Chapter 1

Chaos in Statistical Physics

Rainer Klages

Queen Mary University of London
School of Mathematical Sciences
Mile End Road, London E1 4NS

Max Planck Institute for the Physics of Complex Systems
Nöthnitzer Str. 38, D-01187 Dresden, Germany
e-mail: r.klages@qmul.ac.uk

This chapter introduces to *chaos in dynamical systems* and how this theory can be applied to derive fundamental laws of *statistical physics* from first principles. We first elaborate on the concept of deterministic chaos by defining and calculating Ljapunov exponents and dynamical entropies as fundamental quantities characterizing chaos. These quantities are shown to be related to each other by Pesin's Theorem. Considering open systems where particles can escape from a set asks for a generalization of this theorem which involves fractals, whose properties we briefly describe. We then cross-link this theory to statistical physics by discussing simple random walks on the line, their characterization in terms of diffusivion, and the relation to elementary concepts of Brownian motion. This sets the scene for considering the problem of chaotic diffusion. Here we derive a formula exactly expressing diffusion in terms of the chaos quantities mentioned above.

1. Introduction

This book chapter introduces to *chaos in dynamical systems* and how this theory can be applied to derive fundamental laws of *statistical physics* from first principles. Very intuitively, one may say that the path of a point particle generated by a dynamical system looks "chaotic" if it displays "random-looking" evolution in time and space. A mathematically rigorous definition of chaos was given in Ref.[1] Therein it was discussed that even

very simple one-dimensional maps can exhibit chaotic dynamics. Surprisingly, abstract low-dimensional chaotic dynamics bears similarities to the dynamics of interacting physical many-particle systems. This is the key point that we explore in this chapter.

Over the past few decades it was found that famous statistical physical laws like Ohm's Law for electric conduction, Fourier's Law for heat conduction, and Fick's Law for the diffusive spreading of particles, which a long time ago were formulated phenomenologically, can be derived *from first principles* in chaotic dynamical systems. This sheds new light on the rigorous mathematical foundations of *Nonequlibrium Statistical Physics*, which is the theory of the dynamics of many-particle systems under external gradients or fields. The external forces induce *transport* of physical quantities like charge, energy, or matter. The goal of nonequilibrium statistical physics is to derive macroscopic statistical laws describing such transport starting from the microscopic dynamics for the single parts of many-particle systems. While the conventional theory puts in randomness "by hand" by using probabilistic, or stochastic, equations of motion like random walks or stochastic differential equations, recent developments in the theory of dynamical systems enable to do such derivations starting from deterministic equations of motion. Determinism means that no random variables are involved, rather, randomness is generated by chaos in the underlying dynamics. This is the field of research that will be introduced by this chapter.

This theory also illustrates the emergence of *complexity* in systems under nonequilibrium conditions: Due to the microscopic nonlinear interaction of the single parts in a complex many-particle system novel, non-trivial dynamics, in this case exemplified by universal transport laws, may emerge on macroscopic scales. The dynamics of a complex system, as a whole, is thus different from the sum of its single parts. In the very simplest case, this idea is illustrated by the interaction of a point particle with a "scatterer", where the latter is modeled by a one-dimensional map. This is our vehicle of demonstration in the following, because this simple model can be solved rigorously analytically.

Our chapter consists of two sections: In Sec. 2 we introduce to two important quantities assessing chaos in dynamical systems, Ljapunov exponents and dynamical entropies. The former are widely used in the applied sciences to test whether a given system is chaotic, the latter is motivated by information theory. Cross-links to *ergodic theory* by defining these quantities are explored, which is a core discipline in mathematical dynamical

systems theory. Interestingly, both these different quantities are exactly related to each other by *Pesin's Theorem*. Considering *open system* where particles can escape from a set generates *fractals*, a concept that we will introduce as well.

The latter problem cross-links to Sec. 3, which explores diffusion in chaotic dynamical systems. After briefly introducing to the statistical physical problem of diffusion we outline a rigorous theory that enables one to calculate diffusion coefficients characterizing the spreading of particles from first principles. By combining this approach with a key result for open systems from the previous section we arrive at an exact formula expressing the diffusion coefficient in terms of the two quantities characterizing deterministic chaos introduced before. This important result forms the highlight of our exposition and concludes our chapter.

While Sec. 2 mainly elaborates on textbook material of chaotic dynamical systems,[2–4] Sec. 3 introduces to advanced topics that emerged in research over the past twenty years.[5–7]

2. Deterministic chaos

2.1. *Dynamics of simple maps*

Let us recall the following:

Definition 1. Let $J \subseteq \mathbb{R}, x_n \in J, n \in \mathbb{Z}$. Then

$$F : J \to J \quad , \quad x_{n+1} = F(x_n) \tag{1}$$

is called a *one-dimensional time-discrete map*. $x_{n+1} = F(x_n)$ are sometimes called the *equations of motion* of the dynamical system.

Choosing the initial condition x_0 *determines* the outcome after n discrete time steps, hence we speak of a *deterministic dynamical system*. It works as follows:

$$\begin{aligned} x_1 &= F(x_0) = F^1(x_0), \\ x_2 &= F(x_1) = F(F(x_0)) = F^2(x_0). \end{aligned}$$

$$\Rightarrow \; F^m(x_0) := \underbrace{F \circ F \circ \cdots F(x_0)}_{\text{m-fold composed map}} \quad . \tag{2}$$

In other words, there exists a unique solution to the equations of motion in form of $x_n = F(x_{n-1}) = \ldots = F^n(x_0)$, which is the counterpart of the

flow for time-continuous systems. We will focus on simple piecewise linear maps. The following one serves as a paradigmatic example:[2,3,5,8]

Example 1. The Bernoulli shift (also shift map, doubling map, dyadic transformation)

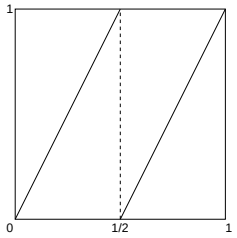


Fig. 1. The Bernoulli shift.

The map shown in Fig. 1 is defined by

$$B : [0,1) \to [0,1) \ , \ B(x) := 2x \bmod 1 \ = \begin{cases} 2x \ , \ 0 \leq x < 1/2 \\ 2x-1 \ , \ 1/2 \leq x < 1 \end{cases} \ . \quad (3)$$

The dynamics of this map can be understood follows, see Fig. 2: Assume we fill the whole unit interval with a uniform distribution of points. We may now decompose the action of the Bernoulli shift into two steps:

(1) The map *stretches* the whole distribution of points by a factor of two, which leads to *divergence* of nearby trajectories.
(2) Then we *cut* the resulting line segment in the middle due to the the modulo operation mod 1, which leads to motion *bounded* on the unit interval.

The Bernoulli shift thus yields a simple example for an essentially nonlinear stretch-and-cut mechanism, as it typically generates *deterministic chaos*.[3] The same mechanisms are encountered in more realistic dynamical systems. We remark that "stretch and fold" or "stretch, twist and fold" provide alternative mechanisms for generating chaotic behavior, see, e.g., the tent map mentioned in Ref.[1] The reader may wish to play around with these ideas in thought experiments, where the sets of points is replaced by kneading dough. These ideas can be made mathematically precise by what is called *mixing*, which is an important concept in the ergodic theory of dynamical systems.[5,9]

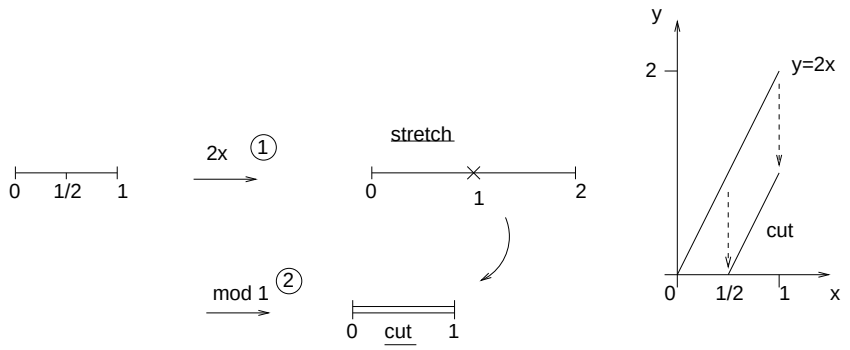


Fig. 2. Stretch-and-cut mechanism in the Bernoulli shift.

2.2. *Ljapunov chaos*

In Ref.[1] Devaney's definition of chaos was discussed, which requires that for a given dynamical system three conditions have to be fulfilled: sensitivity, existence of a dense orbit, and that the periodic points are dense. The *Ljapunov exponent* generalizes the concept of sensitivity in form of a quantity that can be calculated more conveniently, as we will motivate by an example:

Example 2. Ljapunov instability of the Bernoulli shift[3]

Consider two points that are initially displaced from each other by $\delta x_0 := |x_0' - x_0|$ with δx_0 "infinitesimally small" such that x_0, x_0' do not hit different branches of the Bernoulli shift $B(x)$ around $x = 1/2$. We then have

$$\delta x_n := |x_n' - x_n| = 2\delta x_{n-1} = 2^2 \delta x_{n-2} = \ldots = 2^n \delta x_0 = e^{n \ln 2} \delta x_0 \quad . \quad (4)$$

We see that there is an *exponential separation* between two nearby points as we follow their trajectories. The rate of separation $\lambda(x_0) := \ln 2$ is called the (local) *Ljapunov exponent* of the map $B(x)$.

