Kinetic theory of nonlinear diffusion in a weakly disordered nonlinear Schrödinger chain in the regime of homogeneous chaos

D. M. Basko
Université Grenoble 1/CNRS, LPMMC UMR 5493, 25 rue des Martyrs, 38042 Grenoble, France
(Received 12 December 2013; published 21 February 2014)

We study the discrete nonlinear Schröinger equation with weak disorder, focusing on the regime when the nonlinearity is, on the one hand, weak enough for the normal modes of the linear problem to remain well resolved but, on the other, strong enough for the dynamics of the normal mode amplitudes to be chaotic for almost all modes. We show that in this regime and in the limit of high temperature, the macroscopic density ρ satisfies the nonlinear diffusion equation with a density-dependent diffusion coefficient, $D(\rho) = D_0\rho^2$. An explicit expression for D_0 is obtained in terms of the eigenfunctions and eigenvalues of the linear problem, which is then evaluated numerically. The role of the second conserved quantity (energy) in the transport is also quantitatively discussed.

DOI: 10.1103/PhysRevE.89.022921 PACS number(s): 05.45.−a, 63.20.Pw, 63.20.Ry, 72.15.Rn

I. INTRODUCTION

Disordered classical nonlinear chains are convenient model systems where the interplay between Anderson localization and classical nonlinearity can be studied. Indeed, Anderson localization is especially pronounced in one spatial dimension, where even an arbitrarily weak disorder localizes all normal modes of the linear system [1,2]. A nonlinearity couples the normal modes, which may lead to chaotic dynamics [3,4] and destroy Anderson localization [5–8], although periodic solutions may be preserved with some probability [9,10]. Viewed from the nonlinear dynamics side, disordered nonlinear chains represent a class of dynamical systems with an infinite number of degrees of freedom, in which chaos may have a local structure in the real space [11–16], and this structure can be controlled by the disorder strength.

Here we study the discrete nonlinear Schrödinger equation with disorder (DNLSE) [see Eq. (2) below]. It describes several physical systems, such as light propagating in nonlinear photonic lattices [17] or cold bosonic atoms in optical lattices in the mean-field approximation [18]. One of the fundamental problems concerning DNLSE, as well as disordered nonlinear chains in general, is the evolution of an initially localized wave packet at very long times (see Ref. [19] for a review). In a linear system with Anderson localization, the wave packet remains exponentially localized at all times, due to the superposition principle. In the presence of a nonlinearity, the wave packet width was found to increase as a subdiffusive power of time t in the direct numerical simulations of DNLSE [5–8,20–23]: The second moment m_2 of the wave packet was growing as $m_2 \propto t^p$ with $p < 1$ while for the usual diffusive spreading, $p = 1$. Specifically, two regimes of such subdiffusive spreading have been identified numerically: at longest observed times and at low densities, $p \approx 1/3$, but an intermediate range of densities and times has also been found with $p \approx 1/2$. These regimes were called weak and strong chaos, respectively [22,23].

At the same time, rigorous mathematical arguments suggest that at long times the spreading, if any, should be slower than any power of t [24,25]. Analysis of perturbation theory in the nonlinearity suggests that there is a front propagating as $\ln t$ beyond which the wave packet is localized exponentially [26]. These arguments can be reconciled with those of the direct numerical simulations if one assumes that the numerically observed behavior $m_2 \propto t^{1/3}$ is still an intermediate asymptotics, which should eventually break down at very long times. An indication for a slowing down of the power-law subdiffusion has been observed in the scaling analysis of numerical results [27,28] and was related to the scaling of the probability of chaos with the density [29]. A possible mechanism for the breakdown of subdiffusion at long times has been suggested [30]. However, progress in this direction is impeded by the absence of a satisfactory quantitative theory for the observed power-law subdiffusion, which would (i) elucidate the main mechanism responsible for such subdiffusion and (ii) predict quantitatively when it ceases to work.

The purpose of the present work is to construct such a theory for the intermediate spreading regime where $m_2 \propto t^{1/2}$, called "strong chaos" in Refs. [22,23]. However, the arguments of the present work are very much analogous to those used in the derivation of the kinetic equation in the theory of weak wave turbulence [31]. Thus, to avoid the terms "strong" and "weak," which may be confusing, we use the term "homogeneous chaos" to denote the studied regime. The main element of the physical picture of transport developed here is the large localization length of normal modes of the linear problem, which occurs for a weak-enough disorder. In this case, the nonlinearity couples each normal mode to *many* other modes, and, if it is not too weak, the dynamics of almost all modes is chaotic. So, the chaos can be considered homogeneous both in the real space and in the mode space. The same physical image was proposed in Refs. [22,23] to identify the spreading regime with $m_2 \propto t^{1/2}$.

It has already been argued on phenomenological grounds that macroscopic transport of the conserved density ρ in DNLSE and other disordered nonlinear chains can be described by a nonlinear diffusion equation,

$$\frac{\partial\rho}{\partial t} = \frac{\partial}{\partial x}\left(D(\rho)\,\frac{\partial\rho}{\partial x}\right), \tag{1}$$

with a ρ-dependent diffusion coefficient $D(\rho)$ [32,33]. Indeed, for a power-law $D(\rho) = D_0\rho^a$ with some $a > 0$, Eq. (1) gives $m_2 \propto t^{2/(a+2)}$. Such a power-law dependence was found numerically in Ref. [34]. An expression interpolating

between $D(\rho) \propto \rho^2$ at higher densities and ρ^4 at lower densities has been proposed [35], giving a crossover from $m_2 \propto t^{1/2}$ to $t^{1/3}$ at longer times. But, again, a recent matematical work suggests that at lowest densities, $D(\rho)$ should vanish faster than any power of ρ [36]. An explicit expression satisfying this condition has been given for the DNLSE in the limit of strong disorder and very low density, based on the picture of chaos which is very inhomogeneous in space [12]. Detailed quantitative understanding of the mechanisms leading to Eq. (1) with a power-law $D(\rho)$ is still lacking.

In the present paper, starting from the DNLSE with weak disorder, the nonlinear diffusion equation, Eq. (1) with $D(\rho) = D_0\rho^2$, is derived. An explicit expression [Eq. (60)] for D_0 is given in terms of a certain average of eigenfunctions and eigenvalues of the linear part of the DNLSE, which is then evaluated numerically. The resulting value of D_0 is in reasonable agreement with that extracted from the direct numerical simulation of Ref. [22]. The conditions of validity of the present approach are thoroughly discussed, and the corresponding interval of densities is identified (Fig. 5). Its lower boundary agrees with the value at which the crossover between $m_2 \propto t^{1/2}$ and $t^{1/3}$ behavior is observed in Ref. [22].

A specific property of the DNLSE is the presence of *two* conserved quantities: the total norm and the total energy. Its consequence for the thermodynamics of the system is the existence of the so-called nonthermal states which can be described in the microcanonical ensemble but not in the grand canonical one [37]. The consequence for the transport is that Eq. (1) for the norm density ρ is not complete; the complete macroscopic description should involve two coupled equations for the norm and energy. Such coupled transport has received relatively little attention so far: A numerical study for the DNLSE without disorder is available [38], and in the limit of strong disorder, low density, and high temperature, the off-diagonal transport coefficients were estimated to be small [12]. Here, the full 2×2 matrix of the transport coefficients is calculated for weak disorder, intermediate density, and high temperature. The analysis of the full coupled equations shows that in the considered regime the effect of energy transport on the norm transport is small, so Eq. (1) for ρ only is consistent.

The paper is organized as follows. After introducing the model in Sec. II, in Sec. III we derive a Fokker-Planck-type master equation describing diffusive dynamics of the normal mode intensities. Conditions for its validity and its relation to chaos are discussed in Sec. IV. In Sec. V, the macroscopic Eq. (1) is derived from the microscopic master equation, via a Boltzmann-type kinetic equation for average intensities under the assumption of local thermal equilibrium. In Sec. VI, the role of the second conserved quantity (energy) is discussed, and the coupled macroscopic equations are analyzed. Finally, in Sec. VII, the formal expressions for the transport coefficients derived in Secs. V and VI are evaluated for different disorder strength and compared to the numerics of Ref. [22].

II. MODEL AND ASSUMPTIONS

The DNLSE reads as

$$i\,\frac{d\psi_n}{dt} = -\Omega\,(\psi_{n+1} + \psi_{n-1}) + \epsilon_n\psi_n + g|\psi_n|^2\psi_n, \qquad (2)$$

Here $n = 1, \ldots, L$ labels sites of a one-dimensional lattice (the limit $L \to \infty$ will be eventually taken). To each site n a pair of complex conjugate variables ψ_n, ψ_n^* is associated. The on-site frequencies ϵ_n are assumed to be random and uncorrelated and to have the flat box distribution $\epsilon_n \in [-W/2, W/2]$ whose width W characterizes the disorder strength. Ω and g measure the intersite coupling and the strength of the nonlinearity. The sign of g is not important, as Eq. (2) is invariant under the change $g \to -g$, $\psi_n \to (-1)^n\psi_n^*$, $\epsilon_n \to -\epsilon_n$ (for the latter it is important that the distribution of ϵ_n is symmetric).

Equation (2) together with its complex conjugate represent the Hamilton equations for the classical Hamiltonian

$$H = \sum_n \left[\epsilon_n|\psi_n|^2 - \Omega(\psi_n^*\psi_{n+1} + \psi_{n+1}^*\psi_n) + \frac{g}{2}|\psi_n|^4\right] \qquad (3)$$

if $i\psi_n^*$ is treated as the canonical momentum conjugate to the coordinate ψ_n. One could measure time, frequency, action, and energy in the units of $1/\Omega$, Ω, Ω/g, and Ω^2/g, respectively, thus setting $\Omega = 1$, $g = 1$ in Eq. (2) without loss of generality. However, formal expressions are sometimes more transparent physically when written in the dimensional form, so we assume Ω and ψ_n to have the dimensionality of frequency and $(\text{action})^{1/2}$, respectively.

The linear part of the Hamiltonian can be diagonalized by an orthogonal transformation

$$\psi_n(t) = \sum_{\alpha=1}^{L} c_\alpha(t)\,\phi_{\alpha n}, \qquad (4)$$

where $\phi_{\alpha n}$ is the αth eigenfunction of the linear problem

$$\omega_\alpha\phi_{\alpha n} = \epsilon_n\phi_{\alpha n} - \Omega(\phi_{\alpha,n+1} + \phi_{\alpha,n-1}), \qquad (5)$$

corresponding to the eigenvalue ω_α, and c_α is the complex amplitude of the αth normal mode. The normal mode amplitudes satisfy the following equations of motion:

$$i\,\frac{dc_\alpha}{dt} = \omega_\alpha c_\alpha + \sum_{\beta,\gamma,\delta} V_{\alpha\beta\gamma\delta}c_\beta^* c_\gamma c_\delta. \qquad (6)$$

Here we introduced the overlap,

$$V_{\alpha\beta\gamma\delta} = g\sum_n \phi_{\alpha n}\phi_{\beta n}\phi_{\gamma n}\phi_{\delta n}, \qquad (7)$$

which is real and symmetric with respect to any permutations of the mode indices. It is a random quantity, and its statistics will play a crucial role in determining the macroscopic transport properties of the system and their dependence on the disorder strength.

The normal mode frequencies ω_α lie in the interval $|\omega_\alpha| < 2\Omega + W/2$. Their distribution in this interval is determined by the average spectral density per unit length, $\nu_1(\omega)$. All normal mode wave functions are localized, the localization length $\xi(\omega)$ depending on the mode frequency. The behavior of $\nu_1(\omega)$ and $\xi(\omega)$ for weak disorder is discussed in detail in Appendix A1. Here we only note that the localization length is the largest for modes close to the center of the band,

$\xi(\omega = 0) \approx 100\,(\Omega/W)^2$, where $\xi(\omega)$ is of the same order in the most of the band and becomes small near the band edges. The frequency spacing between the modes which are on the same localization segment $\Delta_1(\omega) = [\nu_1(\omega)\,\xi(\omega)]^{-1}$ is the smallest at $\omega = 0$.

For the transport mechanism discussed in the present work, it is crucial that $V_{\alpha\beta\gamma\delta}$, defined in Eq. (7), couples many modes. This is only possible when their localization lengths are much larger than the lattice spacing, otherwise $V_{\alpha\beta\gamma\delta}$ is exponentially suppressed. The representative estimate for the localization length is $\xi(\omega = 0) \equiv \xi$, since the few tightly localized modes near the band edges contribute little to the transport. Thus, the main assumption of the present work which determines the applicability of the whole approach, is $\xi \gg 1$, corresponding to $W \ll 10\,\Omega$. Note that omitting the numerical factor from this condition would result in an unnecessarily severe restriction; in fact, the wave packet spreading with $m_2 \propto t^{1/2}$ was observed in Ref. [22] for $W/\Omega = 4$. At the same time, it turns out that for several quantities determined by the statistics of wave functions, the asymptotic behavior corresponding to weakest disorder is reached at really small $W \lesssim 0.3\,\Omega$, corresponding to $\xi \gtrsim 1000$ (see Sec. IV C, Appendix A2, and Ref. [39]).

The change of variables (4) is useful only if the dynamics of the normal mode amplitudes, c_α, due to the last term on the right-hand side of Eq. (6), is not too fast. The relevant time scale at which the dynamics of the system allows us to resolve the individual modes, is determined by the frequency spacing on one localization segment. The representative value is $\Delta_1(\omega = 0) \equiv \Delta_1 = 2\pi\Omega/\xi$, so the dynamics of c_α should be slow on the time scale $1/\Delta_1$. Since the nonlinear dynamics is faster for larger amplitudes, this condition imposes an upper limit on the typical norm density, $|\psi_n|^2 \sim |c_\alpha|^2 \sim \rho \ll \rho_{\max}$, where the value of $\rho_{\max}$ depends on the disorder strength, and is discussed in Sec. IV C. When the nonlinear dynamics is too fast, the localized normal modes are not a good starting point, so one should start from the freely propagating waves and treat the disorder as a perturbation (in dimensions higher than 1, one starts from the diffusive linear dynamics [40]).

On the other hand, the kinetic approach developed below works if the mode dynamics is sufficiently chaotic. For this, the nonlinearity should be strong enough, so the density should not be too low, $\rho \gg \rho_{\min}$. This condition is put into a quantitative form in Sec. IV A. It will be seen that the two restrictions $\rho \ll \rho_{\max}$ and $\rho \gg \rho_{\min}$ are not in conflict with each other when the condition $\xi \gg 1$ is satisfied.

Finally, the temperature T is assumed to be sufficiently high, $T \gg \Omega\rho$. This implies that there is no correlation between the intensities $|c_\alpha|^2$ and the frequencies ω_α in the initial condition for Eq. (2). It is this situation that was studied in the numerical works [5–8,20–23]. In the opposite case (low T), one should minimize Hamiltonian (3) first and then study the dynamics of the low-energy excitations. The ground state of this classical Hamiltonian at fixed extensive total norm is the Bose condensate, which is spatially nonuniform due to the disorder. The normal modes of Eq. (2), linearized around such a nonuniform condensate (the Bogolyubov modes), differ strongly from the solutions of the eigenvalue problem (5) [41,42]. The assumption of high temperature enables one to disregard the condensate and focus on the normal modes of the linear problem (5).

III. MICROSCOPIC MASTER EQUATION

A. Nonlinear frequency shifts

In the absence of the nonlinear coupling, the solution for the normal mode amplitudes is

$$c_\alpha = \sqrt{I_\alpha}\, e^{-i\theta_\alpha^0 - i\omega_\alpha t}, \tag{8}$$

where I_α and θ_α^0 are the intensity and the initial phase of the mode α, respectively. With the nonlinearity included, among the L^3 terms contributing to the sum on the right-hand side of Eq. (6), there are $2L - 1$ terms oscillating at the frequency ω_α (those for which either $\delta = \alpha, \gamma = \beta$ or $\gamma = \alpha, \delta = \beta$). These, the so-called secular terms, are responsible for the nonlinear shift of the αth frequency,

$$\tilde{\omega}_\alpha = \omega_\alpha + 2\sum_\beta V_{\alpha\beta\beta\alpha} I_\beta. \tag{9}$$

This can be seen by splitting the full Hamiltonian, corresponding to Eq. (6),

$$H = \sum_\alpha \omega_\alpha c_\alpha^* c_\alpha + \frac{1}{2}\sum_{\alpha,\beta,\gamma,\delta} V_{\alpha\beta\gamma\delta} c_\alpha^* c_\beta^* c_\gamma c_\delta \tag{10}$$

into

$$H_0 = \sum_\alpha \omega_\alpha |c_\alpha|^2 + \sum_{\alpha\beta} V_{\alpha\beta\beta\alpha} |c_\alpha|^2 |c_\beta|^2, \tag{11}$$

which is integrable, and all the remaining terms, which represent an integrability-breaking perturbation. The solution of the equations of motion for the Hamiltonian H_0 is given by the same Eq. (8) but with the modified frequencies, Eq. (9), which are nothing but $\tilde{\omega}_\alpha = \partial H_0/\partial(|c_\alpha|^2)$. Since the sum over β in Eq. (9) effectively contains a large number of terms, of the order of $\xi \gg 1$, I_β can be effectively replaced by its average, $\langle I_\beta \rangle \equiv \rho$. Then, using the orthogonality of the eigenmode wave functions $\phi_{\beta n}$, we obtain

$$\tilde{\omega}_\alpha \approx \omega_\alpha + 2g\rho. \tag{12}$$

The precision of this approximation can be quantified by calculating the average relative fluctuation σ^2 of the nonlinear frequency shift,

$$\langle(\tilde{\omega}_\alpha - \omega_\alpha - 2g\rho)^2\rangle = 4\rho^2 \left\langle \sum_\beta V_{\alpha\beta\beta\alpha}^2 \right\rangle \equiv 4g^2\rho^2\sigma^2, \tag{13}$$

where averaging of the quantities that depend on $V_{\alpha\beta\beta\alpha}$ is performed over the disorder realizations, and of those which depend on I_β is over the thermal distribution $\prod_\beta \rho^{-1} e^{-I_\beta/\rho}$ (see Sec. III D below for details). In particular, for this distribution, $\langle I_\beta I_\gamma \rangle = (1 + \delta_{\beta\gamma})\rho^2$, which gives Eq. (13).

Since the typical number of terms contributing to the sum over β is $\sim\xi$ (at larger distances $V_{\alpha\beta\beta\alpha}$ decays exponentially), it is natural to expect the relative fluctuation $\sigma^2 \sim 1/\xi \ll 1$. The same result is obtained by considering the exact

expression

$$\sum_{\beta} V^2_{\alpha\beta\beta\alpha} = g^2 \sum_{\beta} \sum_{n,n'} \phi^2_{\alpha n} \phi^2_{\alpha n'} \phi^2_{\beta n} \phi^2_{\beta n'}, \tag{14}$$

estimating $\phi^2_{\alpha n} \sim 1/\xi$ and assuming each summation to give a factor $\sim\xi$. The numerical test of this argument is performed in Appendix A2.

