Classical and quantum Brownian motion
(Lecture Notes)

László Erdős*
Institute of Mathematics, University of Munich,
Theresienstr. 39, D-80333 Munich, Germany

September 22, 2009

Abstract

Einstein’s kinetic theory of the Brownian motion, based upon light water molecules continuously bombarding a heavy pollen, provided an explanation of diffusion from the Newtonian mechanics. Since the discovery of quantum mechanics it has been a challenge to verify the emergence of diffusion from the Schrödinger equation.

The first step in this program is to verify the linear Boltzmann equation as a certain scaling limit of a Schrödinger equation with random potential. In the second step, one considers a longer time scale that corresponds to infinitely many Boltzmann collisions. The intuition is that the Boltzmann equation then converges to a diffusive equation similarly to the central limit theorem for Markov processes with sufficient mixing. In these lectures I will present the mathematical tools to rigorously justify this intuition.

Contents

1 Overview of rigorous derivations of diffusions 2
1.1 The founding fathers: Brown, Einstein, Boltzmann . . . . . . . . . . . . . . . . 3
1.2 Mathematical models of diffusion . . . . . . . . . . . . . . . . . . . . . . . . 5

2 Recapitulation of stochastic processes 8
2.1 The central limit theorem . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 8
2.2 Markov processes and their generators . . . . . . . . . . . . . . . . . . . . . 13
2.3 Wiener process and its generator . . . . . . . . . . . . . . . . . . . . . . . . 15
2.4 Random jump process on the sphere $S^{d-1}$ and its generator . . . . . . 17

3 Classical mechanics of a single particle 18
3.1 The linear Boltzmann equation . . . . . . . . . . . . . . . . . . . . . . . . . 19

*Partially supported by the EU-IHP Network “Analysis and Quantum” HPRN-CT-2002-0027
1 Overview of rigorous derivations of diffusions

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. This investigation has led to the heat equation, the simplest type of diffusion equations:

$$\partial_t u = \Delta_x u$$  \hspace{1cm} (1.1)
where $u(x, t)$ denotes the temperature at position $x \in \mathbb{R}^d$ and time $t$.

1.1 The founding fathers: Brown, Einstein, Boltzmann

The story of diffusion starts with R. Brown in 1827 who observed almost two centuries ago that the motion of a pollen suspended in water was erratic [4]. He saw a picture similar to Fig. 1 under his microscope. He also noted that this apparently chaotic motion never seems to stop. Some people even speculated that this constant “dynamism” may be a signature of life in the pollen, but then similar behaviour was discovered for “dead” particles (as sand particles) as well.

This picture led to the kinetic explanation by Einstein in 1905 [9] 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 [29] 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.

We should make a remark on Boltzmann’s theory although this will not be the subject of this lecture. Boltzmann’s kinetic theory postulates that in a gas of interacting particles at relatively low density the collisions between the particles are statistically independent (Ansatz of molecular chaos). This enabled him to write up the celebrated (nonlinear) Boltzmann equation for the time evolution of the single particle phase space density $f_t(x, v)$. The main assumption is that the collision rate is simply the product of the particle densities with the
Einstein’s explanation

Figure 2: Einstein’s explanation to Brown’s picture

appropriate velocities, thus he arrived at the following equation:

$$\partial_t f_t(x,v) + v \cdot \nabla_x f_t(x,v) = \int \sigma(v,v_1) \left[ f_t(x,v')f_t(x,v_1') - f_t(x,v)f_t(x,v_1) \right]$$

Here $v, v_1$ is the pair of incoming velocities of the colliding particles and $v', v_1'$ are the outgoing velocities. In case of elastic hard balls, the pair $(v', v_1')$ is uniquely determined by $(v, v_1)$ (plus a randomly chosen contact vector) due to energy and momentum conservation, for other type of collisions the kernel $\sigma(v,v_1)$ describes microscopic details of the collision mechanism. The integration is for all variables apart from $x, v$ and subject to constraints (like momentum and energy conservation) encoded in $\sigma(v,v_1)$.

In addition to the highly non-trivial independence Ansatz, this theory also assumes the molecular nature of gases and fluids in contrast to the more intuitive continuous description. Nowadays we can easily accept that gases and fluids on large scales (starting from micrometer scales – these are called microscopic scales) can be characterized by continuous density functions, while on much smaller scales (nanometers and below) the particle nature dominates. But at Boltzmann’s time the particle picture was strongly debated. The ingenuity of Boltzmann’s theory is that he could combine these two pictures. The Boltzmann equation is an equation of continuous nature (as it operates with density functions) but its justification follows an argument using particles. No wonder that it gave rise to so much controversy especially before experiments were available. Moreover, Boltzmann theory implies irreversibility (entropy production) that was for long thought to be a philosophical than physical question.

After this retour on Boltzmann we return to the diffusion models. Before we discuss individual models, let me mention the key conceptual difficulty behind all rigorous derivation of diffusive equations. Note that the Hamiltonian dynamics (either classical or quantum) is reversible and deterministic. The diffusion equation, and also the Boltzmann equation, are
irreversible: there is permanent loss of information along their evolution (recall that the heat equation is usually not solvable backward in time). Moreover these equations exhibit a certain chaotic nature (Brown’s picture). How can these two contradicting facts be reconciled?

The short answer is that the loss of information is due to a **scale separation** and the integration of microscopic degrees of freedom. On the microscopic (particle) level the dynamics remains reversible. The continuous (fluid) equations live on the macroscopic (fluid) scale: they are obtained by neglecting many degrees of freedom on short scales. Once we switch to the fluid description, the information in these degrees of freedom is lost forever.

This two-level explanation foreshadows that for a rigorous mathematical description one would need a scale separation parameter, usually called \( \varepsilon \) that describes the ratio between the typical microscopic and macroscopic scales. In practice this ratio is

\[
\varepsilon = \frac{1 \text{ Angstrom}}{1 \text{ cm}} = 10^{-8},
\]

but mathematically we will consider the \( \varepsilon \to 0 \) so-called **scaling limit**.

Once the scale separation is put on a solid mathematical ground, we should note another key property of the macroscopic evolution equations we will derive, namely their Markovian character. Both the heat equation (1.1) and the nonlinear Boltzmann equation (1.2) give the time derivative of the macroscopic density at any fixed time \( t \) in terms of the density at the same time. I.e. these evolution equations express a process where the future state depends only on the present state. The dependence of the future on the past is only indirect through the present state. We will call this feature **Markovian property** or in short **Markovity**. Note that colliding particles do build up a memory along their evolution: the state of the recolliding particles will “remember” their previous collisions. This effect will have to be supressed to obtain a Markovian evolution equation, as it was already recognized by Boltzmann. We will thus consider systems where the rate of recollisions is supressed. There are essentially two ways to achieve it: either by going to a low density situation or by introducing a small coupling constant.

Apart from the rigorous formulation of the scaling limit, we thus will have to cope with the main technical difficulty: **controlling memory (recollision) effects**. Furthermore, specifically in quantum models, we will have to control interference effects as well.

Finally, we should remark that Einstein’s model is simpler than Boltzmann’s, as light particles do not interact (roughly speaking, Boltzmann’s model is similar to Fig. 2 but the light particles can also collide with each other and not only with the heavy particle). Thus to verify the Ansatz of molecular chaos is technically easier in Einstein’s model. Still, Einstein’s model addresses the key issue: how does diffusion emerge from Hamiltonian mechanics?

### 1.2 Mathematical models of diffusion

The first step to setup a correct mathematical model is to recognize that some stochasticity has to be added. Although we expect that for a “typical” initial configuration the diffusion equations are correct, this certainly will **not hold for all** initial configuration. We expect that after opening the door between a warm and a cold room, the temperature will equilibrate (thermalize) but certainly there exists a “bad” initial configuration of the participating \( N \sim \)
$10^{23}$ particles such that all “warm” particles will head towards the cold room and vice versa; i.e. the two room temperatures will be exchanged instead of thermalization. Such configuration are extremely rare; their measure in all reasonable sense goes to zero very fast as $N \to \infty$, but nevertheless they are there and prevent us from deriving diffusion equations for all initial configurations.

The stochasticity can be added in different ways and this largely determines the technical complications involved in the derivation of the diffusion. In general, more stochasticity is added, the easier the derivation is.

The easiest if the dynamics itself is stochastic; the typical example being the random walk and its scaling limit, the Wiener process (theorized by Wiener in 1923). In the typical examples, the random process governing the dynamics of the system has no correlation. The whole dynamics is Markovian, but sometimes the dynamics of the one particle densities is not fully Markovian (e.g. stochastic lattice gases), although it typically has strong statistical mixing properties. Of course this dynamics is not Hamiltonian (no energy conservation and no reversibility), and thus it is not particularly surprising that after some scaling limit one obtains a diffusion equation.

The next level of difficulty is a Hamiltonian system with random Hamiltonian. Some physical data in the Hamiltonian (e.g. potential or magnetic field) is randomly selected (describing a disordered system), but then they are frozen once for all and then the Hamiltonian dynamics starts. Typical model is the Lorenz gas, where a single Hamiltonian particle is moving among fixed random scatterers. Recollisions with the same scatterer are possible and this effect potentially ruins the Markovity. Thus this model is usually considered in the weak coupling limit, i.e. the coupling parameter between the test-particle and the random scatterers is set to converge to zero, $\lambda \to 0$, and simultaneously a long time limit is considered. It turns out that if

$$t \sim \lambda^{-2}$$

then one can prove a nontrivial diffusive limit dynamics (see Kesten/Papanicolaou [23] and Komorowski-Ryzhik [25]). The relation (1.3) between the coupling and the time scale is called the van Hove limit. Similar weakening of the recollision can be achieved by the low density limit (see later).

The following level of difficulty is a deterministic Hamiltonian with random initial data of many degrees of freedom. The typical example is Einstein’s model, where the test-particle is immersed in a heat bath. The heat bath is called ideal gas, if it is characterized by an non-interacting Hamiltonian, $H_{bath}$. The initial state is typically a temperature equilibrium state, $\exp(-\beta H_{bath})$ i.e. The initial data of the heat-bath particles are given by this equilibrium measure. Again, some scaling limit is necessary to reduce the chances of recollisions, one can for example consider the limit $m/M \to \infty$, where $M$ and $m$ are the mass of the test-particle and the heat-bath particles, respectively. Since a collision with a single light particle does not have sizable effect on the motion of the heavy particle, the rate of collisions has to be increased in parallel with $m/M \to 0$. Similarly to the van Hove limit, there is a natural scaling limit, where nontrivial limiting dynamics was proven by Dürr-Goldstein-Lebowitz [8]. We remark that the problem becomes enormously more complicated if the heat-bath
particles can interact among each other, i.e. if we are facing a truly interacting many-body dynamics. In this case there is even no need to distinguish between the tracer particle and the heat-bath particles, in the simplest model one just considers identical particles interacting via a two-body interaction and one investigates the single-particle density function $f_t(x,v)$. In a certain scaling limit regime, the model should lead to the nonlinear Boltzmann equation (1.2). This has only been proven for short time more by Lanford [26]. His work appeared in 1975 and since then nobody could extend the proof to include longer time scales. We mention that the complications are partly due to the fact that the nonlinear Boltzmann equation itself does not have a satisfactory long time existence theory.

Finally, the most challenging (classical) model is the deterministic Hamiltonian with a random initial data of a few degrees of freedom, the typical example being the various mathematical billiards. The simplest billiard is the hard-core periodic Lorenz gas (also called Sinai’s billiard [6]), where the scatterers are arranged in a periodic lattice and a single point particle (“billiard ball”) moves among these scatterers according to the usual rules of specular reflections. The methods of proofs here originate more in dynamical system than in statistical mechanics.

All these models have natural quantum mechanical analogues; for the Hamiltonian systems they simply follow from standard quantization of the classical Hamiltonian. These models are summarized in the table below and they are also illustrated in Fig. 3. We note that the quantum analogue of the periodic Lorenz gas is ballistic due to the Bloch waves (a certain diagonalization procedure similar to the Fourier transform), thus in this case the classical and quantum models behave differently; the quantum case being relatively trivial.
2 Recapitulation of stochastic processes

Our goal is to understand the stochastic features of quantum dynamics. A natural way to do it is to compare quantum dynamics with a much simpler stochastic dynamics whose properties are well-known or much easier to establish. In this section the most important ingredients of elementary stochastic processes are collected and we will introduce the Wiener process. I assume basic knowledge of probability theory (random variables, expectation, variance, independence, Gaussian normal random variable, characteristic function).

2.1 The central limit theorem

Let \( v_i, i = 1, 2, \ldots \), be a sequence of independent, identically distributed (denoted by i.i.d. for brevity) random variables with zero expectation, \( \mathbf{E}v_i = 0 \) and finite variance, \( \mathbf{E}v_i^2 = \sigma^2 \). The values of \( v_i \) are either in \( \mathbb{R}^d \) or \( \mathbb{Z}^d \). Think of \( v_i \) being velocities or steps of a walking particle at discrete time \( i \). For simplicity we will work in \( d = 1 \) dimensions, the extension to the general case is easy.

One natural question: Where is the particle after \( n \) steps if \( n \) is big?

Let

\[
S_n := \sum_{i=1}^{n} v_i
\]
Random walk (Wiener)  
Stochastic dynamics

Lorenz gas (random scat.)  
Random Hamiltonian, one body

Einstein’s model  
Deterministic Hamiltonian  
many body random data.

Periodic Lorenz gas (billiard)  
Deterministic Hamiltonian  
one body random data.

Figure 3: Microscopic models for diffusion
be the location of the particle after \( n \) steps. It is clear that the expectation is zero, \( \mathbb{E}S_n = 0 \) and the variance is

\[
\mathbb{E}S_n^2 = \mathbb{E}\left( \sum_{i=1}^{n} v_i \right)^2 = \sum_{i=1}^{n} \mathbb{E}v_i^2 = n\sigma^2
\]

(note that independence via \( \mathbb{E}v_i v_j = 0 \) for \( i \neq j \) has been used).

This suggests to rescale \( S_n \) and define

\[
X_n := \frac{S_n}{\sqrt{n}} = \frac{\sum_{i=1}^{n} v_i}{\sqrt{n}}
\]

then \( \mathbb{E}X_n = 0 \) and \( \mathbb{E}X_n^2 = \sigma^2 \). The question is the distribution of \( X_n \). The surprising fact is that this distribution tends to a universal one (namely to the Gaussian normal distribution) as \( n \to \infty \), no matter what the initial distribution of \( v_i \) was! This is the central limit theorem, one of the fundamental theorems of not just mathematics but of Nature.

**Theorem 2.1.** Under the conditions \( \mathbb{E}v_i = 0 \) and \( \mathbb{E}v_i^2 = \sigma^2 < \infty \) the distribution of the rescaled sum \( X_n \) of the i.i.d random variables \( v_i \) converges to the normal distribution, i.e

\[
X_n \Rightarrow X
\]

in the sense of distributions where \( X \) is a random variable distributed according to \( N(0, \sigma^2) \).

We recall that the density of the normal distribution in \( \mathbb{R}^1 \) is given by

\[
f(x) = \frac{1}{(2\pi\sigma^2)^{1/2}} e^{-\frac{x^2}{2\sigma^2}}
\]

and the convergence is the standard weak convergence of probability measures (or convergence in distribution or convergence in law).

**Definition 2.2.** We say that a sequence of random variables \( X_n \) converge to the random variable \( X \) in distribution, \( X_n \Rightarrow X \) if

\[
\mathbb{E}G(X_n) \to \mathbb{E}G(X) = \int G(x)f(x)dx
\]

for any continuous bounded function \( G : \mathbb{R}^d \to \mathbb{R} \).

