$$\lambda^2\int_{\mathbf{R}^d}\frac{|\widehat{B}(p_{k+1}-p_k)|^2}{|\alpha-\overline{\omega(p_{k+1})}-i\eta|^2}\,dp_{k+1}\leqslant[1+C_0\lambda^{-12\varkappa}(\lambda+|\alpha-\omega(p_k)|^{1/2})], \tag{10.7}$$

$$\begin{aligned}\lambda^2\int_{\mathbf{R}^d}\frac{|\widehat{B}(p_k-p_{k-1})|^2}{|\alpha-\overline{\omega(p_k)}-i\eta|^2}&[1+C_0\lambda^{-12\varkappa}(\lambda+|\alpha-\omega(p_k)|^{1/2})]\,dp_k\\ &\leqslant(1+C\lambda^{1-12\varkappa})[1+C_0\lambda^{-12\varkappa}(\lambda+|\alpha-\omega(p_{k-1})|^{1/2})],\end{aligned} \tag{10.8}$$

etc., with $C:=C_0(1+C_{1/2}\|\widehat{B}^2\|_{2d,0})$. In the last step, we use only (3.13), once for $a=0$ and once for $a=\frac{1}{2}$:

$$\lambda^2\int_{\mathbf{R}^d}\frac{|\widehat{B}(p_2-p_1)|^2}{|\alpha-\overline{\omega(p_2)}-i\eta|^2}[1+C_0\lambda^{-12\varkappa}(\lambda+|\alpha-\omega(p_2)|^{1/2})]\,dp_2\leqslant C. \tag{10.9}$$

Then we integrate $d\alpha\,d\beta$ and finally dp_1 to obtain that

$$E(M)\leqslant C(1+C\lambda^{1-12\varkappa})^k|\log\lambda|^2\leqslant C|\log\lambda|^2, \tag{10.10}$$

by using that $k\leqslant K\ll\lambda^{-1+12\varkappa}$ as $\varkappa<\frac{1}{13}$.

We note that this method also gives a robust bound for the truncated E-value, since the truncation means that Lemma 3.3 is used only $k-1$ times. Summarizing, we have proved the following result.

LEMMA 10.2. *We assume* (7.25) *and that* $\varkappa<\frac{1}{13}$. *Then*

$$\sup_{\sigma\in\mathfrak{S}_k}\sup_{\mathbf{u}}E(\sigma,\mathbf{u})\leqslant C|\log\lambda|^2, \tag{10.11}$$

$$\sup_{\sigma\in\mathfrak{S}_k}\sup_{\mathbf{u}}E_*(\sigma,\mathbf{u})\leqslant C\lambda^2|\log\lambda|^2. \tag{10.12}$$

10.2. Choice of the integration variables

Before we start the proof of Lemma 10.1, we explain the main idea. We use a combination of methods I and II. We will assume in the sequel that $\sigma\neq\mathrm{id}$. The lemma for the trivial case $\sigma=\mathrm{id}$ has been proven in (10.10).

Note that each factor in the integrand in (10.1) is almost singular on a set of codimension 1 of the form

$$\{(p_1,...,p_{k+1},\alpha,\beta):\alpha=\mathrm{Re}\,\omega(p_j)\}\quad\text{or}\quad\{(p_1,...,p_{k+1},\alpha,\beta):\beta=\mathrm{Re}\,\omega([M\mathbf{p}+\tilde{\mathbf{u}}]_j)\}$$

in the high-dimensional space of integration, $(\mathbf{R}^d)^{k+1}\times\mathbf{R}^2$. The singularities are regularized by η and $\operatorname{Im}\omega$, and the two regularizations always have the same sign. The matrix M may enhance the strength of these singularities by forcing these "almost singularity" sets to overlap. For example, in the ladder diagram with $\tilde{\mathbf{u}}\equiv 0$, we have $\tilde{\mathbf{p}}\equiv\mathbf{p}$, hence the singularity sets are pairwise identical if $\alpha=\beta$. Therefore, singularities of quadratic type necessarily occur. It is expected that this is the only mechanism that creates relevant overlaps of singularities. Hence, ideally, one would integrate out the ladder momenta using the precise bound (3.15). This would remove all denominators with indices $j\in I_l$ and the remaining integral should be of $O(\lambda^{2(k-l)})$ with possible logarithmic corrections, i.e. one should gain a factor λ^2 from each non-ladder index.

However, the singularity sets of the remaining denominators may still intersect on higher-codimensional manifolds. For example, since

$$\sup_{\alpha,\beta}\int_{\mathbf{R}^d}\frac{1}{|\alpha-\overline{\omega(p+q)}-i\eta|}\,\frac{1}{|\beta-\omega(p)+i\eta|}\,d\mu(p)\geqslant\frac{C}{|q|+\eta},$$

the integration of these two propagators develops an almost-singularity at $q=0$. We will call such factors *point singularities*, although they are regularized at a very short scale η. After integration, one point singularity is harmless as $(|q|+\eta)^{-1}\in L^1_{\mathrm{loc}}(dq)$; however, several point singularities may accumulate along the procedure whose simultaneous integration may lead to further divergences.

These enhancements of singularities are expected to have contributions of lower order, but their estimate is not easy. Note that every integration variable p_j may appear in many denominators. This interdependence renders the effective L^1-estimate of each integral practically impossible. Schwarz inequality (method II) can remove all correlations between denominators, but the resulting L^2-estimate is of the same order as the main (ladder) term and we would not gain anything from the higher degree of the permutation σ.

The idea is to estimate many, but not all β-denominators in (10.1) in the trivial way (10.4). These denominators are chosen in such a way that the remaining ones can be successively integrated out without ever computing an integrand with more than two propagators and without collecting more than at most one point-singularity factor. This choice of integration variables are determined from $M=M(\sigma)$ by an algorithmic procedure that we describe now.

We first recall that M is a tower matrix, and we define $b(j)$ and $t(j)$ to be the bottom and the top of the tower in the jth column, i.e.

$$b(j):=\max\{h:M_{hj}\neq 0\}\quad\text{and}\quad t(j):=\min\{h:M_{hj}\neq 0\}.$$

We also recall the concepts from Definition 8.3. Any non-peak index $h\in I\setminus I_p$ clearly has the property that h is the bottom of the tower of some column of M, i.e. there exists j such that $b(j)=h$. This column index can be either $j=\tilde{\sigma}^{-1}(h)$ or $j=\tilde{\sigma}^{-1}(h)+1$ or both; ambiguity occurs only if $h\in I_v$ is a valley index. For any ladder or slope or last index, $h\in I_l\cup I_s\cup I_{\mathrm{last}}$, we therefore define $c(h)$ to be the unique column index j such that $b(j)=h$. For $h\in I_{\mathrm{last}}$, i.e. $h=k+1$, we always have $c(k+1)=k+1$. If $h\in I_v$, we define $c(h)$ to be the index of the column whose tower is higher, i.e.

$$c(h):=\begin{cases}\tilde{\sigma}^{-1}(h), & \text{if } t(\tilde{\sigma}^{-1}(h))<t(\tilde{\sigma}^{-1}(h)+1),\\ \tilde{\sigma}^{-1}(h)+1, & \text{otherwise.}\end{cases}$$

The other column index will be denoted by $\tilde{c}(h)$; this concept is defined only for $h\in I_v$. The elements $(h,c(h))$, $h\in I\setminus I_p$, of the matrix M are called *pivot* elements; these will be used for determining the integration variables. The $c(h)$-column will be called the pivot column of h and the corresponding tower in this column is called the pivot tower of h. The definition guarantees that the pivot tower of an index $h\in I_s\cup I_v$ has at least length 2.

