Diffusion of Power in Randomly Perturbed Hamiltonian Partial Differential Equations

E. Kirr * and M.I. Weinstein †

October 10, 2003

Abstract

Abstract We study the evolution of the energy (mode-power) distribution for a class of randomly perturbed Hamiltonian partial differential equations and derive *master equations* for the dynamics of the expected power in the discrete modes. In the case where the unperturbed dynamics has only discrete frequencies (finitely or infinitely many) the mode-power distribution is governed by an equation of discrete diffusion type for times of order $\mathcal{O}(\varepsilon^{-2})$. Here ε denotes the size of the random perturbation. If the unperturbed system has discrete and continuous spectrum the mode-power distribution is governed by an equation of discrete diffusion-damping type for times of order $\mathcal{O}(\varepsilon^{-2})$. The methods involve an extension of the authors' work on deterministic periodic and almost periodic perturbations, and yield new results which complement results of others, derived by probabilistic methods.

Contents

*Department of Mathematics, University of Chicago

†Mathematical Sciences Research, Bell Laboratories - Lucent Technologies, Murray Hill, NJ

1 Introduction

The evolution of an arbitrary initial condition of linear autonomous Hamiltonian partial differential equation (Schrödinger equation),

$$i\partial_t \phi = H_0 \phi, \tag{1}$$

where H_0 is self-adjoint operator, can be studied by decomposing the initial state in terms of the eigenstates (bound modes) and generalized eigenstates (radiation or continuum modes) of H_0. The mode amplitudes evolve independently according to a system of decoupled ordinary differential equations and the energy or power in each mode, the square of the mode amplitude, is independent of time. If the system (1) is perturbed

$$i\partial_t \phi = (H_0 + W(t))\phi, \tag{2}$$

where $W(t)$ respects the Hamiltonian structure ($W^* = W$), then the system of ordinary differential equations typically becomes an infinite coupled system of equations, so-called coupled mode equations. If $W(t)$ has general time-dependence (periodic, almost periodic, random,...), the solutions of the coupled mode equations can exhibit very complex behavior. Of fundamental importance is the question how the mode-powers evolve with t. Kinetic equations, which govern their evolution are called *master equations* [25], [5] and go back to the work of Pauli [20]. A general approach to stochastic systems is presented in [17, 19, 18, 13]; see also [1, 7, 8]. Master equations have been derived in many contexts in statistical mechanics, ocean acoustics and optical wave-propagation in waveguides.

We present a theory of power evolution for (2), for a class of perturbations, $W(t)$, which are random in t. Our theory handles the case where H_0 has spectrum consisting of bound states (finitely or infinitely many discrete eigenvalues) and radiation modes (continuous spectrum). It is a natural extension of the analysis in our work on deterministic periodic, almost periodic and nonlinear systems; see, for example, [9, 11, 10, 24]. Our approach is complementary to the probablistic approach of [7, 8, 19, 18, 13]. The model we consider is well-suited to the study of the effects of an "engineered" perturbation of the system, *e.g.* a prescribed train of light pulses incident on an atomic system, or prescribed distribution of defects encountered by waves propagating along a waveguide; see below. We also give very detailed information on the energy transfer between the subsystems governed by discrete "oscillators" and continuum "radiation field".

In particular, we study the problem

$$i\partial_t \phi \ = \ (\ H_0 \ + \ \varepsilon g(t)\beta \ ) \, \phi, \tag{3}$$

where ε is small, and H_0 and β are self-adjoint operator on the Hilbert space $\mathcal{H}$. H_0 is assumed to support finitely or infinitely many bound states. For example, $H_0 = -\Delta + V(x)$, where V is smooth and sufficiently rapidly decaying as $|x| \to \infty$. β is assumed to be bounded. $g(t)$ is a real valued function of the form of a sequence of short-lived perturbations or "defects"; see figure 1. Our methods can treat the case of more general perturbations, *e.g.* $W(t,x) = \beta(t,x)$, but to simplify the presentation we consider the separable case $W(t) = g(t)\beta(x)$.

Models of the above type arise natural in many contexts. Among them are the interaction between an atom and a train of light pulses [22, and references therein] , a field of great current interest in the control of quantum systems. Such trains of localized perturbations also model sequences of localized defects along waveguides, see [15], [16], introduced by accident or design.

We construct $g(t)$ as follows. Start with $g_0(t)$, a fixed real-valued function with support contained in the interval $[0,T]$ and let $\{d_j\}_{j\geq 0}$ be a nonnegative sequence. Define

$$g(t) \ = \ \sum_{n=0}^{\infty} \ g_0(t - t_n), \ \text{ where} \tag{4}$$

$$\begin{aligned} t_0 \ &= \ d_0 \\ t_n \ &= \ (d_0 + T) + (d_1 + T) + \ldots + (d_{n-1} + T) + d_n, n \geq 1 \end{aligned} \tag{5}$$

denotes the onset of the n^{th} defect.

Note that, if the sequence $\{d_j\}_{j\geq 0}$ is periodic then $g(t)$ is periodic. In this case, the system (3) has already been analyzed by time-independent methods [26] or, more recently and under less restrictive hypothesis, in [9, 11]. For $\{d_j\}_{j\geq 0}$ quasiperiodic or almost periodic (see [2, 4] for a definition) the situation is more delicate. In [11] we treat a general class of almost periodic perturbations of the form:

$$W(t) \ = \ \sum_{j\geq 0} \cos(\mu_j t)\beta_j, \tag{6}$$

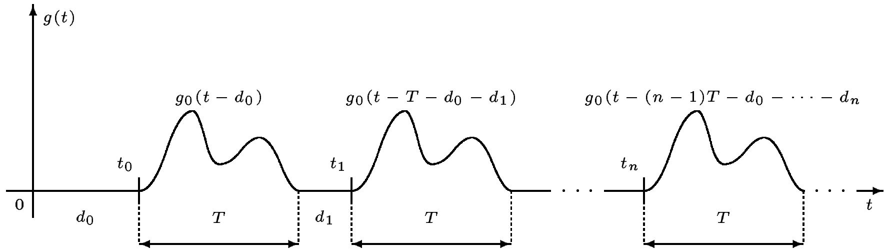

Figure 1: Train of short lived perturbations or "defects". The onset time for the n^{th} defect, t_n, is given by (5).

with appropriate "small denominator" hypotheses on the frequencies $\{\mu_j\}$. We leave it for a future paper [10] to consider the case of almost periodic $\{d_j\}_{j\geq 0}$ and to explore the connection with the results in [11]. We note that a particular case has already been treated in [12, Appendix E].

Note that in [1] and [17] the numbers $d_0, d_1, \ldots,$ are equal to a fixed constant and $g_0(t)$ is random while in our model $d_0, d_1, \ldots,$ are random and $g_0(t)$ is fixed. This is another sense, in which our results complement those in the existing literature.

The paper is divided in two parts. The first part treats stochastic perturbations of Hamiltonian systems with discrete frequencies and then second part extends these results to the case where the unperturbed system has discrete and continuous frequencies. The stochastic perturbation is of order ε and then the vector $P(\tau) \in \ell^1$, whose components are the expected values of the squared discrete mode amplitudes (mode-powers), satisfies on time scales $t = \mathcal{O}(\varepsilon^{-2})$ or equivalently $\tau = \mathcal{O}(1)$, the master equations of diffusion or diffusion-damping type. Specifically, if H_0 has only discrete spectrum (finite or infinite) then

$$\partial_\tau P(\tau) = -BP(\tau), \quad B \geq 0 \tag{7}$$

which has the character of a discrete diffusion equation, *i.e.*

$$\sum_k P_k(\tau) \ = \ \sum_k P_k(0), \quad \frac{d}{d\tau} P \cdot P = -\langle P, BP \rangle \ \leq 0. \tag{8}$$

If H_0 has both discrete and continuous spectra, then

$$\partial_\tau P(\tau) = (-B - \Gamma)P(\tau), \quad B \geq 0, \ \Gamma = \mathrm{diag}(\gamma_k) > 0 \tag{9}$$

for which

$$\sum_k P_k(\tau) \ \leq \ e^{-\gamma\tau} \sum_k P_k(0), \tag{10}$$

where $\gamma = \min_k \ \gamma_k$.

In sections 2 and 3 we study (3) under the hypothesis that H_0 has no continuous spectrum (i.e. no radiation modes) and in section 4 we generalize to the case where H_0 has discrete and continuous spectrum. In section 2 we present the main hypotheses on H_0 and $g_0(t)$ and study the effect of a single short lived perturbation. In section 3 we present our hypotheses on $d_0, d_1, \ldots,$ and analyze the effect of a train of perturbations (3-4). We show that if $d_0, d_1, \ldots,$ are independent random variables with certain distributions, see Hypothesis **(H4)** and Examples 1 and 2, diffusion occurs in the expected value for the powers of the modes. Specifically, if we start with energy in one mode, then, on a time scale of order $1/\varepsilon^2$, one can expect the energy to be distributed among all the modes. In section 4 we analyze equation (3) under the hypothesis that H_0 has both discrete and continuous spectrum (i.e. supports both bound modes and radiation modes). We prove a result similar to the nonradiative case but now bound state- wave resonances lead to loss of power. The effect of our randomly distributed deterministic perturbation is very similar to the one induced by purely stochastic perturbations, see [1, 13, 19], but quite different from the effects of time almost periodic perturbations, see [9, 11]. Section 5 is dedicated to such comparisons.

Notation

1) $\langle x \rangle \ = \ \sqrt{1+x^2}$

2) Fourier Transform:

$$\hat{g}(\xi) \ = \int_{-\infty}^{\infty} e^{-i\xi t} g(t) \ dt \tag{11}$$

3) We write $\zeta + c.c.$ to mean $\zeta + \bar{\zeta}$, where $\bar{\zeta}$ denotes the complex conjugate of ζ.

4) w' denotes the transpose of w.

5) $\lfloor q \rfloor$ denotes the integer part of q.

Acknowledgement: We would like to thank G. C. Papanicolaou and J. L. Lebowitz for helpful discussions concerning this work. E.K. was supported in part by the ASCI Flash Center at the University of Chicago.

2 Short lived perturbation of a system with discrete frequencies

In this section we consider the perturbed dynamical system

$$i\partial_t \phi(t) = H_0 \phi(t) + \varepsilon g_0(t) \beta \phi(t, x), \tag{12}$$

where H_0 has only discrete spectrum and $g_0(t)$ is a short-lived (compactly supported) function. We study the effect of this perturbation on the distribution of energy among the modes of H_0. Here and in section 4 we are extending the results in [23] to multiple bound states but under an additional "incoherence" assumption; see (18).

Hypotheses on H_0, β and $g_0(t)$

(H1) H_0 is a self adjoint operator on a Hilbert space $\mathcal{H}$. It has a pure point spectrum formed by the eigenvalues : $\{\lambda_j\}_{j\geq 1}$ with a complete set of orthonormal eigenvectors: $\{\psi_j\}_{j\geq 1}$:

$$H_0\psi_j = \lambda_j\psi_j \ , \ \langle\psi_i,\psi_j\rangle \ = \ \delta_{ij} \tag{13}$$

(H2) β is a bounded self adjoint operator on $\mathcal{H}$ and satisfies $\|\beta\| = 1$.

(H3) $g_0(t) \in L^2(\mathbb{R})$ is real valued, has compact support contained in $[0,T]$ on the positive real line and its L^1-norm, denoted by $\|g_0\|_1$ is 1. Thus its Fourier transform has L^∞-norm bounded by 1.

Note that one can always take $\|\beta\| = 1$ and $\|g_0\|_1 = 1$ by setting $\varepsilon = \|g_0\|_1 \cdot \|\beta\|$, thus incorporating the size of $g_0\beta$ in ε. Therefore, under assumptions (H2-H3), ε in (12) measures the actual size of the perturbation in the $L^1(\mathbb{R},\mathcal{H})$ norm. Our results are perturbative in ε and are valid for ε sufficiently small.

By the standard contraction method one can show that (12) has an unique solution $\phi(t) \in \mathcal{H}$ for all $t \in \mathbb{R}$. Moreover, because both H_0 and $g_0(t)\beta$ are self adjoint operators, we have for all $t \in \mathbb{R}$:

$$\|\phi(t)\| = \|\phi(0)\|. \tag{14}$$

We can write $\phi(t)$ as a sum of projections onto the complete set of orthonormal eigenvectors of H_0 :

$$\phi(t,x) = \sum_j a_j(t)\psi_j(x), \tag{15}$$

By Parseval's relation

$$\sum_j |a_j(t)|^2 = \|\phi(t)\|^2 \equiv \|\phi(0)\|^2 \tag{16}$$

Now (12) can be rewritten as

$$i\partial_t a_k(t) = \lambda_k a_k(t) + \varepsilon g_0(t) \sum_j a_j(t)\langle\psi_k, \beta\psi_j\rangle, \ \ k \in \{1, 2 \ldots\} \tag{17}$$

where $\langle\cdot,\cdot\rangle$ denotes the scalar product in $\mathcal{H}$.

Hence the equation (12) is equivalent to a weakly coupled linear system in the amplitudes: $a_1, a_2, \ldots,$ (17).

Since the perturbation size is ε we expect, in general, that the change in energy in the k^{th} mode, $|a_k(t)|^2 - |a_k(0)|^2$, to be of order ε. However with a suitable random initial condition we can prove more subtle behavior.

Suppose that there exists an averaging procedure applicable to the amplitudes: $a_1, a_2, \ldots$ of the solutions of (12), denoted by

$$a(t) \mapsto \mathbb{E}(a(t)) \in \mathbb{C}.$$

We now state a fundamental result, applied throughout this paper, for a single defect which is compactly supported in time:

Theorem 2.1. *Assume the conditions* **(H1)-(H3)** *hold and the initial values for (12) are such that*

$$\mathbb{E}\left(a_j(0)\overline{a_k(0)}\right) = 0 \qquad \text{whenever } j \neq k. \tag{18}$$

Then for all $t > \sup\{s \in \mathbb{R} \mid g_0(s) \neq 0\}$ *and* $k \in \{1, 2, \ldots\}$ *we have*

$$P_k(t) - P_k(0) = \varepsilon^2 \sum_j |\alpha_{kj}|^2 |\hat{g}_0(-\Delta_{kj})|^2 (P_j(0) - P_k(0)) + \mathcal{O}(\varepsilon^3), \tag{19}$$

where

$$P_k(t) \equiv \mathbb{E}\left(|a_k(t)|^2\right)$$

denotes the average power in the k^{th}*-mode at time* t*,* $\alpha_{kj} \equiv \langle \psi_k, \beta\psi_j \rangle$*,* $\hat{g}_0$ *denotes the Fourier transform of* g_0 *and* $\Delta_{kj} \equiv \lambda_k - \lambda_j$*.*

Note that (19) can be written in the form:

$$P_k(t) = T_\varepsilon P_k(0) + \mathcal{O}(\varepsilon^3), \tag{20}$$

where

$$T_\varepsilon = \mathbb{I} - \varepsilon^2 B; \qquad B \geq 0 \tag{21}$$

$\mathbb{I}$ is the identity operator (matrix) and B is given by

$$B = (b_{kj})_{1 \leq k,j}, \qquad b_{kj} = \begin{cases} -|\alpha_{kj}|^2 |\hat{g}_0(-\Delta_{kj})|^2, & \text{for } j \neq k, \\ \sum_{l, l \neq k} |\alpha_{kl}|^2 |\hat{g}_0(-\Delta_{kl})|^2, & \text{for } j = k \end{cases} \tag{22}$$

In section 3 we will discuss and use the properties of B and T_ε.

Proof of Theorem 2.1. In the amplitude system, (17), we remove the fast oscillations by letting

$$a_k(t) = e^{-i\lambda_k t} A_k(t), \tag{23}$$

Note that by (16)

$$\sum_j |A_j(t)|^2 \equiv \|\phi(0)\|^2 \tag{24}$$

Now (17) becomes

$$i\partial_t A_k(t) = \varepsilon g_0(t) \sum_j \alpha_{kj} e^{i\Delta_{kj}t} A_j(t), \tag{25}$$

where

$$\Delta_{kj} \equiv \lambda_k - \lambda_j, \tag{26}$$

$$\alpha_{kj} \equiv \langle \psi_k, \beta\psi_j \rangle = \overline{\alpha}_{jk}. \tag{27}$$

The above system leads to the following one in product of amplitudes, $A_k(t)\overline{A}_l(t)$:

$$\begin{aligned} \partial_t(A_k(t)\overline{A}_l(t)) &= i\varepsilon g_0(t) \sum_j \alpha_{jl} e^{i\Delta_{jl}t} A_k(t)\overline{A}_j(t) \\ &- i\varepsilon g_0(t) \sum_j \alpha_{kj} e^{i\Delta_{kj}t} A_j(t)\overline{A}_l(t), \end{aligned} \tag{28}$$

In the particular case $k = l$ we have the power equation for each mode:

$$\partial_t |A_k(t)|^2 = i\varepsilon g_0(t) \sum_j \alpha_{jk} e^{i\Delta_{jk}t} A_k(t)\overline{A}_j(t) + c.c.\ . \tag{29}$$

Note that the sum in (29) commutes with time integral and expected value operators. This is due to (24) and the dominant convergence theorem, see for example [6]. Indeed consider

$$f_m(t) = \sum_{j=1}^{m} \alpha_{jk} e^{i\Delta_{jk}t} A_k(t)\overline{A}_j(t) g_0(t).$$

From (15) we have for all $t \in \mathbb{R}$

$$\lim_{m\to\infty} f_m(t) = \langle \phi(t), \beta\psi_k \rangle a_k(t) g_0(t).$$

From (24) and the Cauchy-Schwarz inequality $|\langle a, b\rangle| \leq \|a\|\ \|b\|$, we have for all $t \in \mathbb{R}$

$$|f_m(t)| \leq \|\phi(0)\|^2 |g_0(t)|. \tag{30}$$

The right hand side of (30) is integrable and the dominant convergence theorem applies. A similar argument is valid for expected values. Therefore, from now on, we are going to commute both time integrals and expected values with summations like the one in (29).

