A Model for Prejudiced Behaviour and a Noisy Dynamical System

Andreas U. Schmidt

DR. ANDREAS U. SCHMIDT, GMD GERMAN NATIONAL RESEARCH CENTER FOR INFORMATION TECHNOLOGY, INSTITUTE FOR SECURE TELECOOPERATION (SIT.MINT), DOLIVOSTRASSE 15, 64293 DARMSTADT, GERMANY
E-mail address: aschmidt@math.uni-frankfurt.de

Contents

ABSTRACT. This report combines two different threads of research: In the first chapter, we make a proposal for a mathematical model of prejudiced behaviour. The heuristic background of this is the study of agents in open marketplaces, *e.g.*, on the internet. We derive our model from a simple outset within the framework of the general theory of learning systems and a few axiomatic assumptions. The model considered turns out to be an example of a noisy dynamical system. Systems of this type have recently attracted much attention in the physical and mathematical communities. We perform a theoretical study of our special system in chapter 2 and carry out some numerical experiments. We stress the analogy of the phenomenology of the system with phase transitions in physical systems. The results that we obtain in chapter 2 give us some insight into the qualitative behaviour that learning systems governed by our prejudiced learning rule will exhibit.

Key words and phrases. Learning system, prejudice, behaviour, deterministic dynamical system, random dynamical system, phase transition, invariant density, Perron–Frobenius operator, Lyapunov exponent, ergodicity, stochastic bifurcation, Poincarè dimension, critical exponent, intermittency, numerical simulation.

Acknowledgements. This research was partially supported by the EU project eBroker within the fifth framework. Thanks to the parallel computing group POPCORN of the GMD-institute SCAI.

ZUSAMMENFASSUNG. Dieser Bericht führt zwei verschiedene Forschungsthemen zusammen: Im ersten Kapitel stellen wir ein mathematisches Modell für vorurteilsbehaftetes Verhalten vor. Den heuristische Hintergrund bildet hierbei die Studie des Verhaltens von Agenten in offenen Marktplätzen wie z. B. internetbasierten. Das Modell leiten wir im Rahmen einfacher Grundlagen der Theorie lernender Systeme aus wenigen axiomatischen Annahmen her. Es erweist sich als ein Beispiel für ein dynamisches System mit Rauschen. Systeme dieses Typs haben in jüngster Zeit in Mathematik und Physik einige Aufmerksamkeit erregt. Wir führen eine theoretische Studie unseres speziellen Systems in Kapitel 2 durch und sammeln empirische Daten mittels numerischer Experimente. Dabei betonen wir besonders die Analogie zum Phänomen der Phasenübergänge in physikalischen Systemen. Die Resultate, die wir erhalten ermöglichen es uns wiederum in Kapitel 1, einigen Einblick in das qualitative Verhalten das lernende Systeme, die unserer Lernregel gehorchen, zeigen werden, zu gewinnen.

Stichworte. Lernendes System, Vorurteil, Verhalten, deterministisches dynamisches System, zufälliges dynamisches System, Phasenübergang, invariante Dichte, Perron–Frobenius Operator, Lyapunov Exponent, Ergodizität, stochastische Bifurkation, Poincarè–Dimension, kritischer Exponent, Intermittenz, numerische Simulation.

Danksagung. Diese Forschung wurde teilweise innerhalb des EU-Projekts eBroker des fünften Rahmenprogramms durchgeführt. Besonderer Dank an die parallell Computer Gruppe POPCORN des GMD-Insituts SCAI.

CHAPTER 1

Modeling Prejudiced Behaviour

1.1. Introduction

We borrow the framework for the behavioural model we will present from basic game theory and reinforcement learning. We defer the reader to [**9, 12**] for details about these vast subjects. But let us nevertheless briefly describe the background needed for our purposes:

The classical game-theoretic subject of *decision making under uncertainty* considers single- or multiplayer games of intelligent agents against nature. Each agent has an *utility function* depending on the action he will take and the (unknown) state of nature. This function describes the payoff of the execution of the corresponding action given that nature is in the corresponding state. The utility function is taken to be the input for the agents *decision rule* which uniquely determines the action to be taken. In the more realistic case when a probability distribution over the states of nature is known to the agent, one is in the realm of *statistical decision theory*. There, this knowledge is fed into a selection scheme which determines one from a set of decision rules accordingly. For example, the well known *Bayesian decision rule* is the combination of selection scheme and decision rule which assigns to each action the average sum of utilities weighted with the known probability distribution, and then chooses the action maximizing this value. The knowledge can be *a priori* or learned by experimentation using some *statistical learning rule*, which can be as simple as taking means over a finite set of experimental outcomes (*Bayesian learning*). When the result of taking a specific action is fed back as the outcome of an experiment into the learning rule and the whole cycle is repeated many times, the agent becomes a *learning automaton*. This is the fundamental object we want to consider.

The Bayesian rule is obviously not the only possible, but there is a multitude of different learning rules and selection schemes. In realistic cases it can be sensible to choose selection schemes which differ significantly from the apparently most rational Bayesian rule. Here, we propose a model which can justifiably be said to represent agents who are *prejudiced* by construction and in behaviour. We derive the heuristics for our construction from a key example, which we want to present before further going into detail.

EXAMPLE 1. *At each time step, the agent executes a business transaction with a certain payoff value, which we normalize to* 1*. Within each transaction, he has to select a specific security protocol from a finite set. Choosing the* k*-th protocol, there is a — generically small — probability* p_k *for the occurrence of a damage* d *which diminishes the payoff to* $1 - d$*. Assume that the damage is symmetrically distributed with small variance around a mean value* $d_k \in [0, 1]$*. The expected payoff is the (true empirical)* ***weight*** $w_k \equiv 1 - r_k \equiv 1 - d_k \cdot p_k$ *of protocol* k *and* r_k *is its* ***risk****. Approximations for these quantities are learned by the agents from the frequency and amount of previous damages. These are fed into a selection scheme which in turn determines the decision rule for protocol selection.*

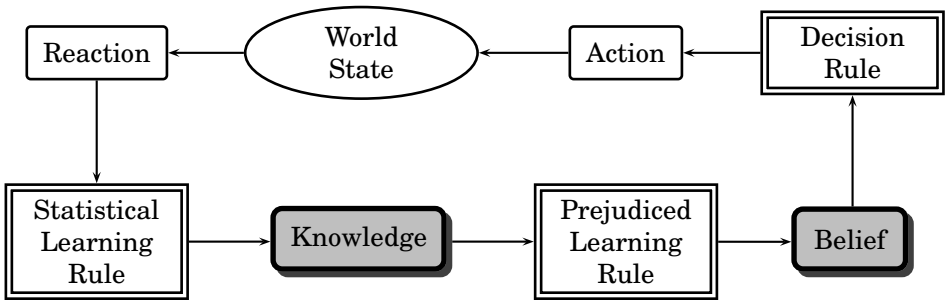

FIGURE 1.1. Information flow in the model for prejudiced behaviour. Shaded rectangles symbolize the agent's internal state, while unshaded ones stand for actual events. Doubly framed rectangles mark the algorithms used by the agent.

The general feature of this model rendering the application of Bayes' rule less attractive is that the damage rate $1/p_k$ and therefore the learning speed of any statistical inference rule for the set $\{r_k\}$ is very low compared to the frequency with that a protocol has to be selected (which is 1 in our simplification). Thus, initial probabilistic fluctuations could result in the — costly — selection of a non-optimal protocol for many steps. On the other hand, the expected damage d_k can be close to 1, resulting in relatively high risk. Thus the agent has high interest in using a reliable risk estimate from the start, hold it stable against fluctuations but nevertheless learn as quickly as possible.

Given these conflicting goals, it can be sensible to use a **prejudiced learning system** for a weight w (we drop the index k for now), by which we mean the following: The agent is given a start value $w(0)$. At time t he infers his (empirical) **knowledge** $\eta(t)$ from the $t+1$ observations made up to this time using some statistical learning rule not further specified. He then determines his **belief** about w at time $t+1$ by a **prejudiced learning rule**

$$w(t+1) = L((w(s))_{s\le t}, \eta(t)), \quad t = 0, 1, \dots$$

The essence of this constructive 'definition' of prejudice is the distinction between *knowledge* and *belief* which seems heuristically sound. Figure 1.1 shows a schematic view of the simple internal structure of the prejudiced agent and it's interaction with the exterior world. Our model almost totally ignores the upper row of the diagram, for example, the 'reaction' we use will consist simply in informing the agent of his error, which for us is nothing but the numerical difference between his belief and the actual state of the world.

EXAMPLE 2. *In the Internet world, the prejudices about the security of business transactions range from "never do it, it can cost your whole wealth" to "transmitting a credit card number online is secure" with more ore less good reasons for both extremes. Furthermore, such opinions are seldom acquired by an individual users experience. The public and individual opinion about these matters is largely determined by gossip, mass media, and exponents of particular interests. Undoubtedly yet, everyone (despite possibly some of the mentioned interested parties, but this might be the authors own prejudice) would value reliable information about the subject very high. Until then, a careful balance between empirical learning and sticking to ones prejudices is a reasonable path to follow.*

1.2. A Prejudiced Learning Rule and its Consequences

The simplest example of a prejudiced learning rule one could think of is taking a weighted average between the initial belief $w(0)$ and $\eta(t)$, *i.e.*, $w(t+1) = \alpha(t)w(0) + (1-\alpha(t))\eta(t)$. Our aim is to construct a rule which somewhat better fulfills the specific requirements of Example 1 but also exhibits some general features of prejudiced learning.

1.2.1. Three Special Cases. We first make realistic assumptions about two extreme cases: a) $w(0) = 0$ expresses the strongest 'prejudice' namely inability to perform the respective protocol (assuming no protocol is dropped by the decision rule). This value should obviously remain constant over time, *i.e.*, $L((0,0,\ldots),\eta(t)) = 0$. b) If, on the other hand, the weight is maximized at some time then the agent has highest interest in correcting it if this belief is wrong, *i.e.*, $|w(t+1)-w(t)|$ should be maximal when $w(t) = 1$. c) As a third case we consider $w(t) = \eta(t)$, *i.e.*, the agents belief accidentally equals his knowledge. Then he has good reason to leave his belief untouched and choose $w(t+1) = w(t)$.

1.2.2. The Logistic Learning Map. Specializing further, we consider $\eta(t) = w = 1 - r$ to be identical to the true weight and constant over time. The most simple prejudiced learning rule satisfying requirements b) and c) above then depends only on the previous time step and is linear in $r(t)$ and the **error** $\Delta(t) \stackrel{\text{def}}{=} r - r(t)$, namely

$$\begin{aligned} \Delta(t+1) &= \alpha \cdot r(t) \cdot \Delta(t) \quad \text{or} \\ r(t+1) &= r - \alpha \cdot r(t) \cdot (r - r(t)) \quad \text{or} \\ w(t+1) &= (1-\alpha) \cdot w + \alpha \cdot w(t) \cdot (1 - w(t) + w), \end{aligned}$$

with a parameter $0 \le \alpha \le 1/r(0)$. It also satisfies a) if the agent chooses $\alpha = 1$, whenever $r(0) = 1$. This learning rule is apparently 'cautious' in tending to maintain a prejudice of high risk (low weight). The parameter value $\alpha = 0$, where the agent gives up his belief for his knowledge in the first time step corresponds to what might be called 'full rationality'. The same is true for the special case $r(0) = 0$. The other extreme is $\alpha = 1/r(0)$ where the agent is 'totally stubborn', sticking to his prejudice. When $r(0) < 1$, a choice $\alpha < 1/r(0)$ assures that the learning map for Δ is contractive (in the first step) so that the agent can expect to reduce his error.

Introducing the **relative error** $\delta(t) \equiv \Delta(t)/r$ we obtain

$$\begin{aligned} \delta(t+1) = \alpha \cdot \frac{\Delta(t)}{r}\left(r - \Delta(t)\right) = \alpha r \cdot \frac{\Delta(t)}{r}\left(1 - \frac{\Delta(t)}{r}\right) = \\ = f_\rho(\delta(t)) \equiv \rho \cdot \delta(t)(1-\delta(t)), \quad \text{with } \rho \equiv \alpha r. \end{aligned}$$

This is the *logistic map*, one of the best-known examples of a discrete time dynamical system exhibiting transition to chaos when varying the *bifurcation parameter* ρ from 0 to 4, in which range it is a self-mapping of the unit interval, see [**7**, setion 7–4] for a detailed description of its qualitative features.

