

# DIFFUSION OF WAVE PACKETS IN A MARKOV RANDOM POTENTIAL

YANG KANG AND JEFFREY SCHENKER

*Dedicated to Jürg Fröhlich and Tom Spencer on the occasions of their 60th birthdays.*

ABSTRACT. We consider the evolution of a tight binding wave packet propagating in a time dependent potential. If the potential evolves according to a stationary Markov process, we show that the square amplitude of the wave packet converges, after diffusive rescaling, to a solution of a heat equation.

## 1. INTRODUCTION

It is generally expected that over long times the amplitude of a wave propagating in a weakly disordered background will be well described by a diffusion, at least in dimension  $d \geq 3$ . This expectation stems from a picture of wave propagation as a multiple scattering process. Scattering off the disordered background results in random phases, and the build up of these phases over multiple scattering events leads eventually to a loss of coherence in the wave. Decoherent propagation of the wave may be understood as a classical superposition of reflections from random obstacles. As long as recurrence does not dominate the evolution, the central limit theorem suggests that the amplitude is given in the long run by a diffusion.

So far it has not been possible to turn this heuristic argument into mathematical analysis without restricting the time scale over which the wave evolution is followed as in [2, 3, 4]. One major obstacle is a lack of control over recurrence: the wave packet may return often to regions visited previously, denying us the independence needed to carry out the central limit argument. Indeed, the phenomenon of Anderson localization indicates that under appropriate conditions recurrence can dominate and cause complete localization of the wave packet. (It is worth noting that, since random walks are highly recurrent in dimensions  $d = 1$  and 2, the above heuristic analysis does not support diffusion in  $d = 1$  or 2, dimensions in which localization is proved ( $d = 1$ ) and expected ( $d = 2$ ) to dominate at any disorder strength.)

A natural way to avoid recurrence difficulties is to bring a time dependence into the disordered background — we suppose that the environment evolves as the packet propagates. Here we consider a stochastic environment evolving independently of the wave packet. A natural assumption in this context is that the background changes in time according to a stationary Markov process. Such evolution equations have been proposed as an effective model for the propagation of wave packets in optical fibers [6].

---

*Date:* August 26, 2008; revised January 22, 2009.

We consider here the simplest example of such a wave equation, namely the tight binding Markov-Schrödinger equation

$$(1.1) \quad \begin{cases} i\partial_t \psi_t(x) = T\psi_t(x) + \lambda v_x(\omega(t))\psi_t(x), \\ \psi_0 \in \ell^2(\mathbb{Z}^d). \end{cases}$$

where  $T$  is a translation invariant hopping operator on  $\ell^2(\mathbb{Z}^d)$ ,  $\omega(t)$  is a Markov process on a measure space  $\Omega$  and  $v_x : \Omega \rightarrow \mathbb{R}$  are measurable functions on  $\Omega$ . An elementary, but important, observation is that, so long as the time dependent generator  $H_{\omega(t)} = T + \lambda v_x(\omega(t))$  is uniformly bounded in time, the non-autonomous problem (1.1) has a unique solution  $\psi_t$  for any initial condition  $\psi_0 \in \ell^2(\mathbb{Z}^d)$ , given for instance by the norm convergent series

$$(1.2) \quad \psi_t = \psi_0 + \sum_{n=1}^{\infty} (-i)^n \int_{\Gamma_n(t)} H_{\omega(r_1)} \cdots H_{\omega(r_n)} \psi_0 dr_1 \cdots dr_n,$$

with  $\Gamma_n(t) = \{(r_1, \dots, r_n) : 0 \leq r_n \leq \dots \leq r_1 \leq t\}$ . The evolution is easily seen to be unitary,  $\|\psi_t\| = \|\psi_0\|$ . Thus a sufficient condition for solutions to (1.1) to exist is that  $\|T\| < \infty$  and  $\sup_x \sup_{\omega} |v_x(\omega)| < \infty$ .

We examine diffusion of the wave packet by considering the mean square amplitude  $\mathbb{E}(|\psi_t(x)|^2)$ , where  $\mathbb{E}(\cdot)$  denotes averaging with respect to the random paths of the Markov process and also an initial distribution for  $\omega(0)$ . Diffusion is characterized by changes in position that scale as the square root of the elapsed time. Thus it is natural to look at the mean square amplitude in the scaling  $t \mapsto \tau t$ ,  $x \mapsto \sqrt{\tau}x$  for a large parameter  $\tau$ . Since  $x$  is a discrete variable, to accomplish this rescaling we need to convolve  $\mathbb{E}(|\psi_t(x)|^2)$  with a function on  $\mathbb{R}^d$ . That is, we look at

$$(1.3) \quad A_t(x) = \sum_{\xi \in \mathbb{Z}^d} u(x - \xi) \mathbb{E}(|\psi_t(\xi)|^2)$$

with  $u$  a fixed ‘‘bump function’’ centered at 0. Let us suppose  $u \geq 0$  and  $\int u dx = 1$ , so that  $A_t(x) \geq 0$  and  $\int A_t(x) dx = \|\psi_0\|^2$ . We interpret diffusion for the mean square amplitude as weak convergence of  $A_t(x)$  under diffusive scaling to a solution of a heat equation. That is, for suitable test functions  $\phi$ ,

$$(1.4) \quad \int_{\mathbb{R}^d} \phi(x) \tau^{\frac{d}{2}} A_{\tau t}(\sqrt{\tau}x) dx \xrightarrow{\tau \rightarrow \infty} \|\psi_0\|^2 \int_{\mathbb{R}^d} \phi(x) \frac{1}{(\pi D t)^{\frac{d}{2}}} e^{-\frac{|x|^2}{Dt}} dx$$

with  $D > 0$ . A sufficient condition for (1.4), which requires no choice of a bump function, is obtained by a Fourier transform:

$$(1.5) \quad \sum_{\xi \in \mathbb{Z}^d} e^{-i\frac{1}{\sqrt{\tau}}\mathbf{k} \cdot \xi} \mathbb{E}(|\psi_{\tau t}(\xi)|^2) \xrightarrow{\tau \rightarrow \infty} \|\psi_0\|^2 e^{-tD|\mathbf{k}|^2}, \quad \forall \mathbf{k} \in \mathbb{R}^d.$$

Following is a brief history of related studies. Ovchinnikov and Erikman obtained diffusion for a Gaussian Markov (‘‘white noise’’) potential [7]. Pillet obtained results on transience of the wave in related models and derived a Feynman-Kac representation which we use below [8]. The evolution (1.1) was considered by Tchermentchanev [9, 10], who used Pillet’s

Feynman-Kac formula to show that position moments such as  $\sum_x |x|^p \mathbb{E}(|\psi_t(x)|^2)$  exhibit diffusive scaling, up to logarithmic corrections. More precisely, he obtained upper and lower bounds of the form

$$(1.6) \quad t^{\frac{p}{2}} \frac{1}{(\ln t)^{\nu_-}} \lesssim \sum_x |x|^p \mathbb{E}(|\psi_t(x)|^2) \lesssim t^{\frac{p}{2}} (\ln t)^{\nu_+}, \quad t \rightarrow \infty.$$

In the present paper we obtain *diffusion* (1.5) and show that (1.6) holds for  $p = 2$  with  $\nu_+ = \nu_- = 0$ . While completing this manuscript, we learned of recent work of De Roeck, Fröhlich and Pizzo on diffusion for a quantum particle weakly coupled to an array of independent free quantum fields [11].

In the next section we state technical conditions which allow us to derive (1.5). These conditions are quite general and cover a large number of models of the form (1.1). However, it may be useful to have at least one example in mind. So, to guide the reader, we close this introduction with a simple example of a potential for which we can derive diffusion. We call this the “flip process.” The state space  $\Omega$  of the Markov potential is just  $\{-1, 1\}^{\mathbb{Z}^d}$ , and  $v_x(\omega) =$  evaluation of the  $x^{\text{th}}$  coordinate. So at any time  $t$ , the potential  $v_x(\omega(t))$  takes only the values  $\pm 1$ . Now suppose the process  $\omega(t)$  is obtained by allowing each coordinate  $v_x(\omega)$  to flip sign at the times  $t_1(x) \leq t_2(x) \leq \cdot$  of a Poisson process, with independent, identical Poisson processes at each site  $x$ . For this potential, our result implies diffusion (1.5) of the wave amplitude.