This simple example can be generalized as follows, leading to the general definition of the Ljapunov exponent for one-dimensional maps F. Consider

$$\delta x_n = |x_n' - x_n| = |F^n(x_0') - F^n(x_0)| =: \delta x_0 e^{n\lambda(x_0)} \; (\delta x_0 \to 0) \quad (5)$$

for which we *presuppose* that an exponential separation of trajectories exists. By assuming that F is differentiable we rewrite this equation to

$$\begin{aligned}\lambda(x_0) &= \lim_{n\to\infty}\lim_{\delta x_0\to 0}\frac{1}{n}\ln\frac{\delta x_n}{\delta x_0}\\ &= \lim_{n\to\infty}\lim_{\delta x_0\to 0}\frac{1}{n}\ln\frac{|F^n(x_0+\delta x_0)-F^n(x_0)|}{\delta x_0}\\ &= \lim_{n\to\infty}\frac{1}{n}\ln\left|\frac{dF^n(x)}{dx}\right|_{x=x_0} \quad . \end{aligned} \tag{6}$$

Using the chain rule we obtain

$$\left.\frac{\mathrm{d}F^n(x)}{\mathrm{d}x}\right|_{x=x_0} = F'(x_{n-1})F'(x_{n-2})\dots F'(x_0) \quad , \tag{7}$$

which leads to

$$\begin{aligned}\lambda(x_0) &= \lim_{n\to\infty}\frac{1}{n}\ln\left|\prod_{i=0}^{n-1}F'(x_i)\right|\\ &= \lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\ln|F'(x_i)| \quad . \end{aligned} \tag{8}$$

This simple calculation motivates the following definition:

Definition 2.[8] Let $F\in C^1$ be a map of the real line. The *local Ljapunov exponent* $\lambda(x_0)$ is defined as

$$\lambda(x_0) := \lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\ln|F'(x_i)| \tag{9}$$

if this limit exists.

Example 3. For the Bernoulli shift $B(x) = 2x \mod 1$ we have $B'(x) = 2\,\forall x\in[0,1)$, $x\neq\frac{1}{2}$, hence trivially

$$\lambda(x) = \frac{1}{n}\sum_{k=0}^{n-1}\ln 2 = \ln 2 \tag{10}$$

at these points.

Note that Definition 2 defines the *local* Ljapunov exponent $\lambda(x_0)$, that is, this quantity may depend on our choice of initial conditions x_0. For the Bernoulli shift this is not the case, because this map has a uniform slope of two except at the point of discontinuity, which makes the calculation trivial. Generally the situation is more complicated. One question is of how

to calculate the local Ljapunov exponent, a second one to which extent it depends on initial conditions. An answer to both these questions is provided by the *global* Ljapunov exponent that we are going to introduce, which does not depend on initial conditions and thus characterizes the stability of the map as a whole.

It is introduced by observing that the local Ljapunov exponent in Definition 2 is defined by a *time average*, where n terms along the trajectory with initial condition x_0 are summed up by averaging over n. That this is not the only possibility to define an average quantity is clarified by the following definition. It requires the concepts of measure and density;[1] if the reader is not familiar with these objects, we recommend Ref.[10] as an introduction.

Definition 3. time and ensemble average[5,9]

Let μ^* be the invariant probability measure of a one-dimensional map F acting on $J \subseteq \mathbb{R}$. Let us consider a function $g : J \to \mathbb{R}$, which we may call an "observable". Then

$$\overline{g(x)} := \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} g(x_k) \quad , \tag{11}$$

$x = x_0$, is called the *time (or Birkhoff) average* of g with respect to F.

$$\langle g \rangle := \int_J d\mu^* g(x) \tag{12}$$

where, if such a measure exists, $d\mu^* = \rho^*(x)\, dx$, is called the *ensemble (or space) average* of g with respect to F. Here $\rho^*(x)$ is the *invariant density* of the map, and $d\mu^*$ is the associated *invariant measure*.[3] Note that $\overline{g(x)}$ may depend on x, whereas $\langle g \rangle$ does not.

If we choose $g(x) = \ln|F'(x)|$ as the observable in Eq. (11) we recover Definition 2 for the local Ljapunov exponent,

$$\lambda(x) := \overline{\ln|F'(x)|} = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \ln|F'(x_k)| \quad , \tag{13}$$

which we may write as $\lambda_t(x) = \lambda(x)$ in order to label it as a time average. If we choose the same observable for the ensemble average Eq. (12) we obtain

$$\lambda_e := \langle \ln|F'(x)| \rangle := \int_J dx \rho^*(x) \ln|F'(x)| \quad . \tag{14}$$

Example 4.

For the Bernoulli shift we have seen that for almost every $x \in [0,1)$ $\lambda_t = \ln 2$. For λ_e we obtain

$$\lambda_e = \int_0^1 dx \rho^*(x) \ln 2 = \ln 2 \quad , \tag{15}$$

taking into account that $\rho^*(x) = 1$.[1] In other words, time and ensemble average are the same for almost every x,

$$\lambda_t(x) = \lambda_e = \ln 2 \quad . \tag{16}$$

This motivates the following fundamental definition:

Definition 4. ergodicity [5,9]

A dynamical system is called *ergodic* if for every g on $J \subseteq \mathbb{R}$ satisfying $\int d\mu^* \, |g(x)| < \infty$

$$\overline{g(x)} = \langle g \rangle \tag{17}$$

for typical x.

For our purpose it suffices to think of a typical x as a point that is randomly drawn from the invariant density $\rho^*(x)$. This definition implies that for ergodic dynamical systems $\overline{g(x)}$ does not depend on x. That the time average is constant is sometimes also taken as a definition of ergodicity.[4,5] To prove that a given system is ergodic is typically a hard task and one of the fundamental problems in the ergodic theory of dynamical systems; see[5,9] for proofs of ergodicity in case of some simple examples. We remark that pure mathematicians define ergodicity in terms of indecomposability.[11]

On this basis, let us get back to Ljapunov exponents. For time average $\lambda_t(x)$ and ensemble average λ_e of the Bernoulli shift we have found that $\lambda_t(x) = \lambda_e = \ln 2$. Definition 4 now states that the first equality must hold whenever a map F is ergodic. This means, in turn, that for an ergodic dynamical system the Ljapunov exponent becomes a *global* quantity characterizing a given map F for a typical point x irrespective of what value we choose for the initial condition, $\lambda_t(x) = \lambda_e = \lambda$. This observation very much facilitates the calculation of λ, as is demonstrated by the following example:

Example 5.

Let us consider the map $A(x)$ displayed in Fig. 3 below:

From the figure we can infer that

$$A(x) := \begin{cases} \frac{3}{2}x \ , \ 0 \leq x < \frac{2}{3} \\ 3x - 2 \ , \ \frac{2}{3} \leq x < 1 \end{cases} \quad . \tag{18}$$

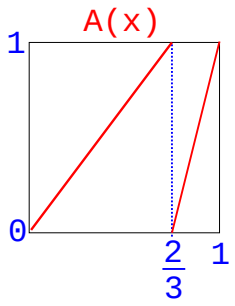


Fig. 3. A simple map for demonstrating the calculation of Ljapunov exponents via ensemble averages.

It is not hard to see that the invariant probability density of this map is uniform, $\rho^*(x) = 1$. The Ljapunov exponent λ for this map is then trivially calculated to

$$\lambda = \int_0^1 \mathrm{d}x \rho^*(x) \ln |A'(x)| = \ln 3 - \frac{2}{3} \ln 2 \quad . \tag{19}$$

By assuming that map A is ergodic (which here is the case), we can conclude that this result for λ represents the value for typical points in the domain of A.

In other words, for an ergodic map the global Ljapunov exponent λ yields a number that assesses whether it is chaotic in the sense of exhibiting an exponential dynamical instability. This motivates the following definition of deterministic chaos:

Definition 5. Chaos in the sense of Ljapunov[3,4,8]

An ergodic map $F : J \to J$, $J \subseteq \mathbb{R}$, F (piecewise) C^1 is said to be *L-chaotic* on J if $\lambda > 0$.

Why did we introduce a definition of chaos that is different from Devaney's definition mentioned earlier? One reason is that often the largest Ljapunov exponent of a dynamical system is easier to calculate than checking for sensitivity. Furthermore, the magnitude of the positive Ljapunov exponent quantifies the strength of chaos. This is the reason why in the applied sciences "chaos in the sense of Ljapunov" became a very popular concept. Note that there is no unique quantifier of deterministic chaos. Many different definitions are available highlighting different aspects of "chaotic behavior", all having their advantages and disadvantages. The

detailed relations between them are usually non-trivial and a topic of ongoing research. We will encounter yet another definition of chaos in the following section.

2.3. *Entropies*

Let us start with a brief motivation outlining the basic idea of entropy production in dynamical systems. Consider again the Bernoulli shift by decomposing its domain $J = [0, 1)$ into $J_0 := [0, 1/2)$ and $J_1 := [1/2, 1)$. For $x \in [0, 1)$ define the *output map* s^1 by[2]

$$s : [0, 1) \to \{0, 1\} \ , \ s(x) := \begin{cases} 0 \ , \ x \in J_0 \\ 1 \ , \ x \in J_1 \end{cases} \tag{20}$$

and let $s_{n+1} := s(x_n)$. Now choose some initial condition $x_0 \in J$. According to the above rule we obtain a digit $s_1 \in \{0, 1\}$. Iterating the Bernoulli shift according to $x_{n+1} = B(x_n)$ then generates a sequence of digits $\{s_1, s_2, \ldots, s_n\}$. This sequence yields nothing else than the binary representation of the given initial condition x_0.[2,3,5] If we assume that we pick an initial condition x_0 at random and feed it into our map without knowing about its precise value, this simple algorithm enables us to find out what number we have actually chosen. In other words, here we have a mechanism of *creation of information* about the initial condition x_0 by analyzing the chaotic orbit generated from it as time evolves.