Definition 10.3. Let $h_1<h_2<\dots$ denote the elements of $I_v\cup I_s$ in increasing order. A slope index $h_\mu\in I_s$, is called *covered* if $t(c(h_{\mu+1}))\leqslant h_\mu$, i.e. if the top of the tower of the next pivot element of a valley or a slope is higher than h_μ. The remaining slope indices are called *uncovered*. The set of covered and uncovered slope indices will be denoted by I_{cs} and I_{us}, respectively.

Figure 6 shows three examples. In the first two cases, the index h_μ is covered by a slope and by a valley. The last case is an uncovered slope index. The shaded boxes indicate the pivot elements.

Apart from the ladder momenta $p_{c(h)}$, $h\in I_l$, we will integrate out the variables $p_{c(h_1)}, p_{c(h_2)}, \dots$, in this order. We will show that, at every integration step, only one α-denominator and one β-denominator, namely

$$|\alpha-\overline{\omega(p_{c(h_\mu)})}-i\eta|^{-1} \quad \text{and} \quad |\beta-\omega(\tilde{p}_{h_\mu})+i\eta|^{-1},$$

will contain the integration variable $p_{c(h_\mu)}$, so we will never have to estimate integrals with more than two propagators.

The significance of the covered slope indices is that their integration step yields a point singularity that will be integrated out immediately at the next non-ladder integration step. Therefore covered slope indices do not give rise to accumulation of point singularities.

The covered slope indices, together with the valley indices and the last index will be used to gain a λ^{const} factor in the integration procedure. They will not accumulate point

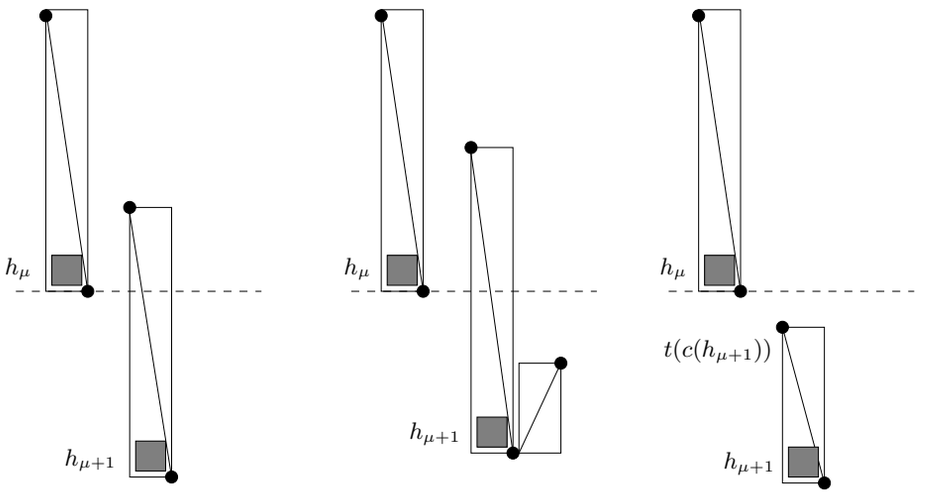

Figure 6. Two examples of a covered slope index h_μ and one of an uncovered index.

singularities. The uncovered slope indices and the peak indices will be treated by the trivial L^∞-bound (10.4) and no gain is obtained from them. That is, we prefer giving up the potential gain of a λ^{const} factor from the uncovered slope indices to dealing with the accumulated point singularities. Finally, the ladder indices will be treated by method II with the Schwarz inequality.

LEMMA 10.4. *The number of uncovered slope indices is at most v, where $v=|I_v|$ is the number of valleys.*

This lemma ensures that treating the uncovered slope indices by the trivial bound is affordable (at the expense of a worse $\varkappa$), since each valley will result in a gain of a λ^{const} factor.

Proof. First, we state two claims from which we easily deduce the lemma, then we prove these claims. Recall that the sequence $h_1<h_2<\dots$ lists all elements of $I_v\cup I_s$.

Claim 1. Let $h_\nu\in I_s$ be the first slope index in the sequence $h_1<h_2<\dots$. Then, there is a peak index h such that $h<h_\nu$.

Claim 2. Let $h_\mu\in I_{\text{us}}$ be an uncovered slope index and suppose that it is not the last entry in the sequence $h_1<h_2<\dots$, i.e. there is a next element $h_{\mu+1}\in I_s\cup I_v$. Then, there is an intermediate index, $h_\mu<h<h_{\mu+1}$, that is a peak index.

By Claim 2, one can assign a distinct peak index to every uncovered slope index h_μ (but the last one), namely the next peak index. Moreover, by Claim 1, there is an additional peak index smaller than any uncovered slope index. Therefore, the number

of uncovered slope indices is at most the number of peaks. The statement of the lemma then follows, since the number of peaks and valleys is the same, $p=v$.

Proof of Claim 1. Suppose that the statement is false. First note that $h_\nu \geqslant 2$, since index 1 cannot be a slope. Then all indices $1, 2, ..., h_\nu - 1$ must be either ladders or valleys. But for any valley index j there is obviously a peak index $j' < j$, since the graph of the permutation goes down from $(0,0)$ to $(k+1, k+1)$, so, following the two branches of the graph emanating upwards starting from a valley, at least one of them must run into a peak that lies higher. Therefore all indices $1, 2, ... h_\nu - 1$ must be ladders and they form a ladder, with the index 0 being the top of this ladder. In particular, $t(c(h)) = h-1$, $h = 1, 2, ..., h_\nu - 1$ (recall that $t(c(h))$ denotes the index of the top row in the pivot tower of h), so every index $0, 1, 2, ..., h_\nu - 2$ is the top of a column of length 1. Moreover, none of these indices can be the top of any other column, since an index which is the top of two different columns must be a peak index. But in that case there is no room for $t(c(h_\nu))$, the top index of the pivot column of h_ν: since h_ν was a slope, $t(c(h_\nu)) \leqslant h_\nu - 2$.

Proof of Claim 2. To see the claim, suppose that none of the indices h strictly between h_μ and $h_{\mu+1}$ is a peak. These indices cannot be slopes or valleys either, since h_μ and $h_{\mu+1}$ are two consecutive elements of $I_s \cup I_v$ and h cannot be the last index, $h \notin I_{\text{last}}$, as $h < h_{\mu+1} < k+1$. Therefore all indices h with $h_\mu < h < h_{\mu+1}$ are ladders.