## 2. STATEMENT OF THE MAIN RESULT: DIFFUSION OF THE AMPLITUDE

**2.1. Assumptions.** We make the following assumptions:

- (1) (*Existence of the Markov process and invariant measure*): We are given a topological space  $\Omega$ , a Borel probability measure  $\mu$ , and a collection  $\{\mathbb{P}_\alpha : \alpha \in \Omega\}$  of probability measures on the path space  $\mathcal{P} = \Omega^{[0, \infty)}$ , taken with the  $\sigma$ -algebra generated by Borel-cylinder sets, such that
  - (a) (*Paths are right continuous and start at  $\alpha$* ): For each  $\alpha \in \Omega$ , with  $\mathbb{P}_\alpha$  probability one, every path  $\omega(\cdot)$  is right continuous and satisfies  $\omega(0) = \alpha$ .
  - (b) (*The Markov property holds*): For any measurable  $\mathcal{A} \subset \mathcal{P}$  we have

$$\int_{\mathcal{P}} \mathbb{P}_{\omega(t)}(\mathcal{A}) d\mathbb{P}_\alpha(\omega(\cdot)) = \mathbb{P}_\alpha(\mathcal{S}_t^{-1}(\mathcal{A}))$$

where  $\mathcal{S}_t$  is the backward shift on  $\mathcal{P}$ ,  $\mathcal{S}_t \omega(\cdot) = \omega(\cdot + t)$ , so

$$\mathcal{S}_t^{-1}(\mathcal{A}) = \{\omega(\cdot) : \omega(\cdot + t) \in \mathcal{A}\}.$$

- (c) (*Invariance of  $\mu$* ): For any measurable  $A \subset \Omega$ ,

$$\int_{\Omega} \mathbb{P}_\alpha(\omega(t) \in A) d\mu(\alpha) = \mu(A)$$

Let  $\mathbb{E}_\alpha(\cdot)$  denote expectation with respect to  $\mathbb{P}_\alpha$ ,

$$(2.1) \quad \mathbb{E}_\alpha(F(\omega(\cdot))) = \int_{\mathcal{P}} F(\omega(\cdot)) d\mathbb{P}_\alpha(\omega(\cdot)),$$

and similarly

$$(2.2) \quad \mathbb{E}(\cdot) = \int_{\Omega} \mathbb{E}_\alpha(\cdot) d\mu(\alpha),$$

which is expectation with respect to the probability measure  $\mathbb{P}(\mathcal{A}) = \int_{\Omega} \mathbb{P}_\alpha(\mathcal{A}) d\mu(\alpha)$  on  $\mathcal{P}$ . By the invariance of  $\mu$  under the Markov process we have

$$(2.3) \quad \mathbb{E}(f(\omega(t))) = \int_{\Omega} f(\alpha) d\mu(\alpha)$$

for any  $t$  and any  $f \in L^1(\Omega)$ .

The Markov property, invariance of  $\mu$ , and right continuity of paths show that

$$(2.4) \quad S_t f(\alpha) = \mathbb{E}_\alpha(f(\omega(t)))$$

defines a strongly continuous contraction semi-group on  $L^2(\Omega)$  (also on  $L^p(\Omega)$  for any  $1 \leq p \leq \infty$ ). Indeed, the Markov property clearly shows this is a semi-group, and from the definition we have

$$(2.5) \quad \langle f, S_t g \rangle_{L^2(\Omega)} = \mathbb{E}(f(\omega(0))g(\omega(t))),$$

so by Cauchy Schwartz and (2.3)

$$(2.6) \quad |\langle f, S_t g \rangle_{L^2(\Omega)}| \leq \|f\|_{L^2(\Omega)} \|g\|_{L^2(\Omega)}$$

from which it follows that  $\|S_t\|_{L^2 \rightarrow L^2} \leq 1$ . The right continuity of the paths under  $\mathbb{P}_\alpha$  now shows that  $S_t$  is strongly continuous, since any  $f \in L^2(\Omega)$  may be approximated by bounded continuous functions and for bounded continuous  $f : \Omega \rightarrow \mathbb{R}$  we have

$$(2.7) \quad \|S_t f - f\|_{L^2(\Omega)}^2 = \mathbb{E}(|f(\omega(t)) - f(\omega(0))|^2) \rightarrow 0$$

by dominated convergence.

The adjoint  $S_t^\dagger$  of  $S_t$  is also a strongly continuous contraction semi-group, given formally by

$$(2.8) \quad S_t^\dagger f(\alpha) = \mathbb{E}(f(\omega(0)) | \omega(t) = \alpha),$$

where the r.h.s. is a conditional expectation. Of particular importance to us is the generator  $B$  of  $S_t^\dagger$  defined by

$$(2.9) \quad B\psi = \lim_{t \rightarrow 0^+} \frac{1}{t} (\psi - S_t^\dagger \psi)$$

on the domain  $\mathcal{D}(B)$  consisting of  $\psi$  such that the limit on the r.h.s. converges in  $L^2$  norm. The generator  $B$  is *maximally dissipative*, meaning  $\text{Re}\langle \psi, B\psi \rangle \geq 0$  for  $\psi \in \mathcal{D}(B)$  and no extension of  $B$  has this property. It follows that the spectrum of  $B$  is contained in the closed right half plane  $\{z : \text{Re } z \geq 0\}$ . The adjoint  $B^\dagger$  of  $B$  is the generator for  $S_t$ .

Note that both  $S_t$  and  $S_t^\dagger$  satisfy

$$(2.10) \quad S_t 1 = S_t^\dagger 1 = 1,$$

where  $1$  denotes the function equal to one everywhere on  $\Omega$ . It follows that  $1 \in \mathcal{D}(B)$  and  $1 \in \mathcal{D}(B^\dagger)$  and that

$$(2.11) \quad B1 = B^\dagger 1 = 0.$$

The orthogonal complement of  $1$  is the space of mean zero functions,

$$(2.12) \quad L_0^2(\Omega) = \left\{ f \in L^2(\Omega) : \int_{\Omega} f(\alpha) d\mu(\alpha) = 0 \right\}.$$

From (2.11) it follows that  $L_0^2(\Omega)$  is an invariant subspace for  $B$  and  $B^\dagger$ . We require that  $B$  is *strictly dissipative* on this space:

- (2) (*Gap condition and sectoriality of the generator*): There is  $T > 0$  such that if  $f \in \mathcal{D}(B)$  and  $\int_{\Omega} f(\alpha) d\mu(\alpha) = 0$ , then

$$\operatorname{Re}\langle f, Bf \rangle_{L^2(\Omega)} \geq \frac{1}{T} \int_{\Omega} |f(\alpha)|^2 d\mu(\alpha).$$

In addition, we require that  $B$  is sectorial, namely there is  $\gamma < \infty$  such that

$$|\operatorname{Im}\langle f, Bf \rangle_{L^2(\Omega)}| \leq \gamma \operatorname{Re}\langle f, Bf \rangle_{L^2(\Omega)}$$

for all  $f \in \mathcal{D}(B)$ .

One consequence of the sectorial condition on the generator is that a precise meaning can be given to the formal relation  $S_t^\dagger = e^{-tB}$  using the Riesz functional calculus — see [1, Chapter II, Section 4].

Finally, we require translation invariance for the hopping operator  $T$ , the Markov process, and the potential  $v_x(\alpha)$ :

- (3) (*Translation invariance of the hopping terms*):  $T$  is a translation invariant hopping operator on  $\ell^2(\mathbb{Z}^d)$ ,

$$T\psi(x) = \sum_y h(x-y)\psi(y),$$

with  $h$  a function such that

- (a) (*Self adjointness of  $T$* ) For every  $x$ ,  $h(-x) = h(x)^*$ .
- (b) (*Non-degeneracy of  $T$* ) For each non-zero vector  $\mathbf{k} \in \mathbb{R}^d$ , there is some  $x \in \mathbb{Z}^d$  such that  $h(x) \neq 0$  and  $\mathbf{k} \cdot x \neq 0$ .
- (c) (*Smoothness of the symbol*)  $\sum_x |x|^2 |h(x)| < \infty$ .

It follows from (c) that  $\widehat{h}(\mathbf{k}) = \sum_x e^{-i\mathbf{k} \cdot x} h(x)$  is a  $C^2$  function on the torus  $\mathbb{T}^d$ . In particular,  $T$  is a bounded operator with  $\|T\|_{\ell^2(\mathbb{Z}^d) \rightarrow \ell^2(\mathbb{Z}^d)} = \max_{\mathbf{k}} |\widehat{h}(\mathbf{k})|$ .

- (4) (*Translation invariance of the process and invariant measure*): There are  $\mu$ -measure preserving maps  $\sigma_x : \Omega \rightarrow \Omega$ ,  $x \in \mathbb{Z}^d$ , such that

$$\sigma_x \circ \sigma_y = \sigma_{x+y},$$

and

$$\mathbb{P}_{\sigma_x(\alpha)}(\mathcal{T}_x(\mathcal{A})) = \mathbb{P}_\alpha(\mathcal{A}),$$

where  $\mathcal{T}_x : \mathcal{P} \rightarrow \mathcal{P}$  is the map  $\mathcal{T}_x(\omega)(\cdot) = \sigma_x(\omega(\cdot))$ .

