Recurrences:
Exploiting Naturally Occurring Analogues

Dissertation zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften (Dr. rer. nat.)
in der Wissenschaftsdisziplin Nichtlineare Dynamik

Eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät
der Universität Potsdam

von
Marco Thiel

Potsdam, August 14, 2004

Meinen Eltern

Contents

Introduction

The paradigm of deterministic chaos has influenced thinking in many fields of science. New mathematical concepts, such as Lyapunov exponents, recurrences and fractal dimensions have been brought out. Nonlinear and especially chaotic systems show rich dynamical structures and sometimes provide an explanation for irregular fluctuations in "real life" systems which do not seem to be inherently stochastic.

Natural systems which are supposed to be (at least partially) chaotic are for example planetary systems, population dynamics and some turbulent flows. One of the most scorching problems is to analyse systems based on observed time series with respect to these new concepts.

Recurrence plots, a rather promising tool of data analysis, have been introduced by Eckman et al. in 1987. They visualise recurrences in phase space and give an overview about the system's dynamics. Two features have made the method rather popular. Firstly they are rather simple to compute and secondly they are putatively easy to interpret. However, the straightforward interpretation of recurrence plots for some systems yields rather surprising results. For example indications of low dimensional chaos have been reported for stock marked data, based on recurrence plots [Holyst et al., 2001].

In this work we exploit recurrences or "naturally occurring analogues" as they were termed by E. Lorenz [Lorenz, 1969], to obtain three key results. One of which is that the most striking structures which are found in recurrence plots are hinged to the correlation entropy and the correlation dimension of the underlying system (chapter 1) [Thiel et al., 2003]. Even though an eventual embedding changes the structures in recurrence plots considerably these dynamical invariants can be estimated independently of the special parameters used for the computation [Thiel et al. 2004a].

The second key result is that the attractor can be reconstructed from the recurrence plot [Thiel et al., 2004b]. This means that it contains all topological information of the system under question in the limit of long time series [Bandt et al., 2002].

The graphical representation of the recurrences can also help to develop new algorithms and exploit specific structures. This feature has helped to obtain the third key result of this study. Based on recurrences to points which have the

same "recurrence structure", it is possible to generate surrogates of the system which capture all relevant dynamical characteristics, such as entropies, dimensions and characteristic frequencies of the system. These so generated surrogates are shadowed by a trajectory of the system which starts at different initial conditions than the time series in question. They can be used then to test for complex synchronisation [Thiel et al., in prep.].
Applications are scattered all over the chapters. However, two especially extensive examples are presented in chapter 8. The first example is an analysis of the stability of extrasolar planetary systems. We use an automatic variant of the algorithm to estimate dynamical invariants from RPs, to determine stable regions in parameter space. The second application evaluates the predictability of annual earth's surface temperature fluctuations. We compare the results for a simulation (GCM-data) and for data derived from observations.

Chapter 1

Basic Principles

In this chapter some concepts which are needed in the course of this work will be introduced. The first concept is the one of phase space, an abstract mathematical space, ever point of which corresponds to a state of the system. A trajectory in this space describes the dynamics of the system under consideration.
Under suitable conditions, it can be shown that trajectories in phase space tend to recur to their pasts. This is the concept of recurrence, the second one introduced in this chapter.
The third concept is the one of generalised entropies. These entropies are introduced in the framework of a thermodynamical description of chaotic systems. They are linked to the system's predictability. All these concepts will be linked in later chapters of this work.

1.1 Phase Space

Consider a purely deterministic system (say a pendulum given by e.g. a set of differential equations). Once the initial conditions are fixed (e.g. position and momentum of the pendulum) the states at all future times are determined as well by the equations of motion. The space of all valid initial conditions is then called state space or phase space. This space has trivially the same dimension as the number of initial conditions needed to describe the system.
Starting at a point in this phase space, the state of the system at any time is given by a trajectory which lies in this abstract space. The system's evolution and its dynamics can be studied, by analysing its trajectory in phase space [Ott, 1993].
The concept of a state of the system is also powerful for nondeterministic cases. A large class of systems can be described by a (possibly infinite) set of states and some transition rules which determine how the system proceeds from one state to another. Markov processes are prominent members of this class. The transition rules are then given by probabilities and the future state is selected randomly according to these probabilities. In a Markov model of order m the

future depends only on the past m states. One can consider a purely deterministic system as a limiting case of a Markov process on a continuum of states. In this case, the transition to the next state given by the deterministic rule occurs with probability 1 and every other transition with probability 0.
Especially in the examples of measured time series we will understand phase space always in this more general meaning, as there is always some observational noise present in such cases.

1.2 Recurrences

Recurrences are a fundamental characteristic of many dynamical systems. The concept of recurrences has been introduced by Henri Poincaré in a seminal work from 1890 [Poincaré, 1890], which won a prize sponsored by King Oscar II of Sweden and Norway on the occasion of his 60th birthday. Therein, Poincare did not only discover the "homoclinic tangle" which lies at the root of the chaotic behaviour of orbits, but he also introduced the concept of recurrences. When speaking about the restricted three body problem he mentioned: "In this case, neglecting some exceptional trajectories, of which the occurrence is infinitely improbable, one can show, that the system recurs infinitely many times as close as one wishes to its initial state."
In a larger context recurrences make part of one of three broad classes of asymptotic invariants: (i) growth of the number of orbits of various kinds and of the complexity of orbit families [1], (ii) types of recurrences, and (iii) asymptotic distribution and statistical behaviour of orbits [Katok & Hasselblatt, 1995]. The first two classes are of purely topological nature. The last class is naturally related to ergodic theory. We will actually deal, among other things, with generalised entropies, Poincaré recurrences and the measure of attractors.
From the different types of recurrences which make part of the second class of invariants, the Poincaré recurrence is of particular interest for this work. The Poincare Recurrence Theorem ([Katok & Hasselblatt, 1995] Theorem 4.1.19) states:
Let T be a measure-preserving transformation of a probability space (X, μ) and let $A \subset X$ be a measurable set. Then for any $N \in \mathcal{N}$

$$\mu(\{x \in A | \{T^n(x)\}_{n \geq N} \subset X \backslash A\}) = 0. \tag{1.1}$$

Here we give the rather short proof of this theorem:
Replacing in Eq. (1.1) T by T^N one sees that it is enough to prove the statement for $N = 1$. The set

$$\tilde{A} := \left\{ x \in A | \{T^n(x)\}_{n \in (N)} \subset X \backslash A = A \cap \left(\bigcap_{n=1}^{\infty} T^{-n}(X \backslash A) \right) \right\}$$

[1] An important invariant of the orbit growth is the topological entropy (see Sec. 1.3).

is measurable. $T^{-n}(\tilde{A}) \cap \tilde{A} = \emptyset$ for every n and hence

$$T^{-n}(\tilde{A}) \cap T^{-m}(\tilde{A}) = \emptyset$$

for all $m, n \in \mathcal{N}$. $\mu(T^{-n}(\tilde{A})) = \mu(\tilde{A}) = 0$ since T preserves μ. Thus $\mu(\tilde{A})$ since $1 = \mu(X) \geq \mu\left(\bigcup_{n=0}^{\infty} T^{-n}(\tilde{A})\right) = \sum_{n=0}^{\infty} \mu\left(T^{-n}(\tilde{A})\right) = \sum_{n=0}^{\infty} \mu(\tilde{A})$.
The Poincaré Recurrence Theorem states that if one has a measure preserving transformation, the trajectory will eventually come back to the neighbourhood of any former point with probability one.
However, the theorem does not tell how long it takes to recur. Especially for high dimensional complex systems the recurrence time might be extremely long. For the earth's atmosphere the recurrence time has been estimated to be about 10^{30} years, which is more than the time the universe exists so far [van den Dool, 1994]. The recurrence time is related to the fractal dimensions of the system [Gao, 1999], which are defined in the next section.

1.3 Information Theory - Rényi Entropies

Entropies are known not only from chaos theory but also from different branches of science such as thermodynamics or information theory. The most prominent examples of entropies are the Boltzmann and Shannon entropy, which are linked to a certain information (or lack of information) one has about the state of a system [Beck & Schlögl, 1993].
To define the Shannon entropy and Shannon information we first introduce the "bit-number". Consider a bit-storage unit of a computer as a switch which has two different positions. If the storage consists of A such switches it can take on

$$N = 2^A \tag{1.2}$$

different states and thus give N different bit patterns. Such a storage could then be used to select a particular one of N numbers. This can obviously be done by a binary number of length A. Eq. (1.2) says that we need

$$A = \frac{\log_2 N}{\log_2 2} = \log_2 N$$

bits to select one of these numbers or events. We now define the bit number by

$$b = \log_2 N. \tag{1.3}$$

Let p_i be the probability of an event i of a sample set of R disjoint events $i = 1, 2, \ldots, R$. We call the set of all p_i, which in general may take on different values, a probability distribution, or simply a distribution, and we shall denote it here

by p.
For the sake of the argument we start with a large sample set of N elementary events of equal probability. This set is divided into subsets $i = 1, 2, \ldots, R$ which have no elements in common. Each subset is a composition of N_i elementary events. The probability to find an elementary event in the i-th subset is then given by the p_i and is the probability of the compound event i. The minimum bit number to select one elementary event out of the large sample set is $\log_2 N$. It has to be independent of whether we select first the subset i in which the elementary event lies and then select select it out of the subset, or whether we select it directly from the large sample set. If the bit number needed for selecting the subset i is denoted by b_i, this statement reads

$$b_i + \log_2 N_i = \log_2 N.$$

Hence, we obtain the result that

$$b_i = -\log_2 p_i$$

is the bit-number an observer is missing in order to know whether the event i will occur, if he only knows the probability p_i of this event. Since, in a long series of observations each event i occurs with the relative frequency p_i the mean value of $-b_i$ is

$$I(p) = \sum_{i=1}^{R} p_i \log_2 p_i. \tag{1.4}$$

This information measure is called Shannon Information. As $0 \leq p_i \leq 1$, $I(p)$ is always non-positive (as it is a lack of information), with its maximum value zero belonging to optimum knowledge. Its negative

$$S(p) = -I(p) \tag{1.5}$$

is called Shannon entropy. It measures the lack of knowledge an observer has about the question of which event of the sample set is to be expected, if he only knows the distribution p.
To demonstrate the distinguished role of the Shannon information, it is worthwhile to note that it can be uniquely derived from reasonable requirements on the properties of an information measure. These are the following axioms by Khinchin [Khinchin, 1957]:

1. $I(p)$ depends only on the probabilities p_i:

$$I(p) = I\left(p_1, \ldots, p_R\right). \tag{1.6}$$

2. The second axiom,

$$I\left(\frac{1}{R}, \ldots, \frac{1}{R}\right) \leq I(p), \tag{1.7}$$

means that $I(p)$ has to take a minimum value for the uniform distribution ($p_i = \frac{1}{R}\ \forall i$).

3. The next axiom,
$$I(p_i, \ldots, p_R) = I(p_i, \ldots, p_R, 0), \tag{1.8}$$
means that $I(p)$ remains unchanged if the sample set is enlarged by a new event with probability zero.

4. To formulate the last axiom, we consider the composition of two subsystems Σ^I and Σ^{II} to one system Σ. We denote the probability distributions of the two subsystems p_i and p_j. If the two subsystems are independent the probability distribution of the compound system factorises with respect to the subsystems
$$p_{i,j} = p_i^I \cdot p_j^{II}$$
Then, $I(p)$ should become additive with respect to the subsystems:
$$I(p) = I\left(p^I\right) + I\left(p^{II}\right)$$
In general, the two subsystems will be correlated and therefore only
$$p_{i,j} = Q(j|i)\, p_i^I$$
will hold. Here $Q(j|i)$ is the conditional probability that the subsystem Σ^{II} is in state j if the subsystem Σ^I is in state i. Then the axiom requires
$$I(p) = I\left(p^I\right) + \sum_i I(Q|i), \tag{1.9}$$
where
$$I(Q|i) = \sum_j Q(j|i) \ln Q(j|i)$$
is the conditional information of the probability distribution $Q(j|i)$ of the events j for fixed i. The meaning of this axiom is that the information has to be independent of how it is collected, whether directly for the compound system or successively for the subsystems.

It can be shown that these four axioms uniquely determine the Shannon information. It is noteworthy that the distinctive property of the Shannon information, namely that it becomes additive for independent subsystems, does not occur as one of the Khinchin axioms. Indeed, if we replaced Khinchin axiom (1.9) by this requirement, we would not obtain the Shannon information in a unique way. It is easy to see that then another information measure, the Rényi information [Rényi, 1970],
$$I_\beta(p) = \frac{1}{\beta - 1} \ln \sum_i (p_i)^\beta \tag{1.10}$$
fulfils the changed axioms as well.
Let us now consider a dynamical system and its trajectory in phase space. We

introduce a partitioning of the phase space $\{A_i\}$. The probability p_i can be interpreted as the probability of the trajectory to be in box A_i. We define $p(i_0, \ldots, i_{N-1})$ as the probability of the occurrence of a sequence of points in the boxes indexed $i_0, \ldots, i_{N-1}$. For finite N there is only a finite number of possible sequences $\omega(N)$. Hence, we can define

$$p_j^{(N)} = p(i_0, \ldots, i_{N-1})$$

using a single index only, with $j = 1, \ldots, \omega(N)$.
Analogously to Eq. (1.5) one can then proceed to the negative dynamical Rényi information

$$H_\beta(p) = -I_\beta(p),$$

and then consider the limit

$$h_\beta = \lim_{N\to\infty} \frac{H_\beta}{N} = \lim_{N\to\infty} \frac{1}{N} \frac{1}{1-\beta} \ln Z_N(\beta),$$

where

$$Z_N(\beta) = \sum_j^\omega (p_j^{(N)})^\beta$$

is the dynamical partition function. The Rényi entropies of order β are then defined as

$$K_\beta = \sup_{\{A\}} h_\beta, \tag{1.11}$$

where the supremum is taken over all partitions. If the partition is a generating one the supremum is already reached. Then we can write

$$K_\beta = \lim_{N\to\infty} \frac{1}{1-\beta} \frac{1}{N} \log \sum_{j=1}^\omega (p_j^{(N)})^\beta,$$

where the logarithm is computed to the base 2 if the entropy is given in bits per iteration and to base e if it is given in nats per iteration.
If no generating partition is known, one usually uses a grid of R boxes of equal size ε and performs the limit $\varepsilon \to 0$ after the limit $N \to \infty$ as been performed. That is, for each R we consider the symbol sequence probabilities $p(i_0, \ldots, i_{N-1}) = p_j^{(N)}$ for the sequence j and define

$$K_\beta = \lim_{\varepsilon\to 0} \lim_{N\to\infty} \frac{1}{1-\beta} \frac{1}{N} \log \sum_{j=1}^\omega (p_j^{(N)})^\beta. \tag{1.12}$$

Note, that the limit $\varepsilon \to 0$ is just taken to avoid possible finite size effects of the partition. The important limit is the limit $N \to \infty$. Two prominent members of the Rényi entropies are the topological entropy K_0 and the Kolmogorov-Sinai

entropy K_1.
The difference between Eq. (1.11) and Eq. (1.12) is similar to the difference between the capacity and the Hausdorff dimension. From a rigorous mathematical point of view, definition Eq. (1.11) is necessary. For the analysis of time series, however, Eq. (1.12) is more appropriate. In most cases of physical relevance, we expect the two definitions to coincide.
The Rényi informations I_β (Eq. (1.10)) also allows to define the Rényi dimensions by

$$D_\beta = \lim_{\varepsilon \to 0} \frac{I_\beta}{\log \varepsilon} = \lim_{\varepsilon \to 0} \frac{1}{\log \varepsilon} \frac{1}{\beta - 1} \log \sum_{i=1}^{R} p_i^\beta. \tag{1.13}$$