We integrate (29) from 0 to $t > \sup\{s \in \mathbb{R} \mid g_0(s) \neq 0\}$ and integrate by parts the right hand side. The result is:

$$\begin{aligned}
|A_k(t)|^2 - |A_k(0)|^2 &= i\varepsilon \sum_j \alpha_{jk} \int_0^t g_0(s) e^{i\Delta_{jk}s} A_k(s)\overline{A}_j(s) + c.c. \\
&= -i\varepsilon \sum_j \alpha_{jk} \int_s^\infty g_0(\tau) e^{i\Delta_{jk}\tau} d\tau A_k(s)\overline{A}_j(s) \,\Big|_{s=0}^{s=t} + c.c. \\
&+ \; i\varepsilon \sum_j \alpha_{jk} \int_0^t \int_s^\infty g_0(\tau) e^{i\Delta_{jk}\tau} d\tau \partial_s \left(A_k\overline{A}_j\right)(s) ds + c.c. \;.
\end{aligned} \tag{31}$$

The boundary terms are

$$\begin{aligned}
&-i\varepsilon \sum_j \alpha_{jk} \int_s^\infty g_0(\tau) e^{i\Delta_{jk}\tau} d\tau A_k(s)\overline{A}_j(s) \,\Big|_{s=0}^{s=t} + c.c. \\
&= \; i\varepsilon \sum_j \alpha_{jk} \hat{g}_0(-\Delta_{jk}) A_k(0)\overline{A}_j(0) + c.c.,
\end{aligned} \tag{32}$$

where $\hat{g}_0$ denotes the Fourier Transform of g_0; see (11). Note that upon taking the average, using (18) and the fact that $\hat{g}_0(0)$ is real, these boundary terms vanish.

Into the last term in (31) we substitute (28):

$$\begin{aligned}
&i\varepsilon \sum_j \alpha_{jk} \int_0^t \int_s^\infty g_0(\tau) e^{i\Delta_{jk}\tau} d\tau \partial_s \left(A_k\overline{A}_j\right)(s) ds = \\
&= \; +|\varepsilon|^2 \sum_{j,p} \alpha_{jk}\alpha_{kp} \int_0^t \int_s^\infty g_0(\tau) e^{i\Delta_{jk}\tau} d\tau g_0(s) e^{i\Delta_{kp}s} A_p(s)\overline{A}_j(s) ds \\
&- \; |\varepsilon|^2 \sum_{j,q} \alpha_{jk}\alpha_{qj} \int_0^t \int_s^\infty g_0(\tau) e^{i\Delta_{jk}\tau} d\tau g_0(s) e^{i\Delta_{qj}s} A_k(s)\overline{A}_q(s) ds.
\end{aligned} \tag{33}$$

We again integrate by parts both terms in (33):

$$\begin{aligned}
&i\varepsilon \sum_j \alpha_{jk} \int_0^t \int_s^\infty g_0(\tau) e^{i\Delta_{jk}\tau} d\tau \partial_s \left(A_k\overline{A}_j\right)(s) ds = \\
&- \; |\varepsilon|^2 \sum_{j,p} \alpha_{jk}\alpha_{kp} \int_u^\infty g_0(s) e^{i\Delta_{kp}s} \int_s^\infty g_0(\tau) e^{i\Delta_{jk}\tau} d\tau ds A_p(u)\overline{A}_j(u) \,\Big|_{u=0}^{u=t} \\
&+ \; |\varepsilon|^2 \sum_{j,q} \alpha_{jk}\alpha_{qj} \int_u^\infty g_0(s) e^{i\Delta_{qj}s} \int_s^\infty g_0(\tau) e^{i\Delta_{jk}\tau} d\tau ds A_k(u)\overline{A}_q(u) \,\Big|_{u=0}^{u=t} \\
&+ \; |\varepsilon|^2 \sum_{j,p} \alpha_{jk}\alpha_{kp} \int_0^t \int_u^\infty \int_s^\infty g_0(\tau) e^{i\Delta_{jk}\tau} d\tau g_0(s) e^{i\Delta_{kp}s} ds \partial_u \left(A_p\overline{A}_j\right)(u) du \\
&- \; |\varepsilon|^2 \sum_{j,q} \alpha_{jk}\alpha_{qj} \int_0^t \int_u^\infty \int_s^\infty g_0(\tau) e^{i\Delta_{jk}\tau} d\tau g_0(s) e^{i\Delta_{qj}s} ds \partial_u \left(A_k\overline{A}_q\right)(u) du.
\end{aligned} \tag{34}$$

Note that the boundary terms calculated at "$u = t$" are zero since $t > \sup\{s \in \mathbb{R} \mid g_0(s) \neq 0\}$. Upon taking the expected value and using (18) the only boundary terms contributing are the ones for which $u = 0$ and $j = p$ in the second row of (34):

$$\begin{aligned}
&\sum_j |\alpha_{kj}|^2 \int_0^\infty \int_s^\infty g_0(\tau) e^{i\Delta_{jk}\tau} d\tau g_0(s) e^{i\Delta_{kj}s} ds \mathbb{E}\left(|A_j(0)|^2\right) + c.c. \\
&= \sum_j |\alpha_{kj}|^2 \mathbb{E}\left(|A_j(0)|^2\right) \cdot 2\Re \int_0^\infty \int_s^\infty g_0(\tau) e^{i\Delta_{jk}\tau} d\tau g_0(s) e^{i\Delta_{kj}s} ds \qquad (35)
\end{aligned}$$

and the ones for which $u = 0$ and $q = k$ in the third row of (34):

$$\begin{aligned}
&\sum_j |\alpha_{kj}|^2 \int_0^\infty \int_s^\infty g_0(\tau) e^{i\Delta_{jk}\tau} d\tau g_0(s) e^{i\Delta_{kj}s} ds \mathbb{E}\left(|A_k(0)|^2\right) + c.c. \\
&= \sum_j |\alpha_{kj}|^2 \mathbb{E}\left(|A_k(0)|^2\right) \cdot 2\Re \int_0^\infty \int_s^\infty g_0(\tau) e^{i\Delta_{jk}\tau} d\tau g_0(s) e^{i\Delta_{kj}s} ds \qquad (36)
\end{aligned}$$

To compute (35-36) we use the lemma:

Lemma 2.1. *If $g_0(t), t \in \mathbb{R}$ is square integrable with compact support included in the positive real line then for all $\lambda \in \mathbb{R}$ the following identity holds*

$$2\Re \int_0^\infty \int_s^\infty g_0(\tau) e^{i\lambda\tau} d\tau g_0(s) e^{-i\lambda s} ds = |\hat{g}_0(-\lambda)|^2.$$

<u>Proof.</u> For any $\lambda \in \mathbb{R}$ we have:

$$\begin{aligned}
&\int_0^\infty \int_s^\infty g_0(\tau) e^{i\lambda\tau} d\tau g_0(s) e^{-i\lambda s} ds \\
&= \lim_{\varepsilon \searrow 0} \int_0^\infty \int_s^\infty g_0(\tau) e^{i(\lambda + i\varepsilon)\tau} d\tau g_0(s) e^{-i\lambda s} ds \\
&= \frac{1}{2\pi} \lim_{\varepsilon \searrow 0} \int_0^\infty g_0(s) e^{-i\lambda s} ds \int_{-\infty}^\infty \hat{g}_0(\mu) \int_s^\infty e^{i(\lambda + \mu + i\varepsilon)\tau} d\tau d\mu \\
&= \frac{i}{2\pi} \lim_{\varepsilon \searrow 0} \int_0^\infty g_0(s) e^{-i\lambda s} ds \int_{-\infty}^\infty \frac{\hat{g}_0(\mu)}{\mu + \lambda + i\varepsilon} e^{is(\mu + \lambda + i\varepsilon)} d\mu \\
&= \frac{i}{2\pi} \lim_{\varepsilon \searrow 0} \int_{-\infty}^\infty \frac{\hat{g}_0(\mu)\hat{g}_0(-\mu)}{\mu + \lambda + i\varepsilon} d\mu \\
&= \frac{1}{2} |\hat{g}_0(-\lambda)|^2 + \frac{i}{2\pi} \text{P.V.} \int_{-\infty}^\infty \frac{|\hat{g}_0(\mu)|^2}{\mu + \lambda} d\mu \qquad (37)
\end{aligned}$$

The last relation in (37) is the Plemelj-Sohotsky's formula for (temperate) distributions:

$$\lim_{\varepsilon \searrow 0} \frac{1}{x + i\varepsilon} = \text{P.V.} \frac{1}{x} - i\pi\delta(x) \stackrel{def}{=} \frac{1}{x + i0}.$$

Note that $\hat{g}_0(\mu) \in C^\infty(\mathbb{R}) \cap L^2(\mathbb{R})$. Since (37) is already decomposed in its real and imaginary part the lemma follows. []

Into the triple integral terms of (34) we again substitute (28). Then one can show that the 1-norm of this correction vector is dominated by $|\varepsilon|^3 \; \|g_0\|_1^3 \; \|\beta\|^3 \; \|\phi(0)\|^2$. Hence, it is of order $\mathcal{O}(|\varepsilon|^3)$.

Thus, after applying Lemma 2.1 to (35-36) and using (31) we arrive at the conclusion of Theorem 2.1. []

3 Diffusion of power in discrete frequency (nonradiative) systems

In the previous section we calculated the effect of a single defect on the the mode-power distribution. In this section we show how to apply this result to prove diffusion of power for the perturbed Hamiltonian system, (2), where $g(t)$ is a random function of the form (4), defined in terms of a random sequence $\{d_j\}_{j\geq 0}$. In particular, the sequence $\{d_j\}_{j\geq 0}$ will be taken to be generated by independent, identically distributed random variables. This will be result in a *mixing the phases* of the complex mode amplitudes, after each defect.

We assume that **(H1-H3)** are satisfied. The following hypothesis ensures that (18) holds before each defect, thus enabling repeated application of Theorem 2.1.

(H4) $d_0, d_1, \ldots$ are independent identically distributed random variables taking only nonnegative values and such that for any $l \in \{0, 1, \ldots\}$ and $j \neq k \in \{1, 2 \ldots\}$ we have

$$\mathbb{E}\left(e^{i(\lambda_j - \lambda_k)d_l}\right) = 0$$

where $\mathbb{E}(\cdot)$ denotes the expected value.

Clearly (H4) requires the eigenvalues to be distinct but aside from these we claim that for any *finitely many, distinct* eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$ there exist a random variable satisfying (H4).

Example 1 (finitely many bound states) Given $\lambda_1, \lambda_2, \ldots, \lambda_m$ distinct choose the random variables $d_l, \; l = 0, 1, \ldots$ to be identically distributed with distribution d :

$$d = \sum_{1\leq j<k\leq m} d_{jk}$$

where d_{jk} are independent random variables such that the distribution of d_{jk} is uniform on the interval $[0, 2\pi/|\lambda_j - \lambda_k|]$. In this case, for any $j' \neq k' \in \{1, 2, \ldots\}$

$$\begin{aligned} \mathbb{E}\left(e^{i(\lambda_{j'}-\lambda_{k'})d}\right) &= \mathbb{E}\left(\prod_{1\le j<k} e^{i(\lambda_{j'}-\lambda_{k'})d_{jk}}\right) \\ &= \prod_{1\le j<k} \mathbb{E}\left(e^{i(\lambda_{j'}-\lambda_{k'})d_{jk}}\right) \\ &= 0 \end{aligned}$$

since $\mathbb{E}\left(e^{i(\lambda_{j'}-\lambda_{k'})d_{j'k'}}\right) = 0$.

Another choice is to consider discrete d_{jk}'s. Namely, take d_{jk} to be the discrete random variable taking each of the values 0 and $\pi/|\lambda_j - \lambda_k|$ with probability 1/2. A concrete example is, in the case we have three eigenvalues $\lambda_1 < \lambda_2 < \lambda_3$, to choose d to be the random variable taking each of the eight values:

$$0, \frac{\pi}{\lambda_2-\lambda_1}, \frac{\pi}{\lambda_3-\lambda_1}, \frac{\pi}{\lambda_3-\lambda_2},$$
$$\pi\left(\frac{1}{\lambda_2-\lambda_1}+\frac{1}{\lambda_3-\lambda_1}\right), \pi\left(\frac{1}{\lambda_2-\lambda_1}+\frac{1}{\lambda_3-\lambda_2}\right), \pi\left(\frac{1}{\lambda_3-\lambda_1}+\frac{1}{\lambda_3-\lambda_2}\right),$$
$$\pi\left(\frac{1}{\lambda_2-\lambda_1}+\frac{1}{\lambda_3-\lambda_1}+\frac{1}{\lambda_3-\lambda_2}\right)$$

with probability 1/8.

(H4) does not restrict us to system with finitely many bound states:

Example 2 (infinitely many bound states) Let the quantum harmonic oscillator in one dimension:

$$H_0 = -\frac{\hbar^2}{2}\partial_x^2 + \omega^2 x^2, \qquad x \in \mathbb{R},$$

be the unperturbed Hamiltonian. Then

$$\lambda_n = \hbar\omega(n + 1/2), \qquad n = 0, 1, 2, \ldots,$$

see for example [14]. Note that (H4) holds provided that we choose d_l, $l = 0, 1, \ldots$ to be identically and uniformly distributed on the interval $[0, 2\pi/(\hbar\omega)]$.

Note on degenerate eigenvalues: As discussed above (H4) cannot be satisfied in the case H_0 admits degenerate eigenvalues. However, at least in some cases, our theory can be applied. In general the degeneracy is a consequence of the symmetries of H_0, i.e. the existence of a self-adjoint operator, say L, commuting with H_0, $[L, H_0] = 0$. To recover our results it is sufficient to assume that β, the "space-like" part of the perturbation, respects the symmetry, i.e. commutes with L. One can now factor out L, i.e. work on the invariant subspaces of L where H_0 is nondegenerate. Along the lines of Example 2 one can consider the quantum harmonic oscillator in three dimensions which has a spherically symmetric Hamiltonian and degenerate eigenvalues, see for example [14]. If β is spherically symmetric then it only couples bound states with the same angular momentum. Hence the problem reduces to subsystems consisting of bound states with the same angular momentum but different energy, therefore nondegenerate. The choice we made in Example 2 will satisfy (H4) in each of the subsystems.

3.1 Power diffusion after a fixed (large) number of defects

Theorem 3.1. *Consider equation (12) with g of the form (4). Assume* **(H1-H4)** *hold. Then the expected value of the power vector after passing a fixed number of perturbations "n" satisfies*

$$P^{(n)} = T_\varepsilon^n P(0) + \mathcal{O}(n\varepsilon^3), \tag{38}$$

where T_ε is given in (21)

$$\begin{aligned} &P_k^{(n)} = \mathbb{E}(|a_k(t)|^2), \ \ k = 1, 2, \ldots \\ &t_{n-1} + T \le t \le t_n, \ \ (t \text{ ranging between the } n^{th} \text{ and } (n+1)^{st} \text{ defects} \end{aligned} \tag{39}$$

Proof. We will prove the theorem by induction on $n \ge 0$, the number of defects traversed. For $n = 0$ the assertion is obvious. Suppose now that for $n \ge 0$ we have

$$P^{(n)} = T_\varepsilon^n P(0) + \mathcal{O}(n\varepsilon^3). \tag{40}$$

We will show

$$P^{(n+1)} = T_\varepsilon^{n+1} P(0) + \mathcal{O}\left((n+1)\varepsilon^3\right) \tag{41}$$

by applying Theorem 2.1 to (40). In order to apply Theorem 2.1 we need to verify that (18) is satisfied before the $n+1^{\text{st}}$ defect. Specifically, we must verify that for any pair $k \ne j$

$$\mathbb{E}\left(a_k(t_{n+1})\bar{a}_j(t_{n+1})\right) \ = \ \mathbb{E}\left(a_k\left(nT + \textstyle\sum_{k=0}^{n+1} d_k\right) \bar{a}_j\left(nT + \textstyle\sum_{k=0}^{n+1} d_k\right)\right) = 0. \tag{42}$$

Using the fact that d_{n+1} is independent of $d_0 + d_1 + \ldots + d_n$, and **(H4)** we have:

$$\begin{aligned} \mathbb{E}\left(a_k\overline{a}_j\left(nT + \textstyle\sum_{k=0}^{n+1} d_k\right)\right) &= \mathbb{E}\left(a_k\overline{a}_j\left(nT + \textstyle\sum_{k=0}^{n} d_k\right) e^{i(\lambda_j - \lambda_k)d_{n+1}}\right) \\ &= \mathbb{E}\left(a_k\overline{a}_j\left(nT + \textstyle\sum_{k=0}^{n} d_k\right)\right) \mathbb{E}\left(e^{i(\lambda_j - \lambda_k)d_{n+1}}\right) = 0. \end{aligned}$$

Thus (42) holds and all the hypothesis of Theorem 2.1 are now satisfied. By applying it and using (40) we have

$$\begin{aligned} P^{(n+1)} &= T_\varepsilon P^{(n)} + \mathcal{O}(\varepsilon^3) \\ &= T_\varepsilon \left(T_\varepsilon^n P(0) + \mathcal{O}(n\varepsilon^3) \right) + \mathcal{O}(\varepsilon^3) \\ &= T_\varepsilon^{n+1} P(0) + \mathcal{O}((n+1)\varepsilon^3). \end{aligned}$$

Hence (40) implies (41). This concludes the induction step and the proof of Theorem 3.1 is now complete. []

In the next two Corollaries we describe the asymptotic behavior of the vector of expected powers when the number of defects n tends to infinity. Note that after a possible reordering of the eigenvectors $\psi_1, \psi_2, \ldots,$ of H_0, the operator B given by (22) might look like[1]:

$$B = \operatorname{diag}\left[B_1, B_2, \ldots, B_q, \ldots\right], \tag{43}$$

where $B_1, B_2, \ldots, B_q, \ldots$ are square matrices (linear operators) of dimensions $m_1, m_2, \ldots, m_q, \ldots,$ $1 \leq m_q \leq \infty,\ q = 1, 2, \ldots$. In linear algebra terms this means that B is reducible. In terms of the dynamical system (38) generated by $T_\varepsilon = \mathbb{I} - \varepsilon^2 B$ it means that, after a possible reordering, the first m_1 bound states of H_0 are isolated from the rest. The same is valid for the next m_2 bound states, etc. To understand the evolution of the full system it is sufficient to analyze each of the isolated subsystems separately. They all evolve according to (38) with $T_\varepsilon = \mathbb{I} - \varepsilon^2 B_q$ and B_q given by (22) but the indices span only a subset of the eigenvectors $\psi_1, \psi_2, \ldots$ of H_0. The main difference is that now B_q is irreducible. In what follows we are focusing on one such subsystem and drop the index q.