In particular, \( X_n \) converges in distribution to \( N(0, \sigma^2) \) if

\[
\mathbb{E}G(X_n) \to \int G(x)f(x)dx
\]

*Sketch of the proof.* For simplicity we work in \( d = 1 \) dimension. For any \( \xi \in \mathbb{R}^d \), let

\[
\varphi(\xi) := \mathbb{E} e^{i\xi v}
\]

10
be the characteristic function (Fourier transform) of the random variable \( v \) (distributed by
the common distribution of all \( v_i \)'s). Then the characteristic function of \( X_n \) is

\[
E e^{i\xi X_n} = E \prod_{j=1}^{n} e^{\frac{\xi}{\sqrt{n}} v_j} = \varphi \left( \frac{\xi}{\sqrt{n}} \right)^n
\]

\[
= \left( \varphi(0) + \varphi'(0) \frac{\xi}{\sqrt{n}} + \frac{1}{2} \varphi''(0) \left( \frac{\xi}{\sqrt{n}} \right)^2 + o(n^{-1}) \right)^n
\]

\[
= \left( 1 - \frac{\sigma^2 \xi^2}{2n} + o(n^{-1}) \right)^n \rightarrow e^{-\frac{1}{2} \sigma^2 \xi^2}
\]

by Taylor expansion and by using that \( \varphi'(0) = iE v = 0 \) and \( \varphi''(0) = -E v^2 = -\sigma^2 \).

We note that \( e^{-\frac{1}{2} \sigma^2 \xi^2} \) is exactly the characteristic function of a normal Gaussian random variable with variance \( \sigma^2 \). Thus the theorem follows from the fact that characteristic functions uniquely determine the distribution (Fourier transform of a measure determines the measure). \( \square \)

Recall that we will have two scales: a microscopic and a macroscopic one. In this example, the microscopic scale corresponds to one step of the walking particle and in the above model \( n \) was the total microscopic time. The macroscopic time will be the natural time scale of the limit process and it will be kept order one even as \( n \to \infty \). Recalling that we introduced \( \varepsilon \ll 1 \) as a scaling parameter. Now we reformulate the central limit theorem in this language.

Let \( T > 0 \) be a macroscopic time (think of it as an unscaled, order one quantity) and let

\[ n := \lfloor T \varepsilon^{-1} \rfloor \]

where \( \lfloor \cdot \rfloor \) denotes the integer part. Let again \( E v_i = 0, \ E v_i^2 = \sigma^2 \). From the central limit theorem it directly follows that

\[ \tilde{X}_T := \frac{1}{\varepsilon^{-1/2}} \sum_{i=1}^{[T \varepsilon^{-1}]} v_i \implies X_T, \text{ as } \varepsilon \to 0 \quad (2.1) \]

where \( X_T \) is a normal Gaussian variable with variance \( T \sigma^2 \) and with density function

\[
f_T(X) = \frac{1}{(2\pi \sigma^2 T)^{1/2}} \exp \left( -\frac{X^2}{2\sigma^2 T} \right)
\]

Note that we normalized only by \( \sqrt{\varepsilon^{-1}} \) and not by \( \sqrt{n} \approx \sqrt{T \varepsilon^{-1}} \), therefore the limit depends on \( T \) but of course not on \( \varepsilon \).

Since the Gaussian function \( f_T(X) \) gives the probability of finding \( X_T \) at location \( x \), we have

\[ f_T(X) dX \approx \text{Prob}\{ \text{finding } \tilde{X}_T \text{ at } X + dX \text{ at time } T \} \]

Note that we used macroscopic space and time coordinates, \( X \) and \( T \). Translating this statement into the original process in microscopic coordinates we see that

\[ f_T(X) dX \approx \text{Prob}\{ \text{finding } S_n \text{ at } \varepsilon^{-1/2}(X + dX) \text{ at time } n \approx \varepsilon^{-1} T \} \]
Note that the space and time are rescaled differently; if \( (X, T) \) denote the macroscopic space/time coordinates and \( (x, t) \) denote the microscopic coordinates, then
\[
x = \varepsilon^{-1/2} X, \quad t = \varepsilon^{-1} T
\]
This is the typical diffusive scaling, where
\[
\text{time} = (\text{distance})^2
\]  
(2.2)
Finally, we point out that the function \( f_T(X) \) satisfies the heat equation with diffusion constant \( \sigma^2 \):
\[
\partial_T f_T(X) = \sigma^2 \Delta_X f_T(X)
\]  
(2.3)
The emergence of the normal distribution and then the heat equation is universal; note that apart from the variance (and the zero expectation) no other information on the distribution of the microscopic step \( v \) was used. The details of the microscopic dynamics are wiped out and only the essential feature, the heat equation (or the normal distribution) with a certain diffusion coefficient, remains visible after the \( \varepsilon \to 0 \) scaling limit.

We remark that essentially the same proof works if \( v_i \)'s have some short time correlation. Assume that \( v_i \) is a stationary sequence of random variables with zero mean, \( \mathbf{E} v_i = 0 \), and let
\[
R(i, j) := \mathbf{E} v_i v_j
\]
the correlation function. By stationarity, \( R(i, j) \) depends only on the difference, \( i - j \), so let us use the notation
\[
R(i - j) := \mathbf{E} v_i v_j
\]
This function is also called the velocity-velocity autocorrelation function.

We just formulate the result here

**Theorem 2.3.** Assume that the velocity-velocity autocorrelation function is summable, i.e.
\[
\sum_{k=-\infty}^{\infty} R(k) < \infty
\]
Then the rescaled sums
\[
\tilde{X}_T^\varepsilon := \varepsilon^{1/2} \sum_{i=1}^{[T\varepsilon^{-1}]} v_i
\]
converge in distribution to a normal random variable,
\[
\tilde{X}_T^\varepsilon \Rightarrow X_T, \quad \varepsilon \to 0
\]
whose density function \( f_T(X) \) satisfies the heat equation
\[
\partial_T f = D \Delta f
\]
with diffusion coefficient

\[ D := \sum_{k=-\infty}^{\infty} R(k) \]

assuming that \( D \) is finite. This relation between the diffusion coefficient and the sum (integral) of the velocity-velocity autocorrelation function is called the Green-Kubo formula.

Finally we mention that all results remain valid in higher dimensions, \( d > 1 \). The role of the diffusion coefficient is taken over by a diffusion matrix \( D_{ij} \) defined by the correlation matrix of the single step distribution:

\[ D_{ij} = \mathbb{E} v^{(i)} v^{(j)} \]

where \( v = (v^{(1)}, v^{(2)}, \ldots, v^{(d)}) \). The limiting heat equation (2.3) in the uncorrelated case is modified to

\[ \partial_t f_T(X) = \sum_{i,j=1}^{d} D_{ij} \partial_i \partial_j f_T(X) \]

### 2.2 Markov processes and their generators

Let \( X_t, t \geq 0 \), be a continuous-time stochastic process, i.e. \( X_t \) is a one-parameter family of random variables, the parameter is usually called time. The state space of the process is the space from where \( X_t \) takes its values, in our case \( X_t \in \mathbb{R}^d \) or \( \mathbb{Z}^d \). As usual, \( \mathbb{E} \) will denote the expectation with respect to the distribution of the process.

**Definition 2.4.** \( X_t \) is a Markov process if for any \( t \geq \tau \)

\[ \text{Dist}(X_t \mid \{X_s\}_{s \in [0,\tau]}) = \text{Dist}(X_t \mid X_\tau) \]

i.e. if the conditional distribution of \( X_t \) conditioned on the family of events \( X_s \) in times \( s \in [0,\tau] \) is the same as conditioned only on the event at time \( \tau \).

In simple words it means that the state of the system at time \( t \), \( X_t \), depends on the past between times \([0, \tau]\) only via the state of the system at the last time of the past interval, \( \tau \). All information about the past is condensed in the last moment.

The process depends on its initial value, \( X_0 \). Let \( \mathbb{E}_x \) denote the expectation value assuming that the process started from the point \( x \in \mathbb{R}^d, \mathbb{Z}^d \), i.e.

\[ \mathbb{E}_x \varphi(X_t) = \mathbb{E}\{\varphi(X_t) \mid X_0 = x\} \]

Here and in the sequel \( \varphi \) will denote functions on the state space, these are also called observables. We will not specify exactly the space of observables.

Markov processes are best described by their generators:

**Definition 2.5.** The generator of the Markov process \( X_t \) is an operator acting on the observables \( \varphi \) and it is given by

\[ (\mathcal{L} \varphi)(x) := \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0+0} \mathbb{E}_x \varphi(X_\varepsilon) \]
This definition is a bit lousy since the function space on which $L$ act is not defined. Typically it is either the space of continuous functions or the space of $L^2$ functions on the state space, but for both concept an extra structure – topology or measure – is needed on the state space. Moreover, the generator is typically an unbounded operator, not defined on the whole function space but only on a dense subset. For this lecture we will leave these technicalities aside, but when necessary, we will think of the state space $\mathbb{R}^d, \mathbb{Z}^d$ equipped with their natural measures and the space of functions we consider will the $L^2$-space.

The key fact is that the generator tells us everything about the process itself. Let us demonstrate it by answering two natural questions in terms of the generator.

**Question 1:** Starting from $X_0 = x$, what is the distribution of $X_t$?

**Answer:** Let $\varphi$ be an observable (function on the state space) and define

$$f_t(x) := \mathbb{E}_x \varphi(X_t) \quad (2.4)$$

We wish to derive an evolution equation for $f_t$:

$$\partial_t f_t(x) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \mathbb{E}_x \left[ \mathbb{E}_{X_\varepsilon} \varphi(\tilde{X}_t) - \mathbb{E}_x \varphi(X_t) \right]$$

since by the Markov property

$$\mathbb{E}_x \varphi(X_{t+\varepsilon}) = \mathbb{E}_x \mathbb{E}_{X_\varepsilon} (\varphi(\tilde{X}_t))$$

where $\tilde{X}_t$ is a new copy of the Markov process started at the (random) point $X_\varepsilon$. Thus

$$\partial_t f_t(x) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \mathbb{E}_x \left[ f_t(X_\varepsilon) - f_t(x) \right] = \frac{d}{d\varepsilon} \bigg|_{\varepsilon=0+} \mathbb{E}_x f_t(X_\varepsilon) = (L f_t)(x)$$

Therefore $f_t$, defined in (2.4) solves the initial value problem

$$\partial_t f_t = L f_t$$

$$f_0(x) = \varphi(x)$$

Formally one can write

$$f_t = e^{tL} \varphi$$

If the observable is a Dirac delta function at a fixed point $y$ in the state space,

$$\varphi(x) = \delta_y(x)$$

then the solution the above initial value problem is called the *transition kernel* of the Markov process and it is denoted by $p_t(x, y)$:

$$\partial_t p_t(\cdot, y) = L p_t(\cdot, y)$$

$$p_0(\cdot, y) = \delta_y$$

14
The intuitive meaning of the transition kernel is
\[ p_t(x, y) \, dy := \text{Prob}\{\text{After time } t \text{ the process is at } y + dy \text{ if it started at } x \text{ at time } 0\} \]

**Question 2:** Suppose that the initial value of the process, \( X_0 = x \), is distributed according to a function \( \psi(x) \). What is the probability that after time \( t \) the process is at \( y \)?

**Answer:**
\[ g_t(y) := \text{Prob}(\text{to be at } y \text{ after time } t) = \int \psi(x) p_t(x, y) \, dx \]

It is a homework exercise to check (at least formally), that \( g_t \) solves the following initial value problem
\[ \frac{\partial_t g_t}{L^* g_t} \]
\[ g_0 = \psi \]

where \( L^* \) denotes the adjoint of \( L \) (with respect to the standard scalar product of the \( L^2 \) of the state space).

### 2.3 Wiener process and its generator

The Wiener process is the rigorous mathematical definition of the Brownian motion. There are various ways to introduce it, we choose the shortest (but not the most intuitive) definition.

First we need the concept of Gaussian process. We recall that the centered Gaussian random variables have the remarkable property that all their moments are determined by the covariance matrix (a random variable is called centered if its expectation is zero, \( \mathbb{E} X = 0 \)). If \((X_1, X_2, \ldots, X_{2k})\) is a centered Gaussian vector-valued random variable, then the higher moments can be computed by Wick’s formula from the second moments (or two-point correlations)

\[ \mathbb{E} X_1 X_2 \ldots X_{2k} = \sum_{\pi \text{ pairing}} \prod_{(i,j) \in \pi} \mathbb{E} X_i X_j \quad (2.5) \]

where the summation is over all possible (unordered) pairings of the indices \( \{1, 2, \ldots, 2k\} \).

A stochastic process \( X_t \) is called Gaussian, if any finite dimensional projection is a Gaussian vector-valued random variable, i.e. for any \( t_1 < t_2 < \ldots \leq t_k \), the vector \((X_{t_1}, X_{t_2}, \ldots, X_{t_n})\) is a Gaussian random variable.

**Definition 2.6.** A continuous Gaussian stochastic process \( W_t \in \mathbb{R}^d, t \geq 0 \), is called the \( d \)-dimensional Wiener process if it satisfies the following

i) \( W_0 = 0 \)

ii) \( \mathbb{E} W_t = 0 \)

iii) \( \mathbb{E} W_s W_t = \min\{s, t\} \)

From Wick’s formula it is clear that all higher moments of \( W_t \) are determined. It is a fact that the above list of requirements determine the Wiener process uniquely.

We can now extend the central limit theorem to processes.
Theorem 2.7. Recall the conditions of 2.1. The stochastic process
\[ \tilde{X}_T^\varepsilon := \varepsilon^{1/2} \sum_{i=1}^{[T\varepsilon^{-1}]} v_i \]
(defined already in (2.1)) converges in distribution (as a stochastic process) to the Wiener process
\[ \tilde{X}_T^\varepsilon \Rightarrow W_T \]
as \( \varepsilon \to 0 \).

The proof of this theorem is not trivial. It is fairly easy to check that the moments converge, i.e. for any \( k \)
\[ E[\tilde{X}_T^\varepsilon]^k \to EW_T^k \]
or even for different times:
\[ E \tilde{X}_{T_1}^\varepsilon \tilde{X}_{T_2}^\varepsilon \ldots \tilde{X}_{T_k}^\varepsilon \to EW_{T_1}W_{T_2} \ldots W_{T_k} \]
but this is not quite enough. The reason is that a continuous time process is a collection of uncountable many random variables and the joint probability measure on such a big space is not necessarily determined by finite dimensional projections. If the process has some additional regularity property, then yes. The correct notion is the stochastic equicontinuity (or Kolmogorov condition) which provides the necessary compactness. We will not go into more details here.

Theorem 2.8. The Wiener process is Markovian with generator
\[ \mathcal{L} = \frac{1}{2} \Delta \]

Idea of the proof. The Markovity can be easily checked from the explicit formula for the correlation function. The most important ingredient is that the Wiener process has independent increments:
\[ E(W_t - W_s)W_u = 0 \]  \( (2.6) \)
if \( u \leq s \leq t \); i.e. the increment in the time interval \([s, t]\) is independent of the past. The formula (2.6) is trivial to check from property (iii).

