On the local nature and scaling of chaos in weakly nonlinear disordered chains

D. M. Basko
Université Grenoble 1/CNRS, LPMMC UMR 5493, 25 rue des Martyrs, 38042 Grenoble, France

The dynamics of a disordered nonlinear chain can be either regular or chaotic with a certain probability. The chaotic behavior is often associated with the destruction of Anderson localization by the nonlinearity. In the present work, it is argued that at weak nonlinearity chaos is nucleated locally on rare resonant segments of the chain. Based on this picture, the probability of chaos is evaluated analytically. The same probability is also evaluated by direct numerical sampling of disorder realizations, and quantitative agreement between the two results is found.

I. INTRODUCTION

Dynamics of classical disordered nonlinear chains is governed by an interplay of two fundamental phenomena: Anderson localization (AL) and chaos. AL, originally introduced as full suppression of diffusion by interference in random lattices [1] and used to explain electronic metal-insulator transitions, turned out to be a universal wave phenomenon and was observed in such diverse systems as microwaves, light, acoustic waves, and ultracold atoms [2]. It is most pronounced in one dimension, where all eigenmodes of a disordered linear system are exponentially localized by arbitrarily weak disorder [3].

By now, AL is well understood for linear systems [4]. In the presence of a nonlinearity, the situation is more complicated, and many fundamental questions are still open. For example, it is unclear what happens to an initially localized wave packet at very long times (see Ref. [5] for a review). In a linear system with AL, it remains exponentially localized at all times. With nonlinearity, the wave packet width was found to increase as a subdiffusive power of time [5–9]. In contrast, a rigorous argument shows that at long times the spreading, if any, must be slower than any power of time [10]. 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 [11]. An indication for a slowing down of the power law has also been seen in the scaling analysis of numerical results [12]. Finally, a possible mechanism of breakdown of subdiffusion at long times is presented in Ref. [13].

One of the difficulties in describing nonlinear system is that their dynamics can be chaotic [14–16]. For Hamiltonian systems, close to integrable and with a finite number of the degrees of freedom, the volume of the chaotic phase space is small as long as the integrability-breaking perturbation (in our case, the coupling between the oscillators or the nonlinearity) is weak, as guaranteed by the Kolmogorov-Arnold-Moser (KAM) theorem.

For a random system, the description of chaotic and regular motion has to be probabilistic. In the pioneering work [17], the existence of a dense set of regular trajectories was proven for a class of disordered weakly-nonlinear lattices. It has been argued for the disordered nonlinear Schrödinger chain (DNLS) that the measure of such set in the phase space is finite [18]. This measure, averaged over the disorder, was recently estimated for DNLS both analytically [19] and numerically [20]. In particular, when the nonlinearity is weak, chaos should appear locally. Namely, for a sufficiently long chain, arbitrarily separated in two segments, the probability to be on a regular trajectory is given by the product of the two individual probabilities for the segments. This naturally follows from two conditions: (i) if any of the segments is chaotic, the whole chain is chaotic, and (ii) the probabilities for the two segments are independent. Moreover, in Ref. [19] an explicit mechanism for chaos generation was proposed: two coinciding resonances for rare combinations of three oscillators. Still, the results of Refs. [18–21], did not agree quantitatively, and the origin of the discrepancy is currently not understood. Evidence for a local origin of chaos has also been found in the simulations of the dynamics of a classical spin chain [21].

In the present work, the probability of chaotic behavior (appearance of a non-zero Lyapunov exponent) is studied for another system of coupled nonlinear oscillators [see Eq. (1) below], which turns out to be simpler to analyze than the DNLS. The probability of chaos is calculated at low energy densities \( \epsilon \) by two different methods: (i) by the analysis of the phase space of an effective Hamiltonian describing a resonant triple of oscillators, in the spirit of Ref. [19], and (ii) by direct numerical sampling of many disorder realizations, and counting those with nonzero Lyapunov exponent, analogously to Ref. [20].

Our main results are the following. (i) The two calculations agree quantitatively, including both the leading low-\( \epsilon \) scaling exponent, and the numerical prefactor, thereby confirming the dominant role of resonant triples in generation of chaos at low energies. (ii) This agreement sets in at unexpectedly small values of \( \epsilon \), while at moderately low \( \epsilon \) the numerical results fall on an intermediate asymptotics. This intermediate asymptotics is not controlled by any small parameter and seems to exist for purely numerical reasons; the microscopic mechanism responsible for it remains unclear at the moment.
II. STATEMENT OF THE PROBLEM

Here we study the model defined by the Hamiltonian

\[ H(p_n, q_n) = \frac{p_n^2}{2m} + \frac{m \omega_n^2 q_n^2}{2} + \frac{g}{4} \sum_{n=1}^{L-1} (q_{n+1} - q_n)^4, \]

where \( q_n, p_n \) are the coordinate and momentum of the \( n \)th oscillator, and \( \omega_n^2 \) are independent random variables uniformly distributed in the interval

\[ \frac{W^2}{2} < \omega_n^2 < \frac{3W^2}{2}, \]

where \( W \) is the disorder strength. This model belongs to the class of models considered in Ref. [17]. It has also been studied in Refs. [12, 22], however, in the latter two works focused on spreading of an initially localized wave packet, rather than on the probability of chaos.

The model of Eq. (1) is different from DNLS in two aspects. First, it has only one conserved quantity (energy) in contrast to the two (energy and total norm) for DNLS. Second, it has only one dimensionless parameter,

\[ \epsilon_n = \frac{g}{m^2 W^4} \left( \frac{p_n^2}{2m} + \frac{m \omega_n^2 q_n^2}{2} \right), \]

controlling both the nonlinearity and the coupling between the oscillators. In DNLS they are controlled separately by two independent dimensionless parameters, and when both are small, it matters which one is smaller. In the FSW case, both limits of weak coupling and weak nonlinearity correspond to \( \epsilon_n \to 0 \), which makes it easier to analyze than DNLS. From now on, we measure momenta, coordinates, and time in the units of \( \sqrt{m^3 W^4/g}, \sqrt{mW^2/g} \), and \( 1/W \), respectively, which is equivalent to setting \( m, W, g = 1 \).