- (5) (*Translation covariance, boundedness and non-degeneracy of the potential*): The functions  $v_x : \Omega \rightarrow \mathbb{R}$  are bounded, translation covariant

$$v_x = v_0 \circ \sigma_x,$$

mean zero

$$\int_{\Omega} v_x(\alpha) d\mu(\alpha) = 0,$$

and there is  $\chi > 0$  such that for all  $x, y \in \mathbb{Z}^d$ ,  $x \neq y$ ,

$$(2.13) \quad \left\| B^{-1}(v_x - v_y) \right\|_{L^2(\Omega)} \geq \chi.$$

Since the Markov process is translation invariant,  $B$  commutes with the translations  $T_x f(\alpha) = f(\sigma_x \alpha)$  of  $L^2(\Omega)$ . Thus (2.13) is equivalent to

$$(2.14) \quad \left\| B^{-1}(v_x - v_0) \right\|_{L^2(\Omega)} \geq \chi.$$

for all  $x \in \mathbb{Z}^d$ ,  $x \neq 0$ .

The condition  $\int v_x d\mu = 0$  can always be obtained by putting the mean of  $v_x$  into the diagonal part of the hopping term. Likewise, by absorbing the normalization into the disorder strength  $\lambda$ , we may assume that  $\|v_x\|_{L^2(\Omega)} = 1$ .

A very general class of models, which includes the flip model described above, is obtained by taking  $\Omega = S^{\mathbb{Z}^d}$  for some set  $S \subset \mathbb{R}$  and supposing that *each coordinate*  $\omega(x)$  of  $\omega \in \Omega$  undergoes an *independent Markov process*, with the processes at different sites *identically distributed*. We then set  $v_x(\omega) = \omega(x)$ . In this case, the generator  $B$  is the sum of the individual generators for the processes at each site — more precisely the Friedrichs extension of that sum defined on the domain of functions depending on only finitely many coordinates. The above conditions are easily translated into conditions on the individual generator of the Markov process on  $S$  for each coordinate  $\omega(x)$ . For these models, the condition (2.13) is *trivial* since, by the independence of different coordinates and translation invariance, we have

$$(2.15) \quad \left\| B^{-1}(v_x - v_y) \right\|_{L^2(\Omega)}^2 = \left\| B^{-1}v_x \right\|_{L^2(\Omega)}^2 + \left\| B^{-1}v_y \right\|_{L^2(\Omega)}^2 = 2 \left\| B^{-1}v_0 \right\|_{L^2(\Omega)}^2, \quad x \neq y.$$

**2.2. Main result.** The wave function  $\psi$  satisfies a *linear* equation, but the square amplitude  $|\psi|^2$  is *quadratic* in  $\psi$ . To obtain a linear equation for the evolution of  $|\psi|^2$ , we consider the density matrix

$$(2.16) \quad \rho_t(x, y) = \psi_t(x)\psi_t(y)^*,$$

which satisfies the evolution equation

$$(2.17) \quad \partial_t \rho_t(x, y) = -i \sum_{\zeta} h(\zeta) [\rho_t(x - \zeta, y) - \rho_t(x, y + \zeta)] - i\lambda (v_x(\omega(t)) - v_y(\omega(t))) \rho_t(x, y).$$

Note that  $|\psi_t(x)|^2 = \rho_t(x, x)$ .

More generally, we may consider the evolution equation (2.17) to be the basic dynamical problem, with an arbitrary initial condition  $\rho_0(x, y)$ . The natural setting is for  $\rho_0$  to be a *density matrix*, namely

$$(2.18) \quad \rho_0 \in \mathcal{DM} := \left\{ \rho : \mathbb{Z}^d \times \mathbb{Z}^d \rightarrow \mathbb{C} : \rho \text{ is the kernel of a non-negative definite, trace class operator on } \ell^2(\mathbb{Z}^d) \right\}.$$

By virtue of the unitarity of the evolution (1.1), the space  $\mathcal{DM}$  is preserved by the evolution (2.17), as is the trace  $\text{tr } \rho_0 = \sum_x \rho_0(x, x)$ .

Under the assumptions outlined in the previous subsection, we have the following

**Theorem 1.** *Any solution to (2.17) with initial condition  $\rho_0 \in \mathcal{DM}$  satisfies*

$$(2.19) \quad \lim_{\tau \rightarrow \infty} \sum_x e^{-i \frac{\mathbf{k}}{\sqrt{\tau}} \cdot x} \mathbb{E}(\rho_{\tau t}(x, x)) = [\text{tr } \rho_0] e^{-t \sum_{i,j} D_{i,j}(\lambda) \mathbf{k}_i \mathbf{k}_j},$$

with  $D_{i,j} = D_{i,j}(\lambda)$  a positive definite matrix. Near  $\lambda = 0$ ,  $D_{i,j}(\lambda)$  has an asymptotic expansion

$$(2.20) \quad D_{i,j}(\lambda) = \frac{1}{\lambda^2} (D_{i,j}^0 + O(\lambda)).$$

If furthermore  $\sum_x |x|^2 \rho_0(x, x) < \infty$ , then

$$(2.21) \quad \lim_{t \rightarrow \infty} \frac{1}{t} \sum_x |x|^2 \mathbb{E}(\rho_t(x, x)) = [\text{tr } \rho_0] \sum_i D_{i,i}.$$

*Remark.* As the proof will show, (2.19) holds also for an initial condition  $\rho_0$  which is the kernel of a non-positive definite trace class operator.

### 3. AUGMENTED SPACE ANALYSIS

**3.1. Augmented space and Pillet's Feynman-Kac formula.** The starting point of our analysis is a ‘‘Feynman-Kac’’ formula due to Pillet [8] which expresses  $\mathbb{E}(\rho_t(x, x))$  as a matrix element of a contraction semigroup on an augmented Hilbert space,

$$(3.1) \quad \mathcal{H} := L^2(\mathbb{Z}^d \times \mathbb{Z}^d \times \Omega),$$

where  $\mathbb{Z}^d \times \mathbb{Z}^d \times \Omega$  is taken with the measure  $M$

$$(3.2) \quad \int f(x, y, \omega) dM(x, y, \omega) = \sum_{x,y} \int_{\Omega} f(x, y, \omega) d\mu(\omega).$$

We think of a vector  $\Psi \in \mathcal{H}$  as a ‘‘random density matrix,’’ at least if  $\Psi(\cdot, \cdot, \omega)$  is the kernel of a non-negative definite trace class operator for  $\mu$  almost every  $\omega$ . We also think of  $\mathcal{H}$  as the tensor products  $\ell^2(\mathbb{Z}^d) \otimes \ell^2(\mathbb{Z}^d) \otimes L^2(\Omega)$  or  $\ell^2(\mathbb{Z}^d \times \mathbb{Z}^d) \otimes L^2(\Omega)$ , using the notations

$$(3.3) \quad [\psi \otimes \phi \otimes f](x, y, \omega) = \psi(x) \phi(y) f(\omega), \quad \psi, \phi \in \ell^2(\mathbb{Z}^d), \quad f \in L^2(\Omega),$$

and

$$(3.4) \quad [\rho \otimes f](x, y, \omega) = \rho(x, y)f(\omega), \quad \rho \in \ell^2(\mathbb{Z}^d \times \mathbb{Z}^d), \quad f \in L^2(\Omega).$$

The Feynman-Kac-Pillet formula basic to our work is

$$(3.5) \quad \mathbb{E}(\rho_t(x, y)) = \langle \delta_x \otimes \delta_y \otimes 1, e^{-t(iK+i\lambda V+B)} \rho_0 \otimes 1 \rangle_{\mathcal{H}},$$

where

$$(3.6) \quad K\Psi(x, y, \omega) = \sum_{\zeta} h(\zeta) [\Psi(x - \zeta, y, \omega) - \Psi(x, y + \zeta, \omega)],$$

$$(3.7) \quad V\Psi(x, y, \omega) = (v_x(\omega) - v_y(\omega)) \Psi(x, y, \omega),$$

and the Markov generator  $B$  acts on  $\mathcal{H}$  as a multiplication operator with respect to the first two coordinates, that is

$$(3.8) \quad B[\rho \otimes f] = \rho \otimes (Bf), \quad \rho \in \ell^2(\mathbb{Z}^d \times \mathbb{Z}^d), \quad f \in L^2(\Omega).$$

In particular, we have

$$(3.9) \quad \mathbb{E}(\rho_t(x, x)) = \langle \delta_x \otimes \delta_x \otimes 1, e^{-tL} \rho_0 \otimes 1 \rangle_{\mathcal{H}},$$

where  $L = iK + i\lambda V + B$ . This equation relates the mean square amplitude of the time dependent dynamics (2.17) to spectral properties of the non-self adjoint operator  $L$ .