1.2.3. Classes of Behaviour. The behaviour of the learning agent is entirely determined by the pair $(r(0), \alpha)$ and the true value of r via the parameter ρ. For the following qualitative discussion we assume $r(0) \in (0,1)$ and $\alpha \in [0, r(0)^{-1})$ to be invariable characteristics of the agents. According to the changing characteristics of f_ρ as ρ varies, we can coarsely divide a population of agents apploying the logistic learning rule into four subpopulations:

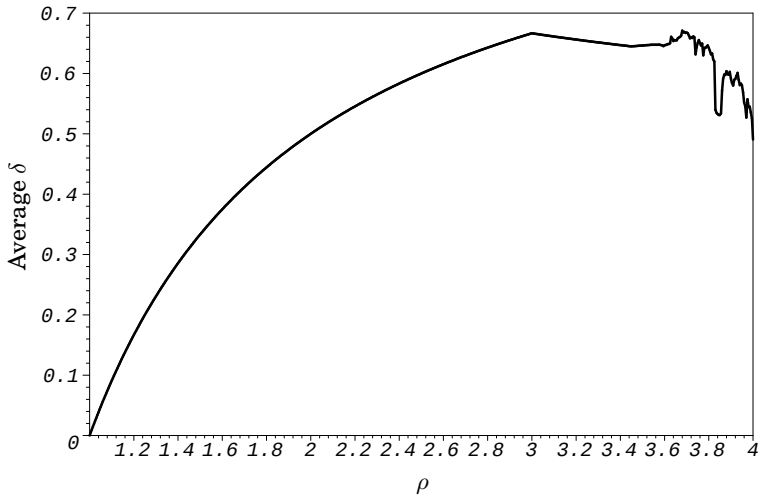

FIGURE 1.2. Average relative error of the logistic learning map. For $\rho > 3.5$ numerical averages over 1000 iterations were taken with 400 initial steps starting from a random value $\delta(0)$ omitted.

Adaptive (A): In the domain $\rho \leq 1$ ($\alpha \leq 1/r$), the iterated map f_ρ approaches 0 and the agent adapts his belief to his knowledge. Since initially $\alpha < 1/r(0)$, this is in particular the case for $r/r(0) \leq 1$. Heuristically this means that a prejudice initially overestimating the risk leads to adaptive behaviour and can therefore be regarded as 'safe'.

Consequently, the following three cases can only occur if the agent underestimates the risk, *i.e.*, $r(0) < r$.

Stubborn (S): The fixpoint 0 of f_ρ becomes unstable at $\rho = 1$. In the range $1 < \rho \leq 3$ ($1/r < \alpha \leq 3/r$), the unique stable fixpoint for δ is $\delta^* = (\rho - 1)/\rho > 0$ (for r it is $r^* = 1/\alpha$). These agents quickly approximate δ^* and stubbornly stick to this underestimation of risk. Due to $\rho = \alpha r < r/r(0)$, the maximum possible δ^* is $1 - r/r(0)$ which is in turn maximized by $2/3$.

Incertain (I): Above $\rho = 3$ ($3/r < \alpha \leq 4/3$), the fixpoint δ^* becomes unstable and first bifurcates into a 2-cycle. That is, the agent oscillates between two underestimations of risk. His behaviour becomes more and more erratic as further bifurcations occur and finally f_ρ becomes chaotic at $\rho = 4$. Instead of further vaguely differentiating, we subsume the case $\rho \in (3, 4]$ under the label 'incertainty'.

Catastrophic: For $\rho > 4$, f_ρ is no longer a self-mapping of $[0, 1]$ and $\delta(t)$ quickly diverges to $-\infty$. We simply exclude this case as unrealistic (as would natural selection presumably do).

It should be noted that by the constructive condition $\rho < r/r(0)$, the more erratic and thus more incertain behaviour in population I is bound to occur with increasing discrepancy between an agents prejudice $r(0)$ and the truth r. This could heuristically be interpreted such that the agent takes a more 'elusive'

strategy when he has to concede an increasing contrast between his belief and reality — a feature which does not sound too unrealistic.

Figure 1.2 shows the expectation value of the relative error δ for populations S and I, which is proportional to the damage an agent has to expect by being stubborn or incertain. Interestingly, the incertain agents perform slightly better than the most stubborn ones.

1.2.4. A Priori Estimates for the Subpopulations. Assuming that the values $(r, r(0), \rho)$ are uniformly distributed in the domain $(0,1)^2 \times \{0 \le \rho < r/r(0)\}$ we can calculate the expectation values of relative sizes of the four subpopulations. Firstly, the probability for an agent to be adaptive is most simply calculated yielding

$$\begin{aligned} \mathrm{P_A} &= \mathrm{P}(r \le r(0)) + \mathrm{P}(\rho \le 1 | r > r(0)) \cdot \mathrm{P}(r > r(0)) = \\ &= \frac{1}{2} + \int_0^1 \mathrm{d}r \int_0^r \mathrm{d}r(0) \frac{r(0)}{r} = 3/4. \end{aligned}$$

Furthermore, given $a > 1$ we can calculate

$$\mathrm{P}(\rho \ge a) = \frac{1}{2}\mathrm{P}(\rho > a | r > r(0)) = \int_0^1 \mathrm{d}r \int_0^r \mathrm{d}r(0)\, \sigma_a(r, r(0)),$$

with

$$\sigma_a(r, r(0)) \equiv \begin{cases} \frac{r(0)}{r}\left(\frac{r}{r(0)} - a\right) & \text{for } a < \frac{r}{r(0)}, \\ 0 & \text{otherwise.} \end{cases}$$

This finally yields

$$\mathrm{P}(\rho \ge a) = \int_0^1 \mathrm{d}r \int_0^{r/a} \mathrm{d}r(0) \left(1 - \frac{a r(0)}{r}\right) = \int_0^1 \mathrm{d}r \frac{r}{2a} = \frac{1}{4a}.$$

From this, we easily derive the probabilities of the other subpopulations: $\mathrm{P_S} = 1/6$, $\mathrm{P_I} = 1/48$, while the relative share of catastrophic agents is $1/16$. Omitting the catastrophic agents and normalizing, we arrive at relative expected quotas of $4/5$, $8/45$, and $1/45$ for the adaptive, stubborn, and incertain subpopulations respectively.

Next, we want to obtain expectation values for the additional damage the agents experience in the respective subpopulations, *i.e.*, the expected values of δ and Δ in the respective parameter ranges. For the stubborn agents, these can be calculated analytically to give

$$\mathrm{E_S}(\delta) = \frac{1}{V_S} \int_0^1 \mathrm{d}r \int_0^r \mathrm{d}r(0) \int_1^{\min(r/r(0),3)} \mathrm{d}\rho\, \delta^* = 1 - \frac{2}{3\ln 3} \approx 0.393,$$

where $V_S = 1/2\ln 3$ is the volume of the integration domain. Similarly we find $\mathrm{E_S}(\Delta) = 2/3\mathrm{E_S}(\delta) \approx 0.262$. In the incertain case, (partially numerical) evaluation of the corresponding integrals yields the expectation values $\mathrm{E_I}(\delta) \approx 0.640$ and $\mathrm{E_I}(\Delta) = 2/3\mathrm{E_I}(\delta) \approx 0.427$. Thus, while individual incertain agents can perform slightly better than the most stubborn ones, the expected error of the incertain population as a whole is nevertheless significantly larger. A total population of A, S, and I-agents with uniform distribution of the parameters in the admissible ranges will have the expected errors $\mathrm{E}(\Delta) \approx 0.0561$ and $\mathrm{E}(\delta) \approx 0.084$.

1.3. Noisy Empirical Data

We return to the general case where $\eta(t)$ is learned from empirical data. It thus underlies statistical fluctuations, which can be written as

$$\eta(t) = r + \Xi(t),$$

where $\Xi(t)$ is a random variable which we assume to be symmetrically and finitely distributed around 0 (such that $\eta(t)$ remains in $[0,1]$). Its variance will decrease over time as prescribed by the convergence of the statistical learning rule adopted. Using relative coordinates $\delta(t)$ and $\xi(t) \equiv \Xi(t) \cdot \rho/r$, we can separate the fluctuations from the learning map:

$$\begin{aligned}\delta(t+1) &= f_\rho(\delta(t)) + \frac{\xi(t)}{\rho} \cdot (\rho(1-\delta(t)) - 1) = \\ &= f_\rho(\delta(t)) + \xi(t)(\delta(t) - \delta^*).\end{aligned}$$

To keep $\delta(t)$ within the range $[0,1]$, we supplement the system by imposing periodic boundary conditions and arrive at

$$\delta(t+1) = f_\rho(\delta(t)) + \xi(t)(\delta(t) - \delta^*) \mod 1. \tag{1}$$

This is an example of a *random dynamical system*. In this special case, it is a **dynamical system with inhomogeneous noise**, where the inhomogenity $\delta(t) - \delta^*$ vanishes at the fixpoint. While systems with homogeneous (additive) noise have been intensively studied, the inhomogeneous case scarcely appears in the literature. The system (1) is very interesting on its own, and thus we undertake a more detailed study of its mathematical theory and its phenomenology in chapter 2.

For the moment, we want to give an upper bound for the random perturbations $\xi(t)$. For this, we assume that ξ is an i.i.d. random variable, uniformly distributed in the range $[-\sigma/2, +\sigma/2]$ with spread $\sigma > 0$, *i.e.*, we ignore the influence of the statistical learning process on the knowledge η. Then to keep $\eta(t)$ within $[0,1]$, Ξ must be a random variable which is uniformly and symmetrically distributed around 0 with spread $\min(r, 1-r)$. Since $\rho = \alpha r \leq r/r(0)$ we find $\sigma \leq 2/r(0) \cdot \min(r, 1-r)$. This shows that the **noise level** σ can in principle be arbitrarily high.

1.4. Empirical Consequences

Let us briefly recapitulate what we have found so far: Starting from the heuristics that under circumstances where learning from experience is rather ineffective and slow compared to the rate at which decisions have to be made, it can be reasonable for a learning agent to behave prejudiced. The key point in our concept of prejudice is the separation of the agents knowledge $\eta(t)$ and its belief $r(t)$. This is reflected by a separation of learning rules into a statistical learning rule for the knowledge (which we haven't considered so far) and a prejudiced learning rule for the belief. For the latter we proposed a simple model in section 1.2, based on a few generic assumptions. This prejudiced learning rule turned out to be equivalent to the well known logistic map. This enabled us to coarsely identify four different classes of behaviour that can be shown by prejudiced agents. Proceeding one step further towards realistic models, we considered in the last section the case where the agent's knowledge exhibits stochastic fluctuations, but we still did not specialize further with respect to the statistical learning rule governing the evolution of the knowledge. We finally showed that this rather generic prejudiced learning system is a dynamical system subjected to a special inhomogeneous form of noise as given

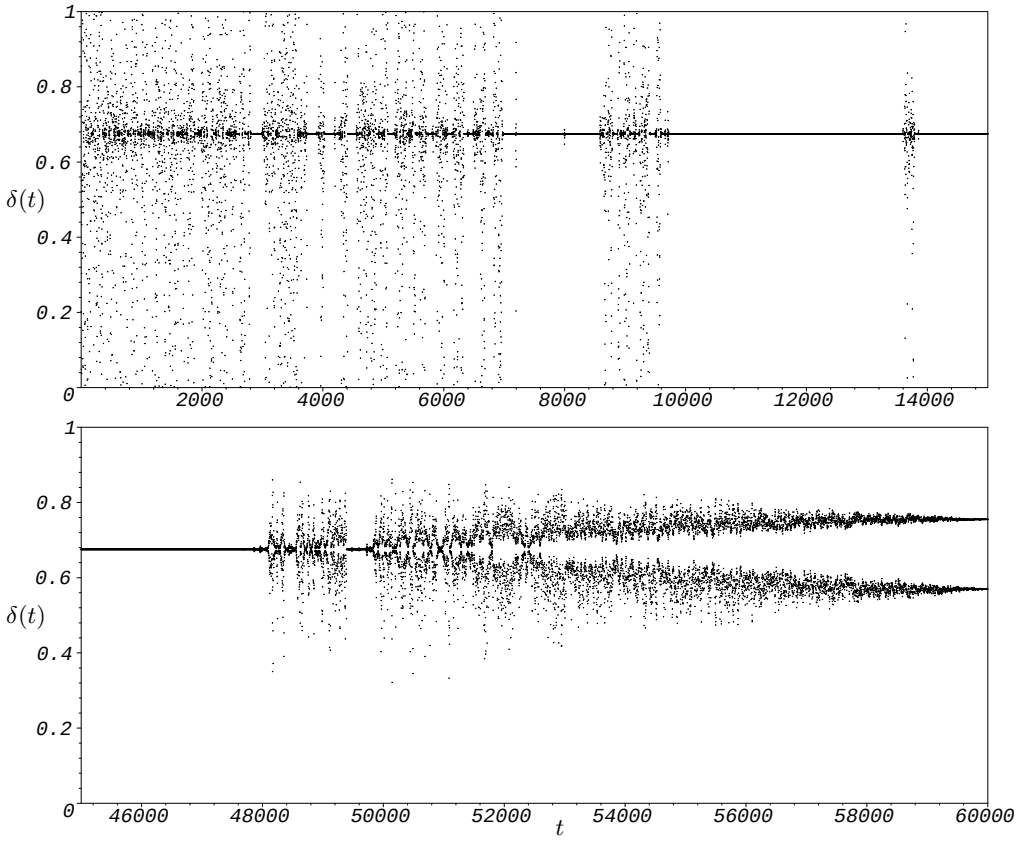

FIGURE 1.3. Example evolution of the relative error $\delta(t)$ under (1) with noise level σ linearly decreasing from 6 to 0 in 60000 timesteps starting from a random initial value.

by formula (1). We have devoted the second chapter of this report to a detailed study of this noisy dynamical system and its rich phenomenology.