For a given realization of disorder \( \{\omega_n^2\} \) and a given initial condition \( \{p_n, q_n\} \), the system trajectory is either regular (quasi-periodic) or chaotic. Characterizing each initial condition by the typical energy \( \epsilon \) per oscillator, we define the probability for the initial condition to be chaotic as

\[ P(\epsilon, L) = \int \Theta(\{p_n, q_n\}, \{\omega_n^2\}) \times \prod_{n=1}^{L} \delta \left( \frac{p_n^2 + \omega_n^2 q_n^2}{2} - \epsilon \right) \frac{dp_n dq_n d\omega_n}{2\pi/\omega_n}, \]

where \( \Theta(\{p_n, q_n\}, \{\omega_n^2\}) = 1 \) if the trajectory is chaotic and zero otherwise. Eq. (4) corresponds to fixing the energies of all oscillators to be \( \epsilon \), and will be referred to as the fixed-\( \epsilon \) ensemble. Alternatively, one can fix the energies only on average, replacing the \( \delta \)-function in Eq. (4) by \( (1/\epsilon) \exp\left(-\frac{(p_n^2 + \omega_n^2 q_n^2)}{2\epsilon}\right) \), i.e., a thermal distribution with temperature \( \epsilon \) (provided that \( \epsilon \ll 1 \), which is our main focus) [23]. In the present paper we will work with the fixed-\( \epsilon \) ensemble which is easier to handle numerically. The corresponding initial conditions can be represented as

\[ p_n = \sqrt{2\epsilon} \cos \phi_n, \quad q_n = \frac{\sqrt{2\epsilon}}{\omega_n} \sin \phi_n, \]

where the phases \( \phi_n \) are uniformly distributed on the interval \([0, 2\pi]\).

The property of locality, mentioned in the introduction, leads to the dependence \( P(\epsilon, L) = 1 - e^{-w(\epsilon)L} \) at sufficiently large \( L \), where the quantity \( w(\epsilon) \), which we call average chaotic fraction (as in Ref. [10]), does not depend on \( L \). The main goal of the present work is to establish its asymptotic behavior \( w(\epsilon \to 0) \), by two methods: (i) relating it to the chaotic phase volume of three resonant oscillators, as in Ref. [19], and (ii) by direct numerical sampling, analogously to Ref. [20]. The latter method also provides a check for the locality hypothesis: if indeed \( P(\epsilon, L) = 1 - e^{-w(\epsilon)L} \), the quantity

\[ \frac{1}{L} \ln \left( \frac{1}{1 - P(\epsilon, L)} \right) \]

does not depend on \( L \), so that all curves for different \( L \) should collapse on a single curve \( w(\epsilon) \). Thanks to the relative simplicity of the system defined by Eq. (1), both calculations can be carried out all the way to the final result, which turns out to be

\[ w(\epsilon \to 0) = Ae^2, \]

where \( A \approx 1.37 \cdot 10^3 \) for the fixed-\( \epsilon \) ensemble.

III. CALCULATION

The reduction of the problem to a few-oscillator configuration [19] is based on the two-resonance picture for weakly non-integrable systems [14,15]. The strongest resonant term in the non-integrable perturbation of an integrable system (the so-called guiding resonance) produces a separatrix in the system phase space. This separatrix is destroyed by another term in the perturbation, which creates a thin stochastic layer in the surrounding part of the phase space. In a disordered system, the main contribution to the chaotic phase space comes from those configurations of disorder and from those regions of the phase space where both perturbation terms are resonant, so one cannot really distinguish between the one responsible for the appearance of the separatrix and the one responsible for its destruction. It is crucial that two resonance conditions should be met simultaneously.

Resonances involving many oscillators are expected to give a subleading contribution to the chaotic phase volume, as the corresponding perturbation terms can be generated in higher orders of perturbation theory, thus resulting in high powers of \( \epsilon \). The minimal number of oscillators needed to generate chaos in the model of Eq. (1)
is two [24]. However, even if \( \omega_1 \approx \omega_2 \), the frequency of the separatrix-destroying perturbation, \( \omega_1 + \omega_2 \), cannot be small for the chosen disorder distribution, \( 1/2 < \omega_n^2 < 3/2 \), which leads to an exponential suppression of the chaotic phase volume [14, 16]. In fact, for the particular case of Eq. (1), the situation is even worse: no separatrix exists in the phase space of the slow motion (see Appendix B). Thus, the dominant contribution to the chaotic phase volume should come from triples of oscillators with all three frequencies close to each other. The oscillators should also be neighboring each other in space, since coupling distant oscillators requires high orders of perturbation theory and results in a high power of \( \epsilon \). Thus, \( w(\epsilon) \) is essentially determined by the probability (per unit length along the chain) to find three neighboring oscillators whose frequencies differ by \( \sim \epsilon \).

This fixes the power \( w(\epsilon) \propto \epsilon^2 \).

To put this argument on a quantitative basis, we assume \( \omega_1 \approx \omega_2 \approx \omega_3 \), and average Eq. (1) over fast oscillations (see Appendix B for details). This gives the effective Hamiltonian of the resonant triple:

\[
H_{\text{tr}} = \Omega |\Psi_1|^2 - (\Omega + \Omega') |\Psi_2|^2 + \Omega' |\Psi_3|^2 + \frac{1}{2} (|\Psi_1|^4 + |\Psi_3|^4 + |\Psi_1 - \Psi_2|^4 + |\Psi_2 - \Psi_3|^4),
\]

written in terms of complex canonical variables \( i\Psi, \Psi_n \), related to \( p_n, q_n \) as

\[
\Psi_n = \frac{e^{i\omega t}}{\sqrt{I}} \left( \frac{p_n}{\sqrt{2}\omega_n} - i\sqrt{2}q_n \right), \quad I = \sum_{n=1}^{3} \frac{p_n^2 + \omega_n^2 q_n^2}{2\omega_n},
\]

where \( \omega = (\omega_1 + \omega_2 + \omega_3)/3 \). The rescaling of \( \Psi \)'s by \( \sqrt{I} \) restricts them to the unit sphere, \( |\Psi_1|^2 + |\Psi_2|^2 + |\Psi_3|^2 = 1 \), which is invariant under the dynamics generated by the Hamiltonian in Eq. (4). The rescaled frequency mismatches \( \Omega, \Omega' \) are defined as

\[
\Omega = \frac{2\omega_1 - \omega_2 - \omega_3}{(9/4)(I/\omega^2)}, \quad \Omega' = \frac{2\omega_3 - \omega_2 - \omega_1}{(9/4)(I/\omega^2)}.
\]