Now we compute the generator using Definition 2.5 (we work in \( d = 1 \) dimension for simplicity)
\[ \frac{d}{d\varepsilon} \bigg|_{\varepsilon=0+0} E_0\varphi(W_\varepsilon) = \frac{d}{d\varepsilon} \bigg|_{\varepsilon=0+0} E_0 \left( \varphi(0) + \varphi'(0)W_\varepsilon + \frac{1}{2}\varphi''(0)W_\varepsilon^2 + \ldots \right) = \frac{1}{2}\varphi''(0) \]
Here we used that \( E_0W_\varepsilon = 0, E_0W_\varepsilon^2 = \varepsilon \) and the dots refer to irrelevant terms that are higher order in \( \varepsilon \). □

Finally we remark that one can easily define a Wiener process with a nontrivial (non-identity) diffusion coefficient matrix \( D_{ij} \). The matrix must be positive definite. The generator is then
\[ \mathcal{L} = \frac{1}{2} \sum_{i,j=1}^{d} D_{ij}\partial_i\partial_j \]
2.4 Random jump process on the sphere $S^{d-1}$ and its generator

The state space of this process is the unit sphere $S = S^{d-1}$ in $\mathbb{R}^d$. We are given a function

$$\sigma(v, u) : S \times S \rightarrow \mathbb{R}_+$$

which is interpreted as jump rate from the point $v$ to $u$.

The process will be parametrized by a continuous time $t$, but, unlike the Wiener process, it will not be continuous, rather piecewise constant with some jumps at a discrete set of times. The good intuitive picture is that the particle jumping on the unit sphere has a random clock (so-called exponential clock) in its pocket. The clock emits a signal with the standard exponential distribution, i.e.

$$\text{Prob}\{\text{there is a signal at } t + dt\} = e^{-t}dt$$

and when it ticks, the particle, being at location $v$, jumps to any other location $u \in S$ according to the distribution $u \rightarrow \sigma(v, u)$ (with the appropriate normalization):

$$\text{Prob}\{\text{The particle from } v \text{ jumps to } u + du\} = \frac{\sigma(v, u)du}{\int \sigma(v, u)du}$$

The transition of the process at an infinitesimal time increment is given by

$$v_{t+\varepsilon} = \begin{cases} u + du & \text{with probability } \varepsilon \sigma(v_t, u)du \\
v_t & \text{with probability } 1 - \varepsilon \int \sigma(v_t, u)du \end{cases}$$

(2.7)

as $\varepsilon \to 0$.

Thus the generator of the process can be computed from Definition 2.5. Let

$$f_t(v) := \mathbb{E}_v \varphi(v_t)$$

assuming that the process starts from some given point $v_0 = v$ and $\varphi$ is an arbitrary observable. Then

$$\partial_t f_t(v) = \lim_{\varepsilon \to 0+0} \frac{1}{\varepsilon} \mathbb{E}_v \left( \mathbb{E}_{v_t} \varphi(v_t) - \mathbb{E}_v \varphi(v_t) \right)$$

$$= \int du \sigma(v, u) \left[ \mathbb{E}_u \varphi(v_t) - \mathbb{E}_v \varphi(v_t) \right]$$

$$= \int du \sigma(v, u) \left[ f_t(u) - f_t(v) \right]$$

where we used the Markov property in the first line and the infinitesimal jump rate from (2.7) in the second line. Note that with probability $1 - \varepsilon \sigma(v_t, u)$ we have $v_\varepsilon = v$. Thus the generator of the random jump process is

$$(\mathcal{L}f)(v) := \int du \sigma(v, u) \left[ f(u) - f(v) \right]$$

Note that it has two terms; the first term is called the gain term the second one is the loss term.
The evolution equation for the time dependent probability density of the jump process,
\[
\partial_t f_t(v) = \int du \sigma(v, u)[f_t(u) - f_t(v)]
\] \hspace{1cm} (2.8)
is called linear Boltzmann equation in velocity space.

The elements of the state space \( S \) will be interpreted as velocities of a moving particle undergoing fictitious random collisions. Only the velocities are recorded.

3 Classical mechanics of a single particle

Classical mechanics of a single particle in \( d \)-dimensions is given by a Hamiltonian (energy function) defined on the phase space \( \mathbb{R}^d \times \mathbb{R}^d \):
\[
H(v, x) := \frac{1}{2}v^2 + U(x)
\] \hspace{1cm} (3.1)
The Hamiltonian equation of motions is the following set of \( 2d \) coupled first order differential equations:
\[
\dot{x}(t) = \partial_v H = v \hspace{1cm} \dot{v}(t) = -\partial_x H = -\nabla U(x)
\] \hspace{1cm} (3.2)
For example, in case of the free evolution, when the potential is zero, \( U \equiv 0 \), we have
\[
x(t) = x_0 + v_0 t
\] \hspace{1cm} (3.3)
i.e. linear motion with a constant speed.

Instead of describing each point particle separately, we can think of a continuum of (non-interacting) single particles, described by a phase space density \( f(x, v) \). This function is interpreted to give the number of particles with given velocity at a given position, more precisely
\[
\int_{\Delta} f(x, v) dx dv = \text{Number of particles at } x \text{ with velocity } v \text{ such that } (x, v) \in \Delta
\]
Another interpretation is that a single particle with velocity \( v_0 \) at \( x_0 \) can be described by the measure
\[
f(x, v) = \delta(x - x_0)\delta(v - v_0)
\]
and it is of course natural to consider more general measures as well.

The system evolves with time, so is the density, i.e. we consider the time dependent phase space density, \( f_t(x, v) \). For example in case of a single particle governed by (3.2) with initial condition \( x(0) = x_0 \) and \( v(0) = v_0 \), the evolution of the phase space density is given by
\[
f_t(x, v) = \delta(x - x(t))\delta(v - v(t))
\]
where \( (x(t), v(t)) \) is the Hamiltonian trajectory computed from (3.2).

It is easy to check that if the point particle trajectories satisfy (3.2), then \( f_t \) satisfies the following evolution equation
\[
(\partial_t + v \cdot \nabla_x) f_t(x, v) = \nabla U(x) \cdot \nabla_v f_t(x, v)
\] \hspace{1cm} (3.4)
which is called the \textit{Liouville equation}. The left hand side is called the \textit{free streaming term}. The solution to

\[(\partial_t + v \cdot \nabla_x) f_t(x, v) = 0\] (3.5)

is given by the linear transport solution

\[f_t(x, v) = f_0(x - vt, v)\]

where \(f_0\) is the phase space density at time zero. This corresponds to the free evolution (3.3) in the Hamiltonian particle representation.

3.1 The linear Boltzmann equation

The linear Boltzmann equation is a phenomenological combination of the free flight equation (3.5) and the jump process on the sphere of the velocity space (2.8)

\[(\partial_t + v \cdot \nabla_x) f_t(x, v) = \int \sigma(u, v) [f_t(x, u) - f_t(x, v)] \, du\] (3.6)

Note that we actually have the adjoint of the jump process (observe that \(u\) and \(v\) are interchanged compared with (2.8)), so the solution determines:

\[f_t(x, v) = \text{Prob\{to be at \((x, v)\) at time \(t\}\}}\]

given the initial probability density \(f_0(x, v)\) (Question 2 in Section 2.2).

The requirement that the velocities are constrained to the sphere corresponds to energy conservation. Of course there is no real Hamiltonian dynamics behind this equation: the jumps are stochastic.

Notice that the free flight plus collision process standing behind the Boltzmann equation is actually a random walk. In the standard random walk the particle “jumps” in a new random direction after every unit time. In the Boltzmann process the jumps occur at random times (“exponential clock”), but the effect on long time is the same. In particular, the long time evolution of the linear Boltzmann equation is diffusion (Wiener process) in position space.

The following theorem formulates this fact more precisely. We recall that to translate the velocity process into position space, one has to integrate the velocity, i.e. will consider

\[x_t = \int_0^t v_s ds\]

\textbf{Theorem 3.1.} Let \(v_t\) be a random velocity jump process given by the generator (2.8). Then

\[X_\varepsilon(T) =: \varepsilon^{1/2} \int_0^{T/\varepsilon} v_t \, dt \to W_T \quad \text{(in distr.)}\]

where \(W_T\) is a Wiener process with diffusion coefficient being the velocity autocorrelation

\[D = \int_0^\infty R(t) dt, \quad R(t) := E v_0 v_t\]

Here \(E\) is with respect to the equilibrium measure of the jump process.
4 Quantum mechanics of a single particle

4.1 Wavefunction, Wigner transform

The state space of a quantum particle in $d$-dimensions is $L^2(\mathbb{R}^d)$. Its elements are called $L^2$-wavefunctions and they are usually denoted by $\psi = \psi(x) \in L^2(\mathbb{R}^d)$. We assume the normalization, $\|\psi\|_2 = 1$. We recall the interpretation of the wave function: the position space density $|\psi(x)|^2 dx$ gives the probability of finding an electron at a spatial domain:

$$\int_{\Delta} |\psi(x)|^2 dx = \text{Prob}\{\text{the particle's position is in } \Delta\}$$

Similarly the Fourier transform of $\psi$,

$$\hat{\psi}(v) := \int_{\mathbb{R}^d} e^{-iv\cdot x} \psi(x) dx$$

determines the momentum space density:

$$\int_{\Delta} |\hat{\psi}(v)|^2 dv = \text{Prob}\{\text{the particle's momentum is in } \Delta\}$$

By the Heisenberg uncertainty principle, one cannot simultaneously determine the position and the momentum of a quantum particle, thus the concept of classical phase space density does not generalize directly to quantum mechanics. Nevertheless one can define a substitute for it, namely the Wigner transform. For any $L^2$-wavefunction $\psi$ we define the Wigner transform of $\psi$ as

$$W_\psi(x, v) := \int_\mathbb{R} \overline{\psi(x + z/2)} \psi(x - z/2) e^{ivz} dz$$

and we still interpret it as “quantum phase space density”.

It is easy to check that $W_\psi$ is always real but in general is not positive (thus it cannot be the density of a positive measure – in coincidence with the Heisenberg principle). However, its marginals reconstruct the position and momentum space densities, as the following formulas can be easily checked:

$$\int W_\psi(x, v) dv = |\psi(x)|^2, \quad \int W_\psi(x, v) dx = |\hat{\psi}(v)|^2$$

In particular

$$\int\int W_\psi(x, v) dv dx = 1 \quad (4.1)$$

for normalized wave functions.

More generally, if $J(x, v)$ is a classical phase space observable, the scalar product

$$\langle J, W \rangle = \int J(x, v) W_\psi(x, v) dx dv$$

20
can be interpreted as the expected value of $J$ in state $\psi$. Recall that “honest” quantum mechanical observables are self-adjoint operators $O$ on $L^2(\mathbb{R}^d)$ and their expected value is given by

$$\langle \psi, O \psi \rangle = \int \overline{\psi(x)} (O \psi)(x) dx$$

For a large class of observable there is a natural relation between observables $O$ and their phase space representations (called symbols) that are functions on the phase space like $J(x,v)$. For example, if $J$ depends only on $x$ or only on $v$, then the corresponding operator is just the standard quantization, i.e.

$$\int J(x) W_\psi(x,v) dx dv = \langle \psi, J \psi \rangle$$

where $J$ is a multiplication operator on the right hand side,

$$\int J(v) W_\psi(x,v) dx dv = \langle \psi, J(-i\nabla) \psi \rangle$$

and similar relations hold for the Weyl quantization of any symbol $J(x,v)$.

We also remark that the map $\psi \rightarrow W_\psi$ is invertible, i.e. one can fully reconstruct the wave function from its Wigner transform. On the other hand, not every real function of two variables $(x,v)$ is the Wigner transform of some wavefunction.

### 4.2 Hamilton operator and the Schrödinger equation

The dynamics is generated by the Hamilton operator

$$H = -\frac{1}{2} \Delta_x + U(x)$$

acting on $\psi \in L^2(\mathbb{R}^d)$. The first term is interpreted as the kinetic energy and it is the quantization of the classical kinetic energy $\frac{1}{2} v^2$ (compare with (3.1)). The momentum operator is $p = -i\nabla$ and we set the mass equal one, so momentum and velocity coincide.

The evolution of a state is given by the Schrödinger equation

$$i \partial_t \psi_t = H \psi_t = \left( -\frac{1}{2} \Delta + U \right) \psi_t$$

with a given initial data $\psi_{t=0} = \psi_0$. Formally the solution is

$$\psi_t = e^{-iH} \psi_0$$

If $H$ is self-adjoint, then the unitary group $e^{-iH}$ can be defined by spectral theorem. Almost all Hamilton operators in mathematical physics are self-adjoint, however the self-adjointness usually requires some proof. We will not worry about this issue here, but we only mention that self-adjointness is more than the symmetry of the operator, because $H$ is typically unbounded and the domain issues become relevant.

Note the complex $i$ in the Schrödinger equation, it plays an absolutely crucial role. It is responsible for the wave-like character of quantum mechanics. The quantum evolution are
Figure 4: Evolution of slower (left) and a faster (right) wave packet
governed by phase and dispersion. Instead of giving precise definitions, look at Fig. 4: the
faster the wave oscillates, the faster it moves.

We can also justify this picture by some sketchy calculation (all these can be made rigorous
in the so-called semiclassical regime). We demonstrate that the free \((U \equiv 0)\) evolution of
\[ \psi_0(x) := e^{i v_0 x} A(x - x_0) \]
(with some fixed \(x_0, v_0\)) after time \(t\) is supported around \(x_0 + v_0 t\). Here we tacitly assume
that \(A\) is not an oscillatory function (e.g. positive) and the only oscillation is explicitly in the
phase \(e^{i v_0 x}\), mimicking a plane wave.

We compute the evolution of \(\psi_0\):

\[
\psi_t(x) = \int e^{i v(x-y)} e^{-itv^2/2} \underbrace{e^{i v_0 y} A(y - x_0)}_{=: \psi_0(y)} dy dv \\
= e^{i v_0 x_0} \int e^{i v(x-x_0)} e^{-itv^2/2} \hat{A}(v - v_0) dv \sim \int e^{i \Phi(v)} \hat{A}(v - v_0) dv
\]
with a phase factor

\[
\Phi(v) := v(x - x_0) - \frac{1}{2} tv^2
\]

We apply the stationary phase argument: the integral is concentrated around the stationary
points of the phase. Thus

\[
0 = \nabla_x \Phi = x - x_0 - tv
\]
gives \(x = x_0 + vt \approx x_0 + v_0 t\) if \(\hat{A}\) is sufficiently localized.

This argument can be made rigorous if a scale separation ansatz is used. We introduce a
small parameter \(\varepsilon\) and assume that

\[
\psi_0(x) = \varepsilon^{d/2} e^{i v_0 x} A(\varepsilon(x - x_0))
\]
i.e. the wave has a slowly varying envelope (amplitude) and fast phase (the normalization
is chosen to keep \(\|\psi_0\| = 1\) independently of \(\varepsilon\).) Such states and their generalizations of the
form

\[
\varepsilon^{d/2} e^{i S(\varepsilon x)/\varepsilon} A(\varepsilon(x - x_0))
\]
are called WKB-states. These states are characterized by a slowly varying amplitude and
wavelength profile and by a fast oscillation. Similar picture holds if the potential is nonzero,
but slowly varying, i.e. \(U(x)\) is replaced with \(U(\varepsilon x)\).
4.3 Semiclassics

The WKB states and the rescaled potential belong to the *semiclassical theory* of quantum mechanics. A quantum mechanical system is in the semiclassical regime whenever all the data (potential, magnetic field, initial profile etc.) live on a bigger scale than the quantum wavelength.

The scale separation parameter is usually explicitly included in the Schrödinger equation, for example:

\[ i\partial_t \psi_t(x) = \left[ -\frac{1}{2}\Delta_x + U(\varepsilon x) \right]\psi_t(x) \]  

This equation is written in microscopic coordinates. It becomes more familiar if we rewrite in macroscopic coordinates

\[(X, T) = (x\varepsilon, t\varepsilon)\]

where (4.3) takes the form

\[ i\varepsilon\partial_T \Psi_T(X) = \left[ -\frac{\varepsilon^2}{2}\Delta_X + U(X) \right]\Psi_T(X) \]

Traditionally \(\varepsilon\) is called \(h\) in this equation and is referred to the Planck constant, whose smallness in standard units justifies taking the limit \(\varepsilon \to 0\) (the so-called *semiclassical limit*). No matter in what coordinates and units, the essential signature of the semiclassical regime is the scale separation between the wavelength and any other lengthscales.

