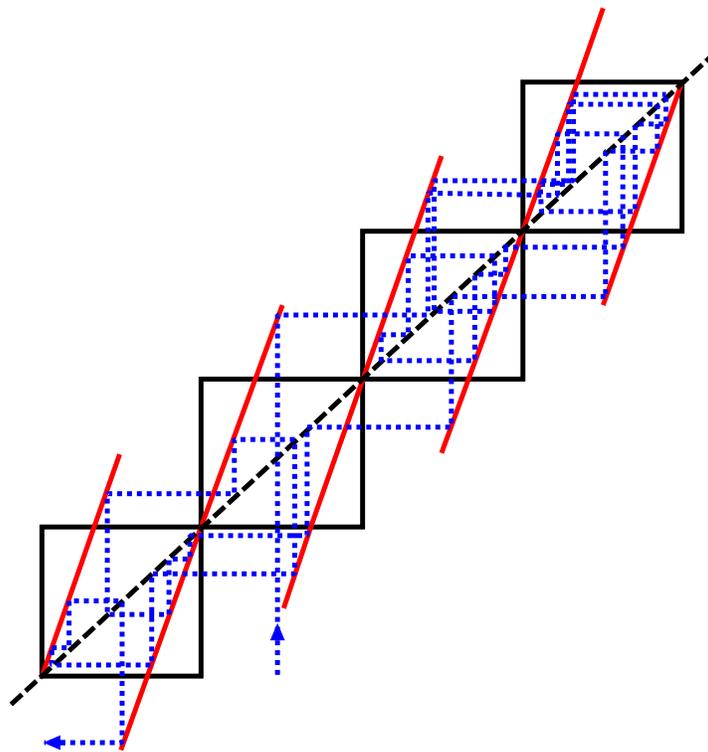


Applied Dynamical Systems

London Taught Course Centre, Lectures 5 – 10
From Deterministic Chaos to Deterministic Diffusion
Version 1.3, 19/03/2010



Rainer Klages
Queen Mary University of London
School of Mathematical Sciences

Abstract

These are easy-to-read lecture notes for a short first-year Ph.D. student course on Applied Dynamical Systems, given at the London Taught Course Centre in Spring 2008, 2009 and 2010. The full course consists of two parts, covering four and six hours of lectures, respectively. The first part taught by Wolfram Just is on *Basic Dynamical Systems*. The present notes cover only the second part, which leads *From Deterministic Chaos to Deterministic Diffusion*. The lecture notes for the first part (plus an additional chapter on intermittency) as well as some exercises for the second part including model solutions are available on the course webpage under http://www.maths.qmul.ac.uk/~klages/teaching/ltcc_ads. Part 2 is essentially self-contained and does not really require knowledge of Part 1 or any other preknowledge on dynamical systems.

Contents

1	Deterministic chaos in one-dimensional maps	2
1.1	Dynamics of simple maps	2
1.2	Ljapunov chaos	4
1.3	Entropies	8
1.4	Open systems, fractals and escape rates	13
2	Deterministic diffusion in one-dimensional maps	20
2.1	What is deterministic diffusion?	20
2.2	Escape rate formalism for deterministic diffusion	24
2.2.1	The diffusion equation	24
2.2.2	Basic idea of the escape rate formalism	25
2.2.3	The escape rate formalism worked out for a simple map	26
2.3	Anomalous diffusion in intermittent maps	31
	Bibliography	34

In these lectures we will continue to study one-dimensional maps as simple examples of dynamical systems. Eventually we will apply these models to investigate a central topic of statistical physics, which is a microscopic understanding of *deterministic diffusion* in many-particle systems. In order to do so, we first introduce two fundamental concepts by which one can quantify *deterministic chaos* in dynamical systems. We then outline basics of diffusion in many-particle systems and show how such processes can be understood starting from microscopic chaos in dynamical systems.

The first chapter elaborates on textbook material of chaotic dynamical systems [Bec93, Ott93], the second chapter covers advanced topics that emerged in research carried out over the past 20 years [Dor99, Gas98, Kla96]. An extended version of these notes (amended by a new chapter on ‘anomalous diffusion’) has been published as Ref. [Kla10]. If you need a more general introduction to dynamical systems as a background reading, which is in the spirit of these lecture notes, please feel free to check out the set of fully worked-out lecture notes of the masters level module *Introduction to Dynamical Systems* [Kla07a].

I thank Phil Howard and Nickolay Korabel for help with some of the figures.

1 Deterministic chaos in one-dimensional maps

We start with a brief reminder about the dynamics of time-discrete one-dimensional dynamical systems. We then quantify chaos in terms of Ljapunov exponents and (metric) entropies by focusing on systems that are closed on the unit interval. These ideas are then generalized to the case of open systems, where particles can escape (in terms of absorbing boundary conditions). Most of the concepts that we are going to introduce carry over, suitably generalized, to higher-dimensional and time-continuous dynamical systems.

It should be remarked that the next section draws on Section 2.2.2 and the beginning of Chapter 3 of [Kla07a], Section 1.2 is an abbreviated version of Sections 5.3 and 7.2 of the same reference, in case you wish to look up further details.

1.1 Dynamics of simple maps

As a warm-up, let us recall the following:

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

$$F : J \rightarrow J \quad , \quad x_{n+1} = F(x_n) \tag{1.1}$$

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*. As you should know already, it works in the following way:

$$\begin{aligned} x_1 &= F(x_0) = F^1(x_0), \\ x_2 &= F(x_1) = F(F(x_0)) = F^2(x_0) (\neq F(x_0)F(x_0) !). \\ \Rightarrow F^m(x_0) &:= \underbrace{F \circ F \circ \dots \circ F(x_0)}_{\text{m-fold composed map}}. \end{aligned} \quad (1.2)$$

In other words, there exists a unique solution to the equations of motion in form of $x_n = F(x_{n-1}) = \dots = F^n(x_0)$, which is the counterpart of the flow for time-continuous systems. Our lectures will mostly focus on simple piecewise linear maps. The following one, which you have seen before, serves as a paradigmatic example:

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

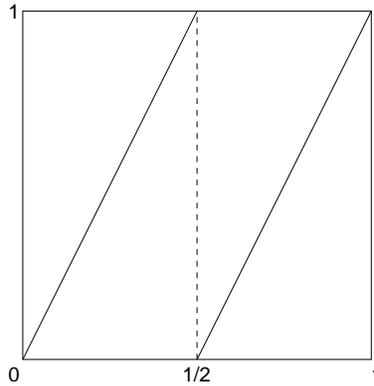


Figure 1.1: The Bernoulli shift.

The Bernoulli shift shown in Fig. 1.1 may be defined by

$$B : [0, 1) \rightarrow [0, 1), \quad B(x) := 2x \bmod 1 = \begin{cases} 2x, & 0 \leq x < 1/2 \\ 2x - 1, & 1/2 \leq x < 1 \end{cases}. \quad (1.3)$$

You already know how to display the iterated dynamics for this type of maps in terms of *orbit webs*, or *cobweb plots*, which may produce quite complicated patterns. Another way of how to think about the dynamics of such maps is as follows, see Fig. 1.2: 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 modulo operation $\bmod 1$, which leads to motion *bounded* on the unit interval.

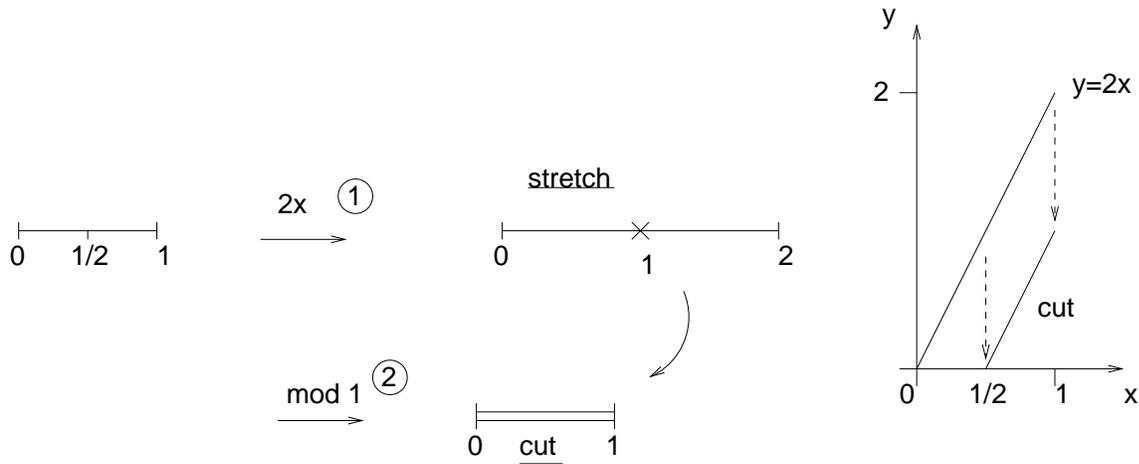


Figure 1.2: Stretch-and-cut mechanism in the Bernoulli shift.

The Bernoulli shift thus yields a simple example for an essentially nonlinear stretch-and-cut mechanism, as it typically generates *deterministic chaos* [Ott93]. Such basic mechanisms are also encountered in more realistic dynamical systems. We may remark that ‘stretch and fold’ or ‘stretch, twist and fold’ provide alternative mechanisms for generating chaotic behavior (check out the tent map!). You may wish to play around with these ideas in thought experiments, where you replace the sets of points by kneading dough. These ideas can be made mathematically precise in form of what is called *mixing* in dynamical systems, which is an important ingredient of the ergodic theory of dynamical systems [Arn68, Dor99].