Now we look at $h^* = t(c(h_{\mu+1}))$, i.e. the index of the top row in the pivot tower of $h_{\mu+1}$ (see the last picture in Figure 6). Since h_μ was uncovered, $h_\mu < h^*$. Note that the pivot tower of $h_{\mu+1} \in I_s \cup I_v$ has length at least 2, therefore $h^* < h_{\mu+1}$, and thus $h^* \in I_l$. Since h^* was also the top of a tower with length at least 2, $h^* - 1$ must be a peak index, as it is a row index right above the top of two towers (namely the towers in the columns $c(h^*)$ and $c(h_{\mu+1})$). Clearly $h_\mu \leqslant h^* - 1 < h_{\mu+1}$, but $h_\mu \neq h^* - 1$ since $h_\mu \in I_s$ and $h^* - 1 \in I_p$. Thus $h^* - 1$ is a peak index strictly between h_μ and $h_{\mu+1}$, which is a contradiction. □

10.3. Integration procedure

Our goal is to estimate (10.1) for $M = M(\sigma)$ when $\sigma \neq \mathrm{id}$. We start with defining

$$|||q||| := \eta + \min\{|q|, 1\}, \quad q \in \mathbf{R}^d. \tag{10.13}$$

This is not a norm, but it satisfies the triangle inequality, $|||p+q||| \leqslant |||p||| + |||q|||$. For any index set $I' \subset I = \{1, 2, ..., k+1\}$, we define the function $\mathcal{U}_{I'}$ as the product of those potential terms, $|\widehat{B}|$, in (10.1) that depend only on momenta $\{p_j : j \in I'\}$. More precisely,

$$\mathcal{U}_{I'}(\mathbf{p}, \tilde{\mathbf{u}}) := \left| \prod_j{}^{*} \widehat{B}(p_{j+1} - p_j) \prod_j{}^{*} \widehat{B}([M\mathbf{p} + \tilde{\mathbf{u}}]_{j+1} - [M\mathbf{p} + \tilde{\mathbf{u}}]_j) \right|,$$

where the star indicates a product on a restricted index set. The first product is taken over all indices j for which $j, j+1 \in I'$. The second product is taken over those j for which $M_{j+1,b} = M_{j,b}$ for all $b \notin I'$.

For any $|I'| \times (k+1)$ matrix M and any vector $\mathfrak{b} = (b_1, b_2, ..., b_{k+1}) \in \mathbf{R}^{k+1}$ we define

$$E(I', M, \mathfrak{b}) := \lambda^{2k} \sup_{\tilde{\mathbf{u}}, v} \int_{-Y}^{Y} \int_{-Y}^{Y} \sup_{p_j : j \notin I'} \int_{(\mathbf{R}^d)^{|I'|}} \mathcal{U}_{I'}(\mathbf{p}, \tilde{\mathbf{u}}) \frac{1}{|||\mathfrak{b} \cdot \mathbf{p} + v|||} \times \Big(\prod_{j \in I'} \frac{1}{|\alpha - \overline{\omega(p_j)} - i\eta|} \, \frac{1}{|\beta - \omega([M\mathbf{p} + \tilde{\mathbf{u}}]_j) + i\eta|} \Big) \Big(\prod_{j \in I'} d\mu(p_j) \Big) \, d\alpha \, d\beta. \tag{10.14}$$

Here $v \in \mathbf{R}^d$ is an additional dummy momentum and $\mathfrak{b} \cdot \mathbf{p}$ is defined, analogously to (8.5), as

$$\mathfrak{b} \cdot \mathbf{p} := b_1 p_1 + b_2 p_2 + ... + b_{k+1} p_{k+1}.$$

We will also use the notation $E(I', M, \varnothing)$ defined exactly by the same formula as (10.14) without the factor $|||\mathfrak{b} \cdot \mathbf{p} + v|||^{-1}$ in the integrand and without the supremum over v. With a slight abuse of notation we will refer to this case as chosing the "empty vector" $\mathfrak{b} = \varnothing$.

For each $h \in I$ we will define an index set $I^{(h)}$, a matrix $M^{(h)}$ and a vector $\mathfrak{b}^{(h)}$ such that $E(I^{(h)}, M^{(h)}, \mathfrak{b}^{(h)})$ is the remaining integral after the hth integration step. Each step consists of eliminating one β-denominator; in the hth step we will eliminate $|\beta - \omega([M\mathbf{p} + \tilde{\mathbf{u}}]_h) + i\eta|^{-1}$ and sometimes one or two additional α-denominators. The α-denominators will always be eliminated by integrating out the corresponding p_j variable. The order of integrations will be such that the integrand in $E(I^{(h)}, M^{(h)}, \mathfrak{b}^{(h)})$ does not depend on any $p_j \notin I^{(h)}$. In particular, the supremum over all $p_j \notin I^{(h)}$ in the definition (10.14) will be superfluous in the integrals appearing in our iteration.

The elimination will be done differently, depending on the type of the index h. If $h \in I_p$, then the trivial L^∞-bound (10.4) will be used. If h is an uncovered slope index, then again the trivial bound will be used, and, additionally, the variable $p_{c(h)}$ will be integrated out. If h is a slope or ladder index, then variable $p_{c(h)}$ will be integrated out. If h is a valley then we integrate out both variables $p_{c(h)}$ and $p_{\bar{c}(h)}$. Finally, in the last step, $h = k+1$, we will integrate out the β variable. We will bookkeep this integration by removing the corresponding h rows from M and we will say in short that these rows have been integrated out. At the end we will have to integrate out the remaining p_j variables, but then only one α-denominator will be present, so the integration can be done easily.

Point singularities are created by integrating out a covered slope index, but there is never more than one singularity present. An existing point singularity will not be touched by the ladder integrations, and it will be changed to another point singularity

when a covered slope index is integrated out. In all other integration steps an existing point singularity will be removed.

Now we define $M^{(h)}$, $I^{(h)}$ and $\mathfrak{b}^{(h)}$. Let $M^{(0)}:=M=M(\sigma)$, $I^{(0)}:=I=\{1,2,...,k+1\}$ and $\mathfrak{b}^{(0)}=\varnothing$. In particular, note that

$$E(M)\leqslant\|\widehat{\psi}_0\|_\infty^2 E(I^{(0)},M^{(0)},\mathfrak{b}^{(0)}). \tag{10.15}$$

For any $h\geqslant 1$, define $M^{(h)}$ to be the matrix $M=M(\sigma)$ with the first h rows, with indices $1,2,...,h$ removed. For $h\geqslant 1$, let

$$I^{(h)}:=I\setminus(\{c(h'):h'\leqslant h \text{ and } h'\in I\setminus I_p\}\cup\{\tilde{c}(h'):h'\leqslant h \text{ and } h'\in I_v\})$$

be the index set of columns that have not been integrated out up to the hth step.

The vectors $\mathfrak{b}^{(h)}=(b_1^{(h)},...,b_{k+1}^{(h)})$ are defined as follows. For any given $h\in I$, let h_μ be the largest element in the listing $h_1<h_2<...$ of $I_v\cup I_s$ (see Definition 10.3) such that $h_\mu\leqslant h$. If there is no such h_μ or $h_\mu\notin I_{\mathrm{cs}}$, then we set $\mathfrak{b}^{(h)}=\varnothing$. If $h_\mu\in I_{\mathrm{cs}}$, then

$$b_j^{(h)}:=\begin{cases} M_{h_\mu,j}, & \text{if } j\neq c(h),\\ 0, & \text{if } j=c(h).\end{cases} \tag{10.16}$$

We collect some information about $\mathfrak{b}^{(h)}$ that directly follow from the definition and from the structure of M.

