Energy Growth in Schrödinger's Equation with Markovian Forcing

M. Burak Erdoğan[1], **Rowan Killip**[2], **Wilhelm Schlag**[3]

[1] Department of Mathematics, University of California, Berkeley, CA 94720-3840, USA
[2] Department of Mathematics, University of Pennsylvania, 209 South 33rd Street, Philadelphia, PA 19104-6395, USA
[3] Department of Mathematics 253-37, Caltech, Pasadena, CA 91125, USA

Abstract: Schrödinger's equation is considered on a one-dimensional torus with time dependent potential $v(\theta, t) = \lambda V(\theta) X(t)$, where $V(\theta)$ is an even trigonometric polynomial and $X(t)$ is a stationary Markov process. It is shown that when the coupling constant λ is sufficiently small, the average kinetic energy grows as the square-root of time. More generally, the H^s norm of the wave function is shown to behave as $t^{s/4}$.

1. Introduction

We study the long term behaviour of a quantum mechanical particle moving on the circle in the presence of a time-dependent potential. The evolution of the wave function ψ is described by the Schrödinger equation

$$i\frac{d\psi}{dt}(\theta, t) = -\frac{d^2\psi}{d\theta^2}(\theta, t) + v(\theta, t)\psi(\theta, t), \tag{1}$$

where we regard the circle as $\mathbf{R}/2\pi\mathbf{Z}$. We are interested in how the kinetic energy,

$$\int_{-\pi}^{\pi} |\partial_\theta \psi(\theta, t)|^2 d\theta = \|\psi\|_{H^1(d\theta)}^2(t) - \|\psi\|_{L^2(d\theta)}^2 \tag{2}$$

grows as a function of time. In fact, we will determine the behaviour of all Sobolev norms. Note that since v will be real, $\|\psi\|_{L^2}$ is conserved.

In two recent papers ([3, 4] see also [5, Appendix 1]), Bourgain studied the case where $v(\theta, t)$ is analytic/smooth in θ and quasi-periodic/smooth in t. In particular, he showed [4] that if

$$\sup_{\theta,t} \left|\partial_\theta^\alpha \partial_t^\beta v(\theta, t)\right| < \infty \quad \text{for all } \alpha, \beta = 0, 1, \ldots,$$

then for any $s, \epsilon > 0$, $\|\psi\|_{H^s} = O(t^\epsilon)$ as $t \to \infty$. (This result holds in any number of space dimensions.) Conversely, it is shown [3] that energy may grow logarithmically

even for t-almost periodic and θ-smooth choices of v. Bourgain also gave an example of a random model exhibiting polynomial growth [4], which we will discuss in due course.

In contradistinction to Bourgain, we are primarily interested in the case where $v(\theta, t)$ is not a smooth function of t. To show that energy growth is a generic phenomenon, there seems no real alternative than to consider a random model. We consider

$$v(\theta, t) = \lambda X(t) V(\theta), \tag{3}$$

where $X(t)$ is a stationary Markov process and $V(\theta)$ is an even trigonometric polynomial and show that if the coupling constant λ is sufficiently small then, on average, the energy grows as the square root of time. One physical interpretation of this model is as a rigid rotator coupled to a classical heat bath.

Let us describe our requirements for the process $X(t)$. We will also take the opportunity to introduce some notation.

Hypothesis 1. *The process $X(t)$ is a stationary Markov process with state space $\mathcal{S} \subseteq \mathbf{R}$ and stationary distribution $d\mu$. It is further assumed that:*

  (i) *the Markov process is generated by a self-adjoint operator B on $L^2(\mathcal{S}; d\mu)$;*
 (ii) *the generator B is positive semi-definite with discrete spectrum;*
(iii) *zero is a simple eigenvalue of B;*
 (iv) *$\int x \, d\mu(x) = 0$ and $\int x^2 \, d\mu = 1$;*
  (v) *the function $u(x) = x$ belongs to the quadratic form domain of B; and*
 (vi) *the co-ordinate operator $x : u(x) \mapsto xu(x)$ is relatively B-bounded with relative bound 0.*

Expectation with respect to this process will be denoted by $\mathbb{E}$, the eigenvalues of B by $0 = \kappa_0 < \kappa_1 \leq \cdots \leq \kappa_N \leq \cdots$ (repeated according to multiplicity) and the corresponding eigenvectors by u_N. Of necessity, $u_0(x) \equiv 1$.

Remarks. 1. By "generator of the Markov process" we mean the operator associated with the forward Kolmogorov equation. Many probability books use the generator associated with the backward equation. As the two operators are the adjoints of one another and we assume that B is self-adjoint, the distinction is moot here.

The meaning of B and μ is perhaps best described with a sample calculation (see also (10) below). The probability that $X(t) \in A$ is equal to $\mu(A)$ and the probability density (with respect to μ) for $X(t)$ given that $X(t) \in A$ is $\chi_A(x)/\mu(A)$. For $s > t$, the probability density for $X(s)$ given that $X(t) \in A$ is $\left[e^{-B(s-t)}\chi_A\right](x)/\mu(A)$ and so the probability that the process passes through the sets A_1, A_2 at times $t_1 < t_2$ is

$$\int \chi_{A_2}(x)\left[e^{-B(t_2-t_1)}\chi_{A_1}\right](x)d\mu(x).$$

2. Regarding (ii) and (iii), the existence of a spectral gap (i.e., that zero is at an isolated point in the spectrum) is essential both for our analysis and, we believe, for the result. However, while ample for our interests, the assumption that the spectrum is discrete is presumably unnecessary.

3. Our arguments remain valid if one chooses $X(t)$ to be the projection under a suitable map of a stationary Markov process on a more general state space. The simplest such modification is to retain the state space $\mathcal{S}$ but consider $Y(t) = f(X(t))$ for some $f : \mathbf{R} \to \mathbf{R}$. Such a change would not affect the arguments presented here provided the analogues of (iv), (v), and (vi) held. That is, provided $\int f \, d\mu = 0$, $\int f^2 \, d\mu = 1$,

    f belongs to the quadratic form domain of B, and $u(x) \mapsto f(x)u(x)$ is a relatively B-bounded operator with relative bound zero. Some remarks about the possibility of treating other functions f are offered at the end of this section.

4. Condition (iv) says that $X(t)$ has mean zero and unit variance.
5. The requirement, (v), that $u(x) = x$ is in the quadratic form domain of B ensures that sample paths of the process are not too rough. Specifically, it controls the high-frequency asymptotics of the power spectrum (see below). Exactly how this enters our analysis is described at the end of Subsect. 3.1.
6. The last condition, (vi), has been chosen to obviate functional analytic minutiae in the definition of L_λ in (14) below. Since λ will later be chosen small, merely assuming x is relatively B-bounded (with any bound) would permit the same elementary analysis.

    The most basic example is the following: take $\mathcal{S} = \{-1, 1\}$ and

$$B = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

in the natural basis for $\ell^2(\mathcal{S})$. With these definitions, $X(t)$ flips between $+1, -1$ with exponential waiting times.

    The Ornstein-Uhlenbeck process also obeys Hypothesis 1. Here $\mathcal{S} = \mathbf{R}$ and

$$B = -\frac{\partial^2}{\partial x^2} + x\frac{\partial}{\partial x}.$$

The stationary distribution is $d\mu = (2\pi)^{-1/2} \exp\{-x^2/2\}\, dx$, the standardized Gaussian. The eigenvalues are $\kappa_N = N$ with corresponding eigenvectors $u_N(x) = (N!)^{-1/2} \mathrm{He}_N(x)$, the L^2-normalized Hermite polynomials (our notation is that of Abramowitz and Stegun [1]). The three term recurrence for these polynomials,

$$x\, u_N(x) = \sqrt{N}\, u_{N-1}(x) + \sqrt{N+1}\, u_{N+1}(x), \tag{4}$$

shows that x is indeed relatively bounded with bound zero.

Theorem. *Let $X(t)$ be a process obeying Hypothesis 1 and let $V(\theta)$ be an even trigonometric polynomial. If the coupling constant λ is sufficiently small then for any initial wave-function $\psi_0 \in H^s$,*

$$\frac{1}{T}\int_0^\infty e^{-t/T}\, \mathbb{E}\left\{\|\psi(\cdot, t)\|_{H^s}^2\right\} dt \asymp \lambda^s T^{s/2}\|\psi_0\|_{L^2}^2 + \|\psi_0\|_{H^s}^2 \tag{5}$$

as $T \to \infty$. The expectation is over possible trajectories of the Markov process $X(t)$. In particular, the energy grows on average as the square-root of time.

Remarks. 1. We write $x \asymp y$ if and only if $x \lesssim y$ and $y \lesssim x$. The notation $x \lesssim y$ means that there exists $C > 0$ so that $x \leq Cy$.
2. Since adding a constant (even if time-dependent) to the Hamiltonian has no physical effect – it just changes the phase of the wave function – we may assume that $\int V(\theta)\, d\theta = 0$.
3. In the interest of clarity, we will not prove the theorem in this generality, but rather in the special case that $V(\theta) = \cos(\theta)$ and $X(t)$ is the Ornstein–Uhlenbeck process. This simplifying assumption will be invoked at the beginning of Sect. 3, where the computational part of the proof begins. Some further remarks on the general case are given at the end of Subsect. 3.1.

4. Extending the method we present to arbitrary (i.e., not necessarily even) trigonometric polynomials, V, however, requires more than just computational stamina; in this setting, the operator H introduced in Sect. 3 is no longer a finite difference operator.

For heuristic reasons that we will describe in a moment, it is natural to believe that the rate of energy growth is determined by the power spectrum of the forcing process. The power spectrum is the non-random function

$$S(\omega) = \lim_{T \to \infty} \frac{1}{4T} \Big| \int_{-T}^{T} e^{i\omega t} X(t) \, dt \Big|^2 \quad \omega \in \mathbf{R}. \tag{6}$$

The limit is in the following sense: for any continuous function ϕ of compact support

$$\frac{1}{4T} \int \Big| \int_{-T}^{T} e^{i\omega t} X(t) \, dt \Big|^2 \phi(\omega) \, d\omega \to \int \phi(\omega) S(\omega) \, d\omega$$

almost surely [9, §X·7].

By the Wiener–Khinchin Theorem, S is the Fourier transform of the two-time correlation function $\mathbb{E}\big(X(t)X(t+\tau)\big)$ so it is easy to show that

$$S(\omega) = \sum_{N} |\hat{x}_N|^2 \frac{\kappa_N}{\omega^2 + \kappa_N^2}, \tag{7}$$

where $\hat{x}_N$ are the Fourier coefficients of x with respect to the B eigenfunctions: $\hat{x}_N = \int x u_N(x) \, d\mu$. This means that $S(\omega) \asymp |\omega|^{-2}$ as $|\omega| \to \infty$. (In fact, $\omega^2 S(\omega) \to \langle x|B|x \rangle$ which is finite by (v) of Hypothesis 1.)

At small coupling, λ, significant energy growth can only occur as the result of a cumulative effect over a long time period. A particle of unit mass and momentum p experiences a cosine potential as being time-dependent with typical angular frequency $\omega = p$. Now if the potential has a time dependent coupling constant given by a process $X(t)$, only the part of that process with characteristic frequency ω will have a cumulative effect – the effect of the parts at other frequencies will average out to zero, at leading order. In this way, it is natural to intuit that the rate at which the particle gains energy is proportional to the amount of power in the process at that frequency. As energy is proportional to momentum squared we expect

$$\frac{d}{dt} p^2 \propto S(p) \asymp p^{-2}$$

which implies $\dot{p} \asymp p^{-3}$. This leads to $p \asymp t^{1/4}$ and so the prediction that energy should grow as the square-root of time.

More generally, the argument of the preceding paragraph suggests that processes with $S(\omega) \asymp |\omega|^{-\gamma}$ give rise to $p \asymp t^{1/(2+\gamma)}$. T. Spencer informed us that V. Zakharov predicted such behaviour during a private discussion of this problem. We are not privy to his reasoning.

Bourgain's random model [4, Part II] is far from being stationary. In essence, his model may be described by $V(\theta) = \cos(\theta)$ and $X(t)$ by $\sum g_j \gamma_j(t)$, where g_j are normalized Gaussian random variables and γ_j are disjointly supported bump functions. Although the functions γ_j are all of approximately unit norm in H^s, they are not evenly spaced. Indeed, for $s = 1$, $\mathrm{supp}(\gamma_j)$ is approximately $[j^{4/5}, (j+1)^{4/5}]$. This means that

$X(t)$ is not uniformly locally H^s. Consequently, it doesn't offer a very clear view, to our eyes at least, of the role of temporal smoothness in the rate of energy growth.