Conversely, if we now assume that we already knew the initial state up to, say, m digits precision and we iterate $p > m$ times, we see that the map simultaneously *destroys information* about the current and future states, in terms of digits, as time evolves. So creation of information about previous states goes along with loss of information about current and future states. This process is quantified by the *Kolmogorov-Sinai (KS) entropy* (also called metric, or measure-theoretic entropy), which measures the exponential rate at which information is produced, respectively lost in a dynamical system, as we will see below.

The situation is similar to the following thought experiment illustrated in Fig. 4: Let us assume we have a gas consisting of molecules, depicted as billiard balls, which is constrained to the left half of the box as shown in (a). This is like having some information about the initial conditions of all gas molecules, which are in a more localized, or ordered, state. If we remove the piston as in (b), we observe that the gas spreads out over the full box until it reaches a uniform equilibrium steady state. We then have less information available about the actual positions of all gas molecules, that

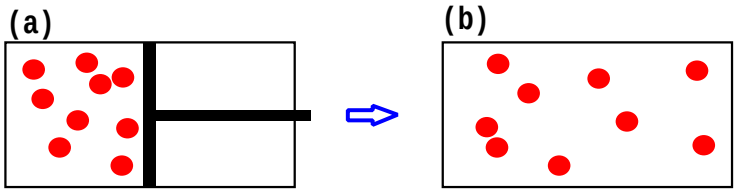


Fig. 4. Schematic representation of a gas of molecules in a box. In (a) the gas is constrained by a piston to the left hand side of the box, in (b) the piston is removed and the gas can spread out over the whole box. This illustrates the basic idea of (physical) entropy production.

is, we have increased the disorder of the whole system. This observation lies at the heart of what is called *thermodynamic entropy production* in the statistical physics of many-particle systems which, however, is usually assessed by quantities that are different from the KS-entropy.

At this point we may not further elaborate on the relation to statistical physical theories. Instead, let us make precise what we mean by KS-entropy starting from the famous *Shannon (or information) entropy*.[3,4] This entropy is defined as

$$H_S := \sum_{i=1}^{r} p_i \ln\left(\frac{1}{p_i}\right) \quad , \tag{21}$$

where p_i , $i = 1, \ldots, r$ are the probabilities for the r possible outcomes of an experiment. Think, for example, of a roulette game, where carrying out the experiment one time corresponds to $n = 1$ in the iteration of an unknown map. H_S then measures the amount of uncertainty concerning the outcome of the experiment, which can be understood as follows:

(1) Let $p_1 = 1$, $p_i = 0$ otherwise. By defining $p_i \ln\left(\frac{1}{p_i}\right) := 0$, $i \neq 1$, we have $H_S = 0$. This value of the Shannon entropy must therefore characterize the situation where the outcome is completely certain.
(2) Let $p_i = 1/r$, $i = 1, 2, \ldots, r$. Then we obtain $H_S = \ln r$ thus characterizing the situation where the outcome is most uncertain because of equal probabilities.

Case (1) thus represents the situation of no information gain by doing the experiment, case (2) corresponds to maximum information gain. These two special cases must therefore define the lower and upper bounds of H_S,

$$0 \leq H_S \leq \ln r \quad . \tag{22}$$

This basic concept of information theory carries over to dynamical systems by identifying the probabilities p_i with invariant probability measures μ_i^* on subintervals of a given dynamical system's phase space. The precise connection is worked out in four steps:[3,5]

1. Partition and refinement:

Consider a map F acting on $J \subseteq \mathbb{R}$, and let μ^* be an invariant probability measure generated by the map. Let $\{J_i\},\ i = 1, \ldots, s$ be a *partition* of J.[8] We now construct a *refinement* of this partition as illustrated by the following example:

Example 6.

Consider the Bernoulli shift displayed in Fig. 5. Start with the partition $\{J_0, J_1\}$ shown in (a). Now create a *refined* partition by iterating these two partition parts backwards according to $B^{-1}(J_i)$ as indicated in (b). Alternatively, you may take the second forward iterate $B^2(x)$ of the Bernoulli shift and then identify the preimages of $x = 1/2$ for this map. In either case the new partition parts are obtained to

$$\begin{aligned} J_{00} &:= \{x : x \in J_0\ ,\ B(x) \in J_0\} \\ J_{01} &:= \{x : x \in J_0\ ,\ B(x) \in J_1\} \\ J_{10} &:= \{x : x \in J_1\ ,\ B(x) \in J_0\} \\ J_{11} &:= \{x : x \in J_1\ ,\ B(x) \in J_1\} \quad . \end{aligned} \tag{23}$$

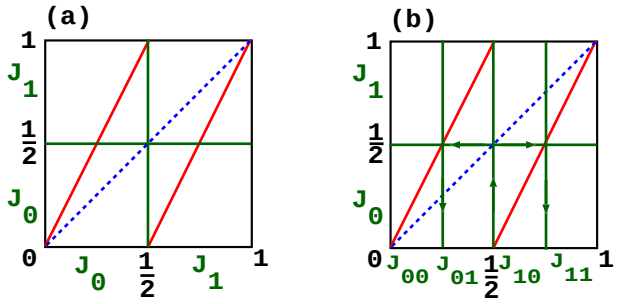


Fig. 5. (a) The Bernoulli shift and a partition of the unit interval consisting of two parts. (b) Refinement of this partition under backward iteration.

If we choose $x_0 \in J_{00}$ we thus know in advance that the orbit emerging from this initial condition under iteration of the map will remain in J_0

at the next iteration. That way, the refined partition clearly yields more information about the dynamics of single orbits.

More generally, for a given map F the above procedure is equivalent to defining

$$\{J_{i_1 i_2}\} := \{J_{i_1} \cap F^{-1}(J_{i_2})\} \quad . \tag{24}$$

The next round of refinement proceeds along the same lines yielding

$$\{J_{i_1 i_2 i_3}\} := \{J_{i_1} \cap F^{-1}(J_{i_2}) \cap F^{-2}(J_{i_3})\} \quad , \tag{25}$$

and so on. For convenience we define

$$\{J_i^n\} := \{J_{i_1 i_2 \ldots i_n}\} = \{J_{i_1} \cap F^{-1}(J_{i_2}) \cap \ldots \cap F^{-(n-1)}(J_{i_n})\} \quad . \tag{26}$$

2. H-function:

In analogy to the Shannon entropy Eq. (21) next we define the function

$$H(\{J_i^n\}) := -\sum_i \mu^*(J_i^n) \ln \mu^*(J_i^n) \quad , \tag{27}$$

where $\mu^*(J_i^n)$ is the invariant measure of the map F on the partition part J_i^n of the nth refinement.

Example 7.

For the Bernoulli shift with uniform invariant probability density $\rho^*(x) = 1$ and associated (Lebesgue) measure $\mu^*(J_i^n) = \int_{J_i^n} dx\, \rho^*(x) = \mathrm{diam}\,(J_i^n)$ we can calculate

$$\begin{aligned} H(\{J_i^1\}) &= -(\frac{1}{2}\ln\frac{1}{2} + \frac{1}{2}\ln\frac{1}{2}) = \ln 2 \\ H(\{J_i^2\}) &= H(\{J_{i_1} \cap B^{-1}(J_{i_2})\}) = -4(\frac{1}{4}\ln\frac{1}{4}) = \ln 4 \\ H(\{J_i^3\}) &= \ldots = \ln 8 = \ln 2^3 \\ &\vdots \\ H(\{J_i^n\}) &= \ln 2^n \end{aligned} \tag{28}$$

3. Take the limit:

We now look at what we obtain in the limit of infinitely refined partition by

$$h(\{J_i^n\}) := \lim_{n\to\infty} \frac{1}{n} H(\{J_i^n\}) \quad , \tag{29}$$

which defines the rate of gain of information over n refinements.

Example 8.

For the Bernoulli shift we trivially obtain

$$h(\{J_i^n\}) = \ln 2 \quad . \tag{30}$$

4. Supremum over partitions:

We finish the definition of the KS-entropy by maximizing $h(\{J_i^n\})$ over all available partitions,

$$h_{KS} := \sup_{\{J_i^n\}} h(\{J_i^n\}) \quad . \tag{31}$$

The last step can be avoided if the partition $\{J_i^n\}$ is *generating* for which it must hold that $\mathrm{diam}\,(J_i^n) \to 0$ $(n \to \infty)$.[4,11] It is quite obvious that for the Bernoulli shift the partition chosen above is generating in that sense, hence $h_{KS} = \ln 2$ for this map.

These considerations suggest yet another definition of deterministic chaos:

Definition 6. Measure-theoretic chaos[4]

A map $F : J \to J$, $J \subseteq \mathbb{R}$, is said to be chaotic in the sense of exhibiting *dynamical randomness* if $h_{KS} > 0$.

Again, one may wonder about the relation between this new definition and our previous one in terms of Ljapunov chaos. Let us look again at the Bernoulli shift:

Example 9.

For $B(x)$ we have calculated the Ljapunov exponent to $\lambda = \ln 2$, see Example 4. Above we have seen that $h_{KS} = \ln 2$ for this map, so we arrive at $\lambda = h_{KS} = \ln 2$.

That this equality is not an artefact due to the simplicity of our chosen model is stated by the following theorem:

Theorem 1. *Pesin's Theorem (1977)*[5]

For closed C^2 Anosov systems the KS-entropy is equal to the sum of positive Ljapunov exponents.