For the fixed-\( \epsilon \) ensemble, Eq. (1), we have \( I = 3\epsilon/\bar{\omega} + O(\epsilon^2) \), and the initial condition should be chosen in the form \( \Psi_1 = e^{-i\varphi/\sqrt{3}}, \Psi_2 = e^{i\varphi + i\varphi'}/\sqrt{3}, \Psi_3 = e^{-i\varphi'}/\sqrt{3} \), where \( 0 < \varphi, \varphi' < 2\pi \) are random phases (the global phase drops out of the result). Then, defining \( \Theta(\varphi, \varphi', \Omega, \Omega') \) to be 1 if the corresponding trajectory is chaotic for given values of \( \varphi, \varphi', \Omega, \Omega' \) and 0 otherwise, we obtain from Eq. (4):

\[
w(\epsilon) = 81\epsilon^2 \int_{-\infty}^{\infty} d\Omega d\Omega' \int_{0}^{2\pi} d\varphi d\varphi' 2\pi \Theta(\varphi, \varphi'; \Omega, \Omega').
\]

Since neither the Hamiltonian [Eq. (7)] nor the initial condition contain any small parameters, the integral over \( \Omega, \Omega' \) in Eq. (10) is dominated by \( |\Omega|, |\Omega'| \approx 1 \), while for larger mismatches the chaotic regions quickly shrink. Thus, the limits of the \( \Omega, \Omega' \)-integration (which are \( \sim 1/\epsilon \)) have been extended to infinity. The factor \( \epsilon^2 \) in front of the integral appears because \( \Omega, \Omega' \propto 1/I \propto 1/\epsilon \), see Eq. (9). The integral is evaluated numerically to be \( 16.9 \pm 0.2 \), leading to \( w(\epsilon \to 0) = 4\epsilon^2 \) with \( \Lambda = 1.37 \cdot 10^3 \), as stated in Eq. (10).

The key element of the numerical procedure is the criterion which enables one to distinguish between regular and chaotic motion. Formally, one studies the largest Lyapunov exponent \( \sigma \), characterizing the mean exponential rate of divergence of two initially close trajectories,

\[
\sigma = \lim_{t \to \infty} \frac{\Lambda(t)}{t}, \quad \Lambda(t) = \ln \frac{d(t)}{d(0)},
\]

where \( d(t) \) is the distance between points belonging to the two trajectories [16]. If \( \sigma = 0 \), the motion is regular, if \( \sigma > 0 \), it is chaotic. In practice, one may integrate the equations of motion for a sufficiently long time \( T \), and consider \( \sigma = 0 \) if \( \Lambda(T)/T < (\text{a few times})/(1/T) \) [21, 22]. Here we use a slightly different criterion, based on the behavior of \( \Lambda(t) \) in the whole integration interval \( 0 < t < T \). Namely, we check how well \( \Lambda(t) \) can be approximated by a logarithmic function, as described in Appendix B.

In Fig. 1 the \( \varphi, \varphi' \)-integral in Eq. (10) is plotted as a function of \( \Omega, \Omega' \). The fact that the chaotic region is confined in all directions, represents a numerical proof of the statement that chaos arises mostly in the regions where two resonant conditions are satisfied simultaneously.

Using the same algorithm, \( w(\epsilon) \) was also calculated by direct numerical sampling of disorder realizations on chains of lengths \( L = 8, 16, 32, 64, 128 \). For \( \epsilon > 10^{-4} \), the integration time \( T = 10^6 \) and averaging over \( 10^4 \) realizations was performed for each \( L \); for smaller \( \epsilon < 10^{-4} \), when the probability \( P_L(\epsilon) \) is also very small, more
than $4 \cdot 10^4$ realizations were required to accumulate at least $N = 20$ events for each data point. Then $1/\sqrt{N}$ was assumed to give the relative uncertainty of the obtained value $w(\epsilon)$. Also, for $\epsilon < 10^{-4}$, the integration time had to be increased to $T = 3 \cdot 10^8$ in order to reliably distinguish between regular and chaotic dynamics (see also Appendix B).

The numerical results, as well as the dependence $w(\epsilon) = (1.37 \cdot 10^3) \epsilon^2$, obtained from Eq. (10), are shown in Fig. 2 by the symbols and the dashed line, respectively. Fig. 2 represents the main result of the present work. Starting from $L = 16$, the collapse of the numerical data is very good, which numerically proves the local origin of chaos. At $\epsilon < 3 \cdot 10^{-4}$ the numerical data fall on the dashed line. At larger $\epsilon$ the data collapses on some intermediate asymptotics for $w(\epsilon)$, which can be approximated by a power law $w(\epsilon) = B \epsilon^2$ with the exponent $\beta = 2.85 \pm 0.1$ and a surprisingly large prefactor $B = (1 - 5) \cdot 10^8$. The precise mechanism responsible for this intermediate asymptotics is not clear at the moment. This mechanism is likely to involve a larger number of oscillators (more than three), since (i) it is in this region of $\epsilon$ that the symbols for short chains ($L = 8$) exhibit a systematic deviation from those for long chains in Fig. 2 in the direction of suppression of the intermediate asymptotics; (ii) for a single-site initial excitation (see Fig. 3 below) the intermediate asymptotics is completely absent, indicating that it requires a certain initial spread.