However, it is possible to fit Bourgain's result into the heuristic described above by introducing a local power spectrum

$$S(x, t) = \mathbb{E} \left| \int X(t + \tau) e^{i\omega\tau} \phi(\tau) \, d\tau \right|^2 ,$$

where ϕ is positive, C^∞, supported by $[-2, 2]$, and equal to 1 on $[-1, 1]$. For the model under discussion, one finds that there are $\asymp t^{1/(2+2s)}$ many γ_j's supported in a unit neighbourhood of t and that each $|\hat{\gamma}(\omega)|^2$ is of size $t^{-(2s+1)/(2s+2)}$ over the interval $[-t^{1/(2s+2)}, +t^{1/(2s+2)}]$ and essentially zero outside this interval. By modifying the above argument to say $p\dot{p} \propto S(p, t)$, we find that $p\dot{p} \asymp t^{-s/(s+1)}$ and so $\|\psi'\| \asymp t^{1/(2s+2)}$ as proved by Bourgain.

Earlier we mentioned that the model under consideration may be interpreted as a rigid rotator coupled to a classical heat bath. We should say that for the case we treat, the bath is at an infinite temperature and so it is reasonable to expect that the energy increases indefinitely. The power spectrum of a finite temperature bath would have exponential tails. (For example, the well known Planck law: $S(\omega) \propto \omega^3/(e^{\omega/T} - 1)$, where T denotes the temperature.) This leads us to ask what would happen in the presence of a finite temperature heat bath.

A finite temperature bath may be modelled by a stationary mixing process all of whose sample paths are analytic in the strip $\mathbf{R} + i[-\alpha, \alpha]$. (For example, $X(t) = \int Y(t + s)(s^2 + 1)^{-1} \, ds$, where $Y(t)$ is the two-state process described above.) Naive physical intuition suggests that thermal equilibrium might be achieved and in particular, that while there may be rare excursions to high energy the energy should remain bounded on average. We believe that this is not the case and that in fact, the energy tends to infinity with probability one. This is what the heuristic given above predicts.

The paper is arranged as follows: in the next section, we show that the combined evolution of the wave-function and the stochastic process is governed by a semigroup and derive an explicit expression for its generator, L_λ. It is further shown that the time averaged quantity discussed in the theorem admits a simple expression in terms of the resolvent of this generator.

In Sect. 3 we study the properties of the resolvent of L_λ and, in particular, show that to leading order in λ it is given by the resolvent of a certain finite difference operator, H. It turns out that H is the generator of a continuous time random walk and so, at leading order in λ, the evolution may be described as a random walk in momentum space (cf. the first remark after Lemma 3.4). Note however that the hopping rate decays as the reciprocal of momentum squared.

The main technical parts of the paper are the study of the long-term properties of the random walk with generator H and the derivation of estimates which permit us to control the discrepancy between this and the true evolution, which is generated by L_λ. The method we employ to achieve the former goal contains elements which appear to be new and which may be of interest to those studying random walks or diffusion.

The division of this technical material between the final two sections, 4 and 5, is not by "random walk" vs. "discrepancy estimates" but rather by the nature of the estimates involved. This is more natural in terms of the proofs.

Section 4 derives weighted-norm inequalities for the resolvent of H (Corollary 4.3) and for the difference between this and the resolvent of L_λ (Corollary 4.4). Section 5 uses

these norm estimates to obtain the pointwise estimates necessary to prove the theorem. The section closes by restating the theorem and finishing its proof.

The material of Sect. 2 and of the beginning of Sect. 3 is essentially standard fare. However, it particularly parallels the work of Tcheremchantsev [12–14] since, in a crude sense, his model resembles the Fourier transform of ours.

The question we address is by no means the only interesting problem of its type. It is natural to ask to what extent the heuristic described above gives the correct behaviour and, in particular, whether Zakharov's prediction is true. While it is presumably impossible to give lower bounds for energy growth purely in terms of the smoothness properties of $V(x, t)$, one might hope that such deterministic upper bounds are possible.

It seems that the methods presented here could be extended to cover certain processes with rougher time dependence; see below. Treating processes with smoother time dependence would necessitate giving up the assumption of the existence of a spectral gap. This is a rather daunting proposition. For the analogous problem on $\mathbf{R}^n/\mathbf{Z}^n$, $n \geq 2$, the operator corresponding to H is rather complicated. (Even its domain, the kernel of L_0, is not simple; it depends on the number of ways an integer can be represented as a sum of n squares.) Studying the behaviour of its Green function would require a better understanding of its structure than we currently possess. Even more difficult is the corresponding problem in $\mathbf{R}^n$. In the case that the potential is supported inside a compact time-independent interval, Bourgain [6] has shown that as a result of dispersion, the energy can be bounded by t^ϵ, $\epsilon > 0$ no matter how rough the time dependence of the potential.

We now wish to offer a few remarks about the possibility of treating processes of the form $f(X(t))$ when f is rather rough. Please note that these remarks are somewhat conjectural. Pursuing them would necessitate a considerable enterprise, one that we do not intend to undertake in the foreseeable future.

Let $X(t)$ be the Ornstein-Uhlenbeck process and consider $Y(t) = f(X(t))$ for some measurable function $f : \mathbf{R} \to \mathbf{R}$ that is not too large at infinity, say $f \in L^\infty(d\mu)$. It is not difficult to show that the power spectrum of $Y(t)$ is given by

$$S(\omega) = \sum_{N=0}^{\infty} |\hat{f}(N)|^2 \frac{N}{N^2 + \omega^2},$$

where $\hat{f}(N) = \langle u_N | f \rangle$ (cf. (7)). If f is in the quadratic form domain of B then $\sum N |\hat{f}(N)|^2 < \infty$ and $S(\omega) \asymp \omega^{-2}$. As promised in Remark 3 after Hypothesis 1, the analysis doesn't change.

If f is rough then it will not belong to the quadratic form domain of B (a second-order elliptic differential operator) and so the sum above will not converge. In the simplest case, $V(\theta) = \cos(\theta)$, this affects the analysis in the following way: the off-diagonal coefficients of the operator H are given by

$$a_n = 2\lambda^2 \sum_{N \geq 1} \frac{(N + \beta)|\hat{f}(N)|^2}{(N + \beta)^2 + (2n + 1)^2} \sim 2\lambda^2 S(2n + 1)$$

(cf. (40)). If $\hat{f}(N) \asymp N^{-\alpha}$ with $\alpha \in (1/2, 1)$, then $S(\omega) \asymp \omega^{-2\alpha}$ and so $a_n \asymp n^{-2\alpha}$. This means that the large-n hopping rate is enhanced relative to the case where f is in the quadratic form domain of B (in that case, $a_n \asymp n^{-2}$). Proceeding formally, we infer that the weight ρ in Sects. 4 and 5 should be changed to $c \min\{n, \sqrt{\beta}\lambda^{-1}n^{1+\alpha}\}$ and so ultimately that, on average, the kinetic energy grows as t^σ with $\sigma = 1/(1 + \alpha)$.

This is exactly the exponent one would predict from the power spectrum asymptotics by the heuristic given above.

2. Reformulation

Just as classical probability distributions are described by normalized positive measures, i.e., normalized positive linear functionals on random variables, so quantum mechanical distributions are described by normalized positive linear functionals on the space of observables (bounded linear operators). These are called density matrices; they are the positive trace-class operators and are normalized to have trace equal to one. The expected value of an observable A is given by $\mathrm{tr}(A\rho)$. For example, a system in the quantum state $|\psi\rangle$ is described by $\rho = |\psi\rangle\langle\psi|$, the projection onto the linear space spanned by $|\psi\rangle$. More generally, a system in states $|\psi_i\rangle$ with probabilities p_i is described by $\rho = \sum_i p_i|\psi_i\rangle\langle\psi_i|$. Notice that the expected value of an observable A is given by $\mathrm{tr}(A\rho) = \sum_i p_i\langle\psi_i|A|\psi_i\rangle$. The only natural choice!

As is usual in this business, we consider the space of density matrices as a subset of the space of Hilbert–Schmidt operators. This affords us the pleasure of working in a Hilbert Space (the inner product is given by $\langle\rho|\sigma\rangle = \mathrm{tr}(\rho^\dagger\sigma)$). We will denote the space of Hilbert–Schmidt operators on $L^2(\mathbf{R}/2\pi\mathbf{Z})$ by $\mathfrak{I}_2\big(L^2(\mathbf{R}/2\pi\mathbf{Z})\big)$ or, more often, $\mathfrak{I}_2$ for short.

The state of the Markov process X is described by its probability density $u(x)$, a non-negative function in $L^2(\mathcal{S}; d\mu)$ with $\int u(x)\, d\mu = 1$. While the quantum system is described by its density matrix ρ, a non-negative operator in $\mathfrak{I}_2$ with $\mathrm{tr}(\rho) = 1$ (in particular, $\rho \in \mathfrak{I}_1$). The state of the combined system of the Markov process and the quantum particle at any time t is described by an element of the Hilbert space $L^2(\mathcal{S}; d\mu)\otimes\mathfrak{I}_2$. We will denote this element by $\mathsf{P}(t)$. For example, when the process and quantum system are independent of one another, $\mathsf{P} = u \otimes \rho$. The most general P is the limit of convex combinations of tensor products.

A particular case is when the quantum system is in an initial state ψ_0 and the process X is in its stationary distribution. This is represented by

$$\mathsf{P}(t = 0) = u_0 \otimes \rho_0\,, \quad \text{where } \rho_0 = |\psi_0\rangle\langle\psi_0| \text{ and } u_0(x) \equiv 1.$$

Note that at positive times, P need no longer be a tensor product. This is because the states of quantum particle and the forcing process become correlated.

To help explain the meaning of P we now describe how to calculate expectations.

A natural class of random variables/observables is $f(X)A$ for some $f \in L^2(d\mu)$ and $A \in \mathfrak{I}_2$. In the independent case, $\mathsf{P} = u \otimes \rho$, it is clear that the average value of $f(X)A$ is given by

$$\int f(x)u(x)\, d\mu(x) \cdot \mathrm{tr}\left(A\rho\right) = \big\langle \bar{f} \otimes A^\dagger \big| \mathsf{P}\big\rangle_{L^2(d\mu)\otimes\mathfrak{I}_2}.$$

Linearity then forces the same choice for general P, namely

$$\mathbb{E}\left\{f(X(t))\langle\psi(t)|A|\psi(t)\rangle\right\} = \big\langle \bar{f} \otimes A^\dagger \big| P(t)\big\rangle_{L^2(\mu)\otimes\mathfrak{I}_2}. \tag{8}$$

Similarly, the average value of more complicated observables such as $\sum_j f_j(X)A_j$ is determined by linearity.

Of course, we are mainly interested in observables that are not Hilbert–Schmidt; for example, the kinetic energy. These can be dealt with using a simple approximation argument.

To determine the evolution equation for P, we perform the following calculation: let $f \in L^2(d\mu)$ be from the domain of B and let $A \in \mathfrak{I}_2$ be such that $[H_0, A] \in \mathfrak{I}_2$, then

$$
\begin{aligned}
\frac{d}{dt} &\, \mathbb{E}\left\{ f(X(t)) \langle \psi(t) | A | \psi(t) \rangle \right\} \\
&= \mathbb{E}\left\{ -[Bf](X(t)) \langle \psi | A | \psi \rangle + f(X(t)) \langle \dot{\psi} | A | \psi \rangle + f(X(t)) \langle \psi | A | \dot{\psi} \rangle \right\} \\
&= \mathbb{E}\left\{ -[Bf](X(t)) \langle \psi | A | \psi \rangle + i f(X(t)) \langle \psi | (H_0 + \lambda X(t) V) A | \psi \rangle \right. \\
&\qquad \left. - i f(X(t)) \langle \psi | A(H_0 + \lambda X(t) V) | \psi \rangle \right\} \\
&= \mathbb{E}\left\{ -[Bf](X(t)) \langle \psi | A | \psi \rangle + i f(X(t)) \langle \psi | [H_0, A] | \psi \rangle \right. \\
&\qquad \left. + i \lambda X(t) f(X(t)) \langle \psi | [V, A] | \psi \rangle \right\},
\end{aligned}
\tag{9}
$$

where we have used a dot to denote the time derivative, H_0 to denote the free Hamiltonian (the Laplacian, $-\frac{d^2}{d\theta^2}$) and V represents the operator of multiplication by the spatial part of the potential, $V(\theta) = \cos(\theta)$. Also, $[A, C] = AC - CA$ denotes the commutator of the operators A, C. In the first equation we used the fact that $\partial_t \mathbb{E}\{ f(X(t)) \} = -\mathbb{E}\{ [Bf](X(t)) \}$. Strictly speaking, this should be $B^\dagger$ since

$$
\begin{aligned}
\frac{d}{dt} \int f(x) u(x, t) \, d\mu(x) &= \int f(x) \frac{\partial u}{\partial t}(x, t) \, d\mu(x) \\
&= -\int f(x) [Bu](x, t) \, d\mu(x) \\
&= -\int [B^\dagger f](x) u(x, t) \, d\mu(x).
\end{aligned}
\tag{10}
$$

However, we assumed B to be self-adjoint so there is no need to distinguish.