(i) Suppose that $\mathfrak{b}^{(h)}\neq\varnothing$, and let h_μ be as above (i.e. $h_\mu\leqslant h$ and it is the largest element with this property). Since $h_\mu\in I_{\mathrm{cs}}$, i.e. $t(c(h_{\mu+1}))\leqslant h_\mu$, we therefore have

$$b_{c(h_{\mu+1})}^{(h)}=\pm 1,\quad h_\mu\leqslant h<h_{\mu+1}\quad\text{and}\quad h_\mu\in I_{\mathrm{cs}}. \tag{10.17}$$

(ii) For any fixed μ, $\mathfrak{b}^{(h)}$ is constant as h runs through the interval $\{h_\mu,...,h_{\mu+1}-1\}$.

(iii) Consider the pivot column $c(j)$ of a ladder index $j\in I_l$. Since the only non-zero entry in the $c(j)$ column of M is $M_{j,c(j)}$, we have that $M_{h_\mu,c(j)}=0$ for all μ. But then we see that the entries corresponding to the ladder pivots are all zero in the $\mathfrak{b}^{(h)}$ vectors:

$$b_{c(j)}^{(h)}=0\quad\text{for any } j\in I_l \text{ and any } h \text{ such that } \mathfrak{b}^{(h)}\neq\varnothing. \tag{10.18}$$

To describe the integration, we explain how each case, depending on the index type of h, is estimated. We also introduce the short notation

$$E(h):=E(I^{(h)},M^{(h)},\mathfrak{b}^{(h)}),\quad h=0,1,...,k+1.$$

Case 1. $h\in I_p$. In this case no integration is done, $I^{(h)}=I^{(h-1)}$ and $\mathfrak{b}^{(h)}=\mathfrak{b}^{(h-1)}$. We estimate $|\beta-\omega([M\mathbf{p}+\tilde{\mathbf{u}}]_h)+i\eta|^{-1}$ trivially as in (10.4), and we obtain that

$$E(h-1)\leqslant\eta^{-1}E(h),\quad h\in I_p. \tag{10.19}$$

Case 2. $h\in I_l$. Let $h, h+1, ..., h+\tau-1\in I_l$ be a ladder, i.e. a maximal sequence of consecutive ladder indices, i.e. $h+\tau\notin I_l$ for some $\tau\geqslant 1$. We will integrate out the whole ladder in this step, so we may assume that $h-1$ is the top of a ladder, $h-1\in I_t$. By definition of ladder, the pivot indices $c(h), c(h+1), ..., c(h+\tau-1)$ are consecutive numbers, which, for definiteness, are assumed to be increasing (the other case is similar). Let $c=c(h)$. Because $h, ..., h+\tau-1$ are all consecutive ladder indices, $[M\mathbf{p}]_j=\mathbf{p}_j+w$ for all $c\leqslant j\leqslant c+\tau-1$, with a vector $w=w(\mathbf{p})$ that does not depend on $\{p_j : c\leqslant j\leqslant c+\tau-1\}$. Note that $\mathfrak{b}$ remains unchanged as k runs through the ladder indices: $\mathfrak{b}^{(k)}=\mathfrak{b}^{(h-1)}$ for all $k=h, h+1, ..., h+\tau-1$. We claim that

$$E(h-1)\leqslant C\lambda^{-2\tau}E(h+\tau-1), \quad h-1\in I_t, \tag{10.20}$$

with a constant C independent of τ. This inequality entails integrating out the variables $p_c, p_{c+1}, ..., p_{c+\tau-1}$ (in this order) to remove the rows $h, h+1, ..., h+\tau-1$. This requires estimating the integral

$$\mathcal{I} := \sup_{\alpha,\beta,w} \int_{(\mathbf{R}^d)^\tau} \prod_{j=c}^{c+\tau-1} \frac{|\widehat{B}(p_{j+1}-p_j)|\,|\widehat{B}(p_{j+1}+\tilde{u}_{j+1}-p_j-\tilde{u}_j)|\,d\mu(p_j)}{|\alpha-\overline{\omega(p_j)}-i\eta|\,|\beta-\omega(p_j+w+\tilde{u}_j)+i\eta|}, \tag{10.21}$$

because the possible point singularity present in $E(I^{(h-1)}, M^{(h-1)}, \mathfrak{b}^{(h-1)})$ does not contain any of the integration variables, by (10.18). Due to the ladder structure, it is easy to see that $\mathcal{U}_{I^{(h-1)}}$ indeed contains all the $\widehat{B}$ factors shown in (10.21).

After Schwarz inequality, we use an iterative integration scheme similar to (10.7)–(10.9) to estimate the τ-fold multiple integral

$$\begin{aligned}
&\int_{(\mathbf{R}^d)^\tau} \prod_{j=c}^{c+\tau-1} \frac{|\widehat{B}(p_{j+1}-p_j)|^2\,d\mu(p_j)}{|\alpha-\overline{\omega(p_j)}-i\eta|^2} \\
&\quad\leqslant [\lambda^{-2}(1+C\lambda^{1-12\varkappa})]^{\tau-1} \\
&\qquad \times \int_{\mathbf{R}^d} [1+C_0\lambda^{-12\varkappa}(\lambda+|\alpha-\omega(p_{c+\tau-1})|^{1/2})] \frac{|\widehat{B}(p_{c+\tau}-p_{c+\tau-1})|^2\,d\mu(p_{c+\tau-1})}{|\alpha-\overline{\omega(p_{c+\tau-1})}-i\eta|^2} \\
&\quad\leqslant C[\lambda^{-2}(1+C\lambda^{1-12\varkappa})]^{\tau-1} \\
&\quad\leqslant C\lambda^{-2(\tau-1)}.
\end{aligned}$$

In the last integration we use (3.13) with $a=0$ and $a=\frac{1}{2}$ and we also use the fact that $\tau\leqslant K\ll\lambda^{-(1-12\varkappa)}$. This gives

$$\mathcal{I}\leqslant C\lambda^{-2\tau}. \tag{10.22}$$

Case 3. $h\in I_{\mathrm{us}}$. We have $I^{(h)}=I^{(h-1)}\setminus\{c(h)\}$ and $\mathfrak{b}^{(h)}=\varnothing$. We start proceeding as in Case 1 by first estimating the denominator $|\beta-\omega([M\mathbf{p}+\tilde{\mathbf{u}}]_h)+i\eta|^{-1}$ trivially as in (10.4). Then observe that the variable $p_{c(h)}$ appears only in one propagator, namely in $|\alpha-\overline{\omega(p_{c(h)})}-i\eta|^{-1}$. Moreover, if $\mathfrak{b}^{(h-1)}\neq\varnothing$, then $p_{c(h)}$ appears also in the point singularity, since $b^{(h-1)}_{c(h)}=\pm1$ by (10.17).