1.2 Ljapunov chaos

You have already been given a definition for what one means if one speaks of *chaos in maps*. This definition, which has been proposed by Devaney [Dev89], classifies chaos in form of three conditions that have to be fulfilled: sensitivity, existence of a dense orbit, and that the periodic points are dense. In the following we will introduce a quantity called *Ljapunov exponent*, which generalizes the concept of sensitivity in form of a quantity that can be calculated more conveniently. Let us motivate the definition of this important quantity by an example:

Example 2 Ljapunov instability of the Bernoulli shift [Ott93, Rob95]

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

$$\delta x_n := |x'_n - x_n| = 2\delta x_{n-1} = 2^2\delta x_{n-2} = \dots = 2^n\delta x_0 = e^{n \ln 2}\delta x_0 \quad . \quad (1.4)$$

We thus 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

$$\delta x_n = |x'_n - x_n| = |F^n(x'_0) - F^n(x_0)| =: \delta x_0 e^{n\lambda(x_0)} \quad (\delta x_0 \rightarrow 0) \quad (1.5)$$

for which we *presuppose* that an exponential separation of trajectories exists.¹ By furthermore assuming that F is differentiable we can rewrite this equation to

$$\begin{aligned} \lambda(x_0) &= \lim_{n \rightarrow \infty} \lim_{\delta x_0 \rightarrow 0} \frac{1}{n} \ln \frac{\delta x_n}{\delta x_0} \\ &= \lim_{n \rightarrow \infty} \lim_{\delta x_0 \rightarrow 0} \frac{1}{n} \ln \frac{|F^n(x_0 + \delta x_0) - F^n(x_0)|}{\delta x_0} \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \ln \left| \frac{dF^n(x)}{dx} \right|_{x=x_0} . \end{aligned} \quad (1.6)$$

Using the chain rule (please convince yourself that this holds true, as a little exercise) we obtain

$$\left. \frac{dF^n(x)}{dx} \right|_{x=x_0} = F'(x_{n-1})F'(x_{n-2}) \dots F'(x_0) \quad , \quad (1.7)$$

which leads to our final result

$$\begin{aligned} \lambda(x_0) &= \lim_{n \rightarrow \infty} \frac{1}{n} \ln \left| \prod_{i=0}^{n-1} F'(x_i) \right| \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \ln |F'(x_i)| \quad . \end{aligned} \quad (1.8)$$

These considerations motivate the following definition:

Definition 2 [All97] *Let $F \in C^1$ be a map of the real line. The local Ljapunov exponent $\lambda(x_0)$ is defined as*

$$\lambda(x_0) := \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \ln |F'(x_i)| \quad (1.9)$$

*if this limit exists.*²

Remark 1

1. If F is not C^1 but *piecewise* C^1 , the definition can still be applied by excluding single points of non-differentiability.
2. If $F'(x_i) = 0 \Rightarrow \cancel{\lambda}(x)$. However, usually this concerns only an ‘atypical’ set of points.

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

$$\lambda(x) = \frac{1}{n} \sum_{k=0}^{n-1} \ln 2 = \ln 2 \quad (1.10)$$

at these points.

¹We emphasize that this is not always the case, see, e.g., Section 17.4 of [Kla07b].

²This definition was proposed by A.M. Ljapunov in his Ph.D. thesis 1892.

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 then of how to calculate the local Ljapunov exponent, and another to which extent it depends on initial conditions. An answer to both these questions is provided in form of 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 in form of a *time average*, where n terms along the trajectory with initial condition x_0 are summed up by averaging over n . However, this is not the only possibility to define an average quantity, as is clarified by the following definition:

Definition 3 *time and ensemble average* [Dor99, Arn68]

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

$$\overline{g(x)} := \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} g(x_k) \quad , \quad (1.11)$$

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

$$\langle g \rangle := \int_J d\mu^* g(x) \quad (1.12)$$

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 as introduced in part one of this course, alternatively see [Kla07a], and $d\mu^*$ is the associated invariant measure. 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. (1.11) we recover Definition 2 for the local Ljapunov exponent,

$$\lambda(x) := \overline{\ln |F'(x)|} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \ln |F'(x_k)| \quad , \quad (1.13)$$

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. (1.12) we obtain

$$\lambda_e := \langle \ln |F'(x)| \rangle := \int_J dx \rho^*(x) \ln |F'(x)| \quad . \quad (1.14)$$

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

$$\lambda_e = \int_0^1 dx \rho^*(x) \ln 2 = \ln 2 \quad , \quad (1.15)$$

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

$$\lambda_t(x) = \lambda_e = \ln 2 \quad . \quad (1.16)$$

This motivates the following fundamental definition:

Definition 4 *ergodicity* [Arn68, Dor99]³

A dynamical system is called ergodic if for every g on $J \subseteq \mathbb{R}$ satisfying $\int d\mu^* |g(x)| < \infty$ ⁴

$$\overline{g(x)} = \langle g \rangle \quad (1.17)$$

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 [Dor99, Bec93]. 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 [Arn68, Dor99, Tod92] for proofs of ergodicity in case of some simple examples.

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 characterising 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. 1.3 below:

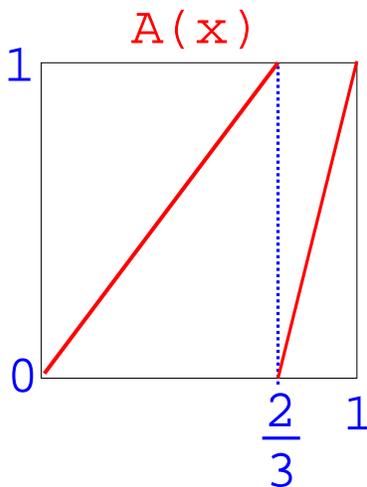


Figure 1.3: A simple map for demonstrating the calculation of Ljapunov exponents via ensemble averages.

³Note that mathematicians prefer to define ergodicity by using the concept of indecomposability [Kat95].

⁴This condition means that we require g to be a Lebesgue-integrable function. In other words, g should be an element of the function space $L^1(J, \mathcal{A}, \mu^*)$ of a set J , a σ -algebra \mathcal{A} of subsets of J and an invariant measure μ^* . This space defines the family of all possible real-valued measurable functions g satisfying $\int d\mu^* |g(x)| < \infty$, where this integral should be understood as a Lebesgue integral [Las94].

From the figure we can infer that

$$A(x) := \begin{cases} \frac{3}{2}x, & 0 \leq x < \frac{2}{3} \\ 3x - 2, & \frac{2}{3} \leq x < 1 \end{cases} \quad (1.18)$$

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 calculated as follows:

$$\lambda = \int_0^1 dx \rho^*(x) \ln |A'(x)| = \int_0^{2/3} dx \ln \frac{3}{2} + \int_{2/3}^1 dx \ln 3 = \frac{2}{3} \ln \frac{3}{2} + \frac{1}{3} \ln 3 = \ln 3 - \frac{2}{3} \ln 2 \quad (1.19)$$

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* [Rob95, Ott93, All97, Bec93]

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

You may wonder why now we have introduced another definition of chaos, which 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.⁵ 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.⁶ 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 (such as the ones between Ljapunov and Devaney chaos) are usually non-trivial and a topic of ongoing research [Kla07a]. We will encounter yet another definition of chaos in the following section.

1.3 Entropies

This section particularly builds upon the presentations in [Ott93, Dor99]; for a more mathematical approach see [Eck85]. Let us start with a brief motivation outlining the basic idea of entropy production in dynamical systems. Consider again the Bernoulli shift by recalling what was introduced as the *output map* s in a previous lecture. That is, let us decompose $J := [0, 1)$ into $J_0 := [0, 1/2)$ and $J_1 := [1/2, 1)$. For $x \in [0, 1)$ define [Sch89]

$$s : [0, 1) \rightarrow \{0, 1\}, \quad s(x) := \begin{cases} 0, & x \in J_0 \\ 1, & x \in J_1 \end{cases} \quad (1.20)$$

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 $x_{n+1} = B(x_n)$

⁵Note that a positive Ljapunov exponent implies sensitivity, but the converse does not hold true [Kla07a].

⁶Here one often silently assumes that a given dynamical system is ergodic. To prove that a system is topologically transitive as required by Devaney's definition is not any easier.