Corollary 3.1. *If the subsystem has a finite number of bound states, say m, then*

$$\lim_{n\to\infty} P^{(n)} = \begin{cases} P(0), & \text{if } n \ll \varepsilon^{-2} \\ e^{-B\tau} P(0) & \text{if } n = \tau\varepsilon^{-2} \\ \frac{E}{m}(1, 1, \ldots, 1)', & \text{if } \varepsilon^{-2} \ll n \ll |\varepsilon|^{-3} \end{cases}, \tag{44}$$

where $E = P_1(0) + P_2(0) + \ldots + P_m(0)$ is the expected total power in the subsystem and it is conserved.

<u>Proof.</u> We use the following properties of the irreducible matrix B:

(B1) B is self adjoint and $B \geq 0$;

[1] For such a decomposition to occur it is sufficient that H_0 and β have common invariant subspaces $\mathcal{H}_1 \subset \mathcal{H}, \mathcal{H}_2 \subset \mathcal{H}, \ldots, \mathcal{H}_q \subset \mathcal{H}, \ldots$

(B2) 0 is a simple eigenvalue for B with corresponding normalized eigenvector

$$r_0 = \frac{1}{\sqrt{m}}\,(1,1,\ldots,1)'. \tag{45}$$

These properties are proved in the Appendix.

Let $\beta_0 = 0, \beta_1, \beta_2, \ldots, \beta_{m-1}$ be the eigenvalues of B counting multiplicity, and let $r_0, r_1, \ldots, r_{m-1}$ be the corresponding orthonormalized eigenvectors. By (B1) and (B2) $\beta_1, \beta_2, \ldots, \beta_{m-1}$ are strictly positive. Let

$$R = [r_0, r_1, \ldots, r_{m-1}]$$

be the matrix whose columns are orthonormalized eigenvectors of B and let R' be its transpose. Then

$$\begin{aligned} R'BR &= \operatorname{diag}[\beta_0, \beta_1, \beta_2, \ldots, \beta_{m-1}] \\ R'R &= \mathbb{I} = RR'. \end{aligned}$$

It follows that

$$\begin{aligned} T_\varepsilon^n &= \left(\mathbb{I} - \varepsilon^2 B\right)^n = R\left[R'\left(\mathbb{I} - \varepsilon^2 B\right)R\right]^n R' \\ &= R\operatorname{diag}\left[(1-\varepsilon^2\beta_0)^n, (1-\varepsilon^2\beta_1)^n, \ldots, (1-\varepsilon^2\beta_{m-1})^n\right] R'. \end{aligned}$$

We now study $\lim_{n\to\infty} T_\varepsilon^n$ for the three asymptotic regimes of (44). Note that for $0 \le k \le m-1$ we have:

$$\begin{aligned} \lim_{n\to\infty, \varepsilon^2 n\to 0} (1-\varepsilon^2\beta_k)^n &= 1 \\ \lim_{n\to\infty, \varepsilon^2 n = \tau} (1-\varepsilon^2\beta_k)^n &= e^{-\beta_k\tau} \\ \lim_{n\to\infty, \varepsilon^2 n\to \infty} (1-\varepsilon^2\beta_k)^n &= 0, \ \beta_k > 0 \\ \lim_{n\to\infty, \varepsilon^2 n\to \infty} (1-\varepsilon^2\beta_k)^n &= 1, \ \beta_k = 0 \end{aligned} \tag{46}$$

Consequently,

$$\lim_{n\to\infty} T_\varepsilon^n = \begin{cases} R\operatorname{diag}[1,1,\ldots,1]R' = \mathbb{I} & \text{if } n \ll \varepsilon^2 \\ R\operatorname{diag}\left[e^{-\beta_0\tau}, e^{-\beta_1\tau}, \ldots, e^{-\beta_{m-1}\tau}\right] = e^{-B\tau} & \text{if } n = \tau\varepsilon^{-2} \\ R\operatorname{diag}[1,0,0,\ldots,0]R' = \text{projection onto } r_0 & \text{if } \varepsilon^{-2} \ll n \ll |\varepsilon|^{-3} \end{cases}, \tag{47}$$

where r_0 is defined in (45).

Substitution of (47) into (38) completes the proof of Corollary 3.1. []

Corollary 3.2. *If the subsystem has an infinite number of bound states, then*

$$\lim_{n\to 0} P^{(n)} = \begin{cases} P(0), & \text{if } n \ll \varepsilon^{-2} \\ e^{-B\tau}P(0) & \text{if } n = \tau\varepsilon^{-2} \end{cases}, \tag{48}$$

For $n \gg \varepsilon^{-2}$ the limit in ℓ^2 is 0, while the limit in ℓ^1 does not exist. More precisely, although the total power in the subsystem is conserved,

$$\sum_{k=1}^{\infty} P_k^{(n)} \;=\; E, \qquad \forall n \geq 0, \tag{49}$$

$\{P^{(n)}\}$ does not converge in ℓ^1 due to an energy transfer to the high modes. In particular, for any fixed $N \geq 1$:

$$\begin{aligned} \lim_{n\to\infty} \sum_{k=N}^{\infty} P_k^{(n)} &= E, \\ \lim_{n\to\infty} \sum_{k=1}^{N} P_k^{(n)} &= 0. \end{aligned} \tag{50}$$

We note that similar results have been obtained in [1] but for different types of random perturbation.

Corollaries 3.1 and 3.2 show that, on time scales of order $1/\varepsilon^2$, the dynamical system is equivalent with

$$\partial_\tau P(\tau) = -BP(\tau). \tag{51}$$

Moreover the definition of $-B$ in (22) together with $-B \leq 0$ and e^{-B} unitary on ℓ^1 implies that the flow (51) is very much like that of a discrete heat or diffusion equation.

In conclusion the number of defects encountered should be comparable with $1/\varepsilon^2$ to have a significant effect. Once they are numerous enough, the defects diffuse the power in the system. If the number of defects is much larger than $1/\varepsilon^2$ the power becomes uniformly distributed among the bound states.

Remark 3.1. *Hyptothesis* **(H4)** *is important. If we do not assume* **(H4)** *then the correction term for each defect is of size ε, since the boundary terms (32) no longer vanish. Consequently the correction term in the main result (38) is $\mathcal{O}(n\varepsilon)$ which on the "diffusion time scale" $n \sim \varepsilon^{-2}$ is very large.*

Proof of Corollary 3.2 In the case of an infinite number of bound states B has the following properties, see the Appendix:

(B1$_\infty$) B is a nonnegative, bounded self adjoint operator on ℓ^2 with spectral radius less or equal to 2;

(B2$_\infty$) 0 is not an eigenvalue for B;

(B3$_\infty$) B is a bounded operator on ℓ^1 with norm $\|B\|_1 \leq 2$;

(B4$_\infty$) For $|\varepsilon| \leq 1$ the operator $T_\varepsilon = (\mathbb{I} - \varepsilon^2 B)$ transforms positive vectors (i.e. all components positive) into positive vectors and conserves their ℓ^1 norm.

We are going to focus first on ℓ^2 results. Based on the spectral representation theorem, see [21], we have for any Borel measurable real function f :

$$f(B) = \int_0^2 f(s) d\mu(s). \tag{52}$$

Here $d\mu(s)$ is the spectral measure induced by B. Note that B2$_\infty$ implies the continuity of $\mu(s)$ at zero.

Now

$$T_\varepsilon^n = \left(\mathbb{I} - \varepsilon^2 B\right)^n = \int_0^2 \left(1 - \varepsilon^2 s\right)^n d\mu(s)$$

and

$$\lim_{n\to\infty} T_\varepsilon^n = \lim_{\varepsilon\to 0} \int_0^2 \left(1 - \varepsilon^2 s\right)^n d\mu(s) = \int_0^2 \lim_{\varepsilon\to 0} \left(1 - \varepsilon^2 s\right)^n d\mu(s). \tag{53}$$

For the last equality we used the dominant convergence theorem with $|1 - \varepsilon^2 s|^n \leq 1$ for $0 \leq s \leq 2$, $|\varepsilon| \leq 1$ and $\int_0^2 1 d\mu(s) = \mathbb{I}$. Using (46), with s replacing β_k, we have that (53) becomes

$$\lim_{n\to\infty} T_\varepsilon^n = \begin{cases} \int_0^2 1 d\mu(s) = \mathbb{I} & \text{if } n \ll \varepsilon^2 \\ \int_0^2 e^{-s\tau} d\mu(s) = e^{-B\tau} & \text{if } n = \tau\varepsilon^{-2} \\ \mu(0+) - \mu(0) = 0 & \text{if } \varepsilon^{-2} \ll n \ll |\varepsilon|^{-3} \end{cases}, \tag{54}$$

where we used (52) and the continuity of $\mu(s)$ at zero .

Plugging (54) in (38) gives the required results in ℓ^2.

For the results in ℓ^1 we use series expansions:

$$\left(\mathbb{I} - \varepsilon^2 B\right)^n = \mathbb{I} + \binom{n}{1} \varepsilon^2(-B) + \binom{n}{2} \varepsilon^4(-B)^2 + \ldots + \binom{n}{n} \varepsilon^{2n}(-B)^n \tag{55}$$

Since $\|B\|_1 \leq 2$, (see property B3$_\infty$), the finite series above is dominated in ℓ^1 operator norm by:

$$1 + 2\varepsilon^2 \binom{n}{1} + (2\varepsilon^2)^2 \binom{n}{2} + \ldots + (2\varepsilon^2)^n \binom{n}{n} = (1 + 2\varepsilon^2)^n \leq e^{2n\varepsilon^2}. \tag{56}$$

As $n \to \infty$ the series in (56) becomes infinite. However, as long as $n \le \tau/\varepsilon^2$, $\tau > 0$ fixed, the sum in (56) is finite and hence that in (55) is convergent. Now for each $k = 1, 2, \ldots$ the $(k+1)^{\text{st}}$ term in the series (55) has the property:

$$\lim_{n\to\infty} \left(\begin{array}{c} n \\ k \end{array} \right) \varepsilon^{2k}(-B)^k = \left\{ \begin{array}{ll} 0 & \text{if } n \ll \varepsilon^{-1} \\ \frac{\tau^k}{k!}(-B)^k & \text{if } n = \tau\varepsilon^{-2} \end{array} \right.$$

Hence by the Weierstrass criterion for absolutely convergent series we have:

$$\lim_{n\to\infty} T_\varepsilon^n = \lim_{n\to\infty} \left(\mathbb{I} - \varepsilon^2 B\right)^n = \left\{ \begin{array}{ll} \mathbb{I} - 0 + 0 - \ldots = \mathbb{I} & \text{if } n \ll \varepsilon^{-1} \\ \mathbb{I} - \tau B + \frac{(\tau B)^2}{2!} - \frac{(\tau B)^3}{3!} + \ldots = e^{-\tau B} & \text{if } n = \tau\varepsilon^{-2} \end{array} \right. \tag{57}$$

It remains to prove that as $n \to \infty, \varepsilon^2 n \to \infty$, $\{P^{(n)}\}$ does not converge in ℓ^1. Let $P^{(0)} \in \ell^1 \cap \ell^2$ denote a vector with positive components, and consider the sequence:

$$P^{(n)} = T_\varepsilon^n P^{(0)} \in \ell^1 \cap \ell^2. \tag{58}$$

By the third part of (54), $\|P^{(n)}\|_2 \to 0$. Assume now that there exists $P \in \ell^1$ such that $\|P^{(n)} - P\|_1 = 0$. Since both ℓ^1 and ℓ^2 convergence imply convergence of each component, we deduce that $P = 0$. On the other hand, by $P^{(n)} = T_\varepsilon P^{(n-1)}$, $\quad n = 1, 2, \ldots$ and property $B4_\infty$, we deduce that $P^{(n)}$ is a positive vector for which $\|P^{(n)}\|_1 = \|P^{(0)}\|_1 \stackrel{def}{=} E > 0$ for all $n \ge 0$. Consequently P is a nonnegative vector with $\|P\|_1 = E > 0$, a contradiction. The proof of the Corollary is now complete. [].

3.2 Power diffusion after a fixed (large) time interval and a random number of defects

As pointed out in its statement, Theorem 3.1 is valid when one measures the power vector after a fixed number of defects "n" regardless of the realizations of the random variables. That is after each realization of $d_0, d_1, \ldots$ the power vector is measured in between the n^{th} and the $(n+1)^{\text{st}}$ defect. Averaging the measurements over all the realizations of $d_0, d_1, d_2, \ldots$ gives the result of Theorem 3.1. What happens if one chooses to measure the power vector at a fixed time "t" (i.e. a fixed distance along the fiber)? The answer is given by the next theorem:

Theorem 3.2. *Consider equation (12) with g of the form (4). Assume that* **(H1-H4)** *are satisfied and that all random variables $d_0, d_1, \ldots,$ have finite mean, variance and third momentum. Fix a time t, $0 \le t \ll 1/|\varepsilon|^3$. Then the expected value of the power vector at a fixed time $P(t)$ satisfies*

$$P(t) = T_\varepsilon^n P(0) + \mathcal{O}(\max\{t\varepsilon^3, \varepsilon^{4/5}\}), \tag{59}$$

where $n = \lfloor t/(T+M) \rfloor$ *denotes the integer part of* $t/(T+M)$, T *is the common time span of the defects and* M *is the mean of the identically distributed random variables* $d_0, d_1, \ldots$.

Corollary 3.3. *In this setting, the conclusions of Corollaries 3.1, 3.2 hold with* n *replaced by* t.

<u>Proof of Theorem 3.2.</u> As before, let $P^{(k)}$ be the expected power vector after exactly "k" defects. Denote by N the random variable counting the number of "defects" up until time t, i.e.

$$(N-1)T + d_0 + \ldots + d_{N-1} < t \leq NT + d_0 + \ldots + d_N. \tag{60}$$

and let $\delta(\varepsilon)$ denote the integer, which grows as ε decreases:

$$\begin{aligned} \tilde{\delta} &= \max\left\{1.39\left(\frac{\rho}{\sigma^2(T+M)}\right)^{2/5}\varepsilon^{-6/5}, \frac{\sigma}{T+M}\sqrt{n\log\left(\varepsilon^{-2}\right)} + \left(\frac{\sigma}{T+M}\right)^2\log\left(\varepsilon^{-2}\right)\right\} \\ \delta &= \lfloor\tilde{\delta}\rfloor + 1, \end{aligned} \tag{61}$$

where M, σ^2, respectively ρ are the mean, variance and the centered third momentum, of the identically distributed variables $d_0, d_1, d_2, \ldots$, and n is the integer part of $t/(T+M)$. Note that for $t \sim \varepsilon^{-3}$ or smaller $\delta \ll \varepsilon^{-2}$. The choice of $\delta(\varepsilon)$ is explained below.

The proof consists of three stages:

1. $P(t) = P^{(n+\delta)} + \mathcal{O}(\varepsilon) + \mathcal{O}(\delta\varepsilon^2)$
2. $P^{(n+\delta)} = P^{(n)} + \mathcal{O}(\delta\varepsilon^2)$
3. $P^{(n)} = T_\varepsilon^n P(0) + \mathcal{O}(n\varepsilon^3)$

where $n = \lfloor t/(T+M) \rfloor$. The last stage is simply Theorem 3.1.

For the second stage one applies again the previous theorem to get:

$$P^{(n+\delta)} = T_\varepsilon^\delta P^{(n)} + \mathcal{O}(\delta\varepsilon^3).$$

Now $T_\varepsilon = I - \mathcal{O}(\varepsilon^2)$ and since $\delta \ll \varepsilon^{-2}$ stage two follows.

The first stage is the trickiest. Without loss of generality we can assume that $t/(T+M)$ is an integer. Indeed, for $n = \lfloor t/(T+M) \rfloor$ we have

$$P(t) - P(n(T+d)) = \mathcal{O}\left(\varepsilon(T+M)\right) = \mathcal{O}(\varepsilon),$$

an error which is already accounted for in this stage.

Suppose first $n-\delta \le N \le n+\delta$, i.e. we condition the expected values to the realization of $|N-\delta| \le 0$. Then the difference between the conditional expected values of the power vector at time t and after $n+\delta$ defects is of order $\mathcal{O}(\varepsilon)+\mathcal{O}(\delta\varepsilon^2)$. This follows from the fact that the condition $n-\delta \le N \le n+\delta$ restricts only the realizations of $d_0, d_1, \dots d_N$ leaving the realizations of $d_{N+1}, \dots d_{n+\delta}$ arbitrary; see (60). Hence, as in stage two, the conditional expected values satisfy:

$$P^{(n+\delta)} = P^{(N+1)} + \mathcal{O}(\delta\varepsilon^2).$$

In addition

$$P^{(N+1)} = P(t) + \mathcal{O}(\varepsilon),$$

since there are at most 2 defects of size ε from "t" up until after the $(N+1)^{th}$ defect.

Let $p(t)$ denote the power vector

$$p(t) = \left(|a_1(t)|^2, |a_2(t)|^2, \dots\right).$$

Recall that by definition $P(t) = \mathbb{E}(p(t))$ and the total power in the system (12) is conserved, *i.e.*

$$\|p(t)\|_1 \stackrel{\text{def}}{=} \sum_k |a_k(t)|^2 \equiv \|p(0)\|_1, \quad t \in \mathbb{R}.$$

Moreover,

$$\begin{aligned} P(t) &= \mathbb{E}\left(p(t) \ : \ |N-n| \le \delta\right) + \mathbb{E}\left(p(t) \ : \ |N-n| > \delta\right) \\ &= P^{(n+\delta)} + \mathcal{O}(\delta\varepsilon^2) + \mathcal{O}(\varepsilon) + \mathcal{O}\left(\|p(0)\|_1 \text{Prob}(|N-n| > \delta)\right) \end{aligned} \tag{62}$$

We claim that for δ given by (61)

$$\text{Prob}(|N-n| > \delta) = \mathcal{O}(\varepsilon) + \mathcal{O}(\delta\varepsilon^2). \tag{63}$$

Indeed, since $t = n(T+M)$

$$\begin{aligned} \text{Prob}(|N-n| > \delta) &= \text{Prob}\left(\sum_{k=0}^{n+\delta}(T+d_k) \le t\right) + \text{Prob}\left(\sum_{k=0}^{n-\delta}(T+d_k) > t\right) \\ &= \text{Prob}\left(\frac{\sum_{k=0}^{n+\delta}(T+d_k) - (n+\delta)(T+M)}{\sigma\sqrt{n+\delta}} \le -\frac{\delta(T+M)}{\sigma\sqrt{n+\delta}}\right) \\ &+ \text{Prob}\left(\frac{\sum_{k=0}^{n-\delta}(T+d_k) - (n-\delta)(T+M)}{\sigma\sqrt{n-\delta}} > \frac{\delta(T+M)}{\sigma\sqrt{n-\delta}}\right). \end{aligned} \tag{64}$$

We are going to show how the choice (61) implies

$$\text{Prob}\left(\frac{\sum_{k=0}^{n-\delta}(T+d_k)-(n-\delta)(T+M)}{\sigma\sqrt{n-\delta}} > \frac{\delta(T+M)}{\sigma\sqrt{n-\delta}}\right) \leq \frac{\varepsilon}{2}+\frac{\delta\varepsilon^2}{2}. \tag{65}$$

The other half of (64):

$$\text{Prob}\left(\frac{\sum_{k=0}^{n+\delta}(T+d_k)-(n+\delta)(T+M)}{\sigma\sqrt{n+\delta}} < -\frac{\delta(T+M)}{\sigma\sqrt{n+\delta}}\right) \leq \frac{\varepsilon}{2}+\frac{\delta\varepsilon^2}{2}. \tag{66}$$

is analogous.