Before we integrate out $p_{c(h)}$, we estimate all $\widehat{B}$ factors that contain this variable by the supremum norm. The number of such factors is at most four, so they can be bounded by a constant factor C. Then, to integrate out $p_{c(h)}$, we either use the bound (with $p=p_{c(h)}$)

$$\sup_{|\alpha|\leqslant Y}\int_{\mathbf{R}^d}\frac{d\mu(p)}{|\alpha-\overline{\omega(p)}-i\eta|}\leqslant C\zeta^{d-2}|\log\eta| \tag{10.23}$$

if $\mathfrak{b}^{(h-1)}=\varnothing$, or we use the bound

$$\sup_{|\alpha|\leqslant Y}\sup_r\int_{\mathbf{R}^d}\frac{1}{|||p-r|||}\frac{d\mu(p)}{|\alpha-\overline{\omega(p)}-i\eta|}\leqslant C\zeta^{d-2}|\log\eta| \tag{10.24}$$

if $\mathfrak{b}^{(h-1)}\neq\varnothing$ (recall that $\zeta=\lambda^{-\varkappa-3\delta}$ from (7.25)). The bound (10.24) follows immediately from (A.1) and (A.5), the bound (10.23) is weaker since $\sup_p|||p|||\leqslant1+\eta$ (the bound (10.23) can be improved to $C|\log\eta|$, but we will not need this fact). In summary, we have proved that

$$E(h-1)\leqslant C\zeta^{d-2}\eta^{-1}|\log\eta|E(h),\quad h\in I_{\mathrm{us}}. \tag{10.25}$$

Case 4. $h\in I_{\mathrm{cs}}$. In this step we will have $I^{(h)}=I^{(h-1)}\setminus\{c(h)\}$, and $\mathfrak{b}^{(h)}$ will be the hth row of M but with the pivot element $(h,c(h))$ changed to zero, i.e. $b^{(h)}_h=0$. Since the index h is among the listing $h_1<h_2<\dots$ of the elements $I_v\cup I_s$ (see Definition 10.3), let $h=h_\mu$ and $h':=h_{\mu+1}$ be the next element of $I_s\cup I_v$. As h is a covered slope index, we have $t(c(h'))\leqslant h$, and thus we know that the $(h,c(h'))$ matrix element of M is ±1. Therefore the variable $p_{c(h')}$ appears non-trivially in $\tilde{p}_h=[M\mathbf{p}+\tilde{\mathbf{u}}]_h$:

$$\tilde{p}_h=[M\mathbf{p}+\tilde{\mathbf{u}}]_h=\pm p_{c(h)}\pm p_{c(h')}\pm\dots\,. \tag{10.26}$$

After estimating those $\widehat{B}$ factors that contain $p_{c(h)}$ by the supremum norm, we will integrate out $p_{c(h)}$ by using the following lemma proved in the appendix.

LEMMA 10.5. *Let* $\lambda^3\leqslant\eta\leqslant\lambda^2$, *and let* $|q|\leqslant C\lambda^{-1}$. *Then we have*

$$\sup_{\alpha,\beta,r}\int_{\mathbf{R}^d}\frac{1}{|||p-r|||}\frac{d\mu(p)}{|\alpha-\overline{\omega(p)}-i\eta|\,|\beta-\omega(p+q)+i\eta|}\leqslant\frac{C\eta^{-1/2}\zeta^{d-3}|\log\eta|^2}{|||q|||}. \tag{10.27}$$

Without point singularity, we have the following improved bound:

$$\sup_{\alpha,\beta}\int\frac{d\mu(p)}{|\alpha-\overline{\omega(p)}-i\eta|\,|\beta-\omega(p+q)+i\eta|}\leqslant\frac{C\zeta^{d-3}|\log\eta|^2}{|||q|||}. \tag{10.28}$$

The first estimate (10.27) is used if $\mathfrak{b}^{(h-1)}\neq\varnothing$, i.e. if there was a non-trivial point singularity present. By (10.17), this non-trivial point singularity must originate from the previous covered slope index, in particular $h_{\mu-1}\in I_{\mathrm{cs}}$ and $b_{c(h)}^{(h-1)}=\pm1$, i.e. the point singularity contains the variable $p_{c(h)}$. The new point singularity, obtained on the right-hand side of (10.27) as $|||q|||^{-1}$, is exactly the linear combination appearing in (10.26) without the integration variable, $p_{c(h)}$. Therefore, the p-dependence of this point singularity is given by $\mathfrak{b}^{(h)}$ (up to an overall sign that is irrelevant, since $|||\cdot|||$ is symmetric). If $\mathfrak{b}^{(h-1)}=\varnothing$, then we can use the second estimate (10.28), and otherwise we follow the same argument as in the case $\mathfrak{b}^{(h-1)}\neq\varnothing$.

The net effect of the elimination in Case 4 is

$$E(h-1)\leqslant C\eta^{-1/2}\zeta^{d-3}|\log\eta|^2E(h),\quad h\in I_{\mathrm{cs}}. \tag{10.29}$$

Case 5. $h\in I_v$. In this step we will have $I^{(h)}=I^{(h-1)}\setminus\{c(h),\tilde{c}(h)\}$ and $\mathfrak{b}^{(h)}=\varnothing$. Recall that $c(h)$ and $\tilde{c}(h)$ are the two columns between which the valley is located, i.e. $\{c(h),\tilde{c}(h)\}=\{\tilde{\sigma}^{-1}(h),\tilde{\sigma}^{-1}(h)+1\}$. We integrate out both $p_{c(h)}$ and $p_{\tilde{c}(h)}$ after estimating the $\widehat{B}$ factors that are involved by their supremum norms. This eliminates two α-denominators and the $|\beta-\omega([M\mathbf{p}+\tilde{\mathbf{u}}]_h)+i\eta|^{-1}$ denominator. It follows from the definition of the valley index that

$$[M\mathbf{p}+\tilde{\mathbf{u}}]_h=\pm p_{c(h)}\mp p_{\tilde{c}(h)}\pm\ldots,$$

i.e. the β-denominator contains both integration variables. If $\mathfrak{b}^{(h-1)}=\varnothing$, i.e. no point singularity is present, then the prototype of this integral can be estimated as follows:

$$\sup_{|\alpha|,|\beta|\leqslant Y}\int_{\mathbf{R}^d}\int_{\mathbf{R}^d}\frac{d\mu(p)\,d\mu(p')}{|\alpha-\overline{\omega(p)}-i\eta|\,|\alpha-\overline{\omega(p')}-i\eta|\,|\beta-\omega(p-p'+q)+i\eta|}\leqslant C\zeta^{2d-5}|\log\eta|^3,$$

assuming that $|q|\leqslant C\lambda^{-1}$. This estimate follows from (10.28) and (10.24).