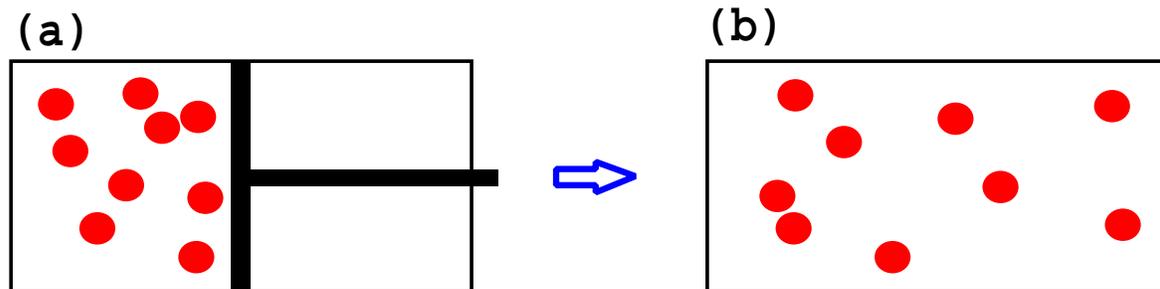


Figure 1.4: Schematic representation of a gas of molecules in a box. In (a) the gas is constrained by a piston to the left hand side of the box, in (b) the piston is removed and the gas can spread out over the whole box. This illustrates the basic idea of (physical) entropy production.

then generates a sequence of digits $\{s_1, s_2, \dots, s_n\}$. You have already seen before that this sequence yields nothing else than the binary representation of the given initial condition x_0 . 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* (1958/59; 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. 1.4: Let us assume we have a gas consisting of molecules, depicted as billiard balls, that 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 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.⁷

At this point we may not further elaborate on this relation to statistical physical theories. Instead, let us make precise what we mean by KS-entropy starting from the famous *Shannon (or information) entropy* (1948) [Ott93, Bec93]. This entropy is defined as

$$H_S := \sum_{i=1}^r p_i \ln \left(\frac{1}{p_i} \right) \quad , \quad (1.21)$$

⁷The connection between thermodynamic and metric entropy production is in fact a matter of ongoing research [Fal05].

where p_i , $i = 1, \dots, r$ 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 $p_i = 1/r$, $i = 1, 2, \dots, r$. 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 ,⁸

$$0 \leq H_S \leq \ln r \quad . \quad (1.22)$$

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:

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.⁹ Let $\{J_i\}$, $i = 1, \dots, s$ be a *partition* of J .¹⁰ We now construct a *refinement* of this partition as illustrated by the following example:

Example 6 Consider the Bernoulli shift displayed in Fig. 1.5. 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

$$\begin{aligned} J_{00} &:= \{x : x \in J_0, B(x) \in J_0\} \\ J_{01} &:= \{x : x \in J_0, B(x) \in J_1\} \\ J_{10} &:= \{x : x \in J_1, B(x) \in J_0\} \\ J_{11} &:= \{x : x \in J_1, B(x) \in J_1\} \quad . \end{aligned} \quad (1.23)$$

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.

⁸A proof employs the convexity of the entropy function and Jensen's inequality; see, e.g., the Wikipedia entry of *information entropy* for details.

⁹If not said otherwise, μ^* holds for the *physical (or natural) measure* of the map in the following [Kla07a].

¹⁰A *partition* of the interval J is a collection of subintervals whose union is J , which are pairwise disjoint except perhaps at the end points [All97].

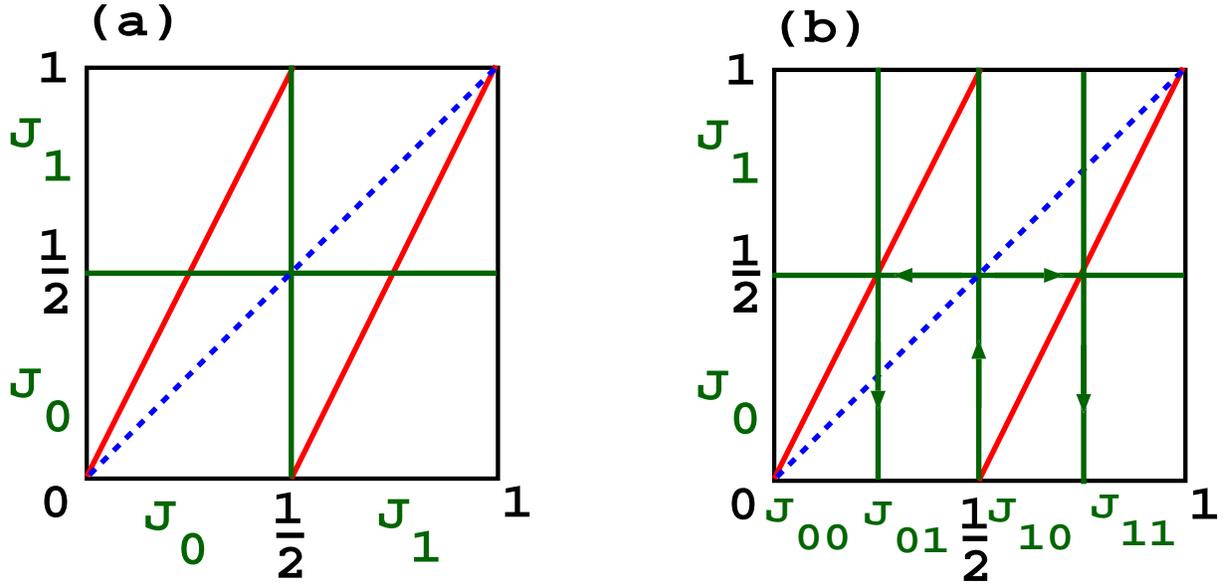


Figure 1.5: (a) The Bernoulli shift and a partition of the unit interval consisting of two parts. (b) Refinement of this partition under backward iteration.

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

$$\{J_{i_1 i_2}\} := \{J_{i_1} \cap F^{-1}(J_{i_2})\} \quad . \quad (1.24)$$

The next round of refinement proceeds along the same lines yielding

$$\{J_{i_1 i_2 i_3}\} := \{J_{i_1} \cap F^{-1}(J_{i_2}) \cap F^{-2}(J_{i_3})\} \quad , \quad (1.25)$$

and so on. For convenience we define

$$\{J_i^n\} := \{J_{i_1 i_2 \dots i_n}\} = \{J_{i_1} \cap F^{-1}(J_{i_2}) \cap \dots \cap F^{-(n-1)}(J_{i_n})\} \quad . \quad (1.26)$$

2. H-function:

In analogy to the Shannon entropy Eq. (1.21) next we define the function

$$H(\{J_i^n\}) := - \sum_i \mu^*(J_i^n) \ln \mu^*(J_i^n) \quad , \quad (1.27)$$

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

Example 7 For the Bernoulli shift with uniform invariant probability density $\rho^*(x) = 1$ and associated (Lebesgue) measure $\mu^*(J_i^n) = \int_{J_i^n} dx \rho^*(x) = \text{diam}(J_i^n)$ we can calculate

$$\begin{aligned} H(\{J_i^1\}) &= -\left(\frac{1}{2} \ln \frac{1}{2} + \frac{1}{2} \ln \frac{1}{2}\right) = \ln 2 \\ H(\{J_i^2\}) &= H(\{J_{i_1} \cap B^{-1}(J_{i_2})\}) = -4\left(\frac{1}{4} \ln \frac{1}{4}\right) = \ln 4 \\ H(\{J_i^3\}) &= \dots = \ln 8 = \ln 2^3 \\ &\vdots \\ H(\{J_i^n\}) &= \ln 2^n \end{aligned} \quad (1.28)$$

3. Take the limit:

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

$$h(\{J_i^n\}) := \lim_{n \rightarrow \infty} \frac{1}{n} H(\{J_i^n\}) \quad , \quad (1.29)$$

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

Example 8 For the Bernoulli shift we trivially obtain¹¹

$$h(\{J_i^n\}) = \ln 2 \quad . \quad (1.30)$$

4. Supremum over partitions:

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

$$h_{KS} := \sup_{\{J_i^n\}} h(\{J_i^n\}) \quad . \quad (1.31)$$

The last step can be avoided if the partition $\{J_i^n\}$ is *generating* for which it must hold that $\text{diam}(J_i^n) \rightarrow 0$ ($n \rightarrow \infty$) [Eck85, Kat95].¹² 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.

Remark 2 [Sch89]

1. For strictly periodic motion there is no refinement of partition parts under backward iteration, hence $h_{KS} = 0$; see, for example, the identity map $I(x) = x$.
2. For stochastic systems all preimages are possible, so there is immediately an infinite refinement of partition parts under backward iteration, which leads to $h_{KS} \rightarrow \infty$.

These considerations suggest yet another definition of deterministic chaos:

Definition 6 *Measure-theoretic chaos* [Bec93]

A map $F : J \rightarrow J$, $J \subseteq \mathbb{R}$, 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:

¹¹In fact, this result is already obtained after a single iteration step, i.e., without taking the limit of $n \rightarrow \infty$. This reflects the fact that the Bernoulli shift dynamics sampled that way is mapped onto a Markov process.

¹²Note that alternative definitions of a generating partition in terms of symbolic dynamics are possible [Bec93].

Theorem 1 *Pesin's Theorem (1977)* [Dor99, Eck85, You02]

For closed C^2 Anosov¹³ systems the KS-entropy is equal to the sum of positive Ljapunov exponents.

A proof of this theorem goes considerably beyond the scope of this course [You02]. 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, $F : J \rightarrow J$, 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. (1.20). 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 $n \rightarrow \infty$ 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.

We may remark that typically the KS-entropy is much harder to calculate for a given dynamical system than positive Ljapunov exponents. Hence, Pesin's theorem is often employed in the literature for indirectly calculating the KS-entropy. Furthermore, here we have described only one type of entropy for dynamical systems. It should be noted that the concept of the KS-entropy can straightforwardly be generalized leading to a whole spectrum of *Renyi entropies*, which can then be identified with topological, metric, correlation and other higher-order entropies [Bec93].