In this context, one should mention the result reported in Ref. [22], where the probability of subdiffusive spreading of an initially localized wave packet was studied for exactly the same model, Eq. (1). Namely, it was argued that an initially localized wave packet either stays localized forever or spreads indefinitely, and the probability of spreading was estimated to be proportional to $\epsilon$ at small $\epsilon$. In the present work, the probability of chaos, was shown to scale as $\epsilon^2$. Since the latter is smaller than the former, the combination of the two results would suggest that the wave packet spreading does not require chaos, which is quite counter-intuitive. However, in the author’s opinion, it is more plausible that the probability of spreading was strongly overestimated in Ref. [22], as discussed below.

First, let us exclude the trivial possibility that the difference between the two results is simply due to the fact that initial conditions studied in Ref. [22] correspond to excitation localized on a single site, while a finite density was assumed in the present work. Indeed, for a finite density the trajectory is chaotic when a resonant triple occurs anywhere on the chain, while for a single-site excitation the condition is simply that the resonant triple includes the excited site. This affects the scaling with $L$, but not with $\epsilon$: the probability of chaos for a single-site excitation scales as $P_0(\epsilon) = A_1(\epsilon) \epsilon^2$ and is independent of $L$. The resonant-triple calculation gives $A_1 \approx 235$ (see Appendix C), which is also confirmed by the direct numerical sampling, as shown in Fig. 3 (Note also the direct crossover from 1 to $A_1 \epsilon^2$ without any intermediate asymptotics).

The analytical arguments of Ref. [22] are based on the assumption that a single resonance is sufficient for spreading. This immediately gives the probability scal-
distribution of $Q = 1/(P - 1)$ [see Eq. (12)] from regular and chaotic trajectories (upper and lower curves, respectively) at times $t = 10^3, 10^6$ [thick red (gray) and thin black curves, respectively] for two chain lengths $L = 11, 21$ (dotted and solid curves, respectively) for $\epsilon = 0.01$.

However, it is not clear why a single resonance should lead to unlimited spreading; indeed, an isolated nonlinear resonance is known to produce just periodic oscillations \cite{14,15}.

The numerical procedure of Ref. \cite{22} was based on the analysis at small $\epsilon$ of the trajectories. Here we omitted oscillating terms, which at large $t > 1$ quickly vanish upon disorder averaging, as well as terms which do not contain potentially small denominators and thus are bounded by $O(\epsilon)$ such as $3\epsilon/(4\omega_n^3)$, originating from the $(q_{n+1} - q_n)^4$ contribution to $E_n^\perp$.

From Eq. (15), one can find the probability of $P > 1.2$ at small $\epsilon$ by noting that $P > 1.2$ may occur when either $\omega_n+1$ or $\omega_n−1$ is close to $\omega_n$. The probability of both occurring simultaneously can be neglected, as well as the probability of $\omega_{n+1}$ being close to $3\omega_n$, as these probabilities are $\propto \epsilon^2$. Then, the probability of $P > 1.2$ is given by

$$Pr\{P > 1.2\} \approx \sum_{\pm} Pr\left\{\frac{9\epsilon^2}{4\omega_n^6} \frac{\omega_n^2 + \omega_{n+1}^2}{(\omega_{n+1} - \omega_n)^2} > 0.2\right\} \approx 2 Pr\left\{|\omega_{n+1} - \omega_n^2| < \frac{3\epsilon}{\sqrt{2}\omega_n^3}\right\} \approx 3\sqrt{8}\epsilon \int_{0.2}^{\omega_n^3} \frac{d\omega_n^2}{\omega_n^2} \approx 20.8\epsilon.$$

The coefficient in front of $\epsilon$ is very close to that obtained numerically for the probability of spreading (see inset of Fig. 3 of Ref. \cite{22}).
The fact that Eq. (16), derived under the assumption of validity of perturbation theory at arbitrarily long times (i.e., that the trajectory is localized), gives practically the same result as the numerics of Ref. [22], means that almost all trajectories counted as spreading in Ref. [22], were in fact localized. Thus, the question of whether the probability of unlimited wave packet spreading scales as \( \propto \epsilon \), or \( \propto \epsilon^2 \), or is identically zero, remains open.

V. CONCLUSIONS

In this work, the probability for a long chain of oscillators, defined by its Hamiltonian, Eq. (1), to be on a chaotic trajectory, was calculated by analyzing the chaotic phase space of rare resonant configurations of three oscillators. The result, \( 1 - e^{-A_\epsilon^2 L} \), agrees with the direct numerical evaluation of Lyapunov exponents for many disorder realizations and initial conditions at sufficiently small values of \( \epsilon \), the energy per oscillator. While the coefficient \( A \) depends on the specific ensemble used for the initial conditions, the power law \( \epsilon^2 \) is determined by the mere fact that chaotic phase volume is dominated by the regions of the phase space where two resonant conditions for three neighboring oscillators are satisfied simultaneously. This fact was numerically verified for the three-oscillator configuration, see Fig. 1.

The following main conclusions can be drawn from the results of present work. (i) The local nature of chaos, established earlier for the discrete nonlinear Schrödinger equation with disorder (DNLS) in Refs. [19, 20] for weak nonlinearity, is further confirmed here for a different model. Namely, the probability of regular dynamics decays with length \( L \) as \( e^{-wL} \), so \( w \) can be called the probability per unit length for the chaos to occur. (ii) Two different ways to calculate this probability, namely, from the analysis of the phase space of a resonant triple, and by direct numerical sampling of disorder realizations, give the same result, but this agreement sets in at unexpectedly small values of \( \epsilon \sim 3 \times 10^{-4} \). At larger \( \epsilon \), an intermediate asymptotics is seen in the numerical results. This fact may be relevant for understanding the disagreement between the results of Refs. [19] and [20] for the chaotic probability in the DNLS; indeed, one cannot exclude existence of similar intermediate regimes in DNLS. However, the results of the present work were obtained for a quite different model with a different number of independent parameters, so no quantitative comparison can be made with the DNLS, and the latter should be studied further.

VI. ACKNOWLEDGEMENTS

The author is grateful to M. Ivanchenko, T. Laptyeva, and S. Flach for clarifying discussions of Ref. [22], and to S. Fishman for a critical reading of the manuscript.