Wigner transform \(W_\psi(x, v)\), written originally in microscopic coordinates, can also be rescaled:

\[ W_\psi^\varepsilon(X, V) := \varepsilon^{-d}W_\psi\left(\frac{X}{\varepsilon}, V\right) \]

Note that apart from the rescaling of the \(X\) variable (\(V\) variable is unscaled!), a normalizing prefactor \(\varepsilon^{-d}\) is added to keep the integral of \(W_\psi^\varepsilon\) normalized:

\[ \int\int W_\psi^\varepsilon(X, V)dXdV = 1 \]
The following theorem is a simple prototype of the theorems we wish to derive in more complicated situations. It shows how the quantum mechanical Schrödinger evolution can be approximated by a classical equation in a certain limiting regime:

**Theorem 4.1 (Textbook).** Consider a sequence of initial states \( \psi_0^\varepsilon \) whose rescaled Wigner transform has a weak limit:

\[
\lim_{\varepsilon \to 0} W_{\psi_0^\varepsilon}(X,V) \to W_0(X,V)
\]

For example, the WKB states given in (4.2) have a weak limit

\[
W_0(X,V) = |A(X)|^2 \delta(V - \nabla S(X))
\]

Let \( \psi_t^\varepsilon \) denote the solution to the semiclassical Schrödinger equation (4.3) with initial state \( \psi_0^\varepsilon \). Then the weak limit of the rescaled Wigner transform of the solution at time \( t = T/\varepsilon \),

\[
W_T(X,V) := \lim_{\varepsilon \to 0} W_{\psi_t/\varepsilon}(X,V)
\]

satisfies the Liouville equation:

\[
(\partial_T + V \cdot \nabla_X) W_T(X,V) = \nabla U(X) \cdot \nabla_V W_T(X,V)
\]  

(4.4)

Recall that the Liouville equation is equivalent to the classical Hamilton dynamics (see the derivation of (3.4)).

In this way, the Liouville equation (4.4) mimics the semiclassical quantum evolution on large scales. Notice that weak limit is taken; this means that the small scale structure of the Wigner transform \( W_{\psi_t^\varepsilon} \) is lost. Before the weak limit, the Wigner transform of \( \psi_t^\varepsilon \) carries all information about \( \psi_t^\varepsilon \), but the weak limit means testing it against functions that live on the macroscopic scale. This phenomenon will later be essential to understand why irreversibility of the Boltzmann equation does not contradict the reversibility of Schrödinger dynamics.

5 Random Schrödinger operators

5.1 Quantum Lorenz gas = Anderson model

Now we introduce our basic model, the quantum Lorenz gas. Physically we wish to model electron transport in a disordered medium. We assume that electrons are non-interacting, thus we can focus on the evolution of a single electron. The disordered medium is modelled by a potential describing random impurities (other models are also possible, e.g. one could include random perturbations in the kinetic energy term). This can be viewed as the quantum analog of the classical Lorenz gas, where a classical particle was moving among random obstacles.

**Definition 5.1.** The single particle system described by the Schrödinger equation on \( \mathbb{R}^d \) or on \( \mathbb{Z}^d \)

\[
i \partial_t \psi_t(x) = H \psi_t(x), \quad H = -\Delta_x + \lambda V_\omega(x)
\]

is called the quantum Lorenz gas, where \( V_\omega(x) \) is a random potential at location \( x \) (here \( \omega \) is an element in the probability space, i.e. it indicates that \( V(x) \) is a random variable).
We usually view
\[ H = H_0 + \lambda V \] (5.1)
where \( H_0 \) is a deterministic Hamiltonian, typically \( H_0 = -\Delta \). We remark that a periodic background potential can be added, i.e.
\[ H = -\Delta_x + U_{\text{per}}(x) + \lambda V_{\omega}(x) \]
can be investigated with similar methods. We recall that the operator with periodic potential and no randomness
\[ H_0 = -\Delta_x + U_{\text{per}}(x) \]
can be more or less explicitly diagonalized by using the theory of Bloch waves and the transport properties are similar to those of the free Laplacian. This is called the \textit{periodic quantum Lorenz gas} and it is the quantum analogue of the Sinai billiard. However, while the Sinai billiard has a very complicated structure, its quantum version is almost trivial (the theory of Bloch waves is not a particularly deep subject and it is even taught in standard physics courses).

We also remark that for the physical content of the model, it practically unimportant whether we work on \( \mathbb{R}^d \) or on its lattice approximation \( \mathbb{Z}^d \). The reason is that we are investigating long time, large distance phenomenon; the short scale structure of the space does not matter. However, technically \( \mathbb{Z}^d \) is much harder (a bit unusual, since typically \( \mathbb{R}^d \) is harder as one has to deal with the ultraviolet regime). If one works on \( \mathbb{Z}^d \), then the Laplace operator is interpreted as the discrete Laplace operator on \( \mathbb{Z}^d \), i.e.
\[ (\Delta f)(x) := 2d f(x) - \sum_{|e|=1} f(x + e) \]
The random potential can be fairly arbitrary, the only important requirement is that it cannot have a long space correlation. For convenience, we will assume i.i.d (=independent, identically distributed) random potential for the lattice case with a standard normalization:
\[ \{V(x) : x \in \mathbb{Z}^d\} \quad \text{i.i.d} \quad \mathbb{E}V(x) = 0, \quad \mathbb{E}V^2(x) = 1 \]
The coupling constant \( \lambda \) can then be used to adjust the strength of the randomness. We can also write the random potential as
\[ V(x) = \sum_{\alpha \in \mathbb{Z}^d} v_\alpha \delta(x - \alpha) \] (5.2)
where \( \{v_\alpha : \alpha \in \mathbb{Z}^d\} \) is a collection of i.i.d. random variables and \( \delta \) is the usual lattice delta function.

For continuous models, the definition of the random potential can for example be the following:
\[ V(x) = \sum_{\alpha \in \mathbb{Z}^d} v_\alpha f(x - \alpha) \]
where \( f \) is a nice (smooth, compactly supported) single site potential profile and \( \{v_\alpha : \alpha \in \mathbb{Z}^d\} \) is a collection of i.i.d. random variables. It is also possible to let the randomness perturb the location of the obstacles instead their strength, i.e.
\[ V(x) = \sum_{\alpha \in \mathbb{Z}^d} f(x - x_\alpha(\omega)) \]
where, for example, \( x_\alpha(\omega) \) is a random point in the unit cell around \( \alpha \in \mathbb{Z}^d \). The combination of these two models is also possible and meaningful.

The lattice model with i.i.d. random potentials (5.2) is called the Anderson model. It was invented by Anderson [3] who was the first to realize that electrons move quite differently in disordered media than in free space. The main phenomenon Anderson discovered was the Anderson localization, claiming that at sufficiently strong disorder (or in \( d = 1, 2 \) at any nonzero disorder) the electron transport stops. We will explain this in more details later, here we just remark that Anderson was awarded the Nobel Prize for this work.

However, in weak disorder and \( d \geq 3 \) dimensions it is expected that electron transport does not stop. This is actually what we experience in real life; the electric wire is conducting although it does not have a perfect periodic lattice structure. However, the nature of the electric transport changes: we expect that on large scales it can be described by a diffusive equation (like the heat equation) instead of a wave-type equation (like the Schrödinger equation).

Note that the free Schrödinger equation is ballistic (due to wave coherence). This means that if we define the mean square displacement by the expectation value of the observable \( x^2 \) at state \( \psi_t \),

\[
\langle x^2 \rangle_t := \int dx |\psi_t(x)|^2 x^2 = \int dx |e^{-it\Delta} \psi_0(x)|^2 x^2,
\]

then it behaves as

\[
\langle x^2 \rangle_t \sim t^2
\]

for large \( t \gg 1 \) as it can be easily checked. In contrast, the mean square displacement for the heat equation scales as \( t \) and not as \( t^2 \) (see (2.2)).

Thus the long time transport of the free Schrödinger equation is very different from the heat equation. We nevertheless claim, that in a weakly disordered medium, the long time Schrödinger evolution can be described by a diffusion (heat) equation. Thus it is an intriguing question how the diffusive character emerges from a wave-type equation.

5.2 Known results about the Anderson model

I will not give a full account of all known results, just mention the main points. The main issue is to study the dichotomic nature of the Anderson model, namely that at low dimension \( (d = 1, 2) \) or at high disorder \( (\lambda \geq \lambda_0(d)) \) the time evolved state remains localized, while at high dimension \( (d \geq 3) \) and at low disorder it is delocalized.

There are several definitions of (de)localization, depending on taste and technical approach. We list three approaches:

i) Spectral approach. If \( H \) has pure point spectrum then the system is in the localized regime, if \( H \) has absolutely continuous spectrum then it is in the delocalized regime. (The singular continuous spectrum, if appears at all, gives rise to anomalous diffusions). It is to be noted that even in the pure point regime the spectrum is dense.

ii) Dynamical approach. One considers the mean square displacement (5.3) and the system is in the localized regime if

\[
\sup_{t \geq 0} \langle x^2 \rangle_t < \infty
\]
(other observables can also be considered).

iii) Conductor or Insulator? This is the approach closest to physics and mathematically it has not been elaborated (for a physical review, see [27]). In this approach one imposes an external voltage to the system and compute the current and the ohmic resistance. (Such approach has been used also mathematically for the quantum Hall effect.)

These criteria to distinguish between localized and delocalized regimes are not fully equivalent (especially in \( d = 1 \) dimensional systems there are many anomalous phenomena due to the singular continuous spectrum), but we do not go into more details.

The mathematically most studied approach is the spectral method. The main results are the following:

i) In \( d = 1 \) dimensions all eigenfunctions are localized for all \( \lambda \neq 0 \). (Goldsheid-Molchanov-Pastur [22])

ii) In \( d \geq 1 \) localization occurs for large \( \lambda \) or for energies near the edges of the spectrum of \( H_0 \) (see (5.1)). This was first proven by the groundbreaking work of Fröhlich-Spencer [19] (multiscale method). Later a different method was found by Aizenman-Molchanov [1] (fractional power method). The spectrum is not only pure point, but the eigenfunctions decay exponentially (strong localization).

iii) Many more localization results were obtained by variants of these methods for various models, including magnetic field, acoustic waves, localized states along sample edges etc. In the last twenty years many people have contributed to these efforts, among others Bourgain, Combes, Hislop, Klein, Klopp, Simon, Wang etc. (apologies for the incomplete list).

Common in all localization results is that the random potential dominates, i.e. in one way or another the free evolution \( H_0 \) is treated as a perturbation.

It is important to understand that the transport in free Schrödinger equation is due to the coherence of the travelling wave. Random potential destroys this coherence; the stronger the randomness is, the surer is the destruction. This picture is essential to understand why random potentials may lead to localization even at energies that belong to the regime of classical transport. For example in \( d = 1 \) any small randomness stops transport, although the classical dynamics is still ballistic at energies that are higher than the maximal potential peak.

5.3 Major open conjectures about the Anderson model

All the previous results were for the localization regime, where one of the main methods (multiscale analysis or fractional power method) is applicable. The complementary regimes remain unproven. Most notably is the following list of open questions:
i) **[Extended states conjecture]** For small $\lambda \leq \lambda_0(d)$ and in dimension $d \geq 3$ the spectrum is absolutely continuous away from the edges of $H_0$. In particular, there exists a threshold value (called *mobility edge*) near the edges of the unperturbed spectrum that separates the two spectral types.

This conjecture has been proven only in three cases:

a) Bethe lattice (infinite binary tree) that roughly corresponds to $d = \infty$ (Klein [24], recently different proofs were obtained in [2] and [18]).

b) Sufficiently decaying random potential, $EV^2_x = |x|^{-\alpha}$, for $\alpha > 1$ (Bourgain [5]). The decay is sufficiently strong such that the typical number of collisions with obstacles is finite.

c) In $d = 2$ with a constant magnetic field the spectrum cannot be everywhere pure point. (Germinet-Klein-Schenker [20])

ii) **[Two dimensional localization]** In $d = 2$ dimensions all eigenfunctions are localized for all $\lambda$. (i.e. the model in $d = 2$ dimensions behaves as in $d = 1$).

iii) **[Quantum Brownian motion conjecture]** For small $\lambda$ and $d \geq 3$, the location of the electron is governed by a heat equation in a vague sense:

\[ \partial_t |\psi_t(x)|^2 \sim \Delta_x |\psi_t(x)|^2 \quad \Rightarrow \quad \langle x^2 \rangle_t \sim t \quad t \gg 1 \]

The precise formulation of the first statement requires a scaling limit. The second statement about the diffusive mean square displacement is mathematically precise, but what really stands behind it is a diffusive equation that on large scales mimics the Schrödinger evolution.

Note that the “Quantum Brownian motion conjecture” is much stronger than the “Extended states conjecture”, since the former more precisely describes how the states extend. All of these three open problems have been outstanding for many years and we seem to be far from their solutions.

To indicate what is proven and what is open in $d \geq 3$, Fig. 6 is the expected phase diagram of the Anderson model in dimensions bigger than three. The picture shows the different spectral types (on the horizontal energy axis) at various disorder.

Our main result, explained in the next sections, is that the “Quantum Brownian motion conjecture” holds in the scaling limit up to times $t \sim \lambda^{-2-\kappa}$. More precisely, we will fix a positive number $\kappa$ and we will consider the family of random Hamiltonians

\[ H = H_\lambda = -\frac{1}{2}\Delta + \lambda V \]

parametrized by $\lambda$, we consider their long time evolution up to times $t \sim \lambda^{-2-\kappa}$ and we take $\lambda \to 0$ limit. Note that the time scale depends on the coupling parameter. This result is of course far from either of the conjectures i) and iii) above, since those conjectures require fixing $\lambda$ (maybe small but fixed) and letting $t \to \infty$. 

28
Expected spectrum at small, nonzero disorder $0 < \lambda << 1$

Figure 6: Phase diagram of the Anderson model in $d = 3$
We emphasize that the typical number of collisions is $\lambda^2 t$. The reason is that the quantum rate of collision with a potential that scales as $O(\lambda)$ is of order $\lambda^2$. This follows from simple scattering theory: if
\[ \tilde{H} = -\Delta + \lambda V_0 \]
is a Hamiltonian with a single bump potential (i.e. $V_0$ is smooth, compactly supported) and $\psi_{in}$ denotes the incoming wave, then after scattering the wave is split into two parts;
\[ e^{-it\tilde{H}}\psi_{in} = \beta e^{it\Delta}\psi_{in} + \psi_{sc}(t), \quad t \gg 1, \]
where $\psi_{sc}(t)$ is the scattered wave (typically spherical if $V_0$ is spherical) and the first term describes the transmitted wave that did not notice the obstacle (apart from its amplitude is scaled down by a factor $\beta$). Elementary calculation then shows that the scattered and transmitted waves are (almost) orthogonal and their amplitudes satisfy
\[ \|\psi_{sc}(t)\|^2 = O(\lambda^2), \quad \beta^2 = 1 - O(\lambda^2) \] (5.4)
Therefore the incoming wave scatters with a probability $O(\lambda^2)$.
Thus up to time $t$, the particle encounters $\lambda^2 t$ collisions. In our scaling
\[ n := \text{Number of collisions} \sim \lambda^2 t \sim \lambda^{-\kappa} \to \infty. \]
The asymptotically infinite number of collisions is necessary to detect the diffusion (Brownian motion, heat equation), similarly as the Brownian motion arises from an increasing number of independent “kicks” (steps in random walk, see Theorem 2.7).