An Anosov system is a diffeomorphism where the expanding and contracting directions in phase space exhibit a particularly "nice", so-called hyperbolic structure.[5] A proof of this theorem goes considerably beyond the scope of this chapter. In the given formulation it applies to higher-dimensional dynamical systems that are "suitably well-behaved" in the sense of exhibiting the Anosov property. Applied to one-dimensional maps, it means that if we consider transformations which are "closed" by mapping an interval onto itself, $F : J \to J$, under certain conditions (which we do not further specify here) and if there is a positive Ljapunov exponent $\lambda > 0$

we can expect that $\lambda = h_{KS}$, as we have seen for the Bernoulli shift. In fact, the Bernoulli shift provides an example of a map that does not fulfill the conditions of the above theorem precisely. However, the theorem can also be formulated under weaker assumptions, and it is believed to hold for an even wider class of dynamical systems.

In order to get an intuition why this theorem should hold, let us look at the information creation in a simple one-dimensional map such as the Bernoulli shift by considering two orbits $\{x_k\}_{k=0}^n$, $\{x'_k\}_{k=0}^n$ starting at nearby initial conditions $|x'_0 - x_0| \leq \delta x_0$, $\delta x_0 \ll 1$. Recall the encoding defined by Eq. (20). Under the first m iterations these two orbits will then produce the very same sequences of symbols $\{s_k\}_{k=1}^m$, $\{s'_k\}_{k=1}^m$, that is, we cannot distinguish them from each other by our encoding. However, due to the ongoing stretching of the initial displacement δx_0 by a factor of two, eventually there will be an m such that starting from $p > m$ iterations different symbol sequences are generated. Thus we can be sure that in the limit of $n \to \infty$ we will be able to distinguish initially arbitrarily close orbits. If you like analogies, you may think of extracting information about the different initial states via the stretching produced by the iteration process like using a magnifying glass. Therefore, under iteration the exponential rate of separation of nearby trajectories, which is quantified by the positive Ljapunov exponent, must be equal to the rate of information generated, which in turn is given by the KS-entropy. This is at the heart of Pesin's theorem.

2.4. *Open systems, fractals and escape rates*

So far we have only studied closed systems, where intervals are mapped onto themselves. Let us now consider an *open system*, where points can leave the unit interval by never coming back to it. Consequently, in contrast to closed systems the total number of points is not conserved anymore. This situation can be modeled by a slightly generalized example of the Bernoulli shift.

Example 10. In the following we will study the map

$$B_a : [0,1) \to [1-a/2, a/2) \ , \ B_a(x) := \begin{cases} ax \ , \ 0 \leq x < 1/2 \\ ax + 1 - a \ , \ 1/2 \leq x < 1 \end{cases} \ , \tag{32}$$

see Fig. 6, where the slope $a \geq 2$ defines a control parameter. For $a = 2$ we recover our familiar Bernoulli shift, whereas for $a > 2$ the map defines an

open system. That is, whenever points are mapped into the escape region of width Δ these points are removed from the unit interval. You may thus think of the escape region as a subinterval that absorbs any particles mapped onto it.

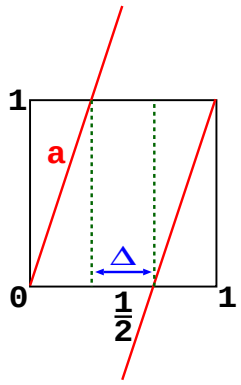


Fig. 6. A generalization of the Bernoulli shift, defined as a parameter-dependent map $B_a(x)$ modeling an open system. The slope a defines a control parameter, Δ denotes the width of the escape region.

We now wish to compute the number of points N_n remaining on the unit interval at time step n, where we start from a uniform distribution of $N_0 = N$ points on this interval at $n = 0$. This can be done as follows: Recall that the probability density $\rho_n(x)$ was defined by

$$\rho_n(x) := \frac{N_{n,j}}{N dx} \quad , \tag{33}$$

where $N_{n,j}$ is the number of points in the interval dx centered around the position x_j at time step n.[1] With $N_n = \sum_j N_{n,j}$ we have that

$$N_1 = N_0 - \rho_0 N \Delta \quad . \tag{34}$$

By observing that for $B_a(x)$, starting from $\rho_0 = 1$ points are always uniformly distributed on the unit interval at subsequent iterations, we can derive an equation for the density ρ_1 of points covering the unit interval at the next time step $n = 1$. For this purpose we divide the above equation by the total number of points N (multiplied with the total width of the unit interval, which however is one), which yields

$$\rho_1 = \frac{N_1}{N} = \rho_0 - \rho_0 \Delta = \rho_0 (1 - \Delta) \quad . \tag{35}$$

This procedure can be reiterated starting now from

$$N_2 = N_1 - \rho_1 N\Delta \tag{36}$$

leading to

$$\rho_2 = \frac{N_2}{N} = \rho_1(1-\Delta) \quad , \tag{37}$$

and so on. For general n we thus obtain

$$\rho_n = \rho_{n-1}(1-\Delta) = \rho_0(1-\Delta)^n = \rho_0 e^{n\ln(1-\Delta)} \quad , \tag{38}$$

or correspondingly

$$N_n = N_0 e^{n\ln(1-\Delta)} \quad , \tag{39}$$

which suggests the following definition:

Definition 7. For an open system with exponential decrease of the number of points,

$$N_n = N_0 e^{-\gamma n} \quad , \tag{40}$$

γ is called the *escape rate*.

In case of our mapping we thus identify

$$\gamma = \ln\frac{1}{1-\Delta} \tag{41}$$

as the escape rate. We may now wonder whether there are any initial conditions that never leave the unit interval and about the character of this set of points. The set can be constructed as exemplified for $B_a(x)$, $a = 3$ in Fig. 7.

Example 11.

Let us start again with a uniform distribution of points on the unit interval. We can then see that the points which remain on the unit interval after one iteration of the map form two sets, each of length 1/3. Iterating now the boundary points of the escape region backwards in time according to $x_n = B_3^{-1}(x_{n+1})$, we can obtain all preimages of the escape region. We find that initial points which remain on the unit interval after two iterations belong to four smaller sets, each of length 1/9, as depicted at the bottom of Fig. 7. Repeating this procedure infinitely many times reveals that the points which never leave the unit interval form the very special set $\mathcal{C}_{B_3}$, which is known as the *middle third Cantor set*.

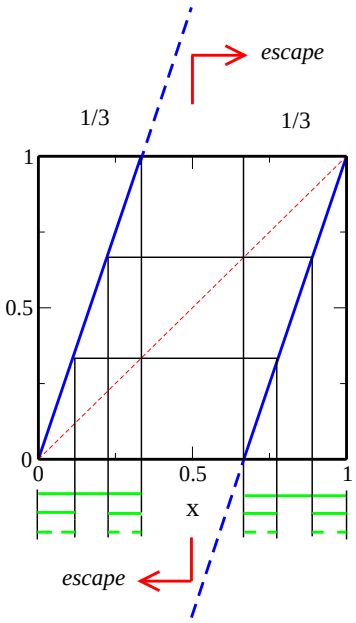


Fig. 7. Construction of the set $\mathcal{C}_{B_3}$ of initial conditions of the map $B_3(x)$ that never leave the unit interval.

Definition 8. Cantor set[3]

A *Cantor set* is a closed set which consists entirely of boundary points each of which is a limit point of the set.

Let us explore some fundamental properties of the set $\mathcal{C}_{B_3}$:[3]

(1) From Fig. 7 we can infer that the total length l_n of the intervals of points remaining on the unit interval after n iterations, which is identical with the Lebesgue measure μ_L of these sets, is

$$l_0 = 1\ ,\ l_1 = \frac{2}{3}\ ,\ l_2 = \frac{4}{9} = \left(\frac{2}{3}\right)^2\ ,\ldots,l_n = \left(\frac{2}{3}\right)^n \quad . \tag{42}$$

We thus see that

$$l_n = \left(\frac{2}{3}\right)^n \to 0\ (n \to \infty) \quad , \tag{43}$$

that is, the total length of this set goes to zero, $\mu_L(\mathcal{C}_{B_3}) = 0$. However, there exist also Cantor sets whose Lebesgue measure is larger than zero.[3] Note that matching $l_n = \exp(-n\ln(3/2))$ to Eq. (41) yields an escape rate of $\gamma = \ln(3/2)$ for this map.

(2) By using the binary encoding Eq. (20) for all intervals of $\mathcal{C}_{B_3}$, thus mapping all elements of this set onto all the numbers in the unit interval, it can nevertheless be shown that our Cantor set contains an *uncountable number of points*.[5]
(3) By construction $\mathcal{C}_{B_3}$ must be the invariant set of the map $B_3(x)$ under iteration, so the invariant measure of our open system must be the measure defined on the Cantor set, $\mu^*(\mathcal{C})$, $\mathcal{C} \in \mathcal{C}_{B_3}$;[11] see the following Example 12 for the procedure of how to calculate this measure.
(4) For the next property we need the following definition:

Definition 9. repeller[4,5]
The limit set of points that never escape is called a *repeller*. The orbits that escape are *transients*, and $1/\gamma$ is the typical duration of them.

From this we can conclude that $\mathcal{C}_{B_3}$ represents the repeller of the map $B_3(x)$.
(5) Since $\mathcal{C}_{B_3}$ is completely disconnected by only consisting of boundary points, its topology is highly singular. Consequently, no invariant density $\rho^*(x)$ can be defined on this set, since this concept presupposes a certain "smoothness" of the underlying topology such that one can meaningfully speak of "small subintervals dx" on which one counts the number of points, see Eq. (33). In contrast, $\mu^*(\mathcal{C})$ is still well-defined, and we speak of it as a *singular measure*.[5]
(6) Fig. 7 shows that $\mathcal{C}_{B_3}$ is *self-similar*, in the sense that smaller pieces of this structure reproduce the entire set upon magnification.[3] Here we find that the whole set can be reproduced by magnifying the fundamental structure of two subsets with a gap in the middle by a constant factor of three. Often such a simple scaling law does not exist for these types of sets. Instead, the scaling may depend on the position x of the subset, in which case one speaks of a *self-affine* structure.[3,4,8]
(7) Again we need a definition:

Definition 10. fractals, qualitatively[4]
Fractals are geometrical objects that possess nontrivial structure on arbitrarily fine scales.