1.4 Open systems, fractals and escape rates

This section draws particularly on [Ott93, Dor99]. 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.

¹³An Anosov system is a diffeomorphism, where the expanding and contracting directions in phase space exhibit a particularly “nice”, so-called hyperbolic structure [Eck85, Dor99].

Example 10 In the following we will study the map

$$B_a : [0, 1) \rightarrow [1 - a/2, a/2), \quad B_a(x) := \begin{cases} ax, & 0 \leq x < 1/2 \\ ax + 1 - a, & 1/2 \leq x < 1 \end{cases}, \quad (1.32)$$

see Fig. 1.6, 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.

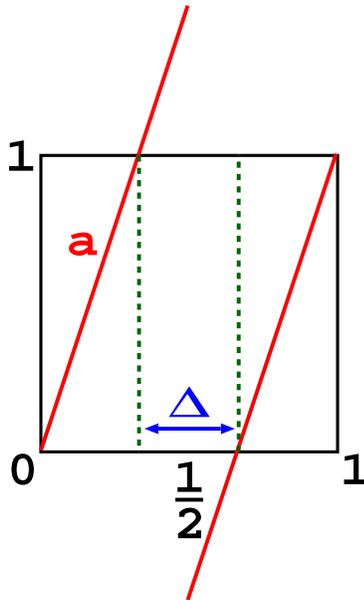


Figure 1.6: A generalization of the Bernoulli shift, defined as a parameter-dependent map $B_a(x)$ modeling an open system. The slope a defines a control parameter, Δ denotes the width of the escape region.

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: You may recall that the probability density $\rho_n(x)$ was defined by

$$\rho_n(x) := \frac{\text{number of points } N_{n,j} \text{ in interval } dx \text{ centered around position } x_j \text{ at time step } n}{\text{total number of points } N \text{ times width } dx}, \quad (1.33)$$

where $N_n = \sum_j N_{n,j}$. With this we have that

$$N_1 = N_0 - \rho_0 N \Delta. \quad (1.34)$$

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 (if you wish, multiplied with the total width of the unit interval, which however is one). This yields

$$\rho_1 = \frac{N_1}{N} = \rho_0 - \rho_0\Delta = \rho_0(1 - \Delta) \quad . \quad (1.35)$$

This procedure can be reiterated starting now from

$$N_2 = N_1 - \rho_1 N \Delta \quad (1.36)$$

leading to

$$\rho_2 = \frac{N_2}{N} = \rho_1(1 - \Delta) \quad , \quad (1.37)$$

and so on. For general n we thus obtain

$$\rho_n = \rho_{n-1}(1 - \Delta) = \rho_0(1 - \Delta)^n = \rho_0 e^{n \ln(1-\Delta)} \quad , \quad (1.38)$$

or correspondingly

$$N_n = N_0 e^{n \ln(1-\Delta)} \quad , \quad (1.39)$$

which suggests the following definition:

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

$$N_n = N_0 e^{-\gamma n} \quad , \quad (1.40)$$

γ is called the escape rate.

In case of our mapping we thus identify

$$\gamma = \ln \frac{1}{1 - \Delta} \quad (1.41)$$

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. 1.7.

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. 1.7. 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*.

Definition 8 *Cantor set [Ott93, Dev89]*

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} [Ott93]:

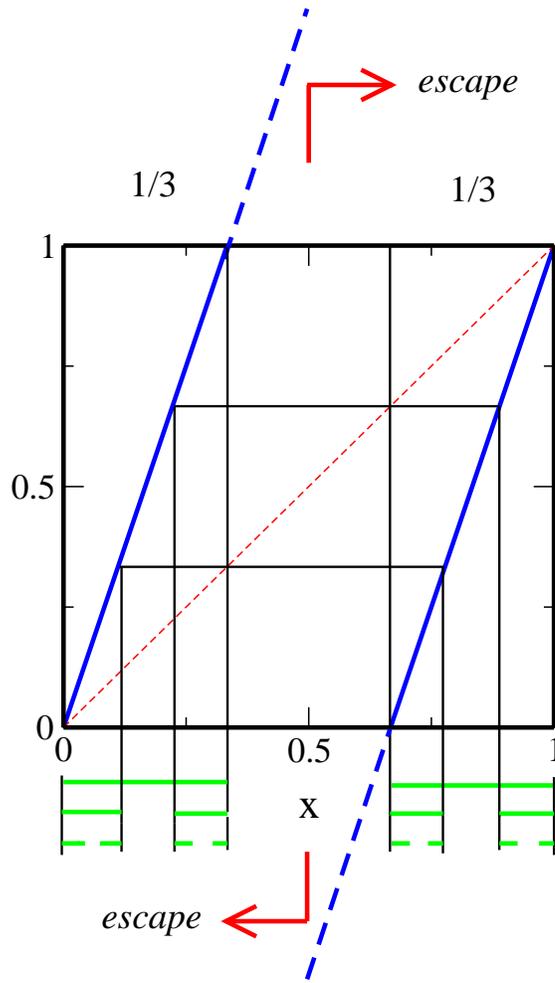


Figure 1.7: Construction of the set \mathcal{C}_{B_3} of initial conditions of the map $B_3(x)$ that never leave the unit interval.

1. From Fig. 1.7 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

$$l_0 = 1, l_1 = \frac{2}{3}, l_2 = \frac{4}{9} = \left(\frac{2}{3}\right)^2, \dots, l_n = \left(\frac{2}{3}\right)^n. \quad (1.42)$$

We thus see that

$$l_n = \left(\frac{2}{3}\right)^n \rightarrow 0 \quad (n \rightarrow \infty), \quad (1.43)$$

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 [Ott93]. Note that matching $l_n = \exp(-n \ln(3/2))$ to Eq. (1.41) yields an escape rate of $\gamma = \ln(3/2)$ for this map.

2. By using the binary encoding Eq. (1.20) 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* [Dor99, Tri95].

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, $\mu^*(\mathcal{C})$, $\mathcal{C} \in \mathcal{C}_{B_3}$ [Kat95]; 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* [Bec93, Dor99]

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. (1.33). In contrast, $\mu^*(\mathcal{C})$ is still well-defined,¹⁴ and we speak of it as a *singular measure* [Kla07a, Dor99].
6. Fig. 1.7 shows that \mathcal{C}_{B_3} is *self-similar*, in the sense that smaller pieces of this structure reproduce the entire set upon magnification [Ott93]. 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 [Man82, Fal90].
7. Again we need a definition:

Definition 10 *fractals, qualitatively* [Bec93, Fal90]

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. A famous example of a fractal in nature, mentioned in the pioneering book by Mandelbrot [Man82], is the coastline of Britain. 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* [Ott93, Bec93], 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 [Tri95, Fal90, Man82, Ott93].

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

¹⁴This is one of the reasons why mathematicians prefer to deal with measures instead of densities.

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. (1.12),(1.14) we have to calculate

$$\lambda(\mathcal{C}_{B_3}) = \int_0^1 d\mu^* \ln |B_3'(x)| \quad . \quad (1.44)$$

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

$$\lambda(\mathcal{C}_{B_3}) = \ln 3 \quad , \quad (1.45)$$

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

$$H(\{\mathcal{C}_i^n\}) := - \sum_{i=1}^{2^n} \mu^*(\mathcal{C}_i^n) \ln \mu^*(\mathcal{C}_i^n) \quad , \quad (1.46)$$

see Eq. (1.27), where \mathcal{C}_i^n denotes the i th part of the emerging Cantor set at the n th level of its construction. We now proceed along the lines of Example 7. From Fig. 1.7 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 [Bec93]. Repeating this procedure yields

$$\begin{aligned} \mu^*(\mathcal{C}_i^2) &= \frac{\frac{1}{9}}{\frac{4}{9}} = \frac{1}{4} \\ &\vdots \\ \mu^*(\mathcal{C}_i^n) &= \frac{(\frac{1}{3})^n}{(\frac{2}{3})^n} = 2^{-n} \end{aligned} \quad (1.47)$$

from which we obtain

$$H(\{\mathcal{C}_i^n\}) = - \sum_{i=1}^{2^n} 2^{-n} \ln 2^{-n} = n \ln 2 \quad . \quad (1.48)$$

We thus see that by taking the limit according to Eq. (1.29) and noting that our partitioning is generating on the fractal repeller $\mathcal{C}_{B_3} = \{\mathcal{C}_i^\infty\}$, we arrive at

$$h_{KS}(\mathcal{C}_{B_3}) = \lim_{n \rightarrow \infty} \frac{1}{n} H(\{\mathcal{C}_i^n\}) = \ln 2 \quad . \quad (1.49)$$

Finally, with Eq.(1.41) and an escape region of size $\Delta = 1/3$ for $B_3(x)$ we get for the escape rate

$$\gamma(\mathcal{C}_{B_3}) = \ln \frac{1}{1 - \Delta} = \ln \frac{3}{2} \quad , \quad (1.50)$$

as we have already seen before.¹⁵

¹⁵Note that the escape rate will generally depend not only on the size but also on the position of the escape interval [Bun08].

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

$$\gamma(\mathcal{C}_{B_3}) = \lambda(\mathcal{C}_{B_3}) - h_{KS}(\mathcal{C}_{B_3}) \quad . \quad (1.51)$$

Again, this equation is no coincidence. It is a generalization of Pesin's theorem to open systems, known as the *escape rate formula* [Kan85]. This equation holds under similar conditions like Pesin's theorem, which is recovered from it in case of no escape [Dor99].