In this final section, we want to indicate briefly, which phenomena can be expected to occur when a prejudiced learning system evolves in a noisy environment. To that end, one would in most realistic cases have to specify the statistical learning rule by which the agent learns his knowledge, and our simple assumption of knowledge variable with additive i.i.d. random fluctuations would break down (apart from cases where a very simple system without statistical learning is realistic.) Nevertheless, since statistical learning theory is a vast subject of its own, and selecting a special learning rule seems to imply an undue loss of generality, we stick to the noisy dynamical system with i.i.d. noise described in the last section. We will consider just one example exhibiting most of the key features that we expect to occur.

Consider an agent with very small initial belief $r(0) = 0.002$ in a situation where it underestimates the real value $r = 0.02$ by a factor of ten. We choose $\alpha \leq 1/r(0) = 500$ to be 308, leading to $\rho = \alpha r = 3.08$. The upper bound for the noise level of the last section yields $\sigma < 20$ and we choose initially $\sigma = 6$. For the evolution shown in figure 1.3 we simulated 60000 timesteps of evolution of the relative error $\delta(t)$ under (1), with noise level σ decreasing linearly from 6 to 0. Note that a noise level of exactly zero cannot be reached in finite time by

statistical learning in a realistic noisy surrounding — the development of σ is chosen only for demonstrational purposes.

We frankly use the notions of chapter 2 in the following description: At very high noise levels above 4.5, the agent shows a behaviour which is absulotely unknown in the noiseless case. Although for most of the time, the agent shows the deterministic behaviour that we classified as stubborn in the noiseless case, there are bursts of highly irregular behaviour with δ spreading over the whole interval. The bursts exhibit lengths of order 100 and appear with decreasing frequency as σ decreases. This pattern has been called *on-off intermittency* and occurs in the region of (ρ, σ)-values that we call the intermittent phase in section 2.5. At values of σ around 4.5, the agent leaves this phase and enters the stable domain **S**, where it shows stubborn behaviour for a wide σ-range. The phase transition between stable and bifurcative phase at lower σ-values which is discussed in section 2.3 takes place around timestep 48000. It can be interpreted in behavioural terms as a transition from stubborn to the simplest kind of incertain behaviour consisting of flipping between two fixed values of δ. The incertainty is additionally perturbed by the noisy environment and up to timestep 52000, *i.e.*, for $\sigma > 0.8$ the agent can return to stubborn behaviour for some time with significant probability. The transition from stubborn to incertain should theoretically occur at the bifurcation point $\sigma_0(3.08) \approx 1.3683$, but seems to be delayed by about 1000 timesteps. This is due to the fact that during his being in the stable phase, the agent has reached a δ very close to the fixed point δ^* (with a difference $|\delta^* - \delta(t)|$ as small as 10^{-20}) and it thus takes some time to leave this 'microscopic' domain.

To conclude, some parts of the behaviour of the prejudiced agent in a noisy environment can be interpreted in terms of the classes of noiseless behaviour (especially stubborn and incertain), and transitions between them. On the other hand, new phenomena unknown in the noiseless case can be expected to occur when models with noisy surrounding and a statistical learning rule are considered.

CHAPTER 2

Phase Transitions of the Logistic Map with Inhomogeneous Noise

2.1. Introduction

In the first chapter, where we proposed a mathematical model for learning systems exhibiting 'prejudiced' behaviour, we encountered the random dynamical system

$$x_{t+1} = f_{\rho,\xi_t}(x_t) \stackrel{\text{def}}{=} \rho x_t(1-x_t) + \xi_t(x_t - x^*) \mod 1, \tag{2}$$

where ξ_t is a random variable, which we assumed to be symmetrically and independently distributed in the interval $[-\sigma/2, \sigma/2]$ (we will assume ξ to be i.i.d. from now on), and $x^* = (\rho-1)/\rho$ is the unique nonzero fixed point of the unperturbed logistic map $f_\rho \stackrel{\text{def}}{=} f_{\rho,0}$. This chapter is devoted to a detailed theoretical and experimental analysis of this system. Its key property is already obvious from (2): The noise is coupled inhomogeneously to the dynamics and vanishes at the fixed point x^*. This property will give special importance to this fixed point, as we will see below.

Noisy alias random dynamical systems in recent years have attracted the interest of physicists, cf. [**11, 10**], as well as that of mathematicians, see [**1**, with a vast bibliography], for two basic reasons: On the physical side, they add a good deal of realism to the theoretical models of many natural phenomena which were hitherto modeled by noiseless dynamical systems, since it has been realized that noise is almost ubiquitous in reality and often significantly affects the behaviour of complicated systems. Mathematically on the other hand, random dynamical systems present an attractive synthesis of the fields of stochastics and general dynamical systems in which many branches of mathematical physics and pure mathematics converge and generate new insights. In both disciplines, it soon became clear that random dynamical systems possess a number of interesting new properties not shown by noiseless systems, most prominently the phenomena *noise induced stability*, *on-off intermittence*, and *stochastic bifurcations* also called *noise induced (phase) transitions* by physicists. We will later on rediscover some of these phenomena in the system (2).

The dynamics of a discrete dynamical system like (2), which is subjected to additive noise, can be captured in it's so called **transition density** as follows (we refer the reader to [**8**] for details about the notions introduced in this section): Denote by $\mathcal{P}_{\rho,\sigma}(x,y)$ the probability density for the mapping of points of the unit interval under the noise map $x \mapsto y \stackrel{\text{def}}{=} \xi(x-x^*) \mod 1$, *i.e.*, $\int_I \int_J \mathcal{P}_{\rho,\sigma}(x,y)\mathrm{d}y\mathrm{d}x$ yields the probability for ending in interval $J \subset [0,1]$ when applying the inhomogeneous noise to the points of interval $I \subset [0,1]$. From this, one easily recovers the transition density of the total random dynamical system (2) to be $\mathcal{P}_{\rho,\sigma}(f_\rho(x), y)$, *i.e.*, this is the probability density for the evolution of points under $f_{\rho,\sigma}$. For our system (2), we have

$$\mathcal{P}_{\rho,\sigma}(x,y) = (\sigma|x-x^*|)^{-1} \cdot \chi_{[x-\frac{\sigma}{2}|x-x^*|, x+\frac{\sigma}{2}|x-x^*|]}(y),$$

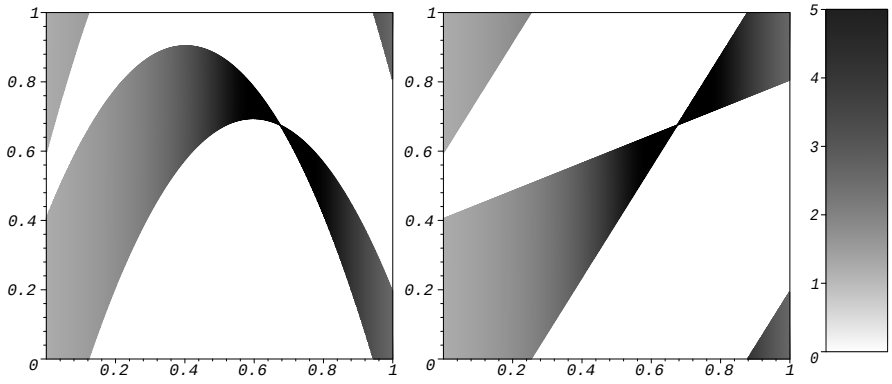

FIGURE 2.1. Transition densities of the logistic map with inhomogeneous noise (left) and of the noise alone (middle) for $\rho = 3.08$ and $\sigma = 1.2$. The used greyscale is shown to the right.

where χ is the characteristic function of the given interval. This density is shown in the right diagram of figure 2.1, while the total transition density $\mathcal{P}_{\rho,\sigma}(f_\rho(x), y)$ is shown in the left picture. Note that in the above formulae, we implicitely assumed periodic boundary conditions, *i.e.*, all variables represent points on the torus S^1.

The transition densities serve to define the **Perron–Frobenius (P–F) operator** of a noisy dynamical system, which is a central tool in its analysis, see **[3, 8]**. In our case, this operator acts on functions $u \in L^1([0,1])$ by

$$\mathbf{PF}(u)(y) \stackrel{\text{def}}{=} \int_0^1 \mathcal{P}_{\rho,\sigma}(f_\rho(x), y) u(x) \mathrm{d}x. \tag{3}$$

Of key importance are the eigenvalues and eigenvectors of **PF**. Especially, the positive and normalized eigenvectors to the highest eigenvalue 1, *i.e.*, probability densities u with $\mathbf{PF}(u) = u$, are called **invariant densities** of the system. An invariant density u defines an associated **invariant measure** μ_u by $\mu_u(A) \stackrel{\text{def}}{=} \int_A u(x)\mathrm{d}x$, where A is any Lebesgue-measurable set. What one really hopes to find is the so-called *physically significant* or *Sinai–Bowen–Ruelle (SBR) measure* of the system. An SBR-measure $\mu_{\text{SBR}}$is defined for an ordinary, *i.e.*, non-random dynamical system given by a deterministic mapping f by the following property: There exists a subset U of the configuration space considered and with positive Lebesgue measure, such that for every continuous function ψ one has

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j=0}^{N-1} \Psi(f^j(x)) = \int \Psi \mathrm{d}\mu_{\text{SBR}}, \tag{4}$$

for all starting points $x \in U$, where f^j denotes the j-th iteration of f, see **[4]**. By the Birkhoff individual ergodic theorem this property is always fulfilled for μ_{SBR}-*almost all* x. The crucial strengthening of the hypothesis lies in the assumption that the ergodic hypothesis can be safely applied for all starting points x in a set of positive Lebesgue measure — a property which indeed renders SBR-measures physically significant. Note that for a random dynamical system, the above property would have to be formulated in the mean with respect to the stochastic perturbation.

Since most physically interesting quantities of a dynamical system are time averages, an ergodic hypothesis is regularly invoked by physicists to calculate them by space averages, *i.e.*, by spatial integration. This makes the existence and uniqueness of SBR-measures an important theoretical issue in the study of random and ordinary dynamical systems. We will not consider these questions but in the next section apply the notions as given to our perturbed logistic map.

2.2. Stable Phase and Lyapunov Exponent

From the explicit form of the transition density $\mathcal{P}_{\rho,\sigma}$ and figure 2.1, it is apparent that the fixed point x^* plays a special role in the dynamics of (2), since it is the only point where the noise vanishes for $\sigma \neq 0$. Indeed, what one can easily infer from a sample of simulations is that there might be a certain range of parameter values (ρ, σ), for which evolutions from almost all starting points in the unit interval eventually end up at x^*. This would mean that the δ-measure concentrated at x^* is the — in this case unique — SBR-measure for the system in this parameter range. First, to see that $\delta(x - x^*)$ is at least an *invariant* density, we perform the following formal calculation:

$$\begin{aligned}\mathbf{PF}(\delta(x-x^*))(y) &= \int_0^1 \mathcal{P}_{\rho,\sigma}(f_\rho(x), y)\delta(x-x^*)\mathrm{d}x \\ &= \mathcal{P}_{\rho,\sigma}(f_\rho(x^*), y) = \mathcal{P}_{\rho,\sigma}(x^*, y) \\ &= \delta(y-x^*).\end{aligned}$$

Here, we used the fact that the density $\mathcal{P}_{\rho,\sigma}(x,\cdot)$ approximates the function δ_{x^*} for $x \to x^*$.