If a point singularity is present, $\mathfrak{b}^{(h-1)}\neq\varnothing$, we know that $b_{c(h)}^{(h-1)}=\pm1$ by (10.17), therefore at least one of the integration variables appears in it. Depending on whether $\tilde{c}(h)$ also appears or not, we have one of the following two prototype estimates:

$$\begin{aligned}\sup_{|\alpha|,|\beta|\leqslant Y}\sup_r\int_{\mathbf{R}^d}\int_{\mathbf{R}^d}\frac{1}{|||p-r|||}\frac{d\mu(p)\,d\mu(p')}{|\alpha-\overline{\omega(p)}-i\eta|\,|\alpha-\overline{\omega(p')}-i\eta|\,|\beta-\omega(p-p'+q)+i\eta|}\\ \leqslant C\eta^{-1/2}\zeta^{2d-5}|\log\eta|^3,\end{aligned} \tag{10.30}$$

or

$$\begin{aligned}\sup_{|\alpha|,|\beta|\leqslant Y}\sup_r\int_{\mathbf{R}^d}\int_{\mathbf{R}^d}\frac{1}{|||p\pm p'-r|||}\frac{d\mu(p)\,d\mu(p')}{|\alpha-\overline{\omega(p)}-i\eta|\,|\alpha-\overline{\omega(p')}-i\eta|\,|\beta-\omega(p-p'+q)+i\eta|}\\ \leqslant C\eta^{-1/2}\zeta^{2d-5}|\log\eta|^3,\end{aligned} \tag{10.31}$$

assuming that $|q|\leqslant C\lambda^{-1}$. Both inequalities are obtained by first applying (10.27) for the dp integral, and then applying (10.24) for the dp' integral.

In summary, the effect of the elimination in Case 5 is that

$$E(h-1)\leqslant C\eta^{-1/2}\zeta^{2d-5}|\log\eta|^3E(h),\quad h\in I_v. \tag{10.32}$$

We eliminate the row indices one by one in increasing order according to Cases 1–5. The total estimate is

$$\begin{aligned}E(0)\leqslant{}&(\eta^{-1})^{|I_p|}\lambda^{-2l}C^{|I_t|}(C\zeta^{d-2}\eta^{-1}|\log\eta|)^{|I_{\mathrm{us}}|}(C\eta^{-1/2}\zeta^{d-3}|\log\eta|^2)^{|I_{\mathrm{cs}}|}\\&\times(C\eta^{-1/2}\zeta^{2d-5}|\log\eta|^3)^{|I_v|}E(k+1),\end{aligned} \tag{10.33}$$

where we recall that I_t is the set of indices that are tops of a ladder; in particular, Case 2 is applied $|I_t|$ times. Since the top of the ladder is not a ladder index and not a valley, it is a peak, a slope, or maybe 0, thus

$$|I_t|\leqslant|I_p|+|I_s|+1=v+s+1\leqslant 2v+s,$$

taking into account that $p=v$ and $v\geqslant 1$ since $\sigma\neq\mathrm{id}$. From (10.33) and from the fact that

$$|I_{\mathrm{cs}}|+|I_{\mathrm{us}}|=|I_s|=s$$

it then follows that

$$E(0)\leqslant C^{v+s+1}\lambda^{-2l}\eta^{-(3v+s+|I_{\mathrm{us}}|)/2}\zeta^{d(2v+s)}E(k+1) \tag{10.34}$$

(where we also used that $|\log\eta|^3\ll\zeta$). Note that the exponent of the constant is comparable with $\deg(\sigma)=k-l=2v+s$. A factor C^l would not be affordable if $l\gg\deg(\sigma)$; this is why the ladder integration had to be done essentially with the precise constant by using (3.15).

Finally, when estimating $E(k+1)$, only the last row of M, $h=k+1\in I_{\mathrm{last}}$, is present. Since $k+1$ is bigger than the largest element of $I_s\cup I_v$, which by Definition 10.3 cannot be a covered slope index, we see from the definition of $\mathfrak{b}^{(h)}$ that $\mathfrak{b}^{(h)}=\varnothing$, i.e. there is no point singularity. Moreover, we have already integrated out $|I_l|+|I_s|+2|I_v|=l+s+2v$ variables. Since $p=v$ and $I=I_l\cup I_s\cup I_v\cup I_p\cup I_{\mathrm{last}}$ is a partition, we see that at the last step there is only one integration variable left, namely p_{k+1}. Considering that the $(k+1)$-th row of M contains only one non-trivial entry, and thus $[M\mathbf{p}+\tilde{\mathbf{u}}]_{k+1}=p_{k+1}+\tilde{u}_{k+1}$, the integral from (10.14) is simplified to

$$\begin{aligned}E(k+1)&=E(I^{(k+1)},M^{(k+1)},\mathfrak{b}^{(k+1)})\\&=\lambda^{2k}\sup_{\tilde{u}}\int_{-Y}^{Y}\int_{-Y}^{Y}\int_{\mathbf{R}^d}\frac{1}{|\beta-\omega(p_{k+1}+\tilde{u})+i\eta|}\frac{d\mu(p_{k+1})}{|\alpha-\overline{\omega(p_{k+1})}-i\eta|}\,d\alpha\,d\beta\\&\leqslant C\lambda^{2k}\zeta^d|\log\eta|^2.\end{aligned} \tag{10.35}$$

Now, combining inequalities (10.15), (10.34) and (10.35), we get that

$$E(M) \leqslant C^{v+s+1}\lambda^{2(k-l)}\eta^{-(3v+s+|I_{\mathrm{us}}|)/2}\zeta^{d(2v+s+1)}|\log\eta|^2,$$

where we recall that the general constant C may depend on $\widehat{\psi}_0$. Using the choice of the parameters from (7.25) and the fact that $\deg(\sigma)=k-l=2v+s\geqslant 2$, we obtain a total λ-power

$$\begin{aligned}2(2v+s)-\big(1+\tfrac{1}{2}\varkappa\big)(3v+s+|I_{\mathrm{us}}|)-(\varkappa+3\delta)d(2v+s+1)&\\ \geqslant v+|I_{\mathrm{cs}}|-\varkappa\big(1+\tfrac{3}{2}d\big)(2v+s)-(2v+s)O(\delta),&\end{aligned}$$

where we used Lemma 10.4 and that $v\geqslant 1$ to estimate

$$\tfrac{1}{2}(3v+s+|I_{\mathrm{us}}|)\leqslant 2v+s \quad\text{and}\quad 2v+s+1\leqslant\tfrac{3}{2}(2v+s).$$

Using Lemma 10.4 again, we see that $v+|I_{\mathrm{cs}}|\geqslant\frac{1}{3}(2v+s)$, i.e. that

$$E(M)\leqslant C(\lambda^{13-(1+3d/2)\varkappa-O(\delta)})^{\deg(\sigma)}|\log\eta|^2,$$

if the exponent is positive, i.e. if $\varkappa<2/(6+9d)$ and $\delta\leqslant\delta(\varkappa)$ is sufficiently small. This completes the proof of Lemma 10.1.

Appendix A. Proof of Lemma 10.5

We can replace $\omega(p)$ with $e(p)$ in (10.27) by using a straightforward resolvent expansion

$$\frac{1}{|\alpha-\omega(p)+i\eta|}\leqslant\frac{1}{|\alpha-\lambda^2\Theta(\alpha)-e(p)+i\eta|}\left[1+\frac{C\lambda^2|\alpha-e(p)|^{1/2}}{|\alpha-\omega(p)+i\eta|}\right]\leqslant\frac{C}{|\tilde{\alpha}-e(p)+i\eta|},\quad\text{(A.1)}$$

with $\tilde{\alpha}=\alpha-\lambda^2\operatorname{Re}\Theta(\alpha)$. Here we used the boundedness and the Hölder continuity of Θ, see (3.5). Therefore, the proof of Lemma 10.5 is reduced to the following lemma.