B. Diffusion in normal mode intensities

The nonsecular terms in the sum on the right-hand side of Eq. (6) (i.e., those contained in $H - H_0$) lead to exchange of energy between the normal modes, so the intensities I_α change with time. Effectively, only those β,γ,δ contribute to the sum, which correspond to the modes separated from the mode α by a distance $\sim\xi$ (for other modes, the overlap $V_{\alpha\beta\gamma\delta}$ is exponentially small). Among those, there are $O(\xi^3)$ terms for which all four indices α, β, γ, and δ differ and $O(\xi^2)$ terms for which at least two indices coincide. For $\xi \gg 1$, these latter terms can be ignored, as their role is analogous to those with differing α, β, γ, and δ, while their number is parametrically smaller.

Let us now consider just one term with different indices, that is, four oscillators described by the Hamiltonian

$$H(\{c_\alpha\}) = \sum_{\alpha=1}^{4} \omega_\alpha |c_\alpha|^2 + 2V_{1234}(c_1^* c_2^* c_3 c_4 + c_1 c_2 c_3^* c_4^*). \tag{15}$$

In principle, there are two other terms coupling the same four oscillators, $c_1^* c_3^* c_2 c_4$ and $c_1^* c_4^* c_2 c_3$ (with their complex conjugates), but we do not consider them for the moment. We have also neglected the nonlinear frequency shifts for simplicity. The first-order (in V_{1234}) correction to the unperturbed solution (8) for c_1 is given by

$$\begin{aligned} \Delta c_1 &= e^{-i\theta_1^0 - i\omega_1 t}\, 2V_{1234}\sqrt{I_2 I_3 I_4}\, e^{i\vartheta_{1234}}\, \frac{e^{1-i\varpi_{1234}t}}{\varpi_{1234}}, \\ \varpi_{1234} &\equiv \omega_1 + \omega_2 - \omega_3 - \omega_4, \\ \vartheta_{1234} &\equiv \theta_1^0 - \theta_2^0 - \theta_3^0 + \theta_4^0. \end{aligned} \tag{16}$$

The solutions for the other three oscillators are obtained by appropriate permutations. The change in the intensities of the oscillators, $\Delta I_\alpha = c_\alpha^*\, \Delta c_\alpha + c_\alpha\, \Delta c_\alpha^*$, is given by

$$\Delta I_1 = 2V_{1234}\sqrt{I_1 I_2 I_3 I_4}\, 2\,\mathrm{Re}\left(e^{-i\vartheta_{1234}}\, \frac{1 - e^{i\varpi_{1234}t}}{\varpi_{1234}} \right), \tag{17a}$$

$$\Delta I_2 = -\Delta I_3 = -\Delta I_4 = \Delta I_1. \tag{17b}$$

The increment ΔI_1 can be positive or negative, depending on the phases θ_α^0.

Let us now consider coupling of the mode $\alpha = 1$ to all other modes. The increment ΔI_α will be a sum of many terms with random signs, so the dynamics of I_α is diffusive. To quantify this dynamics, let us calculate the phase-averaged ΔI_α^2 as follows:

$$\begin{aligned} \langle \Delta I_\alpha^2 \rangle_\theta &= \frac{1}{2} {\sum_{\beta,\gamma,\delta}}' 4V^2_{\alpha\beta\gamma\delta} I_\alpha I_\beta I_\gamma I_\delta\, \frac{2\sin^2(\varpi_{\alpha\beta\gamma\delta}t/2)}{(\varpi_{\alpha\beta\gamma\delta}/2)^2} \\ &\approx \frac{1}{2} {\sum_{\beta,\gamma,\delta}}' 4V^2_{\alpha\beta\gamma\delta} I_\alpha I_\beta I_\gamma I_\delta\, 4\pi t\, \delta(\varpi_{\alpha\beta\gamma\delta}). \end{aligned} \tag{18}$$

where the prefactor $1/2$ prevents double counting which originates from the symmetry $\gamma \leftrightarrow \delta$, and the prime at the sum indicates that the secular terms should be excluded. The key feature of Eq. (18) is that $\langle \Delta I_\alpha^2 \rangle \propto t$, which is indeed a signature of the diffusive dynamics.

The δ function in Eq. (18) deserves some discussion. Formally, the presented derivation is quite analogous to the derivation of the Fermi golden rule for decay into a continuous spectrum, familiar from quantum mechanics [43] (in particular, the short-time limitation of its validity is determined by the inverse bandwidth, $t \gg 1/\Omega$). However, the sum in Eq. (18) is discrete, so the δ function cannot be understood in the strict sense. For any finite $\varpi_{\alpha\beta\gamma\delta}$, each individual term in the sum in the first line of Eq. (18) represents a growth $\propto t^2$ for $t \ll 1/|\varpi_{\alpha\beta\gamma\delta}|$, and saturation (with oscillations) at $t \gg 1/|\varpi_{\alpha\beta\gamma\delta}|$, so it never grows linearly in time. However, when considering the whole sum, only the terms with $|\varpi_{\alpha\beta\gamma\delta}| < 1/t$ contribute to the t^2 growth of ΔI_α^2. The number of such terms decreases $\sim 1/(\Delta_\varpi t)$, where $1/t$ is the typical width of the peak in $\varpi_{\alpha\beta\gamma\delta}$ and Δ_ϖ is the typical spacing between the values of $\varpi_{\alpha\beta\gamma\delta}$ effectively contributing to the sum for various β, γ, and δ (not to be confused with Δ_1, the spacing between single frequencies ω_α, as combinations of three frequencies are more numerous and dense). Thus, Eq. (18) only makes sense for times that are not too long, $t \ll 1/\Delta_\varpi$, when the number of terms contributing to the sum in the first line is large. This large number of terms which enter with random phases also justifies the phase averaging in Eq. (18). This condition of large number of terms is crucial for the validity of the whole approach and will be analyzed in detail in Sec. IV A. Note that even if the initial phases are not random (say, all $\theta_\alpha^0 = 0$), summation over many random frequencies has the same effect of suppressing the oscillating terms.

C. Master equation

To describe the diffusion of normal mode intensities I_α, we introduce the joint distribution function, $\mathcal{F}(\{I_\alpha\})$, which depends on all intensities. Equation (18) together with the constraint

$$\Delta I_\alpha = \Delta I_\beta = -\Delta I_\gamma = -\Delta I_\delta, \tag{19}$$

holding for each coupling term $V_{\alpha\beta\gamma\delta}$ [cf. Eq. (17b)], yield a diffusion equation of the following form:

$$\begin{aligned} \frac{\partial \mathcal{F}}{\partial t} = \frac{1}{8} {\sum_{\alpha,\beta,\gamma,\delta}}' &\left(\frac{\partial}{\partial I_\alpha} + \frac{\partial}{\partial I_\beta} - \frac{\partial}{\partial I_\gamma} - \frac{\partial}{\partial I_\delta} \right) \\ &\times 4V^2_{\alpha\beta\gamma\delta} I_\alpha I_\beta I_\gamma I_\delta\, 2\pi\, \delta(\tilde\omega_\alpha + \tilde\omega_\beta - \tilde\omega_\gamma - \tilde\omega_\delta) \\ &\times \left(\frac{\partial}{\partial I_\alpha} + \frac{\partial}{\partial I_\beta} - \frac{\partial}{\partial I_\gamma} - \frac{\partial}{\partial I_\delta} \right) \mathcal{F}, \end{aligned} \tag{20}$$

where the prefactor $1/8$ is introduced to prevent from double counting which originates from the symmetry $\alpha \leftrightarrow \beta$, $\gamma \leftrightarrow \delta$, and $(\alpha,\beta) \leftrightarrow (\gamma,\delta)$ (indeed, of 24 permutations of four different indices α, β, γ, and δ, only three produce distinct diffusion operators).

In Eq. (20) we have also included the nonlinear frequency shifts in the δ function: Each frequency $\tilde\omega_\alpha$ depends on all intensities I_β, as given by Eq. (9). In fact, all reasoning of

Sec. III B applies also to the case of shifted frequencies. Indeed, since the perturbative expressions of Sec. III B are valid only as long as $\sqrt{\langle \Delta I_\alpha^2 \rangle} \ll I_\alpha$ anyway, in the derivation of Eq. (18) the frequencies $\tilde{\omega}_\alpha$ can be considered fixed and determined by the instantaneous values of the intensities.

D. Equilibrium and relaxation

The master equation (20) conserves the total norm and total (unperturbed) energy,

$$\frac{d\langle \mathcal{N} \rangle_{\mathcal{F}}}{dt} = 0, \quad \mathcal{N} = \sum_\alpha I_\alpha , \tag{21a}$$

$$\frac{d\langle \mathcal{E} \rangle_{\mathcal{F}}}{dt} = 0, \quad \mathcal{E} = \sum_\alpha \omega_\alpha I_\alpha + \sum_{\alpha,\beta} V_{\alpha\beta\beta\alpha} I_\alpha I_\beta , \tag{21b}$$

the latter being ensured by the frequency δ function with the frequencies given by Eq. (9) (the consequences of the finite width of the δ function are discussed in Sec. VI B, and for the moment we proceed formally). Here we introduced the averaging over the distribution function in the following standard way:

$$\langle \cdots \rangle_{\mathcal{F}} = \int (\cdots) \mathcal{F}(\{I_\alpha\}) \prod_\alpha dI_\alpha . \tag{22}$$

Any distribution function $\mathcal{F}$ depending on the intensities through the combinations $\mathcal{N}$,$\mathcal{E}$, will be a stationary solution of the master equation. In particular, this is the case for the thermal distribution function

$$\mathcal{F}_{\rm eq} \propto e^{-(\mathcal{E}-\mu\mathcal{N})/T}, \tag{23}$$

where T is the temperature and μ is the chemical potential. This distribution is the one that maximizes the entropy $\langle \ln(e/\mathcal{F}) \rangle_{\mathcal{F}}$ [for which the H-theorem can be straightforwardly obtained from Eq. (20)], under the constraint of fixed average norm $\langle \mathcal{N} \rangle_{\mathcal{F}}$ and energy $\langle \mathcal{E} \rangle_{\mathcal{F}}$ (provided that the energy does not exceed a certain critical value [37], see the discussion in Sec. VI D). We will be interested in the high-temperature limit

$$T \to \infty, \quad \mu \to -\infty, \quad \frac{\mu}{T} = -\lambda = \text{const}. \tag{24}$$

Then the average intensity $\langle I_\alpha \rangle_{\mathcal{F}_{\rm eq}}$ or, equivalently, the average density $\langle |\psi_n|^2 \rangle_{\mathcal{F}_{\rm eq},\theta}$ are simply given by $\rho = 1/\lambda$ (in the first case, the average is performed over the distribution function, while in the second one the average over the phases is also implied).

Let us assume that all modes are in equilibrium, except one, say, $\alpha = 1$. This corresponds to the distribution function of the form

$$\mathcal{F}(\{I_\alpha\}) = f(I_1) \prod_{\alpha \neq 1} \frac{e^{-I_\alpha/\rho}}{\rho}, \tag{25}$$

where ρ is the equilibrium density, and we have taken the limit (24). Then $f(I_1)$ satisfies the following equation:

$$\frac{1}{\Gamma_1} \frac{\partial f}{\partial t} = \frac{\partial}{\partial I_1} I_1 \left(\rho \frac{\partial}{\partial I_1} + 1 \right) f, \tag{26}$$

where we denoted

$$\Gamma_\alpha = \frac{4\pi}{\rho} {\sum_{\beta,\gamma,\delta}}' V_{\alpha\beta\gamma\delta}^2 \langle \delta(\tilde{\omega}_\alpha + \tilde{\omega}_\beta - \tilde{\omega}_\gamma - \tilde{\omega}_\delta) I_\beta I_\gamma I_\delta \rangle_{\mathcal{F}}. \tag{27}$$

Equation (26) coincides with the Fokker-Planck equation for a damped oscillator subject to external noise (see Appendix B). The eigenfunctions of the right-hand side are $\mathrm{L}_n(I_1/\rho)\, e^{-I_1/\rho}$, where L_n, $n = 0,1,2,\ldots$ denote Laguerre polynomials. The stationary equilibrium solution $e^{-I_1/\rho}$ corresponds to $n = 0$, while the eigenfunctions with $n > 0$ are exponentially decaying in time with the rate given by the corresponding eigenvalues, $n\Gamma_1$. Thus, $1/\Gamma_\alpha$ has the meaning of the typical relaxation time for the intensity I_α.

When the nonlinear frequency shifts are self-averaging, as given by Eq. (12), they disappear from the δ function. Also, since the sum in Eq. (27) is contributed by different β, γ, and δ, for the distribution function (25) we can split $\langle I_\beta I_\gamma I_\delta \rangle_{\mathcal{F}} \to \langle I_\beta \rangle_{\mathcal{F}} \langle I_\gamma \rangle_{\mathcal{F}} \langle I_\delta \rangle_{\mathcal{F}}$ (see also Sec. V A and Appendix D). Then Eq. (27) reduces to

$$\Gamma_\alpha = 4\pi\rho^2 {\sum_{\beta,\gamma,\delta}}' V_{\alpha\beta\gamma\delta}^2\, \delta(\omega_\alpha + \omega_\beta - \omega_\gamma - \omega_\delta). \tag{28}$$

E. Numerical evaluation of the relaxation rate

The frequency δ function appearing in Eqs. (27) and (28) is implemented as a box of a finite width $2w$,

$$\delta(\varpi) \to \delta_w(\varpi) = \frac{\theta(w - |\varpi|)}{2w}, \tag{29}$$

where $\theta(x)$ is the unit step function. This implementation is more efficient for numerical calculations than a smooth function. Using this definition, one can evaluate Γ_α for each mode α using Eq. (28). Two examples of such calculation are shown in Fig. 1.

Rather than looking at mode-specific quantities, we will be interested in some average characteristics. As the properties of an eigenmode are determined by its frequency, it is convenient to work with frequency-resolved averages, so we define

$$\overline{\Gamma}_\omega = \lim_{L\to\infty} \frac{\sum_{\alpha=1}^{L} \delta_\eta(\omega_\alpha - \omega)\, \Gamma_\alpha^{(w)}}{\sum_{\alpha=1}^{L} \delta_\eta(\omega_\alpha - \omega)}, \tag{30}$$

where η is a sufficiently small interval, and averaging over a long chain is assumed to be equivalent to averaging over the disorder realizations. It is important that the limit $L \to \infty$ should be taken prior to $\eta \to 0$.

It should be emphasized that the result of the calculation of individual Γ_α depends on the width w. Namely the same calculation with a larger w produces an analogous picture, but the points in Fig. 1 are more squeezed towards the solid line, that is, Γ_α becomes more self-averaging. For smaller w, the points become more scattered. At the same time, the average $\overline{\Gamma}_\omega$ is not sensitive to the value of w as long as $w \ll \Omega$. However, at sufficiently small w the values of Γ_α are so scattered that the fluctuations become comparable to the average, and the latter is no longer representative. This dependence on w is crucial for the discussion in Sec. IV A below.

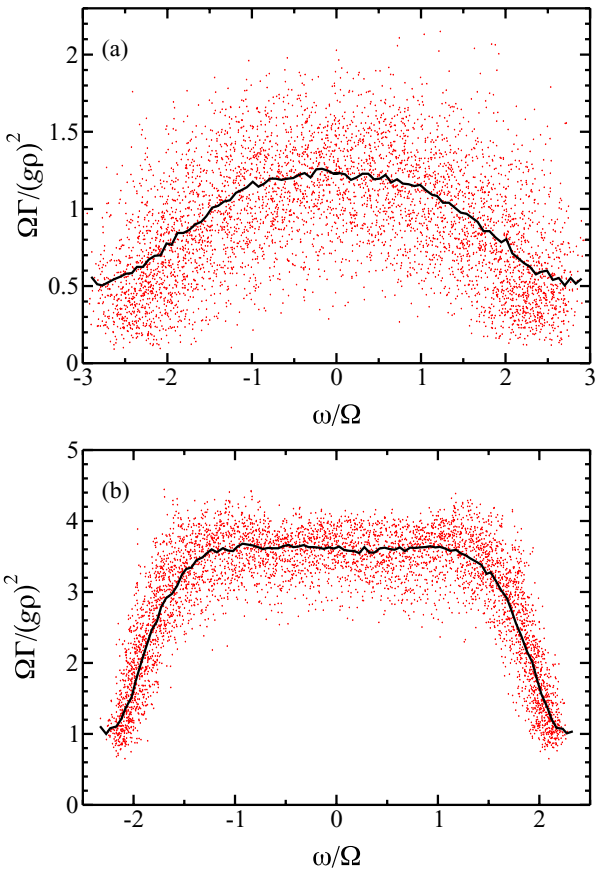

FIG. 1. (Color online) (a) The relaxation rate in a chain of length $L = 5000$ with $W/\Omega = 2.82$ obtained using Eq. (29) with $w = 0.1\,\Omega$ for individual eigenmodes from Eq. (28) (dots) and averaged over the modes at a given frequency, Eq. (30) with $\eta = 0.1\,\Omega$ (solid line). (b) The same for $W/\Omega = 1$ and $w = 0.01\,\Omega$.

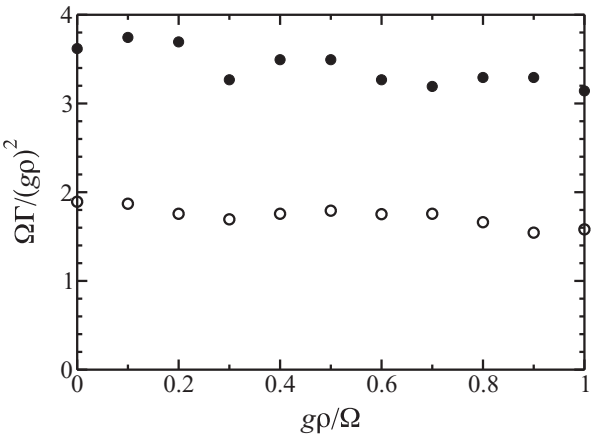

FIG. 2. The relaxation rate as a function of the dimensionless density $g\rho/\Omega$ for $W/\Omega = 1,2$ (filled and open circles, respectively), evaluated from Eq. (27) by numerically averaging over the equilibrium distribution and over the modes around $\omega = 0$ (the width $\eta = 0.1\,\Omega$), the δ function given by Eq. (29) with $w = 0.01\,\Omega$, $\eta = 0.1\,\Omega$.

We also check numerically the dependence $\Gamma_\alpha \propto \rho^2$ by evaluating Eq. (27) [again, implementing the δ function as in Eq. (29) and averaging over the modes at a given frequency as in Eq. (30)], where the intensities I_α are taken from the equilibrium distribution, $\mathcal{F} \propto \prod_\alpha e^{-I_\alpha/\rho}$. As seen from Fig. 2 for two values of the disorder strength, $W/\Omega = 1,2$, the average $\langle\Gamma_\alpha\rangle/\rho^2$ is practically independent of ρ up to the values of ρ as large as Ω/g.