**3.2. Fourier Transform.** To perform a spectral analysis of  $L$ , it is useful to note that  $L$  commutes with a group of translations on  $\mathcal{H}$  — a fact which encodes the distributional invariance of (1.1) under translations. Specifically, if we let  $S_{\xi}$  denote a simultaneous shift of position and disorder,

$$S_{\xi}\Psi(x, y, \omega) = \Psi(x - \xi, y - \xi, \sigma_{\xi}\omega),$$

then we have

**Proposition 2.**

$$S_{\xi}K = KS_{\xi}, \quad S_{\xi}V = VS_{\xi}, \quad \text{and} \quad S_{\xi}B = BS_{\xi}.$$

*Proof.* The first two identities follow directly from the definitions of  $K$  and  $V$ ; the last follows from the assumed translation invariance of the measure  $\mu$ .  $\square$

As  $K$ ,  $V$  and  $B$  commute with a representation of the additive group  $\mathbb{Z}^d$ , we may simultaneously partially diagonalize them by a Fourier transform. In the present context a useful transformation is the following unitary map:

$$(3.10) \quad \widehat{\Psi}(x, \omega, \mathbf{k}) = \sum_{\xi} e^{-i\mathbf{k}\cdot\xi} S_{\xi}\Psi(x, 0, \omega) = \sum_{\xi} e^{-i\mathbf{k}\cdot\xi} \Psi(x - \xi, -\xi, \sigma_{\xi}\omega),$$

$$\widehat{\bullet} : L^2(\mathbb{Z}^{2d} \times \Omega) \rightarrow L^2(\mathbb{Z}^d \times \Omega \times \mathbb{T}_d)$$



where  $\mathbb{T}_d = [0, 2\pi)^d$  is the  $d$  torus. One may easily compute that

$$(3.11) \quad \widehat{K}\widehat{\Psi}(x, \omega, \mathbf{k}) := \widehat{K}\widehat{\Psi}(x, \omega, \mathbf{k}) = \sum_{\zeta} h(\zeta) \left[ \widehat{\Psi}(x - \zeta, \omega, \mathbf{k}) - e^{-i\mathbf{k}\cdot\zeta} \widehat{\Psi}(x - \zeta, \sigma_{\zeta}\omega, \mathbf{k}) \right],$$

$$(3.12) \quad \widehat{V}\widehat{\Psi}(x, \omega, \mathbf{k}) := \widehat{V}\widehat{\Psi}(x, \omega, \mathbf{k}) = (v_x(\omega) - v_0(\omega))\widehat{\Psi}(x, \omega, \mathbf{k}),$$

and

$$(3.13) \quad \widehat{B}\widehat{\Psi}(x, \omega, \mathbf{k}) = B\widehat{\Psi}(x, \omega, \mathbf{k}),$$

where  $B$  is understood to act as a multiplication operator with respect to  $x \in \mathbb{Z}^d$  and  $\mathbf{k} \in \mathbb{T}^d$ .

The operators  $\widehat{K}$ ,  $\widehat{V}$  and  $B$  act fiberwise over the torus  $\mathbb{T}^d$  — i.e., they act as multiplication operators with respect to the coordinate  $\mathbf{k}$ . Thus, eq. (3.5) may be transformed into

$$(3.14) \quad \mathbb{E}(\rho_t(x, y)) = \int_{\mathbb{T}^d} e^{-i\mathbf{k}\cdot y} \langle \delta_{x-y} \otimes 1, e^{-t\widehat{L}_{\mathbf{k}}} \widehat{\rho}_{0;\mathbf{k}} \otimes 1 \rangle_{L^2(\mathbb{Z}^d \times \Omega)} d\ell(\mathbf{k}),$$

where  $\ell$  denotes normalized Lebesgue measure on the torus  $\mathbb{T}^d$ ,

$$(3.15) \quad \widehat{\rho}_{0;\mathbf{k}}(x) = \sum_y e^{-i\mathbf{k}\cdot y} \rho_0(x - y, -y),$$

and

$$(3.16) \quad \widehat{L}_{\mathbf{k}} := i\widehat{K}_{\mathbf{k}} + i\lambda\widehat{V} + B$$

with  $\widehat{V}\phi(x, \omega) = (v_x(\omega) - v_0(\omega))\phi(x, \omega)$  and

$$(3.17) \quad \widehat{K}_{\mathbf{k}}\phi(x, \omega) = \sum_{\zeta} h(\zeta) \left[ \phi(x - \zeta, \omega) - e^{-i\mathbf{k}\cdot\zeta} \phi(x - \zeta, \sigma_{\zeta}\omega) \right].$$

In particular,

$$(3.18) \quad \sum_x e^{-i\mathbf{k}\cdot x} \mathbb{E}(\rho_t(x, x)) = \langle \delta_0 \otimes 1, e^{-t\widehat{L}_{\mathbf{k}}} \widehat{\rho}_{0;\mathbf{k}} \otimes 1 \rangle_{L^2(\mathbb{Z}^d \times \Omega)}.$$

This equation is the starting point for our proof of Theorem 1. It indicates that the diffusive limit on the l.h.s. of (2.19) can be studied via a spectral analysis of the semi-group  $e^{-t\widehat{L}_{\mathbf{k}}}$  for  $\mathbf{k}$  in a neighborhood of 0.

#### 4. SPECTRAL ANALYSIS OF $\widehat{L}_{\mathbf{k}}$ AND THE PROOF OF THEOREM 1

In this section inner products and norms are taken in the space  $L^2(\mathbb{Z}^d \times \Omega)$  unless otherwise indicated.

4.1. **Spectral analysis of  $\widehat{L}_0$ .** To begin, let us consider  $\widehat{L}_k$  with  $k = 0$ . A preliminary observation is that

$$(4.1) \quad \widehat{L}_0 \delta_0 \otimes 1 = 0.$$

Ultimately, this identity is a consequence of the fact that  $\sum_x \mathbb{E}(\text{tr } \rho_t)$  is constant in time.

Let  $P_0$  denote orthogonal projection of  $L^2(\mathbb{Z}^d \times \Omega)$  onto the space  $\mathcal{H}_0 = \ell^2(\mathbb{Z}^d) \otimes \{1\}$  of “non-random” functions,

$$(4.2) \quad P_0 \Psi(x) = \int_{\Omega} \Psi(x, \omega) d\mu(\omega).$$

Then  $P_0^\perp = (1 - P_0)$  is the projection onto mean zero functions

$$(4.3) \quad \mathcal{H}_0^\perp = \left\{ \Psi(x, \omega) : \int_{\Omega} \Psi(x, \omega) d\mu(\omega) = 0 \right\}.$$

The block decomposition of  $\widehat{L}_0$  with respect to the direct sum  $\mathcal{H}_0 \oplus \mathcal{H}_0^\perp$  is of the form

$$(4.4) \quad \widehat{L}_0 = \begin{pmatrix} 0 & i\lambda P_0 \widehat{V} \\ i\lambda \widehat{V} P_0 & i\widehat{K}_0 + B + i\lambda P_0^\perp \widehat{V} P_0^\perp \end{pmatrix}.$$

Indeed, it follows from the definition of  $\widehat{K}_0$  that

$$(4.5) \quad \widehat{K}_0 P_0 = P_0 \widehat{K}_0 = 0,$$

and we have seen in §2 that

$$(4.6) \quad P_0 B = B P_0 = 0.$$

(The identity  $P_0 B = 0$  follows since  $B^\dagger P_0 = 0$ .) Thus,  $\text{ran } P_0^\perp = \mathcal{H}_0^\perp$  is an invariant subspace for  $i\widehat{K}_0 + B$  and  $i\widehat{K}_0 + B$  vanishes on  $\text{ran } P_0 = \mathcal{H}_0$ . Since the potentials  $v_x(\omega)$  are mean zero,  $P_0 \widehat{V} P_0 = 0$  and (4.4) follows.

Using the block decomposition (4.4) we now prove the following

**Lemma 3.** *For each  $\lambda > 0$  there is  $\delta_\lambda > 0$ , with  $\delta_\lambda \geq c\lambda^2$  for  $\lambda$  small, such that*

$$(4.7) \quad \sigma(\widehat{L}_0) = \{0\} \cup \Sigma_+$$

where

- (1) 0 is a non-degenerate eigenvalue, and
- (2)  $\Sigma_+ \subset \{z : \text{Re } z > \delta_\lambda\}$ .

*Proof.* First note that  $\text{Re } \widehat{L}_0 = \text{Re } B \geq 0$  in the sense of quadratic forms. Thus by the sectoriality of  $B$

$$(4.8) \quad \left| \text{Im} \langle \Phi, \widehat{L}_0 \Phi \rangle \right| \leq \left\| \widehat{K}_0 + \lambda \widehat{V} \right\| + |\text{Im} \langle \Phi, B \Phi \rangle| \leq \|\widehat{h}\|_\infty + 2\lambda + \gamma \text{Re} \langle \Phi, \widehat{L}_0 \Phi \rangle,$$

if  $\|\Phi\| = 1$ . It follows that the *numerical range* of  $\widehat{L}_0$ ,  $\text{Num}(\widehat{L}_0) = \{ \langle \Phi, \widehat{L}_0 \Phi \rangle : \|\Phi\| = 1 \}$  is contained in

$$(4.9) \quad \mathcal{N}_+ := \left\{ z : \text{Re } z \geq 0 \text{ and } |\text{Im } z| \leq \|\widehat{h}\|_\infty + 2\lambda + \gamma \text{Re } z \right\}.$$

Since  $\sigma(\widehat{L}_0) \subset \text{Num}(\widehat{L}_0)$ , we may restrict our attention to  $z \in \mathcal{N}_+$ .