Depending on the size of n one has either:

$$\frac{0.8\rho}{\sigma^3\sqrt{n-\delta}} \leq \frac{\delta\varepsilon^2}{2} \tag{67}$$

or:

$$\frac{0.8\rho}{\sigma^3\sqrt{n-\delta}} > \frac{\delta\varepsilon^2}{2}. \tag{68}$$

If (67) holds, which corresponds to large n, we use the central limit theorem with Van Beek rate of convergence, see [6]:

$$\text{Prob}\left(\frac{\sum_{k=0}^{n-\delta}(T+d_k)-(n-\delta)(T+M)}{\sigma\sqrt{n-\delta}} > \frac{\delta(T+M)}{\sigma\sqrt{n-\delta}}\right) \leq \frac{1}{\sqrt{2\pi}}\int_{\frac{\delta(T+M)}{\sigma\sqrt{n-\delta}}}^{\infty} e^{-x^2/2}dx + \frac{0.8\rho}{\sigma^3\sqrt{n-\delta}}.$$

This together with (67), the inequality

$$\frac{1}{\sqrt{2\pi}}\int_a^\infty e^{-x^2/2}dx \leq \frac{e^{-a^2/2}}{2}$$

and the fact that $\delta \geq \frac{\sigma}{(T+M)}\sqrt{n\log\varepsilon^{-2}}$ implies $\frac{\delta(T+M)}{\sigma\sqrt{n-\delta}} \geq 2\ln\varepsilon^{-1}$, proves (65) for the case (67). If (68) holds then we apply Chebyshev inequality:

$$\text{Prob}\left(\frac{\sum_{k=0}^{n-\delta}(T+d_k)-(n-\delta)(T+M)}{\sigma\sqrt{n-\delta}} > \frac{\delta(T+M)}{\sigma\sqrt{n-\delta}}\right) \leq \frac{\sigma^2(n-\delta)}{\delta^2(T+M)^2} \leq \frac{\delta\varepsilon^2}{2},$$

where the latter inequality follows from (68) and

$$\delta \leq 1.39\left(\frac{\rho}{\sigma^2(T+M)}\right)^{2/5}\varepsilon^{-6/5}.$$

From (64), (65) and (66) we get relation (63). The latter plugged into (62) proves the first stage.

Finally, the three stages imply Theorem 3.2 provided that both ε and $\delta\varepsilon^2$ are dominated by $C\max\{n\varepsilon^3, \varepsilon^{3/4}\}$, for an appropriate constant $C > 0$. This follows directly from $\varepsilon \leq 1$ and (61). The proof is now complete. []

4 Diffusion of power in systems with discrete and continuous spectrum

Thusfar we have considered with systems with Hamiltonian, H_0, having only discrete spectrum. We now extend our analysis to the case where H_0 has both discrete and continuous spectrum. Continuous spectrum is associated with radiative behavior and this is manifested in a *dissipative* correction to the operator (21), entering at $\mathcal{O}(\varepsilon^2)$. Therefore, the dynamics on time scales $n \sim \varepsilon^{-2}$ is characterized by diffusion of energy among the discrete modes **and** radiative damping due to coupling of bound modes to the "heat bath" of radiation modes.

The hypotheses on the unperturbed Hamiltonian H_0 are similar to those in [11]. There is one exception though, the singular local decay estimates are replaced by a condition appropriate for perturbations with continuous spectral components, see Hypothesis **(H7')** below. For convenience we list here and label all the hypotheses we use:

(H1') H_0 is self-adjoint on the Hilbert space $\mathcal{H}$. The norm, respectively scalar product, on $\mathcal{H}$ are denoted by $\|\cdot\|$, respectively $\langle\cdot,\cdot\rangle$.

(H2') The spectrum of H_0 is assumed to consist of an absolutely continuous part, $\sigma_{\rm cont}(H_0)$, with associated spectral projection, $\mathbf{P_c}$, spectral measure $dm(\xi)$ and a discrete part formed by isolated eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$ (counting multiplicity) with an orthonormalized set of eigenvectors $\psi_1, \psi_2, \ldots, \psi_m$, i.e. for $k,\ j = 1, \ldots, m$

$$H_0\psi_k = \lambda_k\psi_k,\ \langle\psi_k, \psi_j\rangle = \delta_{kj},$$

where δ_{kj} is the Kronecker-delta symbol.

(H3') Local decay estimates on e^{-iH_0t}: There exist self-adjoint "weights", $w_-,\ w_+$, number $r_1 > 1$ and a constant $\mathcal{C}$ such that

(i) w_+ is defined on a dense subspace of $\mathcal{H}$ and on which $w_+ \geq cI$, $c > 0$

(ii) w_- is bounded, i.e. $w_- \in \mathcal{L}(\mathcal{H})$, such that $Range(w_-) \subseteq Domain(w_+)$

(iii) $w_+\ w_-\ \mathbf{P_c}\ =\ \mathbf{P_c}$ and $\mathbf{P_c}\ =\ \mathbf{P_c}\ w_-\ w_+$ on the domain of w_+

and for all $f \in \mathcal{H}$ satisfying $w_+ f \in \mathcal{H}$ we have

$$\|w_- e^{-iH_0t}\mathbf{P_c} f\| \leq \mathcal{C}\ \langle t\rangle^{-r_1}\|w_+ f\|, \quad t \in \mathbb{R}.$$

The hypotheses on the perturbation are similar to the ones used in the previous sections for discrete systems, namely:

(H4') β is a bounded self adjoint operator on $\mathcal{H}$ and satisfies $\|\beta\| = 1$. In addition we suppose that β is "localized", i.e. $w_+\beta$ and $w_+\beta w_+$ are bounded on $\mathcal{H}$, respectively on $Domain(w_+)$.

(H5') $g_0(t) \in L^2(\mathbb{R})$ is real valued, has compact support contained in $[0,T]$ on the positive real line and its L^1-norm, denoted by $\|g_0\|_1$ is 1. Therefore its Fourier transform, $\hat{g}_0$ is smooth and $\|\hat{g}_0\|_\infty \le 1$.

(H6') $d_0, d_1, \ldots$ are independent identically distributed random variables taking only nonnegative values, with finite mean, M, and such that for any $l \in \{0,1,\ldots\}$ and $j \neq k \in \{1,2\ldots,m\}$ we have

$$\mathbb{E}\left(e^{i(\lambda_j-\lambda_k)d_l}\right) = 0$$

where $\mathbb{E}(\cdot)$ denotes the expected value.

Define the common characteristic (moment generating) function for the random variables $d_0+T, d_1+T, \ldots$

$$\rho(\xi) \;\equiv\; \mathbb{E}\left(e^{-i\xi(d_0+T)}\right) = \mathbb{E}\left(e^{-i\xi(d_1+T)}\right) = \cdots. \tag{69}$$

Note that ρ is a continuous function on $\mathbb{R}$ bounded by 1. Then (H6') is equivalent to

$$\rho(\lambda_k - \lambda_j) = 0$$

for all $j \neq k \in \{1,2,\ldots,m\}$.

We require an additional local decay estimate:

(H7') There exists the number $r_2 > 2$ such that for all $f \in \mathcal{H}$ satisfying $w_+f \in \mathcal{H}$ and all $\lambda_k,\ \lambda_j,\ k,j = 1,\ldots,m$ we have:

$$\|w_- e^{-iH_0t}\rho(H_0-\lambda_k)\hat{g}_0(H_0-\lambda_k)\hat{g}_0(\lambda_j-H_0)\mathbf{P_c}f\| \le \frac{\mathcal{C}\|g_0\|_1^2}{\langle t\rangle^{r_2}}\|w_+f\|, \quad t\in\mathbb{R}.$$

Here $\hat{g}_0$ denotes the Fourier Transform, see (11), and the operators $\rho(H_0-\lambda)\mathbf{P_c}$, $\hat{g}_0(\lambda - H_0)\mathbf{P_c}$ are defined via the spectral theorem:

$$\begin{aligned}
\rho(H_0-\lambda)\mathbf{P_c} &= \int_{\sigma_{\rm cont}(H_0)} \rho(\xi-\lambda)dm(\xi) \\
&= e^{-i(H_0-\lambda)T}\mathbb{E}\left(e^{-i(H_0-\lambda)d_l}\right),\ l=1,2,\ldots, \qquad (70)\\
\hat{g}_0(\lambda-H_0)\mathbf{P_c} &= \int_{\sigma_{\rm cont}(H_0)} \hat{g}_0(\lambda-\xi)dm(\xi) \\
&= \int_0^T g_0(t)e^{-i(\lambda-H_0)t}\mathbf{P_c}dt, \qquad (71)
\end{aligned}$$

where $dm(\xi)$ is the absolutely continuous part of the spectral measure of H_0.

Remark 4.1. Conditions implying (H7') *If $H_0 = -\Delta + V(x)$ is a Schrödinger operator with potential, $V(x)$, which decays sufficiently rapidly as x tends to infinity, then either*

$$\mathbb{E}\left(e^{i\lambda_j d_l}\right) = 0, \quad l = 0, 1, \dots \text{ and } j = 1, 2 \dots, m \tag{72}$$

or

$$\hat{g}_0(\lambda_j) = 0, \; j = 1, \dots, m \tag{73}$$

imply **(H7')**, *provided the mean and variance of the random variables $d_0, d_1, \dots,$ are finite. Note that (72) is equivalent to adding the threshold, $\lambda_0 = 0$, of the continuous spectrum to the set of eigenvalues $\{\lambda_k \; : \; k = 1, 2, \dots, m\}$ for which* **(H6')** *must hold. Hypothesis (73) means that the perturbation should not induce a resonant coupling between the bound states and the threshold generalized eigenfunction associated with $\lambda_0 = 0$.*

In analogy with the case of discrete spectrum, we write the solution of (2) in the form

$$\phi(t, x) = \sum_{j=1}^{m} a_j(t)\psi_j(x) + \mathbf{P_c}\phi(t, x).$$

Recall that the expected power vector $P(t)$ is defined as the column vector

$$P(t) = \left(\mathbb{E}(\overline{a}_1 a_1(t)), \mathbb{E}(\overline{a}_2 a_2(t)), \dots, \mathbb{E}(\overline{a}_m a_m(t))\right).$$

We denote by

$$P^{(n)} = P(t), \qquad t_{n-1} + T \le t < t_n$$

the expected power vector after $n \ge 1$ defects (note that $P(t)$ is constant on the above intervals).

We will show that the change in the power vector induced by each defect can be expressed in terms of a power transmission matrix

$$\begin{aligned} \mathcal{T}_\varepsilon &= T_{disc,\varepsilon} \; - \; \varepsilon^2 \,\mathrm{diag}[\gamma_1, \gamma_2, \dots, \gamma_m] \\ &= \mathbb{I} \; - \; \varepsilon^2 B \; - \; \varepsilon^2 \,\mathrm{diag}[\gamma_1, \gamma_2, \dots, \gamma_m] \end{aligned} \tag{74}$$

Recall that $T_{disc,\varepsilon} = T_\varepsilon = \mathbb{I} - \varepsilon^2 B$, displayed in (21-22), is the power transmission matrix for systems governed by discrete spectrum. Each damping coefficient $\gamma_k > 0, \; k = 1, 2, \dots, m$ results from the interaction between the corresponding bound state and the radiation field. In contrast to the results in [11], there are no contributions from bound state - bound state interactions mediated by the continuous spectrum; these terms cancel out by stochastic averaging.

Remark 4.2. *For sufficiently small ε we have:*

$$\|T_\varepsilon\|_1 = 1 - \varepsilon^2 \min\{\gamma_1, \gamma_2, \dots, \gamma_m\} < 1 \tag{75}$$

The damping coefficients are given by:

$$\gamma_k = \lim_{\eta \searrow 0} \left\| \hat{g}_0(H_0 - \lambda_k) \sqrt{\mathbb{I} - |\rho(H_0 - \lambda_k - i\eta)|^2} \left(\mathbb{I} - \rho(H_0 - \lambda_k - i\eta)\right)^{-1} \mathbf{P_c}[\beta\psi_k] \right\|^2 > 0, \tag{76}$$

for all $k = 1, 2, \dots, m$. Here the operators which are functions of H_0 are defined via the spectral theorem and $\mathbb{I}$ is the identity on $\mathcal{H}$.

The following theorem is a generalization of our previous result on the effect of a single defect on the mode-power distribution, adapted to the case where the Hamiltonian has both discrete and continuous spectrum:

Theorem 4.1. *Consider the Schrödinger equation*

$$i\partial_t \phi \;=\; H_0 \phi \;+\; g(t)\beta\phi, \tag{77}$$

where $g(t)$ is a random function, defined in terms of $g_0(t)$, given by (4). Assume that hypotheses **(H1'-H7')** *hold. Consider initial conditions for (2) such that $w_+ \mathbf{P_c}\phi_0 \in \mathcal{H}$. Then there exists an $\varepsilon_0 > 0$ such that whenever $|\varepsilon| \le \varepsilon_0$ the solution of (2) satisfy:*

$$P^{(n+1)} = \mathcal{T}_\varepsilon P^{(n)} + \mathcal{O}\left(\varepsilon^3\right) + \mathcal{O}\left(\frac{\varepsilon}{\langle nT \rangle^r}\right), \qquad n = 0, 1, 2, \dots, \tag{78}$$

where the matrix $\mathcal{T}_\varepsilon$ is given in (74) and $r = \min\{r_1, r_2 - 1\} > 1$.

By applying this theorem successively we get the change over $n \ge 1$ defects:

$$P^{(n)} = \mathcal{T}_\varepsilon^n P(0) + \sum_{k=0}^{n-1} \mathcal{T}_\varepsilon^k \left(\mathcal{O}(\varepsilon^3) + \mathcal{O}\left(\frac{\varepsilon}{\langle (n-k)T \rangle^r}\right)\right). \tag{79}$$

Using $\|T_\varepsilon\|_1 < 1$ and

$$\sum_{n=1}^{\infty} \langle nT \rangle^{-r} < \infty$$

we can conclude that the last correction term in (79) is of order $\mathcal{O}(\varepsilon)$.[2] As for the other correction term we have two ways in computing its size. The first is based on $\|\mathcal{T}_\varepsilon^k\|_1 < 1$, and gives

$$\sum_{k=0}^{n-1} \mathcal{T}_\varepsilon^k \mathcal{O}(\varepsilon^3) \;=\; \mathcal{O}(n\varepsilon^3).$$

[2]one can actually show that $\sum_{k=0}^{n-1} \mathcal{T}_\varepsilon^k \mathcal{O}\left(\frac{\varepsilon}{\langle (n-k)T \rangle^r}\right) = \mathcal{O}\left(\min\left\{\varepsilon, \frac{\langle nT \rangle^{-r}}{\varepsilon\gamma}\right\}\right)$. However, as $n \to \infty$ the other correction term dominates and the result of Theorem 4.2 cannot be improved.

The second is based on

$$\sum_{k=0}^{n-1} \|\mathcal{T}_\varepsilon^k\|_1 \leq (1 - \|\mathcal{T}_\varepsilon\|_1)^{-1} \leq \frac{1}{\gamma\varepsilon^2},$$

where $\gamma = \min\{\gamma_1, \gamma_2, \dots, \gamma_m\}$, and gives

$$\sum_{k=0}^{n-1} \mathcal{T}_\varepsilon^k \mathcal{O}(\varepsilon^3) \;=\; \mathcal{O}(\varepsilon\gamma^{-1}).$$

We have proved the following theorem:

Theorem 4.2. *Under the assumptions of Theorem 4.1, the expected power vector after n defects, $n = 1, 2, \dots,$ satisfies:*

$$P^{(n)} = \mathcal{T}_\varepsilon^n P(0) + \mathcal{O}\left(\min(\varepsilon\gamma^{-1}, n\varepsilon^3)\right) + \mathcal{O}(\varepsilon).$$

Here, $\mathcal{T}_\varepsilon$ is the diffusion/damping power transmission matrix given in (74).

Moreover, the argument we used in the proof of Theorem 3.2 now gives

Theorem 4.3. *Under the assumptions of Theorem 4.1, the expected power vector at a fixed time $t, 0 \leq t < \infty$ satisfies:*

$$P(t) = \mathcal{T}_\varepsilon^n P(0) + \mathcal{O}(\varepsilon^{4/5}). \tag{80}$$

Here, n is the integer part of $t/(T+M)$, T is the common time span of the defects and M is the mean of the identically distributed random variables $d_0, d_1, \dots$.

The nicer form of the correction term in (80) compared to (59) is due to the fact that $\min(t\varepsilon^3, \varepsilon/\gamma)$ is now dominated by $\mathcal{O}(\varepsilon^{4/5})$.

In analogy with Corollary 3.1 we have, in the present context, the following limiting behavior:

Corollary 4.1. *Under the assumption of theorem 4.1 the following holds:*

$$\lim_{t\to\infty} P(t) = \begin{cases} P(0), & \text{if } t \ll \varepsilon^{-2} \\ e^{-(B+\Gamma)\tau} P(0) & \text{if } t = \tau\varepsilon^{-2} \\ 0, & \text{if } t \gg \varepsilon^{-2},\ \varepsilon \to 0 \end{cases}, \tag{81}$$

where B is displayed in (22) and

$$\Gamma = \operatorname{diag}[\gamma_1, \gamma_2, \dots, \gamma_m] > 0$$

Proof Since $\mathcal{T}_\varepsilon = \mathbb{I} - \varepsilon^2(B+\Gamma)$ and $B+\Gamma$ is self adjoint with

$$B+\Gamma \geq \min\{\gamma_k \ : \ k=1,2,\dots m\} > 0$$

we have

$$\lim_{n\to\infty} \mathcal{T}_\varepsilon^{(n)} = \begin{cases} \mathbb{I}, & \text{if } n \ll \varepsilon^{-2} \\ e^{-(B+\Gamma)\tau} & \text{if } n = \tau\varepsilon^{-2} \\ 0 & \text{if } n \gg \varepsilon^{-2}, \varepsilon \to 0 \end{cases} . \tag{82}$$

This follows from writing $\mathcal{T}_\varepsilon$ in the basis which diagonalizes $B+\Gamma$ and using the fact that all eigenvalues of $B+\Gamma$ are strictly positive, see the proof of Corollary 3.1.