Even though the Rényi informations can diverge in the limit $\varepsilon \to 0$, the corresponding dimensions usually remain finite.
In the course of this work we will find a close connection between the Rényi entropy of second order K_2 and the lines in recurrence plots. Therefore, we concentrate in the following on K_2. Fig. 1.1 illustrates the procedure to compute K_2 based on Eq. (1.12) for two prototypical cases. The left panel represents a

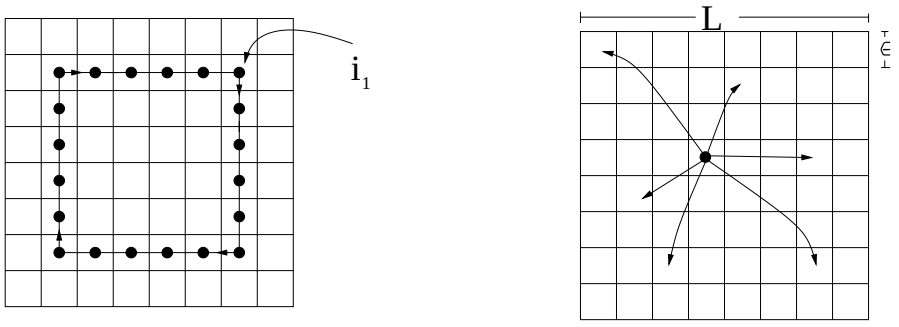

Figure 1.1: **Left panel:** periodic trajectory in a partitioned phase space. **Right panel:** Same for a "random trajectory".

periodic trajectory in its phase space. The sides of each box of the partitioning have fixed length ε. In this discrete system the number of visited boxes does not depend on ε. First, we evaluate K_2 for one step in the future. Then Eq. (1.12) reads

$$K_2(\varepsilon, l = 1) = -\frac{1}{1} \log_2 \sum_{j=1}^{\omega} (p_j)^2.$$

We start at the point marked by i_1. As the trajectory is periodic, there is only one way to move one step forward. This is the only possible future, i.e. it will be realized with probability one. Hence, we find

$$K_2(\varepsilon, l = 1) = -\log_2 \cdot 1^2 = 0.$$

Analogously, one can ask for the number of realized trajectories of length two. As there is only one way to move two steps ahead, also $K(\varepsilon, l = 2)$ is zero. Indeed, K_2 is equal to zero for all trajectories starting at i_1 and for arbitrary length. So, we finally get

$$K_2(\varepsilon, l \to \infty) = 0$$

for periodic systems. The right panel of Fig. 1.1 illustrates a (uniformly distributed) random system. There are $(L/\varepsilon)^2$ boxes in the phase space. The probability to jump to one box is then $p_i = (L/\varepsilon)^{-2} \quad \forall i$. The number of possible future trajectories increases with $(L/\varepsilon)^{2l}$. Hence, one finds going $l \to \infty$ steps ahead

$$\begin{aligned} K_2(\varepsilon, l \to \infty) &= -\lim_{\varepsilon \to 0} \lim_{l \to \infty} \frac{1}{l} \log_2 \sum_{j=1}^{(L/\varepsilon)^{2l}} \left((L/\varepsilon)^{-2l} \right)^2 \\ &= -\lim_{\varepsilon \to 0} \lim_{l \to \infty} \frac{1}{l} \log_2 (L/\varepsilon)^{2l} \cdot (L/\varepsilon)^{-4l} \\ &= \lim_{\varepsilon \to 0} \lim_{l \to \infty} \frac{1}{l} 2l \log_2 \cdot (L/\varepsilon) = \infty. \end{aligned}$$

The divergence of K_2 is typical for stochastic systems. K_2 quantifies the increase of possible future evolutions for increasing time. If the number of possible evolutions increases too slow as in the case of the periodic system, the entropy vanishes. If the number increases too fast, as in the case of the stochastic system, K_2 is infinite. For chaotic systems, one finds non zero but finite values of K_2. Apart from K_2 the so called correlation dimension D_2 will prove to be essential for the description of structures in recurrence plots (Chap. 3).

Chapter 2

Recurrence Plots

2.1 Recurrence Plots and Cross Recurrence Plots

Recurrence plots (RPs) were introduced to simply visualise the behaviour of trajectories in phase space [Eckmann et al., 1987]. Suppose we have a dynamical system represented by the trajectory $\{\vec{x}_i\}$ for $i = 1, \ldots, N$ in a d-dimensional phase space. Then we compute the matrix

$$\mathbf{R}_{i,j} = \Theta(\varepsilon - |\vec{x}_i - \vec{x}_j|), \quad i, j = 1 \ldots N, \tag{2.1}$$

where ε is a predefined threshold and $\Theta(\cdot)$ is the Heaviside function. The norm used in Eq. (2.1) is (in principle) arbitrary. The graphical representation of $\mathbf{R}_{i,j}$ called recurrence plot is obtained encoding the value one as "black" and zero as "white" point.

Fig. 2.1 displays the RPs for three prototypical systems: (a) shows the RP of a sine function. The main diagonal line ($i = j$) is by definition Eq. (2.1) for all systems black. A sine function is topologically a circle in phase space. Imagine that we compare two points which have in time a distance of one period. They trivially coincide. This leads to a line which is parallel to the main diagonal. One might interpret this line as follows: after one period one has two points in phase space which evolve in the same way (to compare the evolution one has to move diagonally in the plot!).

Fig. 2.1 (c) displays the plot for uniform and independent noise. Now, there are no dominant diagonal structures, because even though two points in phase space are neighbouring, their evolution is unpredictable and they will evolve in different ways.

Fig. 2.1 (b) shows the RP for the Rössler system in a chaotic regime. There are some diagonal lines which are nevertheless shorter than in the case of the sine function. Even though two points are neighbouring in phase space, their future developments are different due to the exponential divergence of errors in chaotic systems. It seems obvious that the line lengths are linked to the predictability

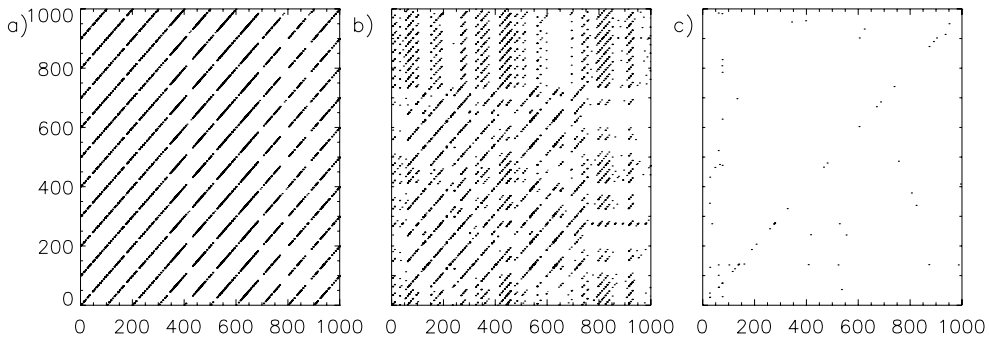

Figure 2.1: (a) Recurrence Plot of a sine function. (b) Recurrence Plot of the Rössler System. (c) Recurrence plot of uniform independent noise.

of the system. Note, that in the upper right of the RP there is a small window with parallel lines, similar to the RP of the sine function. This window indicates an unstable periodic orbit (UPO). UPOs are the "backbone" of chaotic attractors [Ott, 1993], and it is highly desirable to extract them from observed time series [So et al., 1996]. The development of a method for the automatic detection of UPO from recurrence plot is in progress.
The concept of recurrence plots can be extended for the analysis of two systems. One possibility is the definition of a Cross Recurrence Plot [Marwan, 1999] by the matrix

$$\mathbf{CR}_{i,j} = \Theta(\varepsilon - |\vec{x_i} - \vec{y_j}|), \quad i,j = 1 \ldots N. \tag{2.2}$$

However, this definition has some drawbacks. If for example the dimensions of the phase spaces of $\vec{x_i}$ and $\vec{y_i}$ are different, it is impossible to subtract both vectors. From a physical point of view it is impossible if the components of the vectors $\vec{x_i}$ and $\vec{y_i}$ measure different physical quantities, i.e. velocity $[x_1] = m/s$ and temperature $[x_1] = K$. To overcome these problems Joint Recurrence Plots have been introduced [Romano et al., 2004]:

$$\mathbf{JR}_{i,j}^{\vec{x},\vec{y}} = \begin{cases} 1, & \text{if} \quad |\vec{x_i} - \vec{x_j}| < \varepsilon^{\vec{x}} \quad \text{and} \quad |\vec{y_i} - \vec{y_j}| < \varepsilon^{\vec{y}} \\ 0, & \text{else.} \end{cases} \tag{2.3}$$

The JR considers the joint (simultaneous) recurrences of the subsystems as a recurrence of the whole system. This seems to be reasonable for the analysis of two systems. This definition overcomes the above mentioned problems and has further advantages [Romano et al., 2004].
But before moving on, the "standard" quantification of recurrence plots will be elaborated in the next section.

2.2 Recurrence Quantification Analysis

The Recurrence Quantification Analysis was introduced by Zbilut and Webber in [Zbilut & Webber, 1992], [Webber & Zbilut, 1994] and [Zbilut et al., 1998a] with the aim of quantifying the structures found in RPs and hence, go beyond the purely visual classification. Some of the most important recurrence quantification measures are

- **Recurrence Rate** (RR): defined as the percentage of black points in the RP, i.e.

$$RR = \frac{1}{N^2} \sum_{i,j=1}^{N} \Theta(\varepsilon - |\vec{x}_i - \vec{x}_j|). \tag{2.4}$$

  Note, that the definition of RR coincide with the definition of the correlation sum [Grassberger & Procaccia, 1983a].

- **Determinism** (DET): defined as the the percentage of black points belonging to a diagonal line of at least length l_{min} [Marwan, 2003],

$$DET = \frac{\sum_{l=l_{min}}^{N} lP(l)}{\sum_{l=1}^{N} lP(l)}, \tag{2.5}$$

  where $P(l)$ denotes the probability to find a diagonal line of length l in the RP. This measure was introduced to quantify how predictable a system is. For a periodic system, $DET = 1$ and for a purely stochastic system, $DET \to 0$. However, the results depend on the choice of l_{min}, and there exists some ambiguity by choosing this parameter [Marwan, 2003].

- **Divergence** (DIV): defined by

$$DIV = \frac{1}{L_{max}}, \tag{2.6}$$

  where L_{max} is the length of the longest diagonal found in the RP. This measure has been used also to estimate the largest Lyapunov exponent [Trulla et al., 1996].

- **Entropy** ($ENTR$): is the Shannon entropy of the frequency distribution of diagonal lines in the plot,

$$ENTR = -\sum_{l=l_{min}}^{N} p(l) \ln p(l), \tag{2.7}$$

  where $p(l) = P(l)/\sum_{l=l_{min}}^{N} P(l)$. This measure aims to quantify the complexity of the deterministic structure in the system [Marwan, 2003].

The RQA measures introduced up to now can be also computed for each diagonal line parallel to the main diagonal one. Hence, the RQA measures can be obtained in dependence on the distance to the main diagonal [Marwan & Kurths, 2002].

- **Trend** ($TREND$): is a linear regression coefficient over the recurrence rate on each diagonal line parallel to the main diagonal,

$$TREND = \frac{\sum_{i=1}^{\tilde{N}}(i - \tilde{N}/2)(RR_i - < RR_i >)}{\sum_{i=1}^{\tilde{N}}(i - \tilde{N})^2}, \tag{2.8}$$

  where $\tilde{N} < N$, to get rid of the finite size effects [Marwan, 2003]. This measure quantifies the non-stationarity of the system.

- **Ratio** ($RATIO$): is the ration between DET and RR,

$$RATIO = \frac{DET}{RR}. \tag{2.9}$$

  This measure was introduced to determine some transitions between different physiological states, where the RR changes, but not the DET [Webber & Zbilut, 1994].

The measures mentioned before are based on the distribution of diagonal lines. Marwan and Kurths introduced three new measures in [Marwan & Kurths, 2002], that quantify vertical structures (resp. horizontal, because of the symmetry). These are the following:

- **Laminarity** (LAM): analogously to the DET, it is defined as the percentage of black points, that belong to a vertical line of at least length l_{min}:

$$LAM = \frac{\sum_{l=l_{min}}^{N} lP_v(l)}{\sum_{l=1}^{N} lP_v(l)}, \tag{2.10}$$

  where $P_v(l)$ denotes the probability to find a vertical line of length l in the RP. LAM quantifies the occurrence of laminar states in the given trajectory.

- **Trapping Time** (TT): is the mean length of vertical lines

$$TT = \frac{\sum_{l=l_{min}}^{N} lP_v(l)}{\sum_{l=l_{min}}^{N} P_v(l)}, \tag{2.11}$$

  and measures the mean time that the system sticks to a certain state (how long the trajectory will be trapped).

- **Maximal vertical length** (V_{max}): is analogous to the longest diagonal line in the RP.

These last measures based on the distribution of vertical lines, allow to identify chaos-chaos transitions, as well as intermittency [Marwan et al., 2002]. Furthermore, they are rather robust against noise.
These RQA measures quantify efficiently the structures that different dynamical systems exhibit in their RPs. Not only when discussing the structures in RPs in the last section but also considering the RQA measures introduced in this section, one notes that the distribution of diagonal lines $P(l)$ is a crucial point for the analysis of RPs. In Chap. 3 we relate the distribution of diagonals to dynamical invariants. But before we study how observational noise influences the RQA measures and the distribution of the diagonals.

2.3 Influence of Observational Noise

The influence of noise on the results obtained from the RQA (Sec. 2.2) is an important point for the analysis of "real life" systems, which are always corrupted by at least some amount of measurement error. In principle noise can have two effects on the structures in RPs. It can produce spurious (mainly single) black points or it can break up black diagonals by separating two points which really are neighbours in phase space [Thiel et al., 2002]. We now study the influence of these perturbations on the RQA.

2.3.1 Influence of Noise on the RQA

It is crucial to consider randomness as the method of RPs is frequently applied to experimental sets of data [Webber & Zbilut, 1994], [Marwan et al., 2001], [Marwan & Kurths, 2002]. One crucial point for many of the measures used to quantify the structures in RPs, such as DET Eq. (2.5) is the distribution of the lengths of the diagonal lines $P(l)$ of length l that are found in the plot [Marwan, 1999].
The outline of our proceeding is as follows: first we calculate analytically the distribution of the diagonal lines for basic stochastic processes (Gaussian and uniformly distributed noise). Then we estimate the influence of observational noise on the distribution of diagonal lines for an arbitrary underlying process and develop a criterion to choose ε optimally.