Now fix  $z \in \mathcal{N}_+$  and consider the equation

$$(4.10) \quad (\widehat{L}_0 - z) \begin{pmatrix} \phi \otimes 1 \\ \Phi \end{pmatrix} = \begin{pmatrix} -z & i\lambda P_0 \widehat{V} \\ i\lambda \widehat{V} P_0 & P_0^\perp \widehat{L}_0 P_0^\perp - z \end{pmatrix} \begin{pmatrix} \phi \otimes 1 \\ \Phi \end{pmatrix} = \begin{pmatrix} \psi \otimes 1 \\ \Psi \end{pmatrix},$$

for  $(\phi \otimes 1, \Phi) \in \mathcal{H}_0 \oplus \mathcal{H}_0^\perp$  given  $(\psi \otimes 1, \Psi) \in \mathcal{H}_0 \oplus \mathcal{H}_0^\perp$ . By the gap condition on  $B$ ,

$$(4.11) \quad \text{Re } P_0^\perp \widehat{L}_0 P_0^\perp = \text{Re}(i\widehat{K}_0 + B + i\lambda P_0^\perp \widehat{V} P_0^\perp) \geq \frac{1}{T} P_0^\perp,$$

so the second of the two equations (4.10) may be solved provided  $\text{Re } z < \frac{1}{T}$  to give

$$(4.12) \quad \Phi = (P_0^\perp \widehat{L}_0 P_0^\perp - z)^{-1} \left[ \Psi - i\lambda \widehat{V} \phi \otimes 1 \right].$$

Thus the first equation of (4.10) reduces to

$$(4.13) \quad [\Gamma(z) - z] \phi \otimes 1 = \psi \otimes 1 - P_0 \widehat{V} (P_0^\perp \widehat{L}_0 P_0^\perp - z)^{-1} \Psi,$$

with

$$(4.14) \quad \Gamma(z) = \lambda^2 P_0 \widehat{V} (P_0^\perp \widehat{L}_0 P_0^\perp - z)^{-1} \widehat{V} P_0.$$

Thus  $\widehat{L}_0 - z$  is boundedly invertible, for  $\text{Re } z < \frac{1}{T}$ , if and only if  $z \notin \sigma(\Gamma(z))$ . However,

$$(4.15) \quad \begin{aligned} & \text{Re} \langle \phi \otimes 1, \Gamma(z) \phi \otimes 1 \rangle \\ &= \lambda^2 \langle \phi \otimes 1, \widehat{V} (P_0^\perp \widehat{L}_0^\dagger P_0^\perp - z^*)^{-1} (\text{Re } B - \text{Re } z) (P_0^\perp \widehat{L}_0 P_0^\perp - z)^{-1} \widehat{V} \phi \otimes 1 \rangle \\ &\geq \lambda^2 \left( \frac{1}{T} - \text{Re } z \right) \left\| (P_0^\perp \widehat{L}_0 P_0^\perp - z)^{-1} \widehat{V} \phi \otimes 1 \right\|^2 \\ &= \lambda^2 \left( \frac{1}{T} - \text{Re } z \right) \left\| (B^{-1} P_0^\perp (\widehat{L}_0 - z) P_0^\perp)^{-1} B^{-1} \widehat{V} \phi \otimes 1 \right\|^2, \end{aligned}$$

where the inverse of  $B$  is well defined since  $\widehat{V} \phi \otimes 1 \in \mathcal{H}_0^\perp = \text{ran } P_0^\perp$ . Furthermore  $B^{-1}$  is bounded on  $\mathcal{H}_0^\perp$ , with  $\|B^{-1} P_0^\perp\| \leq T$ . Thus  $B^{-1} P_0^\perp (\widehat{L}_0 - z) P_0^\perp$  is bounded,

$$(4.16) \quad \left\| B^{-1} P_0^\perp (\widehat{L}_0 - z) P_0^\perp \right\| \leq 1 + T \left( \|\widehat{h}\|_\infty + 2\lambda + |z| \right),$$

and it follows that

$$\begin{aligned}
(4.17) \quad \operatorname{Re}\langle \phi \otimes 1, \Gamma(z)\phi \otimes 1 \rangle_{\ell^2(\mathbb{Z}^d)} &\geq \lambda^2 \left( \frac{1}{T} - \operatorname{Re} z \right) \frac{1}{\left[ 1 + T \left( \|\widehat{h}\|_\infty + 2\lambda + |z| \right) \right]^2} \left\| B^{-1} \widehat{V} \phi \otimes 1 \right\|^2 \\
&\geq \lambda^2 \chi^2 \left( \frac{1}{T} - \operatorname{Re} z \right) \frac{1}{\left[ 1 + T \left( \|\widehat{h}\|_\infty + 2\lambda + |z| \right) \right]^2} \sum_{x \neq 0} |\phi(x)|^2,
\end{aligned}$$

where we have made use of the non-degeneracy assumption on the potential.

For  $z \in \mathcal{N}_+ \cap \left\{ \operatorname{Re} z < \frac{1}{T} \right\}$  we have

$$(4.18) \quad \|\widehat{h}\|_\infty + 2\lambda + |z| \leq 2\|\widehat{h}\|_\infty + 4\lambda + (1 + \gamma) \frac{1}{T}.$$

Thus

$$(4.19) \quad \operatorname{Re} \Gamma(z) \geq \frac{1 - T \operatorname{Re} z}{T} \left[ \frac{\lambda \chi}{2 + \gamma + 2T \|\widehat{h}\|_\infty + 4T\lambda} \right]^2 (1 - \Pi_0),$$

with  $\Pi_0$  the projection of  $\mathcal{H}_0$  onto  $\delta_0$ . It follows that

- (1)  $\Gamma(0)$  is invertible off the range of  $\Pi_0$ , so 0 is a non-degenerate eigenvalue of  $\widehat{L}_0$ ,
- (2)  $\Gamma(z) - z$  is invertible if  $z \in \mathcal{N}_+$ ,  $z \neq 0$ ,

$$(4.20) \quad \operatorname{Re} z \leq \frac{r}{T}, \quad \text{and} \quad \operatorname{Re} z \leq \frac{1-r}{T} \left[ \frac{\lambda \chi}{2 + \gamma + 2T \|\widehat{h}\|_\infty + 4T\lambda} \right]^2,$$

for some  $r \in (0, 1)$ .

Optimizing the choice of  $r$  gives the following explicit expression for  $\delta_\lambda$ :

$$(4.21) \quad \delta_\lambda = \frac{1}{T} \frac{\lambda^2 \chi^2}{\left( 2 + \gamma + 2T \|\widehat{h}\|_\infty + 4T\lambda \right)^2 + \lambda^2 \chi^2}. \quad \square$$

The spectral gap  $\delta_\lambda$  has consequences for the dynamics of the semi-group:

**Lemma 4.** *Let  $Q_0 =$  orthogonal projection onto  $\delta_0 \otimes 1$  in  $L^2(\mathbb{Z}^d \times \Omega)$ . Then  $e^{-t\widehat{L}_0}(1 - Q_0)$  is a contraction semi-group on  $\operatorname{ran}(1 - Q_0)$ , and for all sufficiently small  $\epsilon > 0$  there is  $C_\epsilon > 0$  such that*

$$(4.22) \quad \left\| e^{-t\widehat{L}_0}(1 - Q_0) \right\| \leq C_\epsilon e^{-t(\delta_\lambda - \epsilon)}$$

*Proof.* That  $e^{-t\widehat{L}_0}(1 - Q_0)$  is a semi-group on  $\operatorname{ran}(1 - Q_0)$  with generator  $\widehat{L}_0^{(0)} = \widehat{L}_0|_{\operatorname{ran}(1 - Q_0)}$  is clear. Since  $\sigma(\widehat{L}_0) \subset \{\operatorname{Re} z \geq \delta_\lambda\}$ , it is known that a sufficient condition for (4.22) to hold is for  $e^{-t\widehat{L}_0}(1 - Q_0)$  to be an analytic semi-group (see [1, Chapter IV Corollary 3.12]). A

necessary and sufficient condition for  $e^{-t\widehat{L}_0}(1-Q_0)$  to be analytic, given that  $\widehat{L}_0$  is maximally dissipative, is that the following estimate

$$(4.23) \quad \left\| \frac{1}{\zeta - \widehat{L}_0^{(0)}} \right\|_{\text{ran}(1-Q_0)} \leq \frac{C}{|\text{Im } \zeta|}$$

holds for all  $\zeta \in \mathbb{C}$  with  $\text{Re } \zeta < 0$ . (See [1, Chapter II Theorem 4.6].)