Equation (9) shows that the time derivative of the average value of an observable $f(X)A$ is given by the average value of another observable,

$$
-[Bf](X)A + i f(X)[H_0, A] + i\lambda X f(X)[V, A].
$$

In order that we obtain the same relation when calculating averages with P, as described in (8), we must have

$$
\begin{aligned}
\left\langle \bar{f} \otimes A^\dagger \left| \frac{d}{dt} \mathsf{P} \right. \right\rangle &= \left\langle -(B\bar{f}) \otimes A^\dagger - i \bar{f} \otimes [H_0, A]^\dagger - i\lambda (x\bar{f}) \otimes [V, A]^\dagger \, \middle| \, \mathsf{P} \right\rangle \\
&= \left\langle -(B\bar{f}) \otimes A^\dagger + i \bar{f} \otimes [H_0, A^\dagger] + i\lambda (x\bar{f}) \otimes [V, A^\dagger] \, \middle| \, \mathsf{P} \right\rangle.
\end{aligned}
\tag{11}
$$

Taking adjoints in $L^2(d\mu) \otimes \mathfrak{I}_2$ and then choosing $f \otimes A$ from a dense subset of this space, we find that

$$
\frac{d}{dt} \mathsf{P} = -\left\{ B \otimes I + i I \otimes [H_0, \cdot] + i\lambda x \otimes [V, \cdot] \right\} \mathsf{P},
\tag{12}
$$

where x denotes the operator $u(x) \mapsto x u(x)$ in $L^2(d\mu)$. In passing from (11) to (12) we used the following observation: For any operator C on $L^2(\mathbf{R}/2\pi\mathbf{Z})$ the adjoint of map

$\rho \mapsto [C, \rho]$ on $\mathfrak{I}_2$ is given by $\rho \mapsto [C^\dagger, \rho]$. The demonstration of this is simple:

$$\langle \sigma \,|[C, \rho]\rangle = \mathrm{tr}\{\sigma^\dagger(C\rho - \rho C)\} = \mathrm{tr}\{(\sigma^\dagger C - C\sigma^\dagger)\rho\} = \langle [C^\dagger, \sigma]\,|\rho\rangle. \qquad (13)$$

In particular, if C is self-adjoint on L^2 then $[C, \cdot]$ is self-adjoint on $\mathfrak{I}_2$.

A more general derivation of the master equation, (12), for a quantum system subjected to Markovian forcing is given in [11]. It is natural to call this a master equation from its similarity to the purely classical and purely quantum mechanical equations of the same name. For a highly readable account of both classical and quantum mechanical stochastic processes with a strong physical motivation, we recommend [10].

To simplify notation, we rewrite the evolution of P as

$$\frac{d}{dt}\mathsf{P} = -L_\lambda \mathsf{P} \quad \text{where} \quad L_\lambda = B \otimes I + iI \otimes [H_0, \cdot] + i\lambda x \otimes [V, \cdot]. \qquad (14)$$

The linear operator L_λ so defined is m-accretive and has domain $D(B) \otimes H^2$. In fact, it is the sum of a non-negative operator, $B \otimes I$, and a skew-adjoint ($A^\dagger = -A$) operator,

$$iI \otimes [H_0, \cdot] + i\lambda x \otimes [V, \cdot].$$

Maximality follows from (vi) of Hypothesis 1. This said, we see that L_λ generates a contraction semigroup so we can write

$$\mathsf{P}(t) = e^{-tL_\lambda}\mathsf{P}(0).$$

Combining the equation above with (8) we infer that for any Hilbert–Schmidt observable A,

$$\frac{1}{T}\int_0^\infty e^{-t/T} \,\mathbb{E}\left\{\langle \psi(t)|A|\psi(t)\rangle\right\} dt$$
$$= \frac{1}{T}\int_0^\infty e^{-t/T}\langle u_0 \otimes A^\dagger\,|\mathsf{P}(t)\rangle dt$$
$$= \frac{1}{T}\int_0^\infty e^{-t/T}\langle u_0 \otimes A^\dagger\,|\,e^{-tL_\lambda}\,|\mathsf{P}(0)\rangle dt$$
$$= \left\langle u_0 \otimes A^\dagger\,\Big|\,\frac{\beta}{L_\lambda + \beta}\,\Big|\mathsf{P}(0)\right\rangle, \qquad (15)$$

where $\beta = \frac{1}{T}$. (Accretiveness ensures that $L_\lambda + \beta$ is invertible.) This trick, which is standard, is behind our choice of time averaging.

As it diagonalizes the Laplacian, it is only natural that we choose to work in a Fourier basis. It will also simplify the study of the H^s norms. Because $V(\theta)$ is assumed even, the Hamiltonian is symmetric under the reflection $\theta \mapsto -\theta$. Therefore, the natural choice of orthonormal basis for $L^2(\mathbf{R}/2\pi\mathbf{Z})$ is

$$|n\rangle = \begin{cases} (2\pi)^{-1/2} & : n = 0 \\ \pi^{-1/2}\cos(n\theta) & : n = 1, 2, \ldots \\ \pi^{-1/2}\sin(|n|\theta) & : n = -1, -2, \ldots \end{cases} \qquad (16)$$

The reflection symmetry of the Hamiltonian actually means that the odd and even subspaces are invariant under the evolution. This is at the origin of the invariant subspaces described in Lemma 3.1 below.

The natural basis for B is in terms of its eigenfunctions u_N, $N = 0, 1, \ldots$. Therefore, we introduce the following orthonormal basis for the Hilbert space $L^2(\mu) \otimes \mathfrak{I}_2$:

$$|N, n, m\rangle = u_N(x) \otimes |n\rangle\langle m| \qquad \text{for } n, m \in \mathbf{Z} \text{ and } N = 0, 1, \ldots \qquad (17)$$

with $|n\rangle$ and $\langle m|$ defined as in (16).

With this notation, we can now reformulate the behaviour of H^s norms in terms of the operator L_λ:

Proposition 2.1. *For a quantum particle initially in the state $\psi(t = 0) = \psi_0 \in H^s$ and for $X(t = 0)$ chosen independently according to the stationary distribution,*

$$\beta \int_0^\infty e^{-\beta t} \, \mathbb{E} \, \|\psi(t)\|_{H^s}^2 \, dt = \sum_n (1 + |n|^{2s}) \Big\langle 0, n, n \Big| \frac{\beta}{L_\lambda + \beta} \Big| \mathsf{P}(0) \Big\rangle, \qquad (18)$$

where $\mathsf{P}(0) = u_0 \otimes |\psi_0\rangle\langle\psi_0|$. Note that β plays the role of $\frac{1}{T}$ in the theorem.

Proof. This follows immediately from (15): For $N > 0$, let

$$A = \sum_{n=-N}^{N} (1 + |n|^{2s}) \, |0, n, n\rangle\langle 0, n, n|$$

and then use the monotone convergence theorem to take $N \to \infty$. $\quad\square$

3. The Operator L_λ

In the previous section, we reduced the study of the average behaviour of H^s norms to the consideration of the resolvent of the operator L_λ. The purpose of this section is to isolate the dominant terms in this resolvent; the main technical estimates appear in the next two sections.

For expository reasons, we make the following:

Hypothesis 2. *For the remainder of this paper we shall consider only the case where $V(\theta) = \cos(\theta)$ and $X(t)$ is the Ornstein–Uhlenbeck process.*

Some remarks about the general case are given after Lemma 3.4.

To a first approximation, we will treat L_λ as a perturbation of L_0, (that is, L_λ with $\lambda = 0$). For this reason, we introduce the notation

$$J = \tfrac{1}{i\lambda}(L_\lambda - L_0) = x \otimes [V, \cdot] \qquad \text{which implies} \qquad L_\lambda = L_0 + i\lambda J. \qquad (19)$$

Notice that J is a self-adjoint operator. For convenience, we will always take $\lambda \geq 0$.

Lemma 3.1. *The basis $|N, n, m\rangle$ of (17) diagonalizes the operator L_0. Specifically,*

$$L_0|N, n, m\rangle = \big\{ B \otimes I + i I \otimes [H_0, \cdot] \big\} |N, n, m\rangle$$
$$= \big\{ N + i(n^2 - m^2) \big\} |N, n, m\rangle. \qquad (20)$$

Also, each of the four subspaces

$$\text{span}\{|N, n, m\rangle : N \geq 0, \ n, m \geq 0\}, \quad \text{span}\{|N, n, m\rangle : N \geq 0, \ n \geq 0, m < 0\},$$
$$\text{span}\{|N, n, m\rangle : N \geq 0, \ n, m < 0\}, \quad \text{span}\{|N, n, m\rangle : N \geq 0, \ n < 0, m \geq 0\}$$

is invariant under J and hence also under $L_\lambda = L_0 + i\lambda J$.

Proof. The first part is self-evident (recall $\kappa_N = N$ for Ornstein-Uhlenbeck). The second is a reflection of the fact that because $V(\theta)$ is even, $V|n\rangle$ and $|n\rangle$ have the same parity: even if $n \geq 0$ and odd if $n < 0$. $\quad \square$

As we are interested in H^s norms, Proposition 2.1 shows that we need only consider how L_λ acts in two of these invariant subspaces, namely span$\{|N, n, m\rangle : N \geq 0, n, m \geq 0\}$, which corresponds to the even part of the wave function, and span$\{|N, n, m\rangle : N \geq 0, n, m < 0\}$, which corresponds to the odd part.

In calculating $(L_\lambda + \beta)^{-1}$ in each of these subspaces, a major role is played by the kernel of L_0. This is given by span$\{|0, n, n\rangle : n \in \mathbf{Z}\}$. To isolate this part we define projections

$$
\begin{aligned}
P & \quad \text{onto} \quad \text{span}\{|0, n, n\rangle : n \geq 0\}, \\
P^\perp & \quad \text{onto} \quad \text{span}\{|N, n, m\rangle : n, m \geq 0 \text{ and either } N > 0 \text{ or } n \neq m\}, \\
Q & \quad \text{onto} \quad \text{span}\{|0, n, n\rangle : n < 0\}, \text{ and} \\
Q^\perp & \quad \text{onto} \quad \text{span}\{|N, n, m\rangle : n, m < 0 \text{ and either } N > 0 \text{ or } n \neq m\}.
\end{aligned}
$$

Notice that $PP^\perp = 0$ and $P + P^\perp$ is the projection onto span$\{|N, n, m\rangle : N \geq 0, n, m \geq 0\}$, the invariant subspace associated with the even part of the wave function. Similarly, $QQ^\perp = 0$ and $Q + Q^\perp$ projects onto span$\{|N, n, m\rangle : N \geq 0, n, m < 0\}$. Further note that the ranges of both P and Q lie within the kernel of L_0.

It is now easier to treat the two parts of L_λ separately. We begin with the "even" part that is the part invariant under $P + P^\perp$. This is followed by a parallel but abbreviated discussion of the "odd" part.

3.1. Even. On the range of $P + P^\perp$, we can write L_λ in block form

$$
\begin{bmatrix} 0 & i\lambda P J P^\perp \\ i\lambda P^\perp J P & P^\perp L_\lambda P^\perp \end{bmatrix}, \tag{21}
$$

where we noted that since $\int V(\theta)\, d\theta = 0$, $PJP = 0$. (This may also be derived from the fact that $\int x\, d\mu = 0$.) Of course, $L_0 P$ is also zero since P is a projection onto a subset of the kernel of L_0.

By the well-known formulae for block inversion,

$$
P(L_\lambda + \beta)^{-1} P = \left\{ \beta + \lambda^2 PJP^\perp R_\lambda^\perp(\beta) P^\perp JP \right\}^{-1}, \tag{22}
$$

where $R_\lambda^\perp(\beta) = (P^\perp L_\lambda P^\perp + \beta)^{-1}$ is the resolvent of the operator appearing in the bottom right corner. Here, as elsewhere in this section, we regard the operator inside the braces on the right-hand side of (22) as an operator acting on the range of P for the purposes of inversion. Similarly, R_λ is the inverse of an operator acting on the range of $P^\perp$.

The formulae for block inversion also show that

$$
P(L_\lambda + \beta)^{-1} P^\perp = -i\lambda \left\{ \beta + \lambda^2 PJP^\perp R_\lambda^\perp(\beta) P^\perp JP \right\}^{-1} PJP^\perp R_\lambda^\perp(\beta). \tag{23}
$$

Note the occurrence of our earlier expression (22) as a factor here. Indeed, control of (23) will be a simple by-product (see Lemma 5.4) of the analysis of (22) to which we now direct our attention. (We do not need the other two matrix entries of $(L_\lambda + \beta)^{-1}$, that

is, those mapping onto the range of $P^\perp$: These are irrelevant to (18) of Proposition 2.1 because $|0, n, n\rangle$ is orthogonal to the range $P^\perp$.)

