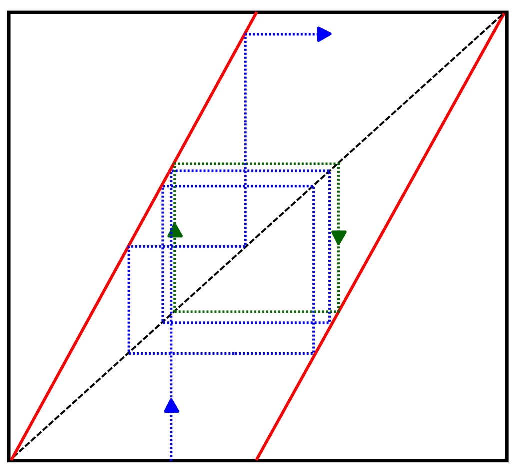
Queen Mary
University of London

Introduction to Dynamical Systems

Lecture Notes for MAS424/MTHM021
Version 1.2, 18/04/2008



Rainer Klages
School of Mathematical Sciences
Queen Mary, University of London
Notes for version 1.1 compiled by Phil Howard

Contents

1 Preliminaries **5**

I What is a dynamical system? **7**

2 Examples of realistic dynamical systems **7**
2.1 Driven nonlinear pendulum 7
2.2 Bouncing ball 9
2.3 Particle billiards 9

3 Definition of dynamical systems and the Poincaré-Bendixson theorem **12**
3.1 Time-continuous dynamical systems 12
3.2 Time-discrete dynamical systems 16
3.3 Poincaré surface of section 19

II Topological properties of one-dimensional maps **23**

4 Some basic ingredients of nonlinear dynamics **24**
4.1 Homeomorphisms and diffeomorphisms 26
4.2 Periodic points 27
4.3 Dense sets, Bernoulli shift and topological transitivity 29
4.4 Stability 35

5 Definitions of deterministic chaos **37**
5.1 Devaney chaos 37
5.2 Sensitivity and Wiggins chaos 38
5.3 Ljapunov chaos 39
5.4 Summary 42

III Probabilistic description of one-dimensional maps **44**

6 Dynamics of statistical ensembles **44**
6.1 The Frobenius-Perron equation 45
6.2 Markov partitions and transition matrices 50

7 Measure-theoretic description of dynamics **60**
7.1 Probability measures 60
7.2 Basics of ergodic theory 69

Bibliography **75**

1 Preliminaries

There exist *two essentially different approaches to the study of dynamical systems*, based on the following distinction:

time-continuous nonlinear differential equations $\rightleftharpoons$ time-discrete maps

One approach starts from time-*continuous* differential equations and leads to time-*discrete* maps, which are obtained from them by a suitable discretization of time. This path is pursued, e.g., in the book by Strogatz [Str94].[1] The other approach starts from the study of time-discrete maps and then gradually builds up to time-continuous differential equations, see, e.g., [Ott93, All97, Dev89, Has03, Rob95]. After a short motivation in terms of nonlinear differential equations, for the rest of this course we shall follow the latter route to dynamical systems theory. This allows a generally more simple way of introducing the important concepts, which can usually be carried over to a more complex and physically realistic context.

As far as the *style of these lectures* is concerned, it is important to say that this course, and thus these notes, are presented not in the spirit of a pure but of an applied mathematician (actually, of a mathematically minded theoretical physicists). That is, we will keep technical mathematical difficulties to an absolute minimum, and if we present any proofs they will be very short. For more complicated proofs or elaborate mathematical subtleties we will usually refer to the literature. In other words, our goal is to give a rough outline of crucial concepts and central objects of this theory as we see it, as well as to establish crosslinks between dynamical systems theory and other areas of the sciences, rather than dwelling upon fine mathematical details. If you wish, you may consider this course as an applied follow-up of the 3rd year course MAS308 Chaos and Fractals.

That said, it is also not intended to present an introduction to the *context and history* of the subject. However, this is well worth studying, the field now encompassing over a hundred years of activity. The book by Gleick [Gle96] provides an excellent starting point for exploring the historical development of this field. The very recent book by Smith [Smi07] nicely embeds the modern theory of nonlinear dynamical systems into the general socio-cultural context. It also provides a very nice popular science introduction to basic concepts of dynamical systems theory, which to some extent relates to the path we will follow in this course.

This course consists of *three main parts*: The introductory Part I starts by exploring some examples of dynamical systems exhibiting both simple and complicated dynamics. We then discuss the interplay between time-discrete and time-continuous dynamical systems in terms of Poincaré surfaces of section. We also provide a first rough classification of different types of dynamics by using the Poincaré-Bendixson theorem. Part II introduces

[1] see also books by Arrowsmith, Percival and Richards, Guckenheimer and Holmes

elementary topological properties of one-dimensional time-discrete dynamical systems, such as periodic points, denseness and stability properties, which enables us to come up with rigorous definitions of deterministic chaos. This part connects with course MAS308 but pushes these concepts a bit further. Part III finally elaborates on the probabilistic, or statistical, description of time-discrete maps in terms of the Frobenius-Perron equation. For this we need concepts like (Markov) partitions, transition matrices and probability measures. We conclude with a brief outline of essentials of ergodic theory. If you are interested in further pursuing these topics, please note that there is a strong research group at QMUL particularly focusing on (ergodic properties of) dynamical systems with crosslinks to statistical physics.[2]

The format of these notes is currently somewhat sparse, and it is expected that they will require substantial annotation to clarify points presented in more detail during the actual lectures. Please treat them merely as a study aid rather than a comprehensive syllabus.

[2] see http://www.maths.qmul.ac.uk/~mathres/dynsys for further information

Part I

What is a dynamical system?

2 Examples of realistic dynamical systems

2.1 Driven nonlinear pendulum

Figure 2.1 shows a pendulum of mass M subject to a torque (the rotational equivalent of a force) and to a gravitational force $\underline{G}$. You may think, for example, of a clock pendulum or a driven swing. The angle with the vertical in a positive sense is denoted by $\theta = \theta(t)$, where $t \in \mathbb{R}$ holds for the time of the system, and we choose $-\pi \leq \theta < \pi$.

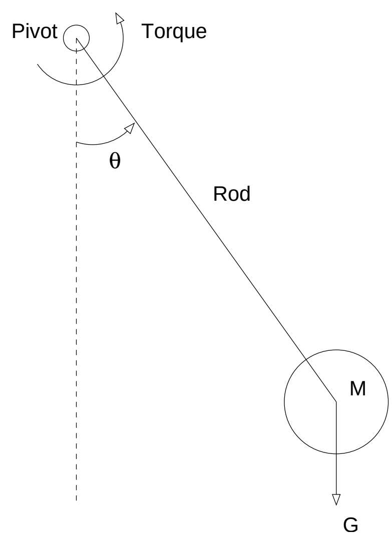


Figure 2.1: Driven pendulum of mass M with a torque applied at the pivot and subject to gravity.

Without worrying too much about how one can use physics to obtain an equation of motion for this system starting from Newton's equation of motion, see [Ott93, Str94] for such derivations, we move straight to the equation itself and merely indicate whence each term arises:

$$\begin{array}{lccccccccc} \text{Equation of motion:} & \ddot{\theta} & + & \nu\dot{\theta} & + & \sin\theta & = & A\sin(2\pi f t) & & \\ & \uparrow & & \uparrow & & \uparrow & & \uparrow & , & (2.1) \\ \text{Balance of forces:} & inertia & + & friction & + & gravity & = & periodic\,torque & & \end{array}$$

where for sake of simplicity we have set the mass M equal to one. Here we write $\dot{\theta} := \frac{d\theta}{dt}$ to denote the derivative of θ with respect to time, which is also sometimes called the angular velocity. In (2.1) θ is an example of a *dynamical variable* describing the state of the system, whereas ν, A, f are called *control parameters*. Here ν denotes the friction coefficient, A the amplitude of the periodic driving and f the frequency of the driving force. In contrast to dynamical variables, which depend on time, the values for the control parameters are chosen once for the system and then kept fixed, that is, they do not vary in time. Equation (2.1) presents an example of a *driven* (driving force), *nonlinear* (because of the sine function, $\sin x \simeq x - x^3/3!$), *dissipative* (because of driving and damping) dynamical system.
It is generally impossible to analytically solve complicated nonlinear equations of motion such as (2.1). However, they can still be integrated by numerical methods (such as Runge-Kutta integration schemes), which allows the production of simulations such as the ones that can be explored in "The Pendulum Lab", a very nice interactive webpage [Elm98]. Playing around with varying the values of control parameters there, one finds the following four different types of characteristic behaviour: This systems has already been studied in many experiments, even by high school pupils!

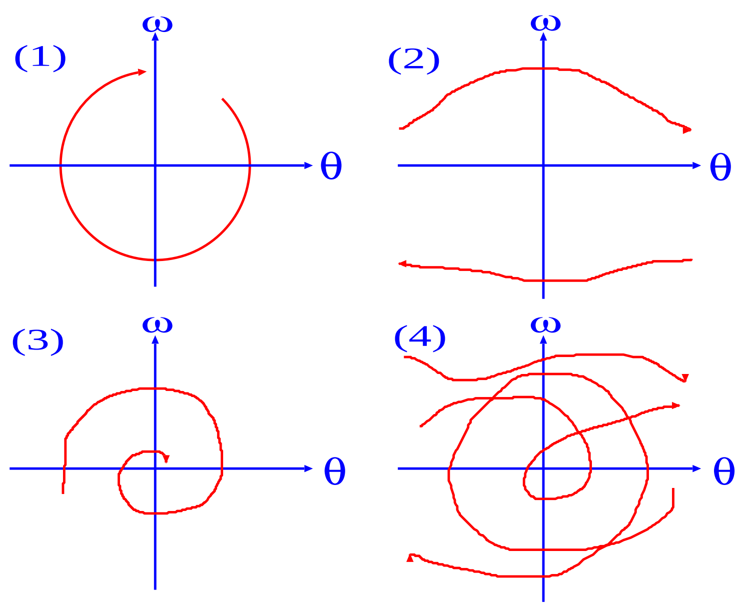


Figure 2.2: Plots of four different types of motion of the driven nonlinear pendulum (2.1) in the (θ, ω)-space under variation of control parameters.

1. No torque $A = 0$, no damping $\nu = 0$, small angles $\theta \ll 1$ leading to

$$\ddot{\theta} + \theta = 0 \quad . \tag{2.2}$$

This system is called the *harmonic oscillator*, and the resulting dynamics is known as *libration*. If we represent the dynamics of this system in a coordinate system where we plot θ and the angular velocity $\omega := \dot{\theta}$ as functions of time, we obtain pictures which look like the one shown in Fig. 2.2 (a).

2. same as 1. except that θ can be arbitrarily large, $-\pi \leq \theta < \pi$: Fixed points exist at $\theta = 0$ (stable) and $\theta = \pi$ (unstable). If the initial condition $\theta(0) = \pi$ is taken with non-zero initial velocity $\omega \neq 0$, continual *rotation* is obtained; see Fig. 2.2 (b).

3. same as 2. except that the damping is not zero anymore, $\nu \neq 0$: The pendulum comes to *rest* at the stable fixed point $\theta = 0$ for $t \to \infty$. This is represented by some spiraling motion in Fig. 2.2 (c).

4. same as 3. except that now we have a periodic driving force, that is, $A, f \neq 0$: In this case we explore the dynamics of the full driven nonlinear pendulum equation (2.1). We observe a 'wild', rather unpredictable, *chaotic-like* dynamics in Fig. 2.2 (d).

We conclude this discussion by mentioning that the driven nonlinear pendulum is a paradigmatic example of a non-trivial dynamical system, which also displays chaotic behavior. It is found in many physical and other systems such as in Josephson junctions (which are microscopic semi-conductor devices), in the motion of atoms and planets and in biological clocks. It is also encountered in many engineering problems.

2.2 Bouncing ball

Another seemingly simple system displaying non-trivial behavior is the bouncing ball schematically depicted in Fig. 2.3: The figure shows an elastic ball of mass M falling under a gravitational force $\underline{G}$ by bouncing off a plate that vibrates with amplitude A and frequency f, thus allowing transfer of energy to and from the ball system. Additionally, the ball experiences a friction ν at the collision.
Without going into further detail here, such as writing down the systems' equations of motion, we may mention that for certain values of the control parameters this system exhibits simple periodic behavior in form of 'frequency locking', where the periodic motion of the bouncing ball and of the vibrating plate are in phase. This is like a ping-pong ball hopping vertically on your oscillating table tennis racket. In other words, here we have a certain type of 'resonance'. However, under smooth variation of control parameters such as amplitude or frequency of the driving one typically observes a transition to periodic motion of increasingly higher order (called "bifurcations") until the motion eventually becomes completely irregular, or "chaotic".[1]

2.3 Particle billiards

A third example is provided by the following system, see Fig. 2.4: Let us consider a hard disk of radius r, whose centre is fixed in the middle of a square box in the plane. We now

[1]see N.Tufillaro's webpage http://www.drchaos.net/drchaos/bb.html or the one by P.Pieranski http://fizyka.phys.put.poznan.pl/~pieransk/BouncigBall.html for further details; see also the book by Tél and Gruiz

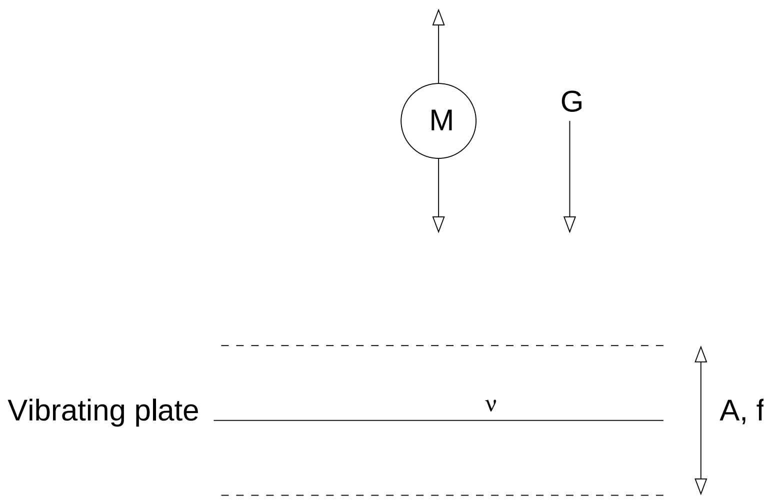


Figure 2.3: A ball of mass M subject to gravity, that elastically bounces off a vibrating plate.

look at the dynamics of a point particle with constant speed that collides elastically with the disk and the walls of the box. That is, we have specular reflection at the walls where in- and outgoing angles are the same, $\theta = \theta'$.[2] This system strongly reminds one of the idealistic case of a billiard table (without friction but with a circular obstacle) and is indeed referred to as a *particle billiard* in the literature. The particular example shown in Fig. 2.4 is known as the *Sinai billiard*.

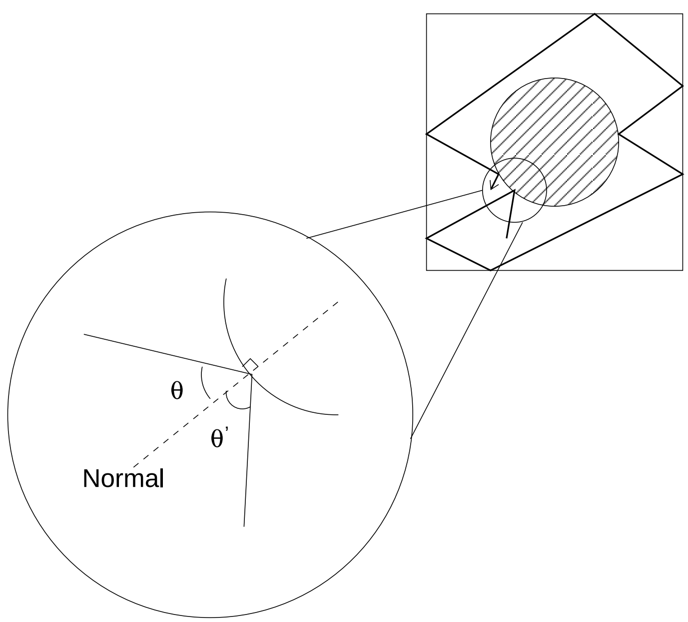


Figure 2.4: The Sinai billiard.

One may now ask the question about *sensitivity of initial conditions*, that is: What happens

[2]see http://www.dynamical-systems.org/sinai/index.html for a simulation (with optional gravity)

to two nearby particles in this billiard (which do not interact with each other) with slightly different directions of their initial velocities? Since at the moment no computer simulation of this case is available to us, we switch to a slightly more complicated dynamical system as displayed in Fig. 2.5, where we can numerically explore this situation [Miy03].

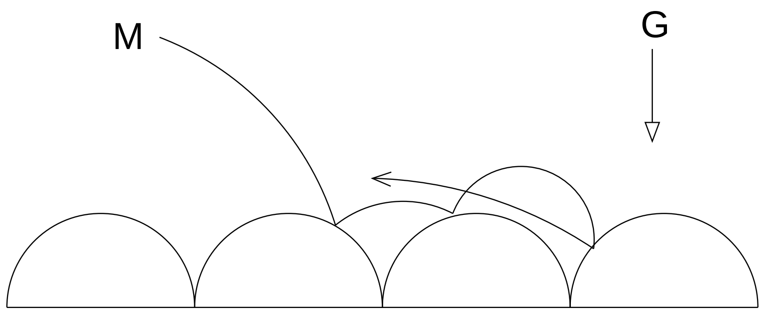


Figure 2.5: A billiard where a point particle bounces off semicircles on a plate.

The dynamical system is defined as follows: we have a series of semicircles periodically continued onto the line, which may overlap with each other. A point particle of mass M now scatters elastically with these semicircles under the influence of a gravitational force $\underline{G}$. In the simulation we study the spreading in time of an ensemble of particles starting from the same point, but with varied velocity angles. The result is schematically depicted in Fig. 2.6: We see that there are two important mechanisms determining the dynamics of the particles, namely a *stretching*, which initially is due to the choice of initial velocities but later on also reflects the dispersing collisions at the scatterers, and a *folding* at the collisions, where the front of propagating particles experiences cusps. This sequence of “stretch” and “fold” generates very complicated structures in the position space of the system, which look like mixing paint.

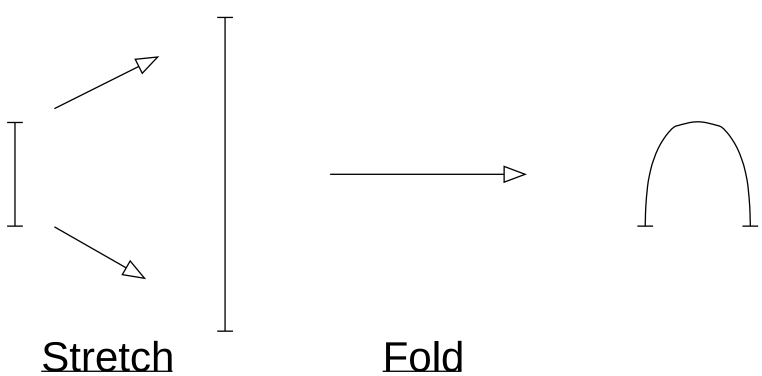


Figure 2.6: Schematic representation of the stretch and fold mechanism of an ensemble of particles in a chaotic dynamical system.

As we shall see later in this course, this is one of the fundamental mechanisms of what is called “chaotic behavior” in nonlinear dynamical systems. The purpose of these lectures is to put the handwaving assessment of the behavior illustrated above for some “realistic” dynamical systems onto a more rigorous mathematical basis such that eventually we can answer the question what it means to say that a system exhibits "chaotic" dynamics. To the end of this course we will also consider the dynamics of statistical ensembles of particles such as in the simulation.

3 Definition of dynamical systems and the Poincaré-Bendixson theorem

Definition 1 *[Las94, Kat95]*[1] *A* dynamical system *consists of a* phase (or state) space P *and a family of transformations* $\underline{\phi}_t : P \to P$*, where the time* t *may be either discrete,* $t \in \mathbb{Z}$*, or continuous,* $t \in \mathbb{R}$*. For arbitrary* states $\underline{x} \in P$ *the following must hold:*

1. $\underline{\phi}_0(\underline{x}) = \underline{x}$ identity *and*
2. $\underline{\phi}_t(\underline{\phi}_s(\underline{x})) = \underline{\phi}_{t+s}(\underline{x}) \; \forall t, s \in \mathbb{R}$ additivity

In other words, a dynamical system may be understood as a mathematical prescription for evolving the state of a system in time [Ott93, All97]. Property 2 above ensures that the transformations $\underline{\phi}_t$ form an Abelian group. As an exercise, you may wish to look up different definitions of dynamical systems on the internet.

3.1 Time-continuous dynamical systems

Definition 2 *Let* $P \subset \mathbb{R}^N, N \in \mathbb{N}, \underline{x} = (x^1, x^2, \ldots, x^N) \in P, t \in \mathbb{R}$*. Then*

$$\underline{F} : P \to P \quad , \quad \dot{\underline{x}} = \underline{F}(\underline{x}(t)) = \underline{F}(\underline{x}) \tag{3.1}$$

is called a vector field*. It can be written as a system of* N first order, autonomous *(i.e., not explicitly time-dependent),* ordinary *differential equations,*

$$\begin{aligned} \frac{\mathrm{d}x^1}{\mathrm{d}t} &= F^1(x^1, x^2, \ldots, x^N) \\ \frac{\mathrm{d}x^2}{\mathrm{d}t} &= F^2(x^1, x^2, \ldots, x^N) \\ &\vdots \\ \frac{\mathrm{d}x^N}{\mathrm{d}t} &= F^N(x^1, x^2, \ldots, x^N) \quad . \end{aligned} \tag{3.2}$$

The formal solution of Eq. (3.1) (if there exists any),

$$\underline{x}(t) = \underline{\phi}_t(\underline{x}(0)) \quad , \tag{3.3}$$

is called the flow *of the vector field.*

[1] see also books by Bronstein, Reitmann

Here $\underline{\phi}_t$ is the transformation that we have already encountered in Definition 1 above. Note that this does not answer the question of how to construct the flow for arbitrary initial conditions $\underline{x}(0)$.

Definition 3 *A* single path *in phase space followed by* $\underline{x}(t)$ *in time is called the* trajectory *or* orbit *of the dynamical system.*[2]

See Fig. 3.1 for an example.

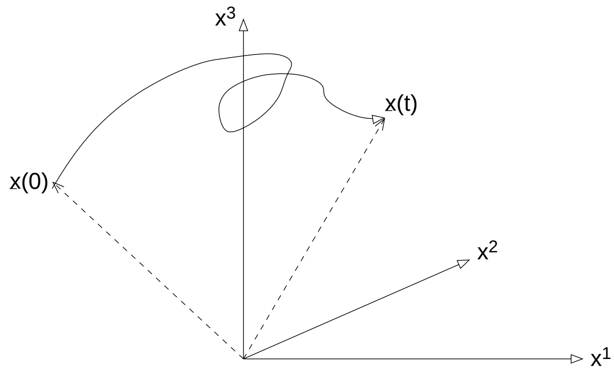


Figure 3.1: Example of a trajectory in 3-dimensional phase space (note that the trajectory is not supposed to cross itself).

In the following we will assume the existence and uniqueness of solutions of the vector field $\underline{F}$. A proof of this exists if $\underline{F}$ is smooth (continuously differentiable, C^1) [Str94].

Corollary 1 *For smooth* $\underline{F}$ *two distinct trajectories cannot intersect, nor can a single trajectory cross itself for* $t < \infty$.

Proof: Assume a single trajectory crosses itself, see Fig. 3.2. Starting with the initial condition at the point of intersection there are two choices of direction to proceed in, which contradicts uniqueness. The same argument applies to two distinct trajectories. q.e.d.

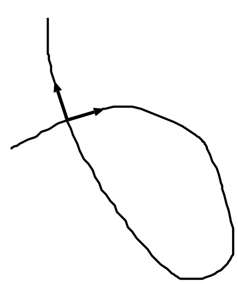


Figure 3.2: A trajectory that crosses itself.

If $\exists$ uniqueness, $\underline{x}(0)$ *determines* the outcome of a flow after time t: we have what is called *determinism*.

[2]In this course we use both words synonymously.

One may now ask about the general 'character' of a flow $\phi_t(\underline{x})$, intuitively speaking: Is it 'simple' or 'complicated'? In the following we give a first, rather qualitative answer.