To have a name for it, we call the hypothetical domain where $\mu_{\mathrm{SBR}} = \mu_{\delta_{x^*}}$ the **stable phase** of (2). The question naturally arises wether there is a theoretical characterization of the stable phase. In the case of ordinary one-dimensional dynamical systems without noise, changes in the dynamics are fully characterized by the behaviour of a single order parameter, namely the so called **Lyapunov exponent**. This exponent λ measures how fast nearby trajectories diverge (if λ is positive) or converge onto an attractive set (if $\lambda < 0$) under the given dynamics. It is defined by the time average

$$\lambda \overset{\mathrm{def}}{=} \lim_{N\to\infty} \frac{1}{N} \sum_{j=0}^{N-1} \ln|f'(x_j)|, \quad x_j \overset{\mathrm{def}}{=} f^j(x_0), \tag{5}$$

for a generic dynamical mapping f and an arbitrary starting value x_0. Whenever an SBR-measure exists, for f this can be rewritten as a space average using (4):

$$\lambda = \int_X \ln|f'(x)|\mathrm{d}\mu_{\mathrm{SBR}}(x).$$

A well known example for how the dynamics can be read off from λ is the logistic map f_ρ, see the overview in [**7**, setion 7–4]. For $1 < \rho < 3$, one has $\lambda < 0$ and this is the domain where the fixed point x^* is attractive. The point $\rho = 3$ is an isolated zero of λ and marks the first **bifurcation** of f_ρ, *i.e.*, the point where the fixed point becomes unstable and a stable two-cycle appears. In the deterministic case, the Lyapunov exponent is considered a good measure for the physical complexity of a dynamical system. Especially if λ is positive for a range of parameter values, then this is an indication of chaotic behaviour.

It is clear that the above definition of the Lyapunov exponent readily generalizes to one-dimensional noisy dynamical systems such as (2) by simply writing

$$\lambda \stackrel{\text{def}}{=} \lim_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N-1} \ln |f'_{\rho,\xi_t}(x_t)|. \tag{6}$$

If we had an SBR-measure at hand, we could also calculate λ via a space average by :

$$\lambda = \left\langle \int_0^1 \ln |f'_{\rho,\xi}(x)| \mathrm{d}\mu_{\mathrm{SBR}}(x) \right\rangle_\xi ,$$

taking into account that in the noisy case we also have to average over the random variable ξ.

It has been questioned wether this Lyapunov exponent plays the same indicative role in the noisy as in the noiseless case, and certain examples have been found where a positive λ does not mark what one would normally call chaotic behaviour, see [**10, 11**]. Nevertheless, for our system (2) we are guided by the heuristics that in the stable phase the asymptotic dynamics is not affected by the noise. Thus we can speculate about the shape of the stability domain and its boundary in terms of (zeroes of) the Lyapunov exponent $\lambda = \lambda(\rho, \sigma)$, depending on the dynamical parameter ρ and the noise level σ. First, the point $(\rho = 3, \sigma = 0)$ marks the 'boundary of the stable phase' in the noiseless case. Assuming that λ is smooth (or at least continuous) in (ρ, σ), there could — in the simplest case — be a nodeline of $\lambda(\rho, \sigma)$ starting at $(3, 0)$. Then, we can speculate:

CONJECTURE. *In the domain $\rho \geq 3$, there exists a stable phase for* (2) *defined by the property $\mu_{SBR} = \mu_{\delta_{x^*}}$. It is characterized by negative values of $\lambda(\rho, \sigma)$ and lies on the side of smaller ρ of a nodeline of $\lambda(\rho, \sigma)$ which extends from $(3, 0)$ to increasing values of σ.*

If we assume that we are in the stable phase, we can calculate the values of what we will call the **stable Lyapunov exponent** $\lambda_{\mathrm{S}}(\rho, \sigma)$ analytically. For our special realization of noise ξ which exhibits an i.i.d. distribution in the interval $[-\sigma/2, \sigma/2]$, the mean in ξ in (6) is nothing but $1/\sigma$ times the integral from $-\sigma/2$ to $\sigma/2$ w.r.t. the Lebesgue measure over the inner integral, *i.e.*,

$$\lambda_{\mathrm{S}}(\rho, \sigma) = \frac{1}{\sigma} \int_{-\sigma/2}^{\sigma/2} \int_0^1 \ln |f'_{\rho,\xi}(x)| \mathrm{d}\mu_{\mathrm{SBR}}(x) \mathrm{d}\xi$$

Inserting the assumption $\mu_{\mathrm{SBR}} = \mu_{\delta_{x^*}}$ yields

$$\begin{aligned} &= \frac{1}{\sigma} \int_{-\sigma/2}^{\sigma/2} \ln |f'_{\rho,\xi}(x^*)| \mathrm{d}\xi \\ &= \frac{1}{\sigma} \int_{-\sigma/2}^{\sigma/2} \ln |\rho(1 - 2x^*) + \xi| \mathrm{d}\xi \\ &= \frac{1}{\sigma} \int_{-\sigma/2}^{\sigma/2} \ln |\xi - \rho + 2| \mathrm{d}\xi \end{aligned}$$

For $\rho - 2 > \sigma/2$ this is

$$= \frac{1}{\sigma} \int_{-\sigma/2}^{\sigma/2} \ln(\rho - 2 + \xi) \mathrm{d}\xi,$$

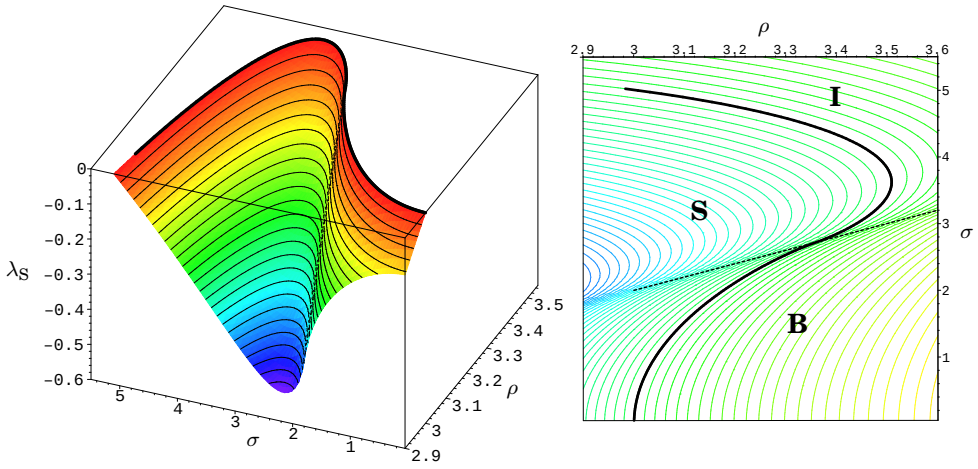

FIGURE 2.2. Theoretical values of the stable Lyapunov exponent λ_S. The solid curve in both pictures is the nodeline $\lambda_S = 0$, while the dotted line in the right picture marks the line $\rho - 2 = \sigma/2$, where the different solutions in (7) connect.

and by a change of variable $\zeta = \rho - 2 + \xi$ we finally find with the abbreviations $\Delta_\pm = \rho - 2 \pm \sigma/2$

$$\begin{aligned}\lambda_S &= \int_{\ln \Delta_-}^{\ln \Delta_+} \frac{\zeta e^\zeta}{\sigma} d\zeta \\ &= \left[\frac{\zeta - 1}{\sigma} e^\zeta\right]_{\ln \Delta_-}^{\ln \Delta_+}.\end{aligned}$$

Similar calculations in the two other cases $\rho - 2 < \sigma/2$ and $\rho - 2 = \sigma/2$ yield the net result

$$\lambda_S(\rho, \sigma) = \frac{1}{\sigma} \begin{cases} (\ln \Delta_+ - 1) \ln \Delta_+ - (\ln \Delta_- - 1) \ln \Delta_-, & \text{if } \Delta_- > 0; \\ (\ln \sigma - 1) \ln \sigma, & \text{if } \Delta_- = 0; \\ (\ln \Delta_+ - 1) \ln \Delta_+ + (\ln |\Delta_-| - 1) \ln |\Delta_-|, & \text{if } \Delta_- < 0. \end{cases} \quad (7)$$

The two pictures in figure 2.2 show that the above formulae lead to a single nodeline for λ_S starting at $(3, 0)$. Also, for $\rho < 3$ one finds that λ_S smoothly approaches the theoretical value $\ln |2 - \rho|$ for the Lyapunov exponent of the noiseless logisitc map if we let σ approach 0 in the range $2 \le \rho \le 3$. On the other hand, λ_S does not reproduce the complicated behaviour of the deterministic Lyapunov exponent $\lambda(\rho, 0)$ for $\rho > 3$. The **turning line** $\Delta_- = 0$ marks turning points of the function $\lambda_S(\rho, \sigma)$ in σ. It intersects with the nodeline in the point $(e/2 + 2, e)$, which we tentatively call **triplepoint**. We marked three different domains in the plan view on the right side of figure 2.2 by the letters **S** for the domain left to the nodeline of λ_S, **B** for the domain right of the nodeline and below the turning line, and **I** for the rest. The meaning of **S** is clear at this point: it stands for the conjectured stability domain. The other two will be discussed shortly.

We will now test the conjecture about the stable phase empirically by comparing λ_S with numerical approximations of the Lyapunov exponent $\lambda(\rho, \sigma)$. From now on, we restrict our attention to the domain $(\rho, \sigma) \in [3, 4] \times [0, 5]$. Numerically, λ was calculated using the classical time averaging formula (5). For

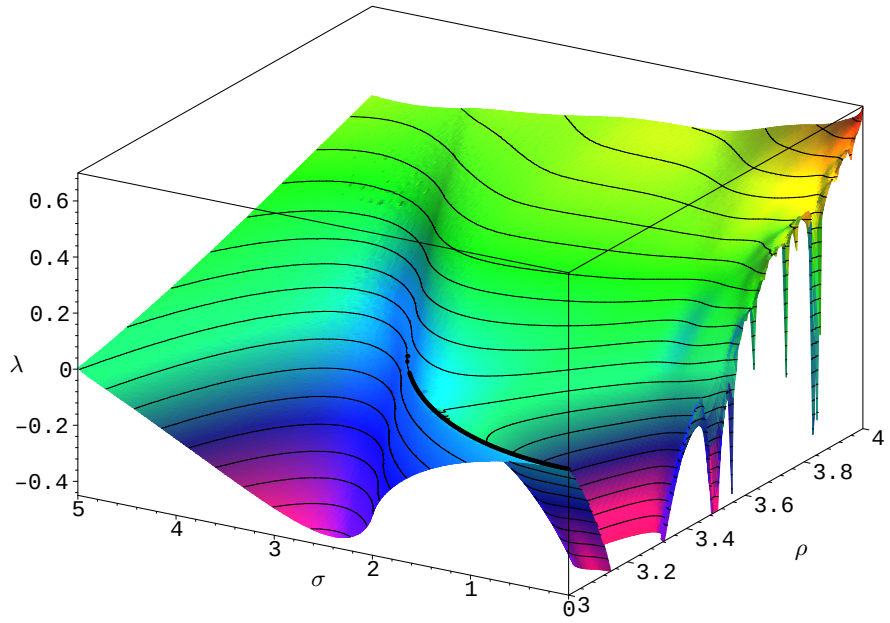

FIGURE 2.3. Perspectivic view of the function $\lambda(\rho, \sigma)$. The values were calculated on a regular grid with resolution $(0.0025, 0.05)$ in (ρ, σ), by numerical simulation of 25 independent evolutions each of length 10^6, and after omitting 10^4 initial steps in order to start from the invariant density. This yields a statistical error significantly below 10^{-2}.

every point, 2.5×10^7 iterations were sampled which leads to a statistical accuracy better than 10^{-3} for the average. The result is shown in figure 2.3. We find that λ is identical to λ_S within the statistical error bounds for the most part of region **S**, which indicates that the stable phase indeed exists. Especially, the nodeline of λ_S (thick curve in figure 2.3) is met by the λ-surface with good accuracy. In region **B**, the Lyapunov exponent approaches that of the logistic map for $\sigma \to 0$ and restores the full bifurcation pattern for $\sigma = 0$ (thence the letter **B**). The turning line of λ_S keeps its significance since it does also mark turning points of the function $\lambda(\rho, \cdot)$. But there are further differences between the areas separated by this line:

A closer examination of the picture unveils that in domain **I**, λ is strictly larger than λ_S, as can, *e.g.*, be guessed from the vanishing of the nodeline below the λ-surface near the triplepoint. Let us introduce the difference $\Delta\lambda \stackrel{\text{def}}{=} \lambda_S - \lambda$ between the stable and the empirical Lyapunov exponents. Figure 2.4 shows a detail view of this quantity in the area connecting domains **B** and **I**. We see that the part of the stability domain above the turning line is not bounded by the nodeline of λ_S but by a curve well within the area **S**. If the positivity of the Lyapunov exponent in domain **I** marks an irregular behaviour of the system, we can interpret this as an intrusion of irregularity into the stable domain. We will come back to this in section 2.5.