For any invertible closed operator  $A$ , we have  $\|A^{-1}\| \leq 1/\inf\{|z| : z \in \text{Num}(A)\}$ , with  $\text{Num}(A)$  the numerical range of  $A$ . Since  $\text{Num}(\widehat{L}_0^{(0)}) \subset \mathcal{N}_+$ , with  $\mathcal{N}_+$  as in (4.9), we have

$$(4.24) \quad \left\| \frac{1}{\zeta - \widehat{L}_0^{(0)}} \right\|_{\text{ran}(1-Q_0)} \leq \frac{1}{\text{dist}(\zeta, \mathcal{N}_+)}.$$

It follows that (4.23) holds for all  $\zeta$  outside a rectangle of the form  $R = [-a, 0] + i[-M, M]$ . On the other hand, by Lemma 3, this rectangle  $R$  is contained in the resolvent set of  $\widehat{L}_0^{(0)}$ . Since  $R$  is compact, we have

$$(4.25) \quad \sup_{\zeta \in R} \left\| \frac{1}{\zeta - \widehat{L}_0^{(0)}} \right\|_{\text{ran}(1-Q_0)} < \infty,$$

which is stronger than (4.23) for  $\zeta \in R$ . □

**4.2. Analytic perturbation theory for  $\widehat{L}_k$ .** Now that we have established a strict spectral gap for  $\widehat{L}_0$ , it follows that the gap persists in the spectrum of  $\widehat{L}_k$  for  $k$  sufficiently small. Indeed,

$$(4.26) \quad [\widehat{L}_k - \widehat{L}_0]\phi(x, \omega) = \sum_{\zeta} h(\zeta)(1 - e^{-i\mathbf{k} \cdot \zeta})\phi(x - \zeta, \sigma_{\zeta}\omega),$$

so

$$(4.27) \quad \left\| \widehat{L}_k - \widehat{L}_0 \right\| \leq c|k|.$$

Thus, an immediately corollary of Lemma 3 is

**Lemma 5.** *If  $|k|$  is sufficiently small, the spectrum of  $\widehat{L}_k$  consists of:*

- (1) *A non-degenerate eigenvalue  $E(k)$  contained in  $H_0 = \{z : |z| < c|k|\}$ .*
- (2) *The rest of the spectrum is contained in the half plane  $H_1 = \{z : \text{Re } z > \delta_{\lambda} - c|k|\}$  such that  $H_0 \cap H_1 = \emptyset$ .*

Furthermore,  $E(k)$  is  $C^2$  in a neighborhood of 0,

$$(4.28) \quad E(\mathbf{0}) = 0, \quad \nabla E(\mathbf{0}) = 0,$$

and

$$(4.29) \quad \begin{aligned} \partial_i \partial_j E(\mathbf{0}) &= 2 \operatorname{Re} \langle \partial_i \widehat{K}_0 \delta_0 \otimes 1, [\widehat{L}_0]^{-1} \partial_j \widehat{K}_0 \delta_0 \otimes 1 \rangle \\ &= 2 \operatorname{Re} \sum_{x, y \in \mathbb{Z}^d} x_i y_j \overline{h(x)} h(y) \langle \delta_x \otimes 1, [\Gamma(0)]^{-1} \delta_y \otimes 1 \rangle, \end{aligned}$$

with

$$(4.30) \quad \Gamma(0) = \lambda^2 P_0 \widehat{V} (P_0^\perp \widehat{L}_0 P_0^\perp)^{-1} \widehat{V} P_0$$

as in the proof of Lemma 3. In particular,  $\partial_i \partial_j E(\mathbf{0})$  is positive definite.

*Proof.* These are all standard facts from analytic perturbation theory — see for instance [5]. A sketch of the proof is as follows. First, it is a general fact that the spectrum moves no further than the norm of the perturbation, so

$$(4.31) \quad \sigma(\widehat{L}_{\mathbf{k}}) \subset \{z : |z| < c|\mathbf{k}|\} \cup \{z : \operatorname{Re} z > \delta_\lambda - c|\mathbf{k}|\}.$$

For sufficiently small  $\mathbf{k}$  the two sets on the r.h.s. are disjoint and we may fit a contour  $\mathcal{C}$  which winds around the origin between them. The (non-Hermitian) Riesz projection

$$(4.32) \quad Q_{\mathbf{k}} = \frac{1}{2\pi i} \int_{\mathcal{C}} \frac{1}{z - \widehat{L}_{\mathbf{k}}} dz$$

is rank one at  $\mathbf{k} = 0$  and continuous as a function of  $\mathbf{k}$ . It follows that  $Q_{\mathbf{k}}$  is rank one as long as  $\sigma(\widehat{L}_{\mathbf{k}})$  does not intersect the contour  $\mathcal{C}$ . Thus for small  $\mathbf{k}$  the only spectrum of  $\widehat{L}_{\mathbf{k}}$  in a neighborhood of zero is a non-degenerate eigenvalue, with associated eigenvector in the one-dimensional range of  $Q_{\mathbf{k}}$ .

Let us call the eigenvalue  $E(\mathbf{k})$ , and the associated normalized eigenvector  $\Phi_{\mathbf{k}}$ . Clearly  $E(\mathbf{0}) = 0$  and  $\Phi_0 = \delta_0 \otimes 1$ . Since

$$(4.33) \quad \widehat{L}_{\mathbf{k}} Q_{\mathbf{k}} = E(\mathbf{k}) Q_{\mathbf{k}}$$

we may compute the derivatives of  $E(\mathbf{k})$  by differentiating  $\widehat{L}_{\mathbf{k}}$  and the projection. In particular,  $\nabla E(\mathbf{k})$  is given by the so-called Feynman-Hellman formula

$$(4.34) \quad \nabla E(\mathbf{k}) = \langle \Phi_{\mathbf{k}}, \nabla \widehat{L}_{\mathbf{k}} \Phi_{\mathbf{k}} \rangle,$$

from which it follows that  $\nabla E(\mathbf{0}) = 0$  since  $\nabla \widehat{L}_{\mathbf{k}} = i \nabla \widehat{K}_{\mathbf{k}}$  is off-diagonal in the position basis on  $\mathcal{H}_0$ . Similarly, we have

$$(4.35) \quad \begin{aligned} \partial_i \partial_j E(\mathbf{k}) &= \langle \Phi_{\mathbf{k}}, \partial_i \partial_j \widehat{L}_{\mathbf{k}} \Phi_{\mathbf{k}} \rangle + \langle \partial_i \widehat{L}_{\mathbf{k}} \Phi_{\mathbf{k}}, (1 - Q_{\mathbf{k}}) \widehat{L}_{\mathbf{k}}^{-1} (1 - Q_{\mathbf{k}}) \partial_j \widehat{L}_{\mathbf{k}} \Phi_{\mathbf{k}} \rangle \\ &\quad + \langle \partial_j \widehat{L}_{\mathbf{k}} \Phi_{\mathbf{k}}, (1 - Q_{\mathbf{k}}) \widehat{L}_{\mathbf{k}}^{-1} (1 - Q_{\mathbf{k}}) \partial_i \widehat{L}_{\mathbf{k}} \Phi_{\mathbf{k}} \rangle \end{aligned}$$

The first term on the r.h.s. vanishes at  $\mathbf{k} = 0$  and the other two combine to give the identities claimed in the Lemma. Since  $[\Gamma(0)]^{-1}$  is positive definite, it follows from the non-degeneracy condition (3b) that  $\partial_i \partial_j E(\mathbf{0})$  is positive definite as well.  $\square$

Again, we obtain dynamical information about the semi-group  $e^{-t\widehat{L}_{\mathbf{k}}}$ :

**Lemma 6.** *If  $\epsilon$  is sufficiently small, then there is  $C_\epsilon < \infty$  such that*

$$(4.36) \quad \left\| e^{-t\widehat{L}_{\mathbf{k}}}(1 - Q_{\mathbf{k}}) \right\| \leq C_\epsilon e^{-t(\delta_\lambda - \epsilon - c|\mathbf{k}|)}$$

for all sufficiently small  $\mathbf{k}$ .

*Proof.* Since  $\text{Num}(\widehat{L}_{\mathbf{k}}) \subset \mathcal{N}_+$ , with  $\mathcal{N}_+$  as in (4.9) in the proof of Lemma 3, this is essentially identical to the proof of Lemma 4. For  $\mathbf{k}$  in a compact neighborhood of the origin, we can choose the bound  $C_\epsilon$  uniform in  $\mathbf{k}$ .  $\square$

**4.3. Proof of Theorem 1.** A first observation is that it suffices to prove diffusion (2.19) under the assumption that the initial density matrix  $\rho_0$  satisfies

$$(4.37) \quad \sum_{x,y} |\rho_0(x,y)| < \infty.$$

To see this, it is useful to note:

- (1) The evolution (2.17) preserves the *trace norm* of  $\rho_0$ .
- (2) Any  $\rho \in \mathcal{DM}$  may be approximated, to arbitrary precision, by an operator  $\rho_0 \in \mathcal{DM}$  satisfying (4.37).