2.3.2 Analytical Calculation of the Distribution of Lengths of Diagonals for Gaussian White Noise

We start by calculating the distribution of diagonals $P(l)$ of length l for a CRP obtained from the two time series $\{x_i\}, \{y_i\}$ without embedding. Suppose that each one is a realization of independent Gaussian noise $\mathcal{WN}(0, \sigma^2)$. The proba-

bility to find a point in the interval $[x, x+\Delta x]$ is given by

$$p(x+\Delta x) = \int_{x}^{x+\Delta x} \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-y^2}{2\sigma^2}} dy.$$

To find a recurrence point at the coordinates (i, j) in the CRP, the condition $|x_i - y_j| < \varepsilon$ must hold. Hence, the probability for the occurrence of a recurrence point is

$$P_{\mathrm{RP}} = \frac{1}{\left(\sqrt{2\pi}\sigma\right)^2} \int_{-\infty}^{\infty} e^{\frac{-y^2}{2\sigma^2}} \int_{y-\varepsilon}^{y+\varepsilon} e^{\frac{-x^2}{2\sigma^2}} dx\, dy = \mathrm{erf}\left(\frac{\varepsilon}{2\sigma}\right). \tag{2.12}$$

For example, for a ratio of $\frac{\varepsilon}{\sigma} = 0.1$ we obtain a probability to find a recurrence point of $P_{RP} = 0.056$.
To calculate the distribution of diagonal lengths, we compute first the probability to find l recurrence points in a row. This probability is - as the time series are independent noise - the product

$$P_{\text{l-RP's}}(l) = P_{\mathrm{RP}}^l.$$

To assure that the line has exactly the length l, we have to claim that the point before the first recurrence point and the point after the last recurrence point are white, i.e. both are not recurrence points. The probability to find such a white point is $(1 - P_{\mathrm{RP}})$. Thus, the probability to find a line of (exactly) length l is given by

$$P(l) = P_{\mathrm{RP}}^l \cdot (1 - P_{\mathrm{RP}})^2 \tag{2.13}$$

This allows to compute the absolute number $N(l)$ of diagonal lines of lengths l in a CRP or RP of total length L using

$$N(l) = L^2 P_{\mathrm{RP}}^l \cdot (1 - P_{\mathrm{RP}})^2 \tag{2.14}$$

for a CRP and

$$N(l) = (L^2 - L) P_{\mathrm{RP}}^l \cdot (1 - P_{\mathrm{RP}})^2 \tag{2.15}$$

for a RP. Numerically, one does not count points on the main diagonal. We take this fact into consideration by subtracting L from the total amount of points in the RP. Simulations confirm Eq. (2.15).

2.3.3 Recurrence Plot of a Function with Observational Noise

In this section we discuss the influence of Gaussian, independent, observational noise on the recurrence plot of a time series. We suppose that we have a given

time series x_i. We further assume that the time series is corrupted by some independent Gaussian observational noise η_i

$$y_i = x_i + \eta_i \tag{2.16}$$

where $\eta_i \sim \mathcal{WN}(0, \sigma^2)$. Then the recurrence matrix $R_{i,j}$ of the underlying time series x_i is

$$R_{i,j} = \Theta\left(\varepsilon - |D_{i,j}|\right) \tag{2.17}$$

with $D_{i,j} = x_i - x_j$. Due to the observational noise that enters the measured time series y_i, we get instead of $R_{i,j}$

$$\tilde{R}_{i,j} = \Theta\left(\varepsilon - \left|\tilde{D}_{i,j}\right|\right) \tag{2.18}$$

with $\tilde{D}_{i,j} = y_i - y_j = x_i - x_j + \eta_i - \eta_j$. The question we address is how this observational noise changes the structures that are found in $\tilde{R}_{i,j}$ with respect to $R_{i,j}$. Therefore, we compute the probability $P_{i,j}$ to find a recurrence point at the coordinates (i, j) if the underlying time series x_i has a distance $D_{i,j}$ at these coordinates.
Given a fixed underlying time series x_i, we measure the same process with different realizations of additive noise $\mathcal{WN}(0, \sigma^2)$. Then we obtain for each time index i an ensemble of measurements $\{y_n\}_i$ that is Gaussian distributed with mean x_i and variance σ^2. As we are interested in the distances $D_{i,j}$, we can assume without of loss of generality that $x_i = 0$ and $x_j = -D_{i,j}$. Then the probability to find a recurrence point at the coordinates i, j is given by

$$\begin{aligned} P_{i,j} &= \frac{1}{\left(\sqrt{2\pi}\sigma\right)^2} \int_{-\infty}^{\infty} e^{-\frac{\eta_i^2}{2\sigma^2}} \int_{\eta_i-\varepsilon}^{\eta_i+\varepsilon} e^{-\frac{(\eta_j - D_{i,j})^2}{2\sigma^2}} d\eta_j \, d\eta_i \\ &= \frac{1}{8}\Big\{ \operatorname{erfc}^2\left(\tfrac{D_{i,j}-\varepsilon}{2\sigma}\right) - \operatorname{erfc}^2\left(-\tfrac{D_{i,j}-\varepsilon}{2\sigma}\right) + \operatorname{erfc}^2\left(-\tfrac{D_{i,j}+\varepsilon}{2\sigma}\right) - \operatorname{erfc}^2\left(\tfrac{D_{i,j}+\varepsilon}{2\sigma}\right)\Big\} \end{aligned} \tag{2.19}$$

where $\operatorname{erfc}(\cdot) = 1 - \operatorname{erf}(\cdot)$. Eq. (2.19) is one of the key results of this section and maps the distance matrix $D_{i,j}$ to a probability matrix $P_{i,j}$ that gives the probability to find a recurrence point in the RP of $y_i = x_i + \eta_i$ at the coordinates i, j for a fixed underlying process and an ensemble of realizations of the observational noise.
The matrix $P_{i,j}$ takes values of the interval $[0, 1]$ in contrast to the usual RP, that maps the distance matrix $D_{i,j}$ to the "binary" matrix $R_{i,j}$, i.e. only two symbols separate large distances (0, if $|x_i - x_j| > \varepsilon$) from small ones (1, if $|x_i - x_j| < \varepsilon$). From now on we will write P and D instead of $P_{i,j}$ and $D_{i,j}$ for convenience. In Fig. 2.2 (a) we represent P of Eq. (2.19) in dependence on $d = \frac{D}{\varepsilon}$ and $s = \frac{\sigma}{\varepsilon}$. We show only the part with $d > 0$ as the picture is symmetric for negative d's. An equivalent representation is given in Fig. 2.2 (b): it shows sections parallel

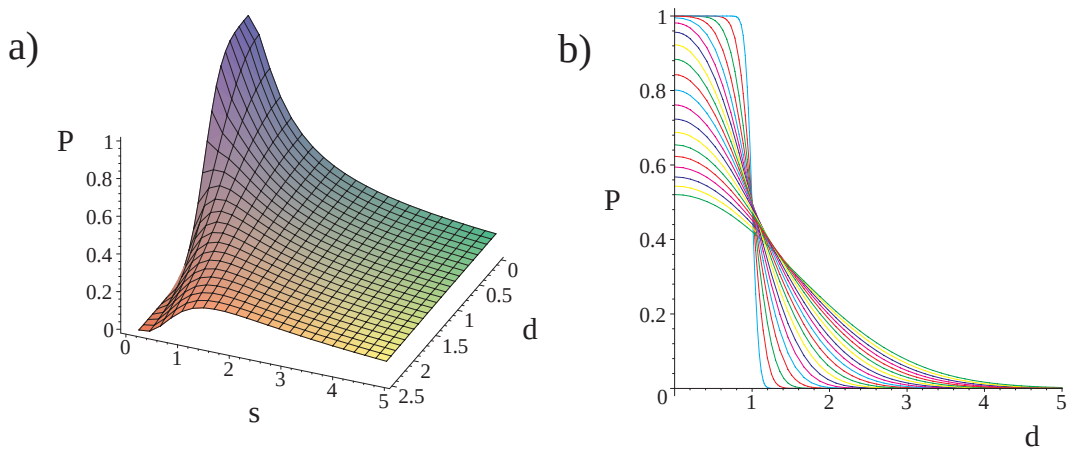

Figure 2.2: (a) dependence of P on $d = \frac{D}{\varepsilon}$ and $s = \frac{\sigma}{\varepsilon}$ for Eq. (2.19). (b) dependence of P on d for different $s \in [0.05, 1]$ in steps of 0.05.

to the "d"-axis for different s, where s takes values from 0.05 to 1 in steps of 0.05. The smaller the values of s the closer comes P to a Heaviside function. In the presence of noise the probabilities differ from this ideal. That means that due to the noise points that are recurrence points for the series x_i , will have a probability lower than one to be properly recognised. Similarly, points that are white in the absence of noise, will be properly recognised with a probability lower than 1. Hence, in the probability plot for systems with observational noise there will be neither black nor white points but only grey points (the main diagonal is not considered).
The striking point is that already a small level of noise will reduce the reliability of the calculation of the distribution of the lengths of diagonals. For a noise level of about 4% the probability to find a black point near the distance $d = 1$ is reduced from 1 to less than 0.6 (see Eq. (2.19)). This means that more than 40% (!) of the recurrence points are not recognised. Hence, the distribution of the diagonals will change enormously. This has crucial consequences for the application of RQA to observed data.

There are three ways to overcome this problem:
i) We can measure the data with very little observational noise. However, this is usually impossible, because the noise often cannot be controlled.
ii) We can develop news statistics. The approach presented in this chapter indicates how to do so.
iii) We can choose an optimal ε, so that these misclassifications will be reduced

as much as possible. Like this, we will develop in Sec. 2.3.4 a criterion based on Eq. (2.19) to calculate such an ε_{opt}.
In the literature some different propositions are made for the choice of ε [1, 2, 3, 4], but they are not founded mathematically [5].
Next we discuss a prototypical case to demonstrate the application of Eq. (2.19).

Logistic Map

We analyse as an example a time series generated by the logistic map

$$x_{n+1} = 4x_n\,(1 - x_n) \tag{2.20}$$

in a chaotic regime. On the left panel in Fig. 2.3 we represent the probability to

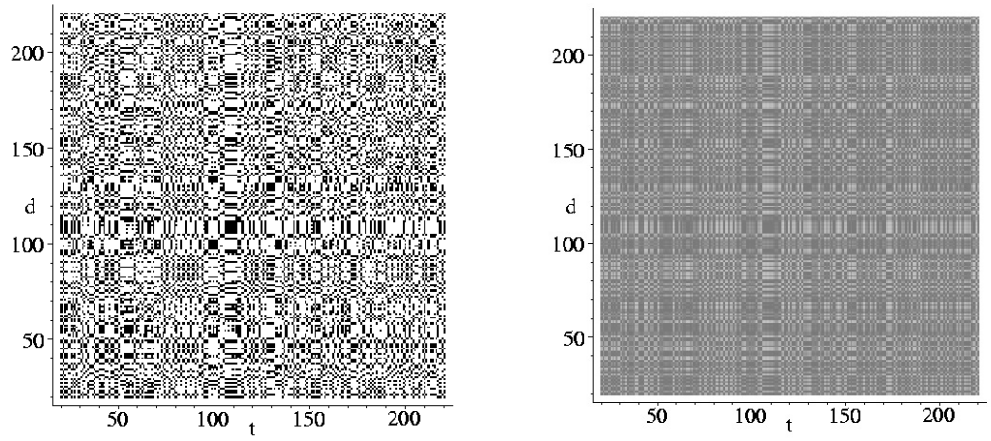

Figure 2.3: Left panel: RP of the logistic map (Eq. (2.20)). Right panel: probability plot of the logistic map (Eq. (2.20)) added by observational Gaussian white noise with standard deviation $\sigma = 0.1\Sigma$, if Σ is the standard deviation of x_n.

find a recurrence point for the logistic map without observational noise. As each "realization" is identical in this deterministic case, there are only probabilities one (=black) and zero (=white).
The right panel shows the same plot for the logistic map with observational noise. The black lines of the left plot seem to be smeared out now, as expected. There is no point that has probability one or zero (except from the main diagonal that always has probability one).
Calculating the distribution of the diagonal lines for both cases, we obtain different results. If we consider the logistic map and take the usually recommended ratio $\frac{\varepsilon}{\sigma} = 1$, already a signal/noise ratio of 10 will decrease the probability to recognise a recurrence point to about 50%. We will see in Sec. 2.3.4 that in this case the distribution of the diagonals will fail to detect the structures of the underlying process. We will also discuss the magnitudes of the changes due to

noise and then propose an optimal choice for ε that allows to recognise properly recurrence points with a much higher probability.

2.3.4 Estimation of the Errors due to Observational Noise

In this section we compute analytically the percentage of recurrence and non-recurrence points that are properly recognised. This will be used to derive a procedure to determine the optimal ε. We show that the results of RQA improve substantially with this optimal choice ε_{opt}.
Fig. 2.2 represents the probabilities to find recurrence points in the presence of noise if the underlying process leads to a difference d. Without any noise we have probability one for $|d| < 1$ and probability zero else. If the density $\rho(\cdot)$ of the increments $x_i - x_j$ is given, we can compute the percentage of recurrence points $p_b(\varepsilon, \sigma)$ that are properly recognised in the presence of observational noise

$$p_b(\varepsilon, \sigma) = \frac{\int_{-\varepsilon}^{\varepsilon} P(D, \varepsilon, \sigma)\rho(D)dD}{\int_{-\varepsilon}^{\varepsilon} \rho(D)dD} \tag{2.21}$$

where

$$P(D, \varepsilon, \sigma) = \frac{1}{8}\left\{\operatorname{erfc}^2\left(\tfrac{D-\varepsilon}{2\sigma}\right) - \operatorname{erfc}^2\left(-\tfrac{D-\varepsilon}{2\sigma}\right) + \operatorname{erfc}^2\left(-\tfrac{D+\varepsilon}{2\sigma}\right) - \operatorname{erfc}^2\left(\tfrac{D+\varepsilon}{2\sigma}\right)\right\} \tag{2.22}$$

is the solution of Eq. (2.19). Analogously, we compute the percentage of properly recognised non-recurrence points

$$p_w(\varepsilon, \sigma) = \frac{\int_{-\infty}^{-\varepsilon} \left[1 - P(D, \varepsilon, \sigma)\right] \rho(D)dD + \int_{\varepsilon}^{\infty} \left[1 - P(D, \varepsilon, \sigma)\right] \rho(D)dD}{1 - \int_{-\varepsilon}^{\varepsilon} \rho(D)dD} \tag{2.23}$$

We have to distinguish two types of errors that can occur (Tab. 2.1). To calculate

Point in the plot	is recognised as recurrence point	is not recognised as recurrence point
is recurrence point	p_b	$1 - p_b$
is not recurrence point	$1 - p_w$	p_w

Table 2.1: Errors and properly recognised recurrence points.

the optimal ε which reduces both types of errors, we further need the percentage or density of recurrence points in the presence of noise. That is the sum of the

properly and miss- recognised recurrence points in the plot divided by the total number of points

$$p_b^{\text{density}} = \int_{-\infty}^{\infty} P(D)\rho(D)dD \tag{2.24}$$

For the non recurrence points we find $p_w^{\text{density}} = 1 - p_b^{\text{density}}$. We compute numerically p_b, p_w for two different sets of data: a shot noise process and a logistic process. See Sec. 2.4 for a detailed discussion. Here we summarise the main results:
The main consequence of our approach is that we can optimise the choice of ε if the standard deviations σ of the Gaussian noise and Σ of the underlying process are given, so that the bias due to noise is minimised. Then we can compute the influence of noise for a given process when only its distribution is known. Simulations show that in many cases it is possible to estimate the influence of noise even for a measured time series i.e. a time series plus noise without knowing the underlying process. This works if the observational noise does not change the distribution of the time series too much.
Even though the results depend on the distribution of the time series, there are some general statements we can make (see Sec. 2.4).

1. For a wide class of processes the choice

$$\varepsilon \approx 5\sigma \tag{2.25}$$

   is appropriate. If ε is smaller, effects of the observational noise will have a dominant influence on the detection of recurrence points. On the other hand if $\varepsilon \approx \Sigma$, where Σ is the standard deviation of the underlying process, the density of recurrence points will be too high to detect detailed structures of the underlying process.