Clearly, (82) and Theorem 4.3 imply the conclusion of the corollary. []

Note that on time scales of order $1/\varepsilon^2$ the dynamical system is now equivalent to:

$$\partial_\tau P(\tau) = (-B-\Gamma)P(\tau),$$

where $-B$ is a diffusion operator, see the discussion after relation (51), while $-\Gamma$ is a damping operator.

It remains to prove Theorem 4.1.

Proof of Theorem 4.1. Consider one realization of the random variables $d_0, d_1, \dots$. For this realization the system (2) is linear, Hamiltonian and deterministic. It is well known that such systems have an unique solution, $\phi(t)$, defined for all $t \geq 0$ and continuously differentiable with respect to t. Moreover

$$\|\phi(t)\| \equiv \|\phi_0\|. \tag{83}$$

We decompose the solution in its projections onto the bound states and continuous spectrum of the unperturbed Hamiltonian:

$$\phi(t,x) = \sum_{j=1}^m a_j(t)\psi_j + \mathbf{P_c}\phi(t) = \phi_b(t) + \phi_d(t), \tag{84}$$

where ϕ_b and ϕ_d are, respectively, the bound and dispersive parts of ϕ:

$$\begin{aligned} \phi_b(t) &= \sum_{j=1}^m a_j(t)\psi_j, \\ \phi_d(t) &= \mathbf{P_c}\phi(t) \end{aligned} \tag{85}$$

and

$$\langle \phi_b(t), \phi_d(t)\rangle \equiv 0. \tag{86}$$

Note that (83) and (86) imply

$$\|\phi_b(t)\| \leq \|\phi_0\|, \qquad \|\phi_d(t)\| \leq \|\phi_0\|, \tag{87}$$

for all $t \geq 0$. Consequently,

$$|a_k(t)| \leq \|\phi_0\|, \tag{88}$$

for all $t \geq 0$.

By inserting (84) into (2) and projecting the later onto the bound states and continuous spectrum we get the coupled system:

$$\begin{aligned} i\partial_t a_k(t) &= \lambda_k a_k(t) + \varepsilon g(t) \langle \psi_k, \beta\phi_b(t)\rangle + \varepsilon g(t) \langle \psi_k, \beta\phi_d(t)\rangle, & (89) \\ i\partial_t \phi_d(t) &= H_0\phi_d(t) \ + \ \varepsilon g(t)\mathbf{P_c}\beta\phi_d(t) \ + \ \varepsilon g(t)\mathbf{P_c}\beta\phi_b(t), & (90) \end{aligned}$$

where $k = 1, 2, \ldots, m$. Duhamel's principle applied to (90) yields

$$\phi_d(t) = e^{-iH_0 t}\phi_d(0) - i\varepsilon \int_0^t g(s)e^{-iH_0(t-s)}\mathbf{P_c}\beta\phi_d(s)ds - i\varepsilon \int_0^t g(s)e^{-iH_0(t-s)}\mathbf{P_c}\beta\phi_b(s)ds. \tag{91}$$

In a manner analogous to the one in [3] we are going to isolate ϕ_d in (91). Consider the following two operators acting on $C(\mathbb{R}^+, Domain(w_+))$ respectively $C(\mathbb{R}^+, \mathcal{H})$, the space of continuous functions on positive real numbers with values in $Domain(w_+)$ respectively $\mathcal{H}$:

$$\begin{aligned} K^+[f](t) &= \int_0^t g(s)w_- e^{-iH_0(t-s)}\mathbf{P_c}\beta w_+ f(s)ds & (92) \\ K[f](t) &= \int_0^t g(s)w_- e^{-iH_0(t-s)}\mathbf{P_c}\beta f(s)ds. & (93) \end{aligned}$$

Then, by applying the w_- operator on both sides of (91) we get:

$$w_-\phi_d(t) = w_- e^{-iH_0 t}\phi_d(0) - i\varepsilon K^+[w_-\phi_d](t) - i\varepsilon K[\phi_b](t). \tag{94}$$

On $C(\mathbb{R}^+, \mathcal{H})$ we introduce the family of norms depending on $\alpha \geq 0$:

$$\|f\|_\alpha = \sup_{t\geq 0} \langle t\rangle^\alpha \|f(t)\| \tag{95}$$

and define the operator norm:

$$\|\mathcal{A}\|_\alpha = \sup_{\|f\|_\alpha \leq 1} \|\mathcal{A}f\|_\alpha. \tag{96}$$

The local decay hypothesis **(H3')** together with **(H4')** and **(H5')** imply:

Lemma 4.1. *If* $0 \le \alpha \le r_1$ *then there exists a constant* C_α *such that*

$$\begin{aligned} \|K^+\|_\alpha &\le C_\alpha \\ \|K\|_\alpha &\le C_\alpha. \end{aligned}$$

Proof of Lemma 4.1. Fix $\alpha,\ 0 \le \alpha \le r_1$ and $f \in C(\mathbb{R}_+, Domain(w_+))$ such that $\|f\|_\alpha \le 1$. Then

$$\begin{aligned} \langle t\rangle^\alpha \|K^+[f](t)\| &= \langle t\rangle^\alpha \left| \int_0^t g(s) w_- e^{-iH_0(t-s)} \mathbf{P_c} \beta w_+ f(s) ds \right| \\ &\le \langle t\rangle^\alpha \int_0^t |g(s)| \|w_- e^{-iH_0(t-s)} \mathbf{P_c} w_-\| \cdot \|w_+ \beta w_+\| \cdot \|f(s)\| ds \\ &\le \langle t\rangle^\alpha \mathcal{C} \|w_+\beta w_+\| \int_0^t \frac{|g(s)|}{\langle t-s\rangle^{r_1}} \|f(s)\| ds, \end{aligned}$$

where we used (H3'). Furthermore, from $\|f\|_\alpha \le 1$ and $\|w_+\beta w_+\|$ bounded, we have

$$\begin{aligned} \langle t\rangle^\alpha \|K^+[f](t)\| &\le C\langle t\rangle^\alpha \int_0^t \frac{|g(s)|}{\langle t-s\rangle^{r_1}\langle s\rangle^\alpha} \langle s\rangle^\alpha \|f(s)\| ds \\ &\le C\langle t\rangle^\alpha \|f\|_\alpha \int_0^t \frac{|g(s)|}{\langle t-s\rangle^{r_1}\langle s\rangle^\alpha} ds \\ &\le C\langle t\rangle^\alpha \sum_{\{j: t_j < t\}} \int_{t_j}^{\min(t, t_j+T)} \frac{|g(s)|}{\langle t-s\rangle^{r_1}\langle s\rangle^\alpha} ds. \end{aligned}$$

By the mean value theorem

$$\int_{t_j}^{\min(t,t_j+T)} \frac{|g(s)|}{\langle t-s\rangle^{r_1}\langle s\rangle^\alpha} ds = \langle t-\tilde{t}_j\rangle^{-r_1} \langle \tilde{t}_j\rangle^{-\alpha} \|g_0\|_1,$$

for some

$$t_j \le \tilde{t}_j \le \min(t, t_j+T). \tag{97}$$

Hence

$$\langle t\rangle^\alpha \|K^+[f](t)\| \le C\langle t\rangle^\alpha \sum_{\{j: \tilde{t}_j < t\}} \langle t-\tilde{t}_j\rangle^{-r_1} \langle \tilde{t}_j\rangle^{-\alpha} \tag{98}$$

We claim that

$$\sum_{\{j: \tilde{t}_j < t\}} \langle t-\tilde{t}_j\rangle^{-r_1} \langle \tilde{t}_j\rangle^{-\alpha} \le D_\alpha \langle t\rangle^{-\alpha} \tag{99}$$

for some constant D_α independent of t. This is a consequence of the fact that we are computing the convolution of two power-like sequences. For a more detailed proof we decompose the sum into two, first running for $\tilde{t}_j \le t/2$ and the second for $t/2 < \tilde{t}_j \le t$. For the former we have :

$$\begin{aligned} \sum_{\{j:\tilde{t}_j<t/2\}} \langle t-\tilde{t}_j\rangle^{-r_1}\langle \tilde{t}_j\rangle^{-\alpha} &\le \left\langle \frac{t}{2}\right\rangle^{-r_1} \sum_{\{j:\tilde{t}_j<t/2\}} \langle \tilde{t}_j\rangle^{-\alpha} \\ &\le \left\langle \frac{t}{2}\right\rangle^{-r_1} \sum_{\{j:jT<t/2\}} \langle jT\rangle^{-\alpha} \\ &\le \left\langle \frac{t}{2}\right\rangle^{-r_1} D_\alpha \left\langle \frac{t}{2}\right\rangle^{\max(0,1-\alpha)} \le D_\alpha \langle t\rangle^{-\alpha}, \end{aligned} \tag{100}$$

since $r_1 > \max(1,\alpha)$ and $\tilde{t}_j \ge t_j \ge (j-1)T$, see (H3'), the hypotheses of this lemma, respectively (97) and (5). The remaining part of the sum is treated similarly:

$$\begin{aligned} \sum_{\{j:t/2<\tilde{t}_j\le t\}} \langle t-\tilde{t}_j\rangle^{-r_1}\langle \tilde{t}_j\rangle^{-\alpha} &\le \left\langle \frac{t}{2}\right\rangle^{-\alpha} \sum_{\{j:t/2<\tilde{t}_j\le t/2\}} \langle t-\tilde{t}_j\rangle^{-r_1} \\ &\le \left\langle \frac{t}{2}\right\rangle^{-\alpha} \sum_{\{k:kT<t/2\}} \langle kT\rangle^{-r_1} \\ &\le \left\langle \frac{t}{2}\right\rangle^{-\alpha} D \le D_\alpha \langle t\rangle^{-\alpha}, \end{aligned} \tag{101}$$

since $r_1 > 1$ and $t - \tilde{t}_j \ge kT$ where k is such that $t_{k+j} = \max\{t_p : \ t_p \le t\}$, see (97) and (5).

Now (100) and (101) imply (99) which replaced in (98) proves the required estimate for the K^+ operator. For the K operator the argument is completely analogous. []

We are going to use Lemma 4.1 for $\alpha = 0$ and $\alpha = r_1$. For C_0 and C_{r_1} defined in the Lemma, let

$$C_K = \max\{C_0, C_{r_1}\}$$

Then, for ε such that $C_K\varepsilon < 1$, the inverse operator $(I - i\varepsilon K^+)^{-1}$ exists and it is bounded in the norms (96) for $\alpha = 0$ and $\alpha = r_1$. Then (94) implies:

$$\begin{aligned} w_-\phi_d(t) &= \left(I - i\varepsilon K^+\right)^{-1} \left[w_- e^{-iH_0 t}\phi_d(0)\right](t) - i\varepsilon \left(I - i\varepsilon K^+\right)^{-1} K[\phi_b](t) \\ &= \mathcal{O}\left(\langle t\rangle^{-r_1}\|w_+\phi_d(0)\|\right) - i\varepsilon K[\phi_b] + \mathcal{O}\left(\varepsilon^2\|K[\phi_b]\|\right). \end{aligned} \tag{102}$$

Thus we have expressed the dispersive part, $\phi_d(t)$ as a functional of the bound state part, $\phi_b(t)$. Substitution of (102) into (89) gives, for $k = 1, 2, \ldots$:

$$\partial_t a_k(t) = -i\lambda_k a_k(t) - i\varepsilon g(t) \sum_{j=1}^{m} a_j(t) \langle \psi_k, \beta\psi_j\rangle$$

$$
\begin{aligned}
&- \quad \varepsilon^2 g(t)\,\langle w_+\beta\psi_k, K[\phi_b](t)\rangle &(103)\\
&+ \quad \varepsilon g(t)\left(\mathcal{O}(\|w_+\phi_d(0)\|\langle t\rangle^{-r_1}) + \mathcal{O}\left(\varepsilon^2\|K[\phi_b]\|\right)\right) \quad k=1,2,\ldots,m.
\end{aligned}
$$

In particular (103) implies

$$
\overline{a}_k(t_n) = e^{i\lambda_k(t_n-t_l)}\overline{a}_k(t_l) + \varepsilon\sum_{p=l}^{n-1} e^{i\lambda_k(t_n-t_p)} D_p(d_0,d_1,\ldots,d_p), \tag{104}
$$

for all $k=0,1,\ldots m, n\geq 2$ and $l<n$. Here each constant D_p depends on the realization of $d_0,d_1,\ldots d_p$ and does not depend on the realization of any other random variable. In addition all D_p are uniformly bounded by a constant depending only on C_K above and the initial condition $\phi(0)$. Hence

$$
a_k(t_n) = e^{-i\lambda_k(t_n-t_l)}a_k(t_l) + \mathcal{O}(\varepsilon|n-l|) \tag{105}
$$

for all $n,l=0,1,2,\ldots,$ and $k=1,2,\ldots,m.$

We multiply both sides of (103) with $\overline{a}_k$, then add the resulting equation to its complex conjugate. Then we integrate from t_n to t_n+T and obtain for $k=1,2,\ldots,m$

$$
\overline{a}_k a_k(t_n+T) - \overline{a}_k a_k(t_n) \;=\; R_1+R_2+R_3, \tag{106}
$$

where

$$
\begin{aligned}
R_1 &= -i\varepsilon\sum_{j=1}^{m}\langle\psi_k,\beta\psi_j\rangle\int_{t_n}^{t_n+T} g(t)\overline{a}_k(t)a_j(t)dt + c.c. &(107)\\
R_2 &= -\varepsilon^2\langle w_+\beta\psi_k, \int_{t_n}^{t_n+T} g(t)\overline{a}_k(t)K[\phi_b](t)dt\rangle + c.c. &(108)\\
R_3 &= \mathcal{O}\left(\varepsilon\langle t_n\rangle^{-r_1}\right) + \mathcal{O}\left(\varepsilon^3\right). &(109)
\end{aligned}
$$

If we neglect the R_2 and R_3 in (106) we are left with R_1, which is precisely the expression associated with the power transfer in systems with discrete spectrum; see Section 2. Moreover R_3 has norm asserted in (78). So, it remains to show of R_2 that

$$
\begin{aligned}
&\mathbb{E}\left(\langle w_+\beta\psi_k, \int_{t_n}^{t_n+T} g(t)\overline{a}_k(t)K[\phi_b](t)dt\rangle + c.c.\right)\\
&= \gamma_k P_k^n + \mathcal{O}(\langle nT\rangle^{-r}) + \mathcal{O}(\varepsilon), &(110)
\end{aligned}
$$

where γ_k is given by (76) and $r=\min\{r_1,r_2-1\}>1$.

We use integration by parts. Let

$$
\tilde{K}[\phi_b](t) \equiv \int_{t_n+T}^{t} g(s)e^{-i\lambda_k(s-t_n)}K[\phi_b](s)ds, \qquad t_n\leq t\leq t_n+T. \tag{111}
$$

and note that $K[\phi_b](t_n+T)=0$. Lemma 4.1 together with

$$g(s)=g_0(s-t_n),\ \ t_n\le s\le t_n+T, \tag{112}$$

imply the existence of a constant C with the property:

$$\|\tilde{K}[\phi_b](t)\|\le C\|g_0\|_1^2=C, \tag{113}$$

uniformly in $t_n\le t\le t_n+T$. Define

$$A_k(t)=a_k(t)e^{i\lambda_k(t-t_n)}, \tag{114}$$

for $k=1,2,\dots,m$. Note that

$$A_k(t_n)=a_k(t_n) \tag{115}$$

From (103) we have

$$|\partial_t A_k(s)|\le C\ |\varepsilon|\ |g_0(s-t_n)| \tag{116}$$

for some constant C independent of s and $t_n\le s\le t_n+T$. Now

$$\begin{aligned}\int_{t_n}^{t_n+T} g(t)\overline{a}_k(t)K[\phi_b](t)dt &= \int_{t_n}^{t_n+T}\overline{A}_k(t)\partial_t\tilde{K}[\phi_b](t)dt\\ &= -\overline{a}_k(t_n)\tilde{K}[\phi_b](t_n)-\int_{t_n}^{t_n+T}\partial_t\overline{A}_k(t)\tilde{K}[\phi_b](t)dt\\ &= \overline{a}_k(t_n)\int_{t_n}^{t_n+T} g(t)e^{i\lambda_k(t-t_n)}K[\phi_b](t)dt+\mathcal{O}(\varepsilon).\end{aligned} \tag{117}$$

To further rewrite (117) we note that for $t_n\le t\le t_n+T$

$$\begin{aligned}K[\phi_b](t) &= \sum_{j=1}^{m}\int_{t_n}^{t} a_j(s)g(s)w_-e^{-iH_0(t-s)}\mathbf{P_c}\beta\psi_j ds\\ &+ \sum_{l=0}^{n-1}\sum_{j=1}^{m}\int_{t_l}^{t_l+T} a_j(s)g(s)w_-e^{-iH_0(t-s)}\mathbf{P_c}\beta\psi_j ds\end{aligned} \tag{118}$$

An integration by parts similar to the one above and use of (112) leads to:

$$\begin{aligned}&\int_{t_l}^{t_l+T} a_j(s)g(s)w_-e^{-iH_0(t-s)}\mathbf{P_c}\beta\psi_j ds =\\ &= a_j(t_l)\int_{t_l}^{t_l+T} g(s)e^{-i\lambda_j(s-t_l)}w_-e^{-iH_0(t-s)}\mathbf{P_c}\beta\psi_j ds+\mathcal{O}\left(\frac{\varepsilon}{\langle t-t_l-T\rangle^{r_1}}\right)\\ &= a_j(t_l)w_-\hat{g}_0(\lambda_j-H_0)e^{-iH_0(t-t_l)}\mathbf{P_c}\beta\psi_j+\mathcal{O}\left(\frac{\varepsilon}{\langle t-t_l-T\rangle^{r_1}}\right),\end{aligned} \tag{119}$$

and

$$\int_{t_n}^{t} a_j(s)g(s)w_-e^{-iH_0(t-s)}\mathbf{P_c}\beta\psi_j ds$$

$$= a_j(t_n)\int_{t_n}^{t} g(s)e^{-i\lambda_j(s-t_n)}w_-e^{-iH_0(t-s)}\mathbf{P_c}\beta\psi_j ds + \mathcal{O}(\varepsilon). \quad (120)$$

By plugging (119-120) in (118) we get

$$\begin{aligned} K[\phi_b](t) &= \sum_{j=1}^{m} a_j(t_n)\int_{t_n}^{t} g(s)e^{-i\lambda_j(s-t_n)}w_-e^{-iH_0(t-s)}\mathbf{P_c}\beta\psi_j ds \\ &+ \sum_{j=1}^{m}\sum_{l=0}^{n-1} a_j(t_l)w_-\hat{g}_0(\lambda_j - H_0)e^{-iH_0(t-t_l)}\mathbf{P_c}\beta\psi_j + \mathcal{O}(\varepsilon), \end{aligned} \quad (121)$$

where to estimate the error we used the fact that the series $\sum_l \langle t - t_l - T\rangle^{-r_1}$ is convergent and uniformly bounded in t.