Definition 4 *We call a flow* simple *if for $t \to \infty$ all $\underline{x}(t)$ are either* fixed points $\underline{x} =$ *constant, respectively $\dot{\underline{x}} = \underline{0}$, or* periodic orbits*, i.e., closed loops $\underline{x}(t + \tau) = \underline{x}(t), \tau \in \mathbb{R}$.*

Two examples of fixed points are shown in Figs. 3.3: In both cases the fixed point is located at the origin of the coordinate system in the two-dimensional plane. However, if we look at initial conditions in an environment around these fixed points and how the trajectory determined by it evolves in time, we may observe different types of behavior: For example, in one case the trajectory 'spirals in' to the fixed point by approaching it time-asymptotically, whereas in the other case the trajectory 'spirals out'.
We remark that these are by far not the only cases of possible dynamics [Str94]. So we see that a fixed point, apart from its mere existence, can have a very different impact onto its environment: In the first case we speak of a *stable* fixed point, whereas in the second case the fixed point is said to be *unstable*.

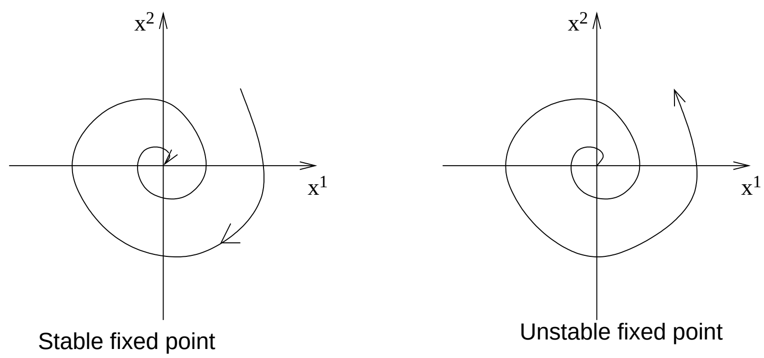


Figure 3.3: stable and unstable fixed points

The same reasoning applies to the simple example of a periodic orbit shown in Fig. 3.4: The left hand side depicts a circular periodic orbit. However, if we choose initial conditions that are not on this circle we may observe, for example, the behavior illustrated on the right hand side of this figure: trajectories 'spiral in' onto the periodic orbit both from the interior of the circle and from the exterior. An isolated closed trajectory such as this periodic orbit is called a *limit cycle*, which in this case is *stable*.
Of course, as for the fixed point the opposite case is also possible, that is, the limit cycle is *unstable* if all nearby trajectories spiral out (you may wish to draw a figure of this). The detailed determination and classification of fixed points and periodic orbits is in the focus of what is called *linear stability analysis* in the literature, see [Ott93, Str94] in case of flows. For maps we will learn right this later in the course.

Definition 5 *We call a flow* complicated *if it is not simple.*

Typically, such a flow cannot be calculated analytically anymore, because it cannot be represented in a simple functional form. In this case solutions can only be obtained numerically. Later on we will introduce 'chaos' as a subset of complicated solutions.[3]

[3] If 'chaos' defines a subset of complicated dynamics, this leaves the possibility of solutions that are

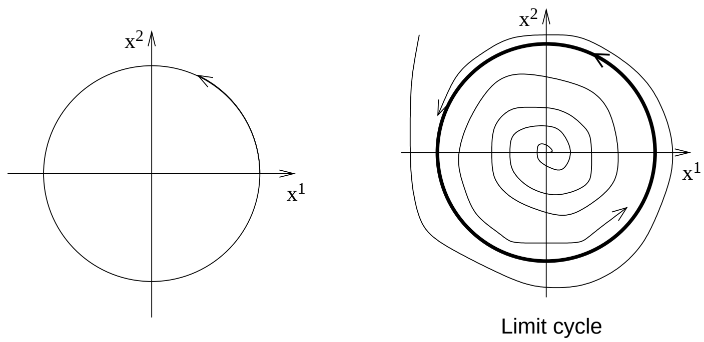


Figure 3.4: left: circular periodic orbit; right: approach to this orbit as a limit cycle

One may now ask the further question of how large the dimensionality N of the phase space has to be for complicated behaviour to be *possible*. The answer is given in terms of the following important theorem:[Str94, Rob95][4]

Theorem 1 *(Poincaré-Bendixson) Let* $\underline{\dot{x}} = \underline{F}(\underline{x})$ *be a smooth vector field acting on an open set containing a closed, bounded set* $R \subset \mathbb{R}^2$*, which is such that all trajectories starting in* R *remain in* R*. Then any trajectory starting in* R *is either a fixed point or a periodic orbit, or it spirals to one as* $t \to \infty$*.*

The proof of this theorem is elaborate and goes beyond the scope of this course. It is based on the idea of uniqueness of solutions and that trajectories cannot cross each other; see, e.g., [All97]. Under these conditions, the different topology in two and three dimensions plays a crucial role as is demonstrated in Fig. 3.5. See also [Str94] for more detailed discussions.

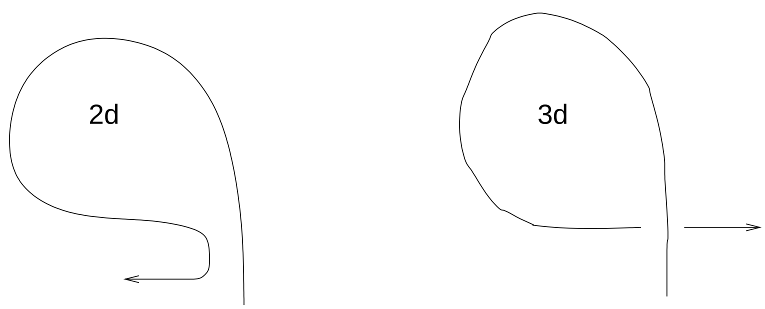


Figure 3.5: Trajectories in 2-d (left) cannot cross whereas in 3-d (right), they can.

Corollary 2 *Let* $\underline{F}(\underline{x}) \in \mathbb{R}^N$ *and the conditions of the Poincaré-Bendixson theorem be fullfilled. Then* $N \leq 2 \Rightarrow$ *solutions* simple*. Consequently,* $N \geq 3 \Rightarrow$ *'anything can happen', i.e., complicated solutions are* possible*.*

This raises the question of how to check these statements for a given differential equation, which we discuss for an example:

neither simple nor chaotic but nevertheless complicated. Such 'weakly chaotic motion' is a very active topic of recent research, see, e.g., Section 17.4 of [Kla07].

[4] see also a book by Hilborn

Example 1 Let us revisit the driven nonlinear pendulum that we have seen before,

$$\ddot{\theta} + \nu\dot{\theta} + \sin\theta = A\sin(2\pi f t) \quad .$$

Rewrite this dynamical system in the form of a vector field by using

$$x^1 := \dot{\theta}, \qquad x^2 := \theta, \qquad x^3 := 2\pi f t \quad .$$

The 'trick' of incorporating the time dependence of the differential equation as a third variable allows us to make the vector field autonomous leading to

$$\begin{aligned} \dot{x}^1 &= \ddot{\theta} = -\nu x^1 - \sin x^2 + A\sin x^3 \\ \dot{x}^2 &= x^1 \\ \dot{x}^3 &= 2\pi f \quad . \end{aligned} \tag{3.4}$$

Therefore $N = 3$: complicated dynamics is *possible*, as we have indeed seen in the simulations.

3.2 Time-discrete dynamical systems

Definition 6 *Let* $P \subset \mathbb{R}^N, N \in \mathbb{N}, \underline{x}_n \in P, n \in \mathbb{Z}$. *Then*

$$\underline{M} : P \to P \quad , \quad \underline{x}_{n+1} = \underline{M}(\underline{x}_n) \tag{3.5}$$

is called a time-discrete map. $\underline{x}_{n+1} = \underline{M}(\underline{x}_n)$ *are sometimes called the* equations of motion *of the dynamical system.*

Choosing the initial condition $\underline{x}_0$ *determines* the outcome after n discrete time steps (hence *determinism*) in the following way:

$$\begin{aligned} \underline{x}_1 &= \underline{M}(\underline{x}_0) = \underline{M}^1(\underline{x}_0), \\ \underline{x}_2 &= \underline{M}(\underline{x}_1) = \underline{M}(\underline{M}(\underline{x}_0)) = \underline{M}^2(\underline{x}_0)(\neq \underline{M}(\underline{x}_0)\underline{M}(\underline{x}_0)!). \\ \Rightarrow \underline{M}^m(\underline{x}_0) &:= \underbrace{\underline{M} \circ \underline{M} \circ \cdots \underline{M}(\underline{x}_0)}_{\text{m-fold composed map}} \quad . \end{aligned} \tag{3.6}$$

In other words, for maps the situation is formally simpler than for differential equations: $\exists$ a unique (why?) solution to the equations of motion in form of $\underline{x}_n = \underline{M}(\underline{x}_{n-1}) = \ldots = \underline{M}^n(\underline{x}_0)$. This is the counterpart of the flow for time-continuous systems.

Example 2 [Ott93] For $N = 1$ let $z_n \in \mathbb{N}_0$ be the number of insects hatching out of eggs in year $n \in \mathbb{N}_0$. Let $r > 0$ be the average number of eggs laid per insect. In an (for the insects) 'ideal' case we have

$$z_{n+1} = r z_n = r^2 z_{n-1} = \ldots = r^{n+1} z_0 = \exp((n+1)\ln r) z_0 \quad , \tag{3.7}$$

where we assume that insects live no longer than for one year. This straightforwardly implies for $r > 1$ an exponentially increasing and for $r < 1$ an exponentially decreasing population.

In 1798 Malthus proposed to incorporate the effect of enemies, or overcrowding, by replacing the control parameter r through the *logistic growth* $r(1-\frac{z_n}{\tilde{z}})$, where $\tilde{z}$ models a limited food supply with $0 \leq z_n \leq \tilde{z}$. That is, if $z_n = \tilde{z}$ the food is exhausted and all insects die. Replacing r after the first equal sign of Eq. (3.7) above leads to the new equation

$$z_{n+1} = r z_n (1 - \frac{z_n}{\tilde{z}}) \quad . \tag{3.8}$$

Dividing both sides by $\tilde{z}$ and defining the new variable $x_n := \frac{z_n}{\tilde{z}}$ yields the famous *logistic map*

$$x_{n+1} = r x_n (1 - x_n) \quad , \tag{3.9}$$

where here we restrict ourselves to $0 \leq x_n \leq 1$, $0 < r \leq 4$.

How do we get information about the dynamics of this map? We could look, for example, at the set of all iterates for a given initial condition x_0, i.e. $\{x_0, x_1, x_2, \ldots, x_n\}$, which defines the *trajectory* or *orbit* of $M(x_0)$.
Note that the same definition carries over to N-dimensional maps. However, in one dimension we have a nice graphical representation of the trajectory in form of the *cobweb plot*, as we may demonstrate for our above example:

Example 3 Cobweb plot for the logistic map restricted to the parameter range $1 < r < 2$ and for $0 \leq x \leq 1$, see Fig. 3.6.

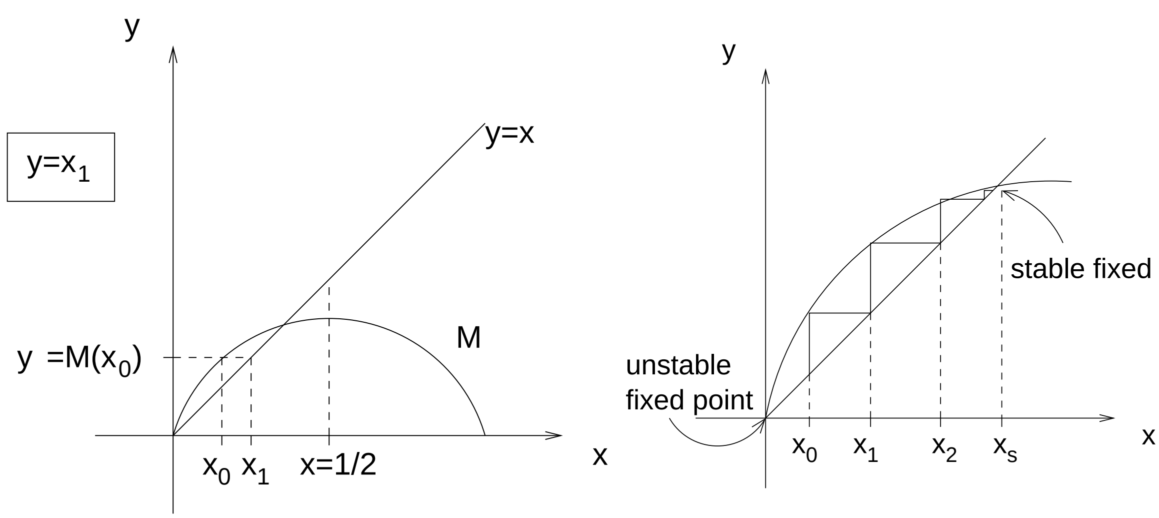


Figure 3.6: left: first step of a cobweb plot for the logistic map with $1 < r < 2$; right: magnification of the left hand side with the result for the iterated procedure.

Constructing a cobweb plot for a map proceeds in five steps according to the following algorithm:

1. Choose an initial point x_0 and draw a vertical line until it intersects the graph of the chosen map $M(x)$. The point of intersection defines the value $y = M(x_0)$ of the map after the first iteration.
2. By using Eq. (3.5) we now identify y with the next position in the domain of the map at time step $n = 1$ which leads to $x_1 = y$.

3. This is represented in the figure by drawing a horizontal line from the previous point of intersection until it intersects with the diagonal $y = x$.
4. Drawing a vertical line from the new point of intersection down to the x-axis makes clear that we have indeed graphically obtained the new position x_1 on the abscissa.
5. However, in practice this last vetical line is neglected for making a cobweb plot. Instead, we immediately construct the next iterate by repeating step 1 to 3 discussed above for the new starting point x_1 instead of x_0, and so on.

In summary, for a cobweb plot we continuously alternate between vertical and horizontal lines that intersect with the graph of the given map and the diagonal; see also p.5 of [All97] and [Has03] for details.

Definition 7 *$\underline{x}$ is a* fixed point *of $\underline{M}$ if $\underline{x} = \underline{M}(\underline{x})$.*

For our one-dimensional example of the logistic map our cobweb plot Fig. 3.6 immediately tells us that we have two fixed points, one at $x_u = 0$ and another one at a (yet unspecified) larger value x_s. Interestingly, the plot furthermore suggests that all points in the neighbourhood of x_u move away from this fixed point and converge to x_s. Hence, we may call x_u an *unstable* or *repelling* fixed point whereas we classify x_s as being *stable* or *attracting*.
In other words, a cobweb plot yields straightforwardly information both about the existence of fixed points, which in one dimension are just the intersections $x = M(x)$, and their stability. Of course, the above statement provides only a qualitative assessment of the stability of fixed points. We will make this mathematically rigorous later on.

Definition 8 *A map $\underline{M}$ is called invertible if $\exists$ an inverse $\underline{M}^{-1}$ with $\underline{x}_n = \underline{M}^{-1}(\underline{x}_{n+1})$.*

Example 4

1. The logistic map is not invertible, since one image has two preimages, as you may convince yourself by a simple drawing.

2. Let us introduce the two-dimensional map

$$\begin{aligned} x_{n+1} &= f(x_n) - ky_n \ , \ k \neq 0 \\ y_{n+1} &= x_n \quad . \end{aligned} \tag{3.10}$$

   Here $f(x)$ is just some function; as an example, let us choose $f(x) = A - x^2, A \in \mathbb{R}$. Eq. (3.10) together with this $f(x)$ defines the famous Hénon map (1976), derived from the (Hamiltonian) equations of motion for a star in a galaxy (1964) [Ott93].

   Let us try to invert this mapping: We get

$$\begin{aligned} x_n &= y_{n+1} \\ y_n &= \frac{1}{k}(f(x_n) - x_{n+1}) \\ &= \frac{1}{k}(f(y_{n+1}) - x_{n+1}) \quad , \end{aligned} \tag{3.11}$$

   so we can conclude that the map is invertible.

We may now be wondering whether there is any condition on N for the existence of 'complicated solutions' in time-discrete maps, in analogy to Corollary 2 for time-continuous dynamical systems. The answer is given as follows:

Corollary 3 *[Ott93] (Poincaré-Bendixson discretized) Let $\underline{x}_{n+1} = \underline{M}(\underline{x}_n)$ be a smooth map acting on an open set containing a closed, bounded set $R \subset \mathbb{R}^N$, which is such that all trajectories starting in R remain in R. Let us look at trajectories starting in R for $n \to \infty$. Let $\underline{M}$ be invertible. Then $N = 1 \Rightarrow$ solutions* simple. *Consequently, $N \geq 2 \Rightarrow$ complicated solutions are* possible. *Let $\underline{M}$ be noninvertible. Then complicated solutions are always* possible.

So the discretized Poincaré-Bendixson theorem leaves us with the possibility that the dynamics even of simple one-dimensional maps is non-trivial, which is the reason why we are going to study them in detail.
If you compare this Corollary with the previous one for flows, you will observe a reduction of the condition on the dimensionality N, implying regular dynamics, by one for invertible maps and by two for noninvertible maps in comparison with flows. This is no coincidence as we will show in the following section. For a proof in case of invertible one-dimensional maps see [Has03, Kat95].

Example 5

1. Consider the logistic map Eq. (3.9) for $0 < r \leq 4$. It is smooth, defined on a compact set but not invertible, hence complicated solutions are possible.

2. Consider Hénon's map Eq. (3.10) for smooth $f(x)$. This map is smooth, however, whether the time asymptotic dynamics is defined on a closed, bounded set R is not clear. On the other hand, the map is invertible and $N = 2$, hence complicated solutions are in any case possible.

3.3 Poincaré surface of section

A *Poincaré surface of section* enables the reduction of an N-dimensional flow to an $(N-1)$-dimensional map. There exist two basic types:
The first one is the Poincaré surface of section *in space*. Fig. 3.7 shows an example for a flow defined by $\dot{\underline{x}} = \underline{F}(\underline{x})$ in $N = 3$-dimensional space.
Let us consider the case, e.g., $x^3 = K$, where K is a constant chosen by convenience. The *Poincaré map* for this system is then defined uniquely by the iteration

$$\begin{pmatrix} x^1_{n+1} \\ x^2_{n+1} \end{pmatrix} = \underline{M} \begin{pmatrix} x^1_n \\ x^2_n \end{pmatrix} \tag{3.12}$$

from the n^{th} to the $(n+1)^{st}$ piercing *under the condition* that $x^3 = K$ is kept fixed.
This definition appears to be fairly simple, however, there are at least two subtleties: Firstly, it is quite easy to produce a Poincaré surface of section for a given differential equation *numerically*, but only in exceptional cases can the corresponding Poincaré map be obtained *analytically*. Secondly, if the times τ between two piercings are constant, then some trivial information about the dynamical system is conveniently separated out by producing a

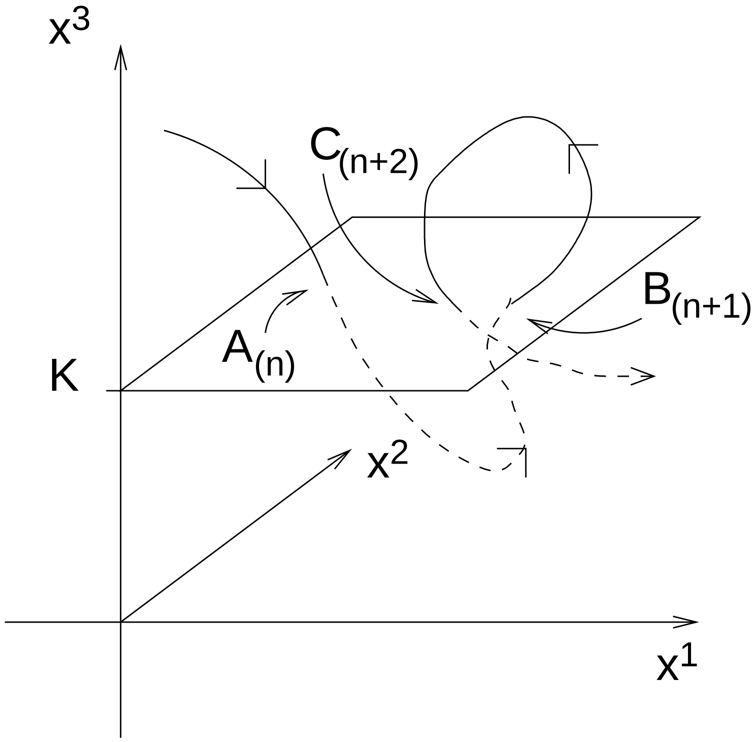


Figure 3.7: A continuous trajectory in space pierces the plane $x^3 = K$ at several points at discrete time n.

Poincaré surface of section. However, most often there is no reason why τ should be constant. If this is not the case, the complete dynamics is defined by what is called a *suspended flow*, or suspension flow,

$$\underline{x}_{n+1} = \underline{M}(\underline{x}_n)\ ,\ \underline{x}_n \in \mathbb{R}^{N-1} \tag{3.13}$$

$$t_{n+1} = t_n + T(\underline{x}_n)\ , T : \mathbb{R}^{N-1} \to \mathbb{R} \tag{3.14}$$

Eq. (3.14) is called the first-return time map.[5] This allows some idea of how the Poincaré-Bendixson theorem 1 applies to reduced-dimension discrete systems: The above suspended flow still provides an exact representation of the underlying time-continuous dynamical system. However, if we refer only to the first Eq. (3.13) by neglecting the second Eq. (3.14), as we did before, we reduce the dimensionality of the whole dynamical system by one. Of course we could also consider the whole suspended flow Eqs. (3.13),(3.14) as an N-dimensional map, however, the set on which Eq. (3.14) is defined is always unbounded and hence we could not relate to the Poincaré-Bendixson theorem.
Still, by construction our Poincaré map Eq. (3.13) is invertible. However, if we neglect any single component of the vector in Eq. (3.13), implying that we further reduce the phase space dimensionality by one, the resulting map will become noninvertible, because we have lost some information. Hence, in view of the Poincaré-Bendixson theorem 1 it is not surprising that despite their minimal dimensionality, one-dimensional noninvertible maps can exhibit complicated time-asymptotic behavior.

[5] Obviously, this equation then replaces the dynamics for the eliminated phase space variable, however, the detailed relation between these two different quantities is typically non-trivial [Gas98, Kat95].

A second version of a Poincaré map is obtained by a Poincaré surface of section *in time*. For this purpose, we sample a time-continous dynamical system not with respect to a constraint in phase space but *at discrete times* $t_n = t_0 + n\tau, n \in \mathbb{N}$. In this case, the map determining Eq. (3.14) of the suspended flow boils thus down to $T(\underline{x}_n) = n\tau$: We have what is called a *stroboscopic sampling* of the phase space.
This variant is very convenient for dynamical systems driven by periodic forces such as, for example, our nonlinear pendulum Eq. (2.1) (why?). However, we wish to illustrate this technique for a simpler system, where this is easier to see. This system provides a rare example for which the Poincaré map can be calculated analytically. In order to do so, we first need to learn about a mathematical object called the (Dirac) 'δ-function', which we may introduce as follows.
We are looking for a 'function' having the following properties:

$$\delta(x) := \begin{cases} 0\,, \ x \neq 0 \\ \infty\,, \ x = 0 \end{cases} \tag{3.15}$$

For example, think of the sequence of functions defined by

$$\delta_\gamma(x) := \frac{1}{\gamma\sqrt{2\pi}} \exp(-\frac{x^2}{2\gamma^2}) \tag{3.16}$$

as shown in Fig. 3.8. It is not hard to see that in the limit of $\gamma \to 0$ this sequence has the desired properties, that is, $\delta_\gamma(x) \to \delta(x)$ $(\gamma \to 0)$.

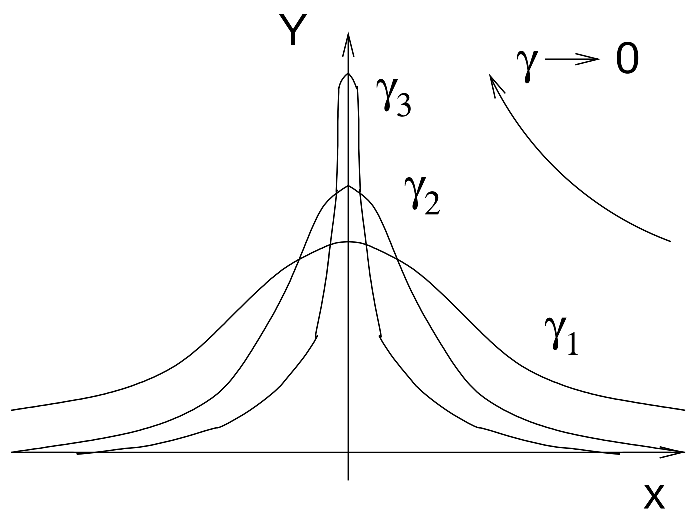


Figure 3.8: A sequence of functions approaching the δ-function.

We remark that many other representations of the δ-function exist. Strictly speaking the 'δ-function' is not a function but rather a functional, respectively a distribution defined on a specific (Schwartz) function space.[6]

[6] A nice short summary of properties of the δ-function, written from a physicist's point of view, is given, for example, in the German version of the book by F.Reif, *Statistische Physik und Theorie der Wärme*, see Appendix A.7.