2. In the literature a frequently applied choice for ε is $\varepsilon \approx \frac{1}{10}\Sigma$ [Marwan, 1999, 4]. Note however that this is appropriate only if $\sigma < \frac{1}{50}\Sigma$ holds, as the condition Eq. (2.25) has to be met.

3. Some processes, e.g. discrete processes with only few states, may demand other choices of ε than Eq. (2.25). In these cases one has to compute Eqs. (2.21, 2.23).

These considerations allow to estimate the errors that will occur in the computation of diagonal lengths, vertical and horizontal lines. This is of special interest as these distributions are the skeleton of the RQA. The idea is to calculate a correction factor for the number of lines of the length l. Analogously to the calculus of the distribution of the diagonal lengths for stochastic processes, we have to consider not only the l black points that form the line, but also the white points

at the beginning and the end of the lines. Hence, for a given process and ratio $\frac{\varepsilon}{\sigma}$ we compute the percentage of properly and recognised recurrence points p_b and non-recurrence points p_w. Now the two white points are properly recognised with probability p_w and the l black points with probability p_b. This leads to the correction factor

$$K(l) = p_w^2 p_b^l \tag{2.26}$$

The meaning of this factor is the following. Given a process x_i and the corresponding distribution of diagonal (horizontal, vertical) lines $P_{\text{lines}}(l)$, then

$$P_{\text{lines}}^{\text{prop}}(l) = K(l) \cdot P_{\text{lines}}(l) \tag{2.27}$$

is the rate of diagonals of length l that one properly recognises, i.e. the lines that are found in the recurrence plot without noise and are still recognised in the plot with noise. Note that $P_{\text{lines}}^{\text{prop}}(l)$ is not the rate that one will actually count in a recurrence plot with noise. There also takes place a spurious detection of diagonals of length l that does not occur in the plot without noise.
Even though the distribution $P_{\text{lines}}(l)$ of the plot without noise does obviously depend on the dynamics of the time series, the correction factor $K(l)$ only depends on the distribution of the time series. It would be the same for a chaotic logistic process, a sine function or noise when they all have the same distribution.
We illustrate how to apply $K(l)$ for the case of the logistic map ($r = 4$, see also Sec. 2.4.1). We first estimate the error we will be faced with computing the distribution of diagonal lines in the presence of 10% noise for two different ratios of $\frac{\varepsilon}{\sigma}$. Table 2.2 gives the percentage of properly recognised diagonals of length l. The first line corresponds to the frequently applied choice of $\varepsilon = 0.1 \cdot \Sigma$ and the second one corresponds to the optimal threshold. In the case of the usual

$\frac{\varepsilon}{\sigma}$	$K(l=1)$	$K(l=2)$	$K(l=3)$	$K(l=4)$	$K(l=5)$	$K(l=6)$
1	46.4%	23.0%	11.4%	5.66%	2.81%	1.39%
5	88.4%	82.2%	76.5%	71.1%	66.1%	61.5%

Table 2.2: Comparison of the probabilities to properly recognise recurrence lines for different $\frac{\varepsilon}{\sigma}$ for the logistic map. $\frac{\varepsilon}{\sigma} = 1$ (usual choice),$\frac{\varepsilon}{\sigma} = 5$ optimal choice.

choice $\frac{\varepsilon}{\sigma} = 1$ the percentage of properly recognised diagonals is too low to draw conclusions about the dynamics from it. For the optimal choice of $\frac{\varepsilon}{\sigma}$ the results are much better. Even though we cannot determine the distribution of diagonals perfectly, we can hope to conserve enough structures to recognise main features of the underlying dynamics.
To illustrate the result of Eq. (2.27) and Tab. 2.2, we compute the distribution of diagonals. The embedding dimension is 1 and the level of noise is 10%. Fig. 2.4 (a) shows the distribution for the usual choice of ε ($\frac{\varepsilon}{\sigma} = 1$). We observe that in this

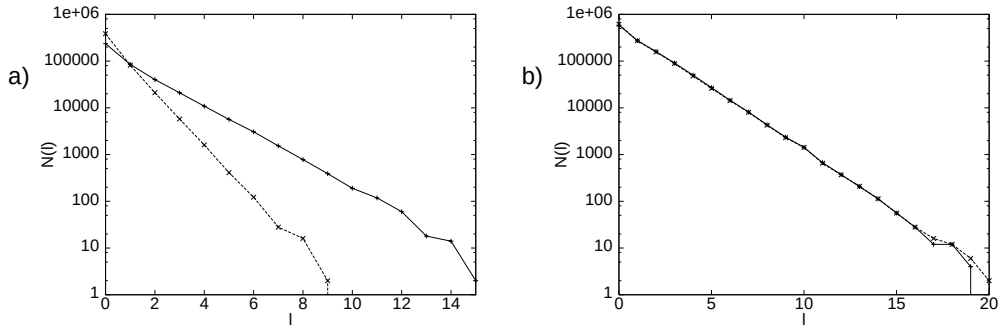

Figure 2.4: (a) Number of diagonals for the logistic map, $N(l) = L^2 P(l)$, analogous to Eq. (2.15). Solid line: without noise, dashed line: with 10% noise. Usual choice of ε. (b) The same but with optimal choice of ε.

case the distribution of the underlying process (solid) is biased if observational noise is present (dashed). Fig. 2.4 (b) shows the distributions for the optimal choice of ε ($\frac{\varepsilon}{\sigma} = 5$). We note that the distributions in this case coincide.
These results are also strongly reflected in the value of DET (see Eq. (2.5)) which is a basic quantity in the RQA (Tab. 2.3). The difference in the values of DET for

choice of ε	DET without noise	DET with noise
usual, $\frac{\varepsilon}{\sigma} = 1$	0.46	0.15
optimal, $\frac{\varepsilon}{\sigma} = 5$	0.57	0.56

Table 2.3: Determinism for usual and optimal choice of ε for the logistic map in the case with and without noise. Parameters: Length of the time series $L = 3000$, embedding dimension $m = 1$, minimal length $l_{\min} = 3$.

the case without noise is due to the different choices of ε. The results show that for the optimal choice of ε the determinism in the case of observational noise is nearly identical with the one of the underlying process whereas the usual choice ($\frac{\varepsilon}{\sigma} = 1$) leads to a strongly biased estimate, i.e. the error is of the order of 300%.

2.3.5 Results for Embedded Time Series

In this section the results of the last section are extended to embedded measured time series

$$\vec{y_i} = \vec{x_i} + \vec{\eta_i}$$

Therefore, Eq. (2.19) has to be extended to d-dimensions. We apply delay embedding [Kantz & Schreiber, 1997], [Takens, 1980] to reconstruct the vector $\vec{D}$ that has the components $(x_i - x_j, x_{i+\tau} - x_{j+\tau}, \ldots, x_{i+(m-1)\tau} - x_{j+(m-1)\tau})$, where m is the embedding dimension and τ the embedding delay (see also Sec. 4.1). So

$\vec{D}$ is the vector that corresponds to the point (i, j) in the CRP. But as for the further considerations it is not necessary to consider the coordinates we write D instead of $D_{i,j}$.

$$P\left(\vec{D}\right) = \frac{1}{\left(\sqrt{2\pi}\sigma\right)^{2m}} \int\limits_{\mathcal{R}^m} e^{-\frac{\vec{\eta}}{2\sigma^2}} \int\limits_{\mathcal{U}_{\vec{\varepsilon}}(\vec{\eta})} e^{-\frac{(\vec{\vartheta}-\vec{D})^2}{2\sigma^2}} d\vec{\vartheta}\, d\vec{\eta} \tag{2.28}$$

The outer integral has to be integrated over the entire space $\mathcal{R}^m$. The inner integral is solved in the ε-environment around $\vec{\eta}$. The integral is more difficult to solve if we consider a sphere around $\vec{\eta}$. It is easier to consider a box with the length of the sides 2ε. If the components of $\vec{D}$ are named D_k the solution of Eq. (2.28) is given by

$$P\left(\vec{D}\right) = \prod_{k=1}^{m} \frac{1}{8}\left\{\operatorname{erfc}^2\left(\tfrac{D_k-\varepsilon}{2\sigma}\right) - \operatorname{erfc}^2\left(-\tfrac{D_k-\varepsilon}{2\sigma}\right) + \operatorname{erfc}^2\left(-\tfrac{D_k+\varepsilon}{2\sigma}\right) - \operatorname{erfc}^2\left(\tfrac{D_k+\varepsilon}{2\sigma}\right)\right\} \tag{2.29}$$

if D_k are considered to be independent. $P\left(\vec{D}\right) = P\left(\vec{D}_{i,j}\right)$ gives the probability to find a recurrence point at the coordinates (i, j), if the underlying process y_i is embedded in a space of dimension m.

Embedding applied to random time series is problematic. However, the optimal choice of ε in the case with embedding is the same as in the case without embedding as the components of $\vec{D}$ enter Eq. (2.29) separately. The recurrence point is properly recognised if the product in Eq. (2.29) is maximal. In a first approximation this is the case if each factor is maximal. This condition leads to the same procedure as described for the case without delay embedding (see Sec. 2.3.4).

We illustrate this result for the Lorenz system (Eqs. (5.1)) with the parameters $\sigma = 16, r = 45.92, b = 4$. The step size for the integration is $h = 0.001$ and the sampling rate $\delta t = 10 \cdot h$. We use an embedding dimension of $m = 3$ and a delay $\tau = 8$ and present the results for the distribution of the diagonal lines and DET (Eq. 2.5) for a time series of length $N = 5,000$ for the case without and with observational noise in Fig. 2.5. The noise was white and Gaussian with $\sigma = 0.1\Sigma$. Then we compare the results for the usual choice of ε (a) with the optimal one (b). The distributions differ rather much for the usual choice whereas for the optimal choice they nearly coincide. The distributions then allow to compute DET ($l_{\min} = 2$) (Tab. 2.4). For the usual choice of ε the values of DET for the case with and without noise differ by a factor of about 7, whereas for the optimal choice the factor is 1.14. So the optimal choice reduces the factor by nearly one order of magnitude. This difference is crucial. In the case of noisy time series with the usual choice of ε the value of DET is so small that it is not distinguishable from a pure stochastic process. The optimal choice of the threshold on the other hand allows to recognise the underlying dynamics.

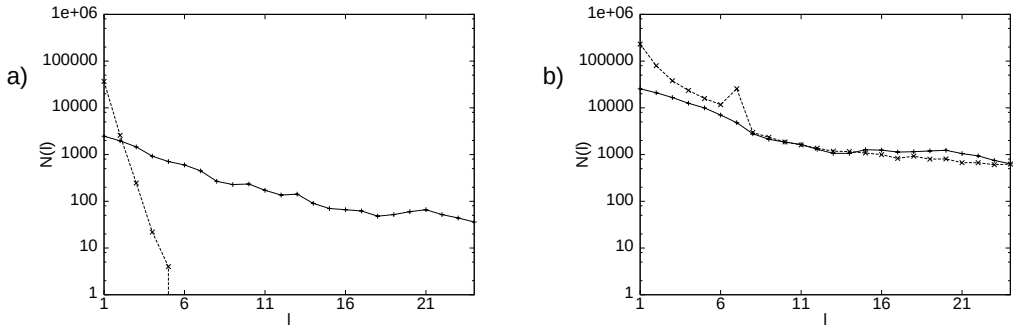

Figure 2.5: (a) Number $N(l)$ of the diagonals for the Lorenz system, solid line: without noise, dashed line: with noise, $\sigma = 0.1\Sigma$. Usual choice of ε. (b) The same but with optimal choice of ε.

choice of ε	DET without noise	DET with noise
usual, $\frac{\varepsilon}{\sigma} = 1$	0.98	0.14
optimal, $\frac{\varepsilon}{\sigma} = 5$	0.99	0.87

Table 2.4: Determinism for usual and optimal choice of ε for the Lorenz system in the case with and without noise. Parameters: Length of the time series $L = 5,000$, embedding dimension $m = 3$, delay $\tau = 8$, minimal length $l_{\min} = 2$.

2.3.6 Uniformly Distributed Noise

In this section we present the results for another class of random numbers (see Sec. 2.3.3). We will consider a uniform distribution in $[-a, a]$. Then, the probability to find a point at x is given by

$$p(x) = \frac{1}{2a} \left[\Theta(x+a) - \Theta(x-a)\right] \tag{2.30}$$

Analogously to the case of the Gaussian distribution, the probability to yield a recurrence point given the distance D of the underlying process is

$$\begin{aligned} P_{\mathrm{RP}} = & \int_{-\infty}^{\infty} p(\eta) \int_{\eta-D-\varepsilon}^{\eta-D+\varepsilon} p(\vartheta)\, d\vartheta\, d\eta \\ = & \tfrac{1}{8a^2} \left[2\Theta\left(D-\varepsilon\right)\left(D-\varepsilon\right)^2 - 2\Theta\left(D+\varepsilon\right)\left(D+\varepsilon\right)^2 \right. \\ & +\Theta\left(D-2a+\varepsilon\right)\left\{\left(D+\varepsilon\right)^2 + 4\left(a^2 - da - \varepsilon a\right)\right\} \\ & -\Theta\left(D-2a-\varepsilon\right)\left\{\left(D-\varepsilon\right)^2 + 4\left(a^2 - da + \varepsilon a\right)\right\} \\ & +\Theta\left(D+2a+\varepsilon\right)\left\{\left(D+\varepsilon\right)^2 + 4\left(a^2 + da + \varepsilon a\right)\right\} \\ & \left. -\Theta\left(D+2a-\varepsilon\right)\left\{\left(D-\varepsilon\right)^2 + 4\left(a^2 + da - \varepsilon a\right)\right\}\right] \end{aligned} \tag{2.31}$$

The following procedure is completely analogous to the one for the Gaussian

distribution. One can simply substitute Eq. (2.31) into the corresponding expressions in the latter case. So we will not discuss it in more detail here.
Applying the first line of Eq. (2.31) to distributions $p(x)$ different from Eq. (2.30), one can compute P_{RP} for arbitrary probability distributions. Even an extension to embedded time series is analogue to the case of the Gaussian distribution.