6 Main result

In this section we formulate our main result precisely, i.e. that “Quantum Brownian motion conjecture” holds in the scaling limit up to times $t \sim \lambda^{-2-\kappa}$. Before going into the explanation of any proof, we will explain why it is a difficult problem.

6.1 Why is this problem difficult?
The dynamics of a single quantum particle among random scatterers (obstacles) is a multiple scattering process. The quantum wave function bumps into an obstacle along its evolution, and according to standard scattering theory, it decomposes into two parts: a wave that goes through the obstacle unnoticing it and a scattering wave. In case of radially symmetric obstacle the scattering wave is radially symmetric. However, the scattered wave later bumps into another obstacle etc, giving rise to a complicated multiple scattering picture, similar to the one on Fig. 7. Obviously the picture becomes more and more complicated as the number of collisions, $n$, increases. Finally, when we perform a measurement, we select a domain in space and we have to compute the wave function at the point. By the superposition principle of quantum mechanics, at that point all elementary scattering waves have to added up; eventually exponentially many (in $\lambda^{-1}$). They are quantum waves, i.e. complex valued functions, so they must be added together with their phases. The cancellations in this sum due to the phases
are essential even to get the right order of magnitude. Thus one way or another one must trace all elementary scattering waves with an enormous precision (the wavelength is order one in microscopic units but the waves travel at distance $\lambda^{-2}$ between two collisions!)

It is important to emphasize that this system is not semiclassical. Despite the appearance of a small parameter, what matters is that the typical wavelength remains comparable with the variational lengthscale of the potential. Thus the process remains quantum mechanical and cannot be described by any WKB method.

6.2 Scales

Since the collision rate is $O(\lambda^2)$, and the typical velocity of the particle is unscaled (order 1), the typical distance between two collisions (the so-called mean free path) is $L = O(\lambda^{-2})$ and time elapsed between consecutive collisions is also $O(\lambda^{-2})$.

6.2.1 Kinetic scale

In order to see some collisions, we have to run the dynamics at least up to times of order $\lambda^{-2}$. Within that time, the typical distance covered is also of order $\lambda^{-2}$ (always understood in microscopic coordinates). Thus the simplest scale that yields a nontrivial nontrivial result is the so-called kinetic scaling

$$t = \lambda^{-2} T, \quad x = \lambda^{-2} X, \quad n = \lambda^2 t = O(1).$$
i.e. the space and time are scaled by $\lambda^2$ and the typical number of collisions is of order 1.

One convenient way to quantify Fig. 7 is that one works in a large quadratic box whose size is at least comparable with the mean free path (otherwise the boundary would have a too big effect). Thus one can consider the box

$$\Lambda = [0, L]^d$$

and put $|\Lambda| = L^d$ obstacles in it, e.g. at every lattice site. For definiteness, we work with the Anderson model, i.e.

$$H = -\Delta + \sum_{\alpha \in \mathbb{Z}^d} v_\alpha \delta(x - \alpha)$$

Then the total number of obstacles is $L^d$. Each elementary scattering wave corresponds to a particular (ordered) sequence of random obstacles. After $n$ collisions we have $\sim |\Lambda|^n$ elementary waves to sum up in Fig. 7, thus

$$\psi_t = \sum_A \psi_A$$  \hspace{1cm} (6.1)

where $A = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ is a sequence of obstacle locations ($\alpha_i \in \mathbb{Z}^d$) and $\psi_A$ describes the elementary scattering wave that has undergone the collisions with obstacles at $\alpha_1, \alpha_2, \ldots, \alpha_n$ in this order. The precise definition of $\psi_A$ will come later.

These waves must be summed up together with their phase. It is easy to see from standard scattering theory that a spherical wave decays as $\lambda|\text{distance}|^{-(d-1)/2}$ (the prefactor $\lambda$ comes from the fact that the total amplitude of the scattering is scaled down by a factor of $\lambda$ due to the coupling constant in front of the potential, see (5.4)). Since the amplitudes multiply, and since the typical mean free path is $\lambda^{-2}$, after $n$ collisions the typical amplitude of $\psi_A$ is

$$|\psi_A| \sim \left[ \lambda \left( \frac{\lambda^2}{\lambda^2} \right)^{d-1} \right]^n = \lambda^{dn}$$

Thus if we try to sum up these waves in (6.1) with neglecting their phases, then

$$\sum_A |\psi_A| \sim |\Lambda|^n \lambda^{dn} = \left( \lambda^{-2d} \right)^n \lambda^{dn} = \lambda^{-dn} \to \infty$$

However, if we can sum up these waves assuming that their phases are independent, i.e. they can be summed up as independent random variables, then it is the variance that is additive (exactly as in the central limit theorem):

$$|\psi_t|^2 = \left| \sum_A \psi_A \right|^2 \approx \sum_A |\psi_A|^2 \sim |\Lambda|^n \lambda^{2dn} = O(1) \hspace{1cm} (6.2)$$

Thus it is essential to extract a strong independence property from the phases. Since phases are determined by the random obstacles, there is some hope that at least most of the elementary waves are roughly independent. This will be our main technical achievement, although it will be formulated in a different language.
6.2.2 Diffusive scale

Now we wish to go beyond the kinetic scale and describe a system with potentially infinitely many collisions. Thus we chose a longer time scale and we rescale space appropriately. We obtain the following **diffusive scaling**:

\[
t = \lambda^{-\kappa} \lambda^{-2} T, \quad x = \lambda^{-\kappa/2} \lambda^{-2} X, \quad n = \lambda^2 t = \lambda^{-\kappa}
\]

Notice that the time is rescaled by an additional factor \(\lambda^{-\kappa}\) compared with the kinetic scaling, and space is rescaled by the square root of this additional factor. This represents the idea that the model scales diffusively with respect to the kinetic scale. The total number of collisions is still \(n = \lambda^2 t\), and now this tends to infinity.

If we try to translate this scaling into the elementary wave picture, then first we have to choose a box that is larger than the largest space scale, i.e. \(L \geq \lambda^{-2 - \kappa/2}\). The total number of elementary waves to sum up is

\[
|\Lambda|^n = (\lambda^{-2 - \frac{\kappa}{2}})^{dn} \sim \lambda^{-\kappa}
\]

i.e. superexponentially huge. Even with the assumption that only the variances have to be summable (see (6.2)), we still get a superexponentially divergent sum:

\[
|\psi_t|^2 \approx \sum_A |\psi_A|^2 \sim |\Lambda|^n 2^{nd} = \left[\lambda^{-2-\kappa/2}\right]^{nd} \lambda^{2nd} = \lambda^{-nd\kappa/2} = \lambda^{-(\text{const})\lambda^{-n}}
\]

We will have to perform a renormalization to prevent this blow-up.

6.3 Kinetic scale: (linear) Boltzmann equation

The main result on the kinetic scale is the following theorem:

**Theorem 6.1. [Boltzmann equation from the kinetic limit]** Let the dimension be at least \(d \geq 2\). Consider the random Schrödinger evolution on \(\mathbb{R}^d\) or \(\mathbb{Z}^d\)

\[
i \partial_t \psi_t = H \psi_t, \quad H = H_\lambda = -\frac{1}{2} \Delta + \lambda V(x)
\]

where the potential is spatially uncorrelated (see more precise conditions below). Let us consider the kinetic rescaling

\[
t = \lambda^{-2} T, \quad x = \lambda^{-2} x,
\]

with setting \(\varepsilon := \lambda^2\) to be the scale separation parameter in space.

Then the weak limit of the rescaled Wigner transform of the solution \(\psi_t\) exists,

\[
\mathbf{E} W_{\psi_{T/\varepsilon}}^\varepsilon (X, V) \to F_T(X, V)
\]

and \(F\) satisfies the linear Boltzmann equation

\[
\left(\partial_T + \nabla e(V) \cdot \nabla X\right) F_T(X, V) = \int dU \sigma(U, V) \left[ F_T(X, U) - F_T(X, V) \right].
\]

Here \(e(V)\) is the dispersion relation of the free kinetic energy operator (it expresses the energy dependence on the velocity, see below).
The velocities $U, V$ are interpreted as incoming and outgoing velocities, respectively, see Fig. 8. The collision kernel always contains an onshell condition, i.e. a term $\delta(e(U) - e(V))$ guaranteeing energy conservation (see the examples below). Thus the right hand side of the linear Boltzmann equation is exactly the generator of the jump process on a fixed energy shell (in case of $e(V) = V^2$ it is on the sphere) as defined in Section 2.4.

This theorem has been proven in various related models.

- [Continuous model on $\mathbb{R}^d$] The random potential is chosen to be a homogeneous Gaussian field, this means that the random variable $V(x)$ is Gaussian and it is stationary with respect to the space translations. We assume that $\mathbb{E} V(x) = 0$, then the distribution of $V(x)$ is uniquely determined by the two-point correlation function

$$R(x - y) := \mathbb{E} V(x)V(y)$$

The dispersion relation of the free Laplacian on $\mathbb{R}^d$ is

$$e(p) := \frac{1}{2} p^2$$

and the Boltzmann collision kernel is computed from the microscopic data as

$$\sigma(U, V) = \delta(e(U) - e(V)) |\hat{R}(U - V)|$$

(6.3)

This result was first obtained by Spohn [30] for short macroscopic time, $T \leq T_0$, and with a different method later extended to arbitrary times in a joint work with Yau [13]

- [Discrete model + Non Gaussian] The method of [13] was extended by Chen in two different directions [7]. He considered the discrete model, i.e. the original Anderson model and he allowed non-Gaussian randomness as well. In the case of Laplacian on $\mathbb{Z}^d$ the dispersion relation is more complicated

$$e(V) := \sum_{i=1}^d 2(1 - \cos(V_i))$$

(where $V = (V_1, \ldots, V_n)$ are the coordinates of $V$), but the collision kernel is simpler;

$$\sigma(U, V) = \delta(e(U) - e(V))$$

(6.4)

i.e. it is simply the uniform measure on the energy surface given by the energy of the incoming velocity.
- **[Phonon model]** In this model the random potential is replaced by a heat bath of non-interacting bosonic particles that interact only with the single electron. *Formally* this model leads to an electron in a time dependent random potential, but the system is still Hamiltonian. This model is the quantum analogue of Einstein’s picture. The precise formulation of the result is a bit more complicated, since the phonon bath has to be mathematically introduced, so we will not do it here. The interested reader can consult with [10].

- **[Wave propagation in random medium]** It is well known that in a system of harmonically coupled oscillators the wave propagate ballistically. If the masses of the oscillators deviate randomly from a constant value, the wave propagation can change. The effect is similar to the disturbances in the free Schrödinger propagation; the ballistic transport is a coherence effect that can be destroyed by small perturbations. This model, following the technique of [13], has been worked out in [28].

- **[Low density models]** There is another way to reduce the effects of possible recollisions apart from introducing a small coupling constant: one can consider a low density of obstacles. The Hamiltonian is given by

\[
H = -\frac{1}{2} \Delta_x + \sum_{\alpha=1}^{M} V_0(x - x_{\alpha}) \quad \text{in a box } [-L, L]^d, \quad L \gg \varepsilon^{-1}
\]

acting on \( \mathbb{R}^d \), where \( \{x_{\alpha}\}_{\alpha=1}^{M} \) denotes a collection of random i.i.d. points, with density

\[
\rho := \frac{M}{L^d} \to 0
\]

The kinetic scaling corresponds to choosing

\[
x, t \sim \varepsilon^{-1}, \quad \varepsilon = \rho
\]

and then we let \( \varepsilon \to 0 \). In this case convergence to the linear Boltzmann equation in the spirit of Theorem 6.1 was proven in [12], [11]. The dispersion relation is \( e(p) = \frac{1}{2}p^2 \) and the collision kernel is

\[
\sigma(U, V) = |T(U, V)|^2 \delta(U^2 - V^2)
\]

where \( T(U, V) \) is the quantum scattering cross section for \(-\frac{1}{2} \Delta + V_0\).

It is amusing the compare the weak-coupling and the low-density sceneries; they indeed describe different physical models (Fig. 9). Although in both cases the result is the linear Boltzmann equation, the microscopic properties of the model influence the collision kernel. In particular, in the low density model the full single-bump scattering information is needed (while the function \( \hat{R} \) (6.3) in the weak coupling model can be viewed as the Born approximation of the full scattering cross section). The scaling is chosen such that the typical number of collisions remain finite in the scaling limit. Note that in both models the wavelength is comparable with the spatial variation of the potential; both live on the microscopic scale. Thus neither model is semiclassical.
Low density scenery

Number of obstacles: $\varepsilon^{-2}$
Density: $O(\varepsilon)$

Weak coupling scenery

Number of obstacles: $\varepsilon^{-3}$
Density: $O(1)$

Figure 9: Low density and weak coupling sceneries

We make one more interesting remark. When compared with the corresponding classical Hamiltonian, the low density model yields the (linear) Boltzmann equation both in classical and quantum mechanics (although the collision kernels are different). The weak coupling model yields Boltzmann equation when starting from quantum mechanics, however it yields a random walk on the energy shell when starting from classical mechanics [23] and [25]. For further references, see [13].

Fig. 10 depicts schematically the microscopic and macroscopic scales in Theorem 6.1. The right side of the picture is under a microscope which is “sees” scales comparable with the wavelength (Angstrom scale). Only one obstacle and only one outgoing wave are pictured. In reality the outgoing wave goes in all directions. On this scale the Schrödinger equation holds:

$$i\partial_t \psi_t(x) = \left[ -\Delta_x + \lambda V(x) \right] \psi_t(x).$$

On the left side we zoomed our “microscope” out. Only a few obstacles are pictured that are touched by a single elementary wave function. All other trajectories are also possible. The final claim is that on this scale the dynamics can be described by the Boltzmann equation

$$(\partial_T + \nabla e(V) \cdot \nabla_{\mathcal{X}}) F_T(\mathcal{X}, V) = \int dU \sigma(U,V) \left[ F_T(\mathcal{X}, U) - F_T(\mathcal{X}', V) \right]$$

Actually the obstacles (black dots) on the left picture are only fictitious obstacles: recall that the Boltzmann equation represents no physical obstacles. It represents a process, where physical obstacles are replaced by a stochastic dynamics: the fictitious particle has an exponential
clock and it randomly decides to change its trajectory i.e. to behave as if there were an obstacle (recall Section 2.4). It is crucial to make this small distinction because it is exactly the fictitious obstacles that allow one to eliminate all recollision problems, hence this guarantees purely the Markov property.

The main message of the result is that the long time \( t = T \varepsilon^{-1} \) Schrödinger evolution can modelled by a finite time \( T \) Boltzmann evolution on the macroscopic scale. Of course the detailed short scale information is lost, explaining irreversibility. The effective limiting equation is classical, but quantum features are retained in the collision kernel.