Appendix A: Few-oscillator configurations

It is convenient to pass to the canonical action-angle variables, \((p_n, q_n) \rightarrow (I_n, \phi_n)\):

\[
p_n = \sqrt{2\omega_n I_n} \cos \phi_n, \quad q_n = \sqrt{2I_n \omega_n} \sin \phi_n. \quad \text{(A1)}
\]

Hamiltonian (1), averaged over the fast oscillations, (that is, with terms depending on \( \phi_n + \phi_{n+1} \) omitted) takes the following form:

\[
H = \sum_{n=1}^{L} \omega_n I_n + \frac{3}{8} \sum_{n=1}^{L-1} \left[ \frac{I_n}{\omega_n} + \frac{I_{n+1}}{\omega_{n+1}} - 2 \sqrt{\frac{I_n I_{n+1}}{\omega_n \omega_{n+1}}} \cos(\phi_n - \phi_{n+1}) \right]^2. \quad \text{(A2)}
\]

1. Two oscillators

Let \( \omega_n \approx \omega_{n+1} \) be strongly different from \( \omega_{n-1}, \omega_{n+2} \). We perform a canonical change of variables of the pair

\[
I = I_n + I_{n+1}, \quad \phi = \frac{\phi_n + \phi_{n+1}}{2},
\]

\[
J = \frac{I_n - I_{n+1}}{2}, \quad \varphi = \phi_n - \phi_{n+1}. \quad \text{(A3)}
\]
and further denote \( J = (I/2) \cos \vartheta \). Then the Hamiltonian of the pair is given by

\[
H = \frac{\omega_1^2 + \omega_2^2}{2} I + (\omega_1^2 - \omega_2^2)(I/2) \cos \vartheta + \\
\frac{3}{4} \frac{I^2}{\omega_1 \omega_2} \left( \omega_1^2, \omega_2^2 \right) \cos \vartheta + \\
\frac{1}{2} \left( \omega_1 \omega_2 \right)^{3/2} \left[ \omega_1 + \omega_2 - (\omega_1 - \omega_2) \cos \vartheta \right] \cos \varphi + \\
\frac{3}{4} \frac{I^2}{\omega_1 \omega_2} \sin^2 \vartheta \cos 2\varphi,
\]

(A4)

where \( \omega_1' = \omega_n + 4I_{n-1}/(\omega_n \omega_{n-1}) \), \( \omega_2' = \omega_{n+1} + 4I_{n+2}/(\omega_{n+1} \omega_{n+2}) \). We can assume \( I_{n-1}, I_{n+2} \) to be constant, as their changes due to weak non-resonant perturbation are small. This Hamiltonian can be approximately rewritten as

\[
H \approx \frac{3}{16} \frac{I^2}{\omega}\left( \Omega \cos \vartheta + 4 \sin \vartheta \cos \varphi + \sin^2 \vartheta \cos 2\varphi \right) + \\
+ \text{const},
\]

(A5a)

\[
\Omega = \frac{8 \omega^2}{3}, \quad \bar{\omega} = \frac{\omega_1 + \omega_2}{2}.
\]

(A5b)

This Hamiltonian has only two elliptic stationary points in the phase space \((\vartheta, \varphi)\) at any value of the rescaled mismatch \( \Omega \). Hence, its phase space does not contain a separatrix at all, so we do not expect any chaos at small \( \epsilon \).

2. Three oscillators

As we are interested in the properties of a resonant triple \( n, n+1, n+2 \) far from the ends of the chain, let us denote the three frequencies by

\[
\omega_1' = \omega_n + \frac{4I_{n-1}}{\omega_n \omega_{n-1}}, \\
\omega_2' = \omega_{n+1}, \\
\omega_3' = \omega_{n+2} + \frac{4I_{n+3}}{\omega_{n+2} \omega_{n+3}},
\]

and assume \( I_{n-1} \) and \( I_{n+3} \) to be constant, since their variation is small in the parameter \( \epsilon \ll 1 \). Moreover, we can also neglect the modification in the probability distributions of \( \omega_1', \omega_3' \) with respect to that of \( \omega_n \), as the difference is again \( \sim \epsilon \). The most important contribution to the chaotic fraction will come from the region where the differences \( \omega_1' - \omega_2', \omega_2' - \omega_3' \), are small compared to the frequencies themselves. Having this in mind, let us further denote

\[
\bar{\omega} = \frac{\omega_1' + \omega_2' + \omega_3'}{3}.
\]

Introducing the complex canonical coordinates \( \psi_i = \sqrt{I_n} e^{-i\varphi_n} \), \( \psi_2 = \sqrt{I_{n+1}} e^{-i\varphi_{n+1}} \), \( \psi_3 = \sqrt{I_{n+2}} e^{-i\varphi_{n+2}} \), whose conjugate momenta are \( i\psi_i^* \), \( i\psi_2^* \), \( i\psi_3^* \), respectively, we can write the Hamiltonian of the triple as

\[
H = \sum_{i=1}^{3} \omega_i' \psi_i^* \psi_i + \frac{3}{8} \psi_1^* \psi_1 \psi_2^* \psi_2 + \frac{3}{8} \psi_3^* \psi_3 \psi_4^* \psi_4 \\
+ \frac{3}{8} \sum_{i=1}^{2} \left( \frac{\psi_i^*}{\sqrt{\omega_i'}} - \frac{\psi_{i+1}^*}{\sqrt{\omega_{i+1}'}} \right)^2 \left( \frac{\psi_i}{\sqrt{\omega_i'}} - \frac{\psi_{i+1}}{\sqrt{\omega_{i+1}'}} \right)^2.
\]

(A8)

In the nonlinear terms (those of the fourth order in \( \psi_i, \psi_i^* \)) it is sufficient to replace \( \omega_1', \omega_2', \omega_3' \), to the leading order in \( \epsilon \). The total norm,

\[
I = \psi_1^* \psi_1 + \psi_2^* \psi_2 + \psi_3^* \psi_3,
\]

(A9)
is conserved for the Hamiltonian \( \{A8\} \), and the corresponding conjugate phase,

\[
\phi = \frac{\phi_n + \phi_{n+1} + \phi_{n+2}}{3},
\]