As we will eventually demonstrate, one can control (22) by replacing $R_\lambda^\perp(\beta)$ by $R_0^\perp(\beta)$ and then treating the error as a small perturbation. With this as our inspiration, we define $H : \text{Range}(P) \to \text{Range}(P)$ by

$$H = \lambda^2 P J P^\perp R_0^\perp(\beta) P^\perp J P. \tag{24}$$

(Notice that H depends on both λ and β.) Equation (22) can now be written as

$$P(L_\lambda + \beta)^{-1} P$$
$$= \left\{ H + \beta + \lambda^2 P J P^\perp \left[R_\lambda^\perp(\beta) - R_0^\perp(\beta) \right] P^\perp J P \right\}^{-1} \tag{25}$$
$$= \sum_{j=0}^\infty (H + \beta)^{-1} \left\{ -\lambda^2 P J P^\perp \left[R_\lambda^\perp(\beta) - R_0^\perp(\beta) \right] P^\perp J P (H + \beta)^{-1} \right\}^j. \tag{26}$$

The convergence of this infinite resolvent series will not be justified until we derive certain estimates in the next section; we merely offer it as a sign-post of where we are headed.

In addition to H, the next two sections employ the operator

$$D = \lambda R_0^\perp(\beta) P^\perp J P \tag{27}$$

which is to be regarded as mapping $\text{Range}(P) \to \text{Range}(P^\perp)$. The importance of this operator will become clearer as we proceed.

We now give a trio of lemmas. The first describes the behaviour of J on the range of $P + P^\perp$, which it leaves invariant. The others give explicit formulas for the way the operators D and H act on their domain, that is, on the range of P.

Lemma 3.2. *For any $N, n, m \geq 0$,*

$$J|N, n, m\rangle = \sqrt{N+1} \Big\{ c_n|N+1, n+1, m\rangle + c_{n-1}|N+1, n-1, m\rangle$$
$$- c_m|N+1, n, m+1\rangle - c_{m-1}|N+1, n, m-1\rangle \Big\}$$
$$+ \sqrt{N} \Big\{ c_n|N, n+1, m\rangle + c_{n-1}|N, n+1, m\rangle$$
$$- c_m|N, n, m+1\rangle - c_{m-1}|N, n, m-1\rangle \Big\}, \tag{28}$$

where the co-efficients c_k are given by

$$c_k = \begin{cases} 0 & : k = -1 \\ \frac{1}{\sqrt{2}} & : k = 0 \\ \frac{1}{2} & : k > 0 \end{cases}.$$

Proof. Recall that $J = x \otimes [V, \cdot]$. The proof is a mundane calculation based on the following two formulae: First, as was stated above, (4),

$$x u_N(x) = \sqrt{N} u_{N-1}(x) + \sqrt{N+1} u_{N+1}(x),$$

and second,

$$\cos(\theta)|n\rangle = c_n|n+1\rangle + c_{n-1}|n-1\rangle = \begin{cases} \frac{1}{\sqrt{2}}|1\rangle & : n = 0 \\ \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{2}|2\rangle & : n = 1 \\ \frac{1}{2}|n+1\rangle + \frac{1}{2}|n-1\rangle & : n \geq 2 \end{cases}$$

in the ket notation of (16).   □

Lemma 3.3. *The operator D acts like a vector-valued first order finite difference operator specifically, one with two components: if $|\psi\rangle = \sum \psi_n|0, n, n\rangle$ then*

$$\langle 1, n+1, n|D|\psi\rangle = \alpha_n[\psi_n - \psi_{n+1}], \tag{29}$$

$$\langle 1, n, n+1|D|\psi\rangle = \bar{\alpha}_n[\psi_{n+1} - \psi_n], \tag{30}$$

where

$$\alpha_n = \frac{\lambda}{2} \frac{1}{1 + \beta + i(2n+1)} \begin{cases} \sqrt{2} & : n = 0 \\ 1 & : n > 0 \end{cases} \tag{31}$$

and all other co-efficients $\langle N, n, m|D|\psi\rangle$ are zero.

Proof. This follows from the previous lemma plus the fact that

$$R_0^{\perp}(\beta)|N, n, m\rangle = \frac{1}{N + \beta + i(n^2 - m^2)}|N, n, m\rangle \quad \square$$

Lemma 3.4. *The operator H is a discrete second order difference operator in divergence form. Indeed, $H = (1 + \beta)D^{\dagger}D$ or, more explicitly, if $|\psi\rangle = \sum \psi_n|0, n, n\rangle$ then*

$$\langle 0, n, n|H|\psi\rangle = \begin{cases} a_0\psi_0 - a_0\psi_1 & n = 0 \\ (a_n + a_{n-1})\psi_n - a_{n-1}\psi_{n-1} - a_n\psi_{n+1} & n \geq 1 \end{cases},$$

where $a_n = 2(1 + \beta)|\alpha_n|^2$.

Proof. Define an operator S by $S|N, n, m\rangle = |N, m, n\rangle$; that is, S swaps n and m. From (28) one sees that $SJP = -JP$ and, by taking adjoints for example, $PJS = -PJ$ so

$$H = \lambda^2 PJP^{\perp}R_0^{\perp}(\beta)P^{\perp}JP \tag{32}$$

$$= \lambda^2 PJSP^{\perp}R_0^{\perp}(\beta)P^{\perp}SJP \tag{33}$$

$$= \lambda^2 PJP^{\perp}R_0^{\perp}(\beta)^{\dagger}P^{\perp}JP. \tag{34}$$

In the last line here we used $SR_0^{\perp}(\beta)S = R_0^{\perp}(\beta)^{\dagger}$ which is easily checked since both are diagonal in the $|N, n, m\rangle$ basis. Now we may write

$$H = \tfrac{1}{2}\lambda^2 PJP^{\perp}[R_0^{\perp}(\beta) + R_0^{\perp}(\beta)^{\dagger}]P^{\perp}JP \tag{35}$$

$$= (1 + \beta)\lambda^2 PJP^{\perp}R_0^{\perp}(\beta)^{\dagger}R_0^{\perp}(\beta)P^{\perp}JP \tag{36}$$

$$= (1 + \beta)D^{\dagger}D. \tag{37}$$

Equation (36) requires some explanation: from (28) we know that the range of $P^\perp J P$ is spanned by vectors of the form $|1, n, m\rangle$ and for such vectors we have

$$\tfrac{1}{2}\big[R_0^\perp(\beta) + R_0^\perp(\beta)^\dagger\big]|1, n, m\rangle = \frac{1 + \beta}{(1 + \beta)^2 + (n^2 - m^2)^2}|1, n, m\rangle \qquad (38)$$

$$= (1 + \beta)R_0^\perp(\beta)^\dagger R_0^\perp(\beta)|1, n, m\rangle. \qquad (39)$$

The explicit formula for H follows easily from $H = (1 + \beta)D^\dagger D$ and the formula for D given in the previous lemma. $\quad\square$

Remarks. 1. If we consider the matrix representation of H in the $|0, n, n\rangle$ basis, the off-diagonal entries are non-positive and the sum of entries in each column (and row) is zero. This makes H the generator of a continuous time Markov chain. Indeed, from the state $|0, n, n\rangle$ there are two possible transitions: to $|0, n+1, n+1\rangle$, which occurs at rate a_n, and to $|0, n-1, n-1\rangle$ with rate a_{n-1}. In this way one can interpret the Markov chain as a continuous time random walk in which the hopping rate decays as $n \to \infty$. The physical interpretation of this is that H, which is a good approximation to $PL_\lambda P$, represents a spatially inhomogeneous random walk in momentum space.
2. For more general V and X, the expression (28) for operator J becomes far more complicated since now both N and either n or m can change by an arbitrary amount. (Notice that in (28), they change by one.) Apart from rendering the explicit formulae unreadable, this generality creates no significant mathematical problems. In particular, the operator D is still a vector-valued first order finite difference operator (cf. Lemma 3.3). Actually, it is better to regard it as a sequence of such operators, $\{D_N : N > 0\}$ corresponding to the final values of N such that D, as defined by (27), is equal to $\sum_N D_N$. From (19),

$$P^\perp J P|0, n, n\rangle = \sum_{N>0, k\neq 0} \hat{x}_N c_{n,k}\big\{|N, n+k, n\rangle - |N, n, n+k\rangle\big\},$$

where $c_{n,k} = \langle n + k|V|n\rangle$ and $\hat{x}_N = \langle u_N|x|u_0\rangle = \int x u_N(x)\, d\mu$. This should be compared with the formulae of Lemma 3.2 where $c_{n,k}$ and $\hat{x}_N$ are only non-zero for two values of k and N respectively. Note also that $\hat{x}_0 = 0$ so there is no need for an $N = 0$ term in the sum. Similarly, $\int V(\theta)\, d\theta = 0$ implies $c_{n,0} = 0$.

Because V is a trigonometric polynomial, only values of k smaller than the degree of V contribute to the sum given above. In this way, each D_N can be regarded as a vector of $2\deg(V)$ finite difference operators:

$$\langle N', n+k, n|D_N|\psi\rangle = \delta_{N,N'}\alpha_{N,n,k}\big[\psi_n - \psi_{n+k}\big],$$
$$\langle N', n, n+k|D_N|\psi\rangle = \delta_{N,N'}\bar{\alpha}_{N,n,k}\big[\psi_{n+k} - \psi_n\big],$$

where $1 \leq k \leq \deg(V)$ and

$$\alpha_{N,n,k} = \frac{\lambda\hat{x}_N}{\kappa_N + \beta + i(2nk + k^2)}c_{n,k}.$$

All other matrix entries of D_N are zero. The derivation of these formulae uses the fact that for $n \geq k \geq 1$, $c_{n+k,-k} = c_{n,k}$.

3. In the general setting, $H = \sum_N (\kappa_N + \beta) D_N^\dagger D_N$. In this way, Lemma 3.4 continues to hold, but with

$$a_n = \sum_{k=1}^{\deg(V)} \sum_{N \geq 1} 2(\kappa_N + \beta)|\alpha_{N,n,k}|^2$$

$$= 2\lambda^2 \sum_{k=1}^{\deg(V)} |c_{n,k}|^2 \sum_{N \geq 1} \frac{(\kappa_N + \beta)|\hat{x}_N|^2}{(\kappa_N + \beta)^2 + (2nk + k^2)^2}. \tag{40}$$

This is how (v) of Hypothesis 1 enters our analysis: $a_n \asymp n^{-2}$ as $n \to \infty$ if and only if $\sum \kappa_N |\hat{x}_N|^2 < \infty$. That is, if and only if $u(x) = x$ belongs to the quadratic form domain of B.

3.2. Odd. For the odd portion of the wave-packet, we must study $Q(L_\lambda + \beta)^{-1} Q$. Equations (21)–(27) continue to hold if P and $P^\perp$ are replaced by Q and $Q^\perp$ respectively. In particular, this defines new operators D and H associated to the odd subspace. The exact form of these operators differs very slightly from those in the even subspace. Indeed, they are slightly simpler as the following replacement for Lemmas 3.3 and 3.4 demonstrates:

Lemma 3.5. *The operator D associated to the odd subspace acts as follows: if $|\psi\rangle = \sum_{n=1}^{\infty} \psi_n |0, -n, -n\rangle$ then*

$$\langle 1, -n-1, -n|D|\psi\rangle = \alpha_n \big[\psi_n - \psi_{n+1}\big], \tag{41}$$

$$\langle 1, -n, -n-1|D|\psi\rangle = \bar{\alpha}_n \big[\psi_{n+1} - \psi_n\big], \tag{42}$$

where

$$\alpha_n = \frac{\lambda}{2} \frac{1}{1 + \beta - i(2n+1)}, \tag{43}$$

and all other co-efficients $\langle N, n, m|D|\psi\rangle$ are zero. Further, $H = (1+\beta)D^\dagger D$ so

$$\langle 0, -n, -n|H|\psi\rangle = (a_n + a_{n-1})\psi_n - a_{n-1}\psi_{n-1} - a_n\psi_{n+1}, \tag{44}$$

where $a_n = 2(1 + \beta)|\alpha_n|^2$ for $n \geq 1$ and $a_0 = 0$.

4. Norm Estimates

This section is devoted to deriving weighted norm estimates involving the operators D and $H = (1+\beta)D^\dagger D$ associated to the even portion of the wave packet (see (24) and (27) in the last section). These are then used to obtain similar estimates for $P(L_\lambda + \beta)^{-1}P$.

The operators D and H associated to the odd portion of the wave so closely resemble those for the even portion, that the same estimates can be derived with only cosmetic changes to the proofs. We will not discuss these operators again in this section.

The derivation of the weighted norm inequalities follows the well-known scheme of Combes–Thomas [7] and Agmon [2]. In the section that follows this, we use these weighted norm results to obtain pointwise estimates on $(H + \beta)^{-1}$ and $P(L_\lambda + \beta)^{-1}P$. However, the standard reduction from norm-estimates to pointwise estimates is unsuitable in our situation; the incorrect prefactor it gives renders it useless. Part of our remedy

to this problem involves obtaining derivative estimates for the Green function. Indeed, as D acts essentially as a (vector-valued) differential operator (cf. Lemma 3.3), this is how we encourage the reader to interpret (50)–(52). We will discuss this further in the next section.