Note that this approximation is not just conceptually interesting but it also gives rise to an enormous computational simplification. Solving the Schrödinger equation directly for long time in an complicated environment is a computationally impossible task. In contrast, the Boltzmann equation needs to be solved on \( O(1) \) times. This is expressed schematically in the “commutative diagram” Fig. 11. The top horizontal line is the true (Schrödinger) evolution. The theorem says that one can “translate” the quantum initial data by Wigner transform and scaling limit into a phase space density \( f(0) = f_0(X,V) \) and one can use the much simpler (and shorter time) Boltzmann evolution to \( f_0 \) to obtain the Wigner transform (after expectation and scaling limit) of the true evolution \( \psi_t \).
6.4 Diffusive scale: Heat equation

We consider the same model with weakly coupled random potential as before. For definiteness we work on the lattice $\mathbb{Z}^d$, the Hamiltonian is given by

$$H = -\Delta + \lambda V$$

where the random potential is

$$V(x) = \sum_{\alpha \in \mathbb{Z}^d} v_\alpha \delta(x - \alpha)$$

where $v_\alpha$ are i.i.d. random variables. Let

$$\psi_t := e^{-itH} \psi_0$$

and we rescale the Wigner transform as before:

$$W_{\psi_t} \to W_{\psi_t}^\varepsilon = \varepsilon^{-d} W_{\psi_t}(X/\varepsilon, v)$$

Our main result is

**Theorem 6.2. [Quantum diffusion]** *Let the dimension be at least $d \geq 3$. Consider the diffusive scaling

$$t = \lambda^{-\kappa} \lambda^{-2} T, \quad x = \lambda^{-\kappa/2} \lambda^{-2} X, \quad \varepsilon = \lambda^{-\kappa/2 - 2},$$

For $d \geq 3$ and $\kappa < 1/5000$ we have as $\lambda \to 0$

$$\int_{\{e(v) = e\}} E W_{\psi_t}^\varepsilon(X, v) \, dv \to f_T(X, e) \quad (weakly) \quad (6.5)$$

and the limiting function satisfies the heat equation

$$\partial_T f_T(X, e) = \nabla_X \cdot D(e) \nabla_X f_T(X, e)$$*
with diffusion matrix given by

\[ D_{ij}(e) = \left\langle \nabla e(v) \otimes \nabla e(v) \right\rangle_e, \quad \left\langle f(v) \right\rangle_e = \text{Av. on } \{ v : e(v) = e \} \]

This result is proven in [17]. The related continuous model is treated in [15, 16]. A concise summary of the ideas can be found in the expository paper [14].

Note again that weak limit means that only macroscopic observables can be controlled:

\[ E(O, W) \to \langle O, f \rangle, \quad O = O(X, e) \in C^\infty \]

In particular, the theorem does not keep track of the velocity any more, since on the largest diffusive scale the particle’s velocity is not well defined (similarly as the Brownian motion has no derivative). Thus we have to integrate out the velocity directions on a fixed energy shell (the energy remains macroscopic observable). This justifies the additional velocity integration in (6.5).

The diffusion coefficient is the velocity autocorrelation matrix obtained from the Boltzmann equation

\[ D(e) = \int_0^\infty \nabla e(v(0)) \otimes \nabla e(v(t)) dt \]

similarly to Theorem (2.3). Since the Boltzmann collision kernel for the discrete case is the uniform measure on the energy surface (6.4), the Boltzmann velocity process has no memory and thus it is possible to compute the velocity autocorrelation just by averaging with respect to the uniform (equilibrium) measure:

\[ D(e) = \int_0^\infty \nabla e(v(0)) \otimes \nabla e(v(t)) dt = \left\langle \nabla e(v) \otimes \nabla e(v) \right\rangle_e, \]

Fig. 12 shows schematically the three scales we discussed. Notice that going from the Boltzmann scale to the diffusive scale is essentially the same as going from random walk to Brownian motion in Wiener’s construction (Theorems 2.3 and 2.7). However it is misleading to try to prove Theorem 6.2 by a two-step limiting argument, first using the kinetic limit (Theorem 6.1) then combining it with Wiener’s argument. There is only one limiting parameter in the problem, one cannot take first the kinetic limit, then the additional \( \lambda^{-\kappa} \to \infty \) limit modelling the long kinetic time scale. The correct procedure is the consider the Schrödinger evolution and run it up to the diffusive time scale. In this way, the intermediate Boltzmann scale can give only a hint what to expect but it cannot be used for proof.

In fact, this is not only a mathematical subtlety. Correlations that are small on the kinetic scale, and were neglected in the first limit, can become relevant on longer time scales. Theorem 6.2 shows that it is not the case; at least up to times \( t \sim \lambda^{2-\kappa} \) the correlations are not strong enough to destroy the diffusive picture coming from the long time limit of the Boltzmann equation (in particular the diffusion coefficient can be correctly computed from the Boltzmann equation). However, this should not be taken for granted, and in fact it is not expected to hold in \( d = 2 \) dimensions for exponentially long times, \( t \sim \exp(\lambda^{-1}) \). Theorem 6.1 on the kinetic limit is valid in \( d = 2 \) as well and Wiener’s argument is dimension independent. On the other hand if the diffusive evolution were true up to any time scales, then in particular the state would be delocalized, in contrast to the conjecture that the random Schrödinger operator in \( d = 2 \) dimensions is always localized.
Diffusive scale: $X, T$

Quantum diffusive limit [E–Salmhofer–Yau]

Quantum kinetic limit [Spohn, E–Yau]

Scaling limit of RW [Wiener]

HEAT EQ. → BOLTZMANN EQ. → SCHRODINGER EQ.

Figure 12: Three scales: diffusive, kinetic and atomic
7 Feynman graphs (without repetition)

Feynman graphs are extremely powerful graphical representations of perturbation expansions. They have been primarily used in many-body theories (they were invented for QED), but they are very convenient for our problem as well. They have been used by physicist to study random Schrödinger operators, a good overview from the physics point of view is [31].

We will derive the Feynman graphs we need for this talk. Recall that

\[ H = -\Delta + \lambda V, \quad V = \sum_{\alpha \in \mathbb{Z}^d} V_\alpha, \quad EV_\alpha = 0 \]

where \( V_\alpha \) is a single-bump potential around \( \alpha \in \mathbb{Z}^d \), for example

\[ V_\alpha(x) = v_\alpha \delta(x - \alpha) \]

We consider the potential as a small perturbation of \(-\Delta\) and we expand the unitary evolution by the identity

\[ \psi_t = e^{-itH} \psi_0 = e^{-it\Delta} \psi_0 - i\lambda \int_0^t e^{-i(t-s)H} Ve^{-is\Delta} \psi_0 \, ds \]

We can iterate this expansion, repeating the same identity for \( e^{-i(t-s)H} \). For example, the next step is

\[ e^{-itH} \psi_0 = e^{-it\Delta} \psi_0 - i\lambda \int_0^t e^{-i(t-s)\Delta} V e^{-is\Delta} \psi_0 \, ds - \lambda^2 \int_0^t ds_1 \int_0^{s_1} ds_2 e^{-i(t-s_1-s_2)H} V e^{-is_2\Delta} Ve^{-is_1\Delta} \psi_0 \]

etc. We see that at each step we obtain a new fully expanded term (second term), characterized by the absence of the unitary evolution of \( H \); only \( e^{-is\Delta} \) appear in these terms. There is always the one last term, that contains the full evolution, and that will be estimated trivially after sufficient number of expansions.

Recalling that each \( V \) is a big summation, we arrive at the following Duhamel formula

\[ \psi_t = \sum_{n=0}^N \sum_A \psi_A + \text{full evolution term}, \quad (7.6) \]

where the summation is over collision histories:

\[ A := (\alpha_1, \alpha_2, \ldots, \alpha_n) \]

and each elementary wave function is given by

\[ \psi_A = (-i\lambda)^n \int e^{-is_0\Delta} V_{\alpha_1} \cdots V_{\alpha_n} e^{-is_n\Delta} \psi_0 \, ds_0 ds_1 \cdots ds_n \]

The full evolution term has a similar structure, but the last unitary evolution \( e^{-is_n\Delta} \) is replaced with \( e^{-is_nH} \).

Each \( \psi_A \) can be symbolically represented as shown on Fig. 13, where lines represent the free propagators and bullets represent the potential terms (collisions). Note the times between collisions is not recorded in the graphical picture since they are integrated out.
\[ \psi_A = \alpha_1 \alpha_2 \alpha_3 \alpha_4 \] with propagator: \[ \phantom{=} = e^{-i\Delta} \]

Figure 13: Representation of a wave function with collision history \( A \)

\[ E \| \psi_t \|^2 = \sum_{\pi} \left( \sum_{\pi(1), \pi(2), \pi(3)} \right) = \sum_{\pi} \text{Val}(\pi) \]

Figure 14: Representation of the \( L^2 \)-norm as sum of Feynman graphs

We need to compute expectation values of quadratic functionals of \( \psi \) like the Wigner transform. For simplicity we will show how the \( L^2 \)-norm can be treated, the Wigner transform is very similar.

Suppose for the moment that there is no repetition in \( A \). To compute \( E \| \psi \|^2 = E \bar{\psi} \psi \), we need to compute

\[ \sum_{A,B} E \langle \psi_B \psi_A \rangle \]

This is zero unless \( B \) is a permutation of \( A \), because individual potentials have zero expectation. Higher order pairs are excluded since we assumed that \( A \) and \( B \) have no repetition. Thus the only nonzero contributions come from pairings of each element of \( A \) with an element of \( B \), in particular, \( |A| = |B| \) is forced.

In conclusion, with the tacit assumption that there are no repetitions, we can express \( E \| \psi \|^2 \) as a summation over permutations between the sets \( A \) and \( B \). Drawing the two horizontal lines that represent \( \psi_A \) and \( \psi_B \) parallel and connecting the paired bullets, we obtain Fig. 14. Each permutation \( \pi \) corresponds to a graph, called the Feynman graph of \( \pi \). We will give an explicit formula for the value of each Feynman graph, here denoted by \( \text{Val}(\pi) \).

7.1 Proof of the Boltzmann equation with Feynman graphs

The main observation is that among the \( n! \) possible permutations, only one, the identity permutation contributes to the limiting Boltzmann equation. Actually it has to be like that, since this is the only graph whose collision history can be interpreted classically, since here \( A = B \) as sets, so the two wavefunctions \( \psi_A \) and \( \psi_B \) visit the same obstacles in the same order. If there were a discrepancy, no classical collision history could be assigned to that pairing.

The Feynman graph of the identity permutation is called the ladder graph, see Fig. 15. Its value can be computed explicitly and it resembles to the \( n \)-th term in the Taylor series of the exponential function.

The ladder with altogether \( n \) pairing corresponds to a classical process with \( n \) collisions. If the collision rate is \( \sigma \lambda^2 \), then the classical probability of \( n \) collisions is given by the (unnor-
\[
\lambda_n \sim t^2 = (\sigma \lambda^2 t^n n! e^{-\text{Im} \theta \lambda^2 t}) \to n \sim \lambda^2 t
\]

Figure 15: Contribution of the ladder graph

normalized Poisson distribution formula

\[
\text{Prob (n collisions)} = \frac{(\sigma \lambda^2 t)^n}{n!}
\]

This is clearly summable for the kinetic scale \( t \sim \lambda^{-2} \), and the typical number of collisions is the value \( n \) giving the largest term, i.e. \( n \sim \lambda^2 t \). The exponential damping factor \( e^{-\text{Im}(\theta) \lambda^2 t} \) in Fig. 15 comes from renormalization (see later in Section 8.3 where the constants \( \sigma \) and \( \theta \) will also be defined).

We note that from the Duhamel formula leading to the Feynman graphs, it is not at all clear that the value of the ladder graphs is order 1. According to the Duhamel representation, they are highly oscillatory integral of the form:

\[
\text{Val(id)} \sim \lambda^{2n} \left( \int_0^t e^{is_1 p_1^2} \int_0^{s_1} e^{is_2 p_2^2} \ldots \int_0^{s_n} \right) \left( \int_0^t e^{-is_1' p_1^2} \int_0^{s_1'} e^{-is_2' p_2^2} \ldots \int_0^{s_n'} \right)
\]

If one estimates them naively, neglecting all oscillations, one obtains

\[
\text{Val(id)} \sim \frac{\lambda^{2n} t^{2n}}{(n!)^2} \quad \text{(naively)}
\]

which is completely wrong.

Assuming that the ladder can be computed precisely, one still has to deal with the remaining \( n! - 1 \) non-ladder graphs. Their explicit formulas will reveal that their values are small, but there are many of them, so it is not clear at all if they are still summable to something negligible.

It is instructive to compare the value of the non-ladder graphs with the ladder graph because that is of order 1. Fig. 16 contains schematic estimates of simple crossing graphs that can be checked by direct computation (once the explicit formula for \( \text{Val(\pi)} \) is established in the next section).

Accepting these formulas, we obtain the following majorizing estimate for the summation of the terms in the Duhamel expansion:

\[
\sum_n \sum_\pi |\text{Val(\pi)}| \leq \sum_n \frac{(\sigma \lambda^2 t)^n n!}{n!} \left( \frac{1}{A=B} + (n! - 1)(\text{small}) \right)
\]

This series clearly converges for short kinetic time \( (\lambda^2 t = T \leq T_0) \) but not for longer times, since the two \( n! \)’s cancel each other and the remaining series is geometric. This was the basic idea of Spohn’s proof [30].
The source of the $1/n!$ is actually the time-ordered multiple time integration; the total volume of the simplex

$$\{(s_1, s_2, \ldots, s_n) : 0 \leq s_1 \leq \ldots \leq s_n \leq t\}$$

is $t^n/n!$. The source of the competing $n!$ is the pairing that comes from taking the expectation, e.g. in case of Gaussian randomness this appears explicitly in the Wick theorem. The reason for this mechanism is the fact that in quantum mechanics the physical quantities are always quadratic in $\psi$, thus we have to pair two collision histories in all possible ways even if only one of them has classical meaning. This additional $n!$ is very typical in all perturbative quantum mechanical argument.

This was only the treatment of the fully expanded terms, the error term in (7.6) needs a separate estimate. Here we simply use the unitarity of the full evolution

$$\| \int_0^t e^{i(t-s)H} \int \int \int V e^{i\Delta s_1} V e^{i\Delta s_2} \ldots \, ds \| \leq \sup_s \| \psi_s \|$$

(7.7)

In this estimate we lose a factor of $t$, but we end up with a fully expanded term $\psi_s$ which can also be expanded into Feynman graphs.

How to go beyond Spohn’s argument valid only up to short macroscopic time? Notice that “most” graphs have many “crosses”, hence they are much smaller due to phase decoherence, then the naive “one-cross” estimate indicates. One can thus distinguish graphs with one, two and more crossings. From the graphs with more crossings, additional $\lambda^2$-factors can be obtained. On the other hand, the combinatorics of the graphs with a few crossings is much smaller than $n!$.

On kinetic scale the following estimate suffices (here $N$ denotes the threshold so that the Duhamel expansion is stopped after $N$ collisions):

$$\mathbb{E}\| \psi_t \|^2 \leq \sum_{n=0}^{N-1} \frac{(\sigma \lambda t)^n}{n!} \left( \frac{1}{\text{ladder}} + \frac{n \lambda^2}{\text{one cross}} + \frac{n! \lambda^4}{\text{rest}} \right)$$
Finally, we can optimize
\[ N = N(\lambda) \sim \frac{(\log \lambda)}{(\log \log \lambda)} \]
to get convergence. This estimate gives the kinetic (Boltzmann) limit for all fixed \( \tilde{T} \) \( (t = T\lambda^{-2}) \). This concludes the outline of the proof of Theorem 6.1.

For the proof of Theorem 6.2, i.e. going to diffusive time scales, one needs to classify all diagrams; it will not be enough to separate a few crosses and treat all other graphs identically.

8 Key ideas of the diffusion proof (Theorem 6.2)

8.1 Stopping rules

The Duhamel formula has the advantage that it can be stopped at various number of terms depending on the collision history. Thus the threshold \( N \) does not have to be chosen uniformly for each term in advance; looking at each term
\[ (-i\lambda)^n \sum_{s_j=t} e^{-i\alpha H} V_{\alpha_1} \cdots V_{\alpha_n} e^{-i\alpha \Delta} \psi_0 \, ds_0 ds_1 \cdots ds_n \]
we can decide if we wish to expand \( e^{-i\alpha H} \) further, or we rather use the unitary estimate (7.7). The decision may depend on the sequence \( (\alpha_1, \ldots, \alpha_n) \).