Now that we have gained some orientation by examining the Lyapunov exponent of the system (2), let us consider the different phenomenologies of the system in the different domains and especially their changing with varying noise parameter σ.

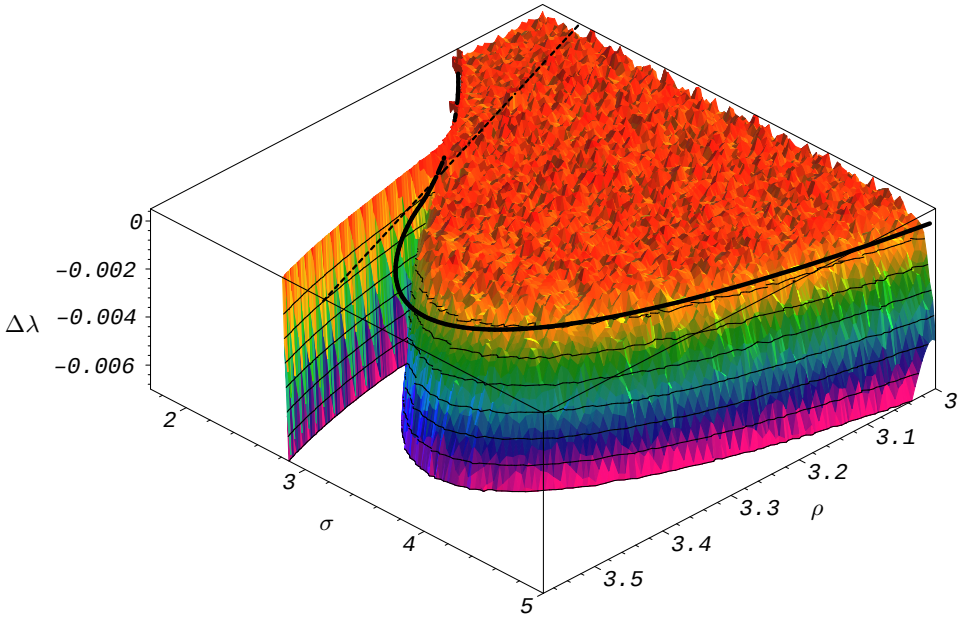

FIGURE 2.4. Magnified view of $\Delta\lambda$.

2.3. Phase Transition S→B

If we let σ approach zero starting from the stable phase, we expect a transition of the behaviour of the system from stability at the fixed point to the deterministic behaviour dictated by the bifurcation parameter ρ. We consider this transition for ρ-values only slightly above 3 where the deterministic behaviour is a 2-cycle with attractor $\{\frac{1}{2\rho}(\rho+1\pm\sqrt{(\rho+1)(\rho-3)})\}$. We denote the smaller value of σ where λ_S crosses 0 by $\sigma_0(\rho)$. Figure 2.5 shows four examples of 'typical' evolutions under (2) for different σ-values for $\rho = 3.08$. In the upper left picture, showing a σ-value only slightly above σ_0, we can see how the system is forced into stability after exhibiting initial fluctuations in the first few thousand steps. This is a striking example of what is called *noise induced stability* in the physical literature, see, *e.g.*, [**5**], although it is not a too surprising phenomenon for our special form of noise. The next three pictures exhibit how the system becomes unstable for $\sigma < \sigma_0$: With decreasing σ, it first exhibits fluctuations into small 2-cycles near x^* which appear and vanish in an intermittent manner. The lifetime of these (noisy) 2-cycles becomes longer as σ further decreases, until almost the whole evolution takes place away from the fixed point. Finally, for vanishing σ the system again gets stabilized, but this time at the deterministic 2-cycle.

To gain a more general insight in how this *noise induced transition* is taking place, we performed a numerical approximation of the invariant densities for (2) at $\rho = 3.08$ and a number of σ-values. This can be done under the assumption that the iterates $\mathbf{PF}^n(u)$ of the density u under the Perron–Frobenius operator of every normalized measure which is absolutely continuous w.r.t. the SBR-measure (and *vice versa*), converge to the unique physically significant invariant density for $n \to \infty$. This is part of the assertion of the Perron–Frobenius–Ruelle Theorem, see [**8, 13**], and assuming it we approximate the physically significant invariant density gathering the statistical distribution of the N-th iteration of M i.i.d. distributed starting points (an approximation

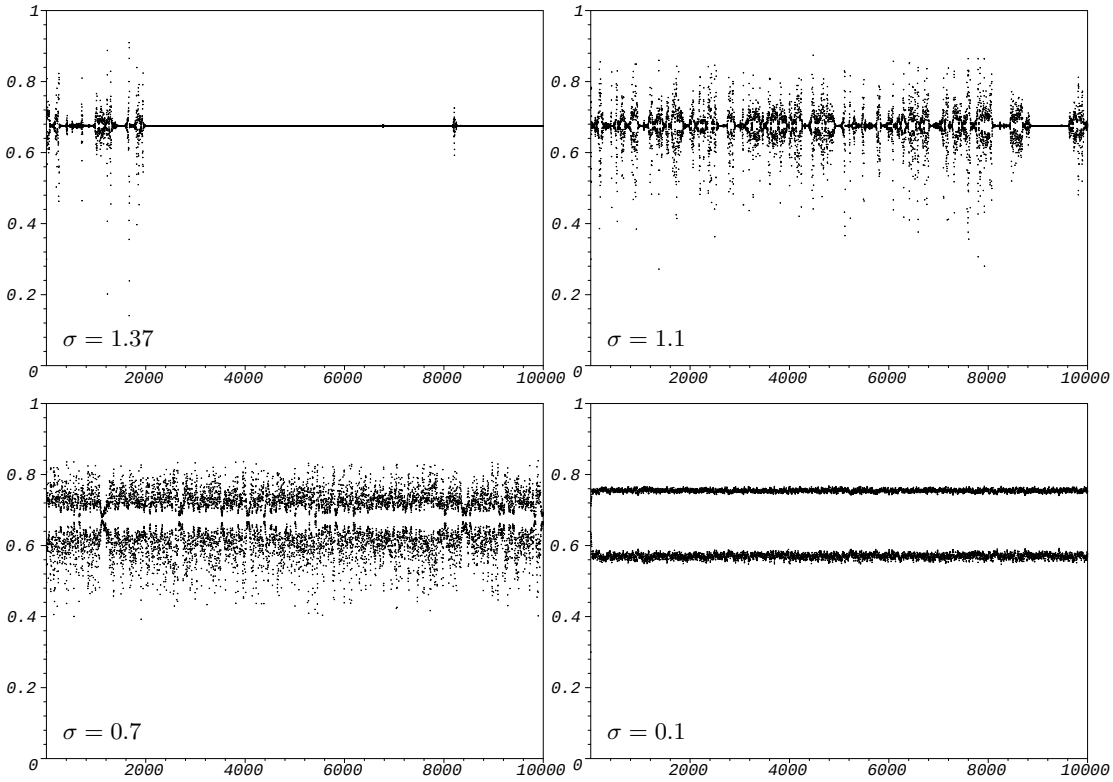

FIGURE 2.5. Examples for evolutions of random initial points under (2) for $\rho = 3.08$ at the four different values of σ which are indicated by the dotted lines in figure 2.6.

of the density 1). Figure 2.6 shows the result for $N = 1000$, $M = 4 \times 10^6$ as a density plot. The first thing one notes is that even for the highest σ-value of 1.375, the density spreads out from the central cell which contains the fixed point x^*. In fact, this cell has an absolute probability of only 0.825. This is due to the fact that the singularity δ_{x^*} is only badly approximated by $\mathbf{PF}^N(1)$ after only thousand iterations, as can be guessed by the upper left picture in figure 2.5, which shows that initial fluctuations around x^* are not marginal after the first few thousand steps.

However for σ below σ_0, we are presented with a smooth decay of the singularity. The invariant density develops near $\sigma \approx 1$ two new maxima which tend to the 2-cycle attractor. This is a typical picture for a so called *stochastic bifurcation*, see [**1**], the general phenomenon of the transition from a noisy attractor with rather simple structure to a more complicated one at lower noise level.

We want to stress the analogy of the noise induced transition found above with phase transitions in physical systems. Especially, our stochastic bifurcation can be interpreted as a *symmetry breaking* phase transition. To show that, we consider the so called **return map** of a discrete dynamical system defined by a map f. It is given by the evolution of the local expansion rates

$$y_i \stackrel{\text{def}}{=} \ln |f'(x_i)|$$

at the i-th step. This notion is directly applicable also in the noisy case. One defines the 'graph' of the evolution of the y_i by looking at the statistical distribution of pairs (y_i, y_{i+1}). Figure 2.7 shows these transition densities for the

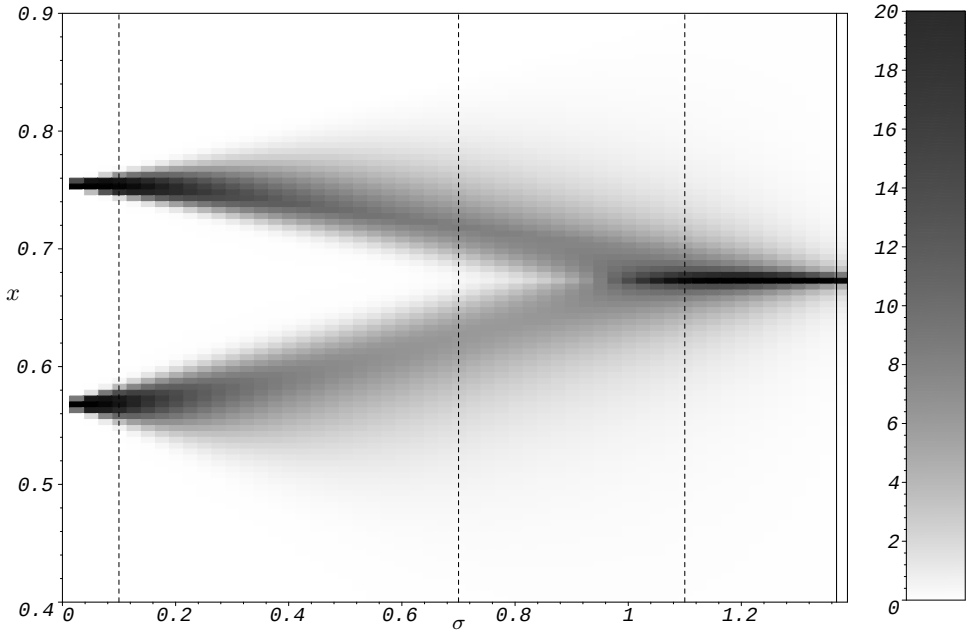

FIGURE 2.6. Density plot of numerical approximations of the invariant densities for (2) at $\rho = 3.08$ and for 55 equidistant σ-values with between 0.025 and 1.375. The approximation took place on a regular partition of $[0, 1]$ into 200 cells. At every value of σ, 4×10^6 values of x_{1000} were sampled. A greyscale is attached to the right. Cells with absolute probabilities below 10^{-3} appear white. $\sigma_0 = 1.3683$ is indicated by a solid vertical line.

return map at $\rho = 3.08$ and σ ranging from 1.5 (upper left picture) to 0 (lower right) in 11 equidistant steps. The value of $\sigma = 1.\overline{36}$ in the second picture lies only slightly below the value $\sigma_0 \approx 1.3683$ where the Lyapunov exponent crosses zero. The vertically reflected angle visible in pictures 3–9 is an artifact due to periodic boundary conditions. For values of σ above σ_0, *i.e.*, in the stable phase, the local expansion rate at time t is nothing but $y_t = \ln(\rho - 2 + \xi_t)$ and the transition density of the return map exhibits the highly symmetrical shape shown in the upper left picture. This symmetry immediately becomes broken when σ passes through σ_0 by a 'leaking' of transition density from the square of stable behaviour into two 'wings' with different shape and statistical weight.