Given a weight function $w : \{0, 1, \dots\} \to (0, \infty)$ we consider two weighted Hilbert spaces: (i) Let ℓ_w^2 denote the range of P endowed with the norm

$$\left\| \sum_n \psi_n |0, n, n\rangle \right\|_{\ell_w^2}^2 = \sum_n w(n) |\psi_n|^2; \tag{45}$$

and (ii) let $\ell_{\tilde{w}}^2$ denote the span of $\{|N, n, m\rangle : N, n, m \geq 0\}$ together with the norm

$$\left\| \sum_{N,n,m} \psi_{N,n,m} |N, n, m\rangle \right\|_{\ell_{\tilde{w}}^2}^2 = \sum_{N,n,m} \tilde{w}(n, m) |\psi_{N,n,m}|^2, \tag{46}$$

where $\tilde{w}(n, m) = \sqrt{w(n)w(m)}$. Notice that by Lemma 3.3, $D : \ell_w^2 \to \ell_{\tilde{w}}^2$ when $w(n) \equiv 1$. In fact, this is true for a much broader class of weights.

The following lemma will be used in Corollary 4.4 to control the sum in (26).

Lemma 4.1. *Consider the weight function* $w(n) = e^{2\rho(n)}$ *with*

$$\rho(n) = c \min\{n, \lambda^{-1} n^2 \sqrt{\beta}\}$$

and $c > 0$ *a (small) constant. There is an operator* $E : \ell_{\tilde{w}}^2 \to \ell_{\tilde{w}}^2$ *such that*

$$-\lambda^2 P J P^\perp \big[R_\lambda^\perp(\beta) - R_0^\perp(\beta) \big] P^\perp J P = D^\dagger E D \tag{47}$$

and, for λ *sufficiently small (not depending on* β*),*

$$\| E \|_{\ell_{\tilde{w}}^2 \to \ell_{\tilde{w}}^2}^2 \lesssim \lambda,$$

where the implicit constant holds uniformly as $c, \lambda, \beta \to 0$.

Proof. From the resolvent formula, one finds

$$\big[R_\lambda^\perp(\beta) - R_0^\perp(\beta) \big] = \sum_{j=1}^\infty (-i\lambda)^j R_0^\perp(\beta) \big\{ P^\perp J P^\perp R_0^\perp(\beta) \big\}^j.$$

This means that (47) holds with

$$E = \sum_{j=0}^\infty (-i\lambda)^{j+1} \big[R_0^\perp(\beta)^\dagger \big]^{-1} R_0^\perp(\beta) \big\{ P^\perp J P^\perp R_0^\perp(\beta) \big\}^j P^\perp J P^\perp$$

and, since $\big[R_0^\perp(\beta)^\dagger \big]^{-1} R_0^\perp(\beta)$ is a unitary operator in $\ell_{\tilde{w}}^2$, that we need only show

$$\| R_0^\perp(\beta) P^\perp J P^\perp \|_{\ell_{\tilde{w}}^2 \to \ell_{\tilde{w}}^2}^2 \lesssim 1. \tag{48}$$

We show this by employing the Schur test. Looking at the explicit formula, (28), for J shows that (in the $|N, n, m\rangle$ basis) the matrix representing J has only finitely many non-zero elements in each row. Indeed the number of non-zero elements is at most eight

(there could be fewer since $c_{-1} = 0$). The same holds for the columns as well since J is self-adjoint in the unweighted space.

These matrix entries are of size $\sqrt{N+1}$ (indeed, J is not bounded); however, we have

$$R_0^\perp(\beta)|N, n, m\rangle = \frac{1}{N + \beta + i(n^2 - m^2)}|N, n, m\rangle$$

and $|N + \beta + i(n^2 - m^2)|^{-1} \leq \frac{1}{2}(N+1)^{-1}$ when $N \neq 0$ or $n \neq m$; that is, for all vectors in its domain. Therefore, each of the at most eight non-zero entries in every row/column of the matrix for $R_0^\perp(\beta)P^\perp J P^\perp$ is bounded by a uniform constant.

The preceding discussion is sufficient to show that $R_0^\perp(\beta)P^\perp J P^\perp$ is bounded in the unweighted space. To obtain the weight estimate one further needs to know that

$$\tilde{w}(n \pm 1, m) = \left\| |N, n \pm 1, m\rangle \right\|_{\ell_{\tilde{w}}^2} \lesssim \left\| |N', n, m\rangle \right\|_{\ell_{\tilde{w}}^2} = \tilde{w}(n, m)$$

and the corresponding result with $m \pm 1$. This is easily checked to be the case because $|\rho(n+1) - \rho(n)| \lesssim 1$ for the weight in question. ☐

Remark. By delving a little deeper into the specifics of the Ornstein–Uhlenbeck process, one can show that the main estimate holds with λ^2 on the right-hand side rather than just λ. This need not be the case for more general processes.

Proposition 4.2. *There exists a (small) constant $c > 0$ so that for all $\lambda < 1$ the following hold:*

$$\left\| (D^\dagger D + \beta)^{-1} \right\|_{\ell_w^2 \to \ell_w^2} \leq 2\beta^{-1}, \tag{49}$$

$$\left\| D(D^\dagger D + \beta)^{-1} D^\dagger \right\|_{\ell_{\tilde{w}}^2 \to \ell_{\tilde{w}}^2} \leq 3, \tag{50}$$

$$\left\| D(D^\dagger D + \beta)^{-1} \right\|_{\ell_w^2 \to \ell_{\tilde{w}}^2} \leq 3\beta^{-1/2}, \tag{51}$$

$$\left\| (D^\dagger D + \beta)^{-1} D^\dagger \right\|_{\ell_{\tilde{w}}^2 \to \ell_w^2} \leq 3\beta^{-1/2}, \tag{52}$$

where $w(n) = e^{2\rho(n)}$ and $\rho(n) = c\min\{n, \lambda^{-1}n^2\sqrt{\beta}\}$.

Proof. The first estimate is standard Combes–Thomas/Agmon fare: Let $e^{\pm\rho}$ denote the multiplication operators associated with the corresponding function of n:

$$e^{\pm\rho}|0, n, n\rangle = e^{\pm\rho(n)}|0, n, n\rangle.$$

Then $\left\| e^\rho|0, n, n\rangle \right\| = \left\| |0, n, n\rangle \right\|_{\ell_w^2}$ and so

$$\left\| (D^\dagger D + \beta)^{-1} \right\|_{\ell_w^2 \to \ell_w^2} = \left\| e^\rho(D^\dagger D + \beta)^{-1}e^{-\rho} \right\|$$
$$= \left\| \left[e^\rho(D^\dagger D + \beta)e^{-\rho} \right]^{-1} \right\|.$$

Now by brute calculation, we have that

$$e^\rho D^\dagger D e^{-\rho} = D^\dagger D - \xi + i\eta,$$

where $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ are self-adjoint operators defined as follows: for $|\psi\rangle = \sum \psi_n |0, n, n\rangle$,

$$\langle 0, n, n | \boldsymbol{\xi} | \psi \rangle = \xi_n \psi_{n+1} + \xi_{n-1} \psi_{n-1},$$

$$\xi_n = 2|\alpha_n|^2 \sinh^2 \left(\frac{\rho(n) - \rho(n+1)}{2} \right) \qquad n \ge 0, \qquad \xi_{-1} = 0,$$

$$\langle 0, n, n | \boldsymbol{\eta} | \psi \rangle = \eta_n \psi_{n+1} + \eta_{n-1} \psi_{n-1},$$

$$\eta_n = |\alpha_n|^2 \sinh \big(\rho(n) - \rho(n+1) \big) \qquad n \ge 0, \qquad \eta_{-1} = 0.$$

Recall that α_n comes from the explicit formula for D (see Lemma 3.3).

By Schur's test,

$$\|\boldsymbol{\xi}\|^2 \le \sup_n \ 2|\xi_n|^2 \le \sup_n \ \frac{\lambda^2}{1+n^2} \sinh^2 \left(\frac{\rho(n) - \rho(n+1)}{2} \right).$$

So, since $|\rho(n+1) - \rho(n)| \lesssim c \min\{1, \lambda^{-1} n \sqrt{\beta}\}$, one can ensure that $\|\boldsymbol{\xi}\| \le \frac{1}{2}\beta$ by choosing c small enough. As $D^\dagger D$ is a positive operator, it follows that

$$\left\| \frac{1}{D^\dagger D + \beta} \right\|_{\ell_w^2 \to \ell_w^2} = \left\| \frac{1}{D^\dagger D - \boldsymbol{\xi} + i\boldsymbol{\eta} + \beta} \right\| \le 2\beta^{-1}$$

which proves (49).

By identical arguments, one can show that for c small enough

$$\left\| \frac{1}{DD^\dagger + \beta} \right\|_{\ell_{\tilde{w}}^2 \to \ell_{\tilde{w}}^2} \le 2\beta^{-1}.$$

Equation (50) now follows from the following commutation formula:

$$D \frac{1}{D^\dagger D + \beta} D^\dagger = \frac{1}{DD^\dagger + \beta} DD^\dagger = 1 - \frac{\beta}{DD^\dagger + \beta},$$

which we learned from Percy Deift (see [8] for other, unrelated, applications).

Corresponding to the weight $\tilde{w}(n, m) = \sqrt{w(n)w(m)}$ we define

$$\tilde{\rho}(n, m) = \tfrac{1}{2}\rho(n) + \tfrac{1}{2}\rho(m)$$

so that $\tilde{w}(n, m) = e^{\tilde{\rho}(n,m)}$. As was the case for e^ρ we regard $e^{\tilde{\rho}}$ as a multiplication operator,

$$e^{\pm \tilde{\rho}} |N, n, m\rangle = e^{\pm \tilde{\rho}(n,m)} |N, n, m\rangle.$$

To prove (51) we will use the following consequence of Lemma 3.3: if $|\psi\rangle = \sum \psi_n |0, n, n\rangle$ then

$$\langle 1, n+1, n | e^{2\tilde{\rho}} D e^{-2\rho} | \psi \rangle = \alpha_n \big[e^{\rho(n+1) - \rho(n)} \psi_n - e^{\rho(n) - \rho(n+1)} \psi_{n+1} \big], \tag{53}$$

$$\langle 1, n, n+1 | e^{2\tilde{\rho}} D e^{-2\rho} | \psi \rangle = \bar{\alpha}_n \big[e^{\rho(n) - \rho(n+1)} \psi_{n+1} - e^{\rho(n+1) - \rho(n)} \psi_n \big], \tag{54}$$

and all other co-efficients are zero. This shows, by Schur's test, that $e^{2\tilde{\rho}} D e^{-2\rho} = D + \zeta$, where ζ is an operator with

$$\big\| e^{-\tilde{\rho}} \zeta e^\rho \big\| \le \sup_n \ 2|\alpha_n| \big| \sinh \big[\tfrac{1}{2}\rho(n+1) - \tfrac{1}{2}\rho(n) \big] \big| \le \beta^{1/2}.$$

The second inequality here follows from our particular choice of ρ when c is sufficiently small. In particular, notice that $|\rho(n + 1) - \rho(n)| \lesssim c \min\{1, \lambda^{-1}n\sqrt{\beta}\}$ and $|\alpha_n| \lesssim \lambda(1 + n)^{-1}$, see (31).

Now,

$$\left\| D\frac{1}{D^\dagger D + \beta} \right\|_{\ell^2_w \to \ell^2_{\tilde{w}}}^2 = \left\| e^{\tilde{\rho}} D \frac{1}{D^\dagger D + \beta} e^{-\rho} \right\|^2 \tag{55}$$

$$= \left\| e^{-\rho} \frac{1}{D^\dagger D + \beta} D^\dagger e^{2\tilde{\rho}} D \frac{1}{D^\dagger D + \beta} e^{-\rho} \right\| \tag{56}$$

$$= \left\| e^{-\rho} \frac{1}{D^\dagger D + \beta} \left[D^\dagger D + D^\dagger \zeta \right] e^{2\rho} \frac{1}{D^\dagger D + \beta} e^{-\rho} \right\| \tag{57}$$

$$\leq 2\beta^{-1} \left\| e^{-\rho} \frac{1}{D^\dagger D + \beta} \left[D^\dagger D + D^\dagger \zeta \right] e^\rho \right\| \tag{58}$$

by employing $e^{2\tilde{\rho}} D = (D + \zeta)e^{2\rho}$ and (49). We continue this chain by using, inter alia, $\frac{1}{D^\dagger D+\beta} D^\dagger D = 1 - \frac{\beta}{D^\dagger D+\beta}$, the triangle inequality, and (49) again,

$$\leq 2\beta^{-1} \left\{ 1 + \left\| e^{-\rho} \frac{\beta}{D^\dagger D + \beta} e^\rho \right\| + \left\| e^{-\rho} \frac{1}{D^\dagger D + \beta} D^\dagger e^{\tilde{\rho}} \right\| \left\| e^{-\tilde{\rho}} \zeta e^\rho \right\| \right\} \tag{59}$$

$$\leq 6\beta^{-1} + \beta^{-1/2} \left\| e^{-\rho} \frac{1}{D^\dagger D + \beta} D^\dagger e^{\tilde{\rho}} \right\| \tag{60}$$

$$\leq 6\beta^{-1} + \beta^{-1/2} \left\| e^{\tilde{\rho}} D \frac{1}{D^\dagger D + \beta} e^{-\rho} \right\| \tag{61}$$

$$\leq 6\beta^{-1} + \beta^{-1/2} \left\| D \frac{1}{D^\dagger D + \beta} \right\|_{\ell^2_w \to \ell^2_{\tilde{w}}}. \tag{62}$$

In passing from (60) to (61) we used the fact that the norm of an operator is equal to the norm of its adjoint. Now, (55)–(62) show that

$$x = \sqrt{\beta} \left\| D \frac{1}{D^\dagger D + \beta} \right\|_{\ell^2_w \to \ell^2_{\tilde{w}}} \qquad \text{obeys} \qquad x^2 \leq 6 + x,$$

from which it follows that $x \leq 3$. This completes the proof of (51).