Set \( K := (\lambda^2 t)\lambda^{-\delta} \) as an absolute upper threshold for the number of collisions in an expanded term (notice that \( K \gg \) typical number of collisions). We stop the expansion if we either have reached \( K \) expanded potential term or we see a repeated term. This procedure has the advantage that repetition terms do not pile up. This stopping rule leads us to the following version of the Duhamel formula:
\[ \psi_t = \sum_{n=0}^{K-1} \psi^{nr}_{n,t} + \int_0^t e^{-i(t-s)H} \left( \psi^{nr}_{K,s} + \sum_{n=0}^{K} \psi^{rep}_{n,s} \right) \, ds \]
(8.8)
with \( \psi^{nr}_{n,t} := \sum_{A: \text{nonrep}} \psi_{A,t} \), \( \psi^{rep}_{n,s} := \sum_{|A|=n} \psi_{A,s} \) first rep at \( \alpha_n \)

Here \( \psi^{nr} \) contains collision histories with no repetition, while the repetition terms \( \psi^{rep} \) contain exactly one repetition at the last obstacle, since once such a repetition is found, the expansion is stopped. In particular, the second sum above runs over \( A = (\alpha_1, \ldots, \alpha_n) \) where \( \alpha_n = \alpha_j \) for some \( j < n \) and this is the first repetition.

The backbone of our argument is the following three theorems:
**Theorem 8.1. [Error terms are negligible]** We have the following estimate

$$E \| \psi_{err}^s \| = o(t^{-2})$$

In particular, by the unitarity estimate this trivially implies that

$$E \left\| \int_0^t e^{-i(t-s)H} \psi_{err}^s ds \right\|^2 = o(1)$$

**Theorem 8.2. [Only the ladder matters]**

$$E \| \psi_{nr}^{n,t} \| ^2 = Val(id) + o(1)$$

$$E W\psi_{nr}^{n,t} = Val_{Wig}(id) + o(1)$$

Here $Val_{Wig}$ denotes the value of the Feynman graphs that obtained by expanding the Wigner transform instead of the $L^2$-norm of $\psi_t$.

**Theorem 8.3. [Wigner transform of the main term]**

$$Val_{Wig}(id)$$

satisfies the heat equation.

We will focus on the non-repetition terms with $n < K$, i.e. on Theorem 8.2. Theorem 8.1 is a fairly long case by case study, but it essentially relies on the ideas behind Theorem 8.2. Some ideas are sketched in Section 10. Finally, Theorem 8.3 is a non-trivial but fairly explicit calculation that we sketch in Section 11.

### 8.2 Feynman diagrams in momentum-space formalism

We recall Fig. 13 for the representation of a single collision history. Now we express the value of a Feynman diagram explicitly in momentum representation. For simplicity we assume that the random potential is Gaussian, so Wick’s formula (2.5) can be applied.

$$\tilde{\psi}_{A,t}(p) = \int dp_j \int_0^t \delta(t - \sum_{j=0}^n s_j) e^{isp_1} \tilde{V}_{\alpha_1}(p - p_1)e^{isp_2} \tilde{V}_{\alpha_2}(p_1 - p_2) \ldots$$

$$= e^{2\eta t} \int_{\eta=1/t} \prod_j dp_j \int_{-\infty}^{\infty} d\alpha \sum_{\prod_j} 1 \alpha - e(p_j) + i\eta \tilde{V}_{\alpha_j}(p_{j-1} - p_j)$$

The $L^2$ norm is computed by doubling this formula. We will use the letters $p_1, \ldots, p_n$ for the momenta in $\psi_A$ and primed momenta for $\psi_B$:

$$E \| \psi_{nr}^{n,t} \| ^2 = E \left\| \sum_{A: \text{nonrep}} \psi_A \right\|^2 = \sum_{A,B} E \langle \psi_A, \psi_B \rangle = \sum_n \sum_{\pi \in \Pi_n} Val(\pi)$$

$$Val(\pi) := e^{2\eta t} \int_{-\infty}^{\infty} d\alpha d\beta e^{(\alpha - \beta)t} \prod_{j=0}^n 1 \alpha - e(p_j) - i\eta \beta - e(p_j') + i\eta$$
Figure 17: Schematic picture of the renormalization

\[ \times \lambda^{2n} \prod_j \delta \left( (p_{j-1} - p_j) - (p'_{\pi(j)-1} - p'_{\pi(j)}) \right) \prod_j dp_j dp'_j \]

Here we used Wick theorem to reduce high order expectation to product of pair expectations:

\[ E \sum_\alpha \hat{V}_\alpha(p) \hat{V}_\alpha(q) = \sum_\alpha e^{i\alpha(p-q)} = \delta(p - q) \]

and we also used that only terms with \( B = \pi(A) \) are nonzero.

This calculation gives that

\[ E \| \psi_t^{nr} \|^2 = \sum_\pi \text{Val}(\pi) \]

as it is pictured in Fig. 14. More concisely,

\[ \text{Val}(\pi) = \int dp dq d\alpha d\beta e^{i\alpha - \beta} \prod_j \frac{\lambda}{\alpha - \omega(p_j) - i\eta} \frac{\lambda}{\beta - \omega(q_j) + i\eta} \Delta_\pi(p, q) , \]

where

\[ \Delta_\pi(p, q) := \prod_j \delta \left( (p_{j-1} - p_j) - (p'_{\pi(j)-1} - p'_{\pi(j)}) \right) \]

contains the product of all delta functions. These delta functions can be obtained from the graph: they express the Kirchoff law at each pair of vertices, i.e. the signed sum of the four momenta attached to any paired vertices must be zero.

### 8.3 Self-energy Renormalization

Simple explicit calculations easily show that not all repetition terms are negligible, in fact immediate recollisions are of order one and they have to be treated differently. A collision sequence \( (\alpha_1, \alpha_2, \ldots) \) has an immediate recollision if \( \alpha_i = \alpha_{i+1} \) for some \( i \). This modification must also be implemented in the stopping rule.

Fortunately immediate repetitions appear very locally in the Feynman graph and they can be resummed according to the schematic picture on Fig. 17. The net effect is that the free propagator \( e(p) = p^2 \) needs to be changed to another propagator \( \omega(p) \) that differs from \( e(p) \) by an \( O(\lambda^2) \) amount.

One can also reorganize how the original Hamiltonian is split into main and perturbation terms:

\[ H = \underbrace{e(p) + \lambda^2 \theta(p) + \lambda V - \lambda^2 \theta(p)}_{\omega(p)} \]
The precise definition of the correction term $\theta(p)$ is determined by the following self-consistent equation:

$$\theta(p) := \int dq \frac{\omega(p) - \omega(q) + i0}{\omega(p) - \omega(q) + i\eta}$$

and its imaginary part can also be computed as

$$\sigma(p) := \eta \int dq \left| \frac{\omega(p) - \omega(q) + i\eta}{\omega(p) - \omega(q) + i\eta} \right|^2 \rightarrow \text{Im} \theta(p) \quad \eta \rightarrow 0$$

The precise formulas are not particularly interesting; the main point is that after renormalization: only the ladder has classical contribution and gives the limiting equation. If one uses the renormalized propagators, then one can assume that no immediate repetitions occur in the expansion (in practice they do occur, but they are algebraically cancelled out by the renormalization of the propagator). After this renormalization will the value of the ladder graph given in Fig. 15 be correct with the exponential damping factor.

### 8.4 Control of crossing terms

For the rest of this discussion, we will focus on Theorem 8.2

$$\mathbf{E}\|\psi_{n,r}\|^2 = \text{Val}(\text{id}) + o(1)$$

The key is a good classification of all Feynman diagrams based upon the complexity of the permutation $\pi : A \rightarrow B$. This complexity is expressed by a degree $d(\pi)$ that is defined as follows:

**Definition 8.4.** Given a permutation $\pi$ on $\{1, 2, \ldots, n\}$, an index $i$ is called **ladder index** if $|\pi(i) - \pi(i - 1)| = 1$ or $|\pi(i) - \pi(i + 1)| = 1$ (the actual definition used in our paper is a bit more involved). The **degree** of $\pi$ is defined as

$$d(\pi) = \text{degree of permutation} = \#\{\text{non-ladder indices}\}$$

Two examples are shown on Fig. 18. The top pictures show the pairing in the Feynman diagrams, the bottom pictures show the graph of the permutation

$$\pi : \{1, 2, \ldots, n\} \rightarrow \{1, 2, \ldots, n\}$$

where the set $\{1, 2, \ldots, n\}$ is naturally embedded into the reals and we conveniently joined the discrete points of the graph of $\pi$. The dotted region encircles the ladder indices: notice that a long antiladder also has many ladder indices. This indeed shows that it is not really the number of total crosses that are responsible for the smallness of the value of the Feynman graph, rather the disordered structure and that is more precisely expressed by the non-ladder indices.

The main philosophy is that most permutations have high degree (combinatorial complexity). The key estimates are the following Lemmas. The first lemma estimates the number of permutations with a given degree; the proof is elementary combinatorics:
Lemma 8.5. The number of permutations with a given degree is estimated as

$$\# \{\pi : d(\pi) = d\} \sim n^d$$

The second lemma is indeed the hard part of the proof; it claims that the amplitude of the Feynman graph decreases polynomially (in $\lambda$) as the degree increases. Thus we can gain a $\lambda$ factor per each non-ladder vertex.

Lemma 8.6. For some positive $\kappa$ we have

$$\text{Val}(\pi) \leq \lambda^{\kappa d(\pi)}$$

Combining these two Lemmas, we can easily control the series $\sum_{\pi} \text{Val}(\pi)$:

$$\sum_{\pi} \text{Val}(\pi) = \sum_{d} \sum_{\pi : d(\pi) = d} \text{Val}(\pi) = \sum_{d} n^d \lambda^{\kappa d} < \infty$$

if $n \leq \lambda^{-\kappa}$. Since $n \sim \lambda^2 t$, get convergence $t \leq \lambda^{-2 - \kappa}$.

We remark that, although for technical reasons we can prove (8.10) only for very small $\kappa$, it should be valid up to $\kappa = 2$ but not beyond. To see this, recall that the best possible estimate for a single Feynman graph is $O(\lambda^{2n})$ and their total number is $n!$. Since

$$\lambda^{2n} n! \approx (\lambda^2 n)^n \approx (\lambda^4 t)^n$$

the summation over all Feynman graphs will diverge if $t \geq \lambda^{-4}$. This means that this method is limited up to times $t \ll \lambda^{-4}$. Going beyond $t \sim \lambda^{-4}$ requires a second resummation procedure, namely the resummation of the so-called four-legged subdiagrams. Similar resummations have been done in many-body problems in euclidean (imaginary time) theories but not in real time.
8.5 An example

In this very concrete example we indicate why the cross is smaller than the ladder. We recall that the value of the Feynman graph is

$$\text{Val}(\pi) = \int d\alpha d\beta e^{i(\alpha - \beta)t} \prod_j \frac{\lambda}{\alpha - \omega(p_j) - i\eta} \frac{\lambda}{\beta - \omega(q_j) + i\eta} \Delta_\pi(p, q),$$

where the delta function comes from momentum conservation. These delta functions in $\Delta_\pi(p, q)$, may enhance singularity overlaps and may increase the value of the integral.

We consider the simplest example for a cross and compare it with the direct pairing (ladder). The notations are found on Fig. 19.

In case of the ladder, the delta functions force all paired momenta to be exactly the same (assuming that $p_0 = q_0$ since we compute the $L^2$-norm of $\psi_t$), i.e.

$$\text{Ladder} \implies p_j = q_j, \quad \forall \ j$$

In contrast, the crossing delta functions yield

$$q_0 = p_0, \quad q_1 = p_0 - p_1 + p_2, \quad q_2 = p_2$$

Now we can compute explicitly:

$$\text{Val(ladder)} = \lambda^4 \int d\alpha d\beta e^{i(\alpha - \beta)t} \int \frac{1}{\alpha - \omega(p_0) - i\eta} \frac{1}{\alpha - \omega(p_1) - i\eta} \frac{1}{\alpha - \omega(p_2) - i\eta} \times \frac{1}{\beta - \omega(p_0) + i\eta} \frac{1}{\beta - \omega(p_1) + i\eta} \frac{1}{\beta - \omega(p_2) + i\eta} dp_0 dp_1 dp_2$$

With a simple analysis one can verify that the main contribution comes from the regime where $|\alpha - \beta| \leq t^{-1}$, thus effectively all singularities overlap (with a precision $1/t$). Since

$$\int \frac{dp}{|\alpha - \omega(p) + i\eta|^2} \sim \frac{1}{|\text{Im}\omega|} \sim \lambda^{-2}$$

thus, effectively, we have

$$\text{Val(ladder)} \sim \lambda^4 (\lambda^{-2})^2 \sim O(1)$$

50
modulo logarithmic corrections. Here (8.12) has been used twice and the last pair of denominators integrated out by $d\alpha d\beta$ collecting a logarithmic term:

$$\int \frac{d\alpha}{|\alpha - \omega(p_0) - i\eta|} = O(|\log \eta|)$$

(recall that $\eta = 1/t$).

By an alternative argument one can simply estimate all (but one) $\beta$-denominators in (8.11) by $\lambda^{-2} (L^\infty$-bound), then integrate out $d\beta$ and then all $\alpha$-denom ($L^1$-bound):

$$\text{Val}(\text{Ladder}) \leq \lambda^4 (\lambda^{-2})^2 (\log \lambda)^4 \sim (\log \lambda)^4$$

The similar calculation for the crossed diagram yields:

$$\text{Val}(\text{cross}) = \lambda^4 \int d\alpha d\beta e^{i(\alpha-\beta)t} \int \frac{1}{\alpha - \omega(p_0) - i\eta} \frac{1}{\alpha - \omega(p_1) - i\eta} \frac{1}{\alpha - \omega(p_2) - i\eta} \frac{1}{\beta - \omega(p_0 + p_1 + p_2) + i\eta} \frac{1}{\beta - \omega(p_0) + i\eta} \frac{1}{\beta - \omega(p_0 - p_1 + p_2) + i\eta} \frac{1}{\beta - \omega(p_2) + i\eta} dp_0 dp_1 dp_2$$

Assuming again that the main contribution is from the regime where $\alpha \sim \beta$ (with precision $1/t$), we notice that spherical singularities of the two middle denominators overlap only at a point singularity

$$\int dp_1 \frac{1}{|\alpha - \omega(p_1) - i\eta|} \frac{1}{|\alpha - \omega(p_0 - p_1 + p_2) + i\eta|} \sim \frac{1}{|p_0 + p_2| + \lambda^2}$$

(see also Fig. 20). In three dimensions the point singularity is harmless and we thus obtain

$$\text{Val}(\text{cross}) \leq \lambda^2 \text{Val}(\text{ladder})$$

(modulo logarithms).

9 Integration of general Feynman graphs

We can reorganize the delta function as follows:

$$\Delta_\pi(p, q) = \prod_j \delta\left(p_{j-1} - p_j - (q_{\pi(j)-1} - q_{\pi(j)})\right) = \prod_i \delta\left(q_i - \sum_j M_{ij} p_j\right)$$
where we expressed all \( q \) momenta in terms of the \( p \)-momenta. The relation is linear, therefore there exists a matrix \( M \) (depending on \( \pi \)) such that

\[
q_i = \sum_j M_{ij} p_j
\]

A little linear algebra and combinatorics reveals that \( M \) is actually a **totally unimodular** matrix, which means that all its subdeterminants are 0 or \( \pm 1 \). This will mean that the Jacobians of the necessary changes of variables are always controlled.