Two important properties of the δ-function that we will use in the following are its normalization,

$$\forall \epsilon > 0 \int_{-\epsilon}^{\epsilon} dx\, \delta(x) = 1 \quad , \tag{3.17}$$

and that for some integrable function $f(x)$

$$\int_{-\infty}^{\infty} dx\, f(x)\delta(x) = f(0) \quad . \tag{3.18}$$

We are now prepared to discuss the following example:

Example 6 The kicked rotor
A rotating bar of inertia I and length l suffers kicks of strength K/l applied periodically at time steps $t = 0, \tau, 2\tau, \ldots\,, \tau \in \mathbb{R}$, see Fig. 3.9.

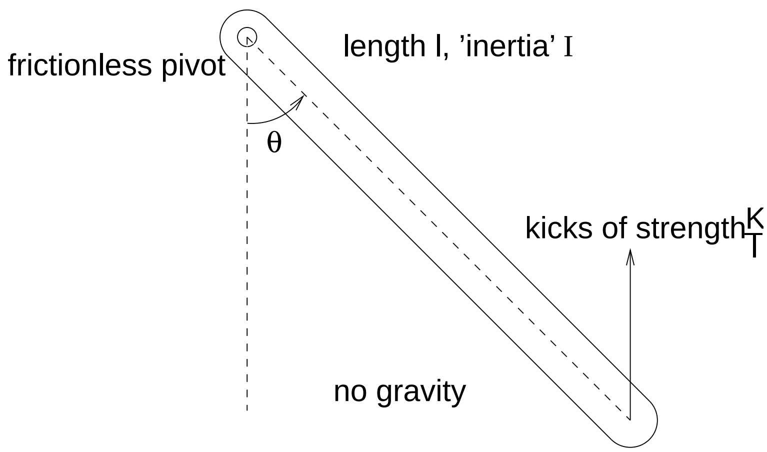


Figure 3.9: The kicked rotor.

The equation of motion for this dynamical system is straightforwardly derived from physical arguments [Ott93], however, we just state it here in form of

$$\ddot{\theta} = k \sin\theta \sum_{m=0}^{\infty} \delta(t - m\tau) \quad , \tag{3.19}$$

where the dynamical variable θ describes the turning angle and $k := K/I$ is a control parameter. As before we can rewrite this differential equation as a vector field,

$$\dot{\theta} = \omega \tag{3.20}$$

$$\dot{\omega} = k \sin\theta \sum_{m=0}^{\infty} \delta(t - m\tau) \quad . \tag{3.21}$$

According to Eq. (3.21) ω is constant during times $t \neq m\tau$ between the kicks but changes *discontinuously* at the kicks, which take place at times $t = m\tau$. Eq. (3.20) then implies that $\theta \sim t$ between the kicks by changing *continuously* at the kicks, reflecting the discontinuous changes in the slope ω.

Let us now construct a suitable Poincaré surface of section in time for this dynamical system. Let us define $\theta_n := \theta(t)$ and $\omega_n := \omega(t)$ at $t = n\tau + 0^+$, where 0^+ is a positive infinitesimal. That is, we look at both dynamical variables right after each kick. We then integrate Eq. (3.20) through the δ-function at $t = (n+1)\tau$ leading to

$$\int_{n\tau+0^+}^{(n+1)\tau+0^+} dt\,\dot{\theta} = \theta_{n+1} - \theta_n = \omega_n \tau \tag{3.22}$$

The same way we integrate Eq. (3.21) to

$$\omega_{n+1} - \omega_n = k \int_{n\tau+0^+}^{(n+1)\tau+0^+} dt\ \sin\theta \sum_m \delta(t - m\tau) = k \sin\theta_{n+1} \quad . \tag{3.23}$$

For the special case $\tau = 1$ we arrive at

$$\theta_{n+1} = \theta_n + \omega_n \tag{3.24}$$

$$\omega_{n+1} = \omega_n + k\sin\theta_{n+1} \quad . \tag{3.25}$$

This two-dimensional mapping is called the *standard map*, or sometimes also the Chirikov-Taylor map. It exhibits a dynamics that is typical for time-discrete Hamiltonian dynamical systems thus serving as a standard example for this class of systems [Ott93].[7] For convenience, the dynamics of (θ_n, ω_n) is often considered mod 2π. We remark in passing that Eqs. (3.24),(3.25) can also be derived from a toy model for cyclotron dynamics, where a charged particle moves under a constant magnetic field and is accelerated by a time-dependent voltage drop.[8]
We may now further reduce the kicked rotor dynamics by feeding Eq. (3.24) into Eq. (3.25) leading to

$$\omega_{n+1} = \omega_n + k\sin(\theta_n + \omega_n) \tag{3.26}$$

By assuming *ad hoc* that $\theta_n \ll \omega_n$, which of course would have to be justified in detail if one wanted to claim the resulting equation to be a realistic model for a kicked rotor,[9] one arrives at

$$\omega_{n+1} = \omega_n + k\sin\omega_n \quad . \tag{3.27}$$

This one-dimensional *climbing sine map* is depicted in Fig. 3.10. It provides another example of a seemingly simple mapping exhibiting very non-trivial dynamics that changes in a very complicated way under parameter variation.

Part II

[7] see, e.g., the book by Lichtenberg and Lieberman for further studies of this specific type of dynamical systems.

[8] see, e.g., a review by J.D.Meiss on symplectic twist maps for some details (Rev. Mod. Phys. **64**, 795 (1992).

[9] This justification is not trivial and depends very much on the choice of the control parameter k; see research papers by Bak et al.

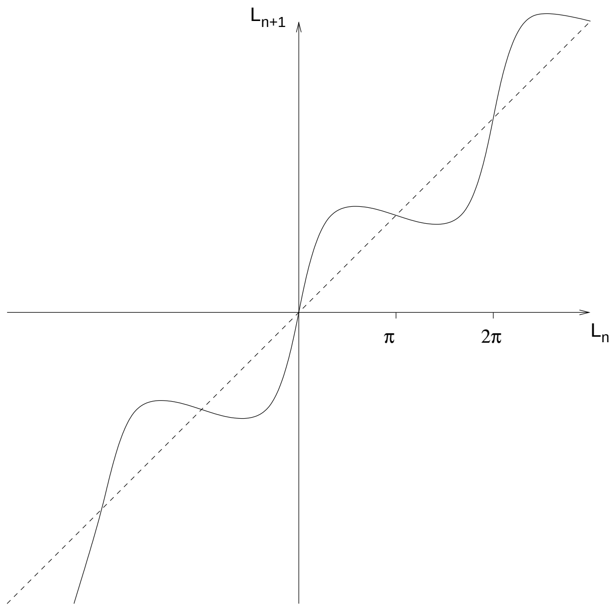


Figure 3.10: The climbing sine map.

Topological properties of one-dimensional maps

4 Some basic ingredients of nonlinear dynamics

One-dimensional maps are the simplest systems capable of chaotic motion. They are thus very convenient for learning about some fundamental properties of dynamical systems. They also have the advantage that they are quite amenable to rigorous mathematical analysis. On the other hand, it is not straightforward to relate them to realistic dynamical systems. However, as we have tried to illustrate in the first part of these lectures, one may argue for such connections by carefully using discretizations of time-continuous dynamics.[1]
Let us start the second part of our lectures with another very simple but prominent example of a one-dimensional map.

[1]The 'physicality' of one-dimensional maps is a delicate issue on which there exist different points of view in the literature [Tél06].

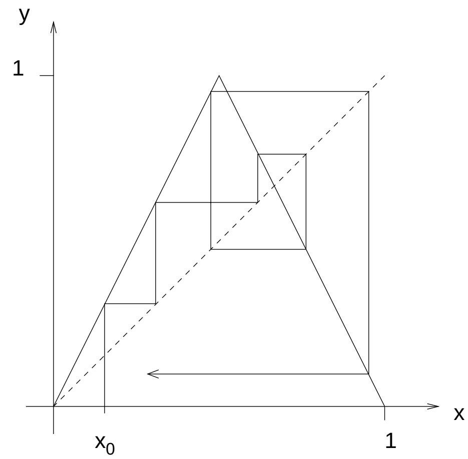


Figure 4.1: The tent map including a cobweb plot.

Example 7 [Ott93] The tent map
The tent map shown in Fig. 4.1 is defined by

$$M : [0,1] \to [0,1] \quad , \quad M(x) := 1 - 2\left|x - \frac{1}{2}\right| = \begin{cases} 2x & , \quad 0 \leq x \leq 1/2 \\ 2 - 2x & , \quad 1/2 < x \leq 1 \end{cases} \tag{4.1}$$

As usual, its equations of motion are given by $x_{n+1} = M(x_n)$. One can easily see that the dynamics is bounded for $x_0 \in [0,1]$.
By definition the tent map is *piecewise linear*. One may thus wonder in which respect such a map can exhibit a possibly chaotic dynamics that is typically associated with *nonlinearity*. The reason is that there exists a point of nondifferentiability, that is, the map is continous but not differentiable at $x = 1/2$. If we wanted to approximate the tent map by a sequence of differentiable maps, we could do so by unimodal functions as sketched in Fig. 4.2 below. We would need to define the maxima of the function sequence and the curvatures around them such that they asymptotically approach the tent map. So in a way, the tent map may be understood as the limiting case of a sequence of nonlinear maps.

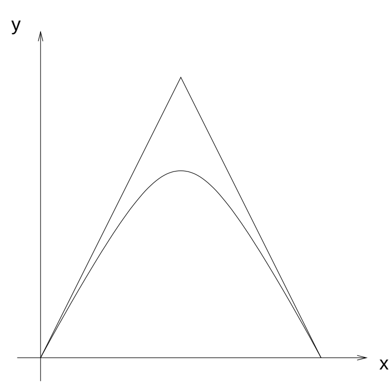


Figure 4.2: Approximation of the piecewise linear nondifferentiable tent map by a sequence of nonlinear differentiable unimodal maps.

A second question may arise about the necessity, or the meaning, of the noninvertibility of the tent map compared with realistic (Hamiltonian) dynamical systems, which are most often invertible. The noninvertibility is necessary in order to model a "stretch-and-fold"-mechanism as we have seen it in the computer simulations for the bouncing ball billiard, see Fig. 4.3 for the tent map: Assume we fill the whole unit interval with a uniform distribution of points. We may now decompose the action of the tent map 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 it *folds* the resulting line segment due to the presence of the cusp at $x = 1/2$, which leads to motion *bounded* on the unit interval.

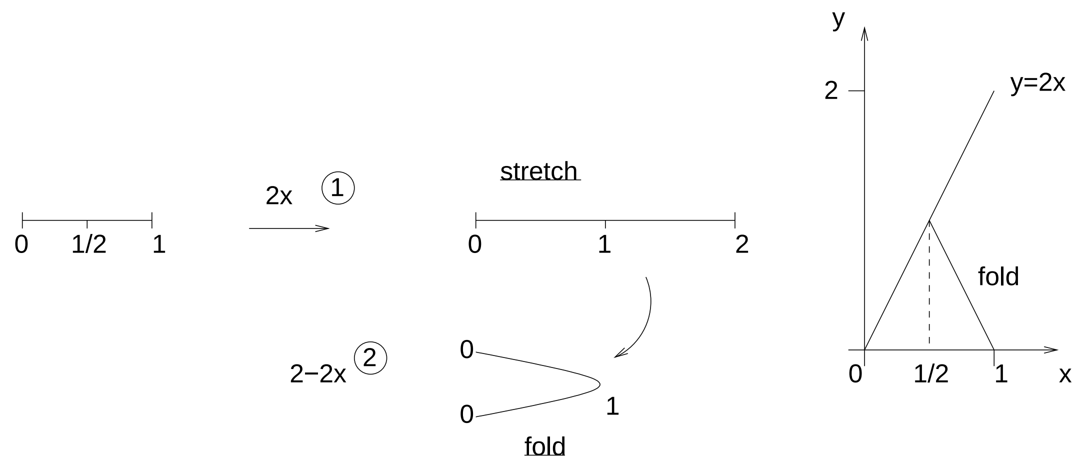


Figure 4.3: Stretch-and-fold mechanism in the tent map.

The tent map thus yields a simple example for an essentially nonlinear stretch-and-fold mechanism, as it typically generates chaos. This mechanism is encountered not only in the bouncing ball billiard but also in many other realistic dynamical systems. We may remark that 'stretch and cut' or 'stretch, twist and fold' provide alternative mechanisms for generating complicated dynamics. You may wish to play around with these ideas in thought experiments, where you replace the sets of points by kneading dough.

From now on we essentially consider one-dimensional time-discrete maps only. However, most of the concepts that we are going to introduce carry over, suitably amended, to n-dimensional and time-continuous dynamical systems as well.

4.1 Homeomorphisms and diffeomorphisms

The following sections draw much upon Ref. [Dev89]; see also [Rob95] for details.
Let $F : I \to J, \quad I, J \subseteq \mathbb{R}, x \mapsto y = F(x)$ be a function.

Definition 9 *$F(x)$ is a* homeomorphism *if F is bijective, continuous and $\exists$ continuous inverse $F^{-1}(x)$.*

Example 8 Let $F : (-\pi/2, \pi/2) \to \mathbb{R}\,,\ x \mapsto F(x) = \tan x$. Then $F^{-1} : \mathbb{R} \to (-\pi/2, \pi/2)\,,\ x \mapsto \arctan x$ and thus it defines a homeomorphism; see Fig. 4.4.

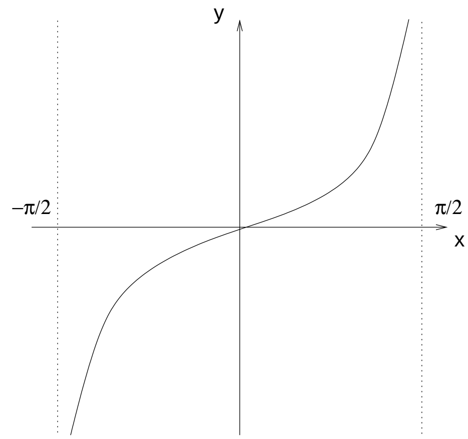


Figure 4.4: The function $\tan x$ is a homeomorphism.

Definition 10 *$F(x)$ is called a C^r-*diffeomorphism *(r-times continuously differentiable) if F is a C^r-homeomorphism such that $F^{-1}(x)$ is also C^r.*

Example 9

1. $F(x) = \tan x$ is a C^∞-diffeomorphism, as one can see by playing around with the functional forms for $\arctan x$ and $\tan x$ under differentiation.

2. $F(x) = x^3$ is a homeomorphism but not a diffeomorphism, because $F^{-1}(x) = x^{1/3}$ and $(F^{-1})'(x) = \frac{1}{3}x^{-2/3}$, so $(F^{-1})'(0) = \infty$.

4.2 Periodic points

Definition 11 *The point x is a* periodic point of period *n of F if $F^n(x) = x$. The least positive n for which $F^n(x) = x$ is called the* prime period *of x.*
$Per_n(F)$ denotes the set of periodic points of (not necessarily prime!) period n.
The set of fixed points *$F(x) = x$ is $Fix(F) = Per_1(F)$.*
The set of all iterates of a periodic point forms a periodic orbit*.*

Remark 1 If x is a fixed point of F, i.e. $F(x) = x$, then $F^2(x) = F(F(x)) = F(x) = x$. It follows $x \in \mathrm{Per}_1(F) \Rightarrow x \in \mathrm{Per}_2(F) \Rightarrow x \in \mathrm{Per}_n(F)\, \forall n \in \mathbb{N}$. Hence the definition of *prime* period.
We furthermore remark that fixed points are the points where the graph $\{(x, F(x))\}$ of F intersects the diagonal $\{(x, x)\}$, as is nicely seen in cobweb plots.

Example 10

1. Let $F(x) = x = \mathrm{Id}(x)$. Then the set of fixed points is determined by $\mathrm{Fix}(F) = \mathbb{R}$; see Fig. 4.5.
2. Let $F(x) = -x$. The fixed points must fulfill $F(x) = x = -x \Rightarrow 2x = 0 \Rightarrow x = 0$, so $\mathrm{Fix}(F) = 0$.

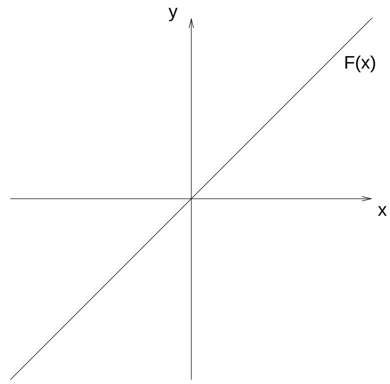


Figure 4.5: Set of fixed points for $F(x) = x$.

The period two points must fulfill $F^2(x) = F(F(x)) = x$. However, $F(x) = -x$ and $F(-x) = x$, so $\mathrm{Per}_2(F) = \mathbb{R}$. But note that Prime $\mathrm{Per}_2(F) = \mathbb{R} \setminus \{0\}$!

Both results could also have been inferred directly from Fig. 4.6.

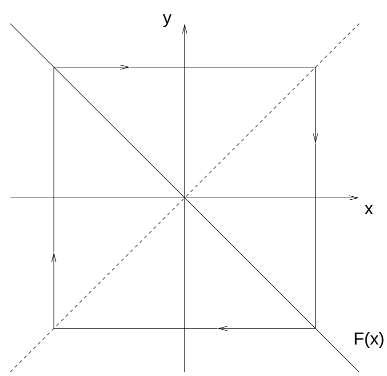


Figure 4.6: Set of fixed points and points of period two for $F(x) = -x$.

Remark 2 In typical dynamical systems the fixed points and periodic orbits are *isolated* with 'more complicated' orbits in between, as will be discussed in detail later on.
There exists also a nice *fixed point theorem*: A *continuous* function F mapping a *compact interval onto itself* has at least one fixed point; see [Dev89] for a proof and for related theorems. The detailed discussion of such theorems is one of the topics of the module 'Chaos and Fractals'.

Example 11 Let $F(x) = \frac{3x-x^3}{2}$.

1. The *fixed points* of this mapping are calculated to $F(x) = x = \frac{3x-x^3}{2} \Rightarrow x - x^3 = x(1-x^2) = 0 \Rightarrow \mathrm{Fix}(F) = \{0, \pm 1\}$.

2. Let us illustrate the dynamics of this map in a *cobweb plot*. For this purpose, note that the extrema are $F'(x) = \frac{1}{2}(3 - 3x^2) = 0 \Rightarrow x = \pm 1$ with $F(1) = 1, F(-1) = -1$.

The roots of the map are $F(x) = 0 = \frac{3x-x^3}{2} \Rightarrow x(3-x^2) = 0 \Rightarrow x \in \{0, \pm\sqrt{3} \simeq 1.73\}$. We can now draw the graph of the map, see Fig. 4.7. The stability of the fixed points can be assessed by cobweb plots of nearby orbits as we have discussed before.

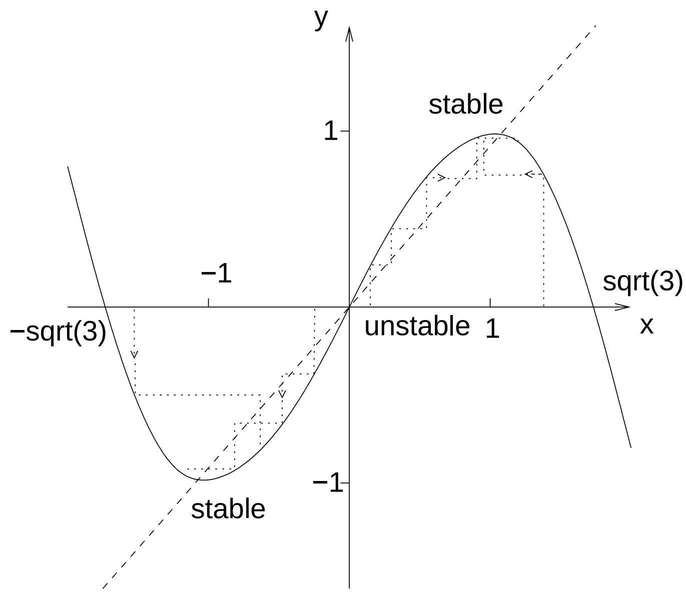


Figure 4.7: Cobweb plot for the map defined by the function $F(x) = \frac{3x-x^3}{2}$.

3. The *points of period 2* are determined by $F^2(x) = x$, so one has to calculate the roots of this polynomial equation. In order to do so it helps to separate the subset of solutions $\text{Fix}(F) \subseteq \text{Per}_2(F)$ from the polynomial equation, which defines the remaining period 2 solutions by division.

Definition 12 *A point x is called* eventually periodic of period n *if x is not periodic but $\exists m > 0$ such that $\forall i \geq m\, F^{n+i}(x) = F^i(x)$, that is, $F^i(x) = p$ is periodic for $i \geq m$, $F^n(p) = p$.*

Example 12

1. $F(x) = x^2 \Rightarrow F(1) = 1$ is a fixed point, whereas $F(-1) = 1$ is *eventually periodic* with respect to this fixed point; see Fig. 4.8.

2. One can easily construct eventually periodic orbits via *backward iteration* as illustrated in Fig. 4.9.

4.3 Dense sets, Bernoulli shift and topological transitivity

Definition 13 *Let I be a set and d a* metric *or* distance function. *Then (I, d) is called a* metric space.

Usually, in these lectures $I \subset \mathbb{R}$ with the Euclidean metric $d(x, y) = |x - y|$, so if not said otherwise we will work on *Euclidean spaces*.

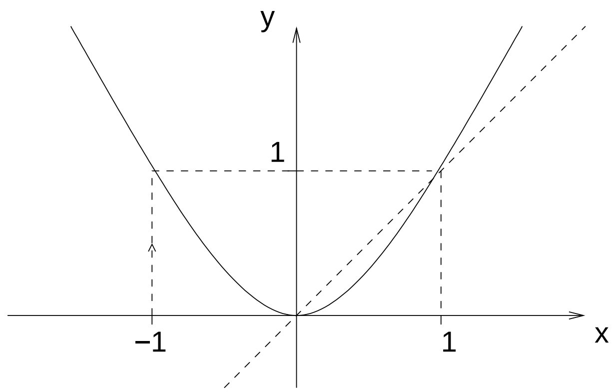


Figure 4.8: An eventually periodic orbit for $F(x) = x^2$.

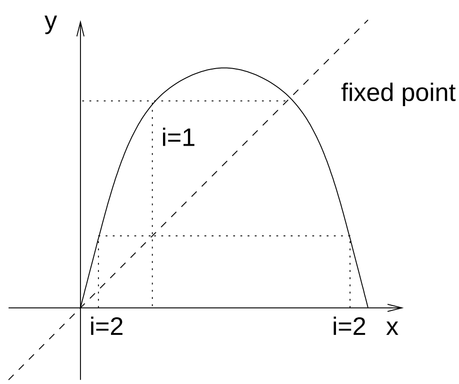


Figure 4.9: Construction of eventually periodic orbits for an example via backward iteration.

Definition 14 *The* epsilon neighbourhood *of a point* $p \in \mathbb{R}$ *is the interval of numbers* $N_\epsilon(p) := \{x \in \mathbb{R} \mid |x - p| < \epsilon\}$ *for given* $\epsilon > 0$*, i.e. the set* $\forall x \in \mathbb{R}$ *within a given distance* ϵ *of* p*, see Fig. 4.10.*

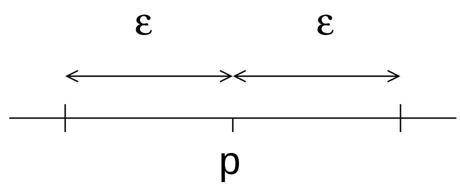


Figure 4.10: Illustration of an ϵ neighbourhood $N_\epsilon(p)$.

Definition 15 *Let* $A, B \subset \mathbb{R}$ *and* $A \subset B$*.* A *is called* dense *in* B *if arbitrarily close to each point in* B *there is a point in* A*, i.e.* $\forall x \in B\ \forall \epsilon > 0\ N_\epsilon(x) \cap A \neq \emptyset$*, see Fig. 4.11.*

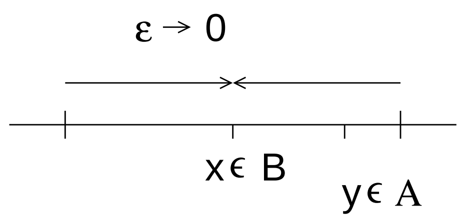


Figure 4.11: Illustration of a set A being dense in B.