In more detail, recall that the *trace norm* of an operator  $A$  is

$$(4.38) \quad \|A\|_{\mathcal{T}_1} = \sup\{|\text{tr} AB| : B \text{ is finite rank and } \|B\| \leq 1\}.$$

Since the evolution (2.17) is given by  $\rho_t = U\rho_0U^\dagger$ , with  $U = U(t, 0)$  the unitary propagator of (1.1), we see that  $\|\rho_t\|_{\mathcal{T}_1} = \|\rho_0\|_{\mathcal{T}_1}$ . On the other hand, given  $\rho_0$  of trace class and  $\epsilon > 0$  we can find  $\tilde{\rho}_0$  satisfying (4.37) and such that  $\|\rho_0 - \tilde{\rho}_0\|_{\mathcal{T}_1} < \epsilon$ . Indeed, we may take

$$(4.39) \quad \tilde{\rho}_0(x,y) = \begin{cases} \rho_0(x,y) & \text{if } |x|, |y| < L \\ 0 & \text{if } |x| \geq L \text{ or } |y| \geq L \end{cases}$$

with  $L$  sufficiently large. Then,

$$(4.40) \quad \left| \sum_x e^{-i\mathbf{k}\cdot x} \mathbb{E}(\rho_t(x,x) - \tilde{\rho}_t(x,x)) \right| \leq \sum_x \mathbb{E}(|\rho_t(x,x) - \tilde{\rho}_t(x,x)|) \leq \mathbb{E}(\|\rho_t - \tilde{\rho}_t\|_{\mathcal{T}_1}) \leq \epsilon.$$

If diffusion (2.19) holds for any  $\tilde{\rho}_0$  satisfying (4.37) we learn that

$$(4.41) \quad \limsup_{\tau \rightarrow \infty} \left| \sum_x e^{-i\frac{1}{\sqrt{\tau}}\mathbf{k}\cdot x} \mathbb{E}(\rho_{\tau t}(x,x)) - (\text{tr } \rho_0) e^{-\sum_{i,j} D_{i,j} \mathbf{k}_i \mathbf{k}_j} \right| \leq 2\epsilon.$$

which gives diffusion for  $\rho_0$  in the limit  $\epsilon \rightarrow 0$ .

Turning now to  $\rho_0$  which satisfies (4.37) we see that

$$(4.42) \quad \widehat{\rho}_{0,\mathbf{k}}(x) = \sum_y \rho_0(x+y,y) e^{i\mathbf{k}\cdot y}$$

defines a function that is uniformly bounded in  $\ell^2(\mathbb{Z}^d)$  as  $\mathbf{k}$  varies through the torus,

$$(4.43) \quad \left[ \sum_x |\rho_{0;\mathbf{k}}(x)|^2 \right]^{\frac{1}{2}} \leq \sum_x |\rho_{0;\mathbf{k}}(x)| \leq \sum_{x,y} |\rho_0(x,y)| := M < \infty.$$

By (3.18) we have

$$(4.44) \quad \sum_x e^{-i\frac{1}{\sqrt{\tau}}\mathbf{k}\cdot x} \mathbb{E}(\rho_{\tau t}(x,x)) = \langle \delta_0 \otimes 1, e^{-\tau t \widehat{L}_{\mathbf{k}/\sqrt{\tau}}} \widehat{\rho}_{0;\frac{1}{\sqrt{\tau}}\mathbf{k}} \otimes 1 \rangle.$$

Letting  $Q_{\mathbf{k}}$  denote the Riesz projection onto the eigenvector of  $\widehat{L}_{\mathbf{k}}$  near zero — see (4.32) in the proof of Lemma 5 —, we have

$$(4.45) \quad \sum_x e^{-i\frac{1}{\sqrt{\tau}}\mathbf{k}\cdot x} \mathbb{E}(\rho_{\tau t}(x,x)) = e^{-\tau t E(\mathbf{k}/\sqrt{\tau})} \langle \delta_0 \otimes 1, Q_{\frac{1}{\sqrt{\tau}}\mathbf{k}} \widehat{\rho}_{0;\frac{1}{\sqrt{\tau}}\mathbf{k}} \otimes 1 \rangle \\ + \langle \delta_0 \otimes 1, e^{-\tau t \widehat{L}_{\mathbf{k}/\sqrt{\tau}}} (1 - Q_{\frac{1}{\sqrt{\tau}}\mathbf{k}}) \widehat{\rho}_{0;\frac{1}{\sqrt{\tau}}\mathbf{k}} \otimes 1 \rangle,$$

for  $\tau$  sufficiently large. By Lemma 6, the second term in (4.45) is exponentially small in the large  $\tau$  limit,

$$(4.46) \quad \left| \langle \delta_0 \otimes 1, (1 - Q_{\frac{1}{\sqrt{\tau}}\mathbf{k}}) e^{-\tau t \widehat{L}_{\mathbf{k}/\sqrt{\tau}}} \widehat{\rho}_{0;\frac{1}{\sqrt{\tau}}\mathbf{k}} \otimes 1 \rangle \right| \\ \leq \left\| (1 - Q_{\frac{1}{\sqrt{\tau}}\mathbf{k}}) e^{-\tau t \widehat{L}_{\mathbf{k}/\sqrt{\tau}}} \right\| \left\| \widehat{\rho}_{0;\frac{1}{\sqrt{\tau}}\mathbf{k}} \otimes 1 \right\| \leq MC_\epsilon e^{-\tau t(\delta_\lambda - \epsilon - c|\mathbf{k}|/\sqrt{\tau})} \rightarrow 0.$$

Regarding the first term in (4.45), we have by Taylor's formula,

$$(4.47) \quad E(\mathbf{k}/\sqrt{\tau}) = \frac{1}{2\tau} \sum_{i,j} \partial_i \partial_j E(\mathbf{0}) \mathbf{k}_i \mathbf{k}_j + o\left(\frac{1}{\tau}\right),$$

since  $E(\mathbf{0}) = \nabla E(\mathbf{0}) = 0$ . Thus

$$(4.48) \quad e^{-\tau t E(\mathbf{k}/\sqrt{\tau})} = e^{-t \frac{1}{2} \sum_{i,j} \partial_i \partial_j E(\mathbf{0}) \mathbf{k}_i \mathbf{k}_j} + o(1).$$

Putting together (4.48) and (4.46) yields

$$(4.49) \quad \sum_x e^{-i\frac{1}{\sqrt{\tau}}\mathbf{k}\cdot x} \mathbb{E}(\rho_{\tau t}(x,x)) = e^{-t \frac{1}{2} \sum_{i,j} \partial_i \partial_j E(\mathbf{0}) \mathbf{k}_i \mathbf{k}_j} \langle \delta_0 \otimes 1, \widehat{\rho}_{0;\frac{1}{\sqrt{\tau}}\mathbf{k}} \otimes 1 \rangle + o(1) \\ = e^{-t \frac{1}{2} \sum_{i,j} \partial_i \partial_j E(\mathbf{0}) \mathbf{k}_i \mathbf{k}_j} \widehat{\rho}_{0;\frac{1}{\sqrt{\tau}}\mathbf{k}}(0) + o(1) \rightarrow e^{-t \frac{1}{2} \sum_{i,j} \partial_i \partial_j E(\mathbf{0}) \mathbf{k}_i \mathbf{k}_j} \widehat{\rho}_{0;\mathbf{0}}(0)$$

since  $Q_{\mathbf{k}}^\dagger \delta_0 \otimes 1 \rightarrow \delta_0 \otimes 1$  as  $\mathbf{k} \rightarrow 0$  and  $\widehat{\rho}_{0;\mathbf{k}}(0)$  is continuous as a function of  $\mathbf{k}$ .

Eq. (4.49) completes the proof of diffusion (2.19), with the diffusion matrix given by  $D_{i,j} = \frac{1}{2} \partial_i \partial_j E(\mathbf{0})$ . From the explicit expression (4.29) for  $\partial_i \partial_j E(\mathbf{0})$  in Lemma 5, we see that the asymptotic form (2.20) holds, with

$$(4.50) \quad D_{i,j}^0 = \text{Re} \left\langle \partial_i \widehat{K}_0 \delta_0 \otimes 1, \left[ P_0 \widehat{V} P_0^\perp \left( i \widehat{K}_0 + B \right)^{-1} P_0^\perp \widehat{V} P_0 \right]^{-1} \partial_j \widehat{K}_0 \delta_0 \otimes 1 \right\rangle.$$



To derive (2.21), diffusive scaling for  $\sum_x \mathbb{E}(|x|^2 \rho_t(x, x))$ , note that by (3.18)

$$(4.51) \quad \frac{1}{t} \sum_x x^2 \mathbb{E}(\rho_t(x, x)) = -\frac{1}{t} \Delta_{\mathbf{k}} \langle \delta_0 \otimes 1, e^{-t\widehat{L}_{\mathbf{k}}} \widehat{\rho}_{0;\mathbf{k}} \otimes 1 \rangle \Big|_{\mathbf{k}=0}.$$