LEMMA A.1. *For any $|q|\leqslant C\lambda^{-1}$ one has*

$$I_1:=\int_{\mathbf{R}^d}\frac{d\mu(p)}{|\alpha-e(p)+i\eta|\,|\beta-e(p+q)+i\eta|}\leqslant\frac{C\zeta^{d-3}|\log\eta|^2}{|||q|||},\quad\text{(A.2)}$$

$$I_2:=\int_{\mathbf{R}^d}\frac{1}{|||p-r|||}\frac{d\mu(p)}{|\alpha-e(p)+i\eta|\,|\beta-e(p+q)+i\eta|}\leqslant\frac{C\eta^{-1/2}\zeta^{d-3}|\log\eta|^2}{|||q|||},\quad\text{(A.3)}$$

uniformly in r, α and β.

Proof. The bound on I_1 follows from a direct calculation and $|q|\leqslant C\lambda^{-1}$:

$$I_1 \leqslant \int_0^\zeta \frac{u^{d-1}\,du}{|\alpha-u^2/2+i\eta|}\int_{-1}^1 \frac{dc}{\bigl|2\beta-(u^2+q^2)-2|q|uc\bigr|+\eta} \leqslant \frac{C\zeta^{d-3}|\log\eta|^2}{|q|}. \tag{A.4}$$

If $\eta\leqslant|q|$, then we use that $|||q|||\leqslant|q|$ to obtain (A.2). If $|q|\leqslant\eta$, then we use Schwarz inequality to separate the denominators and use (3.14) to conclude the proof of (A.2).

To prove (A.3), we first establish the following bound uniformly in α:

$$J := \int_{\mathbf{R}^d} \frac{1}{|||p-r|||}\frac{d\mu(p)}{|\alpha-e(p)+i\eta|} \leqslant C\zeta^{d-2}|\log\eta|, \tag{A.5}$$

which follows by a direct calculation

$$\begin{aligned} J &\leqslant C\int_0^\zeta \frac{u^{d-1}\,du}{|\alpha-u^2/2+i\eta|}\biggl[1+\int_{-1}^1 \frac{dc}{\bigl|u^2+r^2-2|r|uc\bigr|^{1/2}}\biggr] \\ &\leqslant C(\zeta^{d-3}+\zeta^{d-2})\int_0^{\zeta^2}\frac{dv}{|\alpha-v+i\eta|} \leqslant C\zeta^{d-2}|\log\eta|, \end{aligned}$$

with $u=|p|$ and $v=\frac{1}{2}u^2$, and using that $\bigl|u^2+r^2-2|r|uc\bigr|\geqslant|u|^2|1-c^2|$ for $|c|\leqslant1$.

We may assume that $|q|\geqslant\eta$, otherwise we can estimate the β-denominator in (A.3) trivially by η^{-1} and we can conclude with (A.5). We then distinguish two cases. If $|||p-r|||\geqslant\eta^{1/2}$, then we estimate $|||p-r|||$ trivially and we use (A.2).

Now let $|||p-r|||\leqslant\eta^{1/2}$. We split this case into two subcases, $|q|\leqslant2|p|$ and $|q|\leqslant2|p+q|$, whose union clearly covers all values of p.

In the case where $|q|\leqslant2|p|$, we estimate the square root of the β-denominator trivially and use the Schwarz inequality to separate the remaining β-denominator from the point singularity. The corresponding contribution can be estimated by

$$C\eta^{-1/2}\int_{\mathbf{R}^d}\frac{\mathbf{1}(|q|\leqslant2|p|)}{|\alpha-e(p)+i\eta|}\biggl[\frac{1}{|\beta-e(p+q)+i\eta|}+\frac{\mathbf{1}(|||p-r|||\leqslant\eta^{1/2})}{|||p-r|||^2}\biggr]\,d\mu(p).$$

The first term was already estimated in (A.2). The second term is bounded, by the co-area formula, by

$$C\eta^{-1/2}\int_{(|q|/2)^2}^{\zeta^2}\frac{J_a\,da}{|\alpha-a+i\eta|\,|a|^{1/2}}, \quad \text{with } J_a := \int_{\Sigma_a}\frac{\mathbf{1}(|||p-r|||\leqslant\eta^{1/2})}{|||p-r|||^2}\,d\nu(p),$$

where $\Sigma_a:=\{p:e(p)=a\}$ and $d\nu(p)$ is the surface measure. Clearly $J_a\leqslant|\log\eta|$, and we obtain the estimate $C\eta^{-1/2}|\log\eta|^2/|q|$.

In the case where $|q|\leqslant2|p+q|$, we shift $p\mapsto p+q$ and $r\mapsto r-q$, and interchange the role of the α- and β-denominators in the above proof. □

References

[1] Aizenman, M. & Molchanov, S., Localization at large disorder and at extreme energies: an elementary derivation. *Comm. Math. Phys.*, 157 (1993), 245–278.

[2] Aizenman, M., Sims, R. & Warzel, S., Absolutely continuous spectra of quantum tree graphs with weak disorder. *Comm. Math. Phys.*, 264 (2006), 371–389.

[3] Anderson, P., Absences of diffusion in certain random lattices. *Phys. Rev.*, 109 (1958), 1492–1505.

[4] Boldrighini, C., Bunimovich, L. A. & Sinaĭ, Y. G., On the Boltzmann equation for the Lorentz gas. *J. Stat. Phys.*, 32 (1983), 477–501.

[5] Bourgain, J., Random lattice Schrödinger operators with decaying potential: some higher dimensional phenomena, in *Geometric Aspects of Functional Analysis*, Lecture Notes in Math., 1807, pp. 70–98. Springer, Berlin–Heidelberg, 2003.

[6] Brown, R., A brief account of microscopical observations made in the months of June, July, and August 1827, on the particles contained in the pollen of plants and on the general existence of active molecules in organic and inorganic bodies. *Philosophical Magazine* (2), 4 (1828), 161–173.

[7] — Additional remarks on active molecules. *Philosophical Magazine* (2), 6 (1829), 161–166.

[8] Brydges, D., Dimock, J. & Hurd, T. R., The short distance behavior of $(\phi^4)_3$. *Comm. Math. Phys.*, 172 (1995), 143–186.

[9] Brydges, D. & Yau, H.-T., Grad ϕ perturbations of massless Gaussian fields. *Comm. Math. Phys.*, 129 (1990), 351–392.

[10] Bunimovich, L. A. & Sinaĭ, Y. G., Statistical properties of Lorentz gas with periodic configuration of scatterers. *Comm. Math. Phys.*, 78 (1980/81), 479–497.

[11] Chen, T., Localization lengths and Boltzmann limit for the Anderson model at small disorders in dimension 3. *J. Stat. Phys.*, 120 (2005), 279–337.

[12] Denisov, S. A., Absolutely continuous spectrum of multidimensional Schrödinger operator. *Int. Math. Res. Not.*, 74 (2004), 3963–3982.

[13] Dürr, D., Goldstein, S. & Lebowitz, J. L., A mechanical model of Brownian motion. *Comm. Math. Phys.*, 78 (1980/81), 507–530.