2.3.7 Application to Experimental Data

Next, we apply our modified RQA to data of an experimental system of a CO_2 laser with sinusoidal modulation of the cavity losses. More precisely we are dealing with a conventional low-speed longitudinal gas flow CO_2 laser with a Ge intracavity acousto-optic modulator (Mod. AGM-406B1 IntraAction Corp.). The optical cavity, 1.35 m long, is defined by a diffraction grating selecting the P(20) line at 10.6 mm and an outcoupler mirror with a reflectivity of 90%. By applying a sinusoidal signal to the acousto-optic modulator at 100 kHz, which is close to the relaxation oscillation of the system, the laser reaches a chaotic condition after a sequence of subharmonic bifurcations [Arecchi et al., 1982]. The chaotic output intensity is detected by means of a high speed uncooled Hg-Cd-Zn-Te photodetector with a detectivity of 2.3 107 cm Hz-1/W and a bandwidth of 20 MHz (PD-10.6-8 Vigo System Ltd.). Time series of $50,000$ points at a time resolution of 200 ns are recorded on a digital scope (LT 342L Lecroy). Figure 2.6 represents a sample (20 periods) of the data. We perform the analysis to com-

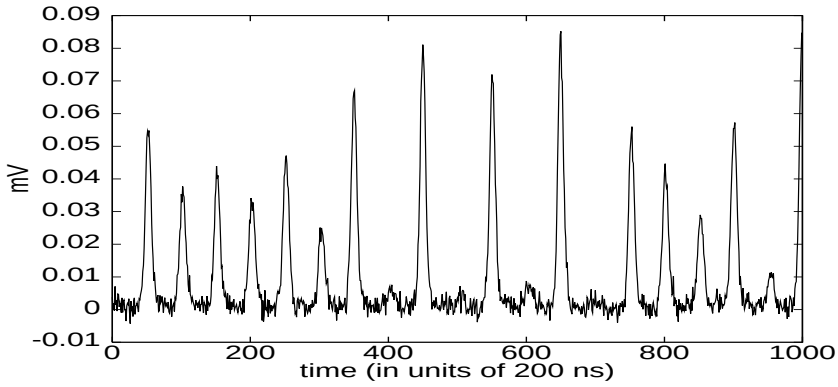

Figure 2.6: Output of the CO_2 laser corrupted by 10% observational Gaussian noise.

pute the distribution of diagonal lines by using the embedding dimension $m = 3$ and the delay $\tau = 10$. The results of the original time series and the time series corrupted by 10% observational Gaussian white noise are compared (Fig. 2.7). The usual choice of the threshold is represented in (a) and the optimal choice in (b). The optimal choice of ε yields an estimation of $N(l)$ which is much closer to the non noise corrupted time series than the usual choice. Tab. 2.5 shows that the values of DET reflect the improved estimation of $N(l)$. The factor between

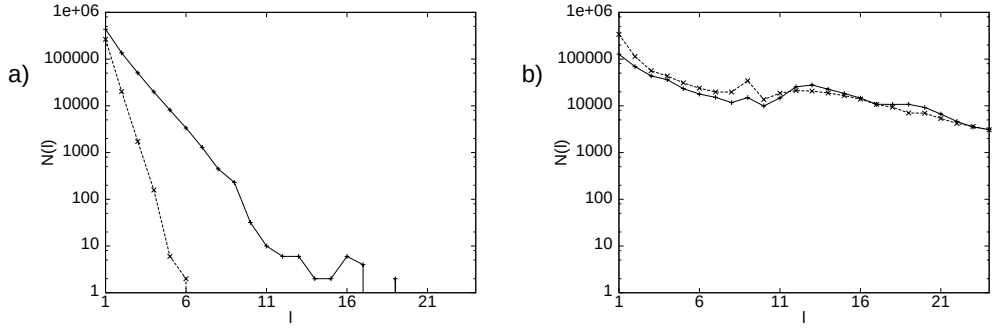

Figure 2.7: (a) Number $N(l)$ of the diagonals for the Laser system, solid line: without noise, dashed line: with 10% noise. Usual choice of ε. (b) The same but with optimal choice of ε.

choice of ε	DET without noise	DET with noise
usual, $\frac{\varepsilon}{\sigma} = 1$	0.57	0.15
optimal, $\frac{\varepsilon}{\sigma} = 5$	0.98	0.94

Table 2.5: Determinism for usual and optimal choice of ε for the Laser system in the case with and without noise. Parameters: Length of the time series $L = 5,000$, embedding dimension $m = 3$, delay $\tau = 10$, minimal length $l_{\min} = 2$.

the values for the determinism in the case without and with noise and the usual choice of ε is 3.8. For the optimal choice of ε the factor is 1.04 and hence very close to unity.
Resuming one can say, that already weak observational noise can change considerably the statistics that usually are regarded. To solve this problem, we have proposed, based on our analytical results, to choose the threshold ε at least five times the standard deviation of the observational Gaussian noise σ. This choice is appropriate in most cases. The presented expressions for the deviations we have to take into account in the case of observational noise, that help to evaluate the reliability of the results of the RQA. However, if the level of observational noise is too high ($\sim 20\%$ of the standard deviation of the underlying process or more) the application of the RQA to the data can lead to pitfalls.
The results motivate the development of new statistics for the RQA, different from simple distributions of diagonal lengths (Chap. 3).
Our results can easily be extended to other classes of noise and hold for all underlying processes. We have already mentioned problems stochasticity causes for embedding. Anyway, the results of this section show that it is possible to consider embedding, too.

2.4 On the Optimal Choice of ε

We now study the influence of observational noise on the correct recognition of recurrence points for two time series (logistic map and shot noise process), Fig. 2.8. Even though the dynamics of both systems is rather different, the optimal choice of ε is similar in both cases. Not without a certain audacity to generalise, we find the thumb rule that choosing $\frac{\varepsilon}{\sigma} \leq 5$, yields rather reliable results.

2.4.1 Logistic Map

First we consider the logistic map Eq. (2.20) in a chaotic regime. The standard deviation of the process is $\Sigma = \frac{3}{10}$ and we use a noise level of 10%, i.e. $\sigma = 0.1\Sigma$. The length of the time series is 65,000. Fig. 2.8 shows that p_b increases with growing $\frac{\epsilon}{\sigma}$. This is plausible as for small $\frac{\varepsilon}{\sigma}$ the noise makes it difficult to recognise

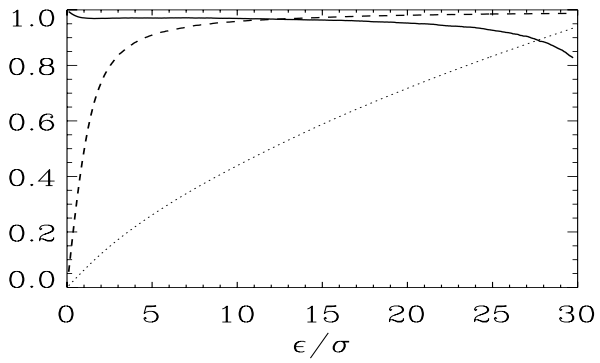

Figure 2.8: Rates of properly recognised recurrence and non recurrence points in the case of the logistic map for different rations of $\frac{\varepsilon}{\sigma}$ (solid resp. dashed). Dotted: RR.

properly small differences $D_{i,j} = x_i - x_j$. If $D_{i,j}$ is small it is probable to find $\tilde{D}_{i,j} > D_{i,j}$, where $\tilde{D}_{i,j} = y_i - y_j = x_i - x_j + \eta_i - \eta_j$ (see Sec. 2.3.3). With growing $\frac{\varepsilon}{\sigma}$ this effect gets less important for the discrimination of recurrence points. So p_b increases for growing $\frac{\varepsilon}{\sigma}$. Analogously, we see that p_w decreases with growing $\frac{\varepsilon}{\sigma}$. The bigger this ratio the more difficult it will be to determine non-recurrence points. Therefore it is reasonable to maximise p_b and p_w. p_b increases till $\frac{\varepsilon}{\sigma} \sim 5$ and then saturates. p_w is nearly constant for $\frac{\varepsilon}{\sigma} < 20$ and then decreases. So in the interval $\frac{\varepsilon}{\sigma} \in [5, 20]$ both probabilities are on a high level, consequently all of the values of $\frac{\varepsilon}{\sigma}$ in this interval are suitable. If we now take into consideration that the percentage of recurrence points is very high for large fractions of $\frac{\varepsilon}{\sigma}$, we tend to choose optimally $\frac{\varepsilon}{\sigma} \approx 5$, because small ratios allow to distinguish smaller changes in the dynamics of the underlying process.
These results are similar for a wide class of processes, e.g. shot noise processes, chaotic oscillators etc.

Chapter 3

Estimation of Dynamical Invariants from RPs

3.1 White Noise Processes

As mentioned in Sec. 2.1 stochastic time series yield RPs which display mainly single points or short diagonal lines. In Sec. 2.3.2 we have calculated analytically the distribution of diagonal lines for independent noise. In this section we compute the corresponding cumulative distribution of diagonals, for reasons which will become clear later. The probability to find a recurrence point in the RP (which is equal to the recurrence rate defined in Eq. (2.4)) is given by

$$RR(\varepsilon) = \lim_{N\to\infty} \frac{1}{N^2} \sum_{i,j=1}^{N} \mathbf{R}_{i,j}, \tag{3.1}$$

and the probability to find a diagonal of at least length l in the RP is given by

$$P_\varepsilon^c(l) = \lim_{N\to\infty} \frac{1}{N^2} \sum_{i,j=1}^{N} \prod_{m=0}^{l-1} \mathbf{R}_{i+m,j+m}, \tag{3.2}$$

where c stands for cumulative. Note that $RR(\varepsilon) = P_\varepsilon(1)$.
We consider a random variable X with probability density $\rho(x)$. Suppose that $\{x_i\}$ for $i = 1, \ldots, N$ is a realization of X and we are interested in the distribution of the distances of each point to all other points of the time series. This can be done by computing the convolution of the density $\rho(\cdot)$

$$R(x) = \rho(x) * \rho(x). \tag{3.3}$$

$RR(\varepsilon)$ is then gained by integrating $R(x)$ over $[-\varepsilon, \varepsilon]$

$$RR(\varepsilon) = \int_{-\varepsilon}^{\varepsilon} R(x)dx = 2\int_{0}^{\varepsilon} R(x)dx. \tag{3.4}$$

Note that $RR(\varepsilon)$ is invariant against shuffling of the data. For $[0,1]$ uniformly distributed noise, $R(x)$ is given by

$$R(x) = \begin{cases} 1-|x| & \text{if } |x|<1 \\ 0 & \text{else} \end{cases} \tag{3.5}$$

and hence the probability $RR(\varepsilon)$ for RPs and CRPs is given by

$$RR(\varepsilon) = 2\varepsilon - \varepsilon^2 + \Theta(\varepsilon - 1)\left[1 - 2\varepsilon + \varepsilon^2\right] \tag{3.6}$$

For Gaussian white noise one finds $RR(\varepsilon) = \operatorname{erf}\left(\frac{\varepsilon}{2\sigma}\right)$, where σ is the standard deviation, in accordance with the results presented in Sec. 2.3.2. Now it is straightforward to compute $P_\varepsilon^c(l)$ in the in RPs. As the noise is independent, we obtain

$$P_\varepsilon^c(l) = RR(\varepsilon)^l. \tag{3.7}$$

The probability to find a recurrence point $RR(\varepsilon)$ is in RPs independent of the preceeding point on the diagonal (except in the main diagonal). Eq. (3.7) shows that the probability to find a line of length l decreases exponentially with l. For our example of uniformly distributed noise we get

$$P_\varepsilon^c(l) = \left(2\varepsilon - \varepsilon^2\right)^l. \tag{3.8}$$

Note that in this case the exponential decay depends on ε. This will be an important different with respect to e.g. chaotic systems presented in the next section.

3.2 Chaotic Systems

We present in this section an approach for chaotic systems. It is an extension of the results presented in [Faure & Korn, 1998] for chaotic maps and also covers general, continuous chaotic flows. The results presented in this section are fundamental for many of the considerations in the course of this work. From now on we will use the maximum norm in Eq. 2.1.
To estimate the distribution of the diagonals in the RP, we start with the correlation integral

$$C(\varepsilon) = \lim_{N\to\infty} \frac{1}{N^2} \times \{\text{number of pairs } (i,j) \text{ with } |\vec{x}_i - \vec{x}_j| < \varepsilon\} \tag{3.9}$$

introduced by Grassberger and Procaccia [Grassberger & Procaccia, 1983b]. Note that the definition of $RR(\varepsilon)$ coincides with the definition of the correlation integral

$$C(\varepsilon) = \lim_{N\to\infty} \frac{1}{N^2} \sum_{i=1}^{N} \Theta\left(\varepsilon - |\vec{x}_i - \vec{x}_j|\right) \overset{\text{Eq. (3.1)}}{=} \lim_{N\to\infty} \frac{1}{N^2} \sum_{i,j=1}^{N} \mathbf{R}_{i,j} = P_b(\varepsilon). \tag{3.10}$$

This fact allows to link the known results about the correlation integral and the structures in RPs.
We consider a trajectory $\vec{x}(t)$ in the basin of attraction of an attractor in the d-dimensional phase space. The state of the system is measured at time intervals τ. Let $\{1, 2, ..., M(\varepsilon)\}$ be a partition of the attractor in boxes of size ε. Then $p(i_1, ..., i_l)$ denotes the joint probability that $\vec{x}(t = \tau)$ is in the box i_1, $\vec{x}(t = 2\tau)$ is in the box i_2, ..., and $\vec{x}(t = l\tau)$ is in the box i_l. The order-2 Rényi entropy [Rényi, 1970, Grassberger, 1983] is then defined by

$$K_2 = -\lim_{\tau\to 0}\lim_{\varepsilon\to 0}\lim_{l\to\infty}\frac{1}{l\tau}\ln\sum_{i_1,\ldots,i_l} p^2(i_1,\ldots,i_l). \tag{3.11}$$

We can approximate $p(i_1, \ldots, i_l)$ by the probability $P_{t,l}(\vec{x}, \varepsilon)$ of finding a sequence of points in boxes of length ε about $\vec{x}(t = \tau)$, $\vec{x}(t = 2\tau)$, ..., $\vec{x}(t = l\tau)$. Assuming that the system is ergodic, which is always the case for chaotic systems as they are mixing, we obtain

$$\sum_{i_1,\ldots,i_l} p^2(i_1,\ldots,i_l) = \frac{1}{N}\sum_{t=1}^{N} p_t(i_1,\ldots,i_l) \sim \frac{1}{N}\sum_{t=1}^{N} P_{t,l}(\vec{x},\varepsilon), \tag{3.12}$$

where $p_t(i_1, \ldots, i_l)$ represents the probability of being in the box i_1 at time $t = \tau$, in the box i_2 at time $t = 2\tau$, ... and in the box i_l at time $t = l\tau$. Furthermore, we can express $P_{t,l}(\vec{x}, \varepsilon)$ by means of the recurrence matrix

$$P_{t,l}(\vec{x},\varepsilon) = \frac{1}{N}\sum_{s=1}^{N}\prod_{m=0}^{l-1}\Theta\left(\varepsilon - |\vec{x}_{t+m} - \vec{x}_{s+m}|\right) = \frac{1}{N}\sum_{s=1}^{N}\prod_{m=0}^{l-1}\mathbf{R}_{t+m,s+m}. \tag{3.13}$$

Hence, we obtain an estimator for the second order Rényi entropy by means of the RP

$$\hat{K}_2(\varepsilon, l) = -\frac{1}{l\tau}\ln\underbrace{\left(\frac{1}{N^2}\sum_{t,s=1}^{N}\prod_{m=0}^{l-1}\mathbf{R}_{t+m,s+m}\right)}_{(*)}. \tag{3.14}$$

Note that $(*)$ is the cumulative distribution of diagonal lines $P_\varepsilon^c(l)$ (Eq. (3.2)). Therefore, if we represent $P_\varepsilon^c(l)$ in a logarithmic scale versus l we should obtain a straight line with slope $-\hat{K}_2(\varepsilon)\tau$ for large l's.
On the other hand, in the G-P algorithm the l-dimensional correlation integral is defined as

$$C_l(\varepsilon) = \lim_{N\to\infty}\frac{1}{N^2}\sum_{t,s=1}^{N}\Theta\left(\varepsilon - \left(\sum_{k=0}^{l-1}|\vec{x}_{i+k} - \vec{x}_{j+k}|^2\right)^{1/2}\right). \tag{3.15}$$

Grassberger and Procaccia [Grassberger & Procaccia, 1984] state that due to the exponential divergence of the trajectories, requiring $\sum_{k=0}^{l-1} |\vec{x}_{i+k} - \vec{x}_{j+k}|^2 \leq \varepsilon^2$ is essentially equivalent to

$$|\vec{x}_{i+k} - \vec{x}_{j+k}| < \varepsilon \quad \text{for } k = 1, \ldots, l. \tag{3.16}$$

Then they make the ansatz:

$$C_l(\varepsilon) \sim \varepsilon^{D_2} \exp(-l\tau K_2). \tag{3.17}$$

Moreover, they make use of the Takens embedding theorem [Takens, 1980] and reconstruct the whole trajectory from l measurements of any single coordinate (see also Sec. 4.1). Hence, they consider

$$\tilde{C}_l(\varepsilon) = \lim_{N\to\infty} \frac{1}{N^2} \sum_{t,s=1}^{N} \Theta \left(\varepsilon - \left(\sum_{k=0}^{l-1} |x_{i+k} - x_{j+k}|^2 \right)^{1/2} \right). \tag{3.18}$$

and assume that $\tilde{C}_l(\varepsilon)$ gives the same estimate $\tilde{C}_l(\varepsilon) \sim \varepsilon^{D_2} exp(-l\tau K_2)$. Then, the G-P algorithm obtains an estimator of K_2 considering

$$\tilde{K}_2(\varepsilon, l) = \frac{1}{\tau} \ln \frac{C_l(\varepsilon)}{C_{l+1}(\varepsilon)}. \tag{3.19}$$

Due to the similarity of the RP approach to the G-P one, we state

$$P_\varepsilon^c(l) \simeq \sum_{i_1,\ldots,i_l} p^2(i_1, \ldots, i_l) \simeq \tilde{C}_l(\varepsilon) \sim \varepsilon^{D_2} \exp(-l\tau K_2). \tag{3.20}$$

The difference between both approaches is that in $P_\varepsilon^c(l)$ we further consider information about l vectors, whereas in $\tilde{C}_l(\varepsilon)$ we have just information about l coordinates. Besides this, in the RP approach l is a length in the plot, whereas in the G-P algorithm it means the embedding dimension. We will comment in detail on the connection between line length and embedding dimension in Sec. 4.4.
In Sec. 3.3 we illustrate these results for the Rössler system.