(A10)
even though depends on time in some complicated way, does not contribute to anything. In the fixed-\( \epsilon \) ensemble, we simply have \( I = 3\epsilon \bar{\omega} + O(\epsilon^2) \). It is convenient to pass to rescaled variables

\[
\psi_i = \sqrt{\frac{I}{\omega}} \psi_i, \quad i = 1, 2, 3,
\]

\[
\Omega = \frac{4}{3} \frac{\bar{\omega}^2}{\bar{\omega}_3'}, \quad \bar{\omega}_3' = \frac{2 \omega_1' - \omega_2' - \omega_3'}{3},
\]

\[
\Omega' = \frac{4}{3} \frac{\bar{\omega}^2}{\bar{\omega}_3'}.
\]

(A11)

Then \( |\psi_i|^2 + |\psi_2|^2 + |\psi_3|^2 = 1 \). The probability density for frequencies is

\[
\nu(\bar{\omega}) = \lim_{\Omega, \Omega' \to 0} \int_{1/2}^{3/2} dx \ dy \ dz \ \delta \left( \sqrt{x} + \sqrt{y} + \sqrt{z} - \bar{\omega} \right) \\
\times \delta \left( \frac{2\sqrt{x} - \sqrt{y} - \sqrt{z}}{9/4} - \Omega \right) \\
\times \delta \left( \frac{2\sqrt{y} - \sqrt{x} - \sqrt{z}}{9/4} - \Omega' \right) \\
= \frac{27}{2} \frac{I^2}{\omega} + O(\epsilon^2).
\]

(A12)

The rescaled Hamiltonian is given by

\[
\frac{H - \bar{\omega} I}{(3/4) \bar{\omega} \omega^2 - 2} = \Omega|\psi_1|^2 + (\Omega + \Omega')|\psi_2|^2 + \Omega'|\psi_3|^2 + \\
\frac{1}{2} |\psi_1|^4 + \frac{1}{2} |\psi_3|^4 + \\
\frac{1}{2} |\psi_1 - \psi_2|^4 + \frac{1}{2} |\psi_2 - \psi_3|^4.
\]

(A13)
For the thermal distribution the initial condition should be taken in the form
\[
\Psi_1 = \sqrt{1-x} e^{i(-2\varphi_1+\varphi_2)/3}, \\
\Psi_2 = \sqrt{x-y} e^{i(\varphi_1+\varphi_2)/3}, \\
\Psi_3 = \sqrt{y} e^{i(\varphi_1-2\varphi_2)/3}.
\] (A14)

Then the chaotic fraction is given by
\[
w(\epsilon) = \int_0^\infty e^{-\omega_1/\epsilon} \frac{I^2 dI}{(\epsilon/\omega)^3} \int \frac{27}{2} \frac{I^2}{\omega} d\omega \int d\Omega d\Omega' \times
\]
\[
\times \int_0^{2\pi} \frac{d\varphi_1}{2\pi} \frac{d\varphi_2}{2\pi} \int_0^1 dx \int_0^\infty dy \Theta(x, y; \varphi_1, \varphi_2; \Omega, \Omega').
\] (A15)

The first two integrals amount to 216 \epsilon^2. The fixed-\epsilon ensemble corresponds to fixed \( x = 2/3, y = 1/3 \), so the chaotic fraction is given by
\[
w(\epsilon) = \int_0^{\sqrt{3}/2} \frac{27}{2} \frac{(3\epsilon)^2}{\omega^3} d\omega \int d\Omega d\Omega' \times
\]
\[
\times \int_0^{2\pi} \frac{d\varphi_1}{2\pi} \frac{d\varphi_2}{2\pi} \Theta(2/3, 1/3; \varphi_1, \varphi_2; \Omega, \Omega').
\] (A16)

The first integral is equal to 81 \epsilon^2, and the integral over \( \Omega, \Omega' \) to 16.9 \pm 0.2.

One can separate the conserved total norm and write the explicit Hamiltonian for the two remaining degrees of freedom. Let us label the actions and phases of the oscillators \( n, n+1, n+2 \) by \( i = 1, 2, 3 \), respectively. Then the corresponding canonical transformation is determined by
\[
I = I_1 + I_2 + I_3, \quad \phi = \frac{\phi_1 + \phi_2 + \phi_3}{3}, \\
\frac{1}{3} J_1 = \frac{2I_1 - I_2 - I_3}{3}, \quad \varphi_1 = \phi_1 - \phi_2, \\
\frac{1}{3} J_2 = \frac{I_1 + I_2 - 2I_3}{3}, \quad \varphi_2 = \phi_2 - \phi_3.
\] (A17)

To ensure \( I_1, I_2, I_3 > 0 \), the variables \( J_1, J_2 \) should lie inside the triangle, shown in Fig. 5. The inverse transformation is
\[
\psi_1 = \sqrt{(1 + J_1)I/3} e^{i(-2\varphi_1+\varphi_2)/3-i\phi}, \\
\psi_2 = \sqrt{(1 - J_1 - J_2)I/3} e^{i(\varphi_1+\varphi_2)/3-i\phi}, \\
\psi_3 = \sqrt{(1 + J_2)I/3} e^{i(\varphi_1-2\varphi_2)/3-i\phi}.
\] (A18)

The rescaled Hamiltonian for the degrees of freedom
\[
\frac{H - \bar{\omega}I}{(I/2\omega)^2} = (2\Omega + \Omega' - 2/3)J_1 + (\Omega + 2\Omega' - 2/3)J_2
\]
\[
- \frac{2}{3} \sqrt{(1 + J_1)(1 - J_1 - J_2)(2 - J_2)} \cos \varphi_1
\]
\[
+ \frac{1}{3} (1 + J_1)(1 - J_1 - J_2) \cos 2\varphi_1
\]
\[
- \frac{2}{3} \sqrt{(1 - J_2)(1 - J_1 + J_2)(2 - J_1)} \cos \varphi_2
\]
\[
+ \frac{1}{3} (1 - J_2)(1 - J_1 + J_2) \cos 2\varphi_2
\]
\[
+ \frac{7}{3} - 2J_1J_2.
\] (A19)

Still, due to the presence of trigonometric functions, the numerical integration of the equations of motion for this Hamiltonian would be less efficient than for the polynomial Hamiltonian \( \mathcal{H} \) with three degrees of freedom.