Based upon (8.9) and (9.14), our task is to successively integrate out all \( p_j \)'s in the multiple integral of the form

\[
Q(M, \mathcal{E}) := \lambda^{2k} \int d\alpha d\beta \int dp \prod_{i=1}^{n} \left| \frac{1}{\alpha - \omega(p_i) - i\eta} \right| \prod_{j=1}^{n} \left| \frac{1}{\beta - \omega(\sum_{j=1}^{n} M_{ij} p_j) + i\eta} \right| \prod_{\mu=1}^{\nu} \left| \frac{1}{\sum_{j=1}^{n} E_{\mu j} p_j} \right| + \lambda^2
\]

Note that in addition to the matrix \( M \), an auxiliary matrix \( \mathcal{E} \) is included to keep track of the possible point singularities arising in integrals like (8.13). As one \( p_j \) is integrated out, the matrices \( M, \mathcal{E} \) need to be updated and effective algorithm is needed to keep track of these changes.

Why is it hard to compute or estimate such a quite explicit integral? The problem is that there is no way to start: every variable may appear in many denominators: apparently there is no “easy” integration variable. Decoupling inequalities like Schwarz or Hölder inequalities cannot be used since they will exactly lose the non-overlapping effect of the crossing. If one tries to decouple these integrals, one obtains the ladder and gains nothing. In contrast to a complicated crossing diagram, the ladder is easily computable because \( p_i = q_i \) means that the integrals decouple.

As a first attempt, we can trivially estimate all (but one) \( \beta \)-denominators, by their \( L^\infty \) norm

\[
\left| \frac{1}{\beta - \omega(\ldots) + i\eta} \right| \leq \frac{1}{\lambda^2 \text{Im}\omega} \sim \lambda^{-2}
\]

and then integrate out all \( p_j \) one by one by using

\[
\int \frac{dp}{|\alpha - \omega(p) + i\eta|} = O(|\log \eta|)
\]

This gives

\[
|Val(\pi)| \sim O(1)
\]

with logarithmic corrections.

This is a standard integration procedure in Feynman diagram theory where momenta are subject to momentum conservation (Kirchoff law) at each vertex. In this case one uses \( L^\infty \) bounds on certain edges in the graph to decouple the rest and use independent \( L^1 \)-bounds for the remaining integration. Typically one constructs a spanning tree in the graph, then each additional edge creates a loop. The Kirchoff delta functions amount to expressing all
tree momenta in terms of loop momenta in a trivial way. Then one uses $L^\infty$-bound on “tree”-denominators to free up all Kirchoff delta functions, integrate out the “tree”-variables freely and then use $L^1$-bound on the “loop”-denominators that are now independent.

In our case simple graph theoretical counting shows that

$$\text{Number of tree momenta} = \text{number of loop momenta} = n$$

Since each $L^\infty$-bound on “tree”-denominators costs $\lambda^{-2}$ by (9.15), the total estimate is of order

$$\lambda^{2n} \lambda^{-2n} = O(1)$$

with actually logarithmic factors. Thus we have no gain from the crossing.

In our new algorithm, we use the $L^\infty$-bound on the for a carefully selected subset of the “tree”-variables: the ones “below the valley”. The selection is shown in Fig. 21. After drawing the graph of the permutation, a structure of valleys and peaks emerges. By carefully examining the relation between this graph and the matrix $M$, one notices that if one estimates only those “tree”-denominators that lie below a valley by the trivial $L^\infty$-bound, then all the remaining denominators can be successively integrated out by selecting the integration variables $p_j$ in an appropriate order (dictated by the graph). Each integration involves no more than three denominators at one time. More precisely, this is correct if there are no ladder indices: from them we do not expect to gain anything.

After an easy bookkeeping we notice that in this way we gain roughly as many $\lambda^2$ factors as many non-valleys we have. It is easy to see that the number of valleys is less than

$$\frac{1}{2}(n - \#\{\text{ladders}\}) = \frac{d(\pi)}{2}$$

so we gain at least $\lambda^{cd(\pi)}$. This proves Lemma 8.6.

10 Feynman graphs with repetitions

The different repetition patterns are shown on Fig. 22. When one of the repetitions shows up (apart from immediate repetition that were renormalized), we stop the expansion to reduce complications (see (8.8)). Actually the stopping rule is a bit more involved, because the repetition pattern has to collect sufficient “repetitions” to render that term sufficiently small even after paying the price for the unitary estimate, but we will no go into these details. Finally, each term we compute by “bare hand”, applying a certain graph surgery to reduce the cases. Two examples of this reduction is shown on Fig. 23 and Fig. 24.

This procedure has a side effect:

$$E \prod V_{\alpha_j}(p_{j+1} - p_j)V_{\alpha_j}(q_{j+1} - q_j) = \sum_{\alpha \neq \alpha_k} \prod_j e^{i\alpha_j[p_{j+1} - p_j - (q_{j+1} - q_j)]}$$

Restriction $\alpha \neq \alpha_k$ destroys the precise delta fn. $\sum_a e^{iap} = \delta(p)$. This requires to employ the
\[ q = p_3 - p_1 - p_3 \ldots \]

Valley \( j \)

Tree mom. with sup. bound
Each ladder gives \( \lambda^2 \)
If no ladder \( \rightarrow \) the rest can be integrated: save \( \lambda^2 \) at each step.
no. of valleys < \( (n\text{-ladder})/2 \)
Total gain: \( \lambda^{(n\text{-ladder})} \)
Difficulty: point singularities \( \rightarrow \) lose an extra 2/3 in the exp.

Figure 21: Valleys and peaks determine the order of integration

![Genuine recollision](image)

Genuine recollision

![Triple collision with a gate](image)

Triple collision with a gate

![Nest](image)

Nest

![Immediate recollision or \( \theta \)](image)

Immediate recollision or \( \theta \)

![All other](image)

All other

Figure 22: Various repetition patterns

![Removal of the gates from a recollision](image)

Removal of the gates from a recollision

Figure 23: Removal of a gate
Lemma 10.1 (Connected Graph Formula). $\mathcal{A}_n$ set of partitions of $\{1, \ldots, n\}$

$$
\sum_{\alpha_l \neq \alpha_k} e^{i\beta_{\alpha l} \beta_{\alpha k}} = \sum_{\mathbf{A} \in \mathcal{A}_n} \prod_{\nu} c(|A_{\nu}|) \delta \left( \sum_{\ell \in A_{\nu}} q_\ell \right) \quad \mathbf{A} = (A_1, A_2, \ldots)
$$

The appearance of nontrivial partitioning sets means that instead of pairs we really have to consider “hyperpairs”, i.e. subsets (called lumps) of size more than 2. Then by a surgery we show how lumps can be broken into usual permutation. Fig. 25 shows two examples of this break-up.

There are many break-ups possible, but the key idea is that we choose the break-up that gives the biggest $d(\pi)$. The following combinatorial statement shows a lower estimate on the maximal degree, demonstrating that every lump can be broken up into an appropriate pairing with relatively high degree.

**Proposition 10.2.** Let $\mathbf{B} = \{B_1, B_2, \ldots\}$ be a partition of vertices.

$$
s(\mathbf{B}) := \frac{1}{2} \sum \{|B_j| : |B_j| \geq 4\}
$$

Then there exists a permutation $\pi$, compatible with $\mathbf{B}$ such that $d(\pi) \geq \frac{1}{2} s(\mathbf{B})$.

Then we reduce the estimate of hyperpairs to previous case involving only pairs.

## 11 Computation of the main term

Fourier Transform of the rescaled Wigner transform $W(X/\varepsilon, V)$

$$
\hat{W}_t(\varepsilon \xi, v) = \overline{\psi_t \left( v + \frac{\varepsilon \xi}{2} \right)} \hat{\psi_t \left( v - \frac{\varepsilon \xi}{2} \right)}
$$

55
Here $\varepsilon = \lambda^{2+\kappa/2}$ space rescaling, $t = \lambda^{-2-\kappa}T$.

Test against a macroscopic observable, compute

$$\langle \mathcal{O}, \hat{E}W_t \rangle = \langle \mathcal{O}(\xi, v), \hat{E}_{t}(\varepsilon \xi, v) \rangle = \int d\varepsilon d\xi \, \mathcal{O}(\xi, v) \, \hat{E}_{t}(\varepsilon \xi, v)$$

Continuity property

$$\langle \mathcal{O}, \hat{E}W_t \rangle - \langle \mathcal{O}, \hat{E}W_\psi \rangle \leq \|\mathcal{O}\|_\infty \|\psi - \phi\|$$

Renormalize, expand $\psi_t$, keep the ladder term. The rest is $o(1)$.

$$\langle \mathcal{O}, \hat{E}W_t \rangle \approx \sum_{k \leq K} \lambda^{2k} \int_\mathbb{R} d\alpha d\beta \, e^{it(\alpha - \beta) + 2\eta}$$

$$\times \int d\xi dv \mathcal{O}(\xi, v) R_\eta \left(\alpha, v + \frac{\varepsilon \xi}{2}\right) R_\eta \left(\beta, v - \frac{\varepsilon \xi}{2}\right)$$

$$\times \prod_{j=2}^k \left[ \int dv_j R_\eta \left(\alpha, v_j + \frac{\varepsilon \xi}{2}\right) R_\eta \left(\beta, v_j - \frac{\varepsilon \xi}{2}\right) \right]$$

$$\times \int dv_1 R_\eta \left(\alpha, v_1 + \frac{\varepsilon \xi}{2}\right) R_\eta \left(\beta, v_1 - \frac{\varepsilon \xi}{2}\right) \hat{E}_0(\varepsilon \xi, v_1),$$

with

$$R_\eta(\alpha, v) := \frac{1}{\alpha - e(v) - \lambda^2\theta(v) + i\eta},$$

We perform each $dv_j$ integral.

The key technical lemma is the following:

**Lemma 11.1.** $f(p) \in C^1$, $a := (\alpha + \beta)/2$, $\lambda^{2+4\kappa} \leq \eta \leq \lambda^{2+\kappa}$ and $|r| \leq \lambda^{2+\kappa/4}$.

$$\int \left(\alpha - e(v) - \lambda^2\theta(v) - i\eta\right) \left(\beta - e(v) + \lambda^2\theta(v) + i\eta\right) \lambda^2 f(v) \, dv$$

$$= -2\pi i \int \frac{\lambda^2 f(v) \delta(e(v) - a)}{(\alpha - \beta) + 2(\nabla e)(v) \cdot r - 2i\lambda^2 \mathcal{I}(a)} \, dv + o(\lambda^{1/4})$$

where

$$\mathcal{I}(a) := \text{Im} \int \frac{dv}{a - e(v) - i0} = \int \delta(e(v) - a) \, dv, \quad \mathcal{I}(\mu) = \text{Im} \theta(\mu)$$

**Idea of the proof.**

$$\frac{dv}{(\alpha - g(v) - i0)(\beta - g(v) + i0) + 1} \approx \frac{1}{\alpha - \beta + g(v + r) - g(v - r)}$$

$$\times \left[ \frac{1}{\beta - g(v) + i0} - \frac{1}{\alpha - g(v) - i0} \right]$$
Change variables $a = (\alpha + \beta)/2$, $b = (\alpha - \beta)/\lambda^2$, choose $\eta \ll t^{-1}$

$$\langle O, E \hat{W}_i \rangle \approx \sum_{k \leq K} \int d\xi \int da \int db \, e^{it\lambda^2b} \left( \prod_{j=1}^{k+1} \int \frac{-2\pi i F^{(j)}(\xi, v_j) \delta(e(v_j) - a)}{b + \lambda^{-2}\varepsilon(\nabla e)(v_j) \cdot \xi - 2itI(a)} \, dv_j \right)$$

$F^{(1)} = \hat{W}_0$, $F^{(k+1)} = O$, rest is $F^{(j)} \equiv 1$.

Let $d\mu_a(v)$ be the probability measure the level surface $\{e(v) = a\}$

$$\int h(v) d\mu_a(v) := \langle h \rangle_a = \frac{\pi}{I(a)} \int h(v) \delta(e(v) - a) \, dv$$

Let $H(v) := \frac{\nabla e(v)}{2\pi I(a)}$

$$2I(a) \sum_{k \leq K} \int d\xi \int d\xi \int db \, e^{i2\lambda^2 I(a)b} \left( \prod_{j=1}^{k+1} \int \frac{-i F^{(j)}(\xi, v_j)}{b + \lambda^{-2}\varepsilon H(v_j) \cdot \xi - i} \, d\mu_a(v_j) \right)$$

Expand the denominator up to second order

$$\int \frac{-i}{b + \varepsilon \lambda^{-2} H(v) \cdot \xi - i} \, d\mu_a(v) = \frac{-i}{b - i} \int \left[ 1 - \frac{\varepsilon \lambda^{-2} H(v) \cdot \xi}{b - i} + \frac{\varepsilon^2 \lambda^{-4}[H(v) \cdot \xi]^2}{(b - i)^2} + O\left( \varepsilon \lambda^{-2}[\xi]^3 \right) \right] d\mu_a(v)$$

After summation the error is $K(\varepsilon \lambda^{-2})^3 = \lambda^{-\kappa}(\lambda^{\alpha/2})^3 = o(1)$.

Linear term cancels by symmetry: $H(v) = -H(-v)$. Define

$$D(a) := 4I(a) \int d\mu_a(v) \, H(v) \otimes H(v)$$

$$\sum_{k \leq K} \int d\xi \int d\xi \int db \, 2I(a) \int \hat{W}_0(\xi, v_1)d\mu_a(v_1) \int O(\xi, v)d\mu_a(v) \times \int db \, e^{i2\lambda^2 I(a)b} \left( \frac{-i}{b - i} \right)^{k+1} \left[ 1 + \frac{\varepsilon^2 \lambda^{-4}[D(a)\xi]}{4I(a)} \frac{1}{(b - i)^2} \right]^{k-1}$$

Note the geometric series.

$$\sum_{k=0}^{\infty} \left( \frac{-i}{b - i} \right)^{k+1} \left[ 1 + \frac{B^2}{(b - i)^2} \right]^{k+1} = \frac{(-i)(b - i)^2 + B^2}{(b - i)^3 + i(b - i)^2 + iB^2}$$

Perform the $db$ integration analytically with $A := 2\lambda^2 I(a)$.

$$(-i) \int db \, e^{itAb} \frac{(b - i)^2 + B^2}{(b - i)^3 + i(b - i)^2 + iB^2} = 2\pi e^{-tA B^2} + o(1)$$
from the dominant residue $b = iB^2$. Compute
\[
tAB^2 = \varepsilon^2 \lambda^{-4-\kappa} T^2 \langle \xi, D(a)\xi \rangle = T^2 \langle \xi, D(a)\xi \rangle,
\]
To get a nontrivial limit: $\varepsilon = O(\lambda^{-2-\kappa/2})$ – Diffusive scaling.

\[
\langle \mathcal{O}, \hat{E}W \rangle \approx \int d\xi \int da T(a) \left( \int \mathcal{O}(\xi, v) d\mu_a(v) \right) \left( \langle W_0 \rangle_a \exp \left( -\frac{T}{2} \langle \xi, D(a)\xi \rangle \right) \right) =: f(T, \xi, a)
\]
where $f(T, \xi, a)$ is the sol. of the heat eq. in F-space. \hfill \square

12 Conclusions

- We rigorously proved Brownian motion from Hamiltonian quantum dynamics with a fixed time independent random scatterer environment.

- We controlled the interferences of random waves in a multiple scattering process with infinitely many collisions with random obstacles in the extended states regime. The result of the two step limit (kinetic + stochastic) is the same as the one-step limit (diffusive limit of QM).

Quantum interferences and memory effects do not become relevant up to our scale. This is expected to hold for any $\kappa$ in $d = 3$, but not expected to hold for $d = 2$ [localization].

- We classified and estimated Feynman graphs up to all orders. We gained an extra $\lambda$-power per each non-ladder vertex compared to the usual $L^\infty/L^1$ bound on propagators with tree/loop momenta.

References


58