An application of this definition is illustrated in the following proposition:

Proposition 1 *The rationals are dense on the unit interval.*

Proof: Let $x \in [0, 1]$. For given $\epsilon > 0$ choose $n \in \mathbb{N}$ sufficiently large such that $10^{-n} < \epsilon$. Let $\{a_1, a_2, a_3, \dots, a_n\}$ be the first n digits of x in decimal representation. Then $|x - 0.a_1a_2a_3 \dots a_n| < 10^{-n} < \epsilon \Rightarrow y := 0.a_1a_2a_3 \dots a_n \in \mathbb{Q}$ is in $N_\epsilon(x)$. q.e.d.

Certainly, much more could be said on the denseness and related properties of rational and irrational numbers in $\mathbb{R}$. However, this is not what we need for the following. We will continue by introducing another famous map:

Example 13 The Bernoulli shift (also shift map, doubling map, dyadic transformation)
The Bernoulli shift shown in Fig. 4.12 is defined by

$$B : [0, 1) \to [0, 1)\ ,\ B(x) := 2x \bmod 1\ = \begin{cases} 2x\ ,\ 0 \le x < 1/2 \\ 2x - 1\ ,\ 1/2 \le x < 1 \end{cases} . \tag{4.2}$$

Proposition 2 *The cardinality* $|Per_n(B)|$*, i.e., the number of elements of* $Per_n(B)$*, is equal to* $2^n - 1$ *and the periodic points are dense on* $[0, 1)$*.*

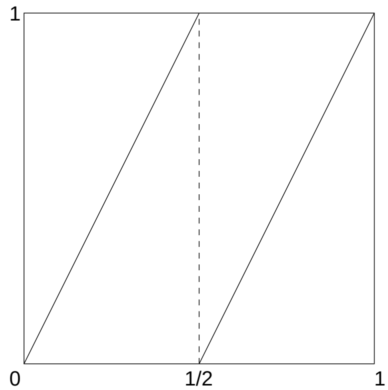


Figure 4.12: The Bernoulli shift.

Proof: Let us prove this proposition by using a more convenient representation of the Bernoulli shift dynamics, which is defined on the circle.[2] Let

$$S^1 := \{z \in \mathbb{C} \mid |z| = 1\} = \{\exp(i2\pi\phi) \mid \phi \in \mathbb{R}\} \tag{4.3}$$

denote the *unit circle* in the complex plane [Rob95, Has03], see Fig. 4.13.

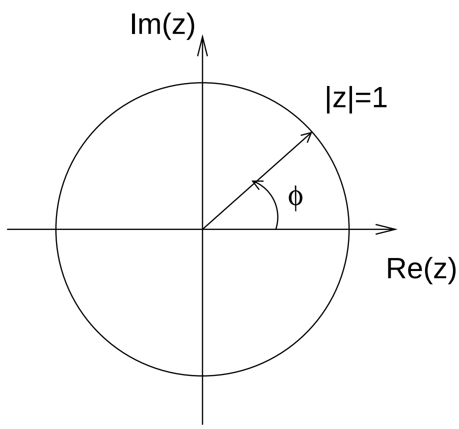


Figure 4.13: Representation of a complex number $z = \cos\phi + i\sin\phi$ on the unit circle.

Then define

$$B : S^1 \to S^1 \ , \ B(z) := z^2 = (\exp(i2\pi\phi))^2 = \exp(i2\pi 2\phi) \quad . \tag{4.4}$$

We have thus lifted the Bernoulli shift dynamics onto $\mathbb{R}$ in form of $\phi \to 2\phi$. Note that the map on the circle is C^0, whereas B in Eq. (4.2) is discontinuous at $x = 1/2$. This is one of the reasons why mathematically it is sometimes more convenient to use the circle representation; see also below.

[2] Strictly speaking one first has to show that Eq. (4.2) is *topologically conjugate* to Eq. (4.4), which implies that many dynamical properties such as the ones we are going to prove are the same. Topological conjugacy is an important concept in dynamical systems theory which is discussed in detail in the module 'Chaos and Fractals', hence we do not introduce it here; see also [Dev89, Ott93, Has03, All97].

Let us now calculate the periodic orbits for $B(z)$. Let $n \in \mathbb{N}$ and

$$B^n(z) = \underbrace{((z^2)^2 \ldots)^2}_{n \text{ times}} = z^{2^n} = z \quad . \tag{4.5}$$

With $z = \exp(2\pi i\phi)$ we get

$$\exp(2^n i2\pi\phi) = \exp(i2\pi\phi) = \exp(i2\pi(\phi + k)) \quad , \tag{4.6}$$

where for the last equality we have introduced a *phase* $k \in \mathbb{Z}$ expressing the possibility that we have k windings around the circle. Matching the left with the right hand side yields $2^n\phi = \phi + k$ and, if we solve for ϕ,

$$\phi = \frac{k}{2^n - 1} \quad . \tag{4.7}$$

Now let us restrict onto $0 \leq \phi < 1$ in order to reproduce the domain defined in Eq. (4.2). This implies the constraint

$$0 \leq k < 2^n - 1, \quad k \in \mathbb{N}_0 \quad . \tag{4.8}$$

Let us discuss this solution for two examples before we state the general case: For $n = 1$ we have according to Eq. (4.8) $0 \leq k < 1$, which implies $\phi = 0 = \mathrm{Fix}(B)$. Analogously, for $n = 2$ we get $0 \leq k < 3$, which implies $\phi \in \{0, \frac{1}{3}, \frac{2}{3}\} = \mathrm{Per}_2(B)$.
In the general case, we thus find that the roots of $z^{2^n-1} = 1$ are periodic points of B of period n. There exist exactly $2^n - 1$ of them, i.e. $|\mathrm{Per}_n(B)| = 2^n - 1$, and they are *uniformly spread* over the circle, see $n = 2$ (if you wish, please convince yourself for higher $n's$), with *equal intervals* of size $\delta_n \to 0 \quad (n \to \infty)$. Hence, the periodic points are dense on $[0, 1)$.

q.e.d.

Definition 16 *[Has03, Rob95] A map $F : J \to J, J \subseteq \mathbb{R}$, is* topologically transitive *on J if there exists a point $x \in J$ such that its orbit is* dense *in J.*

Intuitively, this definition means that F has points which eventually move under iteration from one arbitrarily small neighbourhood of points in J to *any other*. Another interpretation is given by the following theorem:

Theorem 2 *(Birkhoff transitivity theorem) Let J be a compact subset of $\mathbb{R}$ and F be continuous. Then F is topologically transitive if and only if for any two open sets $U, V \subset J \; \exists \; N \in \mathbb{N}$ such that $F^N(U) \cap V \neq \emptyset$.*

In other words, J *cannot be decomposed* into two disjoint sets that remain disjoint under the action of F. The proof of this theorem is too elaborate to be presented in our lectures, see p.205 of [Has03] or p.273 of [Rob95] for details.

Example 14 We present it in the form of

Proposition 3 *The Bernoulli shift $B(x)$ is topologically transitive.*

There are at least two different ways of how to prove this. The idea of the first version is to use the above theorem. For convenience (that is, to avoid the point of discontinuity) we may consider again the map on the circle. Then for whatever subset U we choose $\exists N \in \mathbb{N}$ such that $B^N(U) \supseteq S^1$ as sketched in Fig. 4.14, which is due to the linear expansion of the map. This just needs to be formalized [Has03]. It is no coincidence that this mechanism reminds us of what we have already encountered as "mixing".

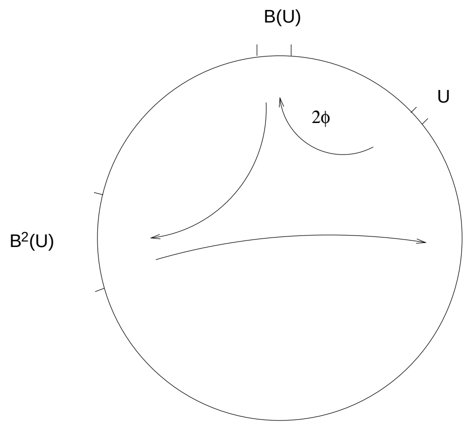


Figure 4.14: Bernoulli shift B operating on a subset of the unit circle.

A second way is to use what is called a *symbolic dynamics*, or *coding* of an orbit. Special cases are, for example, the binary or decimal representations of a real number. However, since symbolic dynamics is covered by the module Chaos and Fractals we do not introduce this concept in our course; see p.42 of [Dev89], p.212 of [Has03] or p.39 of [Kat95] for a proof and for further details.
We have thus learned that the Bernoulli shift has *both* a dense set of periodic orbits and is topologically transitive, which earlier we have associated with 'simple' and 'complicated' dynamics, respectively. So surprisingly this simple map *simultaneously* exhibits both types of dynamics, depending on initial conditions.

Definition 17 *F is called* minimal *if $\forall x \in J$ the orbit is dense in J.*

Example 15

1. The Bernoulli shift is *not* minimal (why?).

2. The rotation on the circle $R_\lambda(\theta)$ as defined on the exercise sheet is minimal if λ is irrational. The proof is a slight extension of what we have already shown in the coursework, that is, for irrational λ there is no periodic orbit [Has03, Kat95, Dev89].

Remark 3

1. Obviously, minimality implies topological transitivity.

2. Minimality is rather rare in dynamical systems, because it allows no existence of periodic orbits.

3. Later on we will encounter ideas that are very similar to topological transitivity by additionally employing what is called a *measure*, leading us to concepts of ergodic theory.

4.4 Stability

In this chapter we make our qualitative assessment of the stability of periodic points quantitative, which we have explored in terms of cobweb plots. We start with the following definition:

Definition 18 *[Dev89, Rob95] Let p be a fixed point of the map $F \in C^1$ acting on $\mathbb{R}$.*
If $|F'(p)| < 1$ then p is called a sink *or* attracting *fixed point.*
If $|F'(p)| > 1$ then p is called a source *or* repelling *fixed point.*
If $|F'(p)| = 1$ then p is called a marginal *(also* indifferent, neutral*) fixed point.*

Remark 4

1. This means that for a sink all points "sufficiently close" to p are attracted to p. Accordingly, for a source all such points are repelled from p. For a marginal fixed point both is possible, and one can further classify this point as being marginally stable, unstable or both, depending on direction.

2. Above we followed Devaney [Dev89] by defining stability for *differentiable* fixed points $F'(p)$ only. One can dig a little bit deeper without requiring differentiability for $F(p)$ by using ϵ-environments and then proving the above inequalities as a theorem [All97].

Definition 19 *[All97] The set $W^s(p)$ of initial conditions x whose orbits converge to a sink p, $\lim_{n\to\infty} F^n(x) = p$, is called the* stable set *or* basin *of p. $W^u(p)$ denotes the* unstable set *of initial conditions, whose orbits are repelled from a source.*

For invertible maps the definition of $W^u(p)$ can be made precise via backward iteration for given x. For non-invertible maps $W^u(p)$ can be defined as the set of all x for which there exists a set of preimages $F(x_{-i}) = x_{-i+1}$ with $\lim_{n\to\infty} F^{-n}(x) = p$ [Rob95].

Example 16 Consider the map defined in Example 11 by the function $F(x) = \frac{3x-x^3}{2}$. Let us recall that $\text{Fix}(F) = \{0, \pm 1\}$. Since $F'(x) = \frac{1}{2}(3 - 3x^2)$ we have $F'(0) = \frac{3}{2} > 1$, so as we already know from cobwep plots in Fig 4.7 it is a source, whereas $F'(\pm 1) = 0 < 1$ are sinks. Fig. 4.15 illustrates that the basin of attraction for $p = 1$ has actually a very complicated, intertwined structure [All97]. This is revealed by iterating backwards the set of initial conditions B_1 leading to the two preimages B_2 and B_2', and so on. Because of symmetry, an analogous topology is obtained for $p = -1$. This points towards the fact that basins of attraction can exhibit *fractal structures*,[3] which indeed is often encountered in dynamical systems.

We now check for the stability of periodic points.

Corollary 4 *Let p be a periodic point of F of period k, $F^k(p) = p$.*
If $|(F^k)'(p)| < 1$ then p is a periodic sink*.*
If $|(F^k)'(p)| > 1$ then p is a periodic source*.*
If $|(F^k)'(p)| = 1$ then p is a marginal periodic point*.*

[3] This is a concept that we do not further discuss in this course, see again the module on Chaos and Fractals.

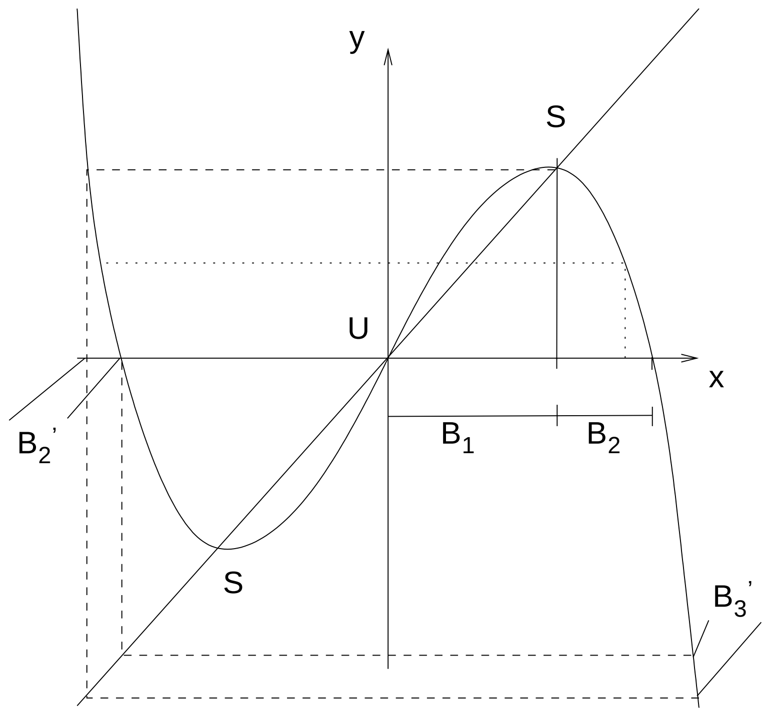


Figure 4.15: An example of a map exhibiting a complicated basin structure, which here is constructed by generating preimages of a given subset of the basin.

Proof: This trivially follows from Definition 18. Define $G(x) := F^k(x)$ and look at $G(p) = p$. Then p is an attracting/repelling/marginal fixed point of $G \iff p$ is an attracting/repelling/marginal fixed point of $F^k \iff p$ is an attracting/repelling/marginal periodic orbit of F. q.e.d.

Definition 20 *Hyperbolicity for one-dimensional maps [Dev89]*
Let p be a periodic point of prime period k. The point p *is said to be* hyperbolic *if* $|(F^k)'(p)| \neq 1$. *A* map $F \in C^1$ *on J is called* hyperbolic *if* $\exists N \in \mathbb{N}$, *such that* $\forall x \in J \forall n \geq N |(F^n)'(x)| \neq 1$.

That is, for hyperbolic maps we wish to rule out any marginal behavior in the nth iterates.

Definition 21 *[All97, Bec93]*
$F \in C^1$ *on J is called* expanding *if* $\forall\, x \in J\, |F'(x)| > 1$.
$F \in C^1$ *on J is called* contracting *if* $\forall\, x \in J\, |F'(x)| < 1$.

Remark 5

1. There is the weaker notion of *expansiveness* for $F \notin C^1$, which differs from expanding [Dev89].

2. It can easily be shown that if F is expanding or contracting it is hyperbolic, see the coursework.

3. If $F \notin C^1$ but the domain can be partitioned into finitely many intervals on which F is C^1, we say a map is *piecewise* C^1 [All97, Has03]. Correspondingly, we speak of *piecewise* expanding/contracting/hyperbolic maps.

4. For contracting maps there exists a *contraction principle* assuring exponential convergence onto a fixed point [Has03].

Example 17

1. The tent map has a point of non-differentiability at $x = 1/2$, hence it is piecewise C^1 and piecewise expanding, which according to the remark above implies piecewise hyperbolicity.

2. Let us consider again the logistic map $L(x) = rx(1-x)$, $x \in [0,1]$, $r > 0$, see also Eq. (3.9). It is $L'(x) = r(1-2x)$ with $L'(1/2) = 0 \, \forall r$. Consequently L cannot be expanding.

5 Definitions of deterministic chaos

We are finally able to define mathematically what previously we already referred to, in a very loose way, as "chaos".

5.1 Devaney chaos

A first definition of deterministic chaos, which was pioneered by Devaney in the first edition of his textbook [Dev89], reads as follows:

Definition 22 *Chaos in the sense of Devaney [Dev89, Has03, Rob95]*
A map $F : J \to J$, $J \subseteq \mathbb{R}$, is said to be D-chaotic *on J if:*

1. *F is topologically transitive.*

2. *The periodic points are dense in J.*

Remark 6 This definition highlights two basic ingredients of a chaotic dynamical system:

1. A *dense set of periodic orbits* related to 'simple' dynamics, which provide an element of regularity.

2. *Topological transitivity* reflecting 'complicated' dynamics and thus an element of irregularity.

Note that according to the Birkhoff Transitivity Theorem 2 topological transitivity also ensures *indecomposability* of the map, which in turn implies that these two sets must be intertwined in a non-trivial way.

Example 18 In Proposition 2 we have seen that the Bernoulli shift has a dense set of periodic orbits. However, Proposition 3 stated that it is also topologically transitive. Consequently it is D-chaotic.

5.2 Sensitivity and Wiggins chaos

Definition 23 *Sensitivity*[1] *[All97, Rob95, Dev89, Has03]*
A map $F : J \to J$, $J \subseteq \mathbb{R}$ *has* sensitive dependence on initial conditions *if* $\exists\, \delta > 0$ *such that* $\forall\, x \in J\, \forall\, N_\epsilon(x)\, \exists\, y \in N_\epsilon(x)$ *and* $\exists\, n \in \mathbb{N}$ *such that* $|F^n(x) - F^n(y)| > \delta$.

Remark 7 Pictorially speaking, this definition means that all neighbours of x (as close as desired) eventually move away at least a distance δ from $F^n(x)$ for n sufficiently large, see Fig. 5.1.

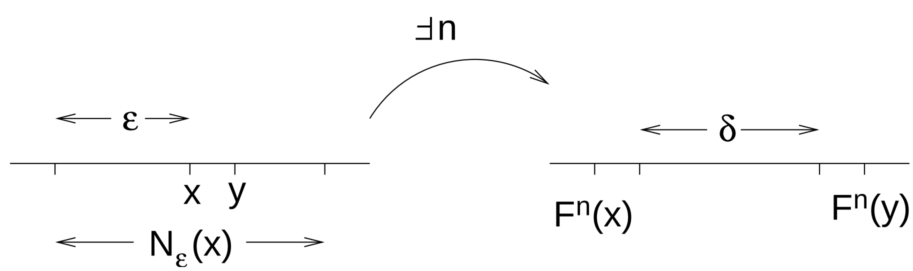


Figure 5.1: Illustration of the idea of sensitive dependence on initial conditions.

Example 19 This definition is illustrated by the following proposition:

Proposition 4 *The Bernoulli shift on* $x \in [0, 1)$, $B(x) = 2x \mod 1$, *is sensitive.*

Proof:
We simplify the proof by choosing y close enough to x such that x, y do not hit different branches of $B(x)$ under iteration. That way, we exclude the technical complication caused by the discontinuity of the map at $1/2$.[2]
We now verify the definition in a 'reverse approach': Let us assume that

$$\delta = |B^m(x) - B^m(y)| = |2^m x - 2^m y| = 2^m |x - y| \quad . \tag{5.1}$$

For given δ, we need to choose y and m such that the definition is fulfilled. It is thus a good idea to solve the above equation for m,

$$\begin{aligned} 2^m &= \frac{\delta}{|x-y|} \; , \; x \neq y \\ m \ln 2 &= \ln \frac{\delta}{|x-y|} \\ m &= \frac{\ln \frac{\delta}{|x-y|}}{\ln 2} \quad . \end{aligned} \tag{5.2}$$

[1] This fundamental definition is originally due to Guckenheimer [Guc79].
[2] This case could be covered by defining the map again on the circle, which for $[0, 1)$ could be done by introducing a "circle metric" [All97]. The proof then proceeds along the very same lines as shown in the following.

Let us now choose $\delta > 0$. Then the above solution stipulates that $\forall x$ and $\forall N_\epsilon(x) \exists y \in N_\epsilon(x)$ such that after $n > m$ iterations

$$|B^n(x) - B^n(y)| = 2^{n-m+m}|x - y| = 2^{n-m}\delta > \delta \quad . \tag{5.3}$$

q.e.d.

The concept of sensitivity enables us to come up with a second definition of chaos:

Definition 24 *Chaos in the sense of Wiggins [Wig92]*
A map $F : J \to J$, $J \subseteq \mathbb{R}$ is said to be W-chaotic *on J if:*

1. *F is topologically transitive.*
2. *It has sensitive dependence on initial conditions.*

Remark 8 Again we can detect two different ingredients in this definition:

1. *Sensitivity* related to some *stretching* mechanism, which eventually implies unpredictability.
2. *Topological transitivity* which means that there exists a dense orbit that, despite stretching, eventually returns close to its initial starting point. This implies that the dynamical systems must exhibit as well some *folding* mechanism.

This interplay between “stretch and fold” is precisely what we have seen in the earlier computer simulations of Chapter 2.

One may now wonder why we have two different definitions of chaos and whether there exists any relation between them. A first answer to this question is given by the following theorem:

Theorem 3 *(Banks et al. [Ban92], Glasner and Weiss [Gla93])*
If F is topologically transitive and there exists a dense set of periodic orbits F is sensitive. Thus D-chaos implies W-chaos.

Example 20 The Bernoulli shift is D-chaotic, consequently it is W-chaotic.

5.3 Ljapunov chaos

As a warm-up, you are encouraged to check the web for different spellings and pronunciations of the Russian name “Ljapunov”.
Let us motivate this section with another application of our favourite map:

Example 21 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 map around $x = 1/2$. We then have

$$\delta x_n := |x'_n - x_n| = 2\delta x_{n-1} = 2^2\delta x_{n-2} = \ldots = 2^n\delta x_0 = e^{n\ln 2}\delta x_0 \quad . \tag{5.4}$$

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 $B(x)$.

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

$$\delta x_n = |x'_n - x_n| = |F^n(x'_0) - F^n(x_0)| =: \delta x_0 e^{n\lambda(x_0)} \; (\delta x_0 \to 0) \tag{5.5}$$

for which we *presuppose* that an exponential separation of trajectories exists.[3] By furthermore assuming that F is differentiable we can rewrite this equation to

$$\begin{aligned} \lambda(x_0) &= \lim_{n\to\infty} \lim_{\delta x_0 \to 0} \frac{1}{n} \ln \frac{\delta x_n}{\delta x_0} \\ &= \lim_{n\to\infty} \lim_{\delta x_0 \to 0} \frac{1}{n} \ln \frac{|F^n(x_0+\delta x_0) - F^n(x_0)|}{\delta x_0} \\ &= \lim_{n\to\infty} \frac{1}{n} \ln \left| \frac{dF^n(x)}{dx} \right|_{x=x_0} \quad . \end{aligned} \tag{5.6}$$

Using the chain rule we obtain

$$\left. \frac{\mathrm{d}F^n(x)}{\mathrm{d}x} \right|_{x=x_0} = F'(x_{n-1})F'(x_{n-2})\dots F'(x_0) \quad , \tag{5.7}$$

which leads to our final result

$$\begin{aligned} \lambda(x_0) &= \lim_{n\to\infty} \frac{1}{n} \ln \left| \prod_{i=0}^{n-1} F'(x_i) \right| \\ &= \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \ln |F'(x_i)| \quad . \end{aligned} \tag{5.8}$$

The last expression defines a *time average* (in the mathematics literature this is sometimes called a *Birkhoff average*), where n terms along the trajectory with initial condition x_0 are summed up by averaging over n. These considerations motivate the following important definition:

Definition 25 *[All97] Let $F \in C^1$ be a map of the real line. The* local Ljapunov number *$L(x_0)$ of the orbit $\{x_0, x_1, \dots, x_{n-1}\}$ is defined as*

$$L(x_0) := \lim_{n\to\infty} \left| \prod_{i=0}^{n-1} F'(x_i) \right|^{\frac{1}{n}} \tag{5.9}$$

if this limit exists. The local Ljapunov exponent *$\lambda(x_0)$ is defined as*

$$\lambda(x_0) := \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \ln |F'(x_i)| \tag{5.10}$$

if this limit exists.[4]

[3] We emphasize that this is not always the case, see, e.g., Section 17.4 of [Kla07].
[4] This definition was proposed by A.M. Ljapunov in his Ph.D. thesis 1892.

Remark 9

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. It holds $\exists\lambda \Leftrightarrow \exists L \neq 0$ or ∞. If either side is true we have $\ln L = \lambda$.

3. If $F'(x_i) = 0 \Rightarrow \nexists\lambda(x)$. However, typically this concerns only a finite set of points.

4. For an expanding map it follows from Definition 21 that $\lambda(x) > 0$. Likewise, for a contracting map $\lambda(x) < 0$.