In case of our Cantor set $\mathcal{C}_{B_3}$, these structures are generated by a simple scaling law. However, generally fractals can be arbitrarily complicated on finer and finer scales. An example of a structure that is trivial, hence not fractal, is a straight line. The fractality of such complicated

sets can be assessed by quantities called *fractal dimensions*,[3,4] which generalize the integer dimensionality of Euclidean geometry. It is interesting how in our case fractal geometry naturally comes into play, forming an important ingredient of the theory of dynamical systems. However, here we do not further elaborate on the concept of fractal geometry and refer to the literature instead.[3,4,8]

Example 12.

Let us now compute all three basic quantities that we have introduced so far, that is: the Ljapunov exponent λ and the KS-entropy h_{ks} *on* the invariant set as well as the escape rate γ *from* this set. We do so for the map $B_3(x)$ which, as we have learned, produces a fractal repeller. According to Eqs. (12),(14) we have to calculate

$$\lambda(\mathcal{C}_{B_3}) = \int_0^1 d\mu^* \ln |B_3'(x)| \quad . \tag{44}$$

However, for typical points we have $B_3'(x) = 3$, hence the Ljapunov exponent must trivially be

$$\lambda(\mathcal{C}_{B_3}) = \ln 3 \quad , \tag{45}$$

because the probability measure μ^* is normalised. The calculation of the KS-entropy requires a bit more work: Recall that

$$H(\{\mathcal{C}_i^n\}) := -\sum_{i=1}^{2^n} \mu^*(\mathcal{C}_i^n) \ln \mu^*(\mathcal{C}_i^n) \quad , \tag{46}$$

see Eq. (27), where $\mathcal{C}_i^n$ denotes the ith part of the emerging Cantor set at the nth level of its construction. We now proceed along the lines of Example 7. From Fig. 7 we can infer that

$$\mu^*(\mathcal{C}_i^1) = \frac{\frac{1}{3}}{\frac{2}{3}} = \frac{1}{2}$$

at the first level of refinement. Note that here we have *renormalized* the (Lebesgue) measure on the partition part $\mathcal{C}_i^1$. That is, we have divided the measure by the total measure surviving on all partition parts such that we always arrive at a proper probability measure under iteration. The measure constructed that way is known as the *conditionally invariant measure* on the Cantor set.[4] Repeating this procedure yields

$$\mu^*(\mathcal{C}_i^2) = \frac{\frac{1}{9}}{\frac{4}{9}} = \frac{1}{4}$$

$$\vdots$$

$$\mu^*(\mathcal{C}_i^n) = \frac{(\frac{1}{3})^n}{(\frac{2}{3})^n} = 2^{-n} \tag{47}$$

from which we obtain

$$H(\{\mathcal{C}_i^n\}) = -\sum_{i=1}^{2^n} 2^{-n} \ln 2^{-n} = n \ln 2 \quad . \tag{48}$$

We thus see that by taking the limit according to Eq. (29) and noting that our partitioning is generating on the fractal repeller $\mathcal{C}_{B_3} = \{\mathcal{C}_i^\infty\}$, we arrive at

$$h_{KS}(\mathcal{C}_{B_3}) = \lim_{n\to\infty} \frac{1}{n} H(\{\mathcal{C}_i^n\}) = \ln 2 \quad . \tag{49}$$

Finally, with Eq.(41) and an escape region of size $\Delta = 1/3$ for $B_3(x)$ we get for the escape rate

$$\gamma(\mathcal{C}_{B_3}) = \ln \frac{1}{1-\Delta} = \ln \frac{3}{2} \quad , \tag{50}$$

as we have already seen before.

In summary, we have that $\gamma(\mathcal{C}_{B_3}) = \ln \frac{3}{2} = \ln 3 - \ln 2$, $\lambda(\mathcal{C}_{B_3}) = \ln 3$, $h_{KS}(\mathcal{C}_{B_3}) = \ln 2$, which suggests the relation

$$\gamma(\mathcal{C}_{B_3}) = \lambda(\mathcal{C}_{B_3}) - h_{KS}(\mathcal{C}_{B_3}) \quad . \tag{51}$$

Again, this equation is no coincidence. It is a generalization of Pesin's theorem to open systems, known as the *escape rate formula*. This equation holds under similar conditions like Pesin's theorem, which is recovered from it if there is no escape.[5]

3. Chaotic diffusion

We now apply the concepts of dynamical systems theory developed in Ref.[1] to a fundamental problem in nonequilibrium statistical physics, which is to understand the microscopic origin of diffusion in many-particle systems. We start with a reminder of diffusion as a simple random walk on the line. Modeling such processes by suitably generalizing the piecewise linear map studied previously, we will see how diffusion can be generated by microscopic deterministic chaos. The main result will be an exact formula relating the diffusion coefficient, which characterizes macroscopic diffusion of particles, to the dynamical systems quantities introduced before.

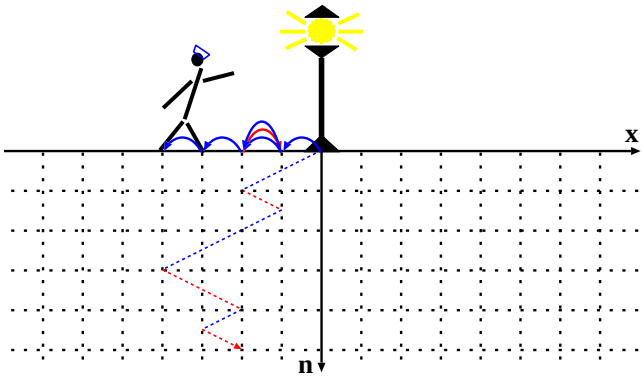


Fig. 8. The "problem of the random walk" in terms of a drunken sailor at a lamppost. The space-time diagram shows an example of a trajectory for such a drunken sailor, where $n \in \mathbb{N}$ holds for discrete time and $x \in \mathbb{R}$ for the position of the sailor on a discrete lattice of spacing s.

3.1. *What is chaotic diffusion?*

In order to learn about chaotic diffusion, we must first understand what ordinary diffusion is all about. Here we introduce this concept by means of a famous example, see Fig. 8: Let us imagine that some evening a sailor wants to walk home, however, he is completely drunk such that he has no control over his single steps. For sake of simplicity let us imagine that he moves in one dimension. He starts at a lamppost at position $x = 0$ and then makes steps of a certain step length s to the left and to the right. Since he is completely drunk he loses all memory between any single steps, that is, all steps are *uncorrelated*. It is like tossing a coin in order to decide whether to go to the left or to the right at the next step. We may now ask for the probability to find the sailor after n steps at position x, i.e., a distance $|x|$ away from his starting point.

The answer to this question is obtained from a calculation for an ensemble of sailors starting from the lamppost and is given in terms of Gaussian probability distributions for the sailor's positions, which are obtained in a suitable scaling limit.[12] Fig. 9 sketches the spreading of such a diffusing distribution of sailors in time. The mathematical reason for the emerging Gaussianity of the probability distributions is nothing else than the central limit theorem.

We may now wish to quantify the speed by which a "droplet of sailors" starting at the lamppost spreads out. This can be done by calculating the

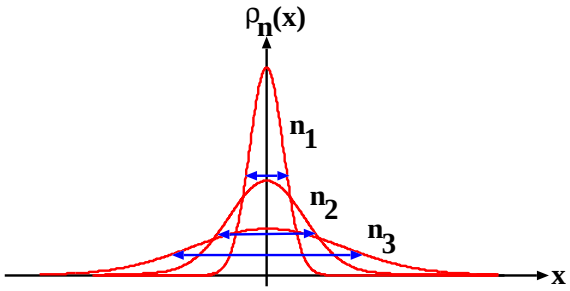


Fig. 9. Probability distribution functions $\rho_n(x)$ to find a sailor after n time steps at position x on the line, calculated for an ensemble of sailors starting at the lamppost, cf. Fig. 8. Shown are three probability densities after different numbers of iteration $n_1 < n_2 < n_3$.

diffusion coefficient for this system. In case of one-dimensional dynamics the diffusion coefficient can be defined by the Einstein formula

$$D := \lim_{n\to\infty} \frac{1}{2n} < x^2 > \quad , \tag{52}$$

where

$$< x^2 > := \int dx\, x^2 \rho_n(x) \tag{53}$$

is the variance, or second moment, of the probability distribution $\rho_n(x)$ at time step n, also called *mean square displacement* of the particles. This formula may be understood as follows: For our ensemble of sailors we may choose $\rho_0(x) = \delta(x)$ as the initial probability distribution with $\delta(x)$ denoting the (Dirac) δ-function, which mimicks the situation that all sailors start at the same lamppost at $x = 0$. If our system is ergodic the diffusion coefficient should be independent of the choice of the initial ensemble. The spreading of the distribution of sailors is then quantified by the growth of the mean square displacement in time. If this quantity grows linearly in time, which may not necessarily be the case but holds true if our probability distributions for the positions are Gaussian in the long-time limit,[7] the magnitude of the diffusion coefficient D tells us how quickly our ensemble of sailors disperses. For further details about a statistical physics description of diffusion we refer to the literature.[12]