The proof of (52) is essentially the same as the above. Indeed, since the proof of (49)–(51) did not require that c be positive, only that it was small, it follows that (51) holds with w and $\tilde{w}$ replaced with their reciprocals. Taking the adjoint of the operator in this modified inequality (51) proves (52). $\quad\square$

Since $H = (1 + \beta)D^\dagger D$, the following follows trivially from this proposition.

Corollary 4.3. *There exists a (small) constant $c > 0$ so that for all $\lambda < 1$ and all $\beta < 1$,*

$$\left\| (H + \beta)^{-1} \right\|_{\ell^2_w \to \ell^2_w} \leq 4\beta^{-1}, \tag{63}$$

$$\left\| D(H + \beta)^{-1} D^\dagger \right\|_{\ell^2_{\tilde{w}} \to \ell^2_{\tilde{w}}} \leq 6, \tag{64}$$

$$\left\| D(H + \beta)^{-1} \right\|_{\ell^2_w \to \ell^2_{\tilde{w}}} \leq 6\beta^{-1/2}, \tag{65}$$

$$\left\| (H + \beta)^{-1} D^\dagger \right\|_{\ell^2_{\tilde{w}} \to \ell^2_w} \leq 6\beta^{-1/2}, \tag{66}$$

where $w(n) = e^{2\rho(n)}$ and $\rho(n) = c \min\{n, \lambda^{-1}n^2\sqrt{\beta}\}$.

The operator H was introduced as the "main part" of $PL_\lambda P$. We now estimate the discrepancy between the two.

Corollary 4.4. *There exists a (small) constant $c > 0$ so that for λ sufficiently small and all $\beta < 1$,*

$$\left\| \left[P \frac{1}{(L_\lambda + \beta)} P - \frac{1}{H + \beta} \right] \right\|_{\ell_w^2 \to \ell_w^2} \lesssim \lambda \beta^{-1}, \tag{67}$$

$$\left\| D \left[P \frac{1}{(L_\lambda + \beta)} P - \frac{1}{H + \beta} \right] D^\dagger \right\|_{\ell_{\tilde{w}}^2 \to \ell_{\tilde{w}}^2} \lesssim \lambda, \tag{68}$$

$$\left\| D \left[P \frac{1}{(L_\lambda + \beta)} P - \frac{1}{H + \beta} \right] \right\|_{\ell_w^2 \to \ell_{\tilde{w}}^2} \lesssim \lambda \beta^{-1/2}, \tag{69}$$

$$\left\| \left[P \frac{1}{(L_\lambda + \beta)} P - \frac{1}{H + \beta} \right] D^\dagger \right\|_{\ell_{\tilde{w}}^2 \to \ell_w^2} \lesssim \lambda \beta^{-1/2}, \tag{70}$$

where $w(n) = e^{2\rho(n)}$ and $\rho(n) = c \min\{n, \lambda^{-1} n^2 \sqrt{\beta}\}$.

Proof. By (26) and then Lemma 4.1,

$$P \frac{1}{(L_\lambda + \beta)} P - \frac{1}{H + \beta}$$

$$= \sum_{j=1}^\infty (H + \beta)^{-1} \left\{ -\lambda^2 PJ P^\perp [R_\lambda^\perp(\beta) - R_0^\perp(\beta)] P^\perp JP (H + \beta)^{-1} \right\}^j$$

$$= \sum_{j=1}^\infty (H + \beta)^{-1} \left\{ D^\dagger ED (H + \beta)^{-1} \right\}^j$$

$$= \sum_{j=0}^\infty (H + \beta)^{-1} D^\dagger E \left\{ D(H + \beta)^{-1} D^\dagger E \right\}^j D(H + \beta)^{-1},$$

where the operator E obeys

$$\|E\|_{\ell_{\tilde{w}}^2 \to \ell_{\tilde{w}}^2}^2 \lesssim \lambda.$$

So, once λ is sufficiently small, (64) from Corollary 4.3 permits us to sum the series and prove (67)–(70). $\square$

5. Pointwise Estimates

Given a weight function $w(n) = e^{2\rho(n)}$ and an estimate such as (63),

$$\left\| (H + \beta)^{-1} \right\|_{\ell_w^2 \to \ell_w^2} \le 4\beta^{-1}$$

there is a standard and simple way to obtain pointwise bounds for the Green function:

$$\langle 0, n, n | (H + \beta)^{-1} | 0, m, m \rangle \tag{71}$$

$$\le \left\| e^{-\rho} | 0, n, n \rangle \right\| \left\| e^\rho (H + \beta)^{-1} e^{-\rho} \right\| \left\| e^\rho | 0, m, m \rangle \right\| \tag{72}$$

$$\le 4\beta^{-1} \exp \left\{ \rho(m) - \rho(n) \right\}. \tag{73}$$

In terms of the exponential behaviour, estimates of this type cannot really be improved. However, as we will show, the factor β^{-1} in front can be significantly improved.

As was remarked after Lemma 3.4, the operator H is the generator of a continuous-time Markov chain. Indeed, since H is the dominant term in the operator L_λ, this shows that to a good approximation, the quantum particle undergoes a random walk in momentum space. Note that the hopping amplitudes for this random walk are $a_n \asymp n^{-2}$ and so the diffusion is slower near infinity than for a homogeneous random walk.

As probability is conserved, for any m,

$$\sum_n \langle 0, n, n | e^{-tH} | 0, m, m \rangle = 1 \qquad \text{for all } t \geq 0, \tag{74}$$

where each summand is positive. (Those unfamiliar with continuous-time Markov chains may see Chapter VI of [9].) From (74),

$$\sum_n \langle 0, n, n | (H + \beta)^{-1} | 0, 0, 0 \rangle = \int_0^\infty e^{-\beta t} \sum_n \langle 0, n, n | e^{-tH} | 0, 0, 0 \rangle \, dt = \beta^{-1}. \tag{75}$$

However, if we use the pointwise estimate that we derived above, (71)–(73), with $\rho(n) = c \min\{n, n^2 \sqrt{\beta}\}$, then, ignoring the λ dependence, we have

$$\sum_n \langle 0, n, n | (H + \beta)^{-1} | 0, 0, 0 \rangle \leq 4\beta^{-1} \sum_n e^{-\rho(n)} \asymp \beta^{-5/4}.$$

The reason for the discrepancy is not in the choice of ρ, but rather that too much was given away in passing from (71) to (72). The weighted norm estimate on $(H + \beta)^{-1}$ is of the correct size, however, the vectors that realize that norm are spread out, not localized as $|0, m, m\rangle$ is. Indeed, H is a differential operator, so the vectors $|\psi\rangle$ for which $\|H|\psi\rangle\|$ is small must have slowly varying co-efficients. Another manifestation of this is the fact that the norm of $D(H + \beta)^{-1}$, which should be regarded as the derivative of the Green function, is $\sqrt{\beta}$ times smaller than the norm of $(H + \beta)^{-1}$. The following lemma shows how one may represent $|0, n, n\rangle$ so as to utilize this extra information.

Lemma 5.1. *Let $w(n) = e^{2\rho(n)}$ with $\rho(n) = c \min\{n, \lambda^{-1} n^2 \sqrt{\beta}\}$ and $c < 1$ a constant. For each $n \geq 0$, there are vectors $|\Lambda_n\rangle \in \ell^2_{\tilde{w}}$ and $|\Omega_n\rangle \in \ell^2_w$ so that*

$$|0, n, n\rangle = D^\dagger |\Lambda_n\rangle + |\Omega_n\rangle$$

and

$$e^{\mp\rho(n)} \left\| e^{\pm\rho} |\Lambda_n\rangle \right\| \lesssim \begin{cases} \lambda^{-1/4} \beta^{-3/8} & : 0 \leq n \leq \lambda^{1/2} \beta^{-1/4} \\ \lambda^{-1/2} \beta^{-1/4} \sqrt{n} & : \lambda^{1/2} \beta^{-1/4} \leq n \leq \lambda\beta^{-1/2} \\ 0 & : \lambda\beta^{-1/2} < n \end{cases} \tag{76}$$

$$e^{\mp\rho(n)} \left\| e^{\pm\rho} |\Omega_n\rangle \right\| \lesssim \begin{cases} \lambda^{-1/4} \beta^{1/8} & : 0 \leq n \leq \lambda^{1/2} \beta^{-1/4} \\ \lambda^{-1/2} \beta^{1/4} \sqrt{n} & : \lambda^{1/2} \beta^{-1/4} \leq n \leq \lambda\beta^{-1/2} \\ 1 & : \lambda\beta^{-1/2} < n \end{cases} \tag{77}$$

uniformly for $\beta, \lambda \in (0, 1)$.

Proof. In the case $\lambda\beta^{-1/2} < n$ we choose $|\Omega_n\rangle = |0, n, n\rangle$ and $|\Lambda_n\rangle = 0$. The inequalities are trivial in this case.

For $n \le \lambda\beta^{-1/2}$, let

$$
|\Lambda_n\rangle = \sum_{k=n}^{n+N-1} \alpha_n^{-1}\left(1 - \tfrac{k-n}{N}\right)|1, k, k+1\rangle,
$$

where $N \approx \min\{\lambda^{1/2}\beta^{-1/4}, \lambda n^{-1}\beta^{-1/2}\}$ is an integer:

$$
N = \begin{cases} \lceil\lambda^{1/2}\beta^{-1/4}\rceil & : 0 \le n \le \lambda^{1/2}\beta^{-1/4} \\ \lceil\lambda n^{-1}\beta^{-1/2}\rceil & : \lambda^{1/2}\beta^{-1/4} \le n \le \lambda\beta^{-1/2} \end{cases}
$$

($\lceil x\rceil$ is the least integer greater than or equal to x). For ease of reading, we use the symbol $\approx$ to indicate that two things are equal modulo the rounding of N to an integer. By Lemma 3.3,

$$
D^\dagger|\Lambda_n\rangle = \sum_{k=n}^{n+N}\left(1 - \tfrac{k-n}{N}\right)\{|0, k, k\rangle - |0, k+1, k+1\rangle\}
$$

$$
= |0, n, n\rangle - \sum_{k=n+1}^{n+N} \tfrac{1}{N}|0, k, k\rangle,
$$

so we choose $|\Omega_n\rangle = \sum_{k=n+1}^{n+N}\tfrac{1}{N}|0, k, k\rangle$.

We now proceed to calculate the norms:

$$
\left\|e^{-\rho}|\Omega_n\rangle\right\|^2 = N^{-2}\sum_{k=n+1}^{n+N} e^{-2\rho(k)}
$$

$$
\le N^{-1}e^{-2\rho(n)}
$$

$$
\lesssim \max\{\lambda^{-1/2}\beta^{1/4}, \lambda^{-1}n\beta^{1/2}\}\,e^{-2\rho(n)},
$$

$$
\left\|e^{\rho}|\Omega_n\rangle\right\|^2 = N^{-2}\sum_{k=n+1}^{n+N} e^{2\rho(k)}
$$

$$
\le N^{-1}e^{2\rho(n+N)}
$$

$$
\lesssim \max\{\lambda^{-1/2}\beta^{1/4}, \lambda^{-1}n\beta^{1/2}\}\,e^{2\rho(n)}.
$$

Here we used the fact that $\rho(n+N) \le \rho(n)+5$. The justification of this is as follows: for $n \le \lambda\beta^{-1/2}$, $\rho(n) = c\lambda^{-1}n^2\sqrt{\beta}$ and then either (a) $0 \le n \le \lambda^{1/2}\beta^{-1/4}$ in which case $\lambda^{-1}N^2\sqrt{\beta} \approx 1$ and $\lambda^{-1}2nN\sqrt{\beta} \le \lambda^{-1/2}2N\beta^{1/4} \approx 2$; or (b) $\lambda^{1/2}\beta^{-1/4} \le n \le \lambda\beta^{-1/2}$ in which case $\lambda^{-1}N^2\sqrt{\beta} \approx \lambda(n^2\sqrt{\beta})^{-1} \le 1$ and $\lambda^{-1}2nN\sqrt{\beta} \approx 2$. This completes the proof of (77).