Expanding the r.h.s. we obtain

$$(4.52) \quad \frac{1}{t} \sum_x x^2 \mathbb{E}(\rho_t(x, x)) = \underbrace{-\frac{1}{t} \langle \delta_0 \otimes 1, e^{-t\widehat{L}_0} \Delta_{\mathbf{k}} \widehat{\rho}_{0;\mathbf{k}} \Big|_{\mathbf{k}=0} \otimes 1 \rangle}_{\text{I}} \\ - \underbrace{\frac{2}{t} \sum_{i=1}^d \langle \delta_0 \otimes 1, \partial_i e^{-t\widehat{L}_{\mathbf{k}}} \Big|_{\mathbf{k}=0} \partial_i \widehat{\rho}_{0;\mathbf{k}} \Big|_{\mathbf{k}=0} \otimes 1 \rangle}_{\text{II}} \\ - \underbrace{\frac{1}{t} \langle \delta_0 \otimes 1, \Delta_{\mathbf{k}} e^{-t\widehat{L}_{\mathbf{k}}} \Big|_{\mathbf{k}=0} \widehat{\rho}_{0;0} \otimes 1 \rangle}_{\text{III}}.$$

The first term is negligible as  $t \rightarrow \infty$ ,

$$(4.53) \quad \text{I} = -\frac{1}{t} \langle \delta_0 \otimes 1, e^{-t\widehat{L}_0} \Delta_{\mathbf{k}} \widehat{\rho}_{0;\mathbf{k}} \Big|_{\mathbf{k}=0} \otimes 1 \rangle = -\frac{1}{t} \sum_x x^2 \rho_0(x, x) = O\left(\frac{1}{t}\right),$$

since  $e^{-t\widehat{L}_0^\dagger} \delta_0 \otimes 1 = \delta_0 \otimes 1$ .

To proceed, we need to recall the formula for differentiating a semi-group,

$$(4.54) \quad \partial_i e^{-t\widehat{L}_{\mathbf{k}}} = -\int_0^t e^{-(t-s)\widehat{L}_{\mathbf{k}}} \left( \partial_i \widehat{L}_{\mathbf{k}} \right) e^{-s\widehat{L}_{\mathbf{k}}} ds.$$

Thus the second term on the r.h.s. of (4.52) is equal to

$$(4.55) \quad \text{II} = -\frac{2}{t} \int_0^t \sum_{i=1}^d \langle \partial_i \widehat{L}_0 \delta_0 \otimes 1, e^{-s\widehat{L}_0} (1 - Q_0) \partial_i \widehat{\rho}_{0;0} \otimes 1 \rangle ds,$$

where we have recalled that

- (1)  $[\partial_i \widehat{L}_{\mathbf{k}}]^\dagger = -\partial_i \widehat{L}_{\mathbf{k}}$ ,
- (2)  $e^{-s\widehat{L}_0} Q_0 = Q_0$ , and
- (3)  $Q_0 \partial_i \widehat{L}_0 = 0$ .

By Lemma 4, we see that II is negligible in the large  $t$  limit:

$$(4.56) \quad |\text{II}| \lesssim \frac{1}{t} \int_0^\infty e^{-s(\delta_\lambda - \epsilon)} ds = O\left(\frac{1}{t}\right).$$

It remains to compute the third term III, for which we have

$$\begin{aligned}
(4.57) \quad \text{III} &= \underbrace{-\frac{1}{t} \int_0^t \langle \Delta_{\mathbf{k}} \widehat{L}_{\mathbf{k}} |_{\mathbf{k}=0} \delta_0 \otimes 1, e^{-s\widehat{L}_0} (1 - Q_0) \widehat{\rho}_{0;\mathbf{0}} \otimes 1 \rangle ds}_{\text{IIIa}} \\
&+ \underbrace{\frac{1}{t} \int_0^t \int_0^s \sum_{i=1}^d \langle \partial_i \widehat{L}_0 \delta_0 \otimes 1, e^{-(s-r)\widehat{L}_0} (1 - Q_0) \partial_i \widehat{L}_0 e^{-r\widehat{L}_0} (1 - Q_0) \widehat{\rho}_{0;\mathbf{0}} \otimes 1 \rangle dr ds}_{\text{IIIb}} \\
&+ \underbrace{\frac{1}{t} \int_0^t \int_0^s \sum_{i=1}^d \langle \partial_i \widehat{L}_0 \delta_0 \otimes 1, e^{-(s-r)\widehat{L}_0} (1 - Q_0) \partial_i \widehat{L}_0 Q_0 \widehat{\rho}_{0;\mathbf{0}} \otimes 1 \rangle dr ds}_{\text{IIIc}}
\end{aligned}$$

Similar to what we had for II, we find that IIIa =  $O\left(\frac{1}{t}\right)$ , and also that

$$(4.58) \quad |\text{IIIb}| \lesssim \frac{1}{t} \int_0^\infty \int_0^s e^{-s(\delta_\lambda - \epsilon)} dr ds = O\left(\frac{1}{t}\right).$$

Finally, for IIIc we have

$$\begin{aligned}
(4.59) \quad \text{IIIc} &= \sum_{i=1}^d \langle \partial_i \widehat{L}_0 \delta_0 \otimes 1, \widehat{L}_0^{-1} (1 - Q_0) \partial_i \widehat{L}_0 Q_0 \widehat{\rho}_{0;\mathbf{0}} \otimes 1 \rangle + O\left(\frac{1}{t}\right), \\
&= (\text{tr } \rho_0) \sum_{i=1}^d \langle \partial_i \widehat{K}_0 \delta_0 \otimes 1, \widehat{L}_0^{-1} \partial_i \widehat{K}_0 \delta_0 \otimes 1 \rangle + O\left(\frac{1}{t}\right),
\end{aligned}$$

since  $\nabla \widehat{L}_{\mathbf{k}} = i \nabla \widehat{K}_{\mathbf{k}}$ . Comparing with the expression (4.29) for  $\partial_i \partial_j E(\mathbf{0})$  yields (2.21).  $\square$

**Acknowledgments.** We wish to express our gratitude for the hospitality extended to both of us by the Erwin Schrödinger Institute and to J.S. by the Isaac Newton Institute. J. S. would like to thank Tom Spencer, Jürg Fröhlich and Wojciech De Roeck for interesting discussions related to this work. We would also like to thank Sergey Denisov for pointing out that the non-degeneracy condition (3b) on the hopping operator was missing from our original manuscript.

## REFERENCES

1. Klaus-Jochen Engel and Rainer Nagel, *One-parameter semigroups for linear evolution equations*, Graduate Texts in Mathematics, vol. 194, Springer-Verlag, New York, 2000, With contributions by S. Brendle, M. Campiti, T. Hahn, G. Metafune, G. Nickel, D. Pallara, C. Perazzoli, A. Rhandi, S. Romanelli and R. Schnaubelt. MR MR1721989 (2000i:47075)
2. László Erdős, Manfred Salmhofer, and Horng-Tzer Yau. *Quantum diffusion for the Anderson model in the scaling limit*. Ann. Henri Poincaré **8** (2007), no. 4, 621–685. MR MR2333778 (2008g:82053)

3. László Erdős, Manfred Salmhofer, and Horng-Tzer Yau. *Quantum diffusion of the random Schrödinger evolution in the scaling limit. II. The recollision diagrams*. *Comm. Math. Phys.* **271** (2007) no. 1, 1–53. MR MR2283953 (2008h:82035)
4. László Erdős, Manfred Salmhofer, and Horng-Tzer Yau. *Quantum diffusion of the random Schrödinger evolution in the scaling limit*. *Acta Math.* **200** (2008), no. 2, 211–277, MR MR2413135
5. Tosio Kato, *Perturbation theory for linear operators*, Classics in Mathematics, Springer-Verlag, Berlin, 1995, Reprint of the 1980 edition. MR MR1335452 (96a:47025)
6. P P Mitra and J B Stark, *Nonlinear limits to the information capacity of optical fibre communications*, *Nature* **411** (2001), no. 6841, 1027–30.
7. A. A. Ovchinnikov and N. S. Érikhman, *Motion of a quantum particle in a stochastic medium*, *Soviet Journal of Experimental and Theoretical Physics* **40** (1974), 733.
8. Claude-Alain Pillet, *Some results on the quantum dynamics of a particle in a Markovian potential*, *Comm. Math. Phys.* **102** (1985), no. 2, 237–254. MR MR820574 (88i:82015)
9. Serguei Tcheremchantsev, *Markovian Anderson model: bounds for the rate of propagation*, *Comm. Math. Phys.* **187** (1997), no. 2, 441–469. MR MR1463837 (98h:82029)
10. Serguei Tcheremchantsev, *Transport properties of Markovian Anderson model*, *Comm. Math. Phys.* **196** (1998), no. 1, 105–131. MR MR1643517 (99g:82075)
11. J. Fröhlich W. De Roeck and A. Pizzo, *Diffusion for a quantum particle coupled to an array of independent thermal fields*, 2008 preprint, arXiv:0810.4537v1.

MICHIGAN STATE UNIVERSITY, EAST LANSING, MI 48823  
*E-mail address:* jeffrey@math.msu.edu