The concept of Ljapunov exponents enables us to come up with a third definition of deterministic chaos:

Definition 26 *Chaos in the sense of Ljapunov [Rob95, Ott93, All97, Bec93]*
A map $F : J \to J$, $J \subseteq \mathbb{R}$, F (piecewise) C^1 is said to be L-chaotic *on J if:*

1. *F is topologically transitive.*

2. *It has a positive Ljapunov exponent for a* typical *initial condition x_0.*

Remark 10

1. "Typical" means, loosely speaking, that this statement applies to any point that we randomly pick on J with non-zero probability.[5]

2. The reason why we require typicality for initial conditions is illustrated in Fig. 5.2. It shows an example of a map with an unstable fixed point at $x = 1/2$, where $\lambda(1/2) > 0$. However, this value of λ is atypical for the map, since all the other points are attracted to the stable fixed points at $\{0, 1\}$, where the Ljapunov exponents are negative. Thus it would be misleading to judge for L-chaos based only on "some" initial condition.

3. So far we have only defined a *local* Ljapunov exponent $\lambda(x)$, $x = x_0$. Later on we will introduce a *global* λ, whose definition straightforwardly incorporates topological transitivity and typicality. The global λ essentially yields a number that assesses whether a map is chaotic in the sense of exhibiting an *exponential* dynamical instability.

4. This is the reason why in the applied sciences "chaos in the sense of Ljapunov" became such a popular concept. However, it seems that mathematicians rather prefer to speak of chaos in the sense of Devaney or Wiggins.

5. $\lambda > 0$ implies that a map is sensitive, however, the other direction is not true. That is, W-chaos is weaker than L-chaos, since trajectories can separate more weakly than exponentially.

6. Note also that W-chaos requires sensitivity $\forall x$, whereas for L-chaos $\lambda > 0$ for one (typical) x is enough. This is so because, in contrast to sensitivity, the Ljapunov exponent defines an average (statistical) quantity.

[5] We will give a rigorous definition of this statement later on, after we know what a measure is, see Def. 48.

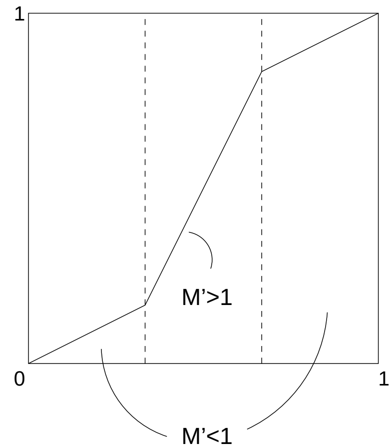


Figure 5.2: Example of a map that has a positive local Ljapunov exponent, whereas for typical initial conditions the Ljapunov exponent is negative.

Definition 27 *Ljapunov exponent for periodic points*
Let $p \in \mathbb{R}$ be a periodic point of period n. Then

$$\lambda(p) = \frac{1}{n} \sum_{i=0}^{n-1} \ln |F'(p_i)| \ , \ p = p_0 \ , \ p_i = F^i(p) \tag{5.11}$$

is the Ljapunov exponent along the periodic orbit $\{p_0, p_1, \ldots, p_{n-1}\}$. For the fixed point case $n = 1$ we have $\lambda(p) = \ln |F'(p)|$.

Example 22

1. 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 $\lambda(x) = \ln 2$ at these points. So here the Ljapunov exponent is the same for almost all x, because the map is uniformly expanding.

2. For the map $F(x) = (3x - x^3)/2$ discussed previously, see Figs. 4.7 and 4.15, we already know that $\text{Fix}(F) = \{0, \pm 1\}$. Let us calculate the map's Ljapunov exponents at these fixed points: With $F'(x) = \frac{3}{2} - \frac{3}{2}x^2$ we have $F'(0) = \frac{3}{2} \Rightarrow \lambda(0) = \ln \frac{3}{2} > 0$ corresponding to a repelling fixed point. For the other two fixed points we get $F'(\pm 1) = 0$, so unfortunately in these (rare) cases the Ljapunov exponent is not defined. The calculation of Ljapunov exponents for period two orbits will be subject of the coursework sheet.

5.4 Summary

We conclude this second part with what may be called a tentative summary: Let us assume that we have a map $F \in C^1$ defined on a compact set and that it is topologically transitive. Figure 5.3 then outlines relationships between different dynamical systems properties, all topologically characterizing "chaotic behavior" by assessing the stability of a dynamical system in slightly different ways.
However, there is one definition in the figure that we have not discussed so far:

Definition 28 *F is* topologically mixing *if for any two open sets $U, V \subset J$ $\exists N \in \mathbb{N}$ such that $\forall n > N$, $n \in \mathbb{N}$, $F^n(U) \cap V \neq \emptyset$.*

You may wish to explore yourself in which respect this property relates to what we have previously encountered as topological transitivity, see Theorem 2.

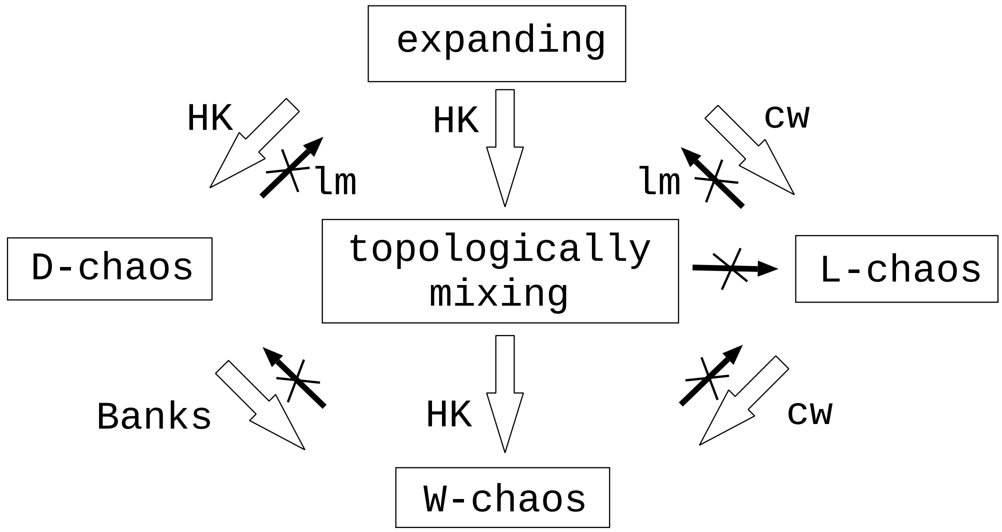


Figure 5.3: Tentative summary of relations between different topological chaos properties.

Figure 5.3 may be understood as follows: Open arrows represent logical implications that hold basically under the conditions stated above,[6] crossed-out bold arrows denote situations where a counterexample can be found ruling out the existence of a respective implication. The label 'HK refers to proofs contained in the book by Hasselblatt and Katok [Has03], 'cw' means that you were supposed to show this in a coursework problem, 'lm' stands for the logistic map, which provides counterexamples for certain values of the control parameter, 'Banks' points to Theorem 3. The arrows without any labels hold for counterexamples that can be found elsewhere in the literature.[7] See also the masters thesis by A.Fotiou [Fot05] for details, which summarizes recent literature on all these relations.

Obviously there are still gaps in this diagram: Of particular importance would be to learn whether W-chaos can imply topological mixing. One may furthermore wonder whether D-chaos and L-chaos can imply topological mixing, whereas it appears unlikely that topological mixing implies D-chaos or that it implies that a map is expanding. For the author of these lecture notes all these questions remain to be answered, and respective hints on solutions would be most welcome.

We may finally remark that apart from the topological chaos properties discussed here, one can come up with further definitions of chaos which, however, are based on more sophisticated concepts of dynamical systems theory that go beyond the scope of these lectures.[8]

[6] In some cases the detailed technical requirements for the validity of the proofs can be more involved.

[7] For W-chaos not implying D-chaos see counterexamples in [Ban92] and in a paper by Assaf, for W-chaos not implying L-chaos see the Pomeau-Manneville map, for topologically mixing not implying L-chaos see some research papers by Prozen and Campbell on irrational triangular billiards.

[8] This relates to notions of dynamical randomness in terms of, e.g., topological and measure-theoretic entropies or symbolic dynamics as they are encountered in what is called the ergodic hierarchy in the literature [Arn68, Ott93, Guc90].

Part III

Probabilistic description of one-dimensional maps

For this part of our lectures we recommend Refs. [Las94, Dor99, Bec93] as supplemetary literature. The first book introduces the topic from a rigorous mathematical point of view, whereas the other two books represent the physicist's approach.

6 Dynamics of statistical ensembles

We first derive a fundamental equation that describes how a whole collection of points in phase space evolves in time under iteration of a mapping.

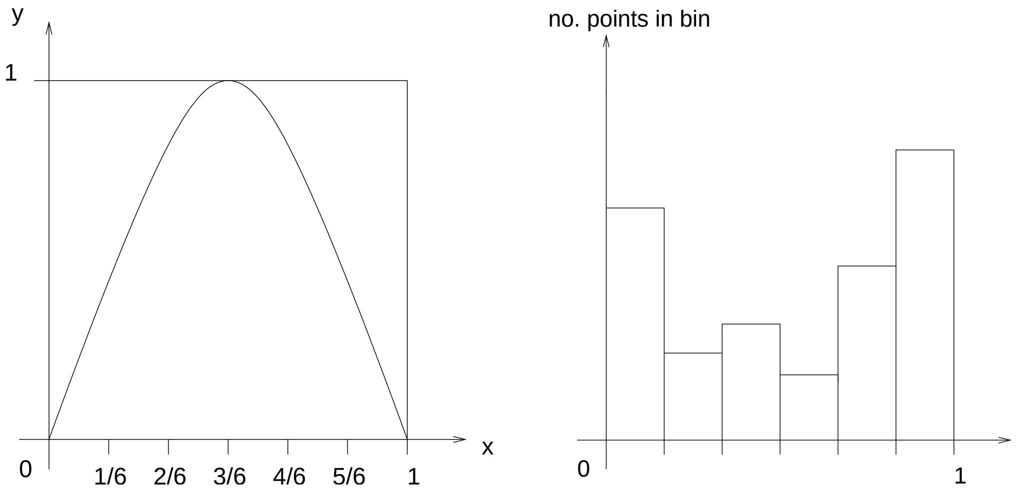


Figure 6.1: Left: The logistic map for control parameter $r = 4$ with $M = 6$ bins; right: Histogram for an ensemble of points iterated n times in the logistic map, computed by using these bins.

6.1 The Frobenius-Perron equation

Let us introduce this equation by means of a simple example:

Example 23 [Las94]
Let us consider again the logistic map Eq. (3.9), say, at the parameter value $r = 4$ for which it reads $L(x) = 4x(1-x)$, $x \in [0,1]$, see Fig 6.1 left for a picture. Let us now sample the nonlinear dynamics of this map by doing statistics in the following way:

1. Take what is called a *statistical ensemble* in phase space, which here are N points with initial conditions $x_0^i,\ i = 1, \ldots, N$ that are, e.g., uniformly distributed on $[0,1]$.[1]
2. *Iterate* all x_0^i according to the usual equation of motion $x_{n+1}^i = L(x_n^i)$.
3. Split $[0,1]$ into M *bins* $[\frac{j-1}{M}, \frac{j}{M}), j = 1, \ldots, M$.
4. Construct *histograms* by counting the fraction $N_{n,j}$ of the N points that are in the jth bin at time step n. This yields a result like the one depicted in Fig. 6.1 right.

This very intuitive approach, which is easily implemented on a computer, motivates the following definition:

Definition 29

$$\rho_n(x_j) := \frac{\textit{number of points } N_{n,j} \textit{ in bin at position } x_j \textit{ at time step } n}{\textit{total number of points } N \textit{ times bin width } \delta x = \frac{1}{M}} \tag{6.1}$$

where $x_j := \frac{1}{2}(\frac{j-1}{M} + \frac{j}{M}) = \frac{2j-1}{2M}$ *is the centre of the bin, defines a* probability density.

[1] The idea of a statistical ensemble is that one looks at a "swarm" of representative points in the phase space of a dynamical system for sampling its statistical properties; see [Bec93] or Wikipedia in the internet for further information.

Remark 11

1. $\rho_n(x_j)$ is *normalised*,

$$\sum_{j=1}^{M} \rho_n(x_j)\delta x = \frac{\sum_j N_{n,j}\delta x}{N\delta x} = \frac{N}{N} = 1 \quad , \tag{6.2}$$

   and by definition $\rho_n(x_j) \geq 0$.

2. For $M \to \infty$ (which implies $\delta x \to 0$) and $N \to \infty$ the sum can be replaced by the continuum limit,

$$\sum_{j=1}^{M} \rho_n(x_j)\delta x \to \int_0^1 \mathrm{d}x \rho_n(x) = 1 \quad , \tag{6.3}$$

   which as normalisation may look more familiar to you.

The question is now whether there exists an equation that describes the *dynamics* of $\rho_n(x)$ in the limits of infinitesimally small bins, $\delta x \to 0$, and an infinite number of points, $N \to \infty$. That is, we want to understand how $\rho_n(x)$ changes with discrete time n by iterating the map. The answer is provided by the *Frobenius-Perron equation*. Here we derive it in an intuitive way, based on the idea of the conservation of the number of points for an ensemble under iteration [Ott93]. Conservation of probability implies that[2]

$$\rho(y)\delta y = \rho(x)\delta x \tag{6.4}$$

with $y = F(x)$ if the map under consideration is one-to-one. However, a chaotic map such as, for example, the Bernoulli shift, is not injective. According to Fig. 6.2 we thus have to modify Eq. (6.4) to

$$\rho(y)\delta y = \sum_{y=F(x^i)} \rho(x^i)\delta x^i \quad . \tag{6.5}$$

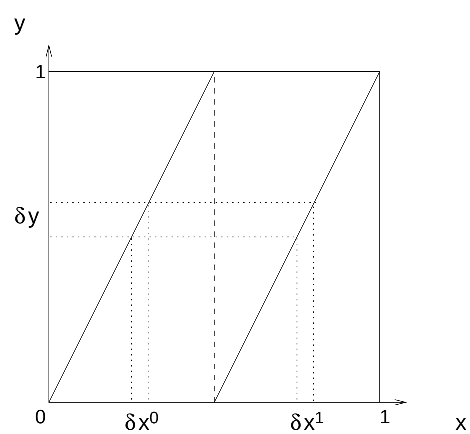


Figure 6.2: Idea of deriving the Frobenius-Perron equation, demonstrated for the example of the Bernoulli shift.

[2] In probability theory this is called transformation of variables.

This leads to

$$\begin{aligned} \rho(y) &= \sum_i \rho(x^i) \frac{\delta x^i}{\delta y} \\ &= \sum_i \rho(x^i) \frac{1}{\delta y / \delta x^i} \\ &= \sum_i \rho(x^i) \frac{1}{|F'(x^i)|} \quad (\delta x^i \to 0) \end{aligned} \tag{6.6}$$

for which we have assumed that the map is at least piecewise differentiable and $F'(x^i) \neq 0$ except at single points. We now identify $\rho_{n+1}(x) = \rho(y)$ and $\rho_n(x) = \rho(x)$ yielding our final result:

Definition 30 *Let F be piecewise C^1 and $\rho_n(x)$ be the corresponding probability density. Then*

$$\rho_{n+1}(x) = \sum_{x=F(x^i)} \rho_n(x^i)|F'(x^i)|^{-1} \tag{6.7}$$

is called the Frobenius-Perron equation *(FPE) of the map.*

Remark 12

1. In physical terms this is an example of a *continuity equation* expressing the conservation of the number of points in the phase space of a dynamical system.

2. This equation has the form of a recursive functional equation and can thus be solved by (numerical) iteration.

3. It can be generalized to arbitrary dimension and to continuous time [Dor99]. In the latter case we arrive at the perhaps more famous *Liouville equation*, so Eq. (6.7) may also be considered as the "Liouville equation for time-discrete maps".

The FPE can be written in the form of [Bec93]

$$\rho_{n+1}(x) = P\rho_n(x) \quad , \tag{6.8}$$

where P is the *Frobenius-Perron operator* acting on ρ_n,

$$P\rho_n(x) := \sum_{x=F(x^i)} \rho_n(x^i)|F'(x_n)|^{-1} \quad . \tag{6.9}$$

Eq. (6.8) makes clear that there exists a formal analogy between the FPE and the equation of motion $x_{n+1} = F(x_n)$ of a map. Note, however, that the latter equation determines the dynamics of *single points*, whereas the FPE describes the dynamics of a *statistical ensemble* of them.
It is easy to show (see coursework) that P exhibits the following two properties:

1. *Linearity*:
$$P(\lambda_1\rho^1 + \lambda_2\rho^2) = \lambda_1 P\rho^1 + \lambda_2 P\rho^2 \; , \; \lambda_1, \lambda_2 \in \mathbb{R} \tag{6.10}$$

2. *Positivity*:

$$P\rho \geq 0 \quad \text{if} \quad \rho \geq 0 \tag{6.11}$$

So we are dealing with a linear positive operator. We could explore further mathematical properties of it along these lines. For this purpose we first would have to define the function space on which P is acting.[3] We could then characterize this operator, for example, in terms of Markov and semigroup properties. However, these are advanced topics of functional analysis on which we do not want to elaborate in these lectures. Ref. [Las94] gives a nice introduction if you are interested in this direction of research.
Focusing onto more applied topics, an important question can be posed on the basis of the following definition.

Definition 31 *Any function $\rho^*(x)$ that is* invariant *under P,*

$$\rho^*(x) = P\rho^*(x) \tag{6.12}$$

is called a fixed point *of P. $\rho^*(x)$ is also called an* invariant density*, which expresses the fact that it does not change with time n.*

A crucial problem is now to *calculate* $\rho^*(x)$ for a given map, which requires to solve the FPE Eq. (6.7). This is of fundamental interest in the theory of dynamical systems, because $\rho^*(x)$ provides a *statistical characteristisation* of a map in terms of what eventually happens to an ensemble of points. However, to obtain analytical solutions is only possible in rather special cases.
In the next section we will learn about a basic method that works for certain classes of piecewise linear maps. However, let us first illustrate the above concept of the Frobenius-Perron operator and invariant densities for the simplest example we can think of:

Example 24 [Bec93]
For the Bernoulli shift $B(x) = 2x \mod 1$ the FPE Eq. (6.7) reads

$$P\rho(x) = \frac{1}{|B'(x^1)|}\rho(x^1) + \frac{1}{|B'(x^2)|}\rho(x^2) \tag{6.13}$$

with

$$\begin{aligned} x &= B(x^i)\,, \quad i = 1,2 \\ &= \begin{cases} 2x^1, & 0 \leq x^1 < \frac{1}{2} \\ 2x^2 - 1, & \frac{1}{2} \leq x^2 < 1 \end{cases} \quad , \end{aligned} \tag{6.14}$$

see Fig. 6.3. However, we can do better by explicitly constructing the Frobenius-Perron operator P for this map, as follows: We first observe that $B'(x) = 2$ (except at $x = 1/2$). We can then calculate the two points x^1 and x^2 as functions of x by *piecewise* inverting Eq. (6.14) to $x^i = B^{-1}(x)$. This leads to

$$\begin{aligned} x^1 &= B^{-1}(x) = \frac{x}{2}\,,\ 0 \leq x^1 < \frac{1}{2} \quad \text{and} \quad 0 \leq x < 1 \\ x^2 &= B^{-1}(x) = \frac{x+1}{2}\,,\ \frac{1}{2} \leq x^2 < 1 \quad \text{and} \quad 0 \leq x < 1 \quad . \end{aligned} \tag{6.15}$$

[3] We wish to work on a Lebesgue space of measurable functions, which is at least L^1.

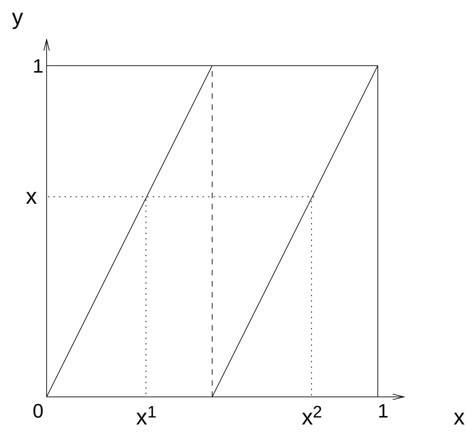


Figure 6.3: Idea of deriving the Frobenius-Perron equation, demonstrated for the example of the Bernoulli shift.

Putting all pieces together we obtain

$$P\rho(x) = \frac{1}{2}\rho\left(\frac{x}{2}\right) + \frac{1}{2}\rho\left(\frac{x+1}{2}\right) \quad , \tag{6.16}$$

which represents the explicit form of the Frobenius-Perron operator for $B(x)$. Let us now look for a fixed point of the FPE of this map, which for this operator must satisfy

$$\rho^*(x) = P\rho^*(x) = \frac{1}{2}\rho\left(\frac{x}{2}\right) + \frac{1}{2}\rho\left(\frac{x+1}{2}\right) \quad . \tag{6.17}$$

A naive guess might be that the invariant density for this map is simply uniform on the whole interval, $\rho^*(x) = c$, where c is a constant (why?). We can easily check this hypothesis by plugging our *ansatz* into the above equation yielding $\rho^*(x) = c = \frac{1}{2}c + \frac{1}{2}c$, which verifies that our guess is correct. Since $\rho^*(x)$ is a probability density it must be normalized, which for a distribution that is uniform on the unit interval implies $\rho^*(x) = 1$. So we can conclude that $B(x)$ has an invariant density of $\rho^*(x) = 1$.

Remark 13

1. We have obtained one fixed point of the FPE, however, is this the only one? That poses the question about uniqueness of the solution. Or in other words, if we choose different initial densities $\rho_0(x)$ and iteratively solve the FPE for this map, is there any dependence on initial densities in that these densities converge to different invariant densities?

   That such a situation can easily occur is illustrated by the "double Bernoulli shift" depicted in Fig. 6.4: Start, for example, with a density that is concentrated on $0 \le x \le 1/2$ only but is zero otherwise. Then it will converge to a uniform density on the same interval while it is zero otherwise. On the other hand, if we choose a density that is concentrated on $1/2 \le x \le 1$ only but is zero otherwise, the invariant density will be uniform on the latter interval while it is zero otherwise. Consequently, we have here a map that does not exhibit a unique invariant density – what we get depends on our choice of the initial density.

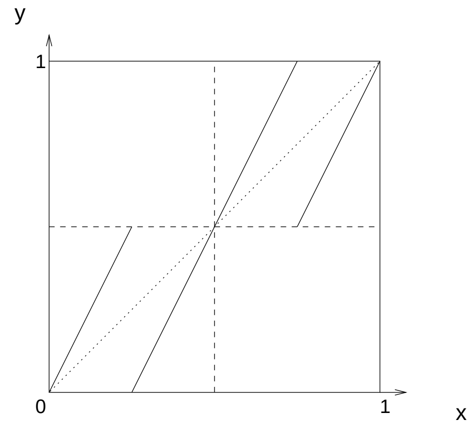


Figure 6.4: The “double Bernoulli shift” as an example of a map that has no unique invariant density.

In general, the question about existence and uniqueness of an invariant density for a given map is difficult to answer [Las94, Bec93]. The result also depends on what kind of density we are looking for, as characterized by further mathematical properties like continuity, differentiability etc. that we may require for our result. We will get back to these questions in the following sections, which will give us some further insight.

2. In case of the Bernoulli shift we have solved the FPE for $\rho^*(x)$ by just guessing a correct solution. In the following section we will present a method by which, for a certain class of maps, one can systematically solve the FPE for obtaining invariant densities.

6.2 Markov partitions and transition matrices

Before we solve the FPE, we need to introduce a number of important concepts.

Definition 32 *partition, formally [Bec93]*
A partition *of the interval $I \subset \mathbb{R}$ is a (finite or countable) set $\{I_i\}$, $i = 1, \dots, M$, of subintervals I_i such that*

1. $I = \bigcup_{i=1}^{M} I_i$, *that is,* $\{I_i\}$ *is a* complete *subdivision of* I
2. $\mathrm{int} I_i \cap \mathrm{int} I_j = \emptyset \; (i \neq j)$, *that is, there is only overlap at the boundaries.*

Alternatively, there is the following verbal definition available:

Definition 33 *partition, verbally [All97]*
A partition *of the interval I is a collection of subintervals (also called parts, or cells of the partition) whose union is I, which are pairwise disjoint except perhaps at the end points.*

Example 25 Consider the unit interval $I = [0, 1]$. Then $\{I_i\}_{i=1,2} = \{I_1, I_2\} = \{[0, \frac{1}{2}], [\frac{1}{2}, 1]\}$ is a partition; see Fig. 6.5. Note that one can find infinitely many different partitions here.