For $\beta < 1$, we have $|\alpha_n| \geq \lambda(1 + n)^{-1}$ and so

$$\left\| e^{-\rho} |\Lambda_n\rangle \right\|^2 \leq \sum_{k=n}^{n+N-1} |\alpha_n|^{-2} \big(1 - \tfrac{k-n}{N}\big)^2 e^{-2\rho(k)} \leq N\lambda^{-2}(n + N)^2 e^{-2\rho(n)},$$

$$\left\| e^{\rho} |\Lambda_n\rangle \right\|^2 \leq \sum_{k=n}^{n+N-1} |\alpha_n|^{-2} \big(1 - \tfrac{k-n}{N}\big)^2 e^{2\rho(k)} \leq N\lambda^{-2}(n + N)^2 e^{2\rho(n+N)}.$$

By using $\rho(n + N) \leq \rho(n) + 5$ these simplify to

$$e^{\mp 2\rho(n)} \left\| e^{\pm\rho} |\Lambda_n\rangle \right\|^2 \lesssim N\lambda^{-2}(n + N)^2.$$

Now, when $n \leq \lambda^{1/2}\beta^{-1/4}$, we have $n \leq N \approx \lambda^{1/2}\beta^{-1/4}$ and so $\lambda^{-2}N(n + N)^2 \lesssim \lambda^{-1/2}\beta^{-3/4}$. When $\lambda^{1/2}\beta^{-1/4} \leq n \leq \lambda\beta^{-1/2}$, we have $n \geq N \approx \lambda\beta^{-1/2}n^{-1}$ from which it follows that $\lambda^{-2}N(n + N)^2 \lesssim \lambda^{-1}n\beta^{-1/2}$. In both cases, this is just what is required to give (76). □

Proposition 5.2. *Let* $\rho(n) = c \min\{n, \lambda^{-1}n^2\sqrt{\beta}\}$ *with* $c > 0$ *sufficiently small. For all* λ *sufficiently small and all* $\beta < 1$,

$$0 < \Big\langle 0, n, n \Big| \frac{1}{H + \beta} \Big| 0, m, m \Big\rangle \leq C_n^{(\beta)} C_m^{(\beta)} e^{-|\rho(n)-\rho(m)|} \tag{78}$$

and

$$\Big| \Big\langle 0, n, n \Big| P \frac{1}{L_\lambda + \beta} P - \frac{1}{H + \beta} \Big| 0, m, m \Big\rangle \Big| \leq \lambda C_n^{(\beta)} C_m^{(\beta)} e^{-|\rho(n)-\rho(m)|}, \tag{79}$$

where

$$C_k^{(\beta)} = C \begin{cases} \lambda^{-1/4}\beta^{-3/8} & : 0 \leq k \leq \lambda^{1/2}\beta^{-1/4} \\ \lambda^{-1/2}\beta^{-1/4}\sqrt{k} & : \lambda^{1/2}\beta^{-1/4} \leq k \leq \lambda\beta^{-1/2} \\ \beta^{-1/2} & : \lambda\beta^{-1/2} < k \end{cases}$$

with some constant C *which does not depend on* β *or on* λ.

Proof. The result depends only on the fact that the operators in question admit the estimates (63)–(66) and (67)–(70). We discuss only the operator H.

By symmetry, it suffices to study the case $n \geq m$. By Lemma 5.1 and Corollary 4.3,

$$\langle 0, n, n | (H + \beta)^{-1} | 0, m, m \rangle$$
$$= \Big\{ \langle \Lambda_n | D + \langle \Omega_n | \Big\} e^{-\rho} e^{\rho} (H + \beta)^{-1} e^{-\rho} e^{\rho} \Big\{ D^\dagger |\Lambda_m\rangle + |\Omega_m\rangle \Big\}$$
$$\leq 6 \Big\{ \| e^{-\rho} |\Lambda_n\rangle \| + \beta^{-1/2} \| e^{-\rho} |\Omega_n\rangle \| \Big\} \Big\{ \| e^{\rho} |\Lambda_m\rangle \| + \beta^{-1/2} \| e^{\rho} |\Omega_m\rangle \| \Big\}$$
$$\lesssim C_n^{(\beta)} C_m^{(\beta)} e^{\rho(m)-\rho(n)}.$$

Positivity in (78) follows from integrating $\langle 0, n, n | e^{-tH} | 0, m, m \rangle > 0$. □

Remark. This new pointwise estimate gives results in line with the conservation of probability (cf, (75)):

$$\sum_n \langle 0, n, n | (H + \beta)^{-1} | 0, 0, 0 \rangle$$

$$\lesssim \sum_n \lambda^{-1/4} \beta^{-3/8} C_n^{(\beta)} e^{-\rho(n)}$$

$$\lesssim \beta^{-1} + \int_{\lambda^{1/2}\beta^{-1/4}}^{\lambda\beta^{-1/2}} \lambda^{-3/4} \beta^{-5/8} e^{-c\lambda^{-1}n^2\sqrt{\beta}} \sqrt{n} \, dn$$

$$\lesssim \beta^{-1} + \int_1^\infty \beta^{-1} e^{-t^2} \sqrt{t} \, dt$$

$$\lesssim \beta^{-1}.$$

In this computation we estimated the sum over $0 \le n \le \lambda^{1/2}\beta^{-1/4}$ by replacing $C_n^{(\beta)}$ by its largest value, namely $\lambda^{-1/4}\beta^{-3/8}$. The sum over $\lambda\beta^{-1/2} < n$ is of a geometric series and so can be computed exactly. Lastly, we changed variables in the integral according to $n = \lambda^{1/2}\beta^{-1/4}t$.

Proposition 5.3. *For all λ sufficiently small, $\beta \le \lambda^2$, $m \ge 0$, and $s \ge 0$,*

$$\sum_n n^{2s} \langle 0, n, n | \frac{1}{H + \beta} | 0, m, m \rangle \asymp \beta^{-1}(m^{2s} + \lambda^s \beta^{-s/2}), \tag{80}$$

and

$$\sum_n n^{2s} \langle 0, n, n | P\frac{1}{L_\lambda + \beta}P - \frac{1}{H + \beta} | 0, m, m \rangle \lesssim \lambda\beta^{-1}(m^{2s} + \lambda^s \beta^{-s/2}), \tag{81}$$

where the constants do not depend on λ or β. Combining these two estimates shows that for λ sufficiently small

$$\sum_n (1 + n^{2s}) \langle 0, n, n | P\frac{1}{L_\lambda + \beta}P | 0, m, m \rangle \asymp 1 + m^{2s} + \lambda^s \beta^{-s/2}. \tag{82}$$

Proof. The proof of (81) is the same as that of the upper bound in (80), differing only in that it uses (79) instead of (78). It is therefore omitted.

We will use the following simple observation repeatedly: for any function f with $f(x) > 0$ and $|f'(x)/f(x)| \lesssim 1$,

$$\sum_{n=a}^b f(n) \lesssim f(x_0) + \int_a^b f(x) \, dx, \tag{83}$$

for any $x_0 \in [a, b]$. The assumption that $\lambda^{-1}\beta^{1/2} \le 1$ is sufficient to ensure the hypothesis on the logarithmic derivative in each of the instances below.

We will also use the following simple inequalities:

$$\int_s^\infty n^\gamma e^{-n} \, dn \lesssim s^\gamma e^{-s}, \qquad \int_0^s n^\gamma e^n \, dn \lesssim s^\gamma e^s, \tag{84}$$

and their corollaries

$$\int_s^\infty n^{\gamma+1} e^{-u^2 n^2} \, dn \lesssim u^{-2} s^\gamma e^{-u^2 s^2}, \qquad \int_0^s n^{\gamma+1} e^{u^2 n^2} \, dn \lesssim u^{-2} s^\gamma e^{u^2 s^2}. \tag{85}$$

In (84) and (85), the implicit constants depend only on $\gamma > -1$.

We now begin the proof of the $\lesssim$ part of (80); it is much like the argument given in the remark above. The first step is to replace the matrix elements of the resolvent of H by the upper bounds given in (78) of Proposition 5.2. Now we need to calculate the sum over n. The reasoning necessary depends on the value of m and so we consider separately three cases:

Case 1. $0 \le m \le \lambda^{1/2} \beta^{-1/4}$. In this regime, $\rho(m) \le c$ and so

$$\sum_n n^{2s} C_n^{(\beta)} C_m^{(\beta)} e^{-|\rho(m)-\rho(n)|} \lesssim \lambda^{-1/4} \beta^{-3/8} \sum_n n^{2s} C_n^{(\beta)} e^{-\rho(n)}$$

$$\lesssim S_1 + S_2 + S_3,$$

where S_1, S_2, and S_3 are given by restricting the sum to those n in $[0, \lambda^{1/2}\beta^{-1/4}]$, $[\lambda^{1/2}\beta^{-1/4}, \lambda\beta^{-1/2}]$, and $[\lambda\beta^{-1/2}, \infty)$ respectively. Using (83) with $x_0 = 0$,

$$S_1 \lesssim \lambda^{-1/2} \beta^{-3/4} \sum_{n=0}^{\lambda^{1/2}\beta^{-1/4}} n^{2s} \lesssim \lambda^s \beta^{-1} \beta^{-s/2}.$$

Next, using (83) with $x_0 = \lambda^{1/2}\beta^{-1/4}$ and (85), we get

$$S_2 \lesssim \lambda^{-3/4} \beta^{-5/8} \sum_{n=\lambda^{1/2}\beta^{-1/4}}^{\lambda\beta^{-1/2}} n^{2s+\frac{1}{2}} e^{-c\beta^{1/2}n^2/\lambda} \lesssim \left(\lambda^{-1/2}\beta^{-3/4} + \beta^{-1}\right) \lambda^s \beta^{-s/2}.$$

But $\beta \le \lambda^2$ and so the above implies $S_2 \lesssim \lambda^s \beta^{-1} \beta^{-s/2}$.

Finally, in S_3, we can extend the sum from $\lambda\beta^{-1/2} \le n < \infty$ to all positive n:

$$S_3 \lesssim \lambda^{-1/4} \beta^{-7/8} \sum_{n=0}^\infty n^{2s} e^{-cn} \lesssim \lambda^{-1/4} \beta^{-7/8} \lesssim \lambda^s \beta^{-1} \beta^{-s/2},$$

where the last inequality uses $\beta < \lambda^2$.

Case 2. $\lambda^{1/2}\beta^{-1/4} \le m \le \lambda\beta^{-1/2}$. In this regime, $\rho(m) = c\lambda^{-1}m^2\beta^{1/2}$ and we proceed as above, i.e., we divide the sum over all n into three pieces S_1, S_2, and S_3 each corresponding to the same interval of ns as in Case 1.

For the sum over $0 \le n \le \lambda^{1/2}\beta^{-1/4}$ we use (83) with $x_0 = 0$, and the fact that $x^{1/4}e^{-cx} \lesssim 1$ uniformly for $x > 0$ and so in particular when $x = \lambda^{-1}m^2\beta^{1/2}$:

$$S_1 \lesssim \lambda^{-3/4} \beta^{-5/8} m^{1/2} e^{-c\lambda^{-1}m^2\beta^{1/2}} \sum_{n=0}^{\lambda^{1/2}\beta^{-1/4}} n^{2s}$$

$$\lesssim \lambda^s \beta^{-s/2} \lambda^{-1/4} \beta^{-7/8} m^{1/2} e^{-c\lambda^{-1}m^2\beta^{1/2}}$$

$$\lesssim \lambda^s \beta^{-1} \beta^{-s/2}.$$

Next, applying (83) twice with $x_0 = m$ and then both parts of (85),

$$S_2 \lesssim \lambda^{-1/2} \beta^{-1/2} m^{1/2} \left\{ \lambda^{-1/2} m^{2s+\frac{1}{2}} + \int_0^m n^{2s+\frac{1}{2}} e^{-c\lambda^{-1}(m^2-n^2)\beta^{1/2}} \, dn \right.$$

$$\left. + \int_m^\infty n^{2s+\frac{1}{2}} e^{-c\lambda^{-1}(n^2-m^2)\beta^{1/2}} \, dn \right\}$$

$$\lesssim \lambda^{-1} m^{2s+1} \beta^{-1/2} + m^{2s} \beta^{-1}$$

$$\lesssim \beta^{-1} m^{2s}.$$

In the last inequality, we used the fact that $m \leq \lambda\beta^{-1/2}$. Finally, using (83) and (84),

$$S_3 \lesssim \lambda^{-1/2} \beta^{-3/4} m^{1/2} \sum_{n=\lambda\beta^{-1/2}}^\infty n^{2s} e^{-cn+c\lambda^{-1}m^2\beta^{1/2}}$$

$$\lesssim \lambda^{-1/2} \beta^{-3/4} m^{1/2} \lambda^{2s} \beta^{-s} e^{-c\lambda\beta^{-1/2}+c\lambda^{-1}m^2\beta^{1/2}}$$

$$\lesssim \beta^{-1} m^{2s}.$$

To obtain the last inequality, consider the cases $s > 1/4$ and $s \leq 1/4$ separately.