2 Deterministic diffusion in one-dimensional maps

We now apply the concepts of dynamical systems theory developed in the previous chapter 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 a formula precisely relating the diffusion coefficient, which characterizes macroscopic diffusion of particles, to the dynamical systems quantities introduced before.

In Section 2.1, which draws upon Section 2.1 of [Kla07b], we explain the basic idea of deterministic diffusion and introduce our model. Section 2.2, which is partially based on [Kla96, Kla99], outlines a method of how to exactly calculate the diffusion coefficient for such types of dynamical systems. Section 2.3 provides the crosslink to part one of this course by outlining how intermittent dynamics leads to what is called anomalous diffusion; see Section 6.2 of [Kla07b] as well as [Kla08] for introductions to this topic.

2.1 What is deterministic diffusion?

In order to learn about deterministic diffusion, we must first understand what ordinary diffusion is all about. Here we introduce this concept by means of a famous example, see Fig. 2.1: 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.

Let us add a short historical note: This “problem of considerable interest” was first formulated by Karl Pearson in a letter to the journal *Nature* in 1905. He asked for a solution, which was provided by Lord Rayleigh referring to older work by himself. Pearson concluded: “The lesson of Lord Rayleigh’s solution is that in open country the most probable place to find a drunken man, who is at all capable of keeping on his feet, is somewhere near his starting point”. This refers to the Gaussian probability distributions for the sailor’s positions, which are obtained in a suitable scaling limit from a Gedankenexperiment with an “ensemble of

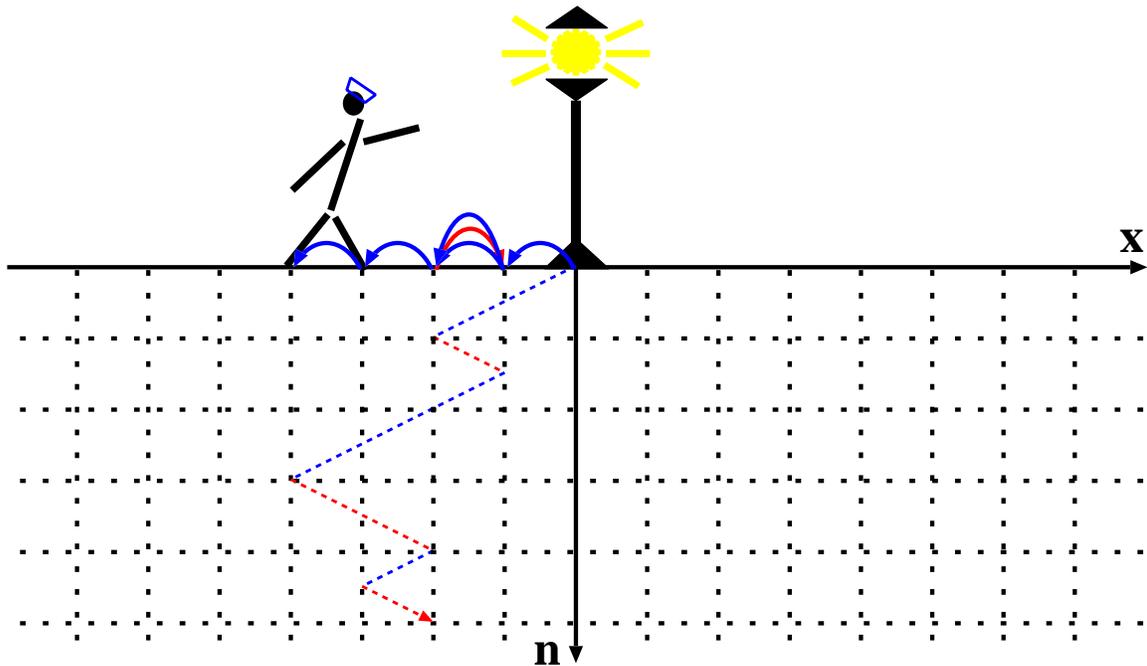


Figure 2.1: The “problem of the random walk” in terms of a drunken sailor at a lamppost. The space-time diagram shows an example of a trajectory for such a drunken sailor, where $n \in \mathbb{N}$ holds for discrete time and $x \in \mathbb{R}$ for the position of the sailor on a discrete lattice of spacing s .

sailors”¹ starting from the lamppost. Fig. 2.2 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; see the first chapter of [Rei65] for a detailed exposition of the drunken sailor’s problem from a statistical physics point of view.

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 *diffusion coefficient* for this system. In case of one-dimensional dynamics the diffusion coefficient can be defined by the Einstein formula (1905)

$$D := \lim_{n \rightarrow \infty} \frac{1}{2n} \langle x^2(n) \rangle \quad , \quad (2.1)$$

where

$$\langle x^2(n) \rangle := \int dx x^2 \rho_n(x) \quad (2.2)$$

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’ [Rei65], 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

¹The term “ensemble” means that we take copies of the original system for sampling its statistical properties.

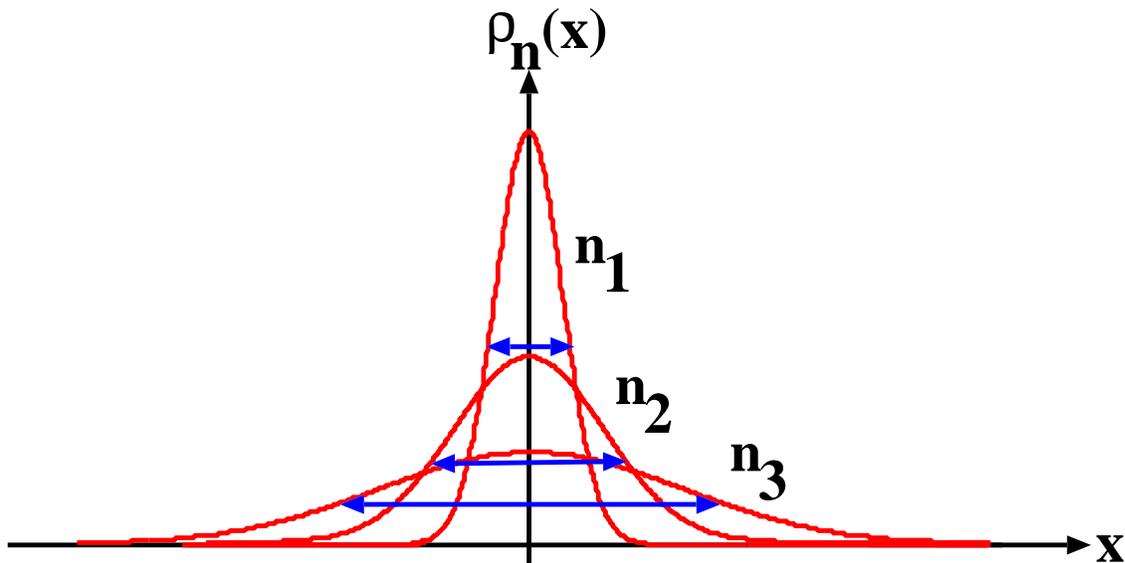


Figure 2.2: Probability distribution functions $\rho_n(x)$ to find a sailor after n time steps at position x on the line, calculated for an ensemble of sailors starting at the lamppost, cf. Fig. 2.1. Shown are three probability densities after different numbers of iteration $n_1 < n_2 < n_3$.

in time. If this quantity grows linearly in time, which may not necessarily be the case but holds true in case our probability distributions for the positions are Gaussian in the long-time limit [Kla07b], the magnitude of the diffusion coefficient D tells us how quickly our ensemble of sailors disperses. Of course, much more could be said about a statistical physics description of diffusion, but we refer to the literature for further details [Rei65].

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 chapters, $x_{n+1} = F(x_n)$. 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. (2.1) exists, we speak of (normal)² *deterministic diffusion* [Dor99, Gas98, Kla07b, Kla96, Sch89].

Fig. 2.3 shows the simple model of deterministic diffusion that we shall study in this chapter. 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

²See Section 2.3 for another type of diffusion, where D is either zero or infinite, which is called *anomalous* diffusion.