IV. CONDITIONS FOR VALIDITY OF THE APPROACH

A. Lower bounds on the density

The arguments of Sec. III B are based on the perturbative expression (16), which is valid only as long as the change in the intensities is small compared to intensities themselves, $\sqrt{\langle\Delta I_\alpha^2\rangle} \ll I_\alpha$. This means that Eq. (16) is valid only at times shorter than the typical intensity relaxation time, $t \ll 1/\Gamma_\alpha$, introduced in Sec. III D, while the dynamics at times $t \sim 1/\Gamma_\alpha$ and longer is accounted for by the master equation. Thus, if at times $t \lesssim 1/\Gamma_\alpha$ the discreteness of the sum in Eq. (18) is still not relevant, the arguments of Sec. III B are self-consistent. In other words, if the effective width of the δ function is taken of the order of Γ_α itself, the sum in Eq. (18) should be contributed by many terms. Note that in this sense the sum in Eq. (18) and that in Eq. (28) are fully analogous. In terms of Δ_ϖ, the typical spacing between the values of $\varpi_{\alpha\beta\gamma\delta}$ effectively contributing to the sum for various β, γ, and δ, the condition of consistency is thus simply $\Delta_\varpi \ll \Gamma_\alpha$, which should hold for most of the modes α. In the time domain, this condition means that the characteristic time $\sim 1/\Delta_\varpi$ needed to decide whether the last term in Eq. (6) is a quasiperiodic function of time or random noise, is longer than the relaxation time of the mode intensities.

The value of Δ_ϖ is determined both by the frequencies $\varpi_{\alpha\beta\gamma\delta}$ *and* by the overlaps $V_{\alpha\beta\gamma\delta}$ [but *not* by the intensities I_α, as long as the nonlinear shifts are self-averaging, Eq. (12)]. In particular, only modes β, γ, and δ which are effectively within the same localization segment as the mode α can give a significant contribution to the sum, since at large distances the overlap $V_{\alpha\beta\gamma\delta}$ is exponentially suppressed. Thus, if there were no correlation between the frequencies and the overlaps $V_{\alpha\beta\gamma\delta}$, we could estimate $\Delta_\varpi \sim \Delta_3$, where $\Delta_3 \approx 0.14\,\Omega/\xi^3$ is the typical spacing between different combinations of three frequencies, $\omega_\delta + \omega_\gamma - \omega_\beta$ (see Appendix A1 for details), for modes which are within the same localizaton length. Note that $\Delta_3 \ll \Delta_1 = 2\pi\Omega/\xi$ for $\xi \gg 1$; indeed, the number of different combinations of three frequencies is much larger than the number of frequencies themselves.

However, the correlation between the frequencies and the overlaps turns out to be quite strong (see Appendix A3 for details), so it is not obvious how to actually define Δ_ϖ. This difficulty can be bypassed by noting that when the sum in Eq. (28) is contributed by many terms, Γ_α should be effectively self-averaging and its fluctuations should be small and vice versa. This is directly related to the dependence of the fluctuations of Γ_α (the vertical spread of the points in Fig. 1) on the effective δ function width w, discussed in Sec. III E above. Thus, the criterion $\Delta_\varpi \ll \Gamma_\alpha$, proposed above, can be replaced by an equivalent one: The fluctuations of Γ_α, evaluated for modes with frequencies $\omega_\alpha \approx \omega$ at $w = \overline{\Gamma}_\omega$,

should be small compared to the average $\overline{\Gamma}_\omega$ itself. Basically, it is the same criterion that was used in Ref. [44] for the validity of the kinetic equation on the metallic side of the quantum many-body localization transition.

Specifically, we focus on $\omega = 0$, assuming it to be representative of the whole band. Indeed, the results shown in Fig. 1 suggest that the dependence on ω is weak as long as ω is away from the band edges (see, however, the discussion in Sec. IV C). We then define the first and the second moments as

$$M_1(w) = \lim_{L\to\infty}\left[\sum_{\alpha=1}^{L}\delta_\eta(\omega_\alpha)\right]^{-1}\sum_{\alpha=1}^{L}\delta_\eta(\omega_\alpha) \times \frac{2\Omega}{g^2}\sum_{\beta,\gamma,\delta}{}' V^2_{\alpha\beta\gamma\delta}\,\delta_w(\varpi_{\alpha\beta\gamma\delta}), \tag{31a}$$

$$M_2(w) = \lim_{L\to\infty}\left[\sum_{\alpha=1}^{L}\delta_\eta(\omega_\alpha)\right]^{-1}\sum_{\alpha=1}^{L}\delta_\eta(\omega_\alpha) \times \left[\frac{2\Omega}{g^2}\sum_{\beta,\gamma,\delta}{}' V^2_{\alpha\beta\gamma\delta}\,\delta_w(\varpi_{\alpha\beta\gamma\delta})\right]^2, \tag{31b}$$

As mentioned in Sec. III E above, when the limit $L \to \infty$ is taken, $M_1(w)$ does not depend on w at $w \ll \Omega$, so we can denote $M_1 \equiv M_1(0)$, and $\overline{\Gamma}_{\omega=0} = 2\pi(g^2\rho^2/\Omega)M_1$. The fluctuations $M_2(w) - M_1^2$ still depend on w, and we denote by $w_{\min}$ the value of w at which the fluctuations are equal to the average,

$$M_2(w_{\min}) - M_1^2 = M_1^2. \tag{32}$$

The dimensionless quantities M_1 and $w_{\min}/\Omega$ depend only on the dimensionless disorder strength W/Ω. We can now define the minimal density $\rho_{\min}$ as the one for which $2\pi(g^2\rho_{\min}^2/\Omega)M_1 = w_{\min}$. Thus, the low-density limit of the validity of the master equation (20) is

$$\frac{g\rho}{\Omega} \gg \sqrt{\frac{w_{\min}}{2\pi\Omega M_1}} \equiv \frac{g\rho_{\min}}{\Omega}. \tag{33}$$

We have determined M_1 and $w_{\min}/\Omega$ numerically for several values of the dimensionless disorder strength W/Ω in the interval $0.5 \leqslant W/\Omega \leqslant 8$, corresponding to $1.5 \leqslant \xi \leqslant 400$. The value of $\eta = 0.1\,\Omega$ was checked to be sufficiently small, and the chain length $L = 16\,000$ was checked to be sufficiently long to not affect the results (the value of $w_{\min}$ could be reliably determined only for $W/\Omega \geqslant 1$, due to computational limitations). The results are shown in Fig. 3. On the weak-disorder side, they can be fitted by

$$M_1 = \frac{0.6 \pm 0.05}{(W/\Omega)^{0.8\pm0.1}}, \tag{34a}$$

$$\frac{w_{\min}}{\Omega} = (2.3 \pm 0.3)\times 10^{-6}\left(\frac{W}{\Omega}\right)^{6.2\pm0.2}. \tag{34b}$$

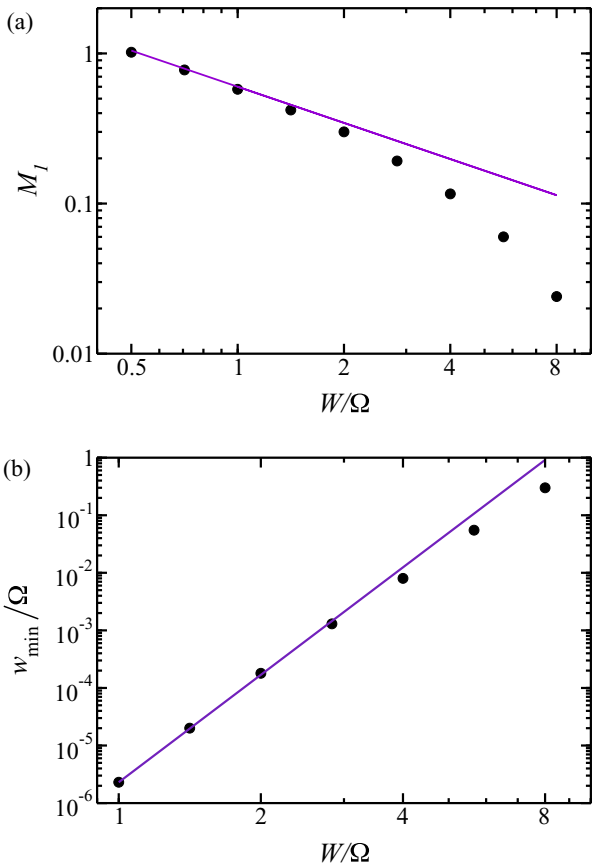

FIG. 3. (Color online) The dots show the values of (a) M_1 and (b) $w_{\min}/\Omega$, calculated numerically. The lines show the dependencies (a) $M_1 = 0.6\,(\Omega/W)^{0.8}$ and (b) $w_{\min}/\Omega = 2.3\times10^{-6}\,(W/\Omega)^{6.2}$, which represent the best fits to the calculated values at low W.

The resulting dependence of $g\rho_{\min}/\Omega$ on W/Ω is shown in Fig. 5 and can be fitted by

$$\frac{g\rho_{\min}}{\Omega} = (0.8 \pm 0.1)\times 10^{-3}\left(\frac{W}{\Omega}\right)^{3.5\pm0.2}. \tag{35}$$

B. Relation to chaos

Along with $\rho_{\min}$, defined above by Eq. (33), we also plot the value $\rho_{1/2}$, defined as the density at which half of the modes of the chain are chaotic and half are not. Even though only the motion of the whole coupled chain can be, strictly speaking, characterized as chaotic or not (in the sense of positive Lyapunov exponent), for given initial conditions one can focus on the dynamics of a single mode assuming the intensities and phases of all other modes to be frozen, so they represent an external (quasiperiodic) force acting on the chosen mode. For the problem of a single nonlinear oscillator under the action of an external force, one can clearly define chaotic and regular motion. Then, for the whole chain with given initial conditions, one can calculate the fraction of modes whose motion is chaotic in the above sense (see Appendix C for the details of numerical implementation). This fraction depends on the density: it vanishes when $\rho \to 0$ (since for a linear system no modes are chaotic) as $\propto\rho^2$ [12,13,15] and approaches

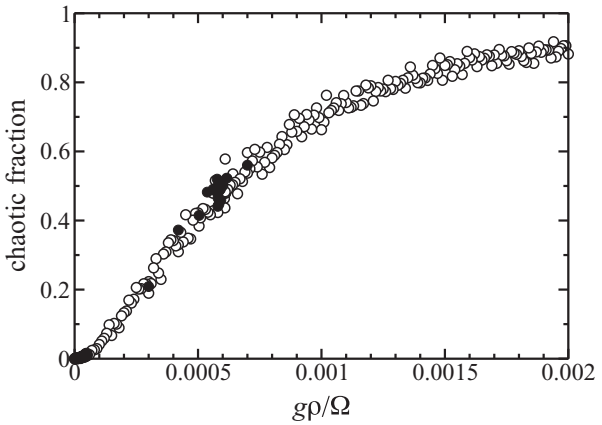

FIG. 4. The fraction of chaotic modes for $W/\Omega = 2$ as a function of the dimensionless density $g\rho/\Omega$. The open and filled circles correspond to $L = 2000$ and $L = 4000$, respectively.

unity for sufficiently large density, as shown in Fig. 4 for $W/\Omega = 2$.

Let $\rho_{1/2}$ be the value of ρ when the fraction is $1/2$. In Fig. 5, we plot the dimensionless quantity $g\rho_{1/2}/\Omega$ versus disorder strength. The extracted dependence,

$$\frac{g\rho_{1/2}}{\Omega} = (4 \pm 0.5) \times 10^{-5} \left(\frac{W}{\Omega}\right)^{3.5\pm0.2}, \qquad (36)$$

has the same power law as that for $\rho_{\min}$ and differs only in the prefactor. This strongly suggests that the two seemingly unrelated criteria (namely $\rho \gg \rho_{\min}$, responsible for the self-consistency of the Fermi golden rule where chaos simply did not enter the discussion at all, and $\rho \gg \rho_{1/2}$, ensuring that most modes are chaotic) are, in fact, closely connected. Thus,

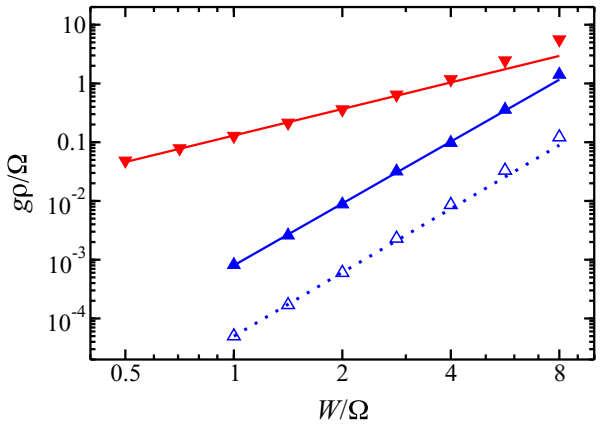

FIG. 5. (Color online) The limits of validity of the master equation approach. The filled upward triangles and the solid line passing through them, which corresponds to $g\rho_{\min}/\Omega = 0.8 \times 10^{-3}\,(W/\Omega)^{3.5}$, represent the low-density limitation expressed by Eq. (33). The open upward triangles and the dotted line passing through them, which correspond to $g\rho_{1/2}/\Omega = 4 \times 10^{5}\,(W/\Omega)^{3.5}$, represent the density at which half of the modes of the chain are chaotic. The filled downward triangles and the solid line passing through them, which correspond to $g\rho_{\max}/\Omega = 0.13\,(W/\Omega)^{1.5}$, represent the high-density limitation expressed by Eq. (37).

the situation analyzed in the present work corresponds to the regime when most of the modes are chaotic. This condition was used in Ref. [22] to define the regime of strong chaos, where the ρ^2 dependence of the diffusion coefficient was observed.

C. Upper bounds on the density

The condition for the validity of the master equation (20) from the high-density side is that the localized normal modes of the linear problem should be well defined on the relaxation time scale $1/\Gamma$, i.e., this time should be longer than the inverse spacing between mode frequencies on the same localization segment, $1/\Delta_1 = \xi/(2\pi\Omega)$. In the opposite case, the discrete localized normal modes are not well resolved and do not represent a good starting basis. Using $\Gamma = 2\pi(g^2\rho^2/\Omega)M_1$, we can write the condition $\Gamma \ll \Delta_1$ as

$$g\rho \ll \sqrt{\frac{\Omega\Delta_1}{2\pi M_1}} \equiv g\rho_{\max}. \qquad (37)$$

The dependence of $g\rho_{\max}/\Omega$ on the disorder strength, obtained from the numerical results for M_1 and Δ_1, is shown in Fig. 5. It can be fitted by the expression

$$\frac{g\rho_{\max}}{\Omega} = (0.13 \pm 0.02)\left(\frac{W}{\Omega}\right)^{1.5\pm0.1}. \qquad (38)$$

It is instructive to approach the same condition from the clean side, treating the disorder as a perturbation. Without disorder, the normal mode wave functions and frequencies are given by

$$\phi_{\alpha n} = \sqrt{\frac{2}{L+1}}\,\sin\frac{\pi\alpha n}{L+1}, \quad \omega_\alpha = -2\Omega\cos\frac{\pi\alpha}{L+1}, \qquad (39)$$

and for $L \gg 1$ it is convenient to introduce the wave vector $k = \pi\alpha/(L+1)$ and velocity $v_k = 2\Omega\sin k$. It is well known that perturbative treatment of the disorder gives the elastic backscattering rate $\Gamma_k^{(\mathrm{bs})} = 2v_k/\xi(\omega_k)$. In other words, the backscattering mean free path $v_k/\Gamma_k^{(\mathrm{bs})}$ is twice shorter than the localization length $\xi(\omega_k)$ at the same frequency $\omega_k = -2\Omega\cos k$ [45]. At the same time, the wave functions (39) of the clean chain can be used to evaluate the mode relaxation rate $\Gamma_k^{(\mathrm{nl})}$ due to the nonlinearity from Eq. (28), whose derivation did not assume any specific form of the wave functions. If $\Gamma_k^{(\mathrm{nl})} \gg \Gamma_k^{(\mathrm{bs})}$, the elastic scattering on the disorder and Anderson localization are not important. Noting that the mean mode spacing on the localization length $\Delta_1 = (\pi/2)\Gamma_k^{(\mathrm{bs})}$, we conclude that the conditions $\Gamma \ll \Delta_1$ mentioned above and $\Gamma_k^{(\mathrm{nl})} \ll \Gamma_k^{(\mathrm{bs})}$ are equivalent, provided that the relaxation rates (28), evaluated on clean and localized wave functions, match.

Equation (28) with the wave functions (39) gives

$$\Gamma_k^{(\mathrm{nl})} = \frac{(g\rho)^2}{2\pi\Omega}\int_{-\pi}^{\pi}\frac{dk'}{|\sin k' - \sin k|} \approx \frac{2(g\rho)^2}{\pi\Omega|\cos k|}\ln\frac{|\cos k|}{\Delta k}, \qquad (40)$$

where Δk is a small uncertainty in the wave vector, needed to cut off the divergence at $k' \to k$. For consistency, it should be taken of the order of the inverse mean free path, $\Delta k \sim \Gamma_k^{(\mathrm{nl})}/v_k$.

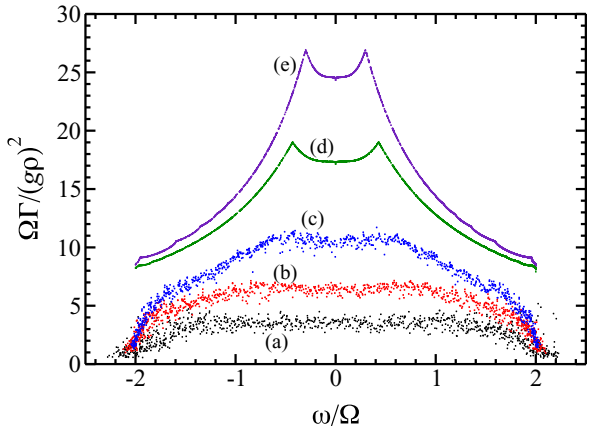

FIG. 6. (Color online) The relaxation rates of individual normal modes in a chain of $L = 5000$ sites. Different sets of points correspond to different disorder strengths: $W/\Omega = 1$ (a), $W/\Omega = 0.5$ (b), $W/\Omega = 0.25$ (c), and $W = 0$ [(d) and (e)]. The spikes on the sets (d) and (e) result from the $k = \pi/2$ singularity being cut off by the finite width of the δ function. This width is $w = 10^{-3}$ for sets (a), (b), (c), and (e) and $w = 3.16 \times 10^{-3}$ for set (d).

Then, for most of the band, one can write with logarithmic precision

$$\Gamma_k^{(\mathrm{nl})} \approx \frac{4(g\rho)^2}{\pi\Omega|\cos k|} \ln\frac{\Omega}{g\rho}. \tag{41}$$

However, at $k \to \pi/2$, expression (40) is divergent. This divergence should be smeared on the scale $|k - \pi/2| \sim \Delta k$, so one can estimate (up to a numerical coefficient)

$$\Gamma_k^{(\mathrm{nl})} \sim g\rho, \quad |k - \pi/2| \lesssim \frac{g\rho}{\Omega}. \tag{42}$$

These results are confirmed by the direct numerical evaluation of Eq. (28) for different values of disorder, including the disorder-free case, as shown in Fig. 6. For the disorder-free case the cutoff scale Δk in the numerical calculation is effectively provided by the δ function width, $\Delta k \sim w/v_k$, as seen by comparing curves (d) and (e) which differ by the value of w used in the calculation.