2.4. Poincarè Dimension and Critical Exponents

We are now going to pursue further the analogy between the phenomenological transitions in our noisy dynamical system and phase transitions in physical systems. The theory of *critical phenomena*, *i.e.*, of the behaviour of physical systems near a *critical point* where two phases become identical, cf. [**7, 2**], gives us hints regarding the observables we should look for in order to see whether the stated analogy is tenable: At the critical point, many physical significant quantities become infinite or zero. Some of the latter play a special role in describing the state of the system in the more disordered phase and are thus called *order parameters*. Their values close to the critical point are governed by universal scaling laws of the system considered and described

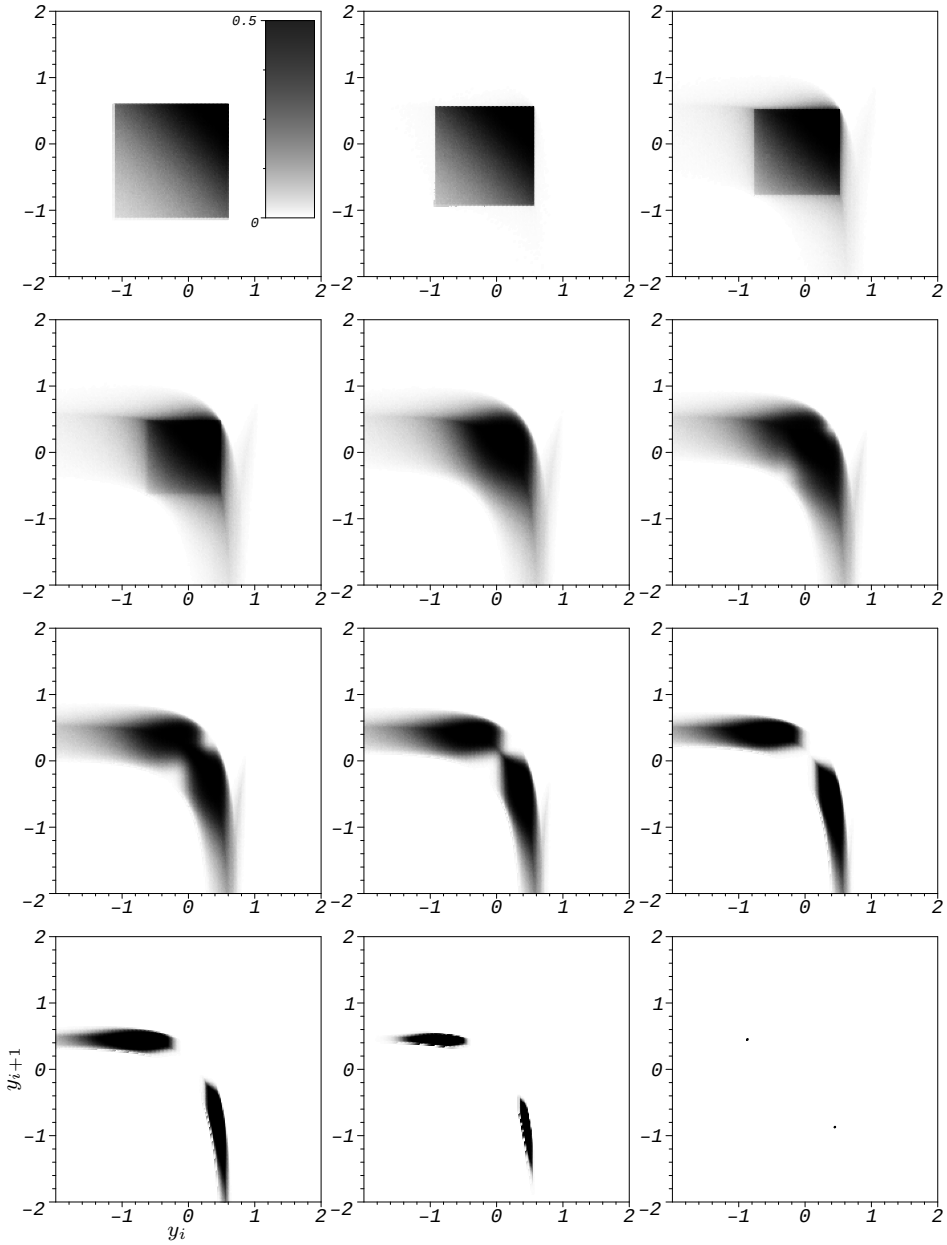

FIGURE 2.7. Transition densities of the return map at $\rho = 3.08$. The used greyscale is inserted in the upper left picture. For every picture, 10 samples of 10^6 iterations (after 10^4 initial iterations) were used to approximate the probability densities on a regular 200×200 partition of $[-2, 2]^2$. Cells with absolute probabilities $\leq 10^{-6}$ were neglected.

by the so called *critical exponents*. We explain the concept by the trivial example of the deterministic logistic map at the first bifurcation point $\rho_c = 3$. This system is described by a single relevant order parameter: The Lyapunov exponent $\lambda(\rho)$. Introducing the scale-free parameter $\tau \stackrel{\text{def}}{=} (\rho - \rho_c)/\rho_c$, the critical exponent of λ is defined as the exponent of ρ of the leading power in the expansion of λ in τ as τ approaches zero from above or below. This can be expressed

by

$$\gamma_\pm \stackrel{\text{def}}{=} \lim_{\tau\to 0_\pm} \frac{\ln\lambda(\tau)}{\ln|\tau|}.$$

For $\rho < \rho_c = 3$ we find $\lambda(\tau) = \ln(\rho-2) = \ln(3\tau+1)$ and thus

$$\gamma_- = \lim_{\tau\to 0_-} \frac{\ln(\ln(3\tau+1)))}{\ln(-\tau)} = 1,$$

by l'Hospital's rule, and a similar calculation yields also $\gamma_+ = 1$. This shows a feature which is conjectured to hold in general: $\gamma_+ = \gamma_-$.

In the noisy case, although the zero of the Lyapunov exponent marks the critical points of the transition from the bifurcative into the stable phase (**B**→**S**), it is not clear whether λ is the only possible order parameter nor what should be its physical interpretation. Therefore, we look for a more physically relevant measure for order in the system near the critical points at $\sigma_0(\rho)$.

It is quite natural to consider the frequency with which the system returns to domains near the attractor x^* of the stable phase, *i.e.*, the *recurrence time statistics* at this fixed point, cf. [**6**]. The **Pioncarè recurrence time**, which is the average time after which the system returns to a small domain $B_r(x^*) \stackrel{\text{def}}{=} \{x \mid |x-x^*| \le r\}$ of radius r around x^*, is given by

$$\overline{T}_{\mathrm{P}}(r) \stackrel{\text{def}}{=} \lim_{N\to\infty} \frac{N}{\#\{x_i \in B_r(x^*),\ 1 \le i \le N\}}.$$

Obviously, this is not a scale-free quantity, so to obtain the desired order parameter, one assumes that $\overline{T}_{\mathrm{P}}$ behaves as

$$\overline{T}_{\mathrm{P}}(r) \propto r^{-D_{\mathrm{P}}} \tag{8}$$

for sufficiently small r. The number D_{P} defined this way is the **Poincarè dimension** of the fixed point x^*. This is the order parameter we will first consider. As the system enters the stable phase with σ approaching σ_0 from below, the $\overline{T}_{\mathrm{P}}(r)$ will approach 1 for all r and thus we expect D_{P} to vanish at the critical point.

Returning to our system (2), we are going to consider the **B**→**S** transition at the four values 3.04, 3.08, 3.12, and 3.16 of ρ. To obtain some initial information we first take a glimpse at the behaviour of the Lyapunov exponent near the four critical points. This is shown in figure 2.8, where λ is plotted against the respective values of the scale-free parameter

$$\tau \stackrel{\text{def}}{=} \frac{(\sigma_0(\rho)-\sigma)}{\sigma_0(\rho)}$$

(defined to be positive in this case) to make the curves comparable. We see that while the behaviours of λ at $\rho = 3.08$ and 3.12 in the left picture are rather regular, the other two ρ-values are a bit different: λ approaches 0 very steep at $\rho = 3.04$ and shows a rapidly changing derivative around σ_0. At $\rho = 3.16$ on the other hand, λ passes a region where it is positive below $\tau \approx 0.1$. We would expect that the positvity of λ in the last case should have some effect on the qualitative features of the transition. Experimentation suggests on the other hand that the transition at $\rho = 3.04$ is not very different from that at $\rho = 3.08$ or 3.12, *i.e.*, we conjecture the transition **B**→**S** to be in one *universality class* of phase transitions for a certain range of ρ-values above 3. This class is possibly left when the nodeline of λ given by $\sigma_0(\rho)$ branches and the system first enters a phase with positive Lyapunov exponent before entering the stable domain. We will see below how justified these conjectures are.

Another thing we see is that the approximate Lyapunov exponents we calculated by finite-time statistics do only badly converge to 0 at $\sigma_0(\rho)$. This is

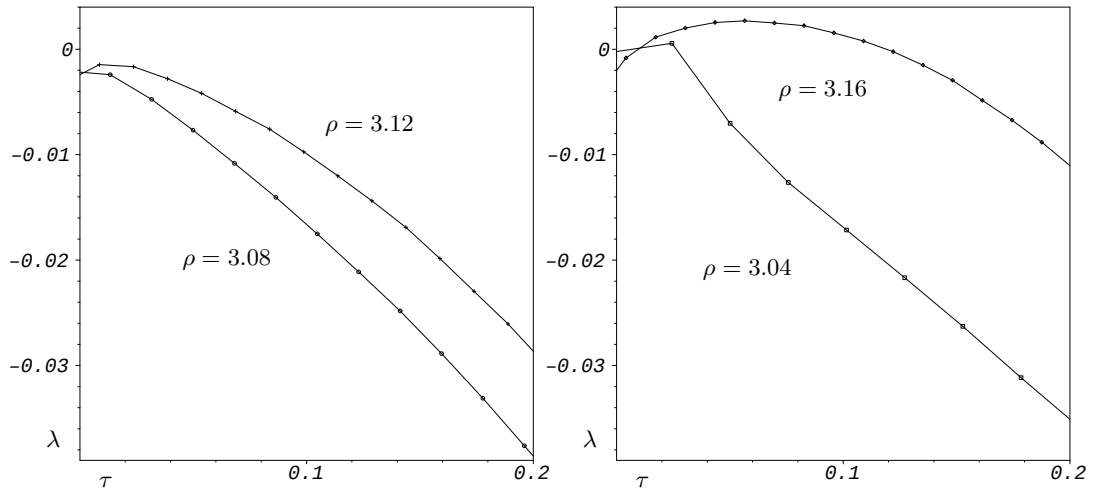

FIGURE 2.8. Lyapunov exponents near the critical points $\sigma_0(\rho)$ for the for indicated values of ρ *vs.* the scale-free parameter τ. Data taken from figure 2.3.

a consequence of the bad statistical behaviour of the system near the divergence of the invariant density which marks the entrance to the stable phase. Such divergencies notoriously show a bad convergence of statistical observables with respect to simulation time. Naturally, this will also cause problems in the computation of D_{P} near the critical point.