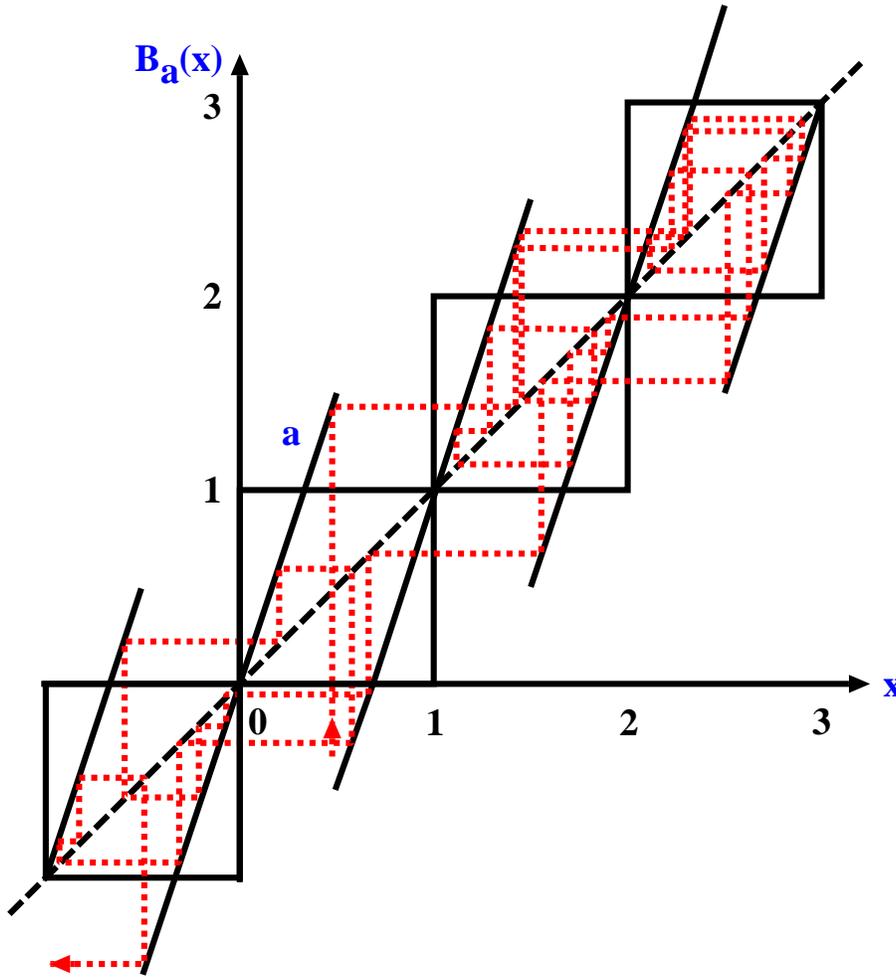


Figure 2.3: A simple model for deterministic diffusion. The dashed line depicts the orbit of a diffusing particle in form of a cobweb plot. The slope a serves as a control parameter for the periodically continued piecewise linear map $B_a(x)$.

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*,

$$B_a(x + 1) = B_a(x) + 1 \quad , \quad (2.3)$$

for which the acronym *old* has been introduced [Kat95]. 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 $x = 0$,

$$B_a(x) = -B_a(-x) \quad , \quad (2.4)$$

so that there is no “drift” in this chain of boxes. The drift case with broken symmetry could be studied as well [Kla07b], but we exclude it here for sake of simplicity.

2.2 Escape rate formalism for deterministic diffusion

Before we can start with this method we need to remind ourselves about the elementary theory of diffusion in form of the diffusion equation. We then outline the basic idea underlying the escape rate formalism, which eventually yields a simple formula expressing diffusion in terms of dynamical systems quantities. Finally, we work out this approach for our deterministic model introduced before.

2.2.1 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 displayed in Fig. 2.2.³ For this purpose, let us reconsider for a moment the situation depicted in Fig. 1.4. 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 .⁴

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

$$j =: -D \frac{\partial \tilde{n}}{\partial x} \quad , \quad (2.5)$$

which after Einstein's formula Eq. (2.1) 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

$$\frac{\partial \tilde{n}}{\partial t} + \frac{\partial j}{\partial x} = 0 \quad . \quad (2.6)$$

This equation provides an example of what is called a continuity equation in physics. It expresses the fact that whenever the particle density \tilde{n} changes in time t , this must be due to a spatial change in the particle flux j . Combining this equation with Fick's first law we obtain *Fick's second law*,

$$\frac{\partial \tilde{n}}{\partial t} = D \frac{\partial^2 \tilde{n}}{\partial x^2} \quad , \quad (2.7)$$

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

³See, e.g., [Rei65] or general undergraduate physics textbooks for the following explanations.

⁴Note the fine distinction between $\tilde{n}(x, t)$ and our previous $\rho_n(x)$, Eq.(1.33), in that here we consider continuous time t for the moment, and all our particles may interact with each other.

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 for our spreading Gaussians of Fig. 2.2,

$$\tilde{n}(x, t) = \frac{1}{\sqrt{4\pi Dt}} \exp\left(-\frac{x^2}{4Dt}\right) . \quad (2.8)$$

Calculating the second moment of this distribution according to Eq. (2.2) would lead us to recover Einstein's definition of the diffusion coefficient Eq. (2.1). Therefore, both this definition and the one provided by Fick's first law are consistent with each other.

2.2.2 Basic idea of the escape rate formalism

We are now fully prepared to put all pieces of what we have learned in the last lectures together 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*, pioneered by Gaspard and others [Gas90, Gas98, Dor99]. It consists of three steps:

Step 1: *Solve the one-dimensional diffusion equation* Eq. (2.7) 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 chapter. 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. (1.40), that is,

$$N(t) = N(0)e^{-\gamma_{de}t} . \quad (2.9)$$

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*

$$\rho_{n+1}(x) = \int dy \rho_n(y) \delta(x - F(y)) \quad (2.10)$$

that you have seen before, which represents the continuity equation for the probability density $\rho_n(x)$ of the map $F(x)$, 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

$$N_n = N_0 e^{-\gamma_{FP}n} . \quad (2.11)$$

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. (1.51).

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*

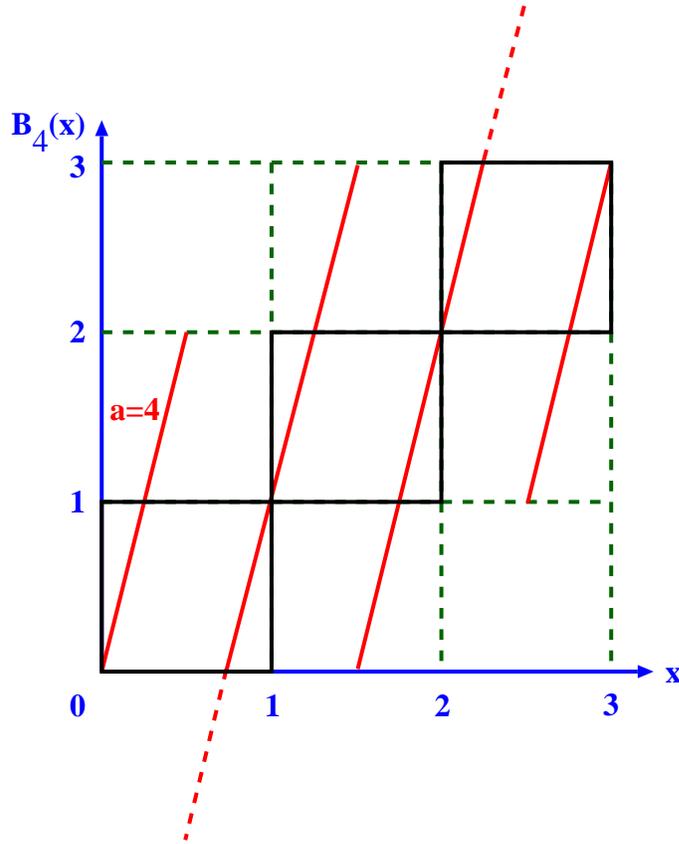


Figure 2.4: Our previous map $B_a(x)$ periodically continued onto the whole real line for the specific parameter value $a = 4$. The example shown depicts a chain of length $L = 3$. The dashed green lines indicate a Markov partition for this map.

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,

$$\gamma_{de} = \gamma_{FP} \quad , \quad (2.12)$$

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 deterministic diffusive model introduced above.

2.2.3 The escape rate formalism worked out for 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. 2.4. 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. (2.7) for the *absorbing boundary conditions*

$$\tilde{n}(0, t) = \tilde{n}(L, t) = 0 \quad , \quad (2.13)$$

which models the situation that particles escape precisely at the boundaries of our one-dimensional domain. Using separation of variables

$$n(x, t) = X(x)T(t) \quad (2.14)$$

we obtain

$$\frac{1}{T} \frac{\partial T}{\partial t} = \frac{1}{X} D \frac{\partial^2 X}{\partial x^2} =: -\gamma \quad . \quad (2.15)$$

Solving

$$\frac{\partial T}{\partial t} = -\gamma T \quad (2.16)$$

to

$$T(t) = T(0)e^{-\gamma t} \quad , \quad (2.17)$$

we see that γ defines the escape rate of the system. It remains to solve

$$\frac{\partial^2 X}{\partial x^2} = -\frac{\gamma}{D} X \quad . \quad (2.18)$$

This can be done by making the ansatz

$$X(x) = a \cos(kx) + b \sin(kx) \quad (2.19)$$

leading to

$$-ak^2 \cos(kx) - bk^2 \sin(kx) = -k^2 X = -\frac{\gamma}{D} X \quad , \quad (2.20)$$

from which we conclude that

$$k^2 = \frac{\gamma}{D} \quad . \quad (2.21)$$

We now need to consider the absorbing boundary conditions stated above, which imply that

$$X(0) = X(L) = 0 \quad . \quad (2.22)$$

This yields

$$X(0) = a = 0 \quad (2.23)$$

and

$$X(L) = b \sin(kL) = 0 \quad , \quad (2.24)$$

which is fulfilled with

$$kL = m\pi \quad , \quad m \in \mathbb{N} \quad , \quad (2.25)$$

taking the anti-symmetry of the sine function into account and that b is just some constant. We thus obtain

$$\gamma = k^2 D = \left(\frac{m\pi}{L} \right)^2 D \quad . \quad (2.26)$$

The solution of the diffusion equation can therefore be expressed in form of the Fourier series

$$\tilde{n}(x, t) = \sum_{m=1}^{\infty} b_m \exp \left(- \left(\frac{m\pi}{L} \right)^2 D t \right) \sin \left(\frac{m\pi}{L} x \right) \quad (2.27)$$

with b_m denoting the Fourier coefficients.