**Appendix B: Numerical criterion for chaos**

The numerical criterion for the chaotic motion used in the present work is based on the analysis of \( \Lambda(t) = \ln \frac{d(t)}{d(0)} \), where \( d(t) \) is the distance between points belonging to two initially close trajectories. To illustrate the idea, we plot several traces \( \Lambda(t) \) in Fig. 6. Looking at them, it is quite easy to guess which curve corresponds to a regular motion, and which one to chaotic. The only exception is the lowest curve in Fig. 6(d). We know it should be chaotic, as the corresponding realization of
\( \{\omega^2_n\} \) contains two neighboring resonances, introduced “by hands”. However, at \( \epsilon = 10^{-5} \) energy exchange between the oscillators is so slow, that following their dynamics up to \( T = 10^6 \) is insufficient to distinguish between regular and chaotic motion. This gives us the lower boundary on \( \epsilon \). Compare also the lowest curve in Fig. 6(a) and the middle curve in Fig. 6(d). The former one is regular and the latter is chaotic, even though they have close values of \( \Lambda(t = T) \). Thus, the fixed cutoff for the Lyapunov exponent (independent of \( L \) and \( \epsilon \)), which was used in Ref. [20], as the criterion for chaos, may introduce systematic errors and affect the scaling.

The main observation from Fig. 6 is that for chaotic trajectories \( \Lambda(t) \) grows linearly on the average, corresponding to a finite Lyapunov exponent, while for regular trajectories

\[
\Lambda(t) = \ln t + q_0, \quad (B1)
\]

up to some weak noise, \( q_0 \) being some realization-dependent constant. This can be understood very simply: if two trajectories lie on neighboring invariant tori, their frequencies are only slightly different, so the phase mismatch between them is accumulated slowly and linearly in time. The distance between the trajectories is proportional to this phase mismatch, as long as the latter is small compared to unity.

To determine how well a given function \( \Lambda(t) \) is approximated by the form \( (B1) \) on an interval \( 0 < t < T \), consider the time average (denoted by overline):

\[
[\Lambda(t) - \ln t - q]^2. \quad (B2)
\]

measures the quality of the fit of \( \Lambda(t) \) by the expression \( (B1) \). \( D[\Lambda(t)] \) for all curves in Fig. 6 is shown on
the corresponding panel. Thus, we consider the motion as regular if \(|D| < 1\) and chaotic otherwise.

When the Lyapunov exponent \(\sigma > 0\), it sets the natural upper time limit for the numerical integration of the differential equations. Indeed, when the double-precision machine zero, \(2^{-52}\), multiplied by \(e^{\sigma t}\) becomes of the order of unity, the integration inevitably deviates from the original trajectory, no matter how good the integration scheme is. This corresponds to \(A = \ln 2^{52} \approx 36\). To check this, we have performed the time-reversal test: at time \(t = t_*\), we invert all the momenta, and integrate up to \(t = 2t_*\). Ideally, the system should return to its initial condition. This was found to be the case if \(A(t_*) < 25 - 30\) (for many trajectories whose Lyapunov exponents differed by one or two orders of magnitude).

Thus, considering larger \(A\) would produce the Lyapunov exponent not for a given trajectory, but its certain average over the phase space. As we are interested not in the trajectory itself, but only in whether it is chaotic or regular, we consider the very fact that \(A(t_*)\) has reached 25 a sufficient evidence for chaos. If this happens within the integration time, \(t < T\), we stop the integration, and simply set \(D|A| = 30\).

One could think that for every realization of the disorder \(\{\omega_n^2\}\) and of the oscillator phases \(\{\phi_n\}\) exists a threshold value \(\epsilon_c\), such that for \(\epsilon < \epsilon_c\) the motion is quasiperiodic, and for \(\epsilon > \epsilon_c\) it is chaotic. Indeed, a natural guess is that upon increase of the integrability-breaking parameter \(\epsilon\) the chaotic region of the phase space should grow. This would make the calculation for the fixed-\(\epsilon\) ensemble more efficient, as one would not need to probe too small or too large values of \(\epsilon\). However, this guess turns out to be wrong, as seen from Fig. \(\text{B.1}\) where we plot \(\ln|D|A|\) as a function of \(\epsilon\) in the vicinity of \(\epsilon = 10^{-3}\) for the same realization of disorder and phases as in Fig. \(\text{B.1}\). The flat regions on the level of \(3.4 \ldots = \ln 30\) correspond to \(A(t)\) reaching 25 for \(t < T\), as discussed in the previous paragraph. One can see that the conclusion about regular/chaotic character of the dynamics is little sensitive to the chosen numerical border \(D = 1\): had we chosen \(D = 5\) or \(D = 1/5\), the result would not change. We also note that for the particular realization, corresponding to Fig. \(\text{B.1}\) in all intervals of \(\epsilon\) where the dynamics is chaotic, the Lyapunov eigenvector is confined to sites from \(n = 21\) to \(n = 26\). This is clearly related to the fact that in this realization \(|\omega_{22} - \omega_{21}| \approx 0.9 \cdot 10^{-3}\), so the observed reentrant behavior of chaos seems to occur for the same guiding resonance.