Thus, upon increasing the disorder at fixed $\rho \ll \Omega/g$, when $1/\xi \gtrsim (g\rho/\Omega)^2$ (up to numerical and logarithmic factors), the localization condition $\Gamma_k^{(\mathrm{nl})} \sim \Gamma_k^{(\mathrm{bs})}$ becomes fulfilled for most of the normal modes, except for those in a relatively narrow frequency range $|\omega| \lesssim (g\rho)^2/(\Omega\xi)$; this range shrinks completely at a stronger disorder such that $1/\xi \sim g\rho/\Omega$. These two conditions correspond to $g\rho_{\max}/\Omega \sim \xi^{-1/2}$ and $g\rho_{\max}/\Omega \sim \xi^{-1}$, respectively, which should be contrasted to $g\rho_{\max}/\Omega \sim \xi^{-0.75}$ following from Eq. (38). This strongly suggests that Eqs. (34a), (34b), (36), and (38) [as well as Eq. (88) below] do not represent the scaling at lowest disorder strengths but rather intermediate asymptotics. The results shown in Fig. 6 suggest that the behaviour of relaxation rates in the disordered system starts to resemble that of the clean one at $W/\Omega > 0.25$, corresponding to extremely large localization lengths $\xi > 1000$. The detailed investigation of this issue requires significant computational effort and is beyond the scope of the present work. We only note that similarly large localization lengths were found to be necessary to reach the weak-deisorder asymptotics in the statistics of a single normal mode wave function [39].

V. MACROSCOPIC DIFFUSION EQUATION

The main task of the present section is to pass from the master equation (20), which is still microscopic as it includes the dynamics of each individual normal mode, to the macroscopic description in terms of the density. In this section, we neglect the nonlinear frequency shifts in Eq. (20). As discussed in Secs. III A and III E, they produce just small corrections, $\sim g\rho/\Omega$, to the main result.

A. Classical Boltzmann equation

Multiplying Eq. (20) by I_α and integrating over all intensities, we obtain an equation for the average as follows:

$$\frac{\partial\langle I_\alpha\rangle_{\mathcal{F}}}{\partial t} = {\sum_{\beta,\gamma,\delta}}' R_{\alpha\beta\gamma\delta}\langle I_\beta I_\gamma I_\delta + I_\alpha I_\gamma I_\delta - I_\alpha I_\beta I_\delta - I_\alpha I_\beta I_\gamma\rangle_{\mathcal{F}}, \tag{43a}$$

$$R_{\alpha\beta\gamma\delta} = 4\pi V^2_{\alpha\beta\gamma\delta}\,\delta(\omega_\alpha + \omega_\beta - \omega_\gamma - \omega_\delta). \tag{43b}$$

Next we note that the intensity of each mode is changed due to random resonant interactions with many other modes, so any two modes effectively see mostly different resonances. Thus, one can neglect the correlations between different intensities and decouple $\langle I_\beta I_\gamma I_\delta\rangle_{\mathcal{F}} \to \langle I_\beta\rangle_{\mathcal{F}}\langle I_\gamma\rangle_{\mathcal{F}}\langle I_\delta\rangle_{\mathcal{F}}$, and so on. This can be done when the number of terms contributing to the sum over modes is large (see Appendix D). The latter condition was discussed in detail in Sec. IV A.

As a result, we obtain a closed kinetic equation for the averages $\langle I_\beta\rangle_{\mathcal{F}} \equiv \bar{I}_\alpha$ (the new notation is introduced for compactness) as follows:

$$\frac{d\bar{I}_\alpha}{dt} = {\sum_{\beta,\gamma,\delta}}' R_{\alpha\beta\gamma\delta}[(\bar{I}_\alpha + \bar{I}_\beta)\bar{I}_\gamma\bar{I}_\delta - \bar{I}_\alpha\bar{I}_\beta(\bar{I}_\gamma + \bar{I}_\delta)]. \tag{44}$$

This kinetic equation conserves the total norm $\sum_\alpha \bar{I}_\alpha$ and the total energy $\sum_\alpha \omega_\alpha \bar{I}_\alpha$ (due to the frequency δ function in $R_{\alpha\beta\gamma\delta}$). The equilibrium state, which nullifies the collision integral in the right-hand side, is given by the Rayleigh-Jeans distribution,

$$\bar{I}_\alpha^{\mathrm{eq}} = \frac{T}{\omega_\alpha - \mu}, \tag{45}$$

or simply $\bar{I}_\alpha^{\mathrm{eq}} = 1/\lambda$ in the limit (24). Equation (45) correctly reproduces the thermodynamics of the chain with the required precision, as discussed in Appendix E. If one sets all $\bar{I}_\alpha = 1/\lambda$ except one, $\alpha = 1$, then Eq. (44) will describe exponential relaxation of $\bar{I}_1$ to its equilibrium value, $\bar{I}_1(t) = 1/\lambda + [\bar{I}_1(0) - 1/\lambda]e^{-\Gamma_1 t}$, as $\Gamma_\alpha = (1/\lambda^2)\sum'_{\beta,\gamma,\delta} R_{\alpha\beta\gamma\delta}$.

As Eq. (44) is nonlinear in the intensities, it is not possible to interpret $p_\alpha = \bar{I}_\alpha(\sum_\beta \bar{I}_\beta)^{-1}$ as a probability, even though $\sum_\alpha p_\alpha = 1$ and $p_\alpha \geqslant 0$ at all times. This invalidates the continuous-time random walk approach to this problem, introduced in Ref. [46] without justification. Reference [46] also ignores the fact that the intensity can leak out of a mode α through many parallel channels, and the corresponding rates add up, as expressed by the sum in Eq. (27).

It is worth noting that Eq. (45) represents the classical limit $\hbar \to 0$ of the Bose-Einstein distribution $N_\alpha = 1/[e^{-\hbar(\omega_\alpha-\mu)/T}-1]$, and Eq. (44) has the same form as the quantum Boltzmann equation for bosonic occupation numbers $N_\alpha = \bar{I}_\alpha/\hbar$ changing due to pair collisions in the limit $N_\alpha \gg 1$. Indeed, in this limit, the bosonic combination $(N_\alpha+1)(N_\beta+1)N_\gamma N_\delta - N_\alpha N_\beta(N_\gamma+1)(N_\delta+1)$ reduces to the one in Eq. (44), while $R_{\alpha\beta\gamma\delta}$ is the rate of a pair collision, as obtained from the Fermi golden rule for quantum Hamiltonian (3) with ψ_n, ψ_n^* treated as bosonic field operators. Such a quantum Boltzmann equation can be applied to describe the dynamics of disordered bosons on the metallic side of the many-body localization transition [47] in the regime, analogous to the interaction-induced nearest-neighbor hopping considered earlier for fermions, with the power-law temperature dependence of the dc conductivity [44,48] (not to be confused with the phonon-induced nearest-neighbor hopping [49,50]). Also, Eq. (44) has a form similar to the kinetic equation used in the theory of weak wave turbulence [31], but with the momenum conservation condition relaxed due to the presence of disorder.

B. Macroscopic density and current

Generally, the diffusion equation (1) is obtained from the continuity equation,

$$\frac{\partial \rho}{\partial t} = -\frac{\partial J}{\partial x}, \tag{46a}$$

supplemented by the Fick's law,

$$J = -D(\rho)\,\frac{\partial \rho}{\partial x}. \tag{46b}$$

Equation (46a) expresses the conservation of the total norm $\mathcal{N} = \int \rho\, dx$, and Eq. (46b) represents the first term of the gradient expansion of the current in local equilibrium, characterized by a spatially dependent density (in the global equilibrium, where the density is constant along the chain, the current must vanish).

The continuous functions $\rho(x,t)$ and $J(x,t)$ entering Eqs. (46a) and (46b) are the macroscopic density and current, whose dependence on the coordinate x and time t is smooth enough compared to some microscopic scales. The microscopic expression for the norm density following directly from Eq. (2), $|\psi_n(t)|^2$, strongly varies from site to site (as the random eigenmode wave functions $\phi_{\alpha n}$ do) and randomly oscillates in time (on the scale $1/\Omega$). If one is not interested in the details of these short-scale oscillations but rather in the slow dynamics of the smooth envelope density, the microscopic density $|\psi_n(t)|^2$ should be coarse grained in space and time on some sufficiently large scales. As we are studying the dynamics of energy and norm exchange between different normal modes, the corresponding length scale is the mode localization length ξ. We are interested in the case when local equilibrium is reached, so the relevant time scale is the inverse of the relaxation rate Γ, introduced in Sec. III D (assumed to be longer that the frequency spacing between the normal modes, $1/\Delta_1$, as discussed in Sec. IV C). Formally, we define

$$\begin{aligned}\rho(x,t) &= \int dt'\,\mathcal{T}(t-t')\sum_n \mathcal{S}(x-n)\langle|\psi_n(t')|^2\rangle_{\mathcal{F}} \\ &= \int dt'\,\mathcal{T}(t-t')\sum_\alpha \mathcal{S}(x-X_\alpha)\,\bar{I}_\alpha(t') + O(\xi^2/\ell^2),\end{aligned} \tag{47}$$

where the "center of mass" of the mode α is defined as

$$X_\alpha = \sum_n n\phi_{\alpha n}^2, \tag{48}$$

and the spatial and temporal smoothing functions $\mathcal{S}(x)$ and $\mathcal{T}(t)$ can be taken, e.g., as Gaussian,

$$\mathcal{S}(x) = \frac{e^{-x^2/(2\ell^2)}}{\sqrt{2\pi\ell^2}}, \quad \mathcal{T}(t) = \frac{e^{-t^2/(2\tau^2)}}{\sqrt{2\pi\tau^2}}. \tag{49}$$

Here the smoothing length and time scales $\ell \gg \xi$ and $\tau \gg 1/\Gamma$, as discussed above; the diffusion equation is then valid at length scales $x \gtrsim \ell$ and $t \gtrsim \tau$. The relation between the first and the second expression in Eq. (47) is discussed in detail in Appendix F.

The macroscopic current is defined in order to identically satisfy the continuity equation,

$$J(x,t) = \int dt'\,\mathcal{T}(t-t')\sum_\alpha \tilde{\mathcal{S}}(x-X_\alpha)\,\frac{d\bar{I}_\alpha(t')}{dt'} + O(\xi^2/\ell^2), \tag{50}$$

where

$$\tilde{\mathcal{S}}(x) \equiv -\int_0^x \mathcal{S}(x')\,dx'. \tag{51}$$

Substituting $d\bar{I}_\alpha/dt$ from Eq. (44) and symmetrizing with respect to $\alpha \leftrightarrow \beta$, $\alpha\beta \leftrightarrow \gamma\delta$, one obtains

$$\begin{aligned}J(x,t) = \int dt'\,\mathcal{T}(t-t')\sum_{\alpha,\beta,\gamma,\delta}{}' \frac{R_{\alpha\beta\gamma\delta}}{4} \\ \times[(\bar{I}_\alpha+\bar{I}_\beta)\bar{I}_\gamma\bar{I}_\delta - \bar{I}_\alpha\bar{I}_\beta(\bar{I}_\gamma+\bar{I}_\delta)] \\ \times[\tilde{\mathcal{S}}(x-X_\alpha)+\tilde{\mathcal{S}}(x-X_\beta) \\ -\tilde{\mathcal{S}}(x-X_\gamma)-\tilde{\mathcal{S}}(x-X_\delta)].\end{aligned} \tag{52}$$

For each term in the sum, it is convenient to introduce the short-hand notations for the "center-of-mass" coordinate and the "displacement,"

$$X_{\alpha\beta\gamma\delta} \equiv \frac{X_\alpha + X_\beta + X_\gamma + X_\delta}{4}, \tag{53a}$$

$$d_{\alpha\beta\gamma\delta} = X_\alpha + X_\beta - X_\gamma - X_\delta. \tag{53b}$$

As a final step, we expand each $\tilde{\mathcal{S}}(x-X_\alpha)$ to the first order around $x - X_{\alpha\beta\gamma\delta}$, which gives

$$\begin{aligned}J(x,t) = \int dt'\,\mathcal{T}(t-t')\sum_{\alpha,\beta,\gamma,\delta}{}' \mathcal{S}(x-X_{\alpha\beta\gamma\delta})\,\frac{R_{\alpha\beta\gamma\delta}}{4} \\ \times d_{\alpha\beta\gamma\delta}[(\bar{I}_\alpha+\bar{I}_\beta)\bar{I}_\gamma\bar{I}_\delta - \bar{I}_\alpha\bar{I}_\beta(\bar{I}_\gamma+\bar{I}_\delta)].\end{aligned} \tag{54}$$

In this expression, the time argument t' is implied for all intensities; it has been omitted for the sake of compactness.

To illustrate how this formalism works for a very simple toy model, in Appendix G it is used to calculate the conductivity of a disordered electric *RC* cirquit.

C. Diffusion coefficient

The diffusion coefficient $D(\rho)$ should be found by calculating the linear response of the current J to an infinitesimal gradient of the density, $\partial\rho/\partial x = -\kappa$. For this, one should look for a stationary solution of Eq. (44) in the form

$$\bar{I}_\alpha = \rho - \kappa X_\alpha + r_\alpha \tag{55}$$

to the linear order order in κ, where r_α is such that

$$\sum_\alpha \mathcal{S}(x - X_\alpha)\, r_\alpha = O(\kappa\xi^2/\ell^2) \tag{56}$$

and does not grow with x. Indeed,

$$\sum_\alpha \frac{d\mathcal{S}(x - X_\alpha)}{dx} = O(\xi^2/\ell^2), \tag{57a}$$

$$\sum_\alpha \frac{d\mathcal{S}(x - X_\alpha)}{dx}\, X_\alpha = 1 + O(\xi^2/\ell^2), \tag{57b}$$

as discussed in Appendix F.

With the substitution (55), the linearized stationary Eq. (44) becomes a system of linear equations for r_α, which can be written as

$$\sum_{\beta,\gamma,\delta}{}' R_{\alpha\beta\gamma\delta}(r_\alpha + r_\beta - r_\gamma - r_\delta) = \kappa \sum_{\beta,\gamma,\delta}{}' R_{\alpha\beta\gamma\delta} d_{\alpha\beta\gamma\delta}. \tag{58}$$

The typical value of $d_{\alpha\beta\gamma\delta}$ is $d_{\alpha\beta\gamma\delta} \sim \xi$, since otherwise the overlap $V_{\alpha\beta\gamma\delta}$ is exponentially suppressed. Moreover, because of the large number of terms contributing to the sum on the right-hand side, the effective self-averaging occurs, so the typical value of sum should be close to its statistical average over the disorder realizations. But the latter is zero because, on the average, the chain is symmetric with respect to translations and spatial inversion; in other words, there are as many terms with $d_{\alpha\beta\gamma\delta} > 0$ as with $d_{\alpha\beta\gamma\delta} < 0$. As a result, the typical value of $r_\alpha + r_\beta - r_\gamma - r_\delta$ is much smaller than $\kappa d_{\alpha\beta\gamma\delta}$.

For the solution given by Eq. (55), the current becomes

$$J(x) = \rho^2 \sum_{\alpha,\beta,\gamma,\delta}{}' \mathcal{S}(x - X_{\alpha\beta\gamma\delta})\, \frac{R_{\alpha\beta\gamma\delta}}{4} \times d_{\alpha\beta\gamma\delta}[\kappa d_{\alpha\beta\gamma\delta} - (r_\alpha + r_\beta - r_\gamma - r_\delta)]. \tag{59}$$

As discussed above, $r_\alpha + r_\beta - r_\gamma - r_\delta$ can be neglected with respect to $\kappa d_{\alpha\beta\gamma\delta}$. Finally, as $\mathcal{S}(x)$ is a slowly varying function with unit integral, the convolution is equivalent to spatial averaging, so the current is $J = \kappa D_0 \rho^2$, with D_0 given by

$$D_0 = \frac{\pi}{L} \sum_{\alpha,\beta,\gamma,\delta}{}' d^2_{\alpha\beta\gamma\delta} V^2_{\alpha\beta\gamma\delta}\, \delta(\omega_\alpha + \omega_\beta - \omega_\gamma - \omega_\delta). \tag{60}$$

In combination with the continuity equation, Eq. (46a), this immediately gives Eq. (1).

VI. ENERGY TRANSPORT

A. General remarks

In Sec. V, the macroscopic nonlinear diffusion equation for the norm density was derived. However, the original nonlinear Schrödinger equation (2), as well as the master equation (20) and the Boltzmann equation (44), have two conserved quantities: norm and energy. The transport of energy was ignored in Sec. V for simplicity. The purpose of the present section is to derive and study the full system of the macroscopic transport equations for the two conserved quantities.

Very generally, the macroscopic norm and energy currents, J and Q, vanish in the state of global thermal equilibrium, characterized by the values of the temperature T and the chemical potential μ, which are constant along the chain. Under the assumption of *local* equilibrium, when μ and T are slowly changing with the position, the currents can be evaluated by performing the expansion in the spatial gradients of μ and T. Restricting the expansion to the first term, one obtains

$$\begin{pmatrix} J \\ Q \end{pmatrix} = \hat{\mathcal{L}}(\mu,T)\, \frac{\partial}{\partial x} \begin{pmatrix} -\mu/T \\ 1/T \end{pmatrix}, \tag{61}$$

where $\hat{\mathcal{L}}$ is a 2×2 matrix which depends on the local values of μ and T. It is symmetric by virtue of the Onsager relations.

Upon substitution of these currents to the corresponding continuity equations, the equations for the macroscopic norm and energy densities, $\rho(x,t)$ and $\varepsilon(x,t)$, are obtained as follows:

$$\frac{\partial}{\partial t} \begin{pmatrix} \rho \\ \varepsilon \end{pmatrix} = \frac{\partial}{\partial x}\, \hat{\mathcal{L}}(\mu,T)\, \frac{\partial}{\partial x} \begin{pmatrix} -\mu/T \\ 1/T \end{pmatrix}. \tag{62}$$

To close the equations, the local equilibrium relations between ρ,ε and μ,T should be supplied. For weak disorder they are approximately the same as for the clean case; the latter is analyzed in Ref. [37]. The explicit expressions in the high-temperature limit are given in Appendix E.