We now substitute (121) into the right hand side of (117) and obtain

$$\begin{aligned} &\int_{t_n}^{t_n+T} g(t)\overline{a}_k(t)K[\phi_b](t)dt = \mathcal{O}(\varepsilon) \\ &+ \sum_{j=1}^{m} \overline{a}_k(t_n)a_j(t_n)\int_{t_n}^{t_n+T} g(t)e^{i\lambda_k(t-t_n)}\int_{t_n}^{t} g(s)e^{-i\lambda_j(s-t_n)}w_-e^{-iH_0(t-s)}\mathbf{P_c}\beta\psi_j dsdt \\ &+ \sum_{j=1}^{m}\sum_{l=0}^{n-1} \overline{a}_k(t_n)a_j(t_l)w_-\hat{g}_0(H_0 - \lambda_k)\hat{g}_0(\lambda_j - H_0)e^{-iH_0(t_n-t_l)}\mathbf{P_c}\beta\psi_j. \end{aligned} \quad (122)$$

Based on (104) we can replace $\overline{a}_k(t_n)a_j(t_l)$ in (122) with

$$\begin{aligned} \overline{a}_k(t_n)a_j(t_l) &= e^{i\lambda_k(t_n-t_l)}\overline{a}_k(t_l)a_j(t_l) + error(l,j) \quad (123) \\ error(l,j) &= \varepsilon\sum_{p=l}^{n-1} e^{i\lambda_k(t_n-t_p)}D_p w_-\hat{g}_0(H_0 - \lambda_k)\hat{g}_0(\lambda_j - H_0)e^{-iH_0(t_n-t_l)}\mathbf{P_c}\beta\psi_j. \end{aligned}$$

Taking into account that $t_n - t_{n-1} = d_n + T$ and the fact that $t_{n-1} - t_l,\ D_p,\ l \le p \le n-1$ do not depend on d_n, the expected value of the error can be rewritten as

$$\begin{aligned} &\mathbb{E}(error(l,j)) = \\ &= \varepsilon\sum_{p=l}^{n-1}\mathbb{E}\left(w_-e^{i(\lambda_k-H_0)(t_n-t_{n-1})}e^{i\lambda_k(t_{n-1}-t_p)}D_p\hat{g}_0(H_0-\lambda_k)\hat{g}_0(\lambda_j-H_0)e^{-iH_0(t_{n-1}-t_l)}\mathbf{P_c}\beta\psi_j\right) \\ &= \varepsilon\sum_{p=l}^{n-1} w_-\rho(H_0-\lambda_k)\mathbb{E}\left(e^{i\lambda_k(t_{n-1}-t_p)}D_p\hat{g}_0(H_0-\lambda_k)\hat{g}_0(\lambda_j-H_0)e^{-iH_0(t_{n-1}-t_l)}\mathbf{P_c}\beta\psi_j\right) \\ &= \varepsilon\sum_{p=l}^{n-1}\mathbb{E}\left(e^{i\lambda_k(t_{n-1}-t_p)}D_p w_-\rho(H_0-\lambda_k)\hat{g}_0(H_0-\lambda_k)\hat{g}_0(\lambda_j-H_0)e^{-iH_0(t_{n-1}-t_l)}\mathbf{P_c}\beta\psi_j\right) \end{aligned} \quad (124)$$

By applying the $\mathcal{H}$ norm to (124), commuting the norm with both summation and expected value and using **(H7')** we get:

$$\|\mathbb{E}(error(l,j))\| \leq |\varepsilon| \frac{\mathcal{C}(n-l)}{\langle t_{n-1}-t_l\rangle^{r_2}} \leq C\ |\varepsilon| \langle (n-l)T\rangle^{1-r_2}. \tag{125}$$

Since $r_2 > 2$ the summation over l and j of all the errors will have an $\mathcal{O}(\varepsilon)$ size. By this argument (122) becomes:

$$\begin{aligned}
\mathbb{E}\left(\left\langle w_+\beta\psi_k, \int_{t_n}^{t_n+T} g(t)\overline{a}_k(t)K[\phi_b](t)dt\right\rangle + c.c.\right) &= \sum_{j=1}^{m} \mathbb{E}(\overline{a}_k(t_n)a_j(t_n)) \\
\cdot \quad & \left\langle w_+\beta\psi_k, \int_{t_n}^{t_n+T} g(t)e^{i\lambda_k(t-t_n)}\int_{t_n}^{t} g(s)e^{-i\lambda_j(s-t_n)}w_-e^{-iH_0(t-s)}\mathbf{P_c}\beta\psi_j dsdt\right\rangle + c.c. \\
+ \quad & \sum_{j=1}^{m}\sum_{l=0}^{n-1} \mathbb{E}(\overline{a}_k(t_l)a_j(t_l)) \\
\cdot \quad & \mathbb{E}\left(\langle w_+\beta\psi_k, w_-\hat{g}_0(H_0-\lambda_k)\hat{g}_0(\lambda_j-H_0)e^{i(\lambda_k-H_0)(t_n-t_l)}\mathbf{P_c}\beta\psi_j\rangle\right) + c.c. \\
+ \quad & \mathcal{O}(\varepsilon).
\end{aligned} \tag{126}$$

But **(H6')** and the technique used to prove (42) imply

$$\mathbb{E}(\overline{a}_k(t_l)a_j(t_l)) = \begin{cases} P_k^{(l)} & \text{for } k=j, \\ 0 & \text{for } k \neq j \end{cases}$$

Moreover, an argument similar to the one we used in (123-125) allows us to replace $P^{(l)}$ by $P^{(n)}$ in (126) and incur an $\mathcal{O}(\varepsilon)$ total error. Then, (126) becomes

$$\begin{aligned}
&\mathbb{E}\left(\left\langle w_+\beta\psi_k, \int_{t_n}^{t_n+T} g(t)\overline{a}_k(t)K[\phi_b](t)dt\right\rangle + c.c.\right) = \\
&= P_k^{(n)}\langle w_+\beta\psi_k, w_-\hat{g}_0(H_0-\lambda_k)\hat{g}_0(\lambda_k-H_0)\mathbf{P_c}\beta\psi_k\rangle \\
&+ P_k^{(n)}\left\langle w_+\beta\psi_k, w_-\hat{g}_0(H_0-\lambda_k)\hat{g}_0(\lambda_k-H_0)\mathbb{E}\left(\sum_{l=0}^{n-1} e^{i(\lambda_k-H_0)(t_n-t_l)}\mathbf{P_c}\right)\beta\psi_k\right\rangle + c.c. \\
&+ \mathcal{O}(\varepsilon).
\end{aligned} \tag{127}$$

We claim that

$$\begin{aligned}
\gamma_k^n \overset{def}{=} & \langle w_+\beta\psi_k, w_-\hat{g}_0(H_0-\lambda_k)\hat{g}_0(\lambda_k-H_0)\mathbf{P_c}\beta\psi_k\rangle \\
+ & \left\langle w_+\beta\psi_k, w_-\hat{g}_0(H_0-\lambda_k)\hat{g}_0(\lambda_k-H_0)\mathbb{E}\left(\sum_{l=0}^{n-1} e^{i(\lambda_k-H_0)(t_n-t_l)}\mathbf{P_c}\right)\beta\psi_k\right\rangle + c.c. \\
= & \gamma_k + \mathcal{O}(\langle nT\rangle^{1-r_2})
\end{aligned} \tag{128}$$

where γ_k is given in (76). (128) replaced in (127) gives (110) which finishes the proof of this Theorem.

To prove (128) we first find a simpler expression for the expected value operator involved. Since $\{d_j\}_{j\geq 0}$ are independent, identically distributed with common characteristic function, $\rho(\xi)$, using the definition of $t_n,\ n\geq 0$, see (5) and the spectral resolution of the operator H_0, see (70), we have:

$$\begin{aligned}
\mathbb{E}\left(e^{i(\lambda_k-H_0)(t_n-t_l)}\mathbf{P_c}\right) &= \int_{\sigma_{\rm cont}(H_0)} \mathbb{E}(e^{i(\lambda_k-\xi)(t_n-t_l)})\ dm(\xi)\\
&= \int_{\sigma_{\rm cont}(H_0)} \mathbb{E}(e^{i(\lambda_k-\xi)\sum_{j=l}^{n-1}(d_j+T)})\ dm(\xi)\\
&= \int_{\sigma_{\rm cont}(H_0)} \prod_{j=l}^{n-1}\mathbb{E}(e^{i(\lambda_k-\xi)(d_j+T)})\ dm(\xi)\\
&= \int_{\sigma_{\rm cont}(H_0)} \rho^{n-l}(\xi-\lambda_k)\ dm(\xi) = \rho^{n-l}(H_0-\lambda_k)\mathbf{P_c}. \qquad (129)
\end{aligned}$$

Hence

$$\begin{aligned}
& w_-\hat{g}_0(H_0-\lambda_k)\hat{g}_0(\lambda_k-H_0)\mathbb{E}\left(\sum_{l=0}^{n-1}e^{i(\lambda_k-H_0)(t_n-t_l)}\mathbf{P_c}\right)\beta\\
&= w_-\hat{g}_0(H_0-\lambda_k)\hat{g}_0(\lambda_k-H_0)\sum_{j=1}^{n}\rho^j(H_0-\lambda_k)\mathbf{P_c}\beta. \qquad (130)
\end{aligned}$$

But each operator term in (130) has its $\mathcal{H}-$norm dominated by:

$$\begin{aligned}
&\|w_-\hat{g}_0(H_0-\lambda_k)\hat{g}_0(\lambda_k-H_0)\rho^j(H_0-\lambda_k)\mathbf{P_c}\beta\| =\\
&= \|w_-\rho(H_0-\lambda_k)\hat{g}_0(H_0-\lambda_k)\hat{g}_0(\lambda_k-H_0)\mathbb{E}(e^{-i(H_0-\lambda_k)(t_{j-1}-t_0)}\mathbf{P_c}\beta)\|\\
&\leq \frac{\mathcal{C}}{\langle t_{j-1}-t_0\rangle^{r_2}}\|w_+\beta\| \leq \langle (j-1)T\rangle^{-r_2}.
\end{aligned}$$

Now $r_2>2$ implies that the sequence $1/\langle jT\rangle^{r_2}$ is summable, and, by the dominant convergence theorem, there exists:

$$\begin{aligned}
\tilde{\gamma}_k &= \langle w_+\beta\psi_k, w_-\hat{g}_0(H_0-\lambda_k)\hat{g}_0(\lambda_k-H_0)\mathbf{P_c}\beta\psi_k\rangle\\
&+ \sum_{j=1}^{\infty}\left\langle w_+\beta\psi_k, w_-\hat{g}_0(H_0-\lambda_k)\hat{g}_0(\lambda_k-H_0)\rho^j(H_0-\lambda_k)\mathbf{P_c}\beta\psi_k\right\rangle + c.c.\\
&= \lim_{n\to\infty}\gamma_k^n.
\end{aligned}$$

Moreover

$$|\tilde{\gamma}_k - \gamma_k^n| = \sum_{j=n+1}^{\infty} \left\langle w_+\beta\psi_k, w_-\hat{g}_0(H_0-\lambda_k)\hat{g}_0(\lambda_k - H_0)\rho^j(H_0-\lambda_k \mathbf{P_c}\beta\psi_k\right\rangle + c.c.$$
$$\leq 2C\sum_{j=n}^{\infty}\langle jT\rangle^{-r_2} \leq D\langle nT\rangle^{1-r_2}. \tag{131}$$

Consider now, for $\eta > 0$,

$$\gamma_k^\eta = \langle w_+\beta\psi_k, w_-\hat{g}_0(H_0-\lambda_k)\hat{g}_0(\lambda_k-H_0)\mathbf{P_c}\beta\psi_k\rangle$$
$$+ \sum_{j=1}^{\infty}\left\langle w_+\beta\psi_k, w_-\hat{g}_0(H_0-\lambda_k)\hat{g}_0(\lambda_k-H_0)\rho^j(H_0-\lambda_k-i\eta)\mathbf{P_c}\beta\psi_k\right\rangle + c.c. \tag{132}$$

On one hand

$$\rho^j(H_0-\lambda_k-i\eta)\mathbf{P_c} = \mathbb{E}(e^{-\eta(t_j-t_0)}e^{-i(H_0-\lambda_k)(t_j-t_0)}\mathbf{P_c}) \tag{133}$$

and, by the dominant convergence theorem, for all $j \geq 1$

$$\lim_{\eta\searrow 0}\rho^j(H_0-\lambda_k-i\eta)\mathbf{P_c} = \rho^j(H_0-\lambda_k)\mathbf{P_c}.$$

On the other hand the series (132) is dominated uniformly in η by a summable series, because:

$$\|w_-\hat{g}_0(H_0-\lambda_k)\hat{g}_0(\lambda_k-H_0)\rho^j(H_0-\lambda_k-i\eta)\mathbf{P_c}\beta\|$$
$$= \left\|\int_0^T\int_0^T duds g_0(s+u)g_0(u)\mathbb{E}\left(e^{-\eta(t_j-t_0)}w_-e^{-i(H_0-\lambda_k)(t_j-t_0-s)}\mathbf{P_c}\beta\right)\right\|$$
$$\leq \frac{\mathcal{C}e^{-\eta jT}}{\langle t_j-t_o-T\rangle^{r_1}}\|g_0\|_1\|w_+\beta\| \leq \langle (j-1)T\rangle^{-r_1}.$$

Here we used **(H3')**, $\|g_0\|_1 = 1$ and $\|w_+\beta\|$ bounded. Therefore, by the Weierstrass criterion:

$$\lim_{\eta\searrow 0}\gamma_k^n = \tilde{\gamma}_k \tag{134}$$

In addition (133) implies

$$\|\rho(H_0-\lambda_k-i\eta)\mathbf{P_c}\| \leq \mathbb{E}\left(e^{-\eta(t_1-t_0)}\|e^{-i(H_0-\lambda_k)(t_1-t_0)}\mathbf{P_c}\|\right)$$
$$\leq e^{-\eta T} < 1.$$

This makes $(\mathbb{I}-\rho(H_0-\lambda_k-i\eta))\mathbf{P_c}$ invertible and given by the Neumann series:

$$(\mathbb{I}-\rho(H_0-\lambda_k-i\eta))^{-1}\mathbf{P_c} = \sum_{j=0}^{\infty}\rho^j(H_0-\lambda_k-i\eta)\mathbf{P_c}. \tag{135}$$

Plugging (135) in (132) we have

$$\begin{aligned}\gamma_k^\eta &= \langle \beta\psi_k, \hat{g}_0(H_0-\lambda_k)\hat{g}_0(\lambda_k - H_0)\mathbf{P_c}\beta\psi_k\rangle \\ &+ \left\langle \beta\psi_k, \hat{g}_0(H_0-\lambda_k)\hat{g}_0(\lambda_k - H_0)\rho(H_0-\lambda_k-i\eta)(\mathbb{I}-\rho(H_0-\lambda_k-i\eta))^{-1}\mathbf{P_c}\beta\psi_k\right\rangle + c.c.\end{aligned}$$

A simple inner product manipulation shows that:

$$\gamma_k^\eta = \left\| \hat{g}_0(H_0-\lambda_k)\sqrt{\mathbb{I}-|\rho(H_0-\lambda_k-i\eta)|^2}\,(\mathbb{I}-\rho(H_0-\lambda_k-i\eta))^{-1}\,\mathbf{P_c}[\beta\psi_k]\right\|^2.$$

Hence

$$\tilde{\gamma}_k = \lim_{\eta\searrow 0}\gamma_k^\eta = \gamma_k, \tag{136}$$

see also (134) and (76).

Finally, (136) and (131) give the claim (128). The theorem is now completely proven. []

5 Comparison to stochastic approach

In this section we want to compare our results with the *stochastic approach* in [17, 19, 18, 13]. We view the results of this paper and those discussed in this section as complementary. The results of this paper apply to the situation when a known localized "defect", g_0, is randomly distributed in a manner which achieves averaged diffusive effect. The results of Papanicolaou *et. al.* apply to a random medium, which is unknown and with assumptions about their distribution. One of the key technical assumptions in this latter work is that the expected value of the randomness, *at any time*, is zero, *i.e.* in our notation $\mathbb{E}(g(t)) = 0$. In the results of this paper, we allow for $\mathbb{E}(g(t))$ to vary with t. Indeed, for our train of pulses (see (4) and figure 1) $\mathbb{E}(g(t)) = 0$ and implies $g_0(t) \equiv 0$, so unless we have the $g_0 \equiv 0$, $\mathbb{E}(g(t))$ is generally different from zero and time-dependent. On the other hand, our hypothesis **(H4)** has no corresponding restriction in Papanicolaou *et. al.*'s theory.

Another important difference is that our result applies on time scales even larger than $1/\varepsilon^2$, where ε is the size of the randomness while the other results apply only on time scales up to $1/\varepsilon^2$. However, it appears that there is a striking similarity between the two results on $1/\varepsilon^2$ time scales. The train of pulses we analyzed is closest to the stochastic process described in [17, Section 2] where both its values in the epochs $[0, d_0+T], [d_0+T, d_0+T+d_1+T]\ldots$, and the epochs are now dependent on the realizations of the *same* random variables, $d_0, d_1, \ldots$. However, if we assume that the radiation modes are not present, the dynamical system we investigate, (89-90), is the one in [17, Section 4], see also [19]. These prevents us to use the formulas in the above papers. Nevertheless, we are going construct another stochastic perturbation, in the spirit of the one in [17, Section 4], for which we can compute the expected power evolution using both theories. We find that the two results coincide but keep in mind that while the example satisfies all our hypothesis it does not satisfy one of theirs, see below.