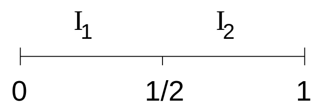


Figure 6.5: Simple example of a partition of the unit interval.

An interesting question is now what happens to partition parts if we let a map act onto them. Let us study this by means of the example sketched in Fig. 6.6, where we apply the map $D(x)$ defined in the figure on the previous sample partition. The map then yields

$$D(I_1) = I \quad = I_1 \cup I_2 \quad \text{and} \quad D(I_2) = I_1 \quad . \tag{6.18}$$

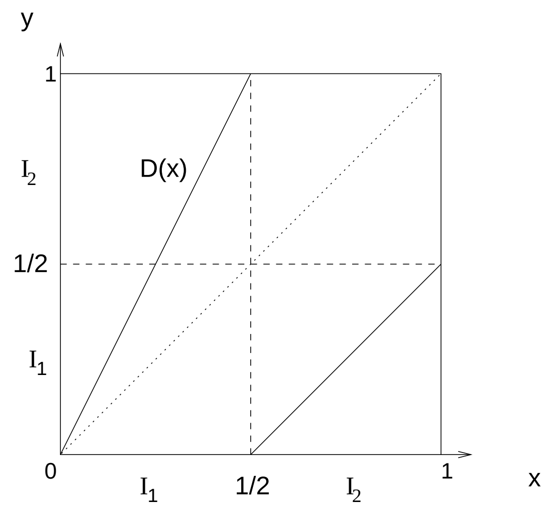


Figure 6.6: The map $D(x)$ applied onto a partition consisting of the two parts $\{I_1, I_2\}$.

Note that here partition parts or unions of them are recovered *exactly* under the action of the map. This is the essence of what is called a *Markov property* of a map:

Definition 34 *Markov partition, verbally*
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*.*

If you like formal definitions, you may prefer the following one:

Definition 35 *Markov partition, formal [Kat95, Guc90]*
Let $\{I_i\}_{i=1,\dots,M}$ *be a partition of some compact interval* $I \subset \mathbb{R}$ *and let* F *be a map acting on* I, $F : I \to I$. *If* $\forall i, j \, [\text{int}F(I_j) \cap \text{int}I_i \neq \emptyset \Rightarrow \text{int}I_i \subseteq \text{int}F(I_j)]$, *then the partition is* Markov*.*

Example 26 Let us check this formal definition for the map $D(x)$ of Fig. 6.6. We just have to verify that the logical implication holds for the four cases

$$\begin{array}{lcll} \mathrm{int}D(I_1)\cap \mathrm{int}I_1 \neq \emptyset & \Rightarrow & \mathrm{int}I_1 \subseteq \mathrm{int}D(I_1). & \checkmark \\ \mathrm{int}D(I_1)\cap \mathrm{int}I_2 \neq \emptyset & \Rightarrow & \mathrm{int}I_2 \subseteq \mathrm{int}D(I_1). & \checkmark \\ \mathrm{int}D(I_2)\cap \mathrm{int}I_1 \neq \emptyset & \Rightarrow & \mathrm{int}I_1 \subseteq \mathrm{int}D(I_2). & \checkmark \\ \mathrm{int}D(I_2)\cap \mathrm{int}I_2 = \emptyset & & \checkmark & \end{array} \tag{6.19}$$

It is elucidating to discuss an example of a map with a partition that is not Markov:

Example 27 Let us look at the map E defined in Fig. 6.7.

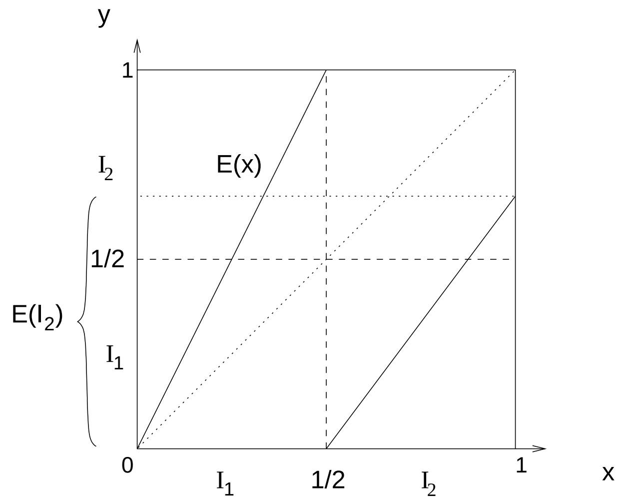


Figure 6.7: An example of a map $E(x)$ with a partition that is not Markov.

The first branch of the map does not cause any problem, since it is identical with the one of our previous map $D(x)$. But let us look at the action of the second branch:

$$\begin{array}{lcll} \mathrm{int}E(I_2)\cap \mathrm{int}I_1 \neq \emptyset & \Rightarrow & \mathrm{int}I_1 \subseteq \mathrm{int}E(I_2). & \checkmark \\ \mathrm{int}E(I_2)\cap \mathrm{int}I_2 \neq \emptyset & \Rightarrow & \mathrm{int}I_2 \subseteq \mathrm{int}E(I_2). & \times \text{ not true!} \end{array} \tag{6.20}$$

Consequently, the partition $\{I_1, I_2\}$ is not Markov for $E(x)$.

Remark 14

1. Maps for which there exists a Markov partition are often called *Markov maps*. However, note that sometimes further conditions such as hyperbolicity, smoothness, etc., are required in the literature.

2. Markov partitions appear to have first been constructed by Sinai in 1968 for the Sinai billiard. For such higher dimensional systems the principle of Markov partitions is the same as the one described above, however, the precise definition is technically a bit more involved than for one-dimensional maps [Kat95, Guc90].

3. Markov partitions are *not unique* for a given map, see the coursework.

4. Markov partitions are important because they define, loosely speaking, an 'invariant frame' according to which the action of a map on statistical ensembles can be described. They enable to map the statistical dynamics of a system onto a Markov process – hence the name –, which is precisely what we will explore in the following.

As usual, we start with a simple example:

Example 28 The map F shown in Fig. 6.8 is defined on a Markov partition (please convince yourself that this is the case). Let N_n^i be the number of points in partition part i at the nth iteration of this map. Let us start with a *uniform distribution* of $N = \sum_i N_n^i$ points on $[0, 1]$ at time step $n = 0$. As one can infer from the left part of Fig. 6.8, the dynamics of N_n^i is then given as follows:

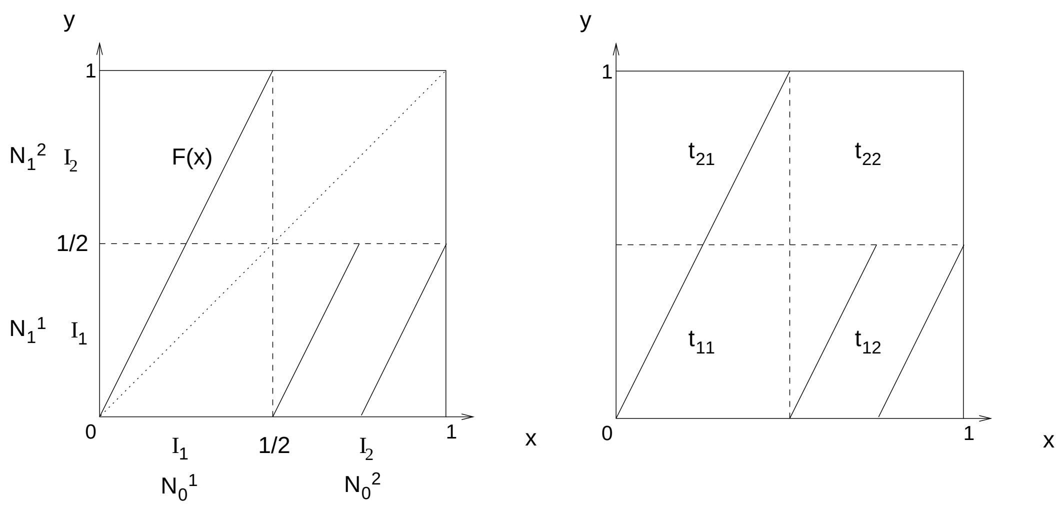


Figure 6.8: Left: A map that enables a matrix representation for the dynamics of a statistical ensemble of points; right: Illustration of where the transition matrix elements come from for this map.

$$
\begin{aligned}
N_1^1 &= \frac{1}{2}N_0^1 + N_0^2 &&= \frac{1}{2}[N_0^1 + 2N_0^2] \\
N_1^2 &= \frac{1}{2}N_0^1 + 0 \cdot N_0^2 &&= \frac{1}{2}N_0^1 \quad .
\end{aligned} \tag{6.21}
$$

This can equivalently be written as

$$
\begin{pmatrix} N_1^1 \\ N_1^2 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} N_0^1 \\ N_0^2 \end{pmatrix} \tag{6.22}
$$

or

$$
\underline{N}_{n+1} = \frac{1}{2}\,\underline{\underline{T}}_F\,\underline{N}_n \quad , \tag{6.23}
$$

where $\underline{N}_n \in \mathbb{R}^2$ is a *column vector* and $\underline{\underline{T}}_F \in M_{2,2}$ a *topological transition matrix* (or *transfer matrix*) acting on $\underline{N}_n$.

Let us focus for a moment on the (topological) structure of the matrix $\underline{\underline{T}}_F$. There is a simple recipe of how to obtain this matrix, as follows:

- Consider a map defined on a Markov partition such as F above.
- Identify the "grid" formed by the Markov partition under iteration of the map with the rows and columns of the desired matrix $\underline{\underline{T}}_F$, however, by turning the grid "upside down".
- Check whether there exist any "links" between partition part i and partition part j under iteration: That is, does the graph of the map intersect the cell labeled by t_{ij}? And if so, how many times?
- Count the number of intersections, which yields the value of the matrix element t_{ij}.

Applied to the above example of map F, this prescription leads directly to the transition matrix $\underline{\underline{T}}_F$ of Eq. (6.22). An alternative representation of a transition matrix for a map is obtained by using the following definition:

Definition 36 *Markov graph [Has03, All97, Rob95]*
Let $F : I \to I$, $I \subset \mathbb{R}$ be a map, $\{I_i\}$ a Markov partition and $\underline{\underline{T}}_F = ((t_{ij}))$, $i, j \in \{1, \dots, M\}$ be the corresponding transition matrix. In a Markov graph *t_{ij} arrows are drawn from j to i.*

Example 29 The Markov graph for the map F of Fig. 6.8 acting on the given Markov partition is sketched in Fig. 6.9.

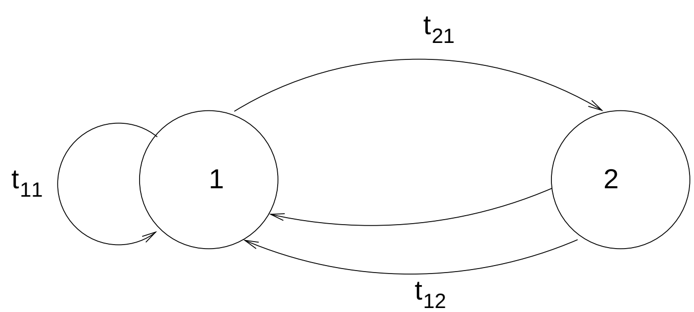


Figure 6.9: Markov graph for the map F on the given Markov partition.

Remark 15

1. What we encounter here is a representation of the action of F in terms of a *symbolic dynamics* of a map. The above Markov graph defines it with respect to the given Markov partition.[4]
2. Note that our definition is such that $t_{ij} \in \mathbb{N}_0$. In other literature you will usually find that $t_{ij} \in \{0, 1\}$.[5] Why our definition is a bit more convenient for practical purposes you will see later on.

[4] Other partitions could be used as well. Symbolic dynamics is a fundamental concept in dynamical systems theory, which would deserve much more detailed explanations. However, since this is again a topic of the module on Chaos and Fractals we do not further discuss it in our course and refer to the literature instead [Bec93, All97].

[5] see, e.g., work by Ruelle and Bowen

3. Markov graphs can be constructed without having any transition matrix $\underline{\underline{T}}_F$ available. It suffices to just look at how a map F acts onto a Markov partition. If you find it simpler, you may thus first sketch the Markov graph and then obtain $\underline{\underline{T}}_F$ from it.

The following definition now makes mathematically precise what we actually mean with a topological transition matrix:

Definition 37 *topological transition matrix (see, e.g., [Guc90])*
Let $\{I_i\}$ be a Markov partition and let the map F be invertible on I_i. Then the topological transition matrix $\underline{\underline{T}}_F = ((t_{ij}))$, $i,j \in \{1,\ldots,M\}$, *is defined by*

$$t_{ij} := \begin{cases} 1, & \mathrm{int} I_i \subseteq \mathrm{int} F(I_j) \\ 0 & \text{otherwise} \end{cases} . \tag{6.24}$$

If F is not *invertible on I_j, let $\{I_{jk}\}, k = 1,\ldots,L$ be a refinement of the Markov partition on I_j such that F is invertible on each I_{jk}. Then*

$$t_{ij} := \begin{cases} \textit{number of parts}\, I_{jk} \textit{ such that}\, \mathrm{int} I_i \subseteq \mathrm{int} F(I_{jk}) \\ 0 \textit{ otherwise} \end{cases} . \tag{6.25}$$

Example 30 Let us apply this definition to the map F already studied above. According to Fig. 6.10 F is invertible on I_1, and we obtain

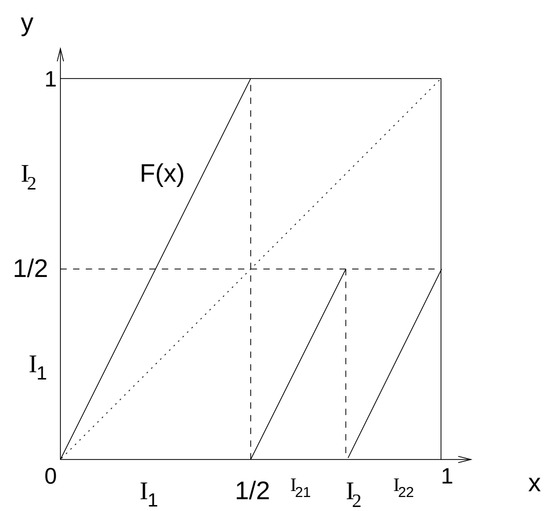


Figure 6.10: Construction of a transition matrix for the map F according to Definition 37.

$$\begin{aligned} \mathrm{int} I_1 \subseteq \mathrm{int} F(I_1) &\Rightarrow t_{11} = 1 \\ \mathrm{int} I_2 \subseteq \mathrm{int} F(I_1) &\Rightarrow t_{21} = 1 \end{aligned} \tag{6.26}$$

However, F is not invertible on I_2, hence we have to use a refinement on which F is invertible such as the one shown in the figure. On this finer partitioning we get

$$\begin{aligned} \mathrm{int} I_1 \subseteq \mathrm{int} F(I_{21}) \,,\ \mathrm{int} I_1 \subseteq \mathrm{int} F(I_{22}) &\Rightarrow t_{12} = 2 \\ \mathrm{int} I_2 \subseteq \mathrm{int} F(I_{21}) \text{ not true} \,,\ \mathrm{int} I_2 \subseteq \mathrm{int} F(I_{22}) \text{ not true} &\Rightarrow t_{22} = 0 \end{aligned} \tag{6.27}$$

Altogether we thus arrive at

$$\underline{\underline{T}}_F = \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} \quad , \tag{6.28}$$

as before.

Let us now reconsider the action of the above transition matrix $\underline{\underline{T}}_F$ onto the given Markov partition. So far you have been provided with several pieces of a puzzle that we can now put together. Recall Eq. (6.23) for the map F of Fig. 6.10, which read

$$\underline{N}_{n+1} = \frac{1}{2}\underline{\underline{T}}_F\,\underline{N}_n \quad .$$

We now divide the whole equation by the product of the total number of points N times $\mathrm{diam}(I_i)$, where according to Fig. 6.10 $\mathrm{diam}(I_i) = 1/2$. This yields

$$\underline{\rho}_{n+1} = \frac{1}{2}\,\underline{\underline{T}}_F\,\underline{\rho}_n \quad , \tag{6.29}$$

where the topological transition matrix $\underline{\underline{T}}_F$ acts onto *probability density vectors* $\underline{\rho}_n, \underline{\rho}_{n+1}$, whose components are constant on the Markov partition parts.
So what did we achieve with Eq. (6.29)? Using Markov partitions for a piecewise linear map enabled us to construct the very simple version of the FPE Eq. (6.7), where the Frobenius-Perron operator is represented as a matrix and the probability densities are piecewise constant step functions in form of vectors. We may expect that such an equation can be solved much more easily than the FPE in its original form. However, before we explore this we remark that the above construction can be performed under fairly general conditions:

Proposition 5 *Let $F : I \to I$, $I \subseteq \mathbb{R}$ be a piecewise linear map with uniformly constant slope $|F'(x)| = const.$. If there exists a topological transition matrix $\underline{\underline{T}}_F$, and by restricting ourselves to probability density vectors $\underline{\rho}_n$ that are constant on Markov partition parts, the FPE of this map can be written as*

$$\underline{\rho}_{n+1} = \frac{1}{|F'(x)|}\,\underline{\underline{T}}_F\,\underline{\rho}_n \quad . \tag{6.30}$$

The proof of this proposition is a bit too lengthy for our lectures, but we outline the basic idea: First of all, note that for our example of map F considered above the corresponding Frobenius-Perron operator $1/2\,\underline{\underline{T}}_F$ actually defines a *stochastic matrix*, where the sum of all entries of a column is equal to one, $\sum_{i=1}^{M} t_{ij}/2 = 1$.[6] However, this case is rather special – consider, for example, map G depicted in Fig. 6.11. Based on the information given in the figure one can show that the constant slope $|G'(x)| \notin \mathbb{N}$. It is then easily seen that the corresponding matrix $\underline{\underline{T}}_G/|G'(x)|$ cannot be a stochastic one, which refers to the typical situation.
Nevertheless, this connection with stochastic matrices points in the right direction: If the conditions of the above proposition are fulfilled there exists a systematic procedure of how to

[6] We remark that there is no unique definition of a stochastic matrix in the literature: some authors define it as above [Kat95], others require that the sum of all row entries is one, instead of the column entries [Bec93]. Some authors identify a stochastic matrix with a *Markov matrix*, others distinguish between them depending on whether the row or the column entries add up to one.

construct the Frobenius-Perron operator in terms of a transition matrix that is generally not a topological but a stochastic one, which acts on probabilities defined on Markov partition parts, see Section 17.5 of Ref. [Bec93] for details. For uniformly constant slope this stochastic matrix can be simplified to a topological transition matrix divided by the absolute value of the slope, which acts onto probability density vectors as given in the proposition.

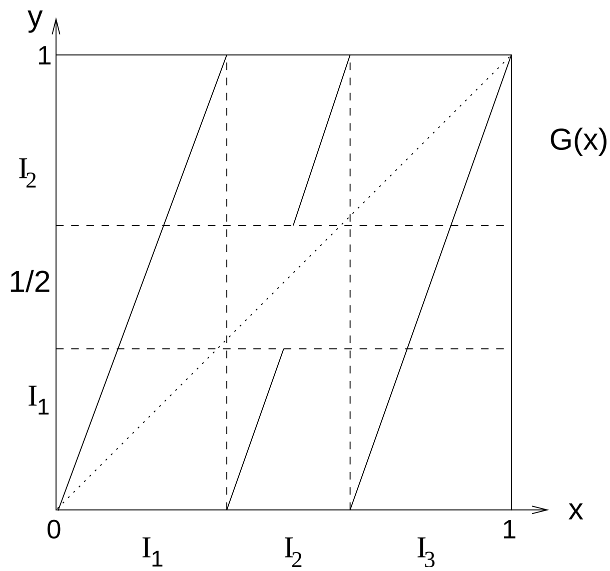


Figure 6.11: Example of a map G yielding a topological transition matrix that is not directly related to a stochastic one.

Remark 16 For a single vector component Eq. (6.30) reads

$$\rho_{n+1}^{i} = c \sum_{j=1}^{M} t_{ij} \, \rho_n^j \quad , \tag{6.31}$$

where $c = const$. Note that all matrix elements t_{ij} defining the transition from states j at time step n to state i at time step $n+1$ are constants. That is, they do not depend on the history of the process at all previous time steps but are defined with respect to the states i and j at time steps $n+1$ and n only, which is the hallmark of a *Markov process*. Proposition 5 thus transforms the dynamics generated by the FPE onto a *Markov chain* [Bec93]. Note, however, that any transition matrix exhibits a specific topological structure, which to some extent characterizes a given map. This structure is encoded in the values of the t_{ij}'s, and summing over all these matrix elements can make the corresponding Markov chain arbitrarily complicated.[7]

If one can transform a Frobenius-Perron operator onto a topological transition matrix, the resulting Frobenius-Perron matrix equation Eq. (6.30) can be used for calculating the invariant density ρ^* in form of the following algorithm:

[7]Interestingly, the topological structure of the transition matrix for a given map can change drastically under tiny variation of control parameters. Such a topological instability of a dynamical system produces a sensitivity of the whole dynamics with respect to control parameters [Kla07].

1. For a given map F defined on a Markov partition construct $\underline{\underline{T}}_F$.
2. Solve the *eigenvalue problem* of $\underline{\underline{T}}_F$

$$\underline{\underline{T}}_F \underline{\rho}_\tau = \tau \underline{\rho}_\tau \quad , \tag{6.32}$$

where τ is an eigenvalue and $\underline{\rho}_\tau$ the corresponding eigenvector. If you did this calculation correctly, you will find that the largest eigenvalue is always equal to $|F'(x)|$,[8]

$$\tau_{\max} = |F'(x)| \quad , \tag{6.33}$$

with eigenvector $\underline{\rho}_{\tau_{\max}}$.

3. It follows: Choose

$$\underline{\rho}_n = c\,\underline{\rho}_{\tau_{\max}} \quad , \tag{6.34}$$

where c is a constant. Then according to the Frobenius-Perron matrix equation Eq. (6.30)

$$\underline{\rho}_{n+1} = \frac{c}{|F'|}\underline{\underline{T}}_F \underline{\rho}_{\tau_{\max}} \overset{\text{Eq. (6.32)}}{=} c\frac{\tau_{\max}}{|F'|}\underline{\rho}_{\tau_{\max}} \overset{\text{Eq. (6.33)}}{=} c\underline{\rho}_{\tau_{\max}} \overset{\text{Eq. (6.34)}}{=} \underline{\rho}_n \quad , \tag{6.35}$$

from which we can draw the important conclusion that

$$c\,\underline{\rho}_{\tau_{\max}} = \underline{\rho}^* \quad , \tag{6.36}$$

that is, the largest eigenvector is proportional to the invariant density.

4. It then remains to determine c by normalisation,

$$\sum_{i=1}^{m} c\,\rho^i_{\tau_{\max}} \cdot \operatorname{diam}(I_i) = 1 \quad . \tag{6.37}$$

This recipe of how to calculate invariant densities has a rigorous mathematical underpinning:

Theorem 4 *Perron-Frobenius Theorem [Kat95, Rob95]*
Let $\underline{\underline{T}}_F$ be such that for a certain power $\underline{\underline{T}}_F^n$ all entries are positive.[9] *Then $\underline{\underline{T}}_F$ has one (up to a scalar) eigenvector $\underline{\rho}_{\tau_{max}}$ with positive components and no other eigenvectors with nonnegative components. The eigenvalue τ_{max} corresponding to $\underline{\rho}_{\tau_{max}}$ is simple, positive and greater than the absolute values of all other eigenvalues.*

The proof of this important theorem is non-trivial, see the references cited above. If applied onto stochastic matrices it directly justifies our 'recipe' given above, see p.157 of [Kat95]. Note that if the conditions of this theorem are fulfilled it guarantees the existence of a unique invariant probability density vector, which gives an answer to the question posed in Remark 13.

Example 31 We now apply the above method for calculating invariant measures of piecewise linear Markov maps to our previous map F introduced in Example 28:

[8] This fact is not trivial but a consequence of the following Proposition 4.
[9] A nonnegative matrix that fulfills this property is called *eventually positive*.

1. The first step in this algorithm we have already performed before, see Examples 28 and 30. They yielded the Frobenius-Perron matrix equation

$$\underline{\rho}_{n+1} = \frac{1}{2} \underline{\underline{T}}_F \, \underline{\rho}_n \quad , \tag{6.38}$$

see Eq. (6.29), where

$$\underline{\underline{T}}_F = \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} \tag{6.39}$$

and the probability density vectors are defined by

$$\underline{\rho}_n = \begin{pmatrix} \rho_n^1 \\ \rho_n^2 \end{pmatrix} \in \mathbb{R}^2 \quad , \tag{6.40}$$

where $\rho_n^i = $ const. on the ith Markov partition part.