B. Neglecting the nonlinear frequency shifts

Let us first neglect the nonlinear frequency shifts and study Eq. (44), which conserves the total energy,

$$\mathcal{E} = \sum_\alpha \omega_\alpha \bar{I}_\alpha = \sum_n \langle h_n \rangle_{\mathcal{F}}, \tag{63}$$

where the on-site energy is defined as

$$h_n = \epsilon_n |\psi_n|^2 - \frac{\Omega}{2}(\psi_n^* \psi_{n-1} + \psi_{n-1}^* \psi_n) - \frac{\Omega}{2}(\psi_n^* \psi_{n+1} + \psi_{n+1}^* \psi_n). \tag{64}$$

The expression for the macroscopic energy density is obtained analogously to Eq. (47):

$$\varepsilon(x,t) = \int dt'\, \mathcal{T}(t - t') \sum_n \mathcal{S}(x - n) \langle h_n(t') \rangle_{\mathcal{F}} = \int dt' \mathcal{T}(t - t') \sum_\alpha \mathcal{S}(x - X_\alpha)\, \omega_\alpha \bar{I}_\alpha(t') + O(\xi^2/\ell^2), \tag{65}$$

where we have used the fact that

$$\sum_n n\phi_{\alpha n}[\epsilon_n\phi_{\alpha n} - \Omega(\phi_{\alpha,n-1} + \phi_{\alpha,n+1})] = \omega_\alpha X_\alpha. \quad (66)$$

The energy current is defined analogously to Eq. (52) as follows:

$$Q(x,t) = \int dt'\, \mathcal{T}(t-t') \sum_{\alpha,\beta,\gamma,\delta}{}' \mathcal{S}(x - X_{\alpha\beta\gamma\delta}) \frac{R_{\alpha\beta\gamma\delta}}{4} \times v_{\alpha\beta\gamma\delta}[(\bar{I}_\alpha + \bar{I}_\beta)\bar{I}_\gamma\bar{I}_\delta - \bar{I}_\alpha\bar{I}_\beta(\bar{I}_\gamma + \bar{I}_\delta)], \quad (67)$$

where the time argument t' of the intensities has been omitted for compactness, and we have denoted

$$v_{\alpha\beta\gamma\delta} = \omega_\alpha X_\alpha + \omega_\beta X_\beta - \omega_\gamma X_\gamma - \omega_\delta X_\delta. \quad (68)$$

The matrix elements of $\hat{\mathcal{L}}$ are obtained by calculating the response of the two currents to the gradients $\chi_1 = \partial(-\mu/T)/\partial x$ and $\chi_2 = \partial(1/T)/\partial x$. If one seeks the solution of the stationary Boltzmann equation in the form

$$\bar{I}_\alpha = \frac{T}{\omega_\alpha - \mu} - \left(\frac{T}{\omega_\alpha - \mu}\right)^2 (\chi_1 X_\alpha + \chi_2\omega_\alpha X_\alpha + r_\alpha), \quad (69)$$

where r_α again satisfies condition (56) and does not grow with X_α, the linearized equation becomes

$$\sum_{\beta,\gamma,\delta}{}' \frac{T^4 R_{\alpha\beta\gamma\delta}(r_\alpha + r_\beta - r_\gamma - r_\delta)}{(\omega_\alpha - \mu)(\omega_\beta - \mu)(\omega_\gamma - \mu)(\omega_\delta - \mu)} = -\sum_{\beta,\gamma,\delta}{}' \frac{T^4 R_{\alpha\beta\gamma\delta}(\chi_1 d_{\alpha\beta\gamma\delta} + \chi_2 v_{\alpha\beta\gamma\delta})}{(\omega_\alpha - \mu)(\omega_\beta - \mu)(\omega_\gamma - \mu)(\omega_\delta - \mu)}, \quad (70)$$

where $d_{\alpha\beta\gamma\delta}$ is defined in Eq. (53b).

In full analogy with Sec. V C, we neglect the contribution of r_α to the currents. Also, we take the limit $T \to \infty$, $\mu/T = -1/\rho$. This limit for $\hat{\mathcal{L}}$ is regular and is given by

$$\hat{\mathcal{L}} = \frac{\rho^4}{L} \sum_{\alpha,\beta,\gamma,\delta}{}' \frac{R_{\alpha\beta\gamma\delta}}{4} \begin{pmatrix} d^2_{\alpha\beta\gamma\delta} & d_{\alpha\beta\gamma\delta}v_{\alpha\beta\gamma\delta} \\ v_{\alpha\beta\gamma\delta}d_{\alpha\beta\gamma\delta} & v^2_{\alpha\beta\gamma\delta} \end{pmatrix}. \quad (71)$$

Corrections to the $T \to \infty$ limit can be obtained by expanding $T/(\omega_\alpha - \mu)$ in ω_α/μ, and the first correction has a relative smallness $\sim\Omega\rho/T$.

The formal construction, presented above, has a caveat. Suppose we choose a different origin for counting the chain sites, that is, $n \to n - n_0$. Then $X_\alpha \to X_\alpha - n_0$ and $v_{\alpha\beta\gamma\delta} \to v_{\alpha\beta\gamma\delta} - n_0(\omega_\alpha + \omega_\beta - \omega_\gamma - \omega_\delta)$. The unphysical dependence on the choice of the origin disappears only when $\varpi_{\alpha\beta\gamma\delta} = \omega_\alpha + \omega_\beta - \omega_\gamma + \omega_\delta$ vanishes exactly. If the frequency δ function entering $R_{\alpha\beta\gamma\delta}$ has a small but finite width, $v_{\alpha\beta\gamma\delta}$ has a component which depends on the origin and makes the limit $L \to \infty$ in Eq. (71) ill defined.

Formally, the problem arises because the original dynamical system, Eq. (2), does not conserve the unperturbed energy $\sum_\alpha \omega_\alpha I_\alpha$ but the total one, Eq. (10). However, on physical grounds, we expect the energy contained in the perturbation terms to be less important than that in the unperturbed part, as long as the nonlinearity is small, $g\rho \ll \Omega$, and then the macroscopic description which neglects the difference between the total and the unperturbed energy should still be meaningful. To obtain such a description, one has to redefine $v_{\alpha\beta\gamma\delta}$, introducing a term which would vanish when $\varpi_{\alpha\beta\gamma\delta} = 0$ but would eliminate the unphysical dependence on the choice of the origin at small but finite $\varpi_{\alpha\beta\gamma\delta}$. We choose

$$v_{\alpha\beta\gamma\delta} = \omega_\alpha X_\alpha + \omega_\beta X_\beta - \omega_\gamma X_\gamma - \omega_\delta X_\delta - \varpi_{\alpha\beta\gamma\delta} X_{\alpha\beta\gamma\delta}. \quad (72)$$

As $X_{\alpha\beta\gamma\delta} \to X_{\alpha\beta\gamma\delta} - n_0$ upon $n \to n - n_0$, Eq. (72) remains invariant. To estimate the error introduced by the last term, we note that since $\hat{\mathcal{L}}$ does not depend on the choice of the origin, one can focus on the region near $n \sim \xi$. Then the modes with $|X_\alpha| \sim \xi$ are important, so the magnitude of the dropped term is $\sim w\xi$ (determined by the δ-function width), while the magnitude of the remaining terms is $\sim\Omega\xi$.

Thus, for the diagonal matrix elements of $\hat{\mathcal{L}}$ we have

$$\mathcal{L}_{11} = D_0\rho^4, \quad \mathcal{L}_{22} = K_0\rho^4, \quad (73)$$

with D_0 given by Eq. (60) and K_0 by the same expression but with the substitution $d_{\alpha\beta\gamma\delta} \to v_{\alpha\beta\gamma\delta}$,

$$K_0 = \frac{\pi}{L} \sum_{\alpha,\beta,\gamma,\delta}{}' v^2_{\alpha\beta\gamma\delta} V^2_{\alpha\beta\gamma\delta}\, \delta(\omega_\alpha + \omega_\beta - \omega_\gamma - \omega_\delta), \quad (74)$$

where $v_{\alpha\beta\gamma\delta}$ is given by Eq. (72).

The off-diagonal elements $\mathcal{L}_{12} = \mathcal{L}_{21}$ involve the product $d_{\alpha\beta\gamma\delta}v_{\alpha\beta\gamma\delta}$ and average to zero. This happens because the statistics is symmetric with respect to $\omega_\alpha \to -\omega_\alpha$, and the off-diagonal matrix elements in Eq. (71) are odd functions of the frequencies. To obtain a nonzero value, one has to include the terms subleading in $1/T$, which gives

$$\mathcal{L}_{12} = \mathcal{L}_{21} = -K_1 \frac{\rho^5}{T}, \quad (75)$$

$$K_1 = \frac{\pi}{L} \sum_{\alpha,\beta,\gamma,\delta}{}' (\omega_\alpha + \omega_\beta + \omega_\gamma + \omega_\delta)\, v_{\alpha\beta\gamma\delta}d_{\alpha\beta\gamma\delta} \times V^2_{\alpha\beta\gamma\delta}\, \delta(\omega_\alpha + \omega_\beta - \omega_\gamma - \omega_\delta). \quad (76)$$

In addition to this, a contribution which remains finite at $T \to \infty$ is obtained if nonlinear frequency shifts are included.

C. Including the nonlinear frequency shifts

The arguments of the previous subsection can be quite straightforwardly generalized to include the nonlinear frequency shifts, Eq. (9). When they are included in the frequency δ function in $R_{\alpha\beta\gamma\delta}$, Eq. (43b), the resulting Boltzmann equation, Eq. (44), conserves the total energy

$$\mathcal{E} = \sum_\alpha \omega_\alpha \bar{I}_\alpha + \sum_{\alpha,\beta} V_{\alpha\beta\beta\alpha}\bar{I}_\alpha\bar{I}_\beta. \quad (77)$$

The equilibrium intensities are given by Eq. (45), however, with shifted frequencies,

$$\bar{I}^{\mathrm{eq}}_\alpha = \frac{T}{\omega_\alpha + 2\sum_\beta V_{\alpha\beta\beta\alpha}\bar{I}^{\mathrm{eq}}_\beta - \mu}. \quad (78)$$

Unless the infinite-temperature limit (24) is taken, this is no longer an explicit expression but a self-consistent equation. Its solution to the first order in $1/T$ is given by

$$\bar{I}^{\mathrm{eq}}_\alpha = -\frac{T}{\mu} - \frac{1}{T}\left(\frac{T}{\mu}\right)^2 \left(\omega_\alpha - 2g\frac{T}{\mu}\right) + O(1/T^2). \quad (79)$$

The energy density, in addition to the expression in Eq. (65), should include the nonlinear contribution,

$$\varepsilon_{\rm nl}(x,t) = \int dt'\,\mathcal{T}(t-t')\sum_{\alpha\beta}\mathcal{S}(x-X_{\alpha\beta}) \times V_{\alpha\beta\beta\alpha}\bar{I}_\alpha(t')\,\bar{I}_\beta(t'), \tag{80a}$$

$$X_{\alpha\beta} \equiv \left(\sum_n \phi_{\alpha n}^2\phi_{\beta n}^2\right)^{-1}\sum_n n\phi_{\alpha n}^2\phi_{\beta n}^2. \tag{80b}$$

The current is given by the same Eq. (67), but instead of $v_{\alpha\beta\gamma\delta}$ from Eq. (72), one should use the nonlinear version,

$$\tilde{v}_{\alpha\beta\gamma\delta} = \omega_\alpha(X_\alpha - X_{\alpha\beta\gamma\delta}) + 2\sum_\eta (X_{\alpha\eta} - X_{\alpha\beta\gamma\delta})V_{\alpha\eta\eta\alpha}\bar{I}_\eta + (\alpha\leftrightarrow\beta) - (\alpha\beta\leftrightarrow\gamma\delta). \tag{81}$$

Now, to find $\mathcal{L}_{21}$, we seek for the solution in the form $\bar{I}_\alpha = \rho - \rho^2\chi_1 X_\alpha$ and calculate the energy current. Noting that

$$\sum_\eta X_{\alpha\eta}V_{\alpha\eta\eta\alpha} = gX_\alpha, \tag{82}$$

we arrive at the final result,

$$\hat{\mathcal{L}} = \rho^4\begin{pmatrix} D_0 & -K_1\rho/T + 2g\rho D_0 \\ -K_1\rho/T + 2g\rho D_0 & K_0\end{pmatrix}. \tag{83}$$

D. Analysis of the coupled transport equations

Let us now investigate what the coupled macroscopic equations for the norm and energy density can give at high temperatures. The behavior of the system can be viewed in the (ρ,ε) plane (more precisely, half-plane, as $\rho \geqslant 0$ by construction), shown in Fig. 7. It is convenient to measure ρ and ε in the natural units Ω/g and Ω^2/g. The properties of the grand-canonical equilibrium mapping $(\mu,T)\to(\rho,\varepsilon)$ define three regions in the (ρ,ε) plane [37]. (i) The region below the $T=0$ line is forbidden: for each fixed value of total norm $\mathcal{N} = \sum_n |\psi_n|^2 = L\rho$, the Hamiltonian (3) has an absolute minimum $H_{\rm min} \equiv L\varepsilon_{\rm min}$, so lower energies are not allowed. This is precisely the $T=0$ line. For weak disorder, $W \ll \Omega$, the $T=0$ line is close to that for the clean case $W=0$, $\varepsilon_{\rm min} = -2\Omega\rho + g\rho^2/2$, and it is the latter one that is shown in Fig. 7. (ii) The region between the $T=0$ line and the $T=\infty$ line (the latter is determined by $\varepsilon = g\rho^2$ regardless of the disorder) corresponds to the usual thermal states: For any ρ,ε in this region the corresponding values of μ and T can be found. (iii) The region above the $T=\infty$ line corresponds to the so-called nonthermal states of the system, which cannot be described by the grand-canonical ensemble with a non-negative temperature (negative temperatures are not allowed in the thermodynamic limit $L\to\infty$, since the upper bound for the Hamiltonian for a fixed total norm is $\propto L^2$). For the disorder-free chain, it has been shown that in the nonthermal region the "excess energy" $\varepsilon - g\rho^2$ tends to condense into localized discrete breathers [37,51–56].

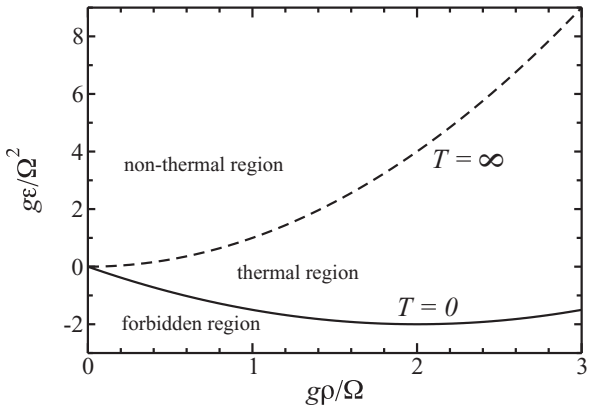

FIG. 7. Three regions in the (ρ,ε) plane (see text for details). The solid and dashed line correspond to zero and infinite temperature, respectively.

The present work is concerned with the regime

$$g\rho \ll \Omega, \quad |\varepsilon| \ll \Omega\rho, \quad \varepsilon \leqslant g\rho^2. \tag{84}$$

In fact, the conditions of validity discussed in Sec. IV imply even stronger restrictions, but for the discussion of this subsection inequalities (84) suffice. Namely they ensure that the thermodynamic relations can be approximated by (see Appendix E for details)

$$\rho = \frac{1}{\lambda} - \frac{2g}{\lambda^3}\frac{1}{T} + O(T^{-2}), \quad \varepsilon = \frac{g}{\lambda^2} - \frac{2\Omega^2}{\lambda^2}\frac{1}{T} + O(T^{-2}), \tag{85}$$

where $\lambda \equiv -\mu/T$. Inequalities (84) imply $T \gg \Omega\rho$, which justifies the expansion in the powers of $1/T$.

The region of the (ρ,ε) plane defined by inequalities (84) is the lower vicinity of the $T=\infty$ line in the left part of Fig. 7. Neglecting the nonlinear frequency shifts, as in Sec. VI B, would correspond to approximating the two parabolas $T = 0,\infty$ by the straight tangents at $\rho=0$, $\varepsilon=0$, and studying the dynamics of ρ on the horizontal tangent to the $T=\infty$ line, $\varepsilon=0$. However, to check how important is the coupling between the norm and the energy transport, one should keep the $g\rho^2$ in the energy and the off-diagonal matrix elements of $\hat{\mathcal{L}}$. The inequality $T \gg \Omega\rho$ is not sufficient to establish which of the two contributions to $\mathcal{L}_{12},\mathcal{L}_{21}$ in Eq. (83) is more important: The first one is proportionl to the small factor $\Omega\rho/T$ and the second one to $g\rho/\Omega$, and it becomes more important than the first only at $T \gg \Omega^2/g$. Thus, both contrubutions are kept in the equations below.

Instead of the temperature, it is convenient to introduce the excess energy

$$u \equiv \varepsilon - g\rho^2 \approx \frac{2\Omega^2\rho^2}{T}. \tag{86}$$

Then, using Eq. (62), Eq. (83) (which is valid to the order $1/T$, other corrections being of the order $1/T^2$ due to the even-odd symmetry in the frequencies), Eq. (85) [which should be inverted keeping the terms $O(1/T)$], and the estimate $K_0 \sim K_1 \sim \Omega^2 D_0$ [a natural guess from Eqs. (60), (74), and (76), which will be verified numerically in Sec. VII], and neglecting terms which have relative smallness $gu/\Omega^2 \ll 1$, one arrives

at the following closed system of equations:

$$\frac{\partial\rho}{\partial t} = \frac{\partial}{\partial x}\left(D_0\rho^2\frac{\partial\rho}{\partial x} + \frac{K_1}{4\Omega^2}\frac{\rho u}{\Omega^2}\frac{\partial u}{\partial x}\right), \tag{87a}$$

$$\frac{\partial u}{\partial t} = 2gD_0\left(\rho\frac{\partial\rho}{\partial x}\right)^2 - \frac{\partial}{\partial x}\left(\frac{2K_0-K_1}{2\Omega^2}\rho u\frac{\partial\rho}{\partial x} - \frac{K_0}{2\Omega^2}\rho^2\frac{\partial u}{\partial x}\right). \tag{87b}$$

Two conclusions can be drawn from these equations. (i) The second term in $\partial\rho/\partial t$ has a relative smallness $u^2/(\Omega\rho)^2$ compared to the first. Thus, in the regime defined by inequalities (84) the use of the nonlinear diffusion equation (1) for the density alone is justified. The origin of this property can be traced back to the symmetry of the spectrum with respect to $\omega_\alpha \to -\omega_\alpha$ (on the average); when all modes are equally populated (at $T\to\infty$) it leads to the relative smallness of $\mathcal{L}_{12},\mathcal{L}_{21}$. If this symmetry is spoiled [e.g., by including $\psi_{n\pm2}$ in the original Eq. (2)], the factor $K_1\rho u/\Omega^2$ would be replaced by $K_2\rho^2$ with $K_2\sim\Omega D_0$; this would produce a term $\sim(\partial/\partial x)(\Omega D_0\rho^2\partial\rho/\partial x)$ in Eq. (87b), and the initial assumption about the smallness of u would be violated after some time. (ii) If one prepares an initial condition with some profile $\rho(x)$, while $u(x)=0$ everywhere, then $\partial u/\partial t>0$, so the system is pulled into the nonthermal region. This fact can be easily understood: if the system tends to an equilibrium state with $\rho(x)=\rho_{\rm eq}$, $\varepsilon(x)=\varepsilon_{\rm eq}$, the equilibrium values must be given by the spatial averages of the initial $\rho(x)$ and $\varepsilon(x)=g\rho^2(x)$; the convexity of the $T=\infty$ line ensures $\varepsilon_{\rm eq}>g\rho_{\rm eq}^2$.