We first examine the case $\rho = 3.08$. The slow convergence of $\overline{T}_{\mathrm{P}}(r)$ for increasing **sampling time** N can be seen by running simulations at the 'exact' numerical value of $\sigma_0(3.08)$, *i.e.*, for $\tau = 0$, with different N. The left picture in figure 2.9 shows the empirical statistics for $\overline{T}_{\mathrm{P}}$ at the critical point for three different values of N ranging over three orders of magnitude. One can derive from this picture, that the convergence of $\overline{T}_{\mathrm{P}}$ with N is not only slow but that $\overline{T}_{\mathrm{P}}$ probably will not converge to 0 at all, an effect that can be attributed to the additional 'numerical noise' which is present in any numerical simulation and whose influence, unfortunately, increases with N. Furthermore, we detect a slight nonlinearity in the run of the $\overline{T}_{\mathrm{P}}$-curves in the bilogarithmic plot, contrary to the assumption made in (8). This renders a direct determination of D_{P} a difficult task. As an example, we have shown the recurrence time statistics for $N = 10^5$ at the point $\sigma = 1.36429$, corresponding to $\tau \approx 0.00291$, in the right picture of figure 2.9 (upper curve drawn with box-shaped symbols).

The course we take to remedy these problems is simple: Instead of considering the 'bare' values of $\overline{T}_{\mathrm{P}}$, we consider only the relative recurrence times $\overline{T}_{\mathrm{P}}(r)/\overline{T}^0_{\mathrm{P}}$, where $\overline{T}^0_{\mathrm{P}}$ denotes the recurrence times at the critical point $\tau = 0$. This amounts in taking the difference curve of the upper two curves in the right picture of figure 2.9. The resulting 'corrected' curve is the lower one in this picture, drawn with crosses. This curve is almost indistinguishable from its linear best fit (solid line). The regression yields in this case a 'corrected' value of 0.00331 for the Poincarè dimension. We will consider exclusively these corrected quantities and consequently denote them also by D_{P}.

The left picture of figure 2.10 shows a bilogarithmic plot of $D_{\mathrm{P}}(\tau)$ determined this way for $\rho = 3.08$ and the three values of N noted above. Apart form some exceptions for the smallest $N = 5 \times 10^3$ and values of $\log_{10} \tau < -3$, the results do not show a strong dependence on N. Thus we can try to determine the critical exponent β for the decay of $D_{\mathrm{P}} \propto \tau^\beta$ by linear regression, and we will

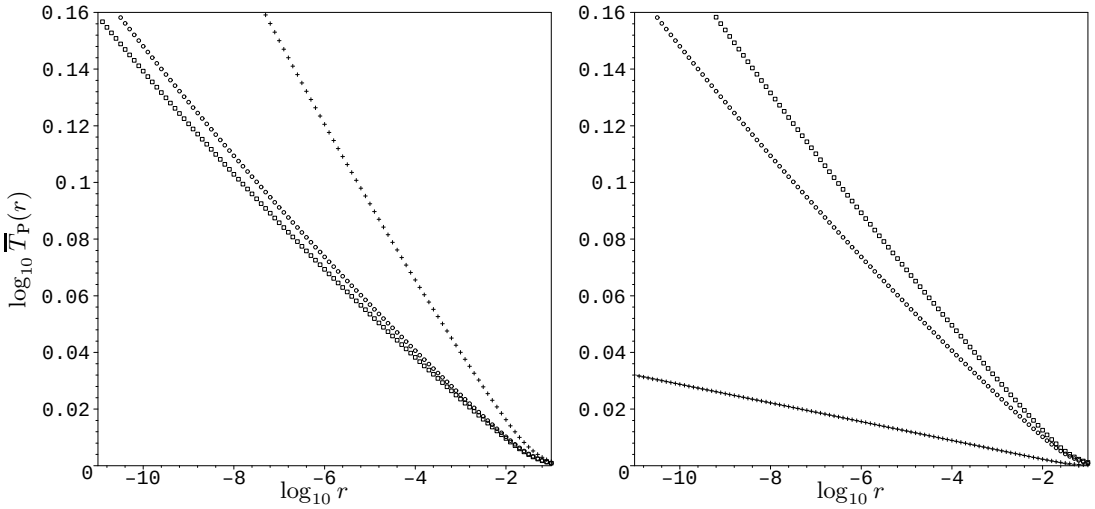

FIGURE 2.9. Left: Recurrence time statistics at the critical point $\sigma_0(3.08)$ for $N = 5 \times 10^3$ (crosses), $N = 10^5$ (circles), $N = 5 \times 10^6$ (boxes). For the first, a total number of $m \cdot N = 2.5 \times 10^8$ points were sampled (m the number of independent runs), while for the second and third $m \cdot N = 5 \times 10^9$. Right: Recurrence time statistics for $\tau \approx 0.00291$ with $N = 10^5$ and $m \cdot N = 5 \times 10^9$ (boxes). The statistics at the critical point is copied from the left picture (circles). The corrected curve (crosses) is almost indistingiushable from its linear fit line. For both pictures, $\overline{T}_{\mathrm{P}}$ was approximated at values of $\log_{10} r$ from -1 to -11 with spacing 0.1.

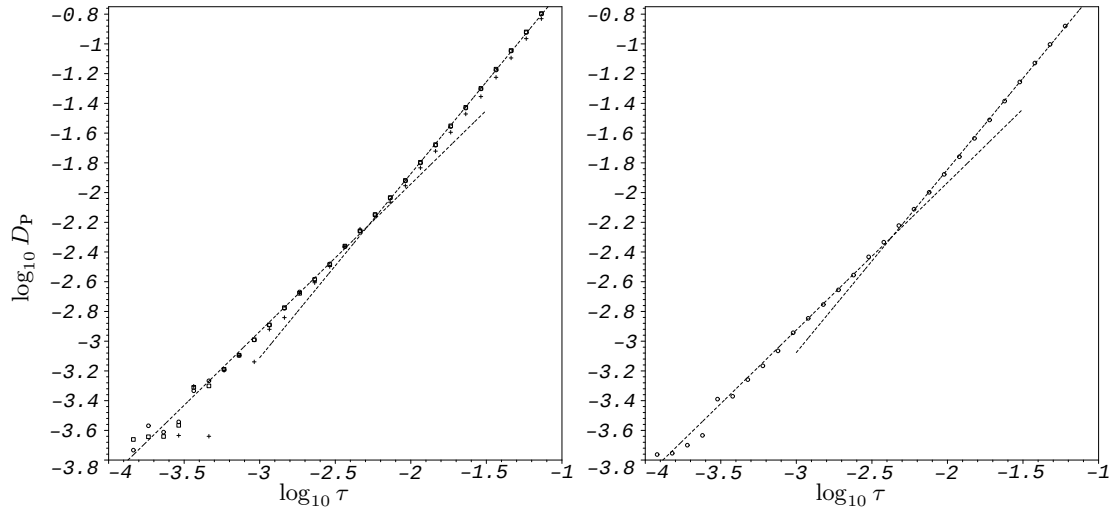

FIGURE 2.10. Left: Poincarè dimensions for $\rho = 3.08$ and $N = 5 \times 10^3$ (crosses), $N = 10^5$ (circles), $N = 5 \times 10^6$ (boxes). Dashed lines are linear regression lines to the data points for $N = 10^5$ above resp. below $\log_{10} \tau = -2.25$. Right: D_{P} for $\rho = 3.12$ with $N = 10^5$ and linear fits.

do so for the data at $N = 10^5$ in the following. One immediately notices that a direct fit to all data points does not lead to a satisfactory result, since there is a certain point between $\log_{10} \tau = -2.0$ and -2.5 at which the inclination of the

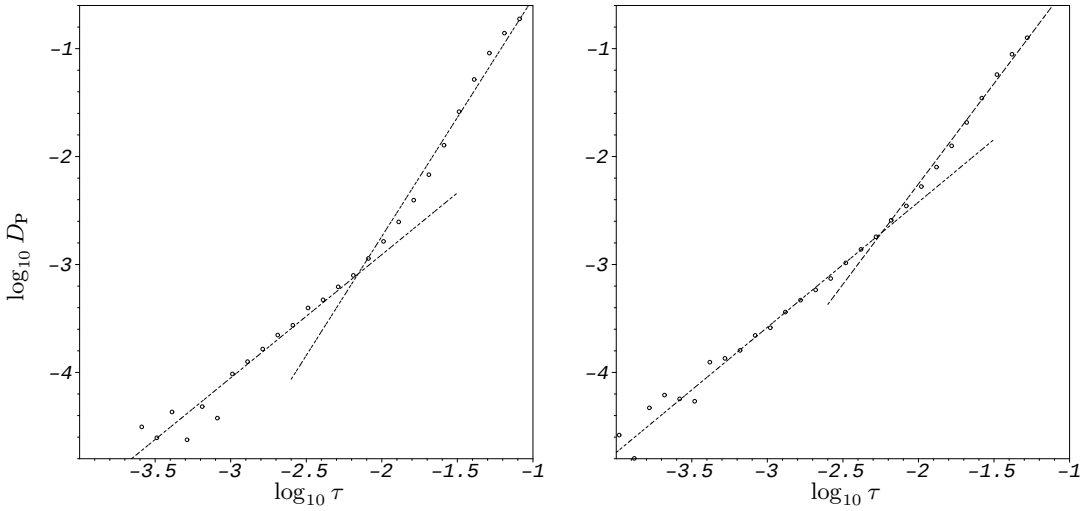

FIGURE 2.11. Left: D_{P} and fit lines for $\rho = 3.04$ and $N = 5 \times 10^6$. Right: The same for $\rho = 3.16$ with $N = 10^5$.

graph connecting the points changes significantly. We therefore separated the data into two sets of points belonging to $\log_{10} \tau$ above resp. below -2.25 and calculated the linear best fits at these two sets separately. The inclinations of the fit lines are approximately 1 for the latter and 1.2 for the former set (to the first two decimals). The right picture in figure 2.10 shows the namely data and fit lines for $\rho = 3.12$. It yields the same approximate gradients within the chosen accuracy.

One should note that the accuracy of our numerical experimentation on the order parameter D_{P} and the critical exponent β is quite weak compared to what physicists achieve in their experiments [**7**]: Values of order parameters are determined for τ well below 10^{-6}. This is not surprising if we recall that physical samples are much larger (10^{23}) than any N we could practically reach. On the other hand, it is often difficult in physics to determine exact values of the parameter τ, typically depending on the temperature.

Let us recall that the critical exponent β is the small-τ limit of $\ln D_{\mathrm{P}} / \ln \tau$. Our results entitle us to conjecture $\beta = 1$, *i.e.*, that the critical exponent for D_{P} at the stochastic bifurcation **B**→**S** is the same as that of the Lyapunov exponent in the noiseless case at the bifurcation point $\rho = 3$.

To explain heuristically why the slope of $\log_{10} D_{\mathrm{P}}$ increases near $\log_{10} \tau = -2.25$, we should take into account which structure of the invariant density the noise induced bifurcation generates: The single peak around x^* is replaced by a two-peaked shape when σ becomes small enough. The additional spread of the density which must forego this replacement can well result in an additional increase of D_{P}. Clearly this argument can only be valid in regions of σ where the density is not diverted too much from the simpler structure so that the implicit assumption on its shape made in (8) is not too false.

To bolster the conjecture $\beta = 1$ we also tried to determine D_{P} at $\rho = 3.04$ and 3.16. The results, corresponding to that of figure 2.10, are shown in figure 2.11. The change of inclination around $\log_{10} \tau = -2.25$ is even more noticeable in these two cases and the initial gradient is larger, namely about 2 (upper fit lines). The lowest values of D_{P} reached are about one order of magnitude

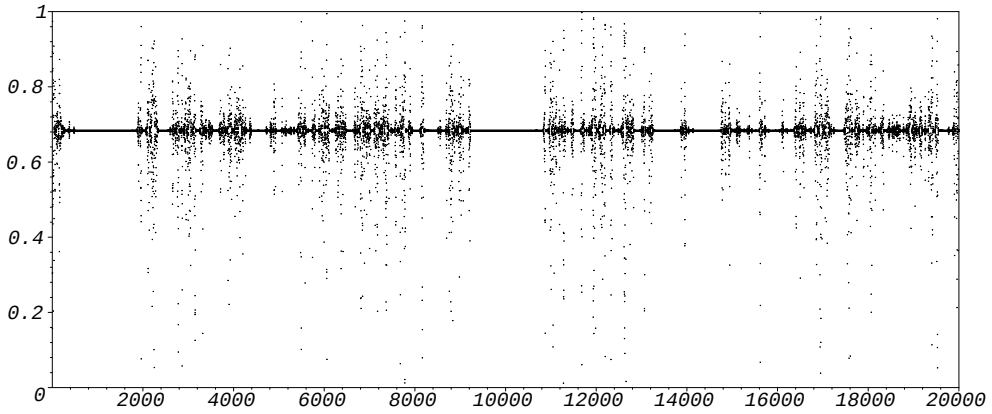

FIGURE 2.12. Example evolution at $(\rho, \sigma) = (3.16, 1.79)$.

smaller than for $\rho = 3.08$ and 3.12, making the determination of β more insecure. Nonetheless, the β-values derived from the lower fit lines in figure 2.11 are also close to 1, such that our conjecture is at least not disproven.