2. Solve now the eigenvalue problem

$$\underline{\underline{T}}_F \underline{\rho}_\tau = \tau \underline{\rho}_\tau \tag{6.41}$$

by calculating

$$\det(\underline{\underline{T}}_F - \tau \underline{\underline{I}}) = 0 \quad , \tag{6.42}$$

where

$$\underline{\underline{I}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{6.43}$$

is the identity matrix. We find

$$\det \begin{pmatrix} 1-\tau & 2 \\ 1 & -\tau \end{pmatrix} = 0, \quad (1-\tau)(-\tau) - 2 = 0$$

$$-\tau + \tau^2 - 2 = 0 \Rightarrow \tau = \frac{1}{2} \pm \sqrt{\frac{1}{4} + 2} = \frac{1}{2} \pm \frac{3}{2} = 2 \text{ or } -1 \quad . \tag{6.44}$$

We see that indeed $\tau_{\max} = |F'(x)|$, as stated to the end of step 2 of our algorithm.

3. We now have to calculate the largest eigenvector by solving the equation

$$\begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \rho_{\tau_{\max}}^1 \\ \rho_{\tau_{\max}}^2 \end{pmatrix} = 2 \begin{pmatrix} \rho_{\tau_{\max}}^1 \\ \rho_{\tau_{\max}}^2 \end{pmatrix} \quad , \tag{6.45}$$

which leads to

$$\begin{aligned} \rho_{\tau_{\max}}^1 + 2\rho_{\tau_{\max}}^2 &= 2\rho_{\tau_{\max}}^1 \\ \rho_{\tau_{\max}}^1 &= 2\rho_{\tau_{\max}}^2 \quad . \end{aligned} \tag{6.46}$$

Assuming that $\rho_{\tau_{\max}}^1 = 1$ gives

$$\rho_{\tau_{\max}}^2 = \frac{1}{2}\rho_{\tau_{\max}}^1 = \frac{1}{2} \tag{6.47}$$

from which we obtain

$$\underline{\rho}_{\tau_{\max}} = \begin{pmatrix} 1 \\ \frac{1}{2} \end{pmatrix} \quad . \tag{6.48}$$

4. The last step is to calculate $\underline{\rho}^*$ by normalisation: It is

$$\underline{\rho}^* = c\underline{\rho}_{\tau_{\max}} \tag{6.49}$$

with

$$1 = \sum_{i=1}^{2} \rho^{*i} \cdot \mathrm{diam}(I_i) = c \sum_{i=1}^{2} \rho^i_{\tau_{\max}} \cdot \mathrm{diam}(I_i) = c\frac{3}{4} \tag{6.50}$$

from which follows $c = 4/3$. Our final result thus reads

$$\underline{\rho}^* = \begin{pmatrix} \frac{4}{3} \\ \frac{2}{3} \end{pmatrix} \quad . \tag{6.51}$$

It is not a bad idea to cross-check whether our solution indeed fulfills the Frobenius-Perron matrix equation Eq. (6.38),

$$\begin{pmatrix} \frac{4}{3} \\ \frac{2}{3} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} \frac{4}{3} \\ \frac{2}{3} \end{pmatrix} = \begin{pmatrix} \frac{2}{3} + \frac{2}{3} \\ \frac{2}{3} \end{pmatrix} \quad . \quad \checkmark \tag{6.52}$$

7 Measure-theoretic description of dynamics

So far we have analyzed the dynamics of statistical ensembles by using probability densities, which represents rather a physicist's approach towards a probabilistic theory of dynamical systems. Mathematicians prefer to deal with *measures*, which are often more well-behaved than probability densities.

In this chapter we introduce these quantities, learn about basic properties of them and eventually briefly elaborate on some ergodic properties of dynamical systems. The first section particularly draws on Ref. [Las94], the second one on Refs. [Dor99, Arn68]. If you have problems to associate any meaning with measures you may wish to take a look into Ref. [All97].

7.1 Probability measures

Let $\rho(x),\ x \in I$, be a probability density, see Definition 29. If $\rho(x)$ exists and is integrable on a subinterval $A \subseteq I$ then

$$\mu(A) := \int_A \mathrm{d}x \rho(x) \tag{7.1}$$

is the *probability* of finding a point in A. $\mu(A)$ is called a *measure* of A, in the sense that we 'assign a positive number to the set A'. Let us demonstrate this idea for a simple example:

Example 32 Let $I = [0,1]$ and $A = [0, \frac{1}{2}]$ with $\rho(x) = 1$ – think, for example, of the Bernoulli shift. Then

$$\mu(A) = \int_0^{\frac{1}{2}} dx\, 1 = \frac{1}{2} \tag{7.2}$$

is the probability of finding a point in A.

We make this basic idea mathematically precise in terms of two definitions:

Definition 38 *[Kat95, Las94]*
*A nonempty countable collection $\mathscr{A}$ of subsets of a set I is called a σ-*algebra[1] *if*

1. *$A \in \mathscr{A} \Rightarrow I \setminus A \in \mathscr{A}$ complementarity*

2. *$A_1, A_2, \ldots, A_N \in \mathscr{A} \Rightarrow \bigcup_k A_k \in \mathscr{A}$ additivity*

Example 33 Consider the three intervals $I = [0,1]$, $A_1 = [0, 1/2)$, $A_2 = [1/2, 1]$. Is $\mathscr{A} = \{A_1, A_2\}$ a σ-algebra? Let us check the two parts of the above definition:

1. $A_1 \in \mathscr{A} \Rightarrow I \setminus A_1 = A_2 \in \mathscr{A}$ ✓ and *vice versa* for A_2.

2. $A_1, A_2 \in \mathscr{A} \Rightarrow A_1 \cup A_2 = I \in \mathscr{A}$ is not the case.

Even for an amended set $\{A_1, A_2, I\}$ the first condition yields $I \setminus I = \emptyset$, which again is not an element of this new set. We thus find that $\tilde{\mathscr{A}} = \{A_1, A_2, I, \emptyset\}$ must be a σ-algebra of I. It is a general feature of σ-algebras that the whole set I as well as the empty set must always be elements of the algebra.

Definition 39 *[Las94]*
A function $\mu : \mathscr{A} \to \mathbb{R}_+ \cup \infty$ defined on a σ-algebra $\mathscr{A}$ is called a measure *if*

1. *$\mu(A) \geq 0 \;\; \forall A \in \mathscr{A}$.*

2. *$\mu(\emptyset) = 0$.*

3. *If $A_1, A_2, \ldots, A_n \in \mathscr{A}$ and $A_i \cap A_j = \emptyset \; (i \neq j)$ then*

$$\mu(\bigcup_k A_k) = \sum_k \mu(A_k) \quad \sigma - additivity \quad . \tag{7.3}$$

Note the analogy between this definition and Kolmogorov's axioms of probability theory.

Example 34 Let $\tilde{\mathscr{A}} = \{A_1, A_2, I, \emptyset\}$ as before. Let $\mu(A)$ be the length of the interval A; for example, if $A = [a, b)$ then $\mu(A) := b - a$. Is this function a measure? Let us check the above definition step by step:

1. $\mu(A) \geq 0 \, \forall A \in \tilde{\mathscr{A}}$ is fulfilled, since a length can never be less than zero.

2. $\mu(\emptyset) = 0$ is also fulfilled.

[1] The symbol σ refers to the fact that condition 2 applies to a countably infinite number of subintervals, whereas for an ordinary algebra condition 2 applies to any two sets only, i.e. $n = 2$.

3. For the third condition we have to check all the different cases:

$$\begin{aligned}
A_1 \cap A_2 = \emptyset \quad &\Rightarrow \quad \mu(A_1 \cup A_2) = \mu(I) = 1 = \mu(A_1) + \mu(A_2) = \frac{1}{2} + \frac{1}{2} \quad \checkmark \\
A_1 \cap I \neq \emptyset \quad & \qquad \checkmark \\
A_2 \cap I \neq \emptyset \quad & \qquad \checkmark \\
A_1 \cap \emptyset = \emptyset \quad &\Rightarrow \quad \mu(A_1 \cup \emptyset) = \mu(A_1) = \frac{1}{2} = \mu(A_1) + \mu(\emptyset) \quad \checkmark \quad ,
\end{aligned} \tag{7.4}$$

since $\mu(\emptyset) = 0$. The same holds for $A_2 \cap \emptyset$ and $I \cap \emptyset$.

We can conclude that $\mu(A)$ is a measure. This specific measure, which is defined by the length of the interval, is called the *Lebesgue measure*. From now on we will denote it by μ_L.

These considerations motivate the following definition:

Definition 40 *[Kat95] A* measure space *is a triple $(I, \mathscr{A}, \mu)$ of a set I, a σ-algebra $\mathscr{A}$ of subsets of I and a measure μ.*

In the following we will always work on measure spaces, mostly without saying it explicitly. For the example above $(I, \mathscr{A}, \mu_L)$ defines a measure space. We will furthermore often require the following property:

Definition 41 *μ is called a* probability measure *on I if $\mu(I) = 1$.*

If we think of μ in terms of Eq. (7.1), this refers to nothing else than normalisation.

Example 35 For any Lebesgue measure μ_L on $I = [0, 1]$ it is $\mu_L(I) = 1$, hence it defines a probability measure. However, note that this does not hold for, say, $\tilde{I} = [0, 2]$.

Remark 17 We emphasize that μ_L is only one example of a measure. There exist plenty of other types of measure such as, for example, the Dirac measure discussed in the coursework [Las94].

Let us get back to our motivation Eq. (7.1). Let $\rho(x)$ be a probability density on $I \subseteq \mathbb{R}$ and let $\mathscr{A}$ be a σ-algebra on the same interval. Then $\mu(A) := \int_A dx \rho(x)$, $A \in \mathscr{A}$, see Eq. (7.1), defines a probability measure on I. That this holds true can be verified straightforwardly by applying Definition 39 and using properties of Riemann integration.
Let us now assume that $\rho(x)$ is generated by a map such as, e.g., the Bernoulli shift, see Fig. 7.1. We have the Frobenius-Perron equation (FPE)

$$\rho_{n+1}(x) = P_F \, \rho_n(x) \tag{7.5}$$

with Frobenius-Perron operator P_F for a map F governing the time evolution of the probability densities $\rho_n(x)$. The above choice of $\mu(A)$ raises the question whether there exists an equation for the evolution of this measure associated with a density that is analogous to the FPE.
In order to answer this question we employ again the principle of conservation of probability as for our derivation of the FPE in Section 6.1. With $A = F \circ F^{-1}(A)$ for a one-dimensional map F, where $F^{-1}(A)$ defines the preimages of the set A, the number of points x_n in $F^{-1}(A)$

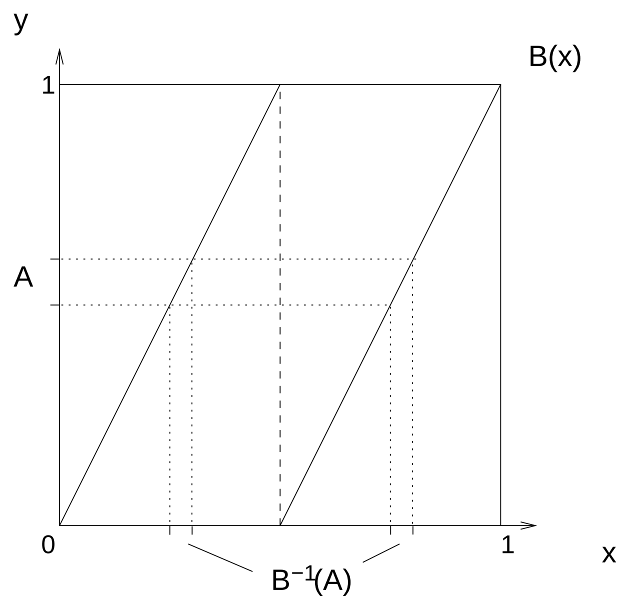


Figure 7.1: Construction of the analogue of the FPE for the measure obtained by integrating a probability density.

must be equal to the number of points x_{n+1} in A. This implies that the probability measures $\mu(A) = \int_A dx \rho(x)$ on image and preimages must be the same under iteration,

$$\mu_{n+1}(A) = \mu_n(F^{-1}(A)) \quad . \tag{7.6}$$

See the Bernoulli shift $B(X) = 2x \mod 1$ in Fig. 7.1 for an illustration. We thus arrive at the measure-theoretic version of the FPE Eq. (6.7) [Bec93]. Alternatively, this equation can be obtained by just integrating the FPE.
Recall that the fixed point of the FPE, $\rho^*(x) = P_F \rho^*(x)$, yielded the *invariant density* of the map F, see Definition 31. This leads to the following analogous defnition for the corresponding probability measure:

Definition 42 *[Bec93]*
$\mu_{n+1}(A) = \mu_n(A) =: \mu^*(A)$ *is called an* invariant measure, *i.e., the measure does not change with time.*

Hence, for our μ-version of the FPE it must hold that $\mu^*(A) = \mu^*(F^{-1}(A))$ [All97]. A map F that fulfills this property is also sometimes called *measure preserving* in the mathematics literature [Las94].
We may now highlight another important property of the probability measure $\mu(A)$ obtained from a density in form of the following definition:

Definition 43 *absolute continuity [Kat95, Las94, dM93]*
Consider a measure space $(I, \mathscr{A}, \mu_L)$*, where* $I \subseteq \mathbb{R}$ *and* μ_L *is the Lebesgue measure. Let* ρ *be a probability density on* I*. If there exists a* continuous

$$\mu(A) := \int_A dx \rho(x) \quad \forall A \in \mathscr{A} \quad , \tag{7.7}$$

μ *is called* absolutely continuous *with respect to* μ_L*.*

Remark 18

1. Note that for the Lebesgue measure in the integral one may also write $dx = d\mu_L = \mu_L(dx)$. This more general notation is often used in the mathematics literature.

2. Loosely speaking one may say that a measure is absolutely continuous if it is given by integrating a density [Kat95, Has03]. However, this only implies continuity for the measure μ if we specify the space of probability density functions on which we are working. The function space should then exclude densities ρ that are singular such as, e.g., the $\delta-$function that we have encountered in Section 3.3, see Eq. (3.15). This is also illustrated in the following examples.

3. Rigorously mathematically speaking, the function space on which we should be working here is the space of Lebesgue-integrable functions [Las94]: Let us replace the Lebesgue measure $\mu_L(dx)$ by some other measure $\nu(dx)$. Then $L^p(I, \mathscr{A}, \nu)$, $1 \leq p < \infty$ is the family of all possible real-valued measurable functions ρ satisfying $\int \nu(dx)\ |\rho(x)|^p < \infty$, where this integral should be understood as a Lebesgue integral.[2] Note that if $\rho \in L^p$ is not Riemann-integrable anymore, it may still be Lebesgue-integrable, that is, the above integral may still exist.

4. Some authors define a measure μ to be absolutely continuous with respect to another measure ν if for all measurable sets $A \in I$ $\mu(A) = 0$ whenever $\nu(A) = 0$ [dM93, Las94]. If $\rho \in L^1$ and ν is σ-finite,[3] then one can prove that this definition is equivalent to Definition 43. This is the contents of the *Radon-Nikodym Theorem*, and $\rho = d\mu/d\nu$ is called the *Radon-Nikodym derivative* of μ with respect to ν [dM93, Las94].

5. If $\rho \in C^0$ and bounded it follows from the fundamental theorem of calculus that μ is C^1-smooth and $\mathrm{d}\mu/\mathrm{d}x = \rho(x)$.

6. If $\rho(x)$ is continuous except at finitely many points and bounded, then there still exists a measure μ according to Eq. (7.7), however, it is not everywhere C^1 anymore.

Let us illustrate the last two points of this remark by two examples:

Example 36

1. Consider the Bernoulli shift $B(x) = 2x \mod 1$ on $[0, 1)$ with $\rho^*(x) = 1$. It is easy to see that there exists a differentiable absolutely continuous invariant measure $\mu^*(x) = \int_0^x d\tilde{x}\ \rho^*(\tilde{x}) = x$.

2. Consider the step function $\rho^*(x)$ sketched in the left part of Fig. 7.2. Integrating over this function according to $\mu^*(x) = \int_0^x d\tilde{x}\rho^*(\tilde{x})$ yields the graph depicted in the right part of this figure, which ist still C^0 but not differentiable at the points of discontinuity of $\rho^*(x)$.

Definition 44 *[Lev94, Las94]*
If the measure μ defined by Eq. (7.7) is not *absolutely continuous, it is called* singular.[4]

[2] I found the entry on Lebesgue integration in the online Wikipedia quite helpful to get an intuition of what Lebesgue integration is about compared with ordinary Riemann integration.

[3] A measure is σ-finite if it can be written as the countable union of sets with finite measure.

[4] One may also say that the measure μ defined by Eq. (7.7) decomposes into an absolutely continuous one μ_a and a singular one μ_s, which is the contents of *Lebesgue's decomposition theorem* [Las94].

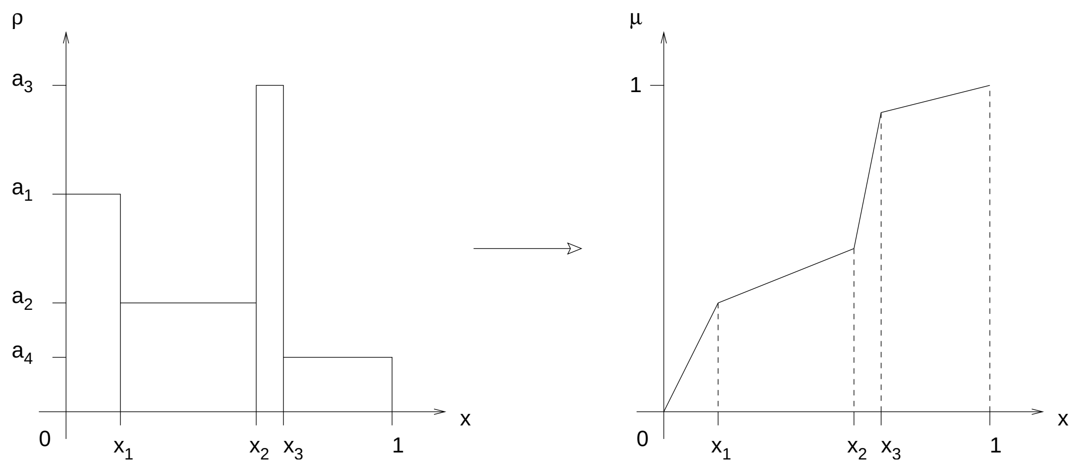


Figure 7.2: A step function density and the corresponding absolutely continuous probability measure obtained by integrating the density.

Example 37 For $B(x)$ let $\tilde{\rho}^*(x) := \frac{1}{2}[\delta(x-\frac{1}{3})+\delta(x-\frac{2}{3})]$ be the invariant density defined on the period 2 orbit $\{\frac{1}{3},\frac{2}{3}\}$ *only*, see the left part of Fig. 7.3. Then $\tilde{\mu}^*(x) = \int_0^x d\tilde{x}\, \rho^*(\tilde{x})$, see the right hand side of this figure, is singular.

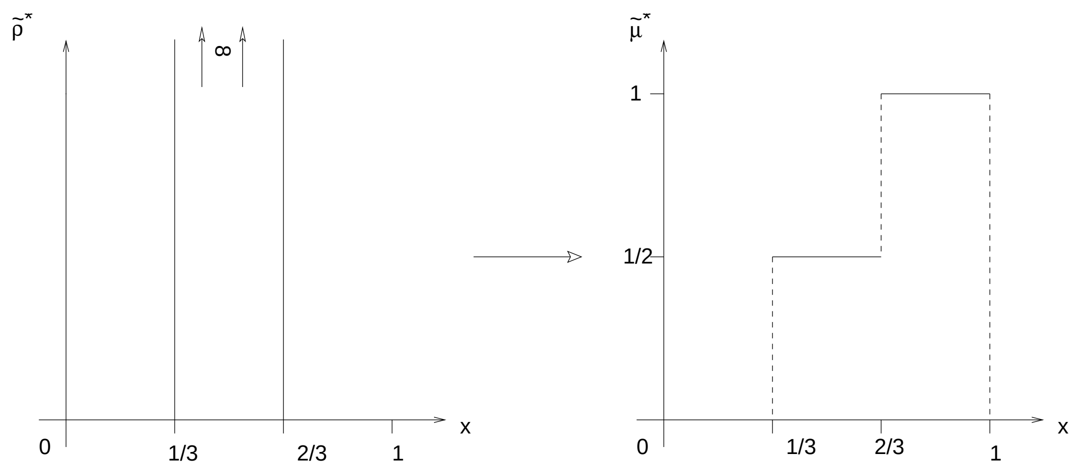


Figure 7.3: A delta function density yields a singular measure via integration.

Definition 45 *[All97]*
A measure μ is called atomic[5] *if all of the measure is contained in a finite or countably infinite set of points.*

Example 38 The measure $\tilde{\mu}^*$ of the above example is atomic,[6] which in fact is better seen if we look at the measure $\tilde{\mu}^*(A)$ on sets $A \subseteq [0,1]$ rather than $\tilde{\mu}^*(x)$.

[5]In measure theory, an atom is a measurable set which has positive measure and contains no ”smaller” set of positive measure.

[6]This measure actually represents a linear combination of the Dirac measure as defined on a coursework sheet.

Remark 19 An atomic measure is singular, however, a singular measure is not necessarily atomic.

For higher-dimensional dynamical systems we commonly have a *combination* of absolutely continuous and singular components of the invariant measure. This leads to another definition, which we first motivate by two examples:

Example 39 [Dor99]

1. The *baker map*: (see also [Las94])

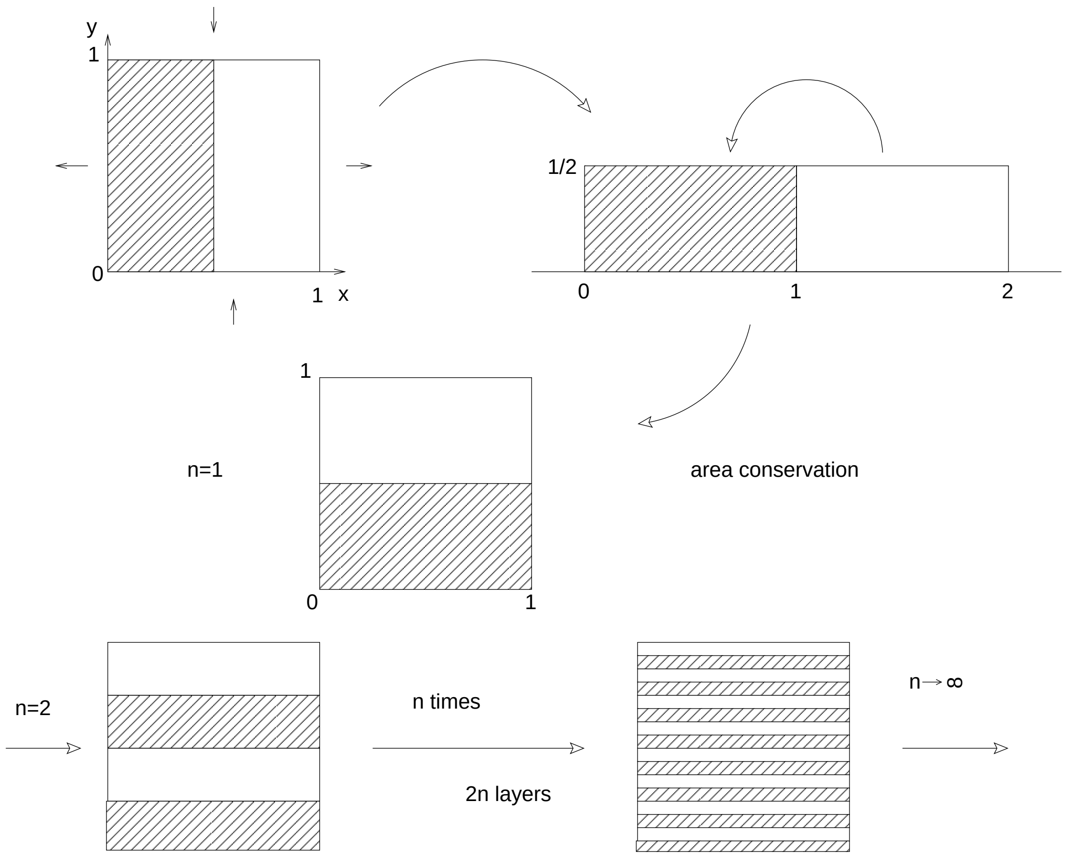


Figure 7.4: The two-dimensional area-preserving baker map and its iterative action onto a set of points initially concentrated on the left half of the unit square.