In contrast to this well-known picture of diffusion as a stochastic random walk, the theory of dynamical systems makes it possible to treat diffusion as a *deterministic dynamical process*. Let us replace the sailor by a point particle. Instead of coin tossing, the orbit of such a particle starting at

initial condition x_0 may then be generated by a *chaotic* dynamical system of the type as considered in the previous sections, $x_{n+1} = F(x_n)$. Note that defining the one-dimensional map $F(x)$ together with this equation yields the *full microscopic equations of motion* of the system. You may think of these equations as a caricature of Newton's equations of motion modeling the diffusion of a single particle. Most importantly, in contrast to the drunken sailor with his memory loss after any time step here the *complete memory* of a particle is taken into account, that is, all steps are fully correlated. The decisive new fact that distinguishes this dynamical process from the one of a simple uncorrelated random walk is hence that x_{n+1} is uniquely determined by x_n, rather than having a random distribution of x_{n+1} for a given x_n. If the resulting dynamics of an ensemble of particles for given equations of motion has the property that a diffusion coefficient $D > 0$ Eq. (52) exists, we speak of *deterministic* or *chaotic diffusion*.[2,5–7]

Fig. 10 shows the simple model of chaotic diffusion that we shall study in the following. It depicts a "chain of boxes" of chain length $L \in \mathbb{N}$, which continues periodically in both directions to infinity, and the orbit of a moving point particle. Let us first specify the map defined on the unit interval, which we may call the box map. For this we choose the map $B_a(x)$ introduced in our previous Example 10. We can now periodically continue this box map onto the whole real line by a *lift of degree one*,

$$B_a(x+1) = B_a(x) + 1 \quad . \tag{54}$$

Physically speaking, this means that $B_a(x)$ continued onto the real line is translational invariant with respect to integers. Note furthermore that we have chosen a box map whose graph is point symmetric with respect to the center of the box at $(x, y) = (0.5, 0.5)$. This implies that the graph of the full map $B_a(x)$ is anti-symmetric with respect to $x = 0$,

$$B_a(x) = -B_a(-x) \quad , \tag{55}$$

so that there is no "drift" in this chain of boxes. The drift case with broken symmetry could be studied as well,[7] but we exclude it here for sake of simplicity.

3.2. *The diffusion equation*

In the last section we have sketched in a nutshell what, in our setting, we mean if we speak of diffusion. This picture is made more precise by deriving an equation that exactly generates the dynamics of the probability densities

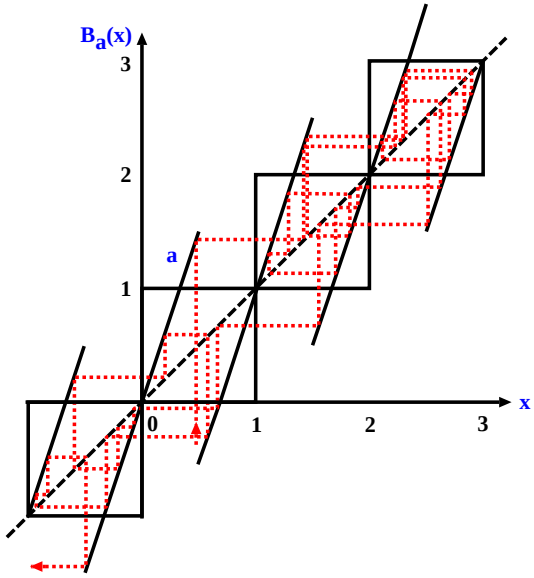


Fig. 10. A simple model for chaotic diffusion. The dashed line depicts the orbit of a diffusing particle in form of a cobweb plot.[8] The slope a serves as a control parameter for the periodically continued piecewise linear map $B_a(x)$.

displayed in Fig. 9.[12] For this purpose, let us reconsider for a moment the situation depicted in Fig. 4. There, we had a gas with an initially very high concentration of particles on the left hand side of the box. After the piston was removed, it seemed natural that the particles spread out over the right hand side of the box as well thus diffusively covering the whole box. We may thus come to the conclusion that, firstly, *there will be diffusion if the density of particles in a substance is non-uniform in space*. For this density of particles and by restricting ourselves to diffusion in one dimension in the following, let us write $\tilde{n} = \tilde{n}(x,t)$, which holds for the number of particles that we can find in a small line element dx around the position x at time step t divided by the total number of particles N.

As a second observation, we see that *diffusion occurs in the direction of decreasing particle density*. This may be expressed as

$$j =: -D\frac{\partial \tilde{n}}{\partial x} \quad , \tag{56}$$

which according to Einstein's formula Eq. (52) may be considered as a second definition of the diffusion coefficient D. Here the flux $j = j(x,t)$ denotes the number of particles passing through an area perpendicular to

the direction of diffusion per time t. This equation is known as *Fick's first law*. Finally, let us assume that no particles are created or destroyed during our diffusion process. In other words, we have *conservation of the number of particles* in form of

$$\frac{\partial \tilde{n}}{\partial t} + \frac{\partial j}{\partial x} = 0 \quad . \tag{57}$$

This *continuity equation* expresses the fact that whenever the particle density $\tilde{n}$ changes in time t, it must be due to a spatial change in the particle flux j. Combining the equation with Fick's first law we obtain *Fick's second law*,

$$\frac{\partial \tilde{n}}{\partial t} = D \frac{\partial^2 \tilde{n}}{\partial x^2} \quad , \tag{58}$$

which is also known as the *diffusion equation*. Mathematicians call the process defined by this equation a *Wiener process*, whereas physicists rather speak of *Brownian motion*. If we would now solve the diffusion equation for the drunken sailor initial density $\tilde{n}(x,0) = \delta(x)$, we would obtain the precise functional form of our spreading Gaussians in Fig. 9,

$$\tilde{n}(x,t) = \frac{1}{\sqrt{4\pi D t}} \exp\left(-\frac{x^2}{4Dt}\right) \quad . \tag{59}$$

Calculating the second moment of this distribution according to Eq. (53) would lead us to recover Einstein's definition of the diffusion coefficient Eq. (52). Therefore, both this definition and the one provided by Fick's first law are consistent with each other.

3.3. *Basics of the escape rate formalism*

We are now fully prepared for establishing an interesting link between dynamical systems theory and statistical mechanics. We start with a brief outline of the concept of this theory, which is called the *escape rate formalism*.[5,6] It consists of three steps:

Step 1: *Solve the one-dimensional diffusion equation* Eq. (58) derived above *for absorbing boundary conditions*. That is, we consider now some type of *open* system similar to what we have studied in the previous section. We may thus expect that the total number of particles $N(t) := \int dx\, \tilde{n}(x,t)$ within the system decreases exponentially as time evolves according to the law expressed by Eq. (40), that is,

$$N(t) = N(0) e^{-\gamma_{de} t} \quad . \tag{60}$$

It will turn out that the escape rate γ_{de} defined by the diffusion equation with absorbing boundaries is a function of the system size L and of the diffusion coefficient D.

Step 2: *Solve the Frobenius-Perron equation*

$$\rho_{n+1}(x) = \int dy\, \rho_n(y)\, \delta(x - F(y)) \quad , \tag{61}$$

which represents the continuity equation for the probability density $\rho_n(x)$ of the map $F(x)$,[1,3–5] for the very same *absorbing boundary conditions* as in Step 1. Let us assume that the dynamical system under consideration is normal diffusive, that is, that a diffusion coefficient $D > 0$ exists. We may then expect a decrease in the number of particles that is completely analogous to what we have obtained from the diffusion equation. That is, if we define as before $N_n := \int dx\, \rho_n(x)$ as the total number of particles within the system at discrete time step n, in case of normal diffusion we should obtain

$$N_n = N_0 e^{-\gamma_{FP} n} \quad . \tag{62}$$

However, in contrast to Step 1 here the escape rate γ_{FP} should be fully determined by the dynamical system that we are considering. In fact, we have already seen before that for open systems the escape rate can be expressed exactly as the difference between the positive Ljapunov exponent and the KS-entropy on the fractal repeller, cf. the escape rate formula Eq. (51).

Step 3: If the functional forms of the particle density $\tilde{n}(x,t)$ of the diffusion equation and of the probability density $\rho_n(x)$ of the map's Frobenius-Perron equation *match in the limit of system size and time going to infinity* — which is what one has to show —, the escape rates γ_{de} obtained from the diffusion equation and γ_{FP} calculated from the Frobenius-Perron equation should be equal,

$$\gamma_{de} = \gamma_{FP} \quad , \tag{63}$$

providing a fundamental link between the statistical physical theory of diffusion and dynamical systems theory. Since γ_{de} is a function of the diffusion coefficient D, and knowing that γ_{FP} is a function of dynamical systems quantities, we should then be able to express D exactly in terms of these dynamical systems quantifiers. We will now illustrate how this method works by applying it to our simple chaotic diffusive model introduced above.

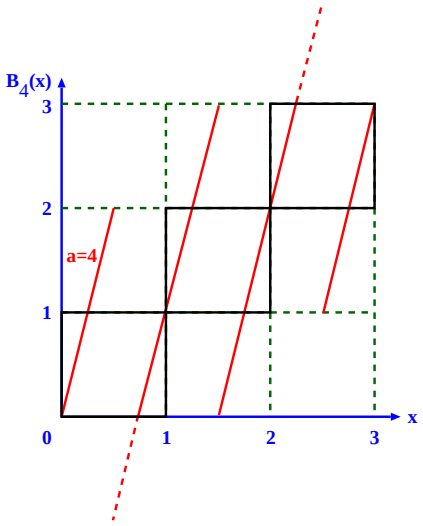


Fig. 11. Our previous map $B_a(x)$ periodically continued onto the whole real line for the specific parameter value $a = 4$. The example shown depicts a chain of length $L = 3$. The dashed quadratic grid indicates a Markov partition for this map.