It is a bit surprising in our analysis so far that the case $\rho = 3.16$ apparently does not differ very much from the other cases considered, although it possesses the distinguishing feature that the system passes a small region of positive Lyapunov exponents before entering the stable phase. to get a grip on this issue, we first consider the phenomenology in that domain. In figure 2.12 we have shown an example evolution near the point where λ reaches its maximum value of about 0.03 (cf. figure 2.8). We have $\sigma_0(3.16) \approx 1.9079$, and thus $\tau \approx 0.06$ in this example and it coarsely corresponds to the upper right picture in figure 2.5. From a comparison of these two pictures it is apparent that the phenomenon of on-off intermittence, *i.e.*, of spikes of irregular behaviour alternating with periods of stability, is much more distinct (the irregularities spread over the whole interval) in the positive-λ domain. In an attempt to discriminate the transition **B**→**S** from the domain $\lambda > 0$ from the other case by recurrence time statistics, we consider what we will call the **reentrance time** at x^* of the system, cf. [**6**]:

$$\overline{T}_{\mathrm{R}}(r) \stackrel{\mathrm{def}}{=} \lim_{N\to\infty} \frac{N}{\#\{x_i \mid x_i \in B_r(x^*) \text{ and } x_{i-1} \notin B_r(x^*),\ 1 \le i \le N\}}.$$

This quantity should become infinite at all scales when approaching σ_0, but empirically it doesn't due to the same statistical convergence problems we encountered in the determination of $\overline{T}_{\mathrm{P}}$. We recourse to the same strategy as above and consider only the 'corrected' values $\overline{T}_{\mathrm{R}}/\overline{T}^0_{\mathrm{R}}$ which then approach 1 as $\tau \to 0$. We could proceed to define a reentrance dimension in analogy to D_{P}, but some experiments show that this quantity also does not enable us to discriminate the two cases in question. Instead, we are going to consider the linear offset $\overline{o}_{\mathrm{R}}$ in the regression $\log_{10} \overline{T}_{\mathrm{R}} = a \log_{10} r + \overline{o}_{\mathrm{R}}$ (a would be the reentrance dimension). It can be viewed as a sort of average of (the logarithm of) $\overline{T}_{\mathrm{R}}$ over the scales $\log_{10} r$.

The experimental results shown in figure 2.13 reveal that reentrance statistics is able to detect the intermittence phenomenon at $\rho = 3.16$. It is distiguished by a much higher $\overline{o}_{\mathrm{R}}$ and a steeper slope towards zero as $\tau \to 0$. Note

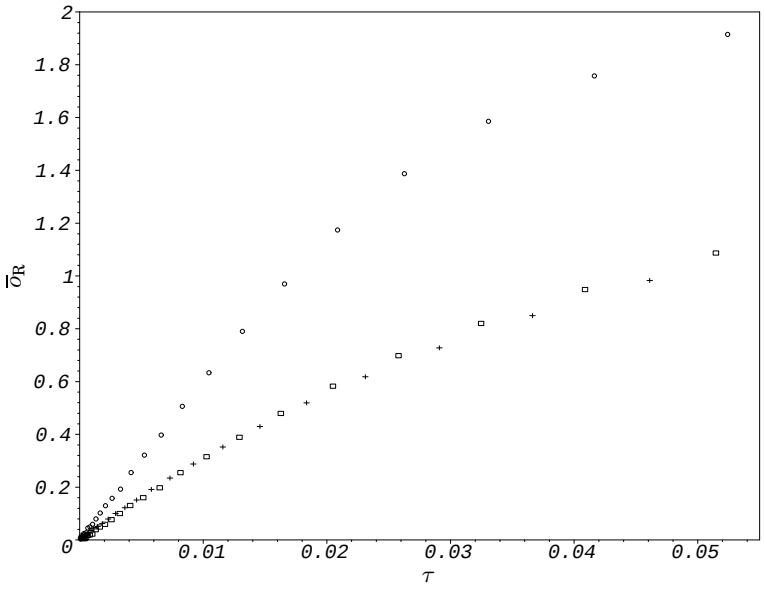

FIGURE 2.13. Reentrance offsets $\overline{\sigma}_{\mathrm{R}}$ for $\rho = 3.04$ (crosses), $\rho = 3.08$ (boxes), and $\rho = 3.16$ (circles). $N = 10^5$.

especially, how the cases $\rho = 3.04$ and 3.16 that were rather similar with respect to Poincarè recurrence statistics are now clearly separated by the runs of the respective $\overline{\sigma}_{\mathrm{R}}$-curves.

2.5. The Strongly Intermittent Phase I

We are not going to consider the phase denoted by **I** in section 2.2 in as much detail as the **B**→**S** transition in the last two sections. **I** is a region of strictly positive Lyapunov exponent and after the discussion of the last section we can guess that phenomenologically it is a phase where the system shows strong on-off intermittency. The single example shown in figure 2.14 is enough to convince us of that fact. The most remarkable fact about the transition from this phase into the stable domain is that it is much less sharply determined than the **B**→**S** transition. Figure 2.4 shows to the contrary, that the intermittence phenomenon seems to intrude into the stable phase. The transition to stability manifests itself in intermittent spikes appearing less frequently and with shorter duration (cf. figure 1.3). This could be analyzed by recurrence and especially reentrance time statistics in a more detailed examination.

Even more remarkable is the phenomenon shown in figure 2.15: At fixed $\sigma = 4.7$ we see that the system returns to the stable Lyapunov exponent at certain ρ-values between $\sigma = 3$ and 3.4. This occurs at about a third of all data points, but for higher σ the convergence to λ_{S} at these points is very slow. The appearance of λ_{S} is stable in σ, *i.e.*, takes place at the same values of ρ at different σ within **I**. One could speculate that these results indicate a sort of 'mixing of phases on a fractal set' but it is certainly worth further investigation and still lacks a theoretical explanation.

2.6. Numerics

The numerical experiments carried out to achieve the results described above were carried out on an IBM 9076 SP2 parallel computer utilizing 34 RS/6000 (POWER2) compute nodes. This system has a peak performance of

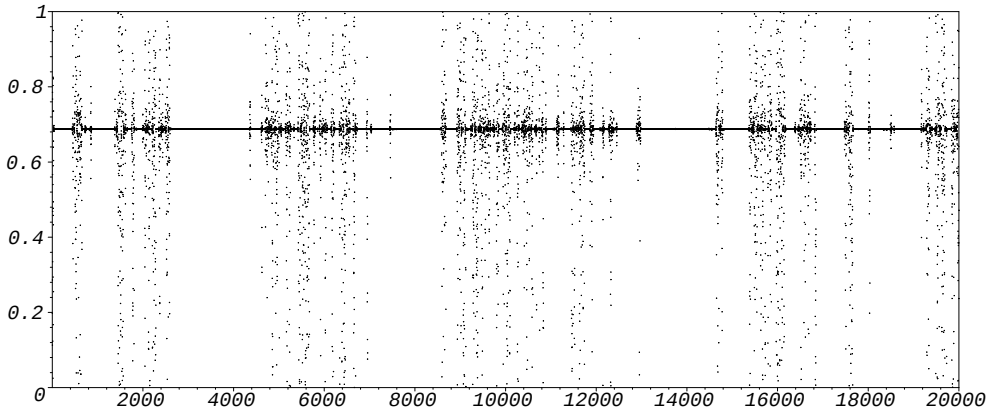

FIGURE 2.14. Example evolution at $(\rho, \sigma) = (3.2, 5.0)$.

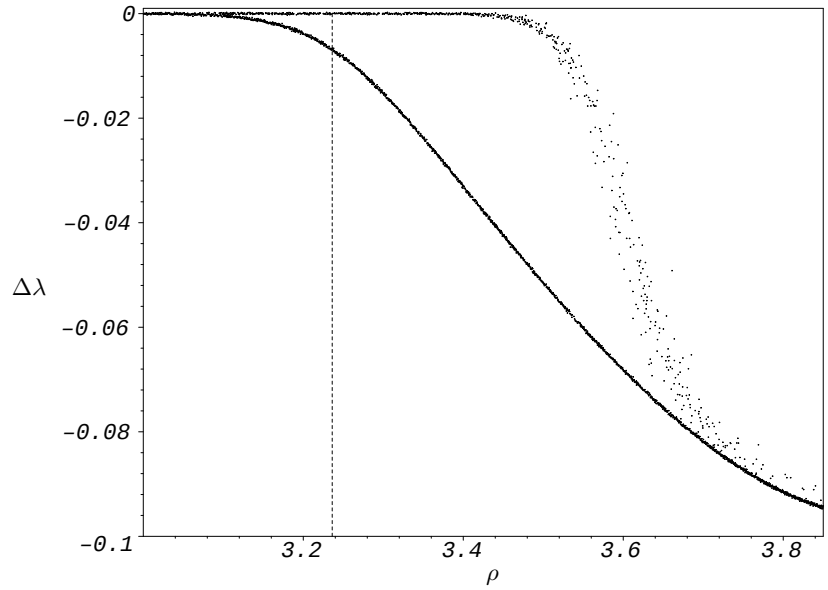

FIGURE 2.15. The Difference $\Delta\lambda$ for $\sigma = 4.7$, determined by 100×10^6 iterations for each point with resolution 2.5×10^{-4} in σ. The dashed line marks the σ-value where λ_S crosses 0.

about 10 GFLOPS. The programming language used was IBM's XL High Performance Fortran 1.4, with the benefit of an extended-precision floating point data type that yields a numerical precision better than 30 significant decimal digits in a single operation.

Visualization was done with the computer algebra system Maple.

Bibliography

1. L. Arnold, *Random Dynamical Systems*, Springer Verlag, Berlin, Heidelberg, New York, 1998. 13, 20
2. J. J. Binney, N. J. Dowrick, A. J. Fisher, and M. E. J. Newman, *The Theory of Critical Phenomena*, Oxford University Press, Oxford, New York, 1992. 21
3. J. P. Crutchfield and N. H. Packard, *Symbolic Dynamics of Noisy Chaos*, Physica D **7** (1983), 201–223. 14
4. Michael Dellnitz and Oliver Junge, *On the Approximation of Complicated Dynamical Behavior*, SIAM Journal on Numerical Analysis **36** (1999), no. 2, 491–515. 14
5. L. Fronzoni, R. Mannella, P. V. E. McClintock, and Frank Moss, *Postponement of Hopf Bifurcation by Multiplivative Coloured Noise*, Physical Reviews A **36** (1987), no. 2, 834–841. 19
6. J. B. Gao, *Recurrence Time Statistics for Chaotic Systems and Their Applications*, Phyical Review Letters **83** (1999), no. 16, 3178–3181. 23, 27
7. Charles E. Hecht, *Statistical Thermodynamics and Kinetic Theory*, W. H. Freeman and Company, New York, 1990. 7, 15, 21, 26
8. Andrzej Lasota and Michael C. Mackey, *Probabilistic Properties of Deterministic Systems*, Cambridge University Press, Cambridge, New York , New Rochelle, Melbourne, Sydney, 1985. 13, 14, 19
9. R. Duncan Lee and Howard Raiffa, *Games and Decisions*, Dover Publications Inc., New York, 1989, Reprint. Originally published: Wiley, New York, 1957. 5
10. V. Loreto, G. Paladin, M. Pasquini, and A. Vulpiani, *Characterization of Chaos in Random Maps*, Physica A **232** (1996), 189–200. 13, 16
11. V. Loreto, G. Paladin, and A. Vulpiani, *Concept of Complexity in Random Dynamical Systems*, Physical Reviews E **53** (1996), no. 3, 2087–2098. 13, 16
12. Richard S. Sutton and Andrew G. Barto, *Reinforcement Learning*, MIT press, Cambridge, Massachusetts, London, England, 1998. 5
13. Michel Zinsmeister, *Thermodynamic Formalism and Holomorphic Dynamical Systems*, SMF/AMS Texts and Monographs, vol. 2, American Mathematical Society, 2000. 19