3.3 The Rössler System

We now apply the algorithm to estimate K_2 to the prototypical Rössler system with standard parameters $a = b = 0.2, c = 5.7$ [Rössler, 1976]

$$\begin{aligned} \dot{x} &= -y - z \\ \dot{y} &= x + ay \\ \dot{z} &= b + (x - c)z. \end{aligned} \tag{3.21}$$

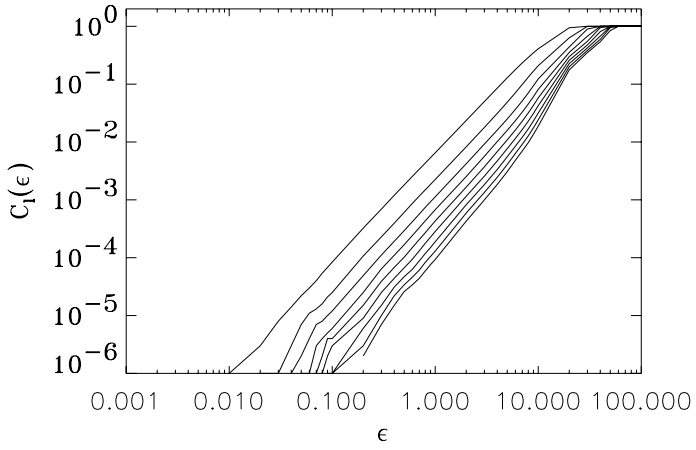

Figure 3.1: G-P algorithm for the Rössler system. l varies from 3 (top) to 27 (bottom) in steps of 3.

to demonstrate its applicability. We generate 15,000 data points based on the Runge Kutta method of fourth order and neglect the first 5,000. The integration step is $h = 0.01$ and the sampling rate is 20.
First, we estimate K_2 by means of the G-P algorithm. Fig. 3.1 shows the results for the correlation integral in dependence on ε. There is one well-expressed scaling region for each embedding dimension l. Then we obtain estimate of K_2 from the vertical distances between the lines (Fig. 3.2), $\tilde{K}_2 = 0.070 \pm 0.003$.

Next, we calculate the cumulative distribution of the diagonal lines of the RP in dependence on the length of the lines l (Fig. 3.3). For large l and small ε the scaling breaks down as there are not enough lines in the RP (see also Sec. 3.4). The most remarkable fact in this figure is the existence of two well differentiated scaling regions. The first one is found for $1 \leq l \leq 84$ and the second one for $l \geq 85$. The existence of two scaling regions is a new and striking point of this analysis that is not observed with the G-P method. The estimate of K_2^f from the slope of the first part of the lines is $K_2^f \approx 0.225 \pm 0.03$ (Fig. 3.4) and the one from the second part is $K_2 \approx 0.0675 \pm 0.004$ (Fig. 3.5). Hence, K_2^f is between 3-4 times higher than K_2. As K_2 is defined for $l \to \infty$ the second slope yields the estimation of the entropy.
However, the slope of the first part of the curve is interesting too, as it is also independent of ε. The region $1 \leq l \leq 84$ characterises the short term dynamics of the system up to three cycles around the attractor and corresponds in absolute units to a time of $t = 16.8$s, as we use a sampling rate of $\delta t = 0.2$s. These three cycles reflect a characteristic period of the Rössler system that we will call *recurrence period* T_{rec}. It is different from the dominant "phase period" T_{ph}, which

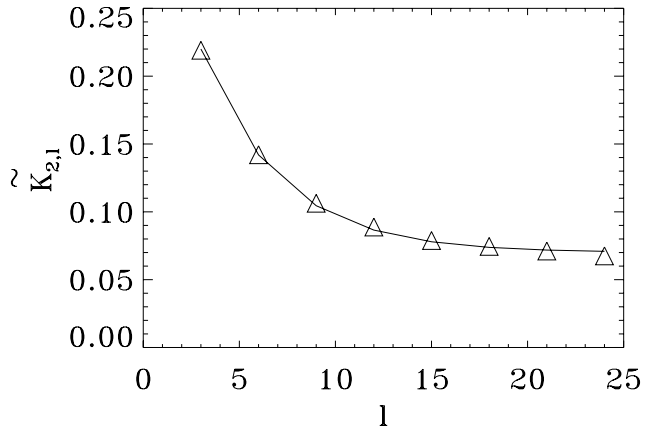

Figure 3.2: Estimation of $K_{2,l}$ for the Rössler system with the G-P algorithm. The line is plotted to guide the eye.

is given by the dominant frequency of the power density spectrum. $T_{\rm rec}$ however, is given by the recurrences to the same state in phase space.
On the other hand, recurrences to a given state in phase space are represented in the plot by vertical (or horizontal, as the plot is symmetric) white lines. Such a white line starts at the coordinates i, j if

$$\mathbf{R}_{i,j+m} = \begin{cases} 1 & \text{if} \quad m = -1 \\ 0 & \text{for} \quad m \in \{0, \ldots, l-1\} \\ 1 & \text{if} \quad m = l. \end{cases} \tag{3.22}$$

The trajectory $\vec{x}_n$ for times $n = j-1, \ldots, j+l$ is compared to the point $\vec{x}_i$. Then the structure given by Eq. (3.22) can be interpreted as follows: at time $n = j-1$ the trajectory falls within an ε-box of $\vec{x}_i$. Then for $n = j, \ldots, j+l-1$ it moves outside of the box, until at $n = j+l$ it recurs to the ε-box of $\vec{x}_i$. Hence, the length of the white line is proportional to the time that the trajectory needs to recur close to $\vec{x}_i$.
In Fig. 3.6 we represent the distribution of white vertical lines in the RP of the Rössler system (Eqs. (3.21)). The period of about 28 points corresponds to $T_{\rm ph}$. However, the highest peak is found at a lag of about 87 points (the second scaling region in Fig. 3.21 sets in at $l = 85$). This means that after this time most of the points recur close to their initial state. This time also defines the recurrence period $T_{\rm rec}$. For the Rössler attractor with standard parameters we find $T_{\rm rec} = 3T_{\rm ph}$. For predictions on time scales below the recurrence period, $\tau^f = 1/K_2^f$ is a better estimate of the prediction horizon than $\tau = 1/K_2$. This result means that the possibility to predict the next value within the epsilon range is in the first part

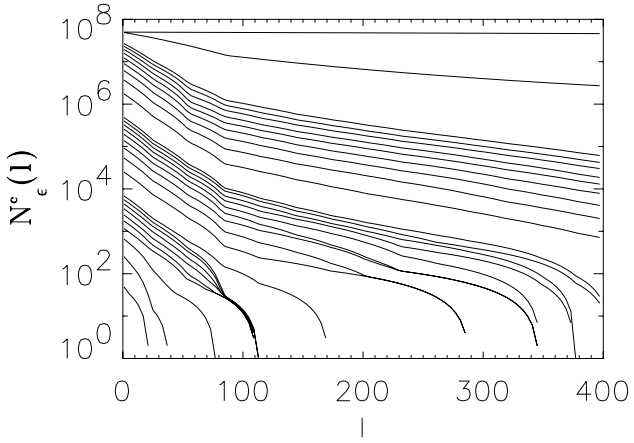

Figure 3.3: RP method for the Rössler system. ε varies logarithmically from 10^{-2} to 10.0 (bottom to top)

by a factor of more than 3 times worse than it is in the second part, i.e. there exist two time scales that characterise the attractor. The first slope is greater than the second one because it is more difficult to predict the next step if we have only information about a piece the trajectory for less than the recurrence period. Once we have scanned the trajectory for more than T_{rec}, the predictability increases and the slope of $P^c_\varepsilon(l)$ in the logarithmic plot decreases. Hence the first slope, as well as the time scale at which the second slope begins, reveal important characteristics of the attractor.

To investigate how the length of the first scaling region depends on the form of the attractor, we have varied the parameter c of the Rössler system with fixed $a = b = 0.1$, so that different types of attractors appear [Alligood et al., 1996]. Especially, we have considered the cases $c = 9$, which yields $T_{\mathrm{rec}} = 2T_{\mathrm{ph}}$, and $c = 30$, which gives $T_{\mathrm{rec}} = 4T_{\mathrm{ph}}$. In both cases the length of the first scaling region corresponds as expected to T_{rec}.

On the other hand, the existence of the two scalings may be linked to the nonhyperbolic nature of the Rössler system for this attractor type, because the resulting two time scales have been also recently found by Anishchenko et al. based on a rather subtle method [Anishchenko et al., 2004]. They report two slopes in the envelope of the correlation function of some oscillatory chaotic systems. The first rather large slope (for small delays) is linked to amplitude fluctuations of the system. The second smaller slope (for rather large delays) is related to the system's phase diffusion. Moreover, based on our numerical simulations the second slope is directly related to K_2. However, the research in this field is still in progress. It is noteworthy that the two slopes are also is detectable in other oscillating

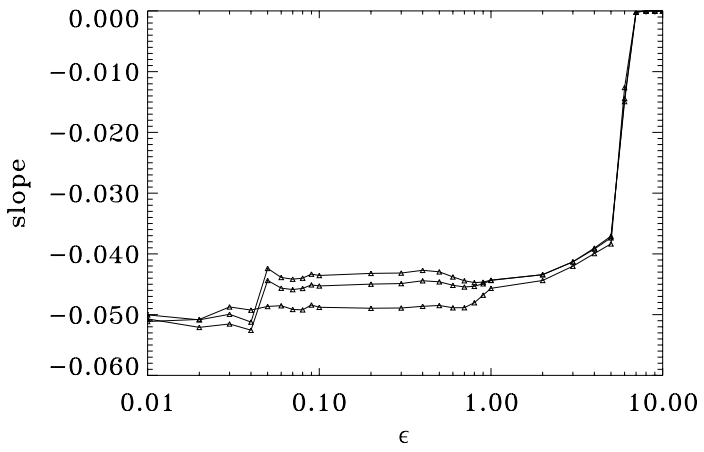

Figure 3.4: RP method for the Rössler system: slope of the curves $N_\varepsilon^c(l)$ in the first region for three different choices of the scaling region in l.

nonhyperbolic systems like the Lorenz system.

3.4 Noncumulative Distribution

In Sec. 3.2 we have shown that the cumulative distribution of diagonal lines $P_\varepsilon^c(l)$ is crucial for the estimation of dynamical invariants. We now study how this cumulative distribution is linked to the noncomulative distribution. $P_\varepsilon^c(l)$ can be calculated based on $P_\varepsilon(l)$ by considering that it counts lines that have at least length l, i.e. a line of length l contains two lines of exactly length $l-1$, three of exactly length $l-2$, and so on, whereas the non-cumulative distribution $P_\varepsilon(l)$ is determined by the occurrence of lines that have exactly length l. Then, we get

$$P_\varepsilon^c(l) = \sum_{i=0}^{\infty} (i+1) P\left(l+i\right). \tag{3.23}$$

Solving this equation for $P_\varepsilon(l)$, we obtain after some algebra

$$P_\varepsilon(l) = P_\varepsilon^c(l) - 2P_\varepsilon^c(l+1) + P_\varepsilon^c(l+2). \tag{3.24}$$

Eqs. (3.23) and (3.24) suggest that if either $P_\varepsilon(l)$ or $P_\varepsilon^c(l)$ decay exponentially the other one will also decay exponentially with the same exponent. Hence, the slope of the logarithmic plots of $P_\varepsilon(l)$ and $P_\varepsilon^c(l)$ can be compared directly. This result can be generalised for the case of m different slopes and for a maximal

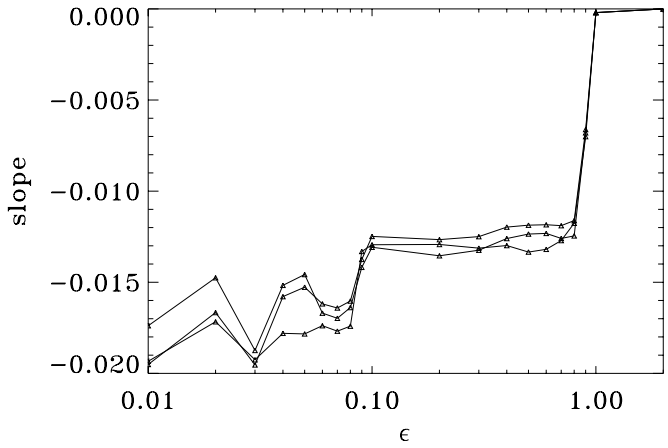

Figure 3.5: RP method for the Rössler system: slope of the curves $N_\varepsilon^c(l)$ in the second region for three different choices of the scaling region in l

considered line length of M. Then we find

$$\begin{aligned} P_\varepsilon(l) &= \sum_{j=1}^{m} a_i \exp(-K_i l) \\ P_\varepsilon^c(l) &= \sum_{j=1}^{m} \sum_{i=0}^{M-l} a_i (i+1) \exp(-K_i l) \\ &= \sum_{j=1}^{m} a_i \left\{ \frac{e^{-K_i l}}{\left(1-e^{-K_i}\right)^2} \right. \\ &\quad \left. - \frac{\left[1+(M-l+1)(1-e^{-K_i})\right] e^{-K_i(M+1)}}{\left(1-e^{-K_i}\right)^2} \right\} \\ &\overset{M\to\infty}{=} \sum_{j=1}^{m} a_i \frac{e^{-K_i l}}{\left(1-e^{-K_i}\right)^2} \end{aligned} \tag{3.25}$$

Eq. (3.25) does not only show that also for m different slopes it is possible to compare $P_\varepsilon(l)$ with $P_\varepsilon^c(l)$, but also explains the finite size effects in Fig. 3.3 for large l. Fig. 3.7 shows the deviations of $P_\varepsilon^c(l)$ from the straight line for $m = 1$, $K_2 = 0.05$ and $M = 150$ resp. $M = 200$. For $M = 150$ the scaling region with respect to l extends from $l = 1, \ldots, \sim 65$ and for $M = 200$ from $l = 1, \ldots, \sim 120$. Eq. (3.25) further shows that the scaling region also depends on K_2. If the largest line length that falls in the scaling region of a system with the correlation entropy K_2^1 is $l_{\max}^1$ resp. $l_{\max}^2$ for a second system with the corresponding entropy K_2^2, then

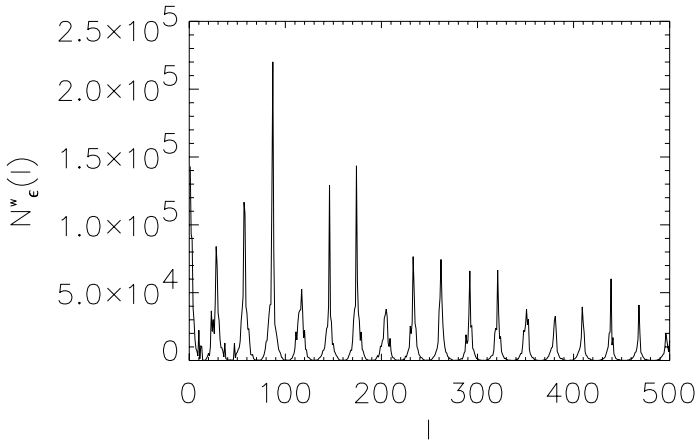

Figure 3.6: Number of vertical white lines in the Recurrence Plot of the Rössler system with standard parameters, $\varepsilon = 0.05$ and based on 60,000 data points.

we find the relation

$$l^1_{\max} < l^2_{\max} \qquad \text{if} \quad K^1_2 < K^2_2,$$

i.e. the larger the correlation entropy the larger is the scaling region.
Eq. (3.25) also allows to conclude that the finite size effect leads to an overestimation of K_2. Fig. 3.7 shows that the absolute value of the slope is larger if the correction for finite M is considered. The best approximation is made for small l because the deviation from the straight line is rather small then.
After these considerations about finite size effects, we estimate in the next section the correlation dimension based on RPs.