In Fig. 7.4 we geometrically define the famous *baker map*, a paradigmatic example of a two-dimensional area-preserving chaotic dynamical system. This map stretches the unit square by a factor of two along the x-axis and squeezes it by a factor of $1/2$ in the y-direction. The resulting rectangle is cut in the middle, and both halfes are put on top of each other recovering the unit square. If we colour the left half of the unit square (or fill this half uniformly with points while leaving the right half empty, respectively) and iterate this map several times we obtain the layered structure shown in the figure.

Note that the corresponding density $\rho_n(x, y)$ is uniform in the x direction, whereas it is a step function along y. However, in the limit of $n \to \infty$ the strips become thinner and thinner such that all points eventually distribute uniformly over the whole square, so there are no "steps" anymore along y. In other words, the invariant density of the baker map is simply constant everywhere. But if we modify the map a little bit the situation changes dramatically leading us to our second example.

2. The *dissipative baker map*:

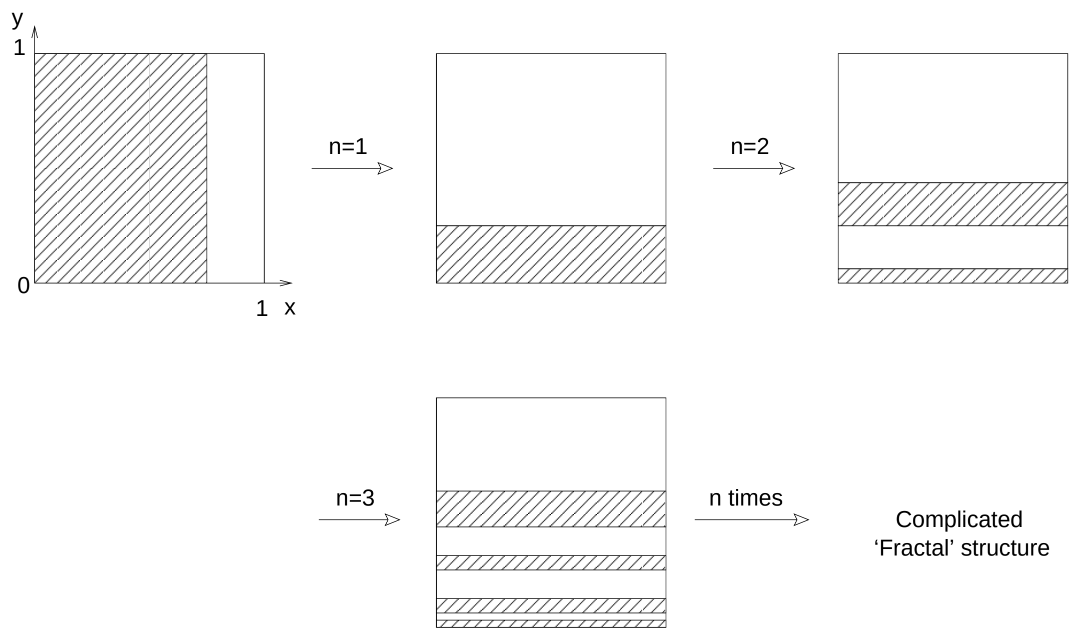


Figure 7.5: The dissipative baker map eventually generates an invariant measure exhibiting a fractal structure in the y direction.

In Fig. 7.5 we introduce a little variant of the previous baker transformation, where we still stretch and squeeze the whole unit square parallel to the x- and to the y-direction, respectively. However, we do this now such that the big left (coloured) rectangle is compressed into the little rectangle at the bottom of the second unit square, whereas the little right (empty) rectangle is expanded into the big empty one on top. If we repeat this action, again a layered structure evolves which is similar to the one of the previous figure. One observes that, as before, the density $\rho_n(x, y)$ is uniform in the x-direction and that in the limit of $n \to \infty$ the strips piling up along y become thinner and thinner. However, in contrast to the previous example it can be shown that in the limit of $n \to \infty$ the invariant density is not uniform anymore on the whole unit square. Strictly speaking, in this case an invariant density does not exist anymore, because a fractal structure was generated in the y direction which makes the density $\rho_n(x, y)$ non-integrable in the limit of $n \to \infty$. However, there still exists an invariant measure $\mu^*(x, y)$. This measure turns out to be absolutely continuous along x with $\mu^*(x) \in C^1$ but it is singular along y with $\mu^*(y) \in C^0$. See [Dor99] for further details.

Definition 46 *SRB measure [You02, Eck85, Bal00]*
An invariant probability measure μ^* is called an SRB measure[7] *if it is* absolutely continuous

[7] The acronym SRB holds for the initials of the mathematicians Ya.Sinai, D.Ruelle and R.Bowen who first introduced this measure.

along unstable manifolds.

Recall that the unstable manifold $W^u(p)$ of a source p of a map can be defined as the set of initial conditions whose orbits are repelled from p, see Definition 19. The same applies to p being an unstable periodic point.[8]

Remark 20

1. Some authors define an SRB measure by asking for slightly stronger C^1-smoothness along unstable manifolds [Eck85].

2. We emphasize that this definition is designed for higher-dimensional dynamical systems, where stable and unstable manifolds typically form complicated structures in phase space. Often the topology of these systems is such that the measure is smooth along unstable manifolds, whereas it is fractal along stable ones. In this case the total measure is not absolutely continuous, however, it is still SRB. So while absolute continuity of the total measure trivially implies that it is SRB, the reverse is not true.

Example 40

1. The Bernoulli shift $B(x) = 2x \mod 1$ is unstable $\forall x \in [0, 1)$, since with $B'(x) = 2$ it is everywhere expanding (except at $x = 1/2$). Hence the whole system lives on an unstable manifold, and $\mu^*(x) = \int_0^x d\tilde{x}\rho^*(\tilde{x}) = \int_0^x \mathrm{d}\tilde{x}1 = x$ is an SRB measure.

2. The baker map of Fig. 7.4 has an SRB measure that is smooth along the x axis, which iidentifies the direction of all unstable manifolds. It is also smooth in the y direction, although this defines the stable manifolds.

3. The dissipative baker map of Fig. 7.5 has an SRB measure that is smooth along the unstable manifolds in the x direction. However, it is singular along the stable manifolds in the y direction.

Remark 21

1. SRB measures are important, because the property of being smooth along at least one direction in phase space enables one to calculate ensemble averages of (physical) observables, see also the following section.

2. An important question is whether an SRB measure exists for a given dynamical system. These measures are supposed to be quite common, however, their existence is usually hard to prove. For so-called Anosov systems an important theorem (sometimes called the SRB theorem [Dor99]) ensures the existence of SRB measures [You02]. The above two baker maps are examples of systems which have 'Anosov-like' properties [Dor99].

[8] For further details about the definition of stable and unstable manifolds see [Bec93, Ott93, Eck85].

7.2 Basics of ergodic theory

In this section we introduce measure-theoretic properties which form the basis of what is called *ergodic theory*, an important branch of dynamical systems theory. Our presentation particularly follows [Dor99, Arn68], for further mathematical details see [Kat95].
Let as usual $F : I \to I\ ,\ \ x_{n+1} = F(x_n)\ ,\ n \in \mathbb{N}_0$ be a one-dimensional map acting on the set $I \subseteq \mathbb{R}$ and μ^* be a corresponding invariant probability measure. Let us consider a function $g : I \to \mathbb{R}\ ,\ g \in C^0$, which we may call an "observable".

Definition 47 *time and ensemble average*

$$\overline{g(x)} := \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} g(x_k) \quad , \tag{7.8}$$

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

$$\langle g \rangle := \int_I d\mu^* g(x) \tag{7.9}$$

where, if it exists, $d\mu^* = \rho^*(x)\,dx$, is called the ensemble (or space) average *of g with respect to F. Note that $\overline{g(x)}$ may depend on x, whereas $\langle g \rangle$ does not.*

Example 41 In Definition 25 we have introduced the local Ljapunov exponent of a map F. If we choose $g(x) = \ln |F'(x)|$ as the observable in Eq. (7.8), we find that

$$\lambda_t(x) := \overline{\ln |F'(x)|} = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \ln |F'(x_k)| \quad , \tag{7.10}$$

so calculating the local Ljapunov exponent actually means taking a time average along a trajectory. If we choose the same observable for the ensemble average Eq. (7.9) we obtain

$$\lambda_e := \langle \ln |F'(x)| \rangle := \int_I dx \rho^*(x) \ln |F'(x)| \quad . \tag{7.11}$$

Let us check these two definitions for the Bernoulli shift $B(x) = 2x \mod 1$, where $B'(x) = 2$ for almost every x: We have that

$$\lambda_t(x) = \frac{1}{n} \sum_{k=0}^{n-1} \ln 2 = \ln 2 \tag{7.12}$$

for almost every $x \in [0, 1)$ and that

$$\lambda_e = \int_0^1 dx \rho^*(x) \ln 2 = \int_0^1 dx 1 \cdot \ln 2 = \ln 2 \quad . \tag{7.13}$$

In other words, time and ensemble average are the same for almost every x,

$$\lambda_t(x) = \lambda_e = \ln 2 \quad . \tag{7.14}$$

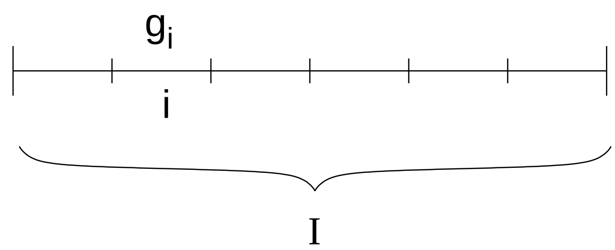


Figure 7.6: Uniform partition of the interval $I \subseteq \mathbb{R}$ and a local average g_i defined on the ith partition part of an observable $g(x)$.

In the following we will show that this result is not a coincidence. We do so in two steps:

Step 1: Let us define a uniform partition on the interval I, see Fig. 7.6, with a total number of N partition parts, where the ith partition part is defined by the subinterval $I_i := [x_i, x_{i+1}]$. Take a long trajectory $\{x_0, x_1, \ldots, x_{n-1}\}$ generated by the map F and let n_i with $\sum_{i=1}^N n_i = n$ be the number of occurrences when the trajectory is in partition part i. Let g_i be the local average of $g(x)$ with respect to the ith partition part,

$$g_i := \frac{1}{n_i} \sum_k g(x_k) \, \forall x_k \in I_i \quad . \tag{7.15}$$

Then

$$\overline{g(x)} = \frac{1}{n} \sum_{k=0}^{n-1} g(x_k) = \frac{1}{n} \sum_{i=1}^{N} n_i g_i = \sum_{i=1}^{N} \frac{n_i}{n} g_i \quad . \tag{7.16}$$

There is now a famous hypothesis relating the number of occurrences n_i in the bin i to the invariant probability measure μ_i^* of the region i. Before we can state it we need the following definition:

Definition 48 *almost everywhere [Kat95, Ott93]*
*A property is said to hold μ-*almost everywhere *if the set of elements for which the property does not hold is of measure zero with respect to a given measure μ. In this case one also says that the property holds for a* typical element.

Hypothesis 1 *Boltzmann's ergodic hypothesis (1871) [Dor99, Tod92]*
The trajectory of a typical point in I spends equal time in regions of equal measure.

If this hypothesis applies we have that

$$\frac{n_i}{n} = \frac{\mu_i^*}{\mu^*(I)} = \mu_i^* \quad , \tag{7.17}$$

which together with Eq. (7.16) implies that

$$\overline{g(x)} = \sum_{i=1}^{N} \mu_i^* g_i \to \int_I d\mu^* g(x) = \langle g \rangle \quad (n, N \to \infty) \quad . \tag{7.18}$$

Step 1 summarizes Boltzmann's intuitive approach to relate time and ensemble averages of observables. His hypothesis forms the rigorous foundation of the whole theory of statistical mechanics [Tod92]. **Step 2** now makes his idea mathematically precise by starting with an important theorem:

Theorem 5 *Birkhoff's ergodic theorem (1931) [Dor99, Kat95, Arn68, Tod92]*
If there exists an invariant probability measure $\mu^*(I)$ with $\int_I d\mu^* |g(x)| < \infty$[9] *then for μ^*-almost every $x \in I$ there exists $\overline{g(x)}$.*

For a proof see [Kat95]. Note that this theorem only ensures the *existence* of the time average. It does not say anything about *uniqueness*, that is, $\overline{g(x)}$ might be different for different initial conditions $x \in I$. However, this is clarified by the following definition:

Definition 49 *ergodicity [Arn68, Dor99]*
A dynamical system is called ergodic *if*

$$\overline{g(x)} = \langle g \rangle \tag{7.19}$$

μ^*-almost everywhere.

This implies that $\overline{g(x)}$ does not depend on x anymore. That the time average is constant is sometimes also taken as a definition of ergodicity in the literature [Dor99, Bec93]. In fact, there is even a third definition, which is based on the following definition:

Definition 50 *decomposability [Arn68]*
Assume I is the disjoint union of two sets I_1, I_2 of positive measure, each of which is invariant under the map F: $F(I_1) = I_1$, $F(I_2) = I_2$. Such a system is called decomposable.

With this definition we have the following proposition:

Proposition 6 *A map F is ergodic if and only if F on I is* indecomposable.

The proof of this proposition is not hard [Dor99, Arn68]. In turn, sometimes this statement is used to *define* ergodicity [Kat95]. The equivalence between time and ensemble average stated by Definition 49 then follows as a proposition. Let us discuss these notions by two examples:

Example 42

1. The "double tent map" shown in Fig. 7.7 is certainly decomposable. According to Proposition 6 it is therefore not ergodic.

2. The Bernoulli shift $B(x)$ is indecomposable and hence ergodic. However, a proof of this property is a bit cumbersome and goes along the lines as the one of ergodicity for the baker map [Dor99].

3. The rotation on the unit circle $R_\alpha : S^1 \to S^1$, $R_\alpha(\theta) := \theta + 2\pi\alpha$, $\alpha \in \mathbb{R}$ as defined on one of the coursework sheets is non-ergodic if α is rational and ergodic if α is irrational. See [Arn68, Dor99, Tod92] for proofs, the one for ergodicity is a bit more elaborate.

[9] Strictly speaking this means that we require g to be Lebesgue-integrabel, $g \in L^1$, see Remark 18.

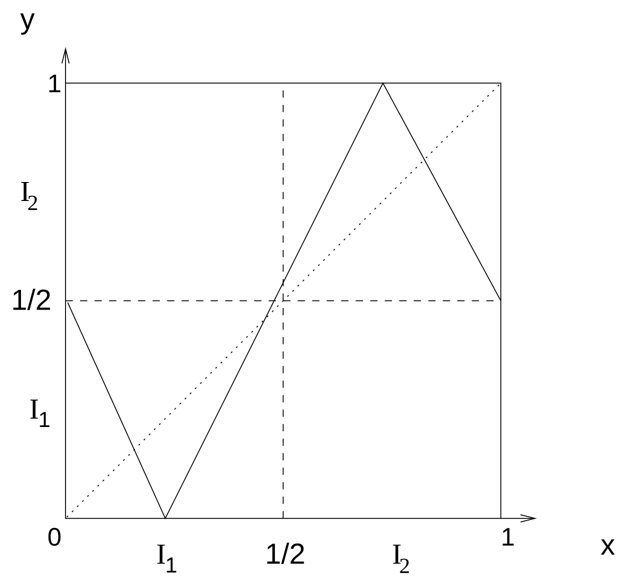


Figure 7.7: An example of a map that is decomposable and hence not ergodic.

Remark 22 You may have noticed that the properties of a map being topologically transitive and being ergodic are of a very similar nature. However, note that ergodicity requires us to use a measure, whereas topological transitivity works without a measure. In fact, for a continuous map it can be shown under fairly general conditions that ergodicity implies topological transitivity, whereas the reverse does not hold [Kat95, Rob95].[10]

Let us now get back to the Ljapunov exponent studied in Example 41. For time average $\lambda_t(x)$ and ensemble average λ_e we have found that

$$\lambda_t(x) = \lambda_e = const. \tag{7.20}$$

for the Bernoulli shift. From Definition 49 we have learned that this result must hold whenever a map F is ergodic. However, this means that in case of ergodicity the Ljapunov exponent λ 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 43 Let us reconsider our specific map F which has already been introduced previously, see Example 28. For convenience we display it again in Fig. 7.8.
In Example 31 the invariant density of this map was calculated to

$$\rho^*(x) = \begin{cases} \frac{4}{3}, & 0 \leq x < \frac{1}{2} \\ \frac{2}{3}, & \frac{1}{2} \leq x < 1 \end{cases} , \tag{7.21}$$

see Eq. (6.51). This motivates us to calculate the Ljapunov exponent of this map by using

[10]For example, the reverse implication does not hold for the Pomeau-Manneville map.

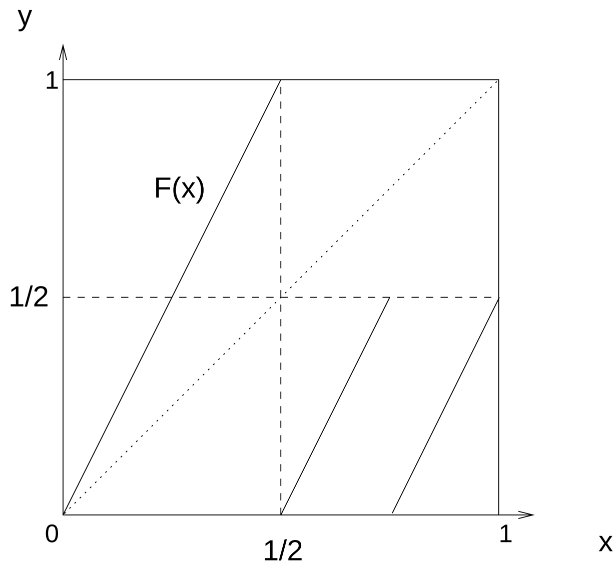


Figure 7.8: A simple map that we have seen before, here used in order to demonstrate the calculation of Ljapunov exponents via ensemble averages.

the ensemble average Eq. (7.11). We obtain

$$\lambda_e = \int_0^1 \mathrm{d}x \rho^*(x) \ln |F'(x)| = \ln 2 \left(\int_0^{\frac{1}{2}} \mathrm{d}x \frac{4}{3} + \int_{\frac{1}{2}}^1 \mathrm{d}x \frac{2}{3} \right) = \ln 2 \left(\frac{4}{3} \cdot \frac{1}{2} + \frac{2}{3} \cdot \frac{1}{2} \right) = \ln 2 \quad . \tag{7.22}$$

Actually, this result trivally follows from the fact that the slope of this map is uniform and that the probability density normalises to one. However, the calculation demonstrates how the definition of λ by ensemble averages could be used in case the slope and the invariant density are more complicated functions. By assuming that the map F 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 F.

We conclude this section by using the concept of ergodicity for a further characterisation of invariant measures:

Definition 51 *physical measure [You02]*
Let $F : I \to I$, $I \subseteq \mathbb{R}$, be a map and μ^* be an invariant probability measure of this map. μ^* is called a physical (or natural) measure *if for an observable $g : I \to \mathbb{R}$, $g \in C^0$ the map is ergodic, that is, $\overline{g(x)} = \langle g \rangle$ for μ^*-almost every $x \in I$.*

Remark 23

1. Note that by restricting ourselves to typical points and considering measures that are defined with respect to integrating probability densities we sort out periodic orbits, which are all of Lebesgue measure zero.

2. It is important to observe that a measure being SRB neither implies that it is physical, nor is the reverse true. For both directions one can construct counterexamples [You02, Bal00, Eck85]. For example, our double tent map considered above has an SRB

measure but not an ergodic one. On the other hand, the map $C(x) = x/2$ contracting onto a fixed point has trivially an ergodic measure on the attracting fixed point which, however, is not SRB.

3. Some authors define an SRB measure via ergodicity, so the definitions of SRB and physical measure are not unique in the literature.

4. An alternative definition of a physical measure first suggested by Kolmogorov yields the physical measure as the measure that emerges in the zero-noise limit of a dynamical system under random perturbations [Eck85].

5. The SRB theorem mentioned before proves that for Anosov systems the SRB measure is equal to the physical one [You02], which identifies an important class of dynamical systems where both measures are identical.

6. You may wonder why physical measures are important at all. This is for the following reason: Pick your favourite (physical) system such as, for example, the driven pendulum that we have encountered at the very beginning of these lectures. You may wish to characterize such a system in a statistical sense by calculating your favourite observable in terms of time or ensemble averages. You would hope that your averages exist (i.e., that your computations converge to specific values in the limit of long times or of large ensembles of points), and also that your results do not depend on the initial conditions you choose. In mathematical terms, you would thus expect that there exists a “nice” measure which guarantees that your averages exist and are the same. If your system exhibits a physical measure, it is the one you will typically “see” by measuring your observable. And being SRB guarantees that you can calculate your observables by projection along the unstable manifolds.

If you are further interested, I would highly recommend that you make yourself familiar with what is called *mixing* in ergodic theory, see, e.g., [Arn68, Dor99].

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).

[Bal00] V. Baladi, *Positive Transfer Operators and Decay of Correlations*, volume 16 of Advanced Series in Nonlinear Dynamics. (World Scientific, Singapore, 2000).

[Ban92] J. Banks, J. Brooks, G. Cairns, G. Davis and P. Stacey, *On Devaney's definition of chaos.* Amer. Math. Monthly **99**, 332–334 (1992).

[Bec93] C. Beck and F. Schlögl, *Thermodynamics of Chaotic Systems*, volume 4 of Cambridge nonlinear science series. (Cambridge University Press, Cambridge, 1993).

[Dev89] R.L. Devaney, *An introduction to chaotic dynamical systems.* (Addison-Wesley, Reading, second edition, 1989).

[dM93] W. de Melo and S. van Strien, *One-dimensional dynamics*, volume 25 of Ergebnisse der Mathematik und ihrer Grenzgebiete, 3.Folge. (Springer, Berlin, 1993).

[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).

[Elm98] F.-J. Elmer. *The Pendulum Lab.* University of Basel, 1998. Online simulation at http://monet.unibas.ch/∼elmer/pendulum.

[Fot05] A. Fotiou, *Deterministic chaos.* Master's Thesis, School of Mathematical Sciences, Queen Mary, University of London, 2005. see http://www.mpipks-dresden.mpg.de/∼rklages/people/msc_fotiou.pdf.

[Gas98] P. Gaspard, *Chaos, scattering, and statistical mechanics.* (Cambridge University Press, Cambridge, 1998).

[Gla93] E. Glasner and B. Weiss, *Sensitive dependence on initial conditions.* Nonlinearity **6**, 1067–1075 (1993).

[Gle96] J. Gleick, *Chaos: Making a New Science.* (Vintage, 1996).

[Guc79] J. Guckenheimer, *Sensitive dependence to initial conditions for one dimensional maps.* Comm. Math. Phys. **70**, 133–160 (1979).

[Guc90] J. Guckenheimer and P. Holmes, *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*, volume 42 of Applied mathematical sciences. (Springer, Berlin, 3rd edition, 1990).

[Has03] B. Hasselblatt and A. Katok, *A first course in dynamics.* (Cambridge University Press, 2003).

[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).

[Kla07] R. Klages, *Microscopic chaos, fractals and transport in nonequilibrium statistical mechanics*, volume 24 of Advanced Series in Nonlinear Dynamics. (World Scientific, Singapore, 2007).

[Las94] A. Lasota and M.C. Mackay, *Chaos, Fractals, and Noise.* (Springer-Verlag, 1994).

[Lev94] R.W. Leven, B.-P. Koch and B. Pompe, *Chaos in dissipativen Systemen.* (Akademie Verlag, Berlin, second edition, 1994).

[Miy03] T. Miyasaka and T. Harayama. *Corrugated floor simulation*, 2003. Simulation courtesy of Dr. Takahisa Harayama and Dr. Tomohiro Miyasaka, ATR labs, Kyoto, Japan.

[Ott93] E. Ott, *Chaos in Dynamical Systems.* (Cambridge University Press, Cambridge, 1993).

[Rob95] C. Robinson, *Dynamical Systems.* (CRC Press, London, 1995).

[Smi07] L. Smith, *Chaos - A very short introduction.* (Oxford University Press, Oxford, 2007).

[Str94] S. H. Strogatz, *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering.* Studies in nonlinearity. (Addison Wesley, Reading, MA, 1994).

[Tél06] T. Tél and M. Gruiz, *Chaotic Dynamics.* (Cambridge University Press, Cambridge, 2006).

[Tod92] M. Toda, R. Kubo and N. Saitô, *Statistical Physics*, volume 1 of Solid State Sciences. (Springer, Berlin, 2nd edition, 1992).

[Wig92] S. Wiggins, *Chaotic transport in dynamical systems*, volume 2 of Interdisciplinary Applied Mathematics. (Springer, New York, 1992).

[You02] L.-S. Young, *What are SRB measures, and which dynamical systems have them?* J. Stat. Phys. **108**, 733–754 (2002).