In addition to **(H1)-(H4)**, suppose the random variables $d_0, d_1, d_2, \ldots,$ can only take values in the interval $[0, d]$ (this is clearly satisfied by the random variables constructed in the previous section for finitely many modes) and denote by $\mu(t)$ the measure induced by their distribution. Consider the positive real axis partitioned into "epochs": $[0, T+d], [T+d, 2(T+d)], \ldots$ of length $T+d$. In each epoch a defect is placed at a distance d_j from the starting point of the j^{th} epoch. Specifically, the first defect is placed at a distance d_0 from $t=0$, the second at a distance d_1 from $t = T+d$, ... Here $d_0, d_1, d_2, \ldots$ are realizations of the random variables $d_0, d_1, d_2, \ldots$. That is, we will now consider equation (12) with the perturbation given by:

$$g(t) = g_0(t - d_0) + g_0(t - (T+d) - d_1) + g_0(t - 2(T+d) - d_2) + \ldots. \qquad (137)$$

see figure 2.

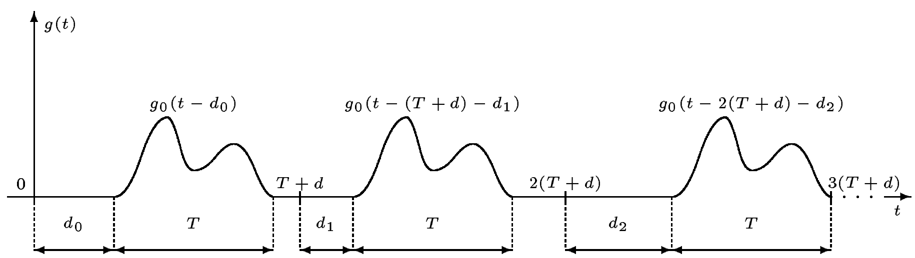

Figure 2: Another train of short lived perturbations

Our result, Theorem 3.1 applies without any modifications since before each perturbation we have:

$$\mathbb{E}\left(a_k \overline{a}_j(l(T+d) + d_l)\right) = \mathbb{E}\left(a_k \overline{a}_j(l(T+d))\right) \mathbb{E}\left(e^{-i\Delta_{kj} d_l}\right) = 0,$$

if $k \neq j$. As for Theorem 3.2 its proof is much simplified and the error estimate improved because we now know how many complete defects are going to appear up until the chosen time "t", namely $n = \lfloor t/(T+d) \rfloor$. The expected power at time t can differ from the one at time $n(T+d)$ by no more than the size of the perturbation, ε, since after each experiment only a part or a single full defect can occur in between this time slots. Hence:

$$P(t) = P^(n) + \mathcal{O}(\varepsilon) = T_\varepsilon^n P(0) + \mathcal{O}(n\varepsilon^3) + \mathcal{O}(\varepsilon), \tag{138}$$

where the integer n is such that $n(T+d) \leq t < (n+1)(T+d)$. To get closer to Papanicolaou *et. al.*'s results, suppose

$$t = \tau/\varepsilon^2, \qquad \tau \geq 0 \text{ is fixed}$$

and pass to the limit $\varepsilon \searrow 0$. We get

$$\lim_{\varepsilon \searrow 0} P(t) = \lim_{\varepsilon \searrow 0} T_\varepsilon^{\left\lfloor \frac{\tau}{(T+d)\varepsilon^2} \right\rfloor} P(0) = \lim_{\varepsilon \searrow 0} (\mathbb{I} - \varepsilon^2 B)^{\left\lfloor \frac{\tau}{(T+d)\varepsilon^2} \right\rfloor} P(0) = e^{\tau \tilde{B}} P(0) \tag{139}$$

where B is given in (22) and

$$\tilde{B} = -\frac{1}{(T+d)} B. \tag{140}$$

Let us now apply Papanicolaou *et. al.* result to the above example. Note that the manner in which the perturbation is constructed makes the example very close to that in [17, Section 4]. But since the stochastic process is not piecewise constant, one has to rely on more general form of their results such as [18, Remark 2 in Section 2]. The ODE system for the amplitude vector, $a(t) = (a_1(t), a_2(t), \ldots)'$, is:

$$\begin{aligned} \partial_t a(t) &= Aa(t) - i\varepsilon g(t)\alpha a(t), \\ a(0) &= a(0), \end{aligned} \tag{141}$$

where

$$\begin{aligned} A &= -i\,\mathrm{diag}[\lambda_1, \lambda_2, \ldots], \\ \alpha &= (\langle \psi_k, \beta\psi_j \rangle)_{1 \leq k,j}\,; \end{aligned}$$

see also (17). This is a special case of system (2.27) in [18] with

$$\tilde{M} \equiv -ig(t)\alpha.$$

Note that hypothesis (2.28) in [18] translates into

$$0 = \mathbb{E}(g(t)) = \int_0^d g_0(t' - s)d\mu(s), \quad \text{for all } t \geq 0 \tag{142}$$

where

$$t' = t - (T+d)\left\lfloor \frac{t}{T+d} \right\rfloor,$$

which generally implies the trivial case $g_0 \equiv 0$. Hence for nontrivial examples Papanicolaou *et. al.*'s theory is not rigorously applicable. We are going to replace (142) with a milder one:

$$\lim_{t \nearrow \infty} \frac{1}{t} \int_0^t \mathbb{E}(g(t))dt = 0, \tag{143}$$

i.e. the time average of the expected value of the perturbation is zero, under which we can formally derive closed coupled power equations. While the results of the stochastic approach do not apply to this example because $\mathbb{E}(g(t)) \neq 0$, our results do apply and it is reasonable to conjecture that there is an extension of the stochastic approach to this case. In the special case, where g is given in (137), the condition (143) reduces to $\hat{g}_0(0) = 0$,

Let us compute their equation for the evolution of the powers, i.e. we prove that their system of equations for the product of amplitudes:

$$\partial_\tau \mathbb{E}(a \otimes \overline{a}) = V \mathbb{E}(a \otimes \overline{a}),$$

where

$$\tau = \frac{t}{\varepsilon^2} \text{ is fixed as } \varepsilon \searrow 0.$$

gives a closed equation in powers, i.e.

$$V_{pq,pq'} = 0, \qquad \text{if } q \neq q' \tag{144}$$

and consequently for the powers $P(\tau) \equiv \operatorname{diag} \mathbb{E}(a \otimes \overline{a}(\tau))$ we have

$$\partial_\tau P(\tau) = \tilde{V} P(\tau), \tag{145}$$

with

$$\tilde{V}_{pq} = V_{pq,pq}. \tag{146}$$

The main point is that $\tilde{V}$ coincides with $\tilde{B}$ in our result (139); see also (140)and (22). Thus the two results agree on time scales of order $1/\varepsilon^2$.

For the formula of V we only have to replace M in [18, equation (2.35)] by its complex conjugate $\overline{M}$ whenever it applies on the right part of the tensor product, i.e.

$$\begin{aligned} V &= \lim_{t \nearrow \infty} \frac{1}{t} \int_{t_0}^{t+t_0} \int_{t_0}^{s} \mathbb{E}\left(M(s)M(\sigma) \otimes I + M(s) \otimes \overline{M}(\sigma)\right) d\sigma ds \\ &+ \lim_{t \nearrow \infty} \frac{1}{t} \int_{t_0}^{t+t_0} \int_{t_0}^{s} \mathbb{E}\left(M(\sigma) \otimes \overline{M}(s) + I \otimes \overline{M}(s)\overline{M}(\sigma)\right) d\sigma ds, \qquad (147) \\ M_{pq} &= -ie^{i\Delta_{pq}t} g(t) \alpha_{pq}; \end{aligned}$$

see also [18, relation (2.32)]. It will be clear from the argument below that the limit in (147) does not depend on t_0 (note that this in in fact a requirement for the validity of the theory) so we are going to work with $t_0 = 0$. Although the computation of V has been done in [19] (denoted there by $\overline{V}$) and then summarized in [13, Section 3] we are not able to use them because they relied on the stationarity of the process, see [19, relation (2.2)] which is not satisfied by our example. Nevertheless we have component wise:

$$\begin{aligned} V_{pq,pq'} &= -\delta_{pq'} \sum_r \alpha_{pr}\alpha_{rq} \lim_{t\nearrow\infty} \frac{1}{t} \int_0^t \int_0^s e^{i\Delta_{pr}s} e^{i\Delta_{rq}\sigma} \mathbb{E}(g(s)g(\sigma))d\sigma ds \\ &+ \alpha_{pq}\alpha_{q'p} \lim_{t\nearrow\infty} \frac{1}{t} \int_0^t \int_0^s e^{i\Delta_{pq}s} e^{i\Delta_{q'p}\sigma} \mathbb{E}(g(s)g(\sigma))d\sigma ds \\ &+ \alpha_{pq'}\alpha_{qp} \lim_{t\nearrow\infty} \frac{1}{t} \int_0^t \int_0^s e^{i\Delta_{qp}s} e^{i\Delta_{pq'}\sigma} \mathbb{E}(g(s)g(\sigma))d\sigma ds \qquad (148) \\ &- \delta_{pq} \sum_r \alpha_{rp}\alpha_{q'r} \lim_{t\nearrow\infty} \frac{1}{t} \int_0^t \int_0^s e^{i\Delta_{rp}s} e^{i\Delta_{q'r}\sigma} \mathbb{E}(g(s)g(\sigma))d\sigma ds, \end{aligned}$$

where we have used $\overline{\alpha}_{kj} = \alpha_{jk}$ due to the self-adjointness of β in $\alpha_{kj} = \langle \psi_k, \beta\psi_j \rangle$ and the fact that $g(t)$ is real valued. Thus it is sufficient to compute

$$\int_0^t \int_0^s e^{i\Delta_{kj}s} e^{i\Delta_{jl}\sigma} \mathbb{E}(g(s)g(\sigma))d\sigma ds.$$

Let us fix t and suppose for the moment that $t = n(T+d)$. Then

$$\begin{aligned} &\int_0^t \int_0^s e^{i\Delta_{kj}s} e^{i\Delta_{jl}\sigma} \mathbb{E}(g(s)g(\sigma))d\sigma ds = \\ &\quad \sum_{m=0}^{n-1} \int_{m(T+d)}^{(m+1)(T+d)} \int_0^s e^{i\Delta_{kj}s} e^{i\Delta_{jl}\sigma} \mathbb{E}(g(s)g(\sigma))d\sigma ds \\ &= \sum_{m=0}^{n-1} \int_{m(T+d)}^{(m+1)(T+d)} \int_{m(T+d)}^s e^{i\Delta_{kj}s} e^{i\Delta_{jl}\sigma} \mathbb{E}(g(s)g(\sigma))d\sigma ds \qquad (149) \\ &+ \sum_{m=0}^{n-1} \int_{m(T+d)}^{(m+1)(T+d)} e^{i\Delta_{kj}s} \mathbb{E}(g(s)) \int_0^{m(T+d)} e^{i\Delta_{jl}\sigma} \mathbb{E}(g(\sigma))d\sigma ds, \end{aligned}$$

where we have used the fact that the random variable $g(s)$ and $g(\sigma)$ are independent unless s and σ are in between the same epochs. Now

$$\int_0^{m(T+d)} e^{i\Delta_{jl}\sigma} \mathbb{E}(g(\sigma))d\sigma = \sum_{r=0}^{m-1} \int_{r(T+d)}^{(r+1)(T+d)} e^{i\Delta_{jl}\sigma} \mathbb{E}(g_0(\sigma - r(T+d) - d_{r+1}))d\sigma$$

$$
\begin{aligned}
&= \sum_{r=0}^{m-1} e^{i\Delta_{jl} r(T+d)} \int_0^{T+d} e^{i\Delta_{jl}\sigma} \int_0^d g_0(\sigma - s) d\mu(s) d\sigma \\
&= \sum_{r=0}^{m-1} e^{i\Delta_{jl} r(T+d)} \int_0^{d} e^{i\Delta_{jl} s} \int_{-s}^{T+d-s} e^{i\Delta_{jl}\sigma} g_0(\sigma) d\sigma d\mu(s) \\
&= \sum_{r=0}^{m-1} e^{i\Delta_{jl} r(T+d)} \int_0^{d} e^{i\Delta_{jl} s} d\mu(s) \int_0^{T} e^{i\Delta_{jl}\sigma} g_0(\sigma) d\sigma \\
&= \sum_{r=0}^{m-1} e^{i\Delta_{jl} r(T+d)} \mathbb{E}(e^{i\Delta_{jl} d_{r+1}}) \hat{g}_0(-\Delta_{jl}) \equiv 0,
\end{aligned}
$$

where we used supp $g_0 \subset [0, T]$, $\mathbb{E}(e^{i\Delta_{jl} d_{r+1}}) = 0$ if $j \neq l$, see **(H4)**, and the fact that $\hat{g}_0(0) = 0$.

The only nonzero terms left in (149) are of the form:

$$
\begin{aligned}
&\int_{m(T+d)}^{(m+1)(T+d)} \int_{m(T+d)}^{s} e^{i\Delta_{kj} s} e^{i\Delta_{jl}\sigma} \mathbb{E}(g(s)g(\sigma)) d\sigma ds \\
&= e^{i\Delta_{kl} m(T+d)} \int_0^{T+d} \int_0^{s} e^{i\Delta_{kj} s} e^{i\Delta_{jl}\sigma} \mathbb{E}(g_0(s - d_m) g_0(\sigma - d_m)) d\sigma ds \\
&= e^{i\Delta_{kl} m(T+d)} \int_0^{T+d} \int_\sigma^{T+d} e^{i\Delta_{kj} s} e^{i\Delta_{jl}\sigma} \int_0^{d} g_0(s-\xi) g_0(\sigma - \xi) d\mu(\xi) ds d\sigma \\
&= e^{i\Delta_{kl} m(T+d)} \int_0^{d} e^{i\Delta_{kl}\xi} \int -\xi^{T+d-\xi} e^{i\Delta_{jl}\sigma} g_0(\sigma) \int_\sigma^{T+d-\xi} e^{i\Delta_{kj} s} g_0(s) ds d\sigma d\mu(\xi).
\end{aligned}
$$

Now, the upper limit of the integrals with respect to s and σ can be replaced by ∞ without modifying their values since supp $g_0 \subset [0, T]$ and $\xi \in [0, d]$. Hence they do not depend on ξ and by computing the integral with respect to the measure $d\mu(\xi)$ first we get:

$$
\int_0^d e^{i\Delta_{kl}\xi} d\mu(\xi) = \mathbb{E}(e^{i\Delta_{kl} d_m}) = \delta_{kl}
$$

Knowing that $k = l$ in order to get a non zero result, we can now compute the integrals with respect to s and σ using (37) without the complex conjugate part. In conclusion we have

$$
\begin{aligned}
&\int_0^t \int_0^s e^{i\Delta_{kj} s} e^{i\Delta_{jl}\sigma} \mathbb{E}(g(s)g(\sigma)) d\sigma ds = \\
&\qquad \left\lfloor \frac{t}{T+d} \right\rfloor \frac{\delta_{kl}}{2} \left(|\hat{g}_0(-\Delta_{kj})|^2 + \frac{i}{\pi} \text{P.V.} \int_{-\infty}^{\infty} \frac{|g(\mu)|^2}{\mu + \Delta_{kj}} d\mu \right) + \mathcal{O}(1)
\end{aligned}
$$

where the correction is needed for $t \neq n(T+d)$. Consequently

$$
\lim_{t \nearrow \infty} \int_0^t \int_0^s e^{i\Delta_{kj} s} e^{i\Delta_{jl}\sigma} \mathbb{E}(g(s)g(\sigma)) d\sigma ds
$$

$$= \frac{\delta_{kl}}{2(T+d)} \left(|\hat{g}_0(-\Delta_{kj})|^2 + \frac{i}{\pi} \text{P.V.} \int_{-\infty}^{\infty} \frac{|g(\mu)|^2}{\mu + \Delta_{kj}} d\mu \right). \tag{150}$$

Replacing (150) in the formula (148) it is easy to see that $V_{pq,pq'} = 0$ unless $q = q'$ and in the later case the first and the fourth terms in (148) are complex conjugate which is true for the second and third terms also. Simple arithmetic leads to

$$\tilde{V}_{pq} \equiv V_{pq,pq} = \tilde{B}_{pq}$$

where $\tilde{B}$ is given by (140) and (22).

In conclusion, on time scales of order $1/\varepsilon^2$ our results for the example in this section coincides with the one obtained by Papanicolaou *et. al.* in the series of papers [17, 19, 18, 13]. As mentioned earlier, although our result applies directly, the stochastic approach requires the $\mathbb{E}(g(t)) = 0$ for all t.

6 Appendix: Properties of the power transmission matrix

In this section we prove the properties of the matrix (linear operator) B we used in Corollaries 3.1 and 3.2. Recall that B is given by (22) and is irreducible, see the discussion before Corollary 3.2. We note that (22) implies in particular that:

1. all the components of B are real;
2. $b_{ii} \geq 0$ for all $i = 1, 2, \ldots$;
3. $b_{ij} \leq 0$ for all $i,\ j,\ i \neq j$;
4. $\sum_j b_{ij} = 0$ or equivalently $b_{ii} = -\sum_{j, j\neq i} b_{ij}$ for all $i = 1, 2, \ldots$.

Lemma 6.1. *If the dimension of B is finite, say m, then B is a nonnegative, self adjoint matrix having 0 as a simple eigenvalue with corresponding normalized eigenvector:*

$$r_0 = \frac{1}{\sqrt{m}}(1, 1, \ldots, 1)'$$

Proof. Since all components of B are real, self adjointness is equivalent to

$$b_{jk} = b_{kj}, \qquad \forall j,\ k,\ j \neq k.$$

From (27) we have

$$|\alpha_{jk}|^2 = |\overline{\alpha}_{kj}|^2 = |\alpha_{kj}|^2,$$

where $\overline{\alpha}$ denotes the complex conjugate of the complex number α.

Now because $\hat{g}_0$ is the Fourier transform of a real valued function, see (11) and (H3), and because $\Delta_{kj} = -\Delta_{jk}$, see (26), we have

$$|\hat{g}_0(-\Delta_{jk})|^2 = |\hat{g}_0(\Delta_{kj})|^2 = \left|\overline{\hat{g}_0(-\Delta_{kj})}\right|^2 = |\hat{g}_0(-\Delta_{kj})|^2 .$$

Hence, for all $j \neq k$

$$b_{jk} = -|\alpha_{jk}|^2 |\hat{g}_0(-\Delta_{jk})|^2 = |\alpha_{kj}|^2 |\hat{g}_0(-\Delta_{kj})|^2 = b_{kj}, \tag{151}$$

rendering B self adjoint[3].