3.5 Correlation Dimension

In this section we show that it is also possible to estimate the correlation dimension by means of the distribution of diagonals in the RP. Eq. (3.20) also suggests how to exploit the vertical distance between $P^c_\varepsilon(l)$ for different ε's. From Eq. (3.20) one can derive

$$\hat{D}_2(\varepsilon) = \ln\left(\frac{P^c_\varepsilon(l)}{P^c_{\varepsilon+\Delta\varepsilon}(l)}\right)\left(\ln\left(\frac{\varepsilon}{\varepsilon+\Delta\varepsilon}\right)\right)^{-1}. \tag{3.26}$$

This is an estimator of the correlation dimension D_2 [Grassberger, 1983]. The result for the Rössler system is represented in Fig. 3.8. The mean value of $\hat{D}_2(\varepsilon)$ is in this case 1.86 ± 0.04. This result is in accordance with the estimation of D_2 by the G-P algorithm given in [Raab & Kurths, 2001], where the value 1.81 is

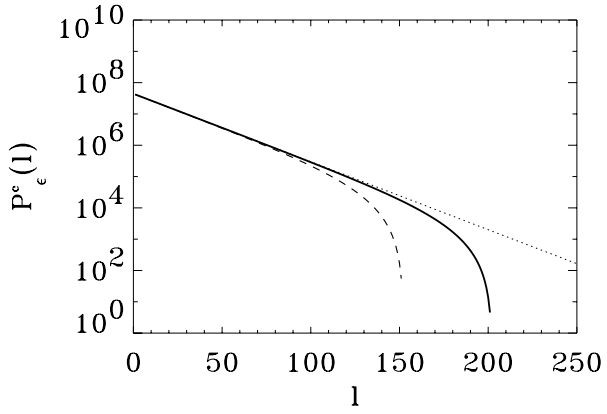

Figure 3.7: The cumulative distribution of diagonals for finite M. The straight line represents the limit for $N \to \infty$. The solid line shows the distribution of diagonals for $M = 200$ and the dashed line for $M = 150$.

obtained. With a modified G-P algorithm, Raab and Kurths obtain a value of 1.89 [Raab & Kurths, 2001]. Note, that to obtain the estimate for D_2 by Eq. (3.26) we have computed the average over lines of length l which correspond to the first. If one restricts the average to l in the second region one obtains the slightly higher value $D_2 = 2.07 \pm 0.01$, which in accordance with the value $D_2 = 2.06 \pm 0.02$ obtained in [Grassberger & Procaccia, 1983b] and [Hübner et al., 1993].
Note, that the extent and the onset of a scaling region in $D_2(\varepsilon)$ may lead to problems in the D_2 estimation [Ding et al., 1993].

3.6 Mutual Information

Both K_2 and D_2 have been estimated from the distribution of the black diagonal lines in an RP. However, one can also study non-diagonal structures. We therefore suggest, an estimator of the generalised mutual information of order 2,

$$I_2(\tau) = 2H_2 - H_2(\tau) \tag{3.27}$$

where

$$H_2 = -\ln \sum_i p_i^2, \qquad H_2(\tau) = -\ln \sum_i p_{i,j}^2(\tau) \tag{3.28}$$

are the generalised Rényi's second order information (also correlation entropy) and its corresponding joint second order information [Pompe, 1993]. This measure can be estimated using the G-P algorithm as follows [Kantz & Schreiber, 1997]

$$\tilde{I}_2(\varepsilon, \tau) = \ln(C_2(\varepsilon, \tau)) - 2\ln(C_1(\varepsilon)). \tag{3.29}$$

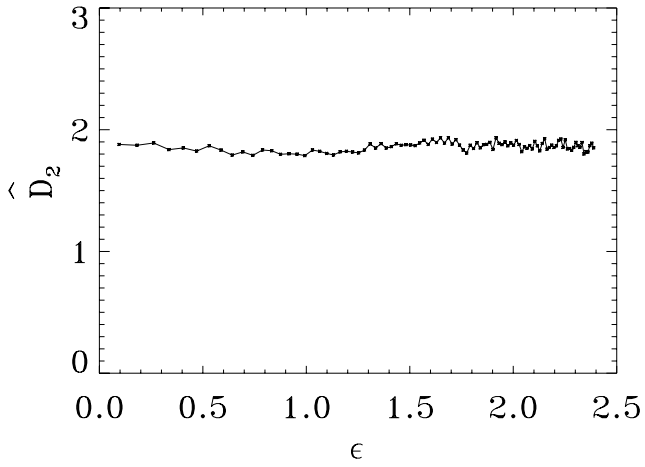

Figure 3.8: Estimation of the correlation dimension D_2 for the Rössler attractor by the RP method. The parameters used for the Rössler system and the integration step are the same as in Sec. 3.3.

Instead, we can estimate $I_2(\tau)$ using the recurrence matrix. Analogously the preceding sections, one can estimate H_2 by

$$\hat{H}_2 = -\ln\left[\frac{1}{N^2}\sum_{i,j=1}^{N}\mathbf{R}_{i,j}\right]. \tag{3.30}$$

Similarly, we can estimate the joint second order information by means of the recurrence matrix

$$\hat{H}_2(\tau) = -\ln\left[\frac{1}{N^2}\sum_{i,j=1}^{N}\mathbf{R}_{i,j}\mathbf{R}_{i+\tau,j+\tau}\right]. \tag{3.31}$$

We compare the estimation of $I_2(\tau)$ based on the G-P algorithm with the one obtained by the RP method in Fig. 3.9. We see, that the RP method yields systematically higher estimates of the mutual information, as in the case of the estimation of the correlation entropy. However, the structure of the curves is qualitatively the same (it is just shifted to higher values by about 0.2). A more exhaustive inspection shows, that the difference is due to the use of the Euclidean norm. The estimate based on the RP method is almost independent of the norm, whereas the estimate based on the G-P algorithm clearly depends on the special choice. If the maximum norm is used (in G-P and RP) both curves coincide.
Note that the estimators for the invariants we propose are different from the ones of the G-P algorithm. Therefore, the obtained values are slightly different, too.

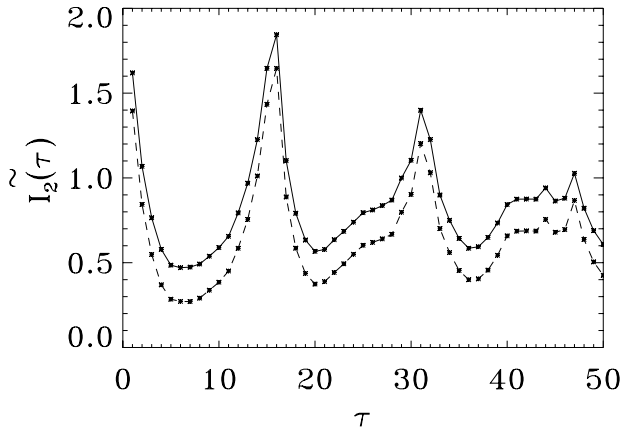

Figure 3.9: Comparison of the estimators of the mutual information for the x-component of the Rössler system computed by the RP method (solid line) and the G-P algorithm (dashed line). The parameters used for the Rössler system and the integration step are the same as in Sec. 3.3.

The three measures that we have proposed, are not only applicable for chaotic systems but also for stochastic ones as the invariants are equally defined for both kinds of systems.

Chapter 4

Influence of Embedding on RPs

4.1 Embedding

In this section we show that delay embedding produces spurious structures in RPs which are not present in the recurrence plot of the real attractor. Typical sets of simulated data are analysed, such as white noise and data from the chaotic Rössler system to show the relevance of this effect.
In the case of experimental data there is often only one component available. Hence, using delay coordinates [Takens, 1980], one can reconstruct the vectors in phase space in the following way. After fixing the embedding dimension d and the delay τ one defines the vectors

$$\vec{x}_i = \left(x_i, x_{i+\tau}, \ldots, x_{i+(d-1)\tau}\right)^{\mathbf{T}}. \tag{4.1}$$

There are some established methods to determine the embedding parameters [Kennel et al., 1992]. To estimate the "optimal" embedding dimension d, the method of false nearest neighbours is generally accepted. The most frequently applied methods for the estimation of τ are based on the autocorrelation function or on the mutual information. However, the choice of a concrete method is still under debate [Grassberger et al., 1991], [Zbilut & Webber, 1992]. Fig. 4.1 presents RPs for our three prototypical examples (uniformly distributed and independent noise, a sine function and the chaotic Rössler system with standard parameters) and different embedding dimensions. The left panel shows the plots for embedding dimension $d = 1$, i.e. no embedding was used. The right panel opposes the same graphics for $d > 1$. These examples make clear that the visual impression that RPs provide can change considerably due to embedding. On the one hand this is expected, as to yield a visualisation of the phase space, the latter one has to be reconstructed. On the other hand we will show in the following sections that delay embedding also produces spurious structures that are not present in the real attractor.

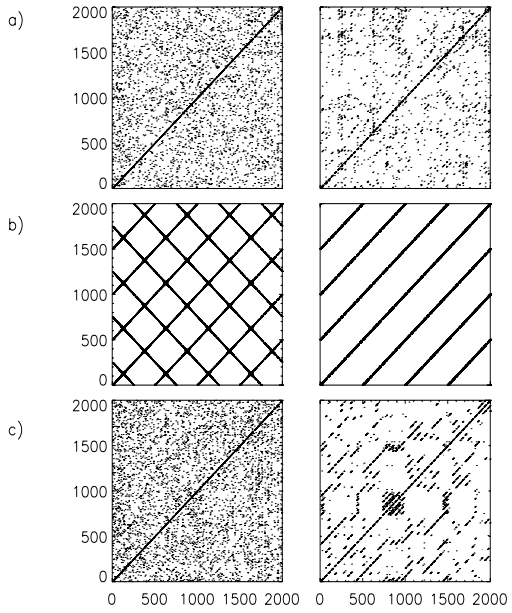

Figure 4.1: RPs for uniformly distributed noise (a), the sine function (b) and the Rössler system (c). The left panel shows the plots for $d = 1$. The right panel represents the plots for $d = 14$, $d = 2$ and $d = 3$, from top to bottom. ε is chosen so that the recurrence rate (Eq. (2.4)) is the same for the embedded and non-embedded time series.

4.2 Correlations Due to Embedding for White Noise

In this section we compute analytically correlations that are induced by the procedure of embedding. Therefore, we apply the method of embedding to independent Gaussian noise – a test process that has no correlations. The correlations we detect afterwards must hence be due to the method of embedding.
Using the embedding dimension d and the delay τ, a vector in phase space is given by

$$\vec{\eta}_i = \sum_{m=0}^{d-1} \eta_{i+m\tau} \vec{e}_m, \tag{4.2}$$

where η_i represents independent Gaussian noise with standard deviation σ and the $\vec{e}_m$ are unit vectors in $\mathcal{R}^d$, i.e. $\vec{e}_m \cdot \vec{e}_n = \delta_{m,n}$. In the RP we have to compute distances of these vectors: $\Delta_{i,j} = |\vec{\eta}_i - \vec{\eta}_j|$. If one moves h steps ahead in time (i.e. on a diagonal in the RP) one finds $\Delta_{i+h,j+h} = |\vec{\eta}_{i+h} - \vec{\eta}_{j+h}|$. Next, we could compute the autocorrelation function of $\Delta_{i,j}$. But to further simplify the calculation we compute the autocorrelation function of $\Delta^2_{i,j}$. This gives by using

the Euclidean norm

$$\mathrm{C}_{\Delta^2}(h, j-i) = \left\langle \left\{ \sum_{m=0}^{d-1} \left[(\eta_{i+m\tau} - \eta_{j+m\tau})^2 \right] - E \right\} \left\{ \sum_{n=0}^{d-1} \left[(\eta_{i+h+n\tau} - \eta_{j+h+n\tau})^2 \right] - E \right\} \right\rangle \tag{4.3}$$

where
$E = \left\langle \sum_{m=0}^{d-1} (\eta_{i+m\tau} - \eta_{j+m\tau})^2 \right\rangle = 2\sigma^2 d\,(1 - \delta_{0,j-i})$ and $\delta_{i,j}$ is the Kronecker delta. $< \cdot >$ denotes the expected value. If we further set $p = j - i$ we can evaluate Eqs. (4.3). Assuming $p > 0$ and $h > 0$ to avoid trivial cases, we find

$$\mathrm{C}_{\Delta^2}(h, p) = \sum_{n=0}^{d-1} (d-n)\,(8\delta_{n\tau,h} + 2\delta_{n\tau,p+h} + \delta_{n\tau,p-h}) \tag{4.4}$$

This formula shows that there are peaks in the correlation function if h and/or $p+h$ and $p-h$ are equal to one of the first $d-1$ multiples of τ. These peaks are not present when embedding is not used. These spurious correlations induced by embedding, lead to modified structures in the RP (see Fig. 4.1 first line, right panel).
In the next section we will investigate numerically the Rössler system, as a paradigmatic and chaotic system.