Case 3. $\lambda\beta^{-1/2} < m$. We proceed similarly: By (83)

$$S_1 \lesssim \lambda^{-1/4} \beta^{-7/8} e^{-cm} \sum_{n=0}^{\lambda^{1/2}\beta^{-1/4}} n^{2s}$$

$$\lesssim \lambda^s \beta^{-s/2} \lambda^{1/4} \beta^{-9/8} e^{-cm}$$

$$\lesssim \lambda^s \beta^{-s/2} \beta^{-1} m^{1/4} e^{-cm}$$

$$\lesssim \beta^{-1} \lambda^s \beta^{-s/2},$$

because $x^{1/4} e^{-cx} \lesssim 1$. Extending the summation region down to $n = 0$, applying (83) with $x_0 = 0$ and then (85),

$$S_2 \lesssim \lambda^{-1/2} \beta^{-3/4} e^{-cm} \int_0^{\lambda\beta^{-1/2}} n^{2s+\frac{1}{2}} e^{c\lambda^{-1}n^2\beta^{1/2}} \, dn$$

$$\lesssim \beta^{-1} \lambda^{2s} \beta^{-s} e^{-c(m-\lambda\beta^{-1/2})}$$

$$\lesssim \beta^{-1} m^{2s},$$

where the last line follows from $m \geq \lambda\beta^{-1/2}$.

To complete the proof of Case 3 and so of the $\lesssim$ part of (80) we estimate S_3 by extending the sum to all $n \geq 0$ and applying both parts of (84):

$$S_3 \lesssim \sum_{n=\lambda\beta^{-1/2}}^\infty n^{2s} e^{-c|m-n|} \lesssim \beta^{-1} m^{2s}.$$

To prove the $\gtrsim$ part of (80) we use

$$\sum_n \langle 0, n, n | (H + \beta)^{-1} | 0, m, m \rangle = \beta^{-1}, \tag{86}$$

which was derived earlier (see (75)).

Let $G(n, m) = \langle 0, n, n|(H + \beta)^{-1}|0, m, m\rangle$ and, given $\epsilon > 0$, define

$$n_0 = \begin{cases} \epsilon\lambda^{1/2}\beta^{-1/4} & : m \leq \lambda^{1/2}\beta^{-1/4} \\ \epsilon m & : m \geq \lambda^{1/2}\beta^{-1/4} \end{cases}. \tag{87}$$

For ϵ sufficiently small (independent of m, β and λ) one can check that

$$\sum_{n<n_0} G(n, m) \leq \tfrac{1}{2}\beta^{-1}, \tag{88}$$

using computations similar to those performed earlier in this proof. By (86), this means that the sum over $n \geq n_0$ must be at least $\frac{1}{2}\beta^{-1}$ and so

$$\sum_{n=0}^{\infty} n^{2s}G(n, m) \geq \sum_{n=n_0}^{\infty} n_0^{2s}G(n, m) \geq \tfrac{1}{2}\beta^{-1}n_0^{2s} \gtrsim \beta^{-1}(m^{2s} + \lambda^s\beta^{-s/2}).$$

The final inequality here should be checked separately for each case in (87). $\quad\square$

Lemma 5.4. *If $w(n) = 1 + n^{2s}$, to be considered as a weight on the range of P, and $W(n, m) = (1 + |n - m|^2)^{-1}\sqrt{w(n)w(m)}$, to be considered as a weight on the range of $P + P^{\perp}$, then*

$$\left\| PJP^{\perp}R_{\lambda}^{\perp}(\beta)P^{\perp}\right\|_{\ell_W^1 \to \ell_w^1} \lesssim 1,$$

when λ is sufficiently small.

Proof. The resolvent formula implies that

$$PJP^{\perp}R_{\lambda}^{\perp}(\beta)P^{\perp} = PJR_0^{\perp}(\beta)P^{\perp}\sum_{j=0}^{\infty}\left\{ -i\lambda P^{\perp}JR_0^{\perp}(\beta)P^{\perp}\right\}^j.$$

Thus, the claim will follow from

$$\left\| JR_0^{\perp}(\beta)P^{\perp}\right\|_{\ell_W^1 \to \ell_W^1} \lesssim 1, \tag{89}$$

whose proof is very much like that of (48) in Lemma 4.1. For variety, we present the argument slightly differently.

Notice that the weight W obeys

$$\frac{W(n \pm 1, m)}{W(N, n, m)} \lesssim 1 \quad \text{and} \quad \frac{W(n, m \pm 1)}{W(N, n, m)} \lesssim 1,$$

and let $\mathfrak{N}$ denote the multiplication operator $\mathfrak{N}|N, n, m\rangle = (1 + N)|N, n, m\rangle$.

By Schur's test, (28), and the above, $J\mathfrak{N}^{-1}$ is a bounded operator on ℓ_W^1. As $R_0^{\perp}(\beta)|N, n, m\rangle = [N + \beta + i(n^2 - m^2)]^{-1}|N, n, m\rangle$, the operator $\mathfrak{N}R_0^{\perp}(\beta)P^{\perp}$ is also bounded on this space. Taking the product of these two bounded operators proves Eq. (89). $\quad\square$

We are now ready to complete the proof of the theorem, which we restate for the reader's convenience.

Theorem. *Let $X(t)$ be a process obeying Hypothesis 1 and let $V(\theta)$ be an even trigonometric polynomial. If the coupling constant λ is sufficiently small then for any initial wave-function $\psi_0 \in H^s$,*

$$\frac{1}{T}\int_0^\infty e^{-t/T}\, \mathbb{E}\left\{\|\psi(\cdot,t)\|_{H^s}^2\right\} dt \asymp \lambda^s T^{s/2}\|\psi_0\|_{L^2}^2 + \|\psi_0\|_{H^s}^2 \tag{90}$$

as $T \to \infty$. The expectation is over possible trajectories of the Markov process $X(t)$. In particular, the energy grows on average as the square-root of time.

Proof. By Proposition 2.1, we need to estimate

$$\sum_n \left(1 + |n|^{2s}\right)\!\Big\langle 0, n, n\Big|\frac{\beta}{L_\lambda + \beta}\Big|\mathsf{P}(0)\Big\rangle, \tag{91}$$

where $\mathsf{P}(0) = u_0 \otimes |\psi_0\rangle\langle\psi_0|$ and $\beta = \frac{1}{T}$. Recall from Sect. 3 that $|0, n, n\rangle$ is in the range of P, if $n \geq 0$, or of Q, if $n < 0$. Recall also that by Lemma 3.1 the ranges of both $P + P^\perp$ and $Q + Q^\perp$ are invariant subspaces for L_λ. Combining these two facts, we can re-write (91) as

$$\sum_{n\geq 0} \left(1 + |n|^{2s}\right)\!\Big\langle 0, n, n\Big|P\frac{\beta}{L_\lambda + \beta}\big[P + P^\perp\big]\Big|\mathsf{P}(0)\Big\rangle$$

$$+ \sum_{n<0} \left(1 + |n|^{2s}\right)\!\Big\langle 0, n, n\Big|Q\frac{\beta}{L_\lambda + \beta}\big[Q + Q^\perp\big]\Big|\mathsf{P}(0)\Big\rangle. \tag{92}$$

As earlier, we will treat the even part of the wave function (the top line above); the odd part may be treated identically.

By (22) and (23),

$$P(L_\lambda + \beta)^{-1}[P + P^\perp] = P(L_\lambda + \beta)^{-1}P\big[P - i\lambda PJP^\perp R_\lambda^\perp(\beta)P^\perp\big]. \tag{93}$$

So let $|\phi_1\rangle = P\,|\mathsf{P}(0)\rangle$ and $|\phi_2\rangle = -i\lambda PJP^\perp R_\lambda^\perp(\beta)P^\perp\,|\mathsf{P}(0)\rangle$. We will write $\phi_j(m) = \langle 0, m, m|\phi_j\rangle$, $j = 1, 2$, for their components. Of course, $\phi_1(m) = |\hat\psi_0(m)|^2$, so by Proposition 5.3, Eq. (82),

$$\sum_{n\geq 0} \left(1 + n^{2s}\right)\!\Big\langle 0, n, n\Big|P\frac{\beta}{L_\lambda + \beta}P\Big|\phi_1\Big\rangle \asymp \sum_m \phi_1(m)(1 + m^{2s} + \lambda^s\beta^{-s/2})$$

$$\asymp \|\psi_0\|_{H^s}^2 + \lambda^s\beta^{-s/2}\|\psi_0\|_{L^2}^2. \tag{94}$$

By similar reasoning,

$$\left|\sum_{n\geq 0} \left(1 + n^{2s}\right)\!\Big\langle 0, n, n\Big|P\frac{\beta}{L_\lambda + \beta}P\Big|\phi_2\Big\rangle\right| \lesssim \sum_m |\phi_2(m)|(1 + m^{2s} + \lambda^s\beta^{-s/2})$$

$$\lesssim \|\phi_2\|_{\ell_w^1} + \lambda^s\beta^{-s/2}\|\phi_2\|_{\ell^1}, \tag{95}$$

where w is as in Lemma 5.4 (we will also use the W notation found there). By this lemma and the fact that $\langle 0, n, m | \mathsf{P}(0) \rangle = \hat{\psi}_0(m)\overline{\hat{\psi}_0(n)}$,

$$
\begin{aligned}
\|\phi_2\|_{\ell^1_w} &\lesssim \lambda \|\mathsf{P}(0)\|_{\ell^1_W} \\
&\lesssim \lambda \sum_{n,m} \frac{1}{1+|n-m|^2} \left\{ w(m)|\hat{\psi}_0(m)|^2 \right\}^{1/2} \left\{ w(n)|\hat{\psi}_0(n)|^2 \right\}^{1/2} \\
&\lesssim \lambda \|\psi_0\|^2_{H^s},
\end{aligned}
$$

where the last inequality used the ℓ^2-boundedness of the matrix with entries $(1 + |n - m|^2)^{-1}$. In the particular case that $s = 0$, the above argument says

$$
\|\phi_2\|_{\ell^1} \lesssim \lambda \|\psi_0\|^2_{L^2}.
$$

Substituting these two estimates for ϕ_2 into (95) gives

$$
\left| \sum_{n \geq 0} (1 + n^{2s}) \left\langle 0, n, n \left| P \frac{\beta}{L_\lambda + \beta} P \right| \phi_2 \right\rangle \right| \lesssim \lambda \|\psi_0\|^2_{H^s} + \lambda^{s+1} \beta^{-s/2} \|\psi_0\|^2_{L^2}. \quad (96)
$$

The combination of (94) and (96) shows that (90) does indeed hold once λ is chosen sufficiently small.   □

References

1. Abramowitz, M., Stegun, I.A.: *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. Superintendent of Documents, Washington, DC: U.S. Government Printing Office, 1964
2. Agmon, S.: *Lectures on Exponential Decay of Solutions of Second-order Elliptic Equations: Bounds on Eigenfunctions of N-body Schrödinger Operators*. Princeton, NJ: Princeton Univ. Press, 1982
3. Bourgain, J.: Growth of Sobolev norms in linear Schrödinger equations with quasi-periodic potential. Commun. Math. Phys. **204**(1), 207–247 (1999)
4. Bourgain, J.: On growth of Sobolev norms in linear Schrödinger equations with smooth time dependent potential. J. Anal. Math. **77**, 315–348 (1999)
5. Bourgain, J.: *Global Solutions of Nonlinear Schrödinger Equations*. Providence, RI: Am. Math. Soc., 1999
6. Bourgain, J.: *On long-time behaviour of solutions of linear Schrödinger euations with smooth time-dependent potential*. Preprint 2002
7. Combes, J.M., Thomas, L.: Asymptotic behaviour of eigenfunctions for multiparticle Schrödinger operators. Commun. Math. Phys. **34**, 251–270 (1973)
8. Deift, P.A.: Applications of a commutation formula. Duke Math. J. **45**(2), 267–310 (1978)
9. Doob, J.L.: *Stochastic Processes*. New York: Wiley, 1953
10. Gardiner, C.W.: *Handbook of Stochastic Methods*. Second edition, Berlin: Springer, 1985
11. Pillet, C.-A.: Some results on the quantum dynamics of a particle in a Markovian potential. Commun. Math. Phys. **102**(2), 237–254 (1985)
12. Tcheremchantsev, S.: Markovian Anderson model. C. R. Acad. Sci. Paris Sér. I Math. **324**(8), 907–912 (1997)
13. Tcheremchantsev, S.: Markovian Anderson model: Bounds for the rate of propagation. Commun. Math. Phys. **187**(2), 441–469 (1997)
14. Tcheremchantsev, S.: Transport properties of Markovian Anderson model. Commun. Math. Phys. **196**(1), 105–131 (1998)