To check how the results are affected by the choice of the integration time \(T\), we take the same realization of disorder as for the trajectories in Fig. \(\text{B.2}\) (the chain length \(L = 32\)), and set “by hands” three nearby frequencies to be

\[
\omega_{12} = \bar{\omega} + \frac{9}{4} \epsilon \frac{\Omega}{\omega^3}, \quad \omega_{13} = \bar{\omega} - \frac{9}{4} \epsilon \frac{\epsilon}{\omega^3} \Omega, \quad \omega_{14} = \bar{\omega},
\]

with a fixed \(\bar{\omega} = \sqrt{0.7}\), two values of \(\epsilon = 10^{-4}, 0.3 \cdot 10^{-4}\), and \(\Omega\) varying in the interval \(-2 < \Omega < 6\). This chain is supposed to be well described by the effective Hamiltonian in Eq. \(\text{B.6}\) with \(\Omega' = 0\). On the two panels of Fig. \(\text{B.2}\) we plot the phase-averaged chaotic fraction as a function of \(\Omega\) for the two above-mentioned values of \(\epsilon\), and for \(T = 10^6, 3 \cdot 10^6, 5 \cdot 10^6\), together with the result of the calculation using Eq. \(\text{B.6}\) (the section of Fig. \(\text{B.1}\) along the line \(\Omega' = 0\)). For both values of \(\epsilon\), as \(T\) is increased, the curves approach the result of Eq. \(\text{B.6}\) up to an overall horizontal shift. This horizontal shift appears because Eq. \(\text{B.6}\) neglects the nonlinear frequency shift for the oscillators with \(n = 12\) and 14 due to their coupling to the oscillators with \(n = 11\) and 15, respectively (the difference between primed and unprimed frequencies in Eq. \(\text{A.10}\)). This is also the reason why the dip at \(\Omega = 0\) for the result of Eq. \(\text{B.6}\) is not resolved on the curves for the long chain. What is important, is that while at \(\epsilon = 10^{-4}\) the integration time \(T = 10^6\) is sufficient, for \(\epsilon = 0.3 \cdot 10^{-4}\) integration up to \(T = 10^6\) clearly underestimates the chaotic fraction, so that \(T = 3 \cdot 10^6\) is required.

Finally, we mention that the numerical integration of the differential equations was performed using (i) the fourth-order Runge-Kutta method for the calculations reported in this section, and (ii) Bulirsch-Stoer method with polynomial extrapolation for the accumulation of statistics. The latter method turns out to require about ten times less CPU time that the former to reach the same level of accuracy. The energy conservation was satisfied extremely well in all calculated trajectories, including those with largest Lyapunov exponents, even at times when the time-reversal test had failed. Thus, it is unlikely that use of an integration scheme which automatically respects some conservation laws of the original equations (e. g., a symplectic integrator which preserves the phase space volume) would lead to more accurate results for a given trajectory: the accuracy is lost primarily in the directions, orthogonal to conservation laws.
Appendix C: Single-site excitation

Let us determine the probability of chaos for an initial condition, localized on a single site,

\[ p_n(0) = 0, \quad q_n(0) = \delta_{n,(L+1)/2} \sqrt{2\epsilon/\omega_n}, \quad (C1) \]

for \( L \) odd. Chaos should occur if the resonant triple contains the site \( n = (L + 1)/2 \). Thus, the probability of chaos \( P_1(\epsilon, L) \) should not depend on \( L \) and scale as \( P_1(\epsilon) = A_1 \epsilon^2 \), at large \( L \) and small \( \epsilon \).

To determine the coefficient \( A_1 \), we again use the effective Hamiltonian of the triple, Eq. (7), and consider two initial conditions:

\[ \Psi_1(0) = \frac{1}{\sqrt{3}}, \quad \Psi_2(0) = \Psi_3(0) = 0, \quad (C2a) \]
\[ \Psi_1(0) = 0, \quad \Psi_2(0) = \frac{1}{\sqrt{3}}, \quad \Psi_3(0) = 0. \quad (C2b) \]

i. e., the excitation to be initially localized on one of the lateral sites of the triple, or on the central site. Let \( \Theta_{1,2}(\Omega, \Omega') \) equal one if the corresponding trajectory is chaotic and zero if it is regular, for each of the two initial conditions, respectively. The functions \( \Theta_{1,2}(\Omega, \Omega') \) are shown in Fig. 9. The chaotic probability is then given by

\[ P_1(\epsilon) = 81 \epsilon^2 \int_{-\infty}^{\infty} d\Omega d\Omega' \left[ 2 \Theta_1(\Omega, \Omega') + \Theta_2(\Omega, \Omega') \right]. \quad (C3) \]

For the initial condition (C2a), the \( \Omega, \Omega' \)-integral is equal to 0.75\pm0.02, while for Eq. (C2b) it is equal to 1.39\pm0.02, which gives \( A_1 \approx 235 \).
FIG. 9. Values of $\Omega, \Omega'$ for which the single-site initial conditions for the resonant triple Hamiltonian, Eq. (7), lie on a chaotic trajectory (black spots) and on a regular trajectory (white areas): (a) and (b) represent the functions $\Theta_1(\Omega, \Omega')$ and $\Theta_2(\Omega, \Omega')$, corresponding to the initial conditions \((C2a), (C2b)\), respectively.

[17] J. Fröhlich, T. Spencer, and C. E. Wayne, J. Stat. Phys. **42**, 247 (1986).

[18] M. Johansson, G. Kopidakis, and S. Aubry, Europhys. Lett. **91**, 50001 (2010).

[19] D. M. Basko, Ann. Phys. **326**, 1577 (2011).

[20] A. Pikovsky and S. Fishman, Phys. Rev. E **83**, 025201(R) (2011).

[21] V. Oganessian, A. Pal, and D. A. Huse, Phys. Rev. B **80**, 115104 (2009).

[22] M. V. Ivanchenko, T. V. Laptyeva, and S. Flach, Phys. Rev. Lett. **107**, 240602 (2011).

[23] Strictly speaking, $\epsilon_n$ does not correspond to the total energy per oscillator, as it does not include $H_{\text{int}}$, the second term of Eq. (1). In particular, this means that (i) the sum of all $\epsilon_n$, corresponding to $H_0$ is not conserved during the evolution, and (ii) Eq. (4) with the $\delta$-function replaced by the exponential is not exactly the thermal distribution. However, $H_0 \sim \epsilon$ and $H_{\text{int}} \sim \epsilon^2$, so $H_{\text{int}}$ represents a small correction at $\epsilon \to 0$, and $\epsilon$ indeed has the meaning of temperature in the thermal ensemble.

[24] This is because the system (1) has only one integral of motion (energy), which should be contrasted with DNLS which has two integrals of motion (energy and norm), so the minimal number of oscillators needed to produce chaos in DNLS is three.

[25] M. Mulansky, K. Ahnert, A. Pikovsky and D. L. Shepelyansky, J. Stat. Phys. **145**, 1256 (2011).