Step 2: Solve the Frobenius-Perron equation Eq. (2.10) for the same absorbing boundary conditions,

$$\rho_n(0) = \rho_n(L) = 0 \quad . \quad (2.28)$$

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

Definition 11 *Markov partition, verbally [Bec93, Kla07a]*

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 green lines in Fig. 2.4 define a Markov partition for the lifted map $B_4(x)$.

Of course there also exists a precise formal definition of Markov partitions, however, here we do not elaborate on these technical details [Kla07a]. 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.⁵ In order to see this, consider an initial density of points that covers, e.g., the interval in the second box of Fig. 2.4 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. (2.10) can be rewritten as

$$\boldsymbol{\rho}_{n+1} = \frac{1}{4} T(4) \boldsymbol{\rho}_n \quad , \quad (2.29)$$

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

$$T(4) = \begin{pmatrix} 2 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 & \cdots & 0 \\ \vdots & & & \vdots & \vdots & & & \vdots \\ 0 & \cdots & 0 & 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 2 \end{pmatrix} . \quad (2.30)$$

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.(2.29) 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

$$\boldsymbol{\rho}_n = |\rho_n(x)\rangle := (\rho_n^1, \rho_n^2, \dots, \rho_n^k, \dots, \rho_n^L)^* \quad , \quad (2.31)$$

⁵Implicitly we herewith choose a specific space of functions, which are tailored for studying the statistical dynamics of our piecewise linear maps.

where “ $*$ ” denotes the transpose and ρ_n^k represents the component of the probability density in the k th box, $\rho_n(x) = \rho_n^k$, $k-1 < x \leq k$, $k = 1, \dots, L$, 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

$$T(4) |\phi_m(x)\rangle = \chi_m(4) |\phi_m(x)\rangle \quad , \quad (2.32)$$

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

$$\begin{aligned} |\rho_n(x)\rangle &= \frac{1}{4} \sum_{m=1}^L \chi_m(4) |\phi_m(x)\rangle \langle \phi_m(x)| \rho_{n-1}(x)\rangle \\ &= \sum_{m=1}^L \exp\left(-n \ln \frac{4}{\chi_m(4)}\right) |\phi_m(x)\rangle \langle \phi_m(x)| \rho_0(x)\rangle \quad , \end{aligned} \quad (2.33)$$

where $|\rho_0(x)\rangle$ 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.(2.31). It remains to solve the eigenvalue problem Eq. (2.32) [Kla99]. The eigenvalue equation for the single components of the matrix $T(4)$ reads

$$\phi_m^k + 2\phi_m^{k+1} + \phi_m^{k+2} = \chi_m \phi_m^{k+1} \quad , \quad 0 \leq k \leq L-1 \quad , \quad (2.34)$$

supplemented by the absorbing boundary conditions

$$\phi_m^0 = \phi_m^{L+1} = 0 \quad . \quad (2.35)$$

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

$$\phi_m^k = a \cos(k\theta) + b \sin(k\theta) \quad , \quad 0 \leq k \leq L+1 \quad . \quad (2.36)$$

The two boundary conditions lead to

$$a = 0 \quad \text{and} \quad \sin((L+1)\theta) = 0 \quad (2.37)$$

yielding

$$\theta_m = \frac{m\pi}{L+1} \quad , \quad 1 \leq m \leq L \quad . \quad (2.38)$$

The eigenvectors are then determined by

$$\phi_m^k = b \sin(k\theta_m) \quad . \quad (2.39)$$

Combining this equation with Eq. (2.34) yields as the eigenvalues

$$\chi_m = 2 + 2 \cos \theta_m \quad . \quad (2.40)$$

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. (2.27) of the diffusion equation reduces to the largest eigenmode,

$$\tilde{n}(x, t) \simeq \exp(-\gamma_{de}t) B \sin\left(\frac{\pi}{L}x\right) \quad , \quad (2.41)$$

where

$$\gamma_{de} := \left(\frac{\pi}{L}\right)^2 D \quad (2.42)$$

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.(2.33) with Eq.(2.39),

$$\begin{aligned} \rho_n(x) &\simeq \exp(-\gamma_{FP}n) \tilde{B} \sin\left(\frac{\pi}{L+1}k\right) \quad , \\ k &= 0, \dots, L+1 \quad , \quad k-1 < x \leq k \end{aligned} \quad (2.43)$$

with an escape rate of this dynamical system given by

$$\gamma_{FP} = \ln \frac{4}{2 + 2 \cos(\pi/(L+1))} \quad , \quad (2.44)$$

which is determined by the largest eigenvalue χ_1 of the matrix $T(4)$, see Eq.(2.33) with Eq.(2.40). We can now see that the functional forms of the eigenmodes of Eqs.(2.41) and (2.43) match precisely.⁶ This in turn leads us to match Eqs. (2.42) and (2.44) leading to

$$D(4) = \left(\frac{L}{\pi}\right)^2 \gamma_{FP} \quad . \quad (2.45)$$

Using the right hand side of Eq. (2.44) and expanding it for $L \rightarrow \infty$, this formula enables us to calculate the diffusion coefficient $D(4)$ to

$$D(4) = \left(\frac{L}{\pi}\right)^2 \gamma_{FP} = \frac{1}{4} \frac{L^2}{(L+1)^2} + \mathcal{O}(L^{-4}) \rightarrow \frac{1}{4} \quad (L \rightarrow \infty) \quad . \quad (2.46)$$

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. (2.44), let us remind ourselves of the escape rate formula Eq. (1.51) for γ_{FP} ,

$$\gamma_{FP} = \gamma(\mathcal{C}_{B_4}) = \lambda(\mathcal{C}_{B_4}) - h_{KS}(\mathcal{C}_{B_4}) \quad , \quad (2.47)$$

which more generally expresses this escape rate in terms of dynamical systems quantities. Combining this equation with the above equation Eq. (2.45) leads to our final result, the *escape rate formula for deterministic diffusion* [Gas90, Gas98]

$$D(4) = \lim_{L \rightarrow \infty} \left(\frac{L}{\pi}\right)^2 [\lambda(\mathcal{C}_{B_4}) - h_{KS}(\mathcal{C}_{B_4})] \quad . \quad (2.48)$$

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

⁶We remark that there are discretization effects in the time and position variables, which are due to the fact that we compare a time-discrete system defined on a specific partition with time-continuous dynamics. They disappear in the limit of time to infinity by using a suitable spatial coarse graining.

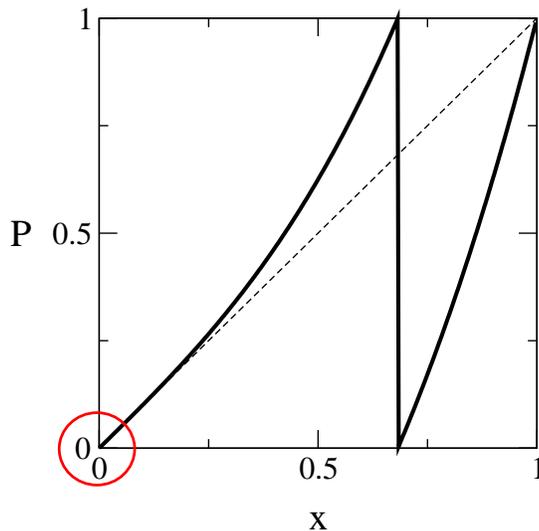


Figure 2.5: The Pomeau-Manneville map Eq. (2.49) for $a = 1$ and $z = 3$. Note that there is a marginal fixed point at $x = 0$ leading to the intermittent behavior depicted in Fig. 2.6.

Remark 3

1. Above we have only considered the special case of the control parameter $a = 4$. Along the same lines, the diffusion coefficient can be calculated for other parameter values of the map $B_a(x)$. Surprisingly, the parameter-dependent diffusion coefficient of this map turns out to be a fractal function of the control parameter [Kla96, Kla99]. This result is believed to hold for a wide class of dynamical systems [Kla07b].
2. The escape rate formula for diffusion does not only hold for simple one-dimensional maps but can be generalized to higher-dimensional time-discrete as well as time-continuous dynamical systems [Gas98, Dor99].
3. This approach can also be generalized to establish relations between chaos and transport for other transport coefficients such as viscosity, heat conduction and chemical reaction rates [Gas98].
4. This is not the only approach connecting transport properties with dynamical systems quantities. In recent research it was found that there exist at least two other ways to establish relations that are different but of a very similar nature; see [Kla07b] for further details.

2.3 Anomalous diffusion in intermittent maps

In the last section we have only explored diffusion for simple piecewise linear maps. One may now wonder what type of diffusive behaviour is encountered if we consider more general maps. Important examples are the nonlinear maps discussed in previous lectures generating *intermittency*. In this section we thus briefly recall what intermittency is about by giving an intuitive explanation for this phenomenon. We will then indicate how this property relates to *anomalous diffusion*, which defines a very active topic in recent research [Kla08].