Interestingly, Eqs. (87a) and (87b) remain mathematically correct even for $u>0$, i.e., in the nonthermal region where the reasoning of this section, based on the assumption of local thermal equilibrium, is not supposed to be valid. Formally, the origin of such behavior can be traced to the fact that Eq. (78) still determines a stationary solution of the kinetic equation even for $T<0$, $-\mu/T=\lambda>0$. This stationary solution, however, no longer corresponds to the maximum of the entropy; the latter is reached in a state with one discrete breather on top of the $T=\infty$ background [51]. Nevertheless, this true equilibrium state will be reached only at very long times [52]. This should be especially true in the regime considered here, when the excess energy u is small. At times that are not too long, the system may equilibrate near a transient metastable state with negative temperature. A straightforward analysis of Eqs. (87a) and (87b), linearized around a homogeneous solution $\rho=0, u=u_0$, shows that this homogeneous solution is stable regardless of the sign of u_0. Such equilibration has been observed numerically [57]. Thus, in the nonthermal region, Eqs. (87a) and (87b) describe the system dynamics near such a metastable state.

VII. NUMERICAL RESULTS FOR THE TRANSPORT COEFFICIENTS

The transport coefficients D_0, K_0, K_1 have been evaluated numerically from Eqs. (60), (74), and (76). The frequency δ functions were approximated by boxes of finite width, as discussed in Sec. III E, and the absence of dependence of the results on w has been checked. The dimensionless combinations $D_0\Omega/g^2$, $K_0/(\Omega g^2)$, $K_1/(\Omega g^2)$ depend only on the dimensionless disorder strength, and the results of their numerical evaluation are shown in Fig. 8. At weak disorder, they can be fitted by

$$\begin{Bmatrix} D_0 \\ K_0/\Omega^2 \\ K_1/\Omega^2 \end{Bmatrix} = \frac{g^2}{\Omega}\left(\frac{\Omega}{W}\right)^{4.9\pm0.2}\begin{Bmatrix} 1.5 \\ 1.0 \\ 1.3 \end{Bmatrix}\times 10^4. \tag{88}$$

One feature of these results is that $K_0\sim K_1\sim D_0\Omega^2$. This follows naturally from Eqs. (60), (74), and (76), if the main contribution comes from modes with frequencies $|\omega|\sim\Omega$. Also, comparing Eqs. (88) and (34a), one can see that $D_0\sim\Gamma\xi^2/\rho^2$, which follows from Eqs. (28) and (60), if the main contribution to D_0 and Γ comes from modes with $|d_{\alpha\beta\gamma\delta}|\sim\xi$.

Does the observed dependence of Γ or D_0 on the disorder strength have a simple explanation? If one assumes that the typical value of $V_{\alpha\beta\gamma\delta}$ for modes located on the same localization segment scales as $V_{\alpha\beta\gamma\delta}\sim g(W/\Omega)^a$ with some exponent a, and the summation over β, γ, and δ in Eq. (28) gives a factor $\sim\xi^3$, then $\Gamma\propto W^{2a-6}$. Several values for the exponent a, ranging from 2 to 4, have been suggested in the literature [30,58–62]. In particular, numerical evaluation of the averages $\overline{|V_{\alpha\beta\gamma\delta}|}$, $\overline{V^2_{\alpha\beta\gamma\delta}}$, gave $a=3.3$ [60] or $a=3.4$ [62]. The result of the present work, $\Gamma\propto W^{-0.8}$ from Eq. (34a), is reproduced if one assumes $a=2.6$. The discrepancy is probably due to the fact that in Refs. [60,62] $V_{\alpha\beta\gamma\delta}$ for all eigenmodes were considered, regardless of their frequencies. At the same time, the sums in Eqs. (28) and (60) are contributed only by those modes for which the frequency mismatch $\varpi_{\alpha\beta\gamma\delta}$ is small, and even within this subset there are correlations between the overlaps and the mode frequencies, as illustrated in Appendix A3. Moreover, even though the points on Fig. 8 seem to fall well on a straight line, one cannot exclude that Eq. (88) still does not represent the true asymptotic behavior at weak disorder (see the discussion in Sec. IV C).

The value of D_0 determined from Eq. (60) can be compared to the one extracted from the rate of wave packet spreading obtained by direct numerical integration of Eq. (2) in Ref. [22].

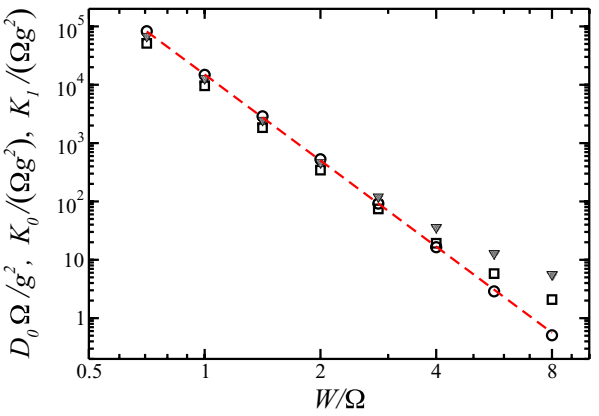

FIG. 8. (Color online) Dependence of the dimensionless transport coefficients $D_0\Omega/g^2$, $K_0/(\Omega g^2)$, and $K_1/(\Omega g^2)$ (circles, squares, and triangles, respectively) on the dimensionless disorder strength W/Ω. The straight line shows the dependence $1.5\times10^4\,(\Omega/W)^{4.9}$.

Equation (1) has a well-known self-similar solution [63,64] (see also Ref. [65] for a comprehensive review),

$$\rho_{\mathcal{N}}(x,t) = \sqrt{\frac{\mathcal{N}}{\pi\sqrt{D_0 t}} - \frac{\pi\, x^2}{4 D_0 t}}, \quad |x| < x_t \equiv \sqrt{\frac{4\mathcal{N}}{\pi}\sqrt{D_0 t}}, \tag{89}$$

and $\rho_{\mathcal{N}}(x,t) = 0$ for $|x| > x_t$. This solution is parametrized by the total norm $\mathcal{N} = \int \rho(x,t)\, dx$, which is determined by the initial conditions and remains conserved in time. Equation (89) also represents the long-time asymptotics of the solution of Eq. (1) for any positive compact initial condition. The second moment for the wave packet described by Eq. (89),

$$m_2(\mathcal{N},t) = \frac{1}{\mathcal{N}} \int_{-\infty}^{\infty} x^2\, \rho_{\mathcal{N}}(x,t)\, dx = \frac{\mathcal{N}}{\pi}\sqrt{D_0 t}, \tag{90}$$

can be directly compared with the numerical result of Ref. [22] for $\log_{10} m_2$, averaged over the disorder realizations,

$$\overline{\log_{10} m_2} = 0.98 + 0.5\, \log_{10}(\Omega t),$$

for $W/\Omega = 4$ and $g\mathcal{N}/\Omega = 0.74 \times 21 \approx 15$. This gives $D_0\Omega/g^2 = 4.0$. It should be noted, however, that $\overline{\log_{10} m_2} \leqslant \log_{10} \overline{m_2}$ as the logarithm is a concave function, so this value is likely to underestimate D_0. At the same time, evaluation of Eq. (60) for $W/\Omega = 4$ gives $D_0\Omega/g^2 = 16 \pm 1$. This can be considered a reasonable agreement, given the fact that $W/\Omega = 4$ is on the borderline of the weak-disorder regime.

Even though the kinetic approach, developed in the present work, does not have intrinsic long-time limitations, in the course of wave packet expansion the density decreases with time, so after some time Eq. (1) with $D(\rho) = D_0\rho^2$ is no longer valid because the density drops below the low-density boundary in Fig. 5. From the data of Ref. [22] it is seen that the wave packet expansion with $m_2 \propto t^{1/2}$ starts to slow down at $m_2 \sim 10^4$, that is, at the average density in the packet $g\rho/\Omega \sim g\mathcal{N}/(4\Omega\sqrt{m_2}) \approx 0.04$, which also agrees with the low-density boundary in Fig. 5.

VIII. CONCLUSIONS

In this paper, we have studied the discrete nonlinear Schrödinger equation in the presence of weak on-site disorder. The nonlinearity was assumed, on the one hand, to be sufficiently weak for the eigenmodes of the linear problem to remain well resolved, but, on the other hand, to be sufficiently strong, for the dynamics of the eigenmode amplitudes to chaotic for almost all modes. It was shown that in this regime the slow dynamics of the eigenmode intensities can be described by a master equation of the Fokker-Planck type. Limits of applicability of the master equation approach have been investigated in detail.

Focusing on the transport of conserved quantities (norm and energy) on macroscopic length and time scales at high temperature, from the master equation we have derived explicit expressions for the macroscopic transport coefficients in terms of the wave functions and frequencies of the eigenmodes of the linear problem. Evaluation of these expressions was performed numerically for different disorder strengths. Analysis of the coupled macroscopic equations for the norm and energy densities have shown that in the considered regime (weak disorder, moderately weak nonlinearity, and high temperature) the effect of the energy transport on the transport of the norm can be neglected, so the norm density ρ satisfies a closed macroscopic equation, which is the nonlinear diffusion equation with the density-dependent diffusion coefficient $D(\rho) = D_0\rho^2$. The numerical value of D_0, obtained from the present theory, is in reasonable agreement with the result of the direct numerical integration of the original nonlinear Schrödinger equation [22].

The density dependence of the diffusion coefficient $D(\rho) = D_0\rho^2$, obtained in the present work, translates into the subdiffusive spreading of an initially localized wave packet with the second moment m_2 growing as $m_2 \propto t^{1/2}$. It is known from numerical simulations that at lower densities or longer times this asymptotics should slow down to $m_2 \propto t^{1/3}$, which would mean $D(\rho) = D_0\rho^4$. Construction of a quantitative theory for this latter regime remains a challenge.

ACKNOWLEDGMENTS

The author is grateful to B. Altshuler, S. Flach, V. Kravtsov, and A. Pikovsky for stimulating discussions.

APPENDIX A: STATISTICS FOR A LINEAR ONE-DIMENSIONAL CHAIN

In this Appendix, several quantities are calculated for the linear eigenvalue problem (5) with the flat box distribution for $\epsilon_n \in [-W/2, W/2]$.

1. Mode spectral density and localization length

The average N-mode spectral density per unit length is defined as

$$\nu_N(\omega) = \lim_{L\to\infty} \frac{1}{L^N} \sum_{\alpha_1,\ldots,\alpha_N=1}^{L} \delta\left(\sum_{i=1}^{N} \omega_{\alpha_i} - \omega\right), \tag{A1}$$

where the δ function should be approximated by a peak of finite width, which should be set to zero *after* the limit $L \to \infty$ is taken. The N-mode frequency spacing within one localization length $\xi(\omega)$ is then $\Delta_N = 1/(\nu_N \xi^N)$.

In the weak-disorder limit, for most frequencies ω, the density of modes can be well approximated by that of the clean chain. $\nu_1(\omega)$ is straightforwardly calculated,

$$\nu_1(\omega) = \int_{-\pi}^{\pi} \frac{dk}{2\pi}\, \delta(\omega - 2\Omega\cos k) = \frac{1}{2\pi}\, \frac{1}{\sqrt{\Omega^2 - (\omega/2)^2}}, \tag{A2}$$

and has square-root type singularities at the band edges, $\omega = \pm 2\Omega$. The main effect of weak disorder is to smear these singularities, as seen from the numerical results shown in Fig. 9(a). The smearing occurs on the frequency scale $||\omega| - 2\Omega| \sim (W/10)^{4/3}\Omega^{-1/3}$.

Given the analytical expression, Eq. (A2), the two-mode spectral density can also be calculated analytically as follows:

$$\begin{aligned}\nu_2(\omega) &= \int_{-\pi}^{\pi} \frac{dk_1}{2\pi}\, \frac{dk_2}{2\pi}\, \delta(\omega - 2\Omega\cos k_1 - 2\Omega\cos k_2) \\ &= \frac{(2/\pi)^2}{4\Omega + |\omega|}\, \mathbf{K}\left(\frac{(4\Omega - |\omega|)^2}{(4\Omega + |\omega|)^2}\right),\end{aligned} \tag{A3}$$

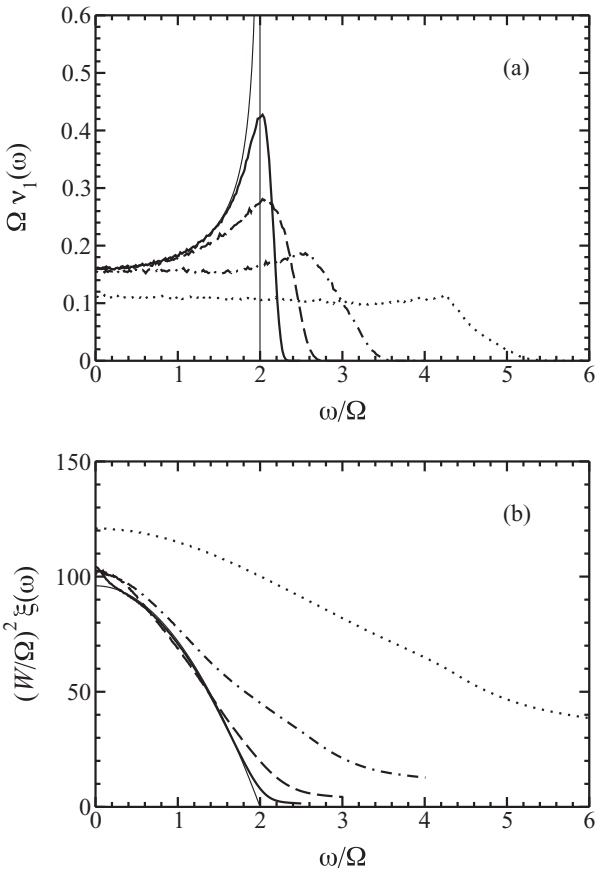

FIG. 9. Single-mode density of states $\nu_1(\omega)$ (a) and the localization length $\xi(\omega)$ (b) for the eigenvalue problem (5), evaluated numerically for $W/\Omega = 1$ (thick solid line), $W/\Omega = 2$ (dashed line), $W/\Omega = 4$ (dash-dotted line), and $W/\Omega = 8$ (dotted line). The thin solid lines correspond to the analytical expressions Eq. (A2) (a) and Eq. (A5) (b).

expressed in terms of the complete elliptic integral $\mathbf{K}(m) = \int_0^{\pi/2}(1 - m\sin^2\phi)^{-1/2}d\phi$. The singularities at $\omega = \pm 2\Omega$ are logarithmic, i.e., weaker than for $\nu_1(\omega)$, as they are smeared by the convolution.

The three-mode spectral density,

$$\nu_3(\omega) = \int_{\max\{-2\Omega,\,-4\Omega+|\omega|\}}^{2} d\omega'\,\nu_1(\omega')\,\nu_2(\omega-\omega'), \qquad \text{(A4)}$$

is regular at $\omega \to 0$, $\Omega\nu_3(0) = 0.1426\ldots$, while at $\omega = \pm 2\Omega$ it has a cusp, $\Omega\nu_3(2\Omega) = 0.1447\ldots$. Thus, in the most relevant interval $|\omega| < 2\Omega$, $\nu_3(\omega)$ is almost a constant. For $2\Omega < |\omega| < 6\Omega$ it monotonously decreases to zero. In general, $\nu_N(\omega)$ have weaker singularities for larger N, and in the limit $N \gg 1$ the shape of $\nu_N(\omega)$ approaches a Gaussian due to the central limit theorem.

For the localization length, the following analytical expression is available in the weak disorder limit [45] as follows:

$$\xi(\omega) = 96\,\frac{\Omega^2}{W^2}\left(1 - \frac{\omega^2}{4\Omega^2}\right). \qquad \text{(A5)}$$

As seen from Fig. 9(b), for $W/\Omega < 2$ this expression works well for most values of ω, except for the band edges and the band center. The latter behavior is a consequence of the well-known anomaly [66–71]. Other anomalies at frequencies corresponding to wave vectors being rational multiples of π [39,72] are beyond the numerical precision of the present calculation.

2. Fluctuations of nonlinear frequency shifts

In this subsection, we present the results for the average relative dispersion of the nonlinear frequency shifts, which was defined in Eq. (13) as the average over all eigenmodes α,

$$\sigma^2 = \frac{1}{L}\sum_{\alpha,\beta}\sum_{n,n'}\phi^2_{\alpha n}\phi^2_{\beta n}\phi^2_{\alpha n'}\phi^2_{\beta n'}, \qquad \text{(A6a)}$$

as well as for

$$\sigma^2_\omega = \frac{1}{L\nu_1(\omega)}\sum_{\alpha,\beta}\delta(\omega-\omega_\alpha)\sum_{n,n'}\phi^2_{\alpha n}\phi^2_{\beta n}\phi^2_{\alpha n'}\phi^2_{\beta n'}, \qquad \text{(A6b)}$$

which restricts the average to eigenmodes α at a given frequency ω.

The numerical results for σ^2 are shown in Fig. 10(a). At weak disorder, the dependence on the disorder strength can be

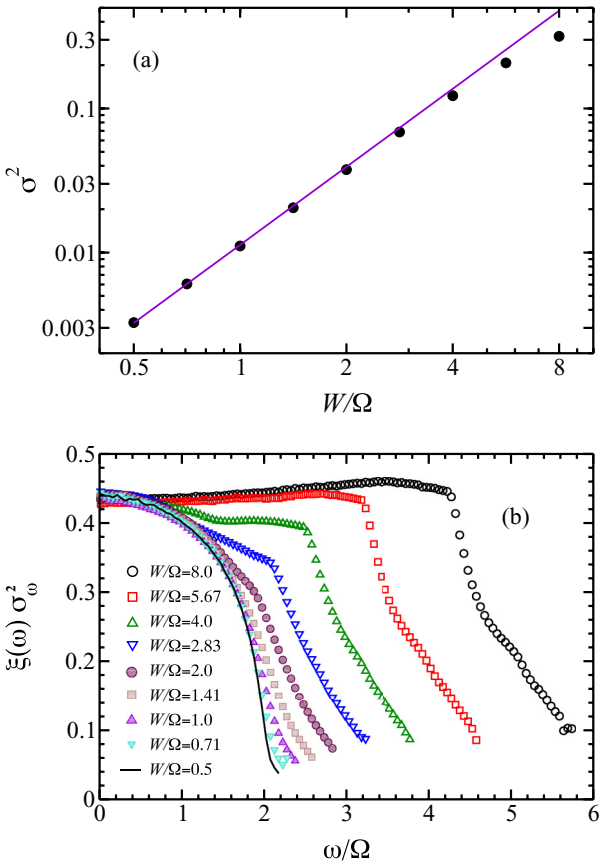

FIG. 10. (Color online) (a) The average relative dispersion σ^2 of the nonlinear frequency shift, Eq. (A6a), as a function of disorder strength (circles), fitted by the dependence in Eq. (A7) (straight line). (b) The rescaled frequency-resolved average relative dispersion, $\sigma^2_\omega\xi(\omega)$, for different disorder strengths.

fitted by the expression

$$\sigma^2 = 0.011 \left(\frac{W}{\Omega}\right)^{1.8}, \tag{A7}$$

that is, $\sigma^2 \propto \xi^{-0.9}$. This dependence is slightly weaker than $\sigma^2 \propto 1/\xi$, which is expected from the simple reasoning of Sec. III A. The reason for this can be understood by looking at the frequency-resolved fluctuation, Eq. (A6b), scaled by the localization length at the same frequency, $\sigma_\omega^2 \xi(\omega)$, plotted in Fig. 10(b). At $W/\Omega \gtrsim 3$, the curves clearly show a different behavior in two distinct frequency intervals, as was also observed in Ref. [73]. At $W/\Omega < 2$ the curves show a clear tendency to collapse on a single universal curve, but the limit is reached quite slowly. Thus, the asymptotics $\sigma^2 \propto 1/\xi$ is expected to set in at small W/Ω, corresponding to extremely large $\xi \gtrsim 1000$. Similar behavior in the statistics of values of a single eigenfunction was observed in Ref. [39]. This is also in agreement with discussion of the relaxation rates in Sec. IV C.