In order to prove that B is nonnegative, consider an arbitrary vector $X = (X_1, X_2, \ldots, X_m)'$ and let $X^* = (\overline{X}_1, \overline{X}_2, \ldots, \overline{X}_m)$ denote its adjoint. Then

$$\begin{aligned} X^*BX &= \sum_{i,j=1}^{m} b_{ij}\overline{X}_i X_j = \sum_{i=1}^{m} b_{ii}|X_i|^2 + \sum_{i,j,i\neq j} b_{ij}\overline{X}_i X_j \\ &= -\sum_{i,j,i\neq j} b_{ij}|X_i|^2 + \sum_{i,j,i\neq j} b_{ij}\overline{X}_i X_j \\ &= \sum_{i,j,i<j} |b_{ij}| \cdot |X_i - X_j|^2 \geq 0, \end{aligned} \tag{152}$$

where we used properties 3. and 4. above. Hence B is nonnegative.

Now, if $Y = Br_0$ then

$$Y_i = \frac{1}{\sqrt{m}} \sum_{i,j}^{m} b_{ij} = 0,$$

by property 4. Consequently 0 is an eigenvalue for B with corresponding eigenvector r_0.

To prove that 0 is a simple eigenvalue we use the irreducibility of B. On the set of components $\{1, 2, \ldots, m\}$ of vectors in $\mathbb{C}^m$ we define the following relation:

Definition 6.1. *We say that components i and i are always coupled to zeroth order.*

We say that components i, j are coupled to first order if $b_{ij} \neq 0$.

We say that components i, j are coupled to n^{th}, $n \geq 2$ order if there exists a sequence of components $k_1, k_2, \ldots, k_{n-1}$, such that the pairs $1, k_1$; $k_1, k_2; \ldots; k_{n-1}, j$; are all coupled to first order.

We say that components i, j are coupled if they are coupled to any order.

[3]Identity (151) does not rely upon B having a finite dimension. Therefore it is valid even when B has infinite dimension.

It is easy to show that "to be coupled" is an equivalence relation on the set of components $\{1, 2, \ldots, m\}$. Hence it induces a partition of the components.

Claim 1. If B is irreducible the above partition is trivial.

Indeed, if we assume contrary the partition is formed by at least two proper subsets of the set of components $\{1, 2, \ldots, m\}$. By a reordering of the components, i.e. a reordering of the standard basis vectors in $\mathbb{C}^m$, we can assume assume that the partition is formed by:

$$\{1, 2, \ldots, m_1\},\ \{m_1+1, m_1+2, \ldots, m_2\}, \ldots$$

Then $b_{ij} = 0$ whenever i, j fall in different subsets of the partition, otherwise they would be coupled. Consequently, B has the form:

$$B = \operatorname{diag}\left[B_1, B_2, \ldots\right],$$

where B_1 is a $m_1 \times m_1$ matrix, B_2 is a $m_2 \times m_2$ matrix, etc. But these contradict the irreducibility of B, see also the discussion before Corollary 3.1.

Claim 2. If $X = (X_1, X_2, \ldots, X_m)'$ is a zero eigenvector for B and i, j are coupled then $X_i = X_j$.

Indeed, $X^*BX = 0$ because $BX = 0$, and (152) implies

$$\sum_{i,j,i<j} |b_{ij}| \cdot |X_i - X_j|^2 = 0. \tag{153}$$

If i, j are coupled to the first order then by definition $b_{ij} \neq 0$ and we must have $X_i = X_j$ in order for (153) to hold. By induction on the order of coupling one obtains the result of the claim.

Finally, Claim 1 and the irreducibility of B imply that all components are coupled. Then Claim 2 implies that all components of a zero eigenvector must be equal. Hence all zero eigenvectors are parallel to r_0. Since B is self adjoint this means that 0 is a simple eigenvalue. []

Lemma 6.2. *If B is infinite dimensional, then B is a bounded linear operator on ℓ^1 with $\|B\|_1 \leq 2$. In addition, for $|\varepsilon| \leq 1$, the operator $T_\varepsilon = \mathbb{I} - \varepsilon^2 B$ transforms positive vectors (i.e. vectors with all components positive) into positive vectors and conserves their ℓ^1 norm.*

Proof. It is well known that $B = (b_{ij})_{1\le i,k<\infty}$ is a bounded linear operator on ℓ^1 iff there exists a constant $C \ge 0$ such that:

$$\sum_{i=1}^{\infty} |b_{ij}| \le C, \qquad \forall j = 1, 2, \dots \tag{154}$$

In this case$\|B\|_1 \le C$. We are going to show that for B given by (22) we can choose $C = 2$ in (154).

Indeed, let us fix an arbitrary $j \in \{1, 2, \dots\}$ and consider the j^{th} vector in the standard basis of ℓ^1 :

$$X = (X_1, X_2, \dots)', \quad X_i = \begin{cases} 0 & \text{if } i \ne j \\ 1 & \text{if } i = j \end{cases} \tag{155}$$

Let

$$\begin{aligned} A &= (a_{ij})_{1\le i,j<\infty} \\ a_{ij} &= \alpha_{ij}\ \hat{g}_0(-\Delta_{ij}) = \langle \psi_i, \beta\psi_j\rangle \int_{-\infty}^{\infty} g_0(t) e^{i(\lambda_i-\lambda_j)t} dt. \end{aligned} \tag{156}$$

By a direct calculation we have

$$\begin{aligned} \sum_{i=1}^{\infty} |b_{ij}| &= \sum_{i=1}^{\infty} \left| \overline{X}_i \sum_{k,p=1}^{\infty} a_{ik} a_{kp} X_p - \overline{\sum_{k=1}^{\infty} a_{ik} X_k} \sum_{k=1}^{\infty} a_{ik} X_k \right| \\ &= \sum_{i=1}^{\infty} \left| \overline{X}_i (A \cdot AX)_i - \left(\overline{AX}\right)_i (AX)_i \right| \\ &\le \sum_{i=1}^{\infty} |X_1| \cdot |(A \cdot AX)_i| + \sum_{i=1}^{\infty} |(AX)_i|^2. \end{aligned} \tag{157}$$

Clearly $X \in \ell^2$, $\|X\|_2 = 1$. We are going to prove below that:

Claim 3. A is a bounded operator on ℓ^2 with $\|A\|_2 \le 1$.

Hence

$$\sum_{i=1}^{\infty} |(AX)_i|^2 = \|AX\|_2^2 \le \|X\|_2^2 = 1, \tag{158}$$

while using Cauchy-Buniakowski-Schwarz inequality we have:

$$\begin{aligned} \sum_{i=1}^{\infty} |X_1| \cdot |(A \cdot AX)_i| &\le \left(\sum_{i=1}^{\infty} |X_i|^2 \right)^{1/2} \left(\sum_{i=1}^{\infty} |(A \cdot AX)_i|^2 \right)^{1/2} \\ &= \|X\|_2 \cdot \|A \cdot AX\|_2 \le \|A\|_2^2 \|X\|_2 \\ &\le 1. \end{aligned} \tag{159}$$

By plugging in (158) and (159) into (157) we get

$$\sum_{i=1}^{\infty} |b_{ij}| \leq 2 \tag{160}$$

and, since j was arbitrary, (154) holds with $C = 2$. Consequently, B is a bounded linear operator on ℓ^1 with norm $\|B\|_1 \leq 2$.

Consider now

$$T_\varepsilon = \mathbb{I} - \varepsilon^2 B, \qquad T_\varepsilon = (t_{ij})_{1 \leq i,j < \infty}.$$

Then for $i \neq j$,

$$t_{ij} = -\varepsilon^2 b_{ij} \geq 0.$$

Note that by (151) we also have:

$$t_{ji} = t_{ij}, \qquad \forall i, j. \tag{161}$$

On the other hand

$$t_{ii} = 1 - \varepsilon^2 b_{ii} = 1 - \varepsilon^2 \sum_{j, j \neq i} |b_{ij}|,$$

where we used properties 3. and 4. above. Moreover

$$\sum_{j=1}^{\infty} |b_{ij}| = |b_{ii}| + \sum_{j, j \neq i} |b_{ij}| = \left| - \sum_{j, j \neq i} b_{ij} \right| + \sum_{j, j \neq i} |b_{ij}| = 2 \sum_{j, j \neq i} |b_{ij}|.$$

Using now (160) and (151) we have

$$\sum_{j, j \neq i} |b_{ij}| = \frac{1}{2} \sum_{j=1}^{\infty} |b_{ij}| = \frac{1}{2} \sum_{j=1}^{\infty} |b_{ji}| \leq 1. \tag{162}$$

Hence

$$t_{ii} = 1 - \varepsilon^2 \sum_{j, j \neq i} |b_{ij}| \geq 1 - \varepsilon^2 \geq 0, \quad \text{if } |\varepsilon| \leq 1.$$

We also have:

$$\sum_{j=1}^{\infty} t_{ij} = t_{ii} + \sum_{j, j \neq i} t_{ij} = 1 - \varepsilon^2 \sum_{j, j \neq i} |b_{ij}| + \varepsilon^2 \sum_{j, j \neq i} |b_{ij}| = 1,$$

and by (161)

$$\sum_{i=1}^{\infty} t_{ij} = \sum_{i=1}^{\infty} t_{ji} = 1. \tag{163}$$

Now let

$$X = (X_1, X_2, \ldots)' \in \ell^1, \qquad X_j > 0 \ \forall j = 1, 2, \ldots$$

Then

$$(T_\varepsilon X)_i = \sum_{j=1}^{\infty} t_{ij} X_j > 0$$

since all terms in the sum are nonnegative with at least one being strictly positive. Moreover

$$\|T_\varepsilon X\|_1 = \sum_{i=1}^{\infty} |(T_\varepsilon X)_i| = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} t_{ij} X_j = \sum_{j=1}^{\infty} X_j \sum_{i=1}^{\infty} t_{ij} = \sum_{j=1}^{\infty} X_j = \|X\|_1,$$

where we exchanged the order of summation because we are dealing with convergent series with nonnegative terms and we also used (163).

The Lemma is now finished provided we prove Claim 3. Let

$$X = (X_1, X_2, \ldots)' \in \ell^2, \qquad \|X\|_2 = 1$$

be arbitrary and denote by

$$Y(t) = \sum_{j=1}^{\infty} e^{i\lambda_j t} X_j \psi_j, \quad Y(t) \in \mathcal{H}, \quad \|Y(t)\| \equiv 1.$$

Then

$$\begin{aligned} |X^* A X| &= \left| \sum_{j,k=1}^{\infty} a_{jk} \overline{X}_j X_k \right| = \left| \int_{-\infty}^{\infty} g_0(t) \langle \textstyle\sum_{j=1}^{\infty} e^{i\lambda_j t} X_j \psi_j, \sum_{k=1}^{\infty} e^{i\lambda_k t} X_k \beta \psi_k \rangle dt \right| \\ &= \left| \int_{-\infty}^{\infty} g_0(t) \langle Y(t), \beta Y(t) \rangle dt \right| \le \int_{-\infty}^{\infty} |g_0(t)| \cdot |\langle Y(t), \beta Y(t) \rangle| dt \\ &\le \int_{-\infty}^{\infty} |g_0(t)| \cdot \|\beta\|_{\mathcal{H}} \|Y(t)\|^2 dt = \|\beta\|_{\mathcal{H}} \int_{-\infty}^{\infty} |g_0(t)| dt = \|\beta\|_{\mathcal{H}} \|g_0\|_{L^1} = 1 \end{aligned}$$

where, at the very end, we used (H2) and (H3). []

Lemma 6.3. *If B is infinite dimensional, then B is a bounded, linear, self adjoint, nonnegative operator on ℓ^2 with spectral radius less or equal to 2. Moreover, 0 is not an eigenvalue for B.*

<u>Proof.</u> Because of (151) B is symmetric on ℓ^2. Consider the 2-form induced by B on ℓ^2:

$$\begin{aligned} X^* B X &= \sum_{i,j=1}^{\infty} b_{ij} \overline{X}_i X_j = \sum_{i=1}^{\infty} b_{ii} |X_i|^2 + \sum_{i=1}^{\infty} \sum_{j, j \ne i}^{\infty} b_{ij} \overline{X}_i X_j \\ &= -\sum_{i=1}^{\infty} \sum_{j, j \ne i}^{\infty} b_{ij} \bar{|} X_i|^2 + \sum_{i=1}^{\infty} \sum_{j, j \ne i}^{\infty} b_{ij} \overline{X}_i X_j = \sum_{i=1}^{\infty} \sum_{j, j \ne i}^{\infty} |b_{ij}| (|X_i|^2 - \overline{X}_i X_j), \quad (164) \end{aligned}$$

where we used properties 3. and 4. above. Hence

$$\begin{aligned} X^*BX &= \sum_{i=1}^{\infty}\sum_{j,j\neq i}^{\infty}|b_{ij}|\left(|X_i|^2+\frac{|X_i|^2+|X_j|^2}{2}\right)=2\sum_{i=1}^{\infty}\sum_{j,j\neq i}^{\infty}|b_{ij}|\cdot|X_i|^2 \\ &\leq 2\sup_i\left(\sum\nolimits_{j,j\neq i}|b_{ij}|\right)\sum_{i=1}^{\infty}|X_i|^2\leq 2\|X\|_2^2, \end{aligned}$$

where we used (162). So the 2-form induced by B is bounded. Since B is symmetric this implies that B is a self adjoint, bounded operator on ℓ^2 with $\|B\|_2 \leq 2$. Therefore its spectral radius is less or equal to 2.

Moreover, (164) implies

$$X^*BX = \sum_{i<j}|b_{ij}|\cdot|X_i - X_j|^2 \geq 0.$$

On one hand this shows that B is nonnegative. On the other hand, together with obvious generalizations of Claims 1 and 2 in Lemma 6.1 for the case of infinitely but countable components, it shows that if a zero eigenvector for the irreducible operator B exists then the eigenvector should have all components equal. However such a vector is not in ℓ^2 unless it is trivial. Therefore 0 is not an eigenvalue for B. []

References

[1] L. Bunimovich, H. R. Jauslin, J. L. Lebowitz, A. Pellegrinotti, and P. Nielaba. Diffusive energy growth in classical and quantum driven oscillators. *J. Statist. Phys.*, 62(3-4):793–817, 1991.

[2] C. Corduneanu. *Almost periodic functions.* Interscience Publishers [John Wiley & Sons], New York-London-Sydney, 1968.

[3] O. Costin and A. Soffer. Resonance theory for Schrödinger operators. *Comm. Math. Phys.*, 224(1):133–152, 2001.

[4] H. L. Cycon, R. G. Froese, W. Kirsch, and B. Simon. *Schrödinger operators with application to quantum mechanics and global geometry.* Springer-Verlag, Berlin, 1987.

[5] E.B. Davies. *Quantum theory of open systems.* Academic Press, New York, 1976.

[6] R. Durrett. *Probability: theory and examples.* Duxbury Press, Belmont, CA, 1996.

[7] J. Garnier. Asymptotic behavior of the quantum harmonic oscillator driven by a random time-dependent electric field. *J. Statist. Phys.*, 93(1-2):211–241, 1998.

[8] J. Garnier. Light propagation in square law media with random imperfections. *Wave Motion*, 31(1):1–19, 2000.

[9] E. Kirr and M. I. Weinstein. Parametrically excited Hamiltonian partial differential equations. *SIAM J. Math. Anal.*, 33(1):16–52 (electronic), 2001.

[10] E. Kirr and M. I. Weinstein. Almost periodic trains of pulses and their interaction with hamiltonian systems. In preparation, 2003.

[11] E. Kirr and M. I. Weinstein. Metastable states in parametrically excited multimode Hamiltonian systems. *Commun. Math. Phys.*, 236(2):335–372, 2003.

[12] Eduard-Wilhelm Kirr. *Resonances in Hamiltonian Systems.* Ph.d. thesis, University of Michigan, 2002.

[13] Werner Kohler and George C. Papanicolaou. Wave propagation in a randomly inhomogenous ocean. In *Wave propagation and underwater acoustics (Workshop, Mystic, Conn., 1974)*, pages 153–223. Lecture Notes in Phys., Vol. 70. Springer, Berlin, 1977.

[14] L. D. Landau and E. M. Lifshitz. *Quantum mechanics: non-relativistic theory. Course of Theoretical Physics, Vol. 3.* Pergamon Press Ltd., London-Paris, 1958.

[15] D. Marcuse. *Theory of Dielectric Optical Waveguides.* Academic Press, 1974.

[16] P. D. Miller, A. Soffer, and M. I. Weinstein. Metastability of breather modes of time-dependent potentials. *Nonlinearity*, 13(3):507–568, 2000.

[17] G. C. Papanicolaou and R. Hersh. Some limit theorems for stochastic equations and applications. *Indiana Univ. Math. J.*, 21:815–840, 1971.

[18] G. C. Papanicolaou and W. Kohler. Asymptotic theory of mixing stochastic ordinary differential equations. *Comm. Pure Appl. Math.*, 27:641–668, 1974.

[19] George C. Papanicolaou. A kinetic theory for power transfer in stochastic systems. *J. Mathematical Phys.*, 13:1912–1918, 1972.

[20] W. Pauli. *Festschrift zum 60. Geburstage Sommerfelds.* Hirzel, Leipzig, 1928.

[21] Michael Reed and Barry Simon. *Methods of modern mathematical physics. I. Functional analysis.* Academic Press, New York, 1972.

[22] C. O. Reinhold, S. Yoshida, J. Burgdörfer, B. E. Tannian, C. L. Stokely, and Dunning F. B. Designing rydberg wavepackets using trains of half-cycle pulses. *J. Phys. B: At. Mol. Opt. Phys.*, 34(17):L551–L558, 2001.

[23] A. Soffer and M. I. Weinstein. Ionization and scattering for short-lived potentials. *Lett. Math. Phys.*, 48(4):339–352, 1999.

[24] A. Soffer and M. I. Weinstein. Resonances, radiation damping and instability in Hamiltonian nonlinear wave equations. *Invent. Math.*, 136(1):9–74, 1999.

[25] L. Van Hove. *"Master equations and approach to equilibrium for quantum systems", in Fundamental problems in statistical mechanics, edited by E.G.D. Cohen.* North Holland, Amsterdam, 1962.

[26] Kenji Yajima. Scattering theory for Schrödinger equations with potentials periodic in time. *J. Math. Soc. Japan*, 29(4):729–743, 1977.