3.4. *The escape rate formalism applied to a simple map*

Let us consider the map $B_a(x)$ lifted onto the whole real line for the specific parameter value $a = 4$, see Fig. 11. With L we denote the chain length. Proceeding along the above lines, let us start with

Step 1: Solve the one-dimensional diffusion equation Eq. (58) for the *absorbing boundary conditions*

$$\tilde{n}(0,t) = \tilde{n}(L,t) = 0 \quad , \tag{64}$$

which models the situation that particles escape precisely at the boundaries of our one-dimensional domain. A straightforward calculation yields

$$\tilde{n}(x,t) = \sum_{m=1}^{\infty} b_m \exp\left(-\left(\frac{m\pi}{L}\right)^2 Dt\right) \sin\left(\frac{m\pi}{L}x\right) \tag{65}$$

with b_m denoting the Fourier coefficients.

Step 2: Solve the Frobenius-Perron equation Eq. (61) for the same absorbing boundary conditions,

$$\rho_n(0) = \rho_n(L) = 0 \quad . \tag{66}$$

In order to do so, we first need to introduce *Markov partitions* for our map $B_4(x)$:

Definition 11. Markov partition, verbally[4]

For one-dimensional maps acting on compact intervals a partition is called *Markov* if *parts of the partition* get mapped again onto *parts of the partition*, or onto *unions of parts of the partition*.

Example 13.

The dashed quadratic grid in Fig. 11 defines a Markov partition for the lifted map $B_4(x)$.

Having a Markov partition at hand enables us to rewrite the Frobenius-Perron equation in form of a matrix equation, where a Frobenius-Perron matrix operator acts onto probability density vectors defined with respect to this special partitioning. In order to see this, consider an initial density of points that covers, e.g., the interval in the second box of Fig. 11 uniformly. By applying the map onto this density, one observes that points of this interval get mapped two-fold onto the interval in the second box again, but that there is also escape from this box which uniformly covers the third and the first box intervals, respectively. This mechanism applies to any box in our chain of boxes, modified only by the absorbing boundary conditions at the ends of the chain of length L. Taking into account the stretching of the density by the slope $a = 4$ at each iteration, this suggests that the Frobenius-Perron equation Eq. (61) can be rewritten as

$$\boldsymbol{\rho}_{n+1} = \frac{1}{4} T(4) \, \boldsymbol{\rho}_n \quad , \tag{67}$$

where the L x L-transition matrix $T(4)$ must read

$$T(4) = \begin{pmatrix} 2 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 & \cdots & 0 \\ \vdots & & & \vdots & \vdots & & & \vdots \\ 0 & \cdots & 0 & 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 2 \end{pmatrix} \quad . \tag{68}$$

Note that in any row and in any column we have three non-zero matrix elements except in the very first and the very last rows and columns, which reflect the absorbing boundary conditions. In Eq.(67) this transition matrix $T(4)$ is applied to a column vector $\boldsymbol{\rho}_n$ corresponding to the probability density $\rho_n(x)$, which can be written as

$$\boldsymbol{\rho}_n = |\rho_n(x)> := (\rho_n^1, \rho_n^2, \ldots, \rho_n^k, \ldots, \rho_n^L)^* \quad , \tag{69}$$

where "$*$" denotes the transpose and ρ_n^k represents the component of the probability density in the kth box, $\rho_n(x) = \rho_n^k$, $k-1 < x \leq k$, $k = 1, \ldots, L$, ρ_n^k being constant on each part of the partition. We see that this transition matrix is symmetric, hence it can be diagonalized by spectral decomposition. Solving the eigenvalue problem

$$T(4)\,|\phi_m(x)> = \chi_m(4)\,|\phi_m(x)> \quad , \tag{70}$$

where $\chi_m(4)$ and $|\phi_m(x)>$ are the eigenvalues and eigenvectors of $T(4)$, respectively, one obtains

$$\begin{aligned} |\rho_n(x)> &= \frac{1}{4}\sum_{m=1}^{L} \chi_m(4)\,|\phi_m(x)><\phi_m(x)|\rho_{n-1}(x)> \\ &= \sum_{m=1}^{L} \exp\left(-n\ln\frac{4}{\chi_m(4)}\right)\,|\phi_m(x)><\phi_m(x)|\rho_0(x)> \quad , \end{aligned} \tag{71}$$

where $|\rho_0(x)>$ is the initial probability density vector. Note that the choice of initial probability densities is restricted by this method to functions that can be written in the vector form of Eq.(69). It remains to solve the eigenvalue problem Eq. (70).[7] The eigenvalue equation for the single components of the matrix $T(4)$ reads

$$\phi_m^k + 2\phi_m^{k+1} + \phi_m^{k+2} = \chi_m \phi_m^{k+1} \quad , \quad 0 \leq k \leq L-1 \quad , \tag{72}$$

supplemented by the absorbing boundary conditions

$$\phi_m^0 = \phi_m^{L+1} = 0 \quad . \tag{73}$$

This equation is of the form of a discretized ordinary differential equation of degree two, hence we make the ansatz

$$\phi_m^k = a\cos(k\theta) + b\sin(k\theta) \quad , \quad 0 \leq k \leq L+1 \quad . \tag{74}$$

The two boundary conditions lead to

$$a = 0 \quad \text{and} \quad \sin((L+1)\theta) = 0 \tag{75}$$

yielding

$$\theta_m = \frac{m\pi}{L+1} \quad , \quad 1 \leq m \leq L \quad . \tag{76}$$

The eigenvectors are then determined by

$$\phi_m^k = b\sin(k\theta_m) \quad . \tag{77}$$

Combining this equation with Eq. (72) yields as the eigenvalues

$$\chi_m = 2 + 2\cos\theta_m \quad . \tag{78}$$

Step 3: Putting all details together, it remains to match the solution of the diffusion equation to the one of the Frobenius-Perron equation: In the limit of time t and system size L to infinity, the density $\tilde{n}(x,t)$ Eq. (65) of the diffusion equation reduces to the largest eigenmode,

$$\tilde{n}(x,t) \simeq \exp\left(-\gamma_{de} t\right) B \sin\left(\frac{\pi}{L} x\right) \quad , \tag{79}$$

where

$$\gamma_{de} := \left(\frac{\pi}{L}\right)^2 D \tag{80}$$

defines the escape rate as determined by the diffusion equation. Analogously, for discrete time n and chain length L to infinity we obtain for the probability density of the Frobenius-Perron equation, Eq.(71) with Eq.(77),

$$\begin{aligned} \rho_n(x) \simeq \exp\left(-\gamma_{FP} n\right) \tilde{B} \sin\left(\frac{\pi}{L+1} k\right) \quad , \\ k = 0, \ldots, L+1 \quad , \quad k-1 < x \leq k \end{aligned} \tag{81}$$

with an escape rate of this dynamical system given by

$$\gamma_{FP} = \ln \frac{4}{2 + 2\cos(\pi/(L+1))} \quad , \tag{82}$$

which is determined by the largest eigenvalue χ_1 of the matrix $T(4)$, see Eq.(71) with Eq.(78). We can now see that the functional forms of the eigenmodes of Eqs.(79) and (81) match precisely. This allows us to match Eqs. (80) and (82) leading to

$$D(4) = \left(\frac{L}{\pi}\right)^2 \gamma_{FP} \quad . \tag{83}$$

Using the right hand side of Eq. (82) and expanding it for $L \to \infty$, this formula enables us to calculate the diffusion coefficient $D(4)$ to

$$D(4) = \left(\frac{L}{\pi}\right)^2 \gamma_{FP} = \frac{1}{4} \frac{L^2}{(L+1)^2} + \mathcal{O}(L^{-4}) \ \to \ \frac{1}{4} \quad (L \to \infty) \quad . \tag{84}$$

Thus we have developed a method by which we can exactly calculate the deterministic diffusion coefficient of a simple chaotic dynamical system. However, more importantly, instead of using the explicit expression for γ_{FP} given by Eq. (82), let us remind ourselves of the escape rate formula Eq. (51) for γ_{FP},

$$\gamma_{FP} = \gamma(\mathcal{C}_{B_4}) = \lambda(\mathcal{C}_{B_4}) - h_{KS}(\mathcal{C}_{B_4}) \quad , \tag{85}$$

which more geneally expresses this escape rate in terms of dynamical systems quantities. Combining this equation with the above equation Eq. (83) leads to our final result, the *escape rate formula for chaotic diffusion*[6]

$$D(4) = \lim_{L\to\infty} \left(\frac{L}{\pi}\right)^2 [\lambda(\mathcal{C}_{B_4}) - h_{KS}(\mathcal{C}_{B_4})] \quad . \tag{86}$$

We have thus established a fundamental link between quantities assessing the chaotic properties of dynamical systems and the statistical physical property of diffusion.