3. Correlations between frequencies and overlaps

To see how the overlaps $V_{\alpha\beta\gamma\delta}$ which contribute to Γ_α are correlated with the frequencies of the corresponding modes, we calculate the following quantity:

$$R(\omega,\omega') = \frac{4\pi\rho^2}{L\overline{\Gamma}_\omega} {\sum_{\alpha,\beta,\gamma,\delta}}' V^2_{\alpha\beta\gamma\delta} \delta(\omega_\alpha + \omega_\beta - \omega_\gamma - \omega_\delta) \times \frac{\delta(\omega_\alpha - \omega)}{\nu_1(\omega)} \frac{\delta(\omega_f - \omega')}{\nu_1(\omega')}, \tag{A8}$$

where $\omega_f = \omega_\gamma$ if $|\omega_\gamma - \omega_\alpha| < |\omega_\delta - \omega_\alpha|$ and $\omega_f = \omega_\delta$ in the opposite case (ω_f is introduced in order to take care of the symmetry $\omega_\gamma \leftrightarrow \omega_\delta$ by choosing the one which is closer to ω_α). Thus defined, $R(\omega,\omega')$ represents the relative weight with which partner modes at some frequency ω' contribute to $\overline{\Gamma}_\omega$, with the normalization

$$\int R(\omega,\omega')\, \nu_1(\omega')\, d\omega' = 1. \tag{A9}$$

$R(\omega = 0,\omega')$ for different disorder strengths is shown in Fig. 11. At weak disorder, the symbols have a tendency to collapse on a single curve. At $W/\Omega \gtrsim 3$, $R(0,\omega')$ has a singularity at $|\omega - \omega'| \to 0$. The present data are not sufficient to establish its precise character (power-law or logarithmic); one can only conclude that the singularity is not stronger than $|\omega - \omega'|^{-0.5}$ and is thus integrable. At the same time, in the disorder-free system, the singularity is $|\omega - \omega'|^{-1}$, as was seen in Sec. IV C. Again, as in Sec. IV C and Appendix A 2, it appears that the values of disorder corresponding for which the results are presented in Fig. 11 are still too large to be in the true weak-disorder limit.

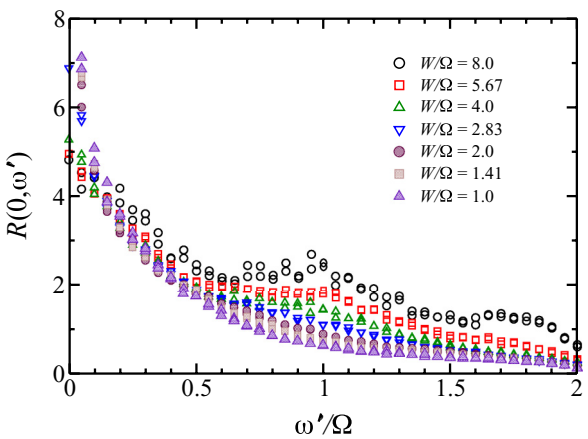

FIG. 11. (Color online) The relative weight of modes with frequency ω' in the decay rate $\overline{\Gamma}_{\omega=0}$ of the modes at $\omega =$ for different disorder strengths.

APPENDIX B: DAMPED OSCILLATOR SUBJECT TO NOISE

Consider a single oscillator, described by the complex amplitude c which satisfies the equation of motion,

$$i\,\frac{dc}{dt} = (\omega + g|c|^2)c, \tag{B1}$$

where ω is the frequency of linear oscillations and g is the anharmonicity. The solution of Eq. (B1) is

$$c(t) = \sqrt{I}\, e^{-i(\omega + gI)t - i\theta^0}. \tag{B2}$$

Let the oscillator now be subject to a white noise and friction,

$$\frac{dc}{dt} = -i(\omega + g|c|^2)c - \frac{\Gamma}{2}\, c + \eta_x(t) + i\,\eta_y(t), \tag{B3}$$

where the noise amplitudes satisfy

$$\langle \eta_i(t)\, \eta_j(t')\rangle = \nu\delta_{ij}\delta(t - t'), \quad i,j = x,y, \tag{B4}$$

where ν measures the strength of the noise and Γ is the friction coefficient. Writing $c = x + iy$, one can introduce the probability distribution function $P(x,y)$ in the complex plane of c and write down the Fokker-Planck equation, corresponding to the Langevin equation (B3),

$$\frac{\partial P}{\partial t} = -\frac{\partial}{\partial x}[\omega + g(x^2 + y^2)]yP + \frac{\partial}{\partial y}[\omega + g(x^2 + y^2)]xP + \frac{\Gamma}{2}\left[\frac{\partial}{\partial x}\,xP + \frac{\partial}{\partial y}\,yP\right] + \frac{\nu}{2}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)P. \tag{B5}$$

If $P(x,y)$ is interpreted as the density in a cloud of particles, the first two lines of Eq. (B5) correspond to the clockwise rotation of the cloud around the origin, the third line to the uniform squeezing towards the origin, and the last line to the uniform spread of the cloud. In the action-angle variables, $x + iy = \sqrt{I}\, e^{-i\theta}$, the same equation becomes

$$\frac{\partial P}{\partial t} = -(\omega + gI)\,\frac{\partial P}{\partial \theta} + \frac{\Gamma\rho}{4I}\,\frac{\partial^2 P}{\partial \theta^2} + \Gamma\,\frac{\partial}{\partial I}\, I\left(\rho\,\frac{\partial}{\partial I} + 1\right)P, \tag{B6}$$

where we introduced $\rho = 2\nu/\Gamma$ [which is the equilibrium average value of I, since the stationary solution of Eq. (B6) is $e^{-I/\rho}$] and $\Gamma\rho I$ plays the role of the action-dependent diffusion coefficient. Upon averaging over the phase θ, Eq. (B6) becomes identical to Eq. (26).

APPENDIX C: NONLINEAR OSCILLATOR UNDER A QUASIPERIODIC FORCE

Consider a single oscillator, described by the complex amplitude c which satisfies the equation of motion,

$$i\frac{dc}{dt} = (\omega + g|c|^2)c + \sum_k f_k e^{-i\omega_k t - i\theta_k}, \tag{C1}$$

where the real f_k,θ_k are the amplitude and the phase of the kth external force oscillating at frequency ω_k. Without loss of generality, we assume $g > 0$, $f_k > 0$. When will the motion of this oscillator be chaotic?

Let us first analyze the case when only one term is present. Substituting

$$c(t) = \sqrt{I(t)}\, e^{-i\omega_k t - i\theta_k - i\theta(t)}, \tag{C2}$$

we arrive at equations of motion which have a Hamiltonian form as follows:

$$\frac{d\theta}{dt} = \omega + gI - \omega_k + \frac{1}{\sqrt{I}} f_k \cos\theta = \frac{\partial \mathcal{H}}{\partial I}, \tag{C3a}$$

$$\frac{dI}{dt} = 2\sqrt{I}\, f_k \sin\theta = -\frac{\partial \mathcal{H}}{\partial \theta}, \tag{C3b}$$

$$\mathcal{H}(I,\theta) = -\varpi_k I + \frac{gI^2}{2} + 2\sqrt{I}\, f_k \cos\theta, \tag{C3c}$$

where we denoted $\varpi_k \equiv \omega_k - \omega$. The stationary points are located at $\theta = 0$ or π, and I can be found from the cubic equation

$$\frac{(gI - \varpi_k)^2}{f_k^2} = \frac{1}{I}. \tag{C4}$$

At $\varpi_k^3 < (27/4)gf_k^2$, there is only one elliptic stationary point at $\theta = \pi$, and the phase portrait of the oscillator motion in the (Re c, Im c) plane has the same topology as in the absence of the external force. At $\varpi_k^3 > (27/4)gf_k^2$, two more stationary points appear at $\theta = 0$, one elliptic and one hyperbolic. The phase portrait of the oscillator in this case is shown in Fig. 12(a). It has a separatrix, which corresponds to the standard case of the nonlinear resonance [3,74].

Let us study this separatrix in more detail. As the stationary points are at $\theta = 0,\pi$ (corresponding to Im $c = 0$) it is convenient to study the Hamiltonian $\mathcal{H}$ as a function of the dimensionless variable $x = \sqrt{g/\varpi_k}$ Re c,

$$\mathcal{H}(x) = \frac{2\varpi_k^2}{g}\left(\frac{x^4}{4} - \frac{x^2}{2} + Fx\right), \quad F \equiv \sqrt{\frac{gf^2}{\varpi_k^3}}, \tag{C5}$$

whose behavior is determined by a single dimensionless parameter F. Let $x_1 < x_2 < x_3$ be the stationary points, i.e., the roots of the cubic equation $\mathcal{H}'(x) = 0$ (the prime indicates the derivative). Noting that $\partial^2\mathcal{H}(I,\theta)/\partial\theta^2 \propto -\cos\theta$ changes sign at $x = 0$, we obtain that the points $x_{1,2}$ are elliptic, while the point x_3 is hyperbolic.

The points $x_\pm$ at which the separatrix crosses the negative real semiaxis of c (that is, $\theta = \pi$) should be found from the equation $\mathcal{H}(x) = \mathcal{H}(x_3)$. Since

$$\mathcal{H}(x) - \mathcal{H}(x_3) = \frac{2\varpi_k^2}{g}(x - x_3)^2\left[\frac{(x + x_3)^2}{4} - \frac{F}{2x_3}\right], \tag{C6}$$

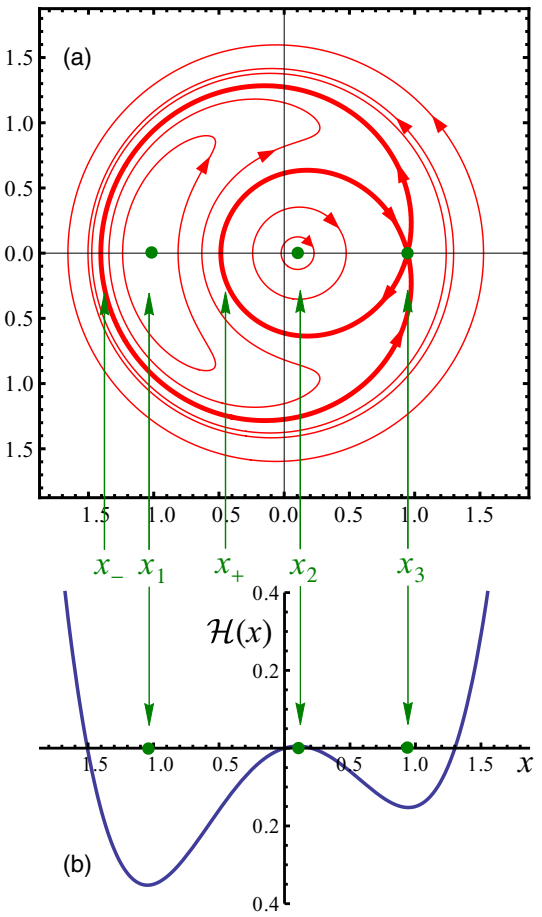

FIG. 12. (Color online) (a) The phase portrait corresponding to the Hamiltonian (C3c) (the thick line shows the separatrix, the the full circles show the stationary points) and (b) the function $\mathcal{H}(x)$ for $\mathcal{F} = 0.1$.

they are given by $x_\pm = -x_3 \pm \sqrt{2F/x_3}$. Both solutions are negative when $\mathcal{H}(x_2) < 0$, which is the case when $F < \sqrt{2/27}$.

An exact analytic expression for x_3 (which depends on the dimensionless parameter F) is not available. Analyzing the cubic parabola $\mathcal{H}'(x)$, one can see that x_1,x_2,x_3 satisfy

$$-\sqrt{4/3} < x_1 < -1, \quad 0 < x_2 < \sqrt{1/3} < x_3 < 1, \tag{C7}$$

where $0,\pm1$ are the roots at $F = 0$, the points $\pm1/\sqrt{3}$ are the extrema of $\mathcal{H}'(x)$, and $-\sqrt{4/3}$ is the solution of $\mathcal{H}'(x) = \mathcal{H}'(1/\sqrt{3})$. Numerically, x_1,x_2,x_3 can be efficiently found by use of the Newton's method starting from $-\sqrt{4/3},0,1$, respectively. It turns out that x_3 is well approximated by

$$x_3 \approx 1/\sqrt{3} + \sqrt{2\sqrt{3} - 3}\,\sqrt{\sqrt{4/27} - F}, \tag{C8}$$

whose relative error does not exceed 1.4% in the whole available range $0 < F < \sqrt{4/27}$. It is this approximate expression that is used in the numerical calculation.

Let us now consider several forces acting on the oscillator. Treating each of them separately, for those of them which have $\varpi_k^3 > (27/4)gf_k^2$ and thus produce a separatrix, we can define the values I_{3k} (the action corresponding to the hyperbolic point) and $I_{\pm,k}$ (the actions corresponding to the points where the separatrix passes through $\theta = \pi$). If the separatrices are

well separated, that is, the width $I_{-,k} - I_{+,k}$ of the separatrix is much smaller than the typical distance $|I_{3k} - I_{3k'}|$ for different k,k', the different resonances do not interfere. If the opposite happens, that is, for some k,k' the intervals $I_{+,k} < I < I_{-,k}$ and $I_{+,k'} < I < I_{-,k'}$ overlap, the motion becomes chaotic, and the chaotic region roughly corresponds to the interval $\min\{I_{+,k}, I_{+,k'}\} < I < \max\{I_{-,k}, I_{-,k'}\}$. This is essentially the Chirikov's criterion [3,74,75].

Thus, for each normal mode α, whose intensity I_α is determined by the initial condition and the anharmonicity is given by $V_{\alpha\alpha\alpha\alpha} = g\sum_n \phi_{\alpha n}^4$, we determine the forces $f_{\alpha\beta\gamma\delta} = V_{\alpha\beta\gamma\delta}\sqrt{I_\beta I_\gamma I_\delta}$ and the corresponding separatrix intervals $I_{+,\beta\gamma\delta} < I < I_{-,\beta\gamma\delta}$. If among various terms at least two distinct triples β,γ,δ can be found such that $I_{+,\beta\gamma\delta} < I_\alpha < I_{-,\beta\gamma\delta}$, the mode α is counted as chaotic.

APPENDIX D: DECOUPLING OF HIGHER MOMENTS

Let us neglect the correlations in zero approximation, $\langle I_\alpha I_\beta\rangle \to \langle I_\alpha\rangle\langle I_\beta\rangle$, and check whether in the next approximation the cumulant $\langle I_\alpha I_\beta\rangle - \langle I_\alpha\rangle\langle I_\beta\rangle$ for $\alpha \neq \beta$ is smaller than the main average $\langle I_\alpha\rangle\langle I_\beta\rangle$. The master equation, Eq. (20), gives

$$\frac{d}{dt}(\langle I_\alpha I_\beta\rangle - \langle I_\alpha\rangle\langle I_\beta\rangle) = 2{\sum_{\gamma,\delta}}'(R_{\alpha\beta\gamma\delta} - 2R_{\alpha\gamma\delta\beta})\bar I_\alpha \bar I_\beta \bar I_\gamma \bar I_\delta, \tag{D1}$$

containing a double sum. At the same time, for the main average, $d(\bar I_\alpha \bar I_\beta)/dt$, from Eq. (44) one straightforardly obtains an expression involving a triple sum. Since each summation involves a large number of terms, the cumulant is smaller then the main average.

APPENDIX E: THERMODYNAMICS AT HIGH TEMPERATURES

Let us denote $\mu/T \equiv -\lambda$ for brevity. Then the partition function,

$$Z(\lambda,T) = \int e^{-H_\lambda/T}\prod_n \frac{d\,\mathrm{Re}\,\psi_n\, d\,\mathrm{Im}\,\psi_n}{\pi}, \tag{E1}$$

for the Hamiltonian

$$H_\lambda = -\Omega\sum_n(\psi_n^*\psi_{n+1} + \psi_{n+1}^*\psi_n) + \sum_n(\epsilon_n + \lambda T)|\psi_n|^2 + \frac{g}{2}\sum_n|\psi_n|^4 \tag{E2}$$

can be straightforwardly calculated in the high-temperature limit as follows:

$$\frac{\ln Z}{L} = \ln\frac{1}{\lambda} - \frac{g}{\lambda^2}\frac{1}{T} + \left(\frac{\Omega^2+\overline{\epsilon_n^2}}{\lambda^2} + \frac{5}{2}\frac{g^2}{\lambda^4}\right)\frac{1}{T^2} + O(T^{-3}), \tag{E3}$$

where $\overline{\epsilon_n^2} = W^2/12$ is the second moment of the disorder potential, $L\to\infty$ is the chain length, and the limit $T\to\infty$ is taken at constant λ. Differentiating this expression with respect to λ and $1/T$, we obtain the norm and energy densities,

$$\rho(\lambda,T) = \frac{1}{\lambda} - \frac{2g}{\lambda^3}\frac{1}{T} + O(T^{-2}), \tag{E4a}$$

$$\varepsilon(\lambda,T) = \frac{g}{\lambda^2} - \left(2\frac{\Omega^2+\overline{\epsilon_n^2}}{\lambda^2} + 5\frac{g^2}{\lambda^4}\right)\frac{1}{T} + O(T^{-2}). \tag{E4b}$$

The $T\to\infty$ limit is reached at the line $\varepsilon = g\rho^2$. Since the temperature $T > 0$ (otherwise the partition function diverges, as the Hamiltonian is not bounded from above), the states of the system for which $\varepsilon > g\rho^2$ are non-Gibbsian; that is, it is impossible to find λ and T which would produce such ρ and ε in the grand-canonical ensemble [37]. It is convenient to introduce the amount of "nonthermal" energy in the system,

$$u \equiv g\rho^2 - \varepsilon = \left(2\frac{\Omega^2+\overline{\epsilon_n^2}}{\lambda^2} + \frac{g^2}{\lambda^4}\right)\frac{1}{T} + O(T^{-2}). \tag{E5}$$

Let us now see what thermodynamic relations are obtained from Eq. (45), the equilibrium solution of the kinetic equation (44) which neglects the nonlinear shifts. Expanding in $1/T$, we obtain

$$\rho(\lambda,T) = \frac{1}{L}\sum_{\alpha=1}^{L}\bar I_\alpha = \frac{1}{\lambda} + O(T^{-2}), \tag{E6a}$$

$$\varepsilon(\lambda,T) = \frac{1}{L}\sum_{\alpha=1}^{L}\omega_\alpha\bar I_\alpha = \frac{1}{L}\sum_{\alpha=1}^{L}\frac{\omega_\alpha^2}{\lambda^2}\frac{1}{T} + O(T^{-2}) = \frac{2\Omega^2+\overline{\epsilon_n^2}}{\lambda^2 T} + O(T^{-2}), \tag{E6b}$$

where the average of ω_α^2 is calculated by noting that it is the trace of the square of the linear operator on the right-hand side of Eq. (5). Equations (E6a) and (E6b) differ from Eqs. (E4a) and (E4b) by the absence of the nonlinear terms proportional to g but also by the coefficients at $\overline{\epsilon_n^2}$. The origin of this latter difference can be traced back to nonvanishing correlations, $\langle I_\alpha I_\beta\rangle \neq \langle I_\alpha\rangle\langle I_\beta\rangle$, when $\alpha = \beta$. Indeed, Eq. (E4b) keeps track of all correlations, while Eq. (E6b) follows from Eq. (44) which was obtained by neglecting correlations. Still, the relative error of Eq. (E6b) with respect to Eq. (E4b) is $\overline{\epsilon_n^2}/\Omega^2 \sim 1/\xi \ll 1$, which was precisely the justification for neglecting the correlations in Eq. (44).