[14] — Asymptotic motion of a classical particle in a random potential in two dimensions: Landau model. *Comm. Math. Phys.*, 113 (1987), 209–230.

[15] Einstein, A., Zur Theorie der Brownschen Bewegung. *Ann. der Physik*, 19 (1906), 371–381.

[16] Erdős, L., Linear Boltzmann equation as the long time dynamics of an electron weakly coupled to a phonon field. *J. Stat. Phys.*, 107 (2002), 1043–1127.

[17] Erdős, L., Salmhofer, M. & Yau, H.-T., Towards the quantum Brownian motion, in *Mathematical Physics of Quantum Mechanics*, Lecture Notes in Physics, 690, pp. 233–257. Springer, Berlin–Heidelberg, 2006.

[18] — Quantum diffusion for the Anderson model in the scaling limit. *Ann. Henri Poincaré*, 8 (2007), 621–685.

[19] — Quantum diffusion of the random Schrödinger evolution in the scaling limit. II. The recollision diagrams. *Comm. Math. Phys.*, 271 (2007), 1–53.

[20] Erdős, L. & Yau, H.-T., Linear Boltzmann equation as the weak coupling limit of a random Schrödinger equation. *Comm. Pure Appl. Math.*, 53 (2000), 667–735.

[21] Feldman, J., Knörrer, H. & Trubowitz, E., A representation for fermionic correlation functions. *Comm. Math. Phys.*, 195 (1998), 465–493.

[22] — Convergence of perturbation expansions in Fermionic models. II. Overlapping loops. *Comm. Math. Phys.*, 247 (2004), 243–319.

[23] Feldman, J., Magnen, J., Rivasseau, V. & Sénéor, R., Bounds on completely convergent Euclidean Feynman graphs. *Comm. Math. Phys.*, 98 (1985), 273–288.
[24] — Bounds on renormalized Feynman graphs. *Comm. Math. Phys.*, 100 (1985), 23–55.
[25] — A renormalizable field theory: the massive Gross–Neveu model in two dimensions. *Comm. Math. Phys.*, 103 (1986), 67–103.
[26] — Construction and Borel summability of infrared Φ^4_4 by a phase space expansion. *Comm. Math. Phys.*, 109 (1987), 437–480.
[27] Feldman, J., Salmhofer, M. & Trubowitz, E., Perturbation theory around nonnested Fermi surfaces. I. Keeping the Fermi surface fixed. *J. Stat. Phys.*, 84 (1996), 1209–1336.
[28] — Regularity of interacting nonspherical Fermi surfaces: the full self-energy. *Comm. Pure Appl. Math.*, 52 (1999), 273–324.
[29] 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 (2006), 184–221.
[30] Fröhlich, J. & Spencer, T., Absence of diffusion in the Anderson tight binding model for large disorder or low energy. *Comm. Math. Phys.*, 88 (1983), 151–184.
[31] Gallavotti, G., Rigorous theory of the Boltzmann equation in the Lorentz model. Nota interna n. 358, Physics Department, Università "La Sapienza", Roma, (1972), pp. 1–9. `mp_arc@math.utexas.edu`, #93–304.
[32] Gawędzki, K. & Kupiainen, A., Gross–Neveu model through convergent perturbation expansions. *Comm. Math. Phys.*, 102 (1985), 1–30.
[33] — Massless lattice φ^4_4 theory: rigorous control of a renormalizable asymptotically free model. *Comm. Math. Phys.*, 99 (1985), 197–252.
[34] Gol′dsheĭd, I. J., Molchanov, S. A. & Pastur, L. A., A random homogeneous Schrödinger operator has a pure point spectrum. *Funktsional. Anal. i Prilozhen.*, 11 (1977), 1–10 (Russian); English translation in *Funct. Anal. Appl.*, 11 (1977), 1–8.
[35] Ho, T. G., Landau, L. J. & Wilkins, A. J., On the weak coupling limit for a Fermi gas in a random potential. *Rev. Math. Phys.*, 5 (1993), 209–298.
[36] Kesten, H. & Papanicolaou, G. C., A limit theorem for stochastic acceleration. *Comm. Math. Phys.*, 78 (1980), 19–63.
[37] Kirsch, W. & Martinelli, F., On the essential selfadjointness of stochastic Schrödinger operators. *Duke Math. J.*, 50 (1983), 1255–1260.
[38] Klein, A., Absolutely continuous spectrum in the Anderson model on the Bethe lattice. *Math. Res. Lett.*, 1 (1994), 399–407.
[39] Komorowski, T. & Ryzhik, L., Diffusion in a weakly random Hamiltonian flow. *Comm. Math. Phys.*, 263 (2006), 277–323.
[40] Lanford, O. E., On a derivation of the Boltzmann equation, in *International Conference on Dynamical Systems in Mathematical Physics* (Rennes, 1975), Astérisque, 40, pp. 117–137. Soc. Math. France, Paris, 1976.
[41] Lee, P. A. & Ramakrishnan, T. V., Disordered electronic systems. *Rev. Mod. Phys.*, 57 (1985), 287–337.
[42] Lukkarinen, J. & Spohn, H., Kinetic limit for wave propagation in a random medium. *Arch. Ration. Mech. Anal.*, 183 (2007), 93–162.
[43] Perrin, J. B., Mouvement brownien et réalité moléculaire. *Annales de chimie et de physiqe*, VIII 18 (1909), 5–114.
[44] Rodnianski, I. & Schlag, W., Classical and quantum scattering for a class of long range random potentials. *Int. Math. Res. Not.*, 5 (2003), 243–300.
[45] Salmhofer, M. & Wieczerkowski, C., Positivity and convergence in fermionic quantum field theory. *J. Stat. Phys.*, 99 (2000), 557–586.

[46] Schlag, W., Shubin, C. & Wolff, T., Frequency concentration and location lengths for the Anderson model at small disorders. *J. Anal. Math.*, 88 (2002), 173–220.
[47] Seiler, E., *Gauge Theories as a Problem of Constructive Quantum Field Theory and Statistical Mechanics*. Lecture Notes in Physics, 159. Springer, Berlin–Heidelberg, 1982.
[48] Spohn, H., Derivation of the transport equation for electrons moving through random impurities. *J. Stat. Phys.*, 17 (1977), 385–412.
[49] — The Lorentz process converges to a random flight process. *Comm. Math. Phys.*, 60 (1978), 277–290.
[50] Vollhardt, D. & Wölfle, P., Diagrammatic, self-consistent treatment of the Anderson localization problem in $d\leqslant 2$ dimensions. *Phys. Rev. B*, 22 (1980), 4666–4679.

László Erdős
Institute of Mathematics
University of Munich
Theresienstrasse 39
D-80333 Munich
Germany
lerdos@mathematik.uni-muenchen.de

Manfred Salmhofer
Max Planck Institute for Mathematics
and Theoretical Physics
University of Leipzig
Inselstrasse 22
D-04103 Leipzig
Germany
salmhofer@itp.uni-leipzig.de

Horng-Tzer Yau
Department of Mathematics
Harvard University
Cambridge, MA 02138
U.S.A.
htyau@math.harvard.edu

Received April 21, 2006
Received in revised form April 2, 2007