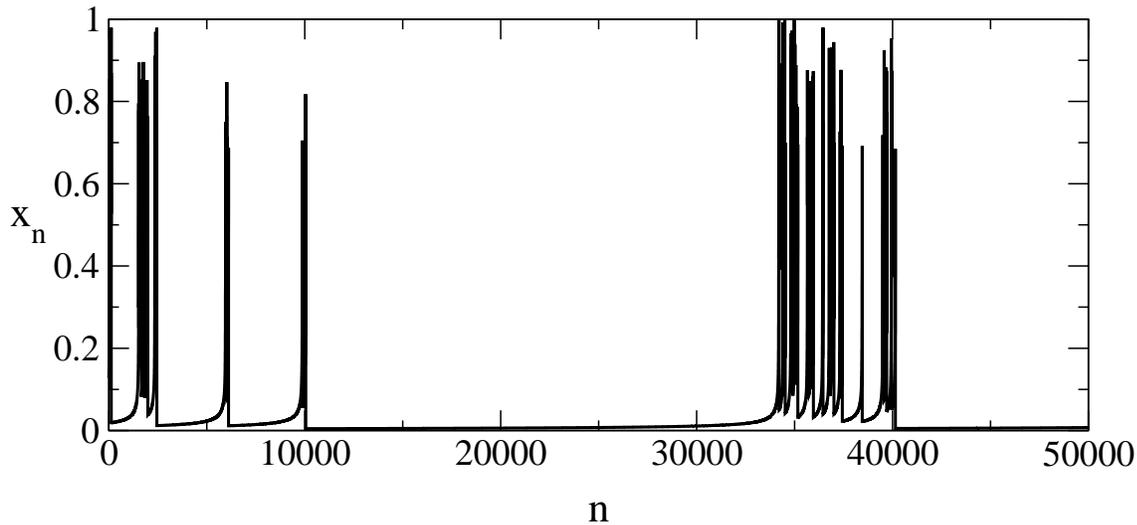


Figure 2.6: Phenomenology of intermittency in the Pomeau-Manneville map Fig. 2.5: The figure shows the time series, i.e., position x_n versus discrete time step n , of an orbit generated by the map Eq. (2.49) for a typical initial condition x_0 .

Let us consider an example of an intermittent map that is slightly different from what you have seen previously. It is the famous *Pomeau-Manneville map*, introduced in 1980 as a very simple model of intermittency [Pom80],

$$P_{a,z}(x) = x + ax^z \pmod{1} \quad , \quad (2.49)$$

see Fig. 2.5, where as usual the dynamics is defined by $x_{n+1} = P_{a,z}(x_n)$. This map has two control parameters, $a \geq 1$ and the exponent of nonlinearity $z \geq 1$. For $a = 1$ and $z = 1$ this map just reduces to our familiar Bernoulli shift, however, for $z > 1$ it provides a nontrivial nonlinear generalization of it. The nontriviality is due to the fact that in this case the stability of the fixed point at $x = 0$ becomes *marginal* (sometimes also called indifferent, neutral), $P'_{a,z}(0) = 1$. Since the map is smooth around $x = 0$, the dynamics resulting from the left lap of the map is largely determined by the stability of this fixed point, whereas the right lap is just of Bernoulli shift-type yielding ordinary chaotic dynamics. There is thus a competition in the dynamics between these two different branches, which explains the intermittent behavior depicted in Fig. 2.6: One can observe that long periodic *laminar phases* determined by the marginal fixed point around $x = 0$ are interrupted by *chaotic bursts* reflecting the “Bernoulli shift-like part” of the map with slope $a > 1$ around $x = 1$. This phenomenology is the hallmark of what one calls intermittent dynamics.

Following Section 2.1 it is now straightforward to define a spatially extended version of the Pomeau-Manneville map: For this purpose we just continue $P_{a,z}(x) = x + ax^z$, $0 \leq x < \frac{1}{2}$, cf. Eq. (2.49), onto the real line by the translation $P_{a,z}(x+1) = P_{a,z}(x) + 1$, see Eq. (2.3), under reflection symmetry $P_{a,z}(-x) = -P_{a,z}(x)$, see Eq. (2.4). The resulting model is displayed in Fig. 2.7.

As before, we may now be interested in the type of deterministic diffusion generated by this dynamical system. Surprisingly, by calculating the mean square displacement Eq. (2.2) for this model either analytically or from computer simulations one finds that for $z > 2$

$$\langle x^2 \rangle = K n^\alpha \quad , \quad \alpha < 1 \quad (n \rightarrow \infty) \quad . \quad (2.50)$$

Bibliography

- [All97] K.T. Alligood, T.S. Sauer and J.A. Yorke, *Chaos - An introduction to dynamical systems*. (Springer, New York, 1997).
- [Arn68] V.I. Arnold and A. Avez, *Ergodic problems of classical mechanics*. (W.A. Benjamin, New York, 1968).
- [Bec93] C. Beck and F. Schlögl, *Thermodynamics of Chaotic Systems*, volume 4 of Cambridge nonlinear science series. (Cambridge University Press, Cambridge, 1993).
- [Bun08] L. Bunimovich and A. Yurchenko, *Where to place a hole to achieve a maximal escape rate*. preprint arXiv.org:0811.4438, 2008.
- [Dev89] R.L. Devaney, *An introduction to chaotic dynamical systems*. (Addison-Wesley, Reading, second edition, 1989).
- [Dor99] J.R. Dorfman, *An introduction to chaos in nonequilibrium statistical mechanics*. (Cambridge University Press, Cambridge, 1999).
- [Eck85] J.-P. Eckmann and D. Ruelle, *Ergodic theory of chaos and strange attractors*. Rev. Mod. Phys. **57**, 617–656 (1985).
- [Fal90] K. Falconer, *Fractal geometry*. (Wiley, New York, 1990).
- [Fal05] M. Falcioni, L. Palatella and A. Vulpiani, *Production rate of the coarse-grained Gibbs entropy and the Kolmogorov-Sinai entropy: A real connection?* Phys. Rev. E **71**, 016118/1–8 (2005).
- [Gas90] P. Gaspard and G. Nicolis, *Transport properties, Lyapunov exponents, and entropy per unit time*. Phys. Rev. Lett. **65**, 1693–1696 (1990).
- [Gas98] P. Gaspard, *Chaos, scattering, and statistical mechanics*. (Cambridge University Press, Cambridge, 1998).
- [Kan85] H. Kantz and P. Grassberger, *Repellers, semi-attractors, and long-lived chaotic transients*. Physica D **17**, 75–86 (1985).
- [Kat95] A. Katok and B. Hasselblatt, *Introduction to the modern theory of dynamical systems*, volume 54 of Encyclopedia of Mathematics and its Applications. (Cambridge University Press, Cambridge, 1995).
- [Kla96] R. Klages, *Deterministic diffusion in one-dimensional chaotic dynamical systems*. (Wissenschaft & Technik-Verlag, Berlin, 1996).

-
- [Kla99] R. Klages and J.R. Dorfman, *Simple deterministic dynamical systems with fractal diffusion coefficients*. Phys. Rev. E **59**, 5361–5383 (1999).
- [Kla07a] R. Klages, *Introduction to dynamical systems*. Lecture notes, see <http://www.maths.qmul.ac.uk/~klages/teaching/mas424>, 2007.
- [Kla07b] R. Klages, *Microscopic chaos, fractals and transport in nonequilibrium statistical mechanics*, volume 24 of Advanced Series in Nonlinear Dynamics. (World Scientific, Singapore, 2007).
- [Kla08] R. Klages, G. Radons and I.M. Sokolov, Eds., *Anomalous transport*. (Wiley-VCH, Berlin, 2008).
- [Kla10] R. Klages, *From Deterministic Chaos to Anomalous Diffusion*. in: H.G. Schuster, Eds., Reviews of Nonlinear Dynamics and Complexity, volume 3, pages 169–227, Wiley-VCH, New York, 2010.
- [Las94] A. Lasota and M.C. Mackay, *Chaos, Fractals, and Noise*. (Springer-Verlag, 1994).
- [Man82] B.B. Mandelbrot, *The fractal geometry of nature*. (W.H. Freeman and Company, San Francisco, 1982).
- [Ott93] E. Ott, *Chaos in Dynamical Systems*. (Cambridge University Press, Cambridge, 1993).
- [Pom80] Y. Pomeau and P. Manneville, *Intermittent transition to turbulence in dissipative dynamical systems*. Commun. Math. Phys. **74**, 189–197 (1980).
- [Rei65] F. Reif, *Fundamentals of statistical and thermal physics*. (McGraw-Hill, Auckland, 1965). see also the amended German translation: Statistische Physik und Theorie der Wärme, 3rd edition, de Gruyter, Berlin, 1987.
- [Rob95] C. Robinson, *Dynamical Systems*. (CRC Press, London, 1995).
- [Sch89] H.G. Schuster, *Deterministic chaos*. (VCH Verlagsgesellschaft mbH, Weinheim, 2nd edition, 1989).
- [Tod92] M. Toda, R. Kubo and N. Saitô, *Statistical Physics*, volume 1 of Solid State Sciences. (Springer, Berlin, 2nd edition, 1992).
- [Tri95] C. Tricot, *Curves and fractal dimension*. (Springer, Berlin, 1995).
- [You02] L.-S. Young, *What are SRB measures, and which dynamical systems have them?* J. Stat. Phys. **108**, 733–754 (2002).