If we use the equilibrium solution which includes the nonlinear frequency shifts, Eq. (79), and calculate the average action and energy densities, it reproduces Eq. (E4a) for ρ, while for ε it gives

$$\varepsilon(\lambda,T) = \frac{g}{\lambda^2} - \left(\frac{2\Omega^2+\overline{\epsilon_n^2}}{\lambda^2} + 4\frac{g^2}{\lambda^4}\right)\frac{1}{T} + O(T^{-2}). \tag{E7}$$

This expression differs from Eq. (E4b) by the numerical coefficients at $\overline{\epsilon_n^2}/\lambda^2$ (discussed above) and at g^2/λ^4. The latter produces a relative error $(g\rho/\Omega)^2 \ll 1$ in the coefficient at $1/T$.

APPENDIX F: SPATIOTEMPORAL SMOOTHING

Let us express $|\psi_n(t)|^2$ in Eq. (47) in terms of the normal mode wave functions $\phi_{\alpha n}$ and amplitudes $c_\alpha(t) = \sqrt{I_\alpha(t)}\, e^{-i\theta_\alpha(t)}$ as follows:

$$|\psi_n(t)|^2 = \sum_{\alpha\beta} \phi_{\alpha n}\phi_{\beta n}\sqrt{I_\alpha(t)\, I_\beta(t)}\, e^{-i\theta_\alpha(t)+i\theta_\beta(t)}. \tag{F1}$$

Only terms with $\alpha = \beta$ survive the convolution with $\mathcal{T}(t)$ when $\tau \gg 1/\Delta_1$. Then, averaging over $\mathcal{F}$, we obtain

$$\rho(x,t) = \int dt'\, \mathcal{T}(t-t') \sum_{\alpha,n} \mathcal{S}(x-n)\, \phi_{\alpha n}^2\, \bar{I}_\alpha(t'). \tag{F2}$$

In the spatial sum, we expand $\mathcal{S}(x-n)$ around $x - X_\alpha$ to the second order as follows:

$$\begin{aligned}\sum_n \mathcal{S}(x-n)\,\phi_{\alpha n}^2 = {} & \sum_n \mathcal{S}(x-X_\alpha)\,\phi_{\alpha n}^2 \\ & - \frac{d\mathcal{S}(x-X_\alpha)}{dx} \sum_n (n-X_\alpha)\phi_{\alpha n}^2 \\ & + \frac{d^2\mathcal{S}(x-X_\alpha)}{dx^2} \sum_n \frac{(n-X_\alpha)^2}{2}\,\phi_{\alpha n}^2.\end{aligned} \tag{F3}$$

The second term vanishes by the definition of X_α, and in the last term the first factor gives $1/\ell^2$, while the sum over n is $\sim\xi^2$.

Equations (57a) and (57b) follow from the fact that for any smooth function $A(x)$, depending on x on the scale $\ell \gg \xi$,

$$\sum_\alpha A(X_\alpha) = \int A(x)\, dx[1 + O(\xi^2/\ell^2)]. \tag{F4}$$

This can be seen by first noting that

$$\sum_\alpha A(X_\alpha) = \sum_n A(n) \sum_\alpha \phi_{\alpha n}^2[1 + O(\xi^2/\ell^2)], \tag{F5}$$

obtained analogously to the previous paragraph. Then

$$\sum_n A(n) = \int A(x)\, dx, \tag{F6}$$

with exponential in ℓ precision if $A(x)$ is infinitely differentiable.

APPENDIX G: TOY MODEL OF AN ELECTRIC *RC* CIRQUIT

It is instructive to see how the formalism of Sec. V works for a very simple toy model, that of an electric *RC* cirquit, shown in Fig. 13. The dynamics of the charge q_n on the nth capacitor is governed by the equation

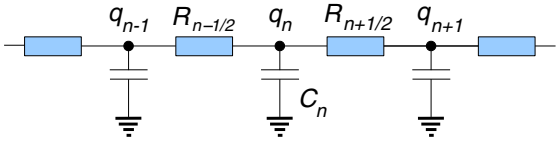

FIG. 13. (Color online) The electric circuit with resistors and capacitors, described by Eq. (G1). The bottom plate of each capacitor is grounded.

$$\begin{aligned}\frac{dq_n}{dt} = {} & \frac{1}{R_{n+1/2}}\left(\frac{q_{n+1}}{C_{n+1}} - \frac{q_n}{C_n}\right) \\ & + \frac{1}{R_{n-1/2}}\left(\frac{q_{n-1}}{C_{n-1}} - \frac{q_n}{C_n}\right).\end{aligned} \tag{G1}$$

Indeed, $\varphi_n \equiv q_n/C_n$ is the electrostatic potential on the upper plate of the nth capacitor, and $(\varphi_{n-1} - \varphi_n)/R_{n-1/2}$ is the current flowing through the resistor between the capacitors n and $n-1$. In equilibrium, the potential is constant along the chain.

For the simple model of Eq. (G1), the exact relation between the current and the macroscopic charge density or potential gradient can be derived straightforwardly. Indeed, noticing that in a stationary situation the currents through all resistors should be the same, we the potential drop is determined by

$$\frac{\varphi_{n+1} - \varphi_n}{R_{n+1/2}} = -J = \text{const}, \tag{G2}$$

so the potential difference between any two capacitors n and $n+\ell$ is given by

$$\varphi_{n+\ell} - \varphi_n = -J \sum_{n'=n}^{n+\ell-1} R_{n'+1/2}. \tag{G3}$$

Taking $\ell \gg 1$, from this we can calculate the macroscopic potential gradient

$$\frac{\partial\varphi}{\partial x} \approx \frac{\varphi_{n+\ell} - \varphi_n}{\ell} = -J\overline{R}, \tag{G4}$$

where $\overline{R}$ is the average resistance,

$$\overline{R} = \lim_{\ell\to\infty} \frac{1}{\ell} \sum_{n'=n}^{n+\ell-1} R_{n'+1/2}. \tag{G5}$$

The "thermodynamic equation of state," relating the charge density ρ to the potential, is

$$\rho = \frac{1}{\ell}\sum_{n'=n}^{n+\ell} q_{n'} = \frac{1}{\ell}\sum_{n'=n}^{n+\ell} C_{n'}\varphi = \overline{C}\varphi, \tag{G6}$$

where $\overline{C}$ is the average capacitance, defined analogously to $\overline{R}$. As a result, we obtain the sought relation between the current and the gradient of the potential or charge density as follows:

$$J = -\frac{1}{\overline{R}}\frac{\partial\varphi}{\partial x} = -\frac{1}{\overline{R}\,\overline{C}}\frac{\partial\rho}{\partial x}. \tag{G7}$$

Let us now see how the approach of Sec. V works for Eq. (G1). The macroscopic density and current are defined as

$$\rho(x,t) = \int dt'\, \mathcal{T}(t-t') \sum_n \mathcal{S}(x-n)\, q_n(t'), \tag{G8a}$$

$$\begin{aligned}J(x,t) = {} & \int dt'\, \mathcal{T}(t-t') \sum_n \tilde{\mathcal{S}}(x-n)\,\frac{dq_n(t')}{dt'} \\ = {} & \int dt'\, \mathcal{T}(t-t') \sum_n \tilde{\mathcal{S}}(x-n)\Bigg[\frac{1}{R_{n+1/2}} \\ & \times\left(\frac{q_{n+1}}{C_{n+1}} - \frac{q_n}{C_n}\right) + \frac{1}{R_{n-1/2}}\left(\frac{q_{n-1}}{C_{n-1}} - \frac{q_n}{C_n}\right)\Bigg]\end{aligned}$$

$$= \int dt' \mathcal{T}(t-t') \sum_n \frac{\tilde{\mathcal{S}}(x-n) - \tilde{\mathcal{S}}(x-n-1)}{R_{n+1/2}} \times \left(\frac{q_{n+1}}{C_{n+1}} - \frac{q_n}{C_n} \right)$$
$$\approx \int dt' \mathcal{T}(t-t') \sum_n \frac{\mathcal{S}(x-n-1/2)}{R_{n+1/2}} \times \left[\frac{q_n(t')}{C_n} - \frac{q_{n+1}(t')}{C_{n+1}} \right]. \quad \text{(G8b)}$$

To find the current response to a small gradient of the potential, let us look for a stationary solution of Eq. (G1) in the form

$$q_n = C_n(\varphi_{\rm eq} - \chi n + r_n), \quad \text{(G9)}$$

with the requirement $\overline{r_n} = 0$. The corrections r_n can be found from the equations

$$\frac{r_{n+1} - r_n}{R_{n+1/2}} - \frac{r_n - r_{n-1}}{R_{n-1/2}} = \frac{\chi}{R_{n+1/2}} - \frac{\chi}{R_{n-1/2}}, \quad \text{(G10)}$$

which are satisfied when

$$\frac{r_{n+1} - r_n}{R_{n+1/2}} = \frac{\chi}{R_{n+1/2}} + C. \quad \text{(G11)}$$

The constant C should be chosen to ensure $\overline{r_n} = 0$. This condition yields $C = -\chi/\overline{R}$ and the solution

$$r_n = r_0 + \sum_{n'=0}^{n-1} \chi \left(1 - \frac{R_{n'+1/2}}{\overline{R}} \right). \quad \text{(G12)}$$

With this solution, the current is obtained from Eq. (G8b), which gives $J = \chi/\overline{R}$.

[1] M. B. Gertsenshtein and V. B. Vasil'ev, Theory Probab. Appl. **4**, 391 (1959).
[2] N. F. Mott and W. D. Twose, Adv. Phys. **10**, 107 (1961).
[3] G. M. Zaslavsky, *Chaos in Dynamic Systems* (Harwood, Newark, 1985).
[4] A. J. Lichtenberg and M. A. Lieberman, *Regular and Chaotic Dynamics*, 2nd ed (Springer-Verlag, New York, 1992).
[5] D. L. Shepelyansky, Phys. Rev. Lett. **70**, 1787 (1993).
[6] M. I. Molina, Phys. Rev. B **58**, 12547 (1998).
[7] G. Kopidakis, S. Komineas, S. Flach, and S. Aubry, Phys. Rev. Lett. **100**, 084103 (2008).
[8] A. S. Pikovsky and D. L. Shepelyansky, Phys. Rev. Lett. **100**, 094101 (2008).
[9] J. Fröhlich, T. Spencer, and C. E. Wayne, J. Stat. Phys. **42**, 247 (1986).
[10] M. Johansson, G. Kopidakis, and S. Aubry, Europhys. Lett. **91**, 50001 (2010).
[11] V. Oganesyan, A. Pal, and D. A. Huse, Phys. Rev. B **80**, 115104 (2009).
[12] D. M. Basko, Ann. Phys. **326**, 1577 (2011).
[13] A. Pikovsky and S. Fishman, Phys. Rev. E **83**, 025201(R) (2011).
[14] M. Mulansky, K. Ahnert, A. Pikovsky, and D. L. Shepelyansky, J. Stat. Phys. **145**, 1256 (2011).
[15] D. M. Basko, Phys. Rev. E **86**, 036202 (2012).
[16] Ch. Skokos, I. Gkolias, and S. Flach, Phys. Rev. Lett. **111**, 064101 (2013).
[17] Y. Lahini, A. Avidan, F. Pozzi, M. Sorel, R. Morandotti, D. N. Christodoulides, and Y. Silberberg, Phys. Rev. Lett. **100**, 013906 (2008).
[18] E. Lucioni, B. Deissler, L. Tanzi, G. Roati, M. Zaccanti, M. Modugno, M. Larcher, F. Dalfovo, M. Inguscio, and G. Modugno, Phys. Rev. Lett. **106**, 230403 (2011).
[19] S. Fishman, Y. Krivolapov, and A. Soffer, Nonlinearity **25**, R53 (2012).
[20] S. Flach, D. O. Krimer, and Ch. Skokos, Phys. Rev. Lett. **102**, 024101 (2009).
[21] Ch. Skokos and S. Flach, Phys. Rev. E **82**, 016208 (2010).
[22] T. V. Laptyeva, J. D. Bodyfelt, D. O. Krimer, Ch. Skokos, and S. Flach, Europhys. Lett. **91**, 30001 (2010).
[23] J. D. Bodyfelt, T. V. Laptyeva, Ch. Skokos, D. O. Krimer, and S. Flach, Phys. Rev. E **84**, 016205 (2011).
[24] G. Benettin, J. Fröhlich, and A. Giorgilli, Commun. Math. Phys. **119**, 95 (1988).
[25] W.-M. Wang and Z. Zhang, J. Stat. Phys. **134**, 953 (2009).
[26] S. Fishman, Y. Krivolapov, and A. Soffer, Nonlinearity **22**, 2861 (2009).
[27] M. Mulansky, K. Ahnert, and A. Pikovsky, Phys. Rev. E **83**, 026205 (2011).
[28] M. Mulansky and A. Pikovsky, New J. Phys. **15**, 053015 (2013).
[29] M. Mulansky, arXiv:1310.8116.
[30] E. Michaely and S. Fishman, Phys. Rev. E **85**, 046218 (2012).
[31] V. E. Zakharov, V. S. L'vov, and G. E. Falkovich, *Kolmogorov Spectra of Turbulence I: Wave Turbulence* (Springer-Verlag, Berlin, 1992).
[32] S. Flach, Chem. Phys. **375**, 548 (2010).
[33] M. Mulansky and A. Pikovsky, Europhys. Lett. **90**, 10015 (2010).
[34] S. Flach, M. Ivanchenko, and N. Li, Pramana J. Phys. **77**, 1007 (2011).
[35] T. V. Laptyeva, J. D. Bodyfelt, and S. Flach, Physica D **256**, 1 (2013).
[36] F. Huveneers, Nonlinearity **26**, 837 (2013).
[37] K. Ø. Rasmussen, T. Cretegny, P. G. Kevrekidis, and N. Grønbech-Jensen, Phys. Rev. Lett. **84**, 3740 (2000).
[38] S. Iubini, S. Lepri, and A. Politi, Phys. Rev. E **86**, 011108 (2012).
[39] V. E. Kravtsov and V. I. Yudson, Ann. Phys. **326**, 1672 (2011).
[40] G. Schwiete and A. M. Finkel'stein, Phys. Rev. Lett. **104**, 103904 (2010); Phys. Rev. A **87**, 043636 (2013).
[41] N. Bilas and N. Pavloff, Eur. Phys. J. D **40**, 387 (2006).
[42] V. Gurarie, G. Refael, and J. T. Chalker, Phys. Rev. Lett. **101**, 170407 (2008).
[43] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics, Non-Relativistic Theory* (Pergamon Press, Oxford, 1991).
[44] D. M. Basko, I. L. Aleiner, and B. L. Altshuler, Ann. Phys. (NY) **321**, 1126 (2006).
[45] D. J. Thouless, J. Phys. C **6**, L49 (1973).
[46] A. Iomin, Phys. Rev. E **80**, 037601 (2009); **81**, 017601 (2010).

[47] I. L. Aleiner, B. L. Altshuler, and G. V. Shlyapnikov, Nat. Phys. **6**, 900 (2010).
[48] D. M. Basko, Phys. Rev. Lett. **91**, 206801 (2003).
[49] A. A. Gogolin, V. I. Melnikov, and E. I. Rashba, Zh. Eksp. Teor. Fiz. **69**, 327 (1975) [Sov. Phys. JETP **42**, 168 (1975)].
[50] B. Shapiro, in *Percolation Structures and Processes*, Annals of the Israel Physical Society, edited by G. Deutscher, R. Zallen, and J. Adler, Vol. 5 (Institute of Physics, Jerusalem, 1983).
[51] B. Rumpf, Phys. Rev. E **69**, 016618 (2004).
[52] M. Johansson and K. Ø. Rasmussen, Phys. Rev. E **70**, 066610 (2004).
[53] B. Rumpf, Europhys. Lett. **78**, 26001 (2007).
[54] B. Rumpf, Phys. Rev. E **77**, 036606 (2008).
[55] B. Rumpf, Physica D **238**, 2067 (2009).
[56] S. Iubini, R. Franzosi, R. Livi, G.-L. Oppo, and A. Politi, New J. Phys. **15**, 023032 (2013).
[57] M. Mulansky, K. Ahnert, A. Pikovsky, and D. L. Shepelyansky, Phys. Rev. E **80**, 056212 (2009).
[58] D. L. Shepelyansky, Phys. Rev. Lett. **73**, 2607 (1994).
[59] Y. Imry, Europhys. Lett. **30**, 405 (1995).
[60] K. Frahm, A. Müller-Groeling, J.-L. Pichard, and D. Weinmann, Europhys. Lett. **31**, 169 (1995).
[61] I. V. Ponomarev and P. G. Silvestrov, Phys. Rev. B **56**, 3742 (1997).
[62] D. O. Krimer and S. Flach, Phys. Rev. E **82**, 046221 (2010).
[63] Y. B. Zeldovich and A. Kompaneets, *Collected Papers of the 70th Anniversary of the Birth of Academician A. F. Ioffe* (Izd. Akad. Nauk SSSR, Moscow, 1950).
[64] G. I. Barenblatt, Prikl. Mat. Mekh. **16**, 67 (1952).
[65] J. L. Vazquez, *The Porous Medium Equation: Mathematical Theory* (Oxford University Press, Oxford, 2007).
[66] L. P. Gorkov and O. N. Dorokhov, Solid State Commun. **20**, 789 (1976)
[67] G. Czycholl, B. Kramer, and A. MacKinnon, Z. Phys. B **43**, 5 (1981).
[68] M. Kappus and F. Wegner, Z. Phys. B **45**, 15 (1981).
[69] H. Schomerus and M. Titov, Phys. Rev. B **67**, 100201(R) (2003).
[70] L. I. Deych, M. V. Erementchouk, A. A. Lisyansky, and B. L. Altshuler, Phys. Rev. Lett. **91**, 096601 (2003).
[71] V. E. Kravtsov and V. I. Yudson, Phys. Rev. B **82**, 195120 (2010).
[72] B. Derrida and E. Gardner, J. Phys. (Paris) **45**, 1283 (1984)
[73] S. Johri and R. N. Bhatt, Phys. Rev. Lett. **109**, 076402 (2012).
[74] B. V. Chirikov, Phys. Rep. **52**, 263 (1979).
[75] B. V. Chirikov, J. Nucl. Energy Part C: Plasma Phys. **1**, 253 (1960).