4. Exercises and solutions

4.1. *Exercises*

(1) Prove Eq. (7).
(2) Consider the map defined by the function

$$E(x) := \begin{cases} 2x+1 & , \quad -1 \le x < -1/2 \\ 2x & , \quad -1/2 \le x < 1/2 \\ 2x-1 & , \quad 1/2 \le x \le 1\,. \end{cases} \tag{87}$$

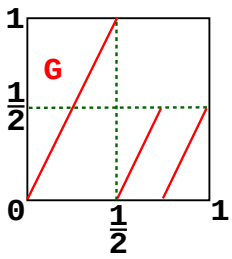
Draw the graph of this map. Is E ergodic? Prove your answer.
(3) Consider the Frobenius-Perron equation

$$\rho_{n+1}(x) = \sum_{x=G(x^i)} \rho_n(x^i)|G'(x^i)|^{-1} =: P\rho_n(x) \tag{88}$$

for a map G defined on the real line, where $\rho_n(x)$ are probability densities at time step $n \in \mathbb{N}_0$ and P defines the Frobenius-Perron operator.
(a) Show that P is a linear and positive operator.
(b) Construct P for the map G defined in the figure below and verify that

$$\rho^*(x) = \begin{cases} 4/3 & , \quad 0 \le x < 1/2 \\ 2/3 & , \quad 1/2 \le x \le 1 \end{cases} \tag{89}$$

is an invariant density of the above Frobenius-Perron equation.



(c) By assuming that G is ergodic, calculate the Ljapunov exponent λ for this map.

(4) Consider the asymmetric tent map

$$S(x) := \begin{cases} ax & , \quad 0 \leq x < 1/a \\ b - bx & , \quad 1/a \leq x \leq 1 \end{cases} \tag{90}$$

with $1/a + 1/b = 1$.

(a) Calculate the Ljapunov exponent λ for this map.
(b) Show that for the H-function defined by Eq. (27) in the lecture notes it holds $H(\{J_i^2\}) = 2H(\{J_i^1\})$. By assuming that $\forall n \in \mathbb{N} H(\{J_i^n\}) = nH(\{J_i^1\})$ calculate the KS-entropy h_{KS}. Compare your result for h_{KS} with the one obtained for λ.

(5) Consider the map $H(x) = 5x \mod 1$ on a domain which has "holes" where points escape from the unit interval. Let these escape regions be defined by the two subintervals $(0.2, 04)$ and $(0.6, 0.8)$.
(a) Sketch the map and the first two steps in the construction of its fractal repeller $\mathcal{R}_H$.
(b) Calculate the escape rate $\gamma(\mathcal{R}_H)$.
(c) Calculate the Ljapunov exponent $\lambda(\mathcal{R}_H)$
(d) Calculate the KS-entropy $h_{KS}(\mathcal{R}_H)$.
(e) By using these results, verify the escape rate formula for this map.

(6) Verify Eq. (84) by using Eqs. (83), (82).

4.2. *Solutions*

(1) Applying the chain rule we get

$$\begin{aligned} (F^n)'(x) &= (F(F^{n-1}))'(x) = F'(F^{n-1}(x))(F^{n-1})'(x) \\ &= \ldots = F'(x_{n-1})F'(x_{n-2}) \ldots F'(x_0) \end{aligned} \tag{91}$$

with $x = x_0$.

(2) We leave the drawing of the map to the reader.
Choose for g the indicator function $g : [-1, 1] \to \{0, 1\}$ with

$$g(x) := \begin{cases} 0 & , \quad -1 \leq x < 0 \\ 1 & , \quad 0 \leq x \leq 1 \, . \end{cases} \tag{92}$$

We need to check Def. 4. For g we have

$$\int_{-1}^{1} d\mu^* \, |g(x)| = \int_0^1 d\mu^* = \int_0^1 dx \, \rho^*(x) < \infty \, , \tag{93}$$

because μ^* is a probability measure. Hence the assumption is fulfilled, but for $-1 \leq x < 0$ we have $\overline{g(x)} = 0$ while for $0 \leq x \leq 1$ $\overline{g(x)} = 1$. Hence $\overline{g(x)}$ depends on x, consequently $\overline{g(x)} \neq const.$ in contradiction to Def. 4, which implies that the map E is not ergodic.

(3) (a) Linearity of an operator P is defined by

$$P(\alpha_1\rho^1 + \alpha_2\rho^2) = \alpha_1 P\rho^1 + \alpha_2 P\rho^2 \tag{94}$$

with $\alpha_1, \alpha_2 \in \mathbb{R}$ and probability densities ρ^1, ρ^2. Positivity means $P\rho(x) \geq 0$. The proofs of these two properties for the Frobenius-Perron operator defined by Eq. (88) is then straightforward.

(b) From the figure we can infer

$$G(x) := \begin{cases} 2x & , \quad 0 \leq x < 1/2 \\ 2x - 1 & , \quad 1/2 \leq x < 3/4 \\ 2x - 3/2 & , \quad 3/4 \leq x \leq 1 \end{cases} \tag{95}$$

and $G'(x) = 2 \forall x (x \neq 1/2, 3/4)$. We construct the piecewise inverse functions

$$\begin{aligned} x^1 &= G^{-1}(x) = x/2\,,\ 0 \leq x^1 < 1/2\,,\ 0 \leq x < 1 \\ x^2 &= G^{-1}(x) = (x+1)/2\,,\ 1/2 \leq x^2 < 3/4\,,\ 0 \leq x < 1/2 \\ x^3 &= G^{-1}(x) = (x+3/2)/2\,,\ 3/4 \leq x^1 < 1\,,\ 0 \leq x \leq 1/2\,. \end{aligned} \tag{96}$$

Plugging these results into Eq. (88) yields

$$P\rho(x) = 1/2\rho(x/2) + 1/2\rho((x+1)/2) + 1/2\rho(x/2 + 3/4)\,. \tag{97}$$

Feeding $\rho^*(x)$ given by Eq. (89) into this equation shows $P\rho^*(x) = \rho^*(x)$.

(c) The calculation is analogous to the one of Example 5 yielding $\lambda = \ln 2$.

(4) (a) See again Example 5; in this case the solution is $\lambda = 1/a \ln a + 1/b \ln b$.

(b) The calculation follows Sec. 2.3:

i. It is convenient to choose as a partition the one generated by the backward iteration of the critical point at $x_c = 1/a$.

ii.

$$\begin{aligned} H(\{J_i^1\}) &= 1/a \ln a + 1/b \ln b \\ H(\{J_i^2\}) &= 1/a^2 \ln a^2 + 1/(ab) \ln(ab) + 1/b^2 \ln b^2 + 1/(ba) \ln(ba) \\ &= 2H(\{J_i^1\}) \end{aligned} \tag{98}$$

iii. Using the stated assumption we get

$$h(\{J_i^n\}) = \lim_{n\to\infty} \frac{1}{n} H(\{J_i^n\}) = H(\{J_i^1\})\,. \tag{99}$$

iv. Since the partition above is generating we find

$$h_{KS} = h(\{J_i^n\}) = 1/a \ln a + 1/b \ln b = \lambda \tag{100}$$

according to Pesin's theorem.

(5) (a) The sketch can be performed in analogy to Fig. 7 and is left to the reader.
(b) Since $\rho^*(x) = 1$, just consider the Lebesgue measure of the sets $\{R_i^n\}$, i.e., the total lengths l_n, which are $l_0 = 1$, $l_1 = 3/5$, $l_2 = 9/25$. It follows that $l_n = (3/5)^n = \exp(-n \ln(5/3))$, hence $\gamma(\mathcal{R}_H) = \ln(5/3)$.
(c) It is easy to find $\lambda(\mathcal{R}_H) = \ln 5$.
(d) The calculation is in analogy to Ex. 12 and yields $h_{KS}(\mathcal{R}_H) = \ln 3$.
(e) We thus have the escape rate formula $\gamma(\mathcal{R}_H) = \ln(5/3) = \lambda(\mathcal{R}_H) - h_{KS}(\mathcal{R}_H)$.

(6) With $\cos x \simeq 1 - x^2/2$ and $\ln(1 \pm x) \simeq \pm x$ we have

$$\gamma_{FP} \simeq \ln \frac{4}{2 + 2 - (\pi/(L+1))^2} \simeq \frac{1}{4} \left(\frac{\pi}{L+1} \right)^2 , \tag{101}$$

which leads to Eq. (84).

References

1. D. Arrowsmith, *Applied Dynamical Systems*, In eds. S.Bullet, T.Fearn, and F.Smith, *Dynamical and complex systems*, vol. 5, *LTCC Advanced Mathematics Series*. World Scientific, Singapore (2016).
2. H. Schuster, *Deterministic chaos*, 2nd edn. VCH Verlagsgesellschaft mbH, Weinheim (1989).
3. E. Ott, *Chaos in Dynamical Systems*. Cambridge University Press, Cambridge (1993).
4. C. Beck and F. Schlögl, *Thermodynamics of Chaotic Systems*. vol. 4, *Cambridge nonlinear science series*, Cambridge University Press, Cambridge (1993).
5. J. Dorfman, *An introduction to chaos in nonequilibrium statistical mechanics*. Cambridge University Press, Cambridge (1999).
6. P. Gaspard, *Chaos, scattering, and statistical mechanics*. Cambridge University Press, Cambridge (1998).
7. R. Klages, *Microscopic chaos, fractals and transport in nonequilibrium statistical mechanics*. vol. 24, *Advanced Series in Nonlinear Dynamics*, World Scientific, Singapore (2007).

8. K. Alligood, T. Sauer, and J. Yorke, *Chaos - An introduction to dynamical systems*. Springer, New York (1997).
9. V. Arnold and A. Avez, *Ergodic problems of classical mechanics*. W.A. Benjamin, New York (1968).
10. A. Lasota and M. Mackay, *Chaos, Fractals, and Noise*. Springer-Verlag (1994).
11. A. Katok and B. Hasselblatt, *Introduction to the modern theory of dynamical systems*. vol. 54, *Encyclopedia of Mathematics and its Applications*, Cambridge University Press, Cambridge (1995).
12. F. Reif, *Fundamentals of statistical and thermal physics*. McGraw-Hill, Auckland (1965).

Subject Index