4.3 Correlations due to Embedding for the Rössler System

In this section we compute numerically the correlation of the squared distances using delay coordinates and the original coordinates (x, y, z) for the Rössler system with standard parameters (Eq. (3.21)). Fig. 4.2 shows results of a simulation for the original coordinates (a) and for the embedded ones using the x-component(b), with $d = 3$ and $\tau = 8$. Obviously, the correlation structure has changed considerably. The correlation structure for the embedded time series does not reflect characteristic patterns of the Rössler system, but it is mainly influenced by the embedding. In this case the overlaying high frequency is given by the inverse of the delay time $\tau = 8$. This difference in the correlation structure strongly influences the structures in RPs that are computed based on delay embedding. Fig. 4.2 (c) shows the conditional probability to find a black point h unit time steps after another black point for the embedded Rössler system. Obviously, the form of the curves of Figs. 4.2 (b) and (c) is qualitatively the same. Our results show that the procedure of delay embedding induces spurious correlations in the

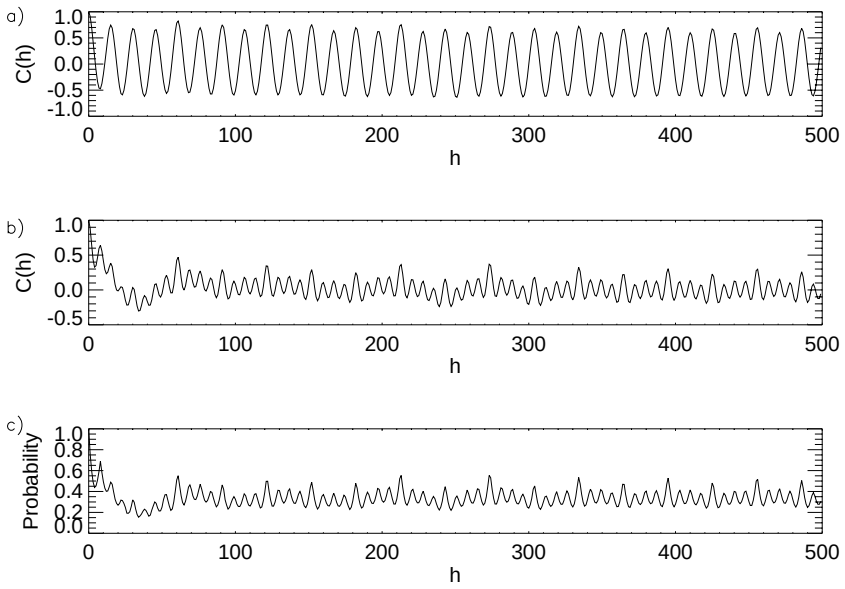

Figure 4.2: (a): Autocorrelation function of the squared distances for the real coordinates of the Rössler system for $p = 10$. (b): Same for the embedded coordinates $d = 3, \tau = 8$. For the embedding the x-component is used. (c): Probability to find a black point in the RP h points after another one. Same parameters as (b).

data and hence in the RPs. This is also reflected in the RQA measures. Tab. 4.1 summarises the results for four typical RQA measures: RR (Eq. (2.4)), DET (Eq. (2.5)), RATIO (Eq. (2.9)) and ENTR (Eq. (2.7)).
To compare all cases we have chosen the threshold ε so that RATE ≈ 0.01. The first line of Tab. 4.1 summarises the results for the real coordinates, the rest of the lines for different embedding parameters. Note, that the measures yield rather different values depending on the embedding parameters.
As we will see in the next section K_2 and D_2 are independent of the embedding parameters which are used. This makes them rather appropriate for the analysis of observed data.

4.4 Independence of the Embedding

In this section, we show that the correlation entropy K_2 and correlation dimension D_2 can be estimated independently from the embedding parameters used. They can indeed be estimated even if no embedding is used at all.
We start our consideration with the cumulative distribution of diagonal lines

embed. param.	DET	RATIO	ENTR
"real coord."	0.95	96.3	0.0067
$d = 1$	0.04	3.58	0.0670
$d = 3, \tau = 8$	0.72	67.1	0.0169
$d = 3, \tau = 6$	0.83	76.4	0.0142
$d = 3, \tau = 25$	0.50	46.5	0.0288
$d = 6, \tau = 8$	0.75	68.7	0.0129

Table 4.1: Comparison of four RQA measures calculated for the Rössler system and for different embedding parameters. In all cases the x-component is used for the embedding.

$P_\varepsilon^c(l)$ (Eq. 3.13) in the RP:

$$P_\varepsilon^c(l) = \frac{1}{N} \sum_{s=1}^{N} \prod_{m=0}^{l-1} \mathbf{R}_{t+m,s+m} = \frac{1}{N} \sum_{s=1}^{N} \prod_{m=0}^{l-1} \Theta\left(\varepsilon - |\vec{x}_{t+m} - \vec{x}_{s+m}|\right) \tag{4.5}$$

To make clear that we use the reconstructed phase space - as we only have observed one component - we write for the cumulative distribution $P_{\varepsilon,\tau}^d$ from now on, where d is the embedding dimension and τ the delay used for the reconstruction. Choosing the maximum norm in Eq. (4.5), $P_{\varepsilon,\tau}^d(l)$ reads

$$P_{\varepsilon,\tau}^d(l) = \frac{1}{N^2} \sum_{i,j=1}^{N} \prod_{m=0}^{l-1} \Theta\left(\varepsilon - \max_{k=0,\ldots,d-1} |x_{i+m+k\tau} - x_{j+m+k\tau}|\right), \tag{4.6}$$

with the embedding dimension d and the delay τ. Obviously,

$$\begin{aligned} &\prod_{m=0}^{l-1} \Theta\left(\varepsilon - \max_{k=0,\ldots,d-1} |x_{i+m+k\tau} - x_{j+m+k\tau}|\right) \\ &= \Theta\left(\varepsilon - \max_{\substack{m=0,\cdots,l-1 \\ k=0,\ldots,d-1}} |x_{i+m+k\tau} - x_{j+m+k\tau}|\right) \end{aligned} \tag{4.7}$$

holds. Eq. (4.7) can be interpreted as testing if the conditions

$$|x_{i+m+k\tau} - x_{j+m+k\tau}| < \varepsilon \qquad \forall \begin{matrix} m = 0, \cdots, l-1 \\ k = 0, \ldots, d-1 \end{matrix} \tag{4.8}$$

are met. The terms on both sides of Eq. (4.7) are one if all conditions are simultaneously met and zero otherwise. Hence, $P_{\varepsilon,\tau}^d(l)$ in Eq. (4.6) can be interpreted as an estimate of the probability that all the conditions Eq. (4.8) are simultaneously met, i.e. the probability to find a diagonal line of at least length l when the embedding dimension d and the threshold ε are used.

Note that Eqs. (4.8) are a set of $l \cdot d$ conditions that are in general not independent. If for example $m + k\tau = m' + k'\tau$, one of the two conditions

$$\begin{aligned} \varepsilon &> |x_{i+m+k\tau} - x_{j+m+k\tau}| \qquad \text{or} \\ \varepsilon &> |x_{i+m'+k'\tau} - x_{j+m'+k'\tau}| \end{aligned}$$

is redundant. If l is sufficiently large, i.e. $l > \tau$, we can condense the conditions Eqs. (4.8) and find the $l + (d-1)\tau$ relations

$$|x_{i+m} - x_{j+m}| < \varepsilon \quad \forall\, m = 0, \ldots, l - 1 + (d-1)\tau. \tag{4.9}$$

These conditions have to be met to find a line of at least length $l - 1 + (d-1)\tau$ if the time series x_i is not embedded, i.e.

$$P_\varepsilon^1(l - 1 + (d-1)\tau) = \frac{1}{N^2} \sum_{i,j=1}^{N} \prod_{m=0}^{l-1+(d-1)\tau} \Theta\left(\varepsilon - |x_{i+m+k\tau} - x_{j+m+k\tau}|\right) \tag{4.10}$$

Note, that the further condition $l > \tau$ has to be met. More general, one finds

$$P_{\varepsilon,\tau}^d(l) = P_{\varepsilon,\tau}^1(l - 1 + (d-1)\tau) = P_{\varepsilon,\tau}^1\Big(\underbrace{[l \mp \Delta d]}_{l'} - 1 + (\underbrace{[d \pm \Delta d]}_{d'} - 1)\tau \Big) = P_{\varepsilon,\tau}^{d'}(l') \tag{4.11}$$

provided that $l, l' > \tau$ and $d, d' \geq 1$. Eqs. (4.11) show that the decay of $P_{\varepsilon,\tau}^d(l)$ is essentially the same for different embedding dimensions and delays. The curve is only shifted to larger l's if the dimension is decreased. The condition for $P_{\varepsilon,\tau}^d(l) = P_{\varepsilon,\tau}^{d'}(l')$ is

$$l + (d-1)\tau = l' + (d'-1)\tau'. \tag{4.12}$$

Note, that this relation only holds due to the special choice of the maximum norm. As the estimator of K_2 (Eq. (3.14))

$$\hat{K}_2(\varepsilon, l) = -\frac{1}{l\tau} \ln P_\varepsilon^c(l) \approx -\frac{1}{l\tau} \ln P_{\varepsilon,\tau}^d(l) \tag{4.13}$$

is based on the slope of $\ln P_\varepsilon^c(l)$ resp. $\ln P_{\varepsilon,\tau}^d(l)$ for large l, this estimation is independent from the embedding parameters. By a similar argument, we show that the estimator of the correlation dimension by means of recurrences is also independent from the embedding parameters. Eq. (3.26) proposed the following estimator for D_2:

$$\hat{D}_2(\varepsilon) = \ln\left(\frac{P_\varepsilon^c(l)}{P_{\varepsilon+\Delta\varepsilon}^c(l)}\right) \left(\ln\left(\frac{\varepsilon}{\varepsilon + \Delta\varepsilon}\right)\right)^{-1}, \tag{4.14}$$

where $P^c_\varepsilon(l)$ is the cumulative distribution of the real coordinates. Substituting

$$\begin{aligned} \ln\left(P^d_{\varepsilon_1,\tau}(l)\right) - \ln\left(P^d_{\varepsilon_2,\tau}(l)\right) &= \ln(P^{d'}_{\varepsilon_1,\tau}(l')) - \ln(P^{d'}_{\varepsilon_2,\tau}(l')) \\ \text{i.e.} \qquad \ln\left(\frac{P^d_{\varepsilon_1,\tau}(l)}{P^d_{\varepsilon_2,\tau}(l)}\right) &= \ln\left(\frac{P^{d'}_{\varepsilon_1,\tau}(l')}{P^{d'}_{\varepsilon_2,\tau}(l')}\right) \end{aligned}$$

which follows from Eq. (4.11) in Eq. (4.14), we see that the estimate D_2 is independent of the choice of d and τ.

4.5 The Rössler System

Next we apply the algorithm to the prototypical Rössler system (Eqs. (3.21)). Tab.4.2 summarises the results obtained for the estimates of D_2 and K_2 based on the RP method for original and embedded coordinates, and estimated by the Grassberger-Procaccia algorithm. The values for the correlation entropy are identical within the error bounds whereas the estimates of the correlation dimension are slightly higher for the RP method than for the G-P algorithm.

system	K_2	D_2
"original coordinates"	0.0675 ± 0.004	2.07 ± 0.01
Embedding, x-component, arbitrary d, τ	0.067 ± 0.007	2.06 ± 0.06
G-P algorithm	0.070 ± 0.003	1.81 ± 0.02

Table 4.2: Estimates of K_2 and D_2 for the Rössler system (Eqs. (3.21)) estimated by RPs based on the original coordinates, on the embedded ones for arbitrary (!) embedding parameters and the same values estimated by the Grassberger-Procaccia algorithm (embedding dimensions $d = 5$ to $d = 25$).

The independence of the embedding dimension d and the delay τ is a very important point for the analysis of observed time series. Even though the embedding dimension d and the delay τ may be difficult to determine, the slope of $P^d_{\varepsilon,\tau}(l)$ for large l will be independent of the special choice of the parameters. Hence, the estimates for K_2 and D_2 do not depend on these parameters. This is illustrated in Fig. 4.3 for the Rössler system. The graph shows the number of diagonals $N^c_\varepsilon(l) = N^2 \cdot P^c_\varepsilon(l)$ in a logarithmical scale for different embedding parameters (see caption). K_2 is given by the slope of $N^c_\varepsilon(l)$ for large l. For $l > 100$ the graphs are (approximately) parallel and so the estimate of K_2 is independent of the special choice of the embedding parameters in accordance with Eq. (4.11).
This independence of d and τ is a remarkable fact, as this means that these quantities are not only invariant under delay embedding, i.e. when the (minimal) embedding dimension $d_{\min}$ is sufficiently large, but they can even be estimated without any embedding. For $d_{\min}$ an lower and an upper limit is known. If the

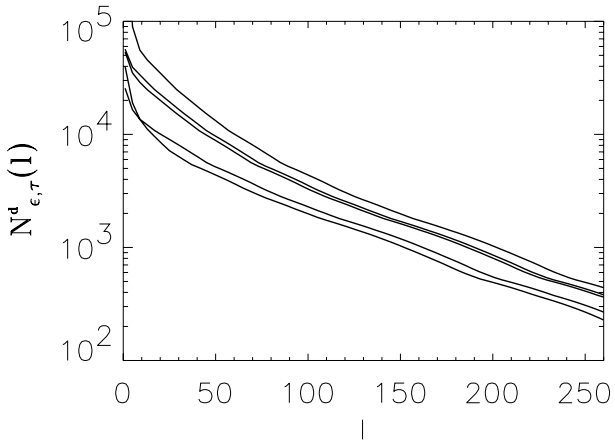

Figure 4.3: Comparison of the cumulative number of diagonal lines of length l ($N^d_{\varepsilon,\tau} = N^2 \cdot P^d_{\varepsilon,\tau}(l)$) for the Rössler system for different embedding parameters ($\{d = 1\}$, $\{d = 3, \tau = 6\}$, $\{d = 3, \tau = 8\}$, $\{d = 6, \tau = 8\}$ and $\{d = 3, \tau = 25\}$ from top to bottom). For all cases the x-component was used for the embedding.

underlying dynamical system has a dimension D, i.e. the number of coordinates in the original phase space is D, then

$$D \leq d_{\min} \leq 2D_F + 1 \tag{4.15}$$

must hold. (In principle this also sets a lower limit for the line length which has to be considered in the RP based method. However, usually much longer line lengths are considered, so that this condition is always fulfilled.)
Takens' theorem states that if a dimension $d \geq d_{\min}$ is used for the delay embedding, the attractor can be reconstructed and the dynamical invariants can be estimated. Even more difficult than fixing the dimension is the choice of τ, as mentioned above. However Eq. (4.12) shows that $P^d_{\varepsilon,\tau}(l)$ is up to a shift in l independent of the special choice of both embedding parameters.
Our argument shows that at least some of the dynamical invariants, e.g. K_2 and D_2, can be estimated without any embedding. This is perhaps the most important advantage of using RPs in the analysis of measured data.

4.6 The Mackey Glass System

The independence of the estimates of the correlation dimension and entropy are of special advantage for the analysis of infinite dimensional systems. A dynamical system is infinite dimensional if an infinite set of independent numbers is required to specify an initial condition, i.e. its phase space dimension is infinite. The

Mackey-Glass equation,

$$\frac{d}{dt}x(t) = \frac{ax(t-\delta)}{1+x^c(t-\delta)} - bx(t), \tag{4.16}$$

is an example of such a system [Mackey & Glass, 1977], [Farmer, 1982]. It is used as a model for the investigation of blood production. In this study we set $a = 0.2$, $b = 0.1$, $c = 10$ and $\delta = 17$. The integration step is $h = 0.01$. Fig. 4.4 (a) shows a section of the time series. Fig. 4.4 (b) represents a phase portrait of the attractor. To estimate the correlation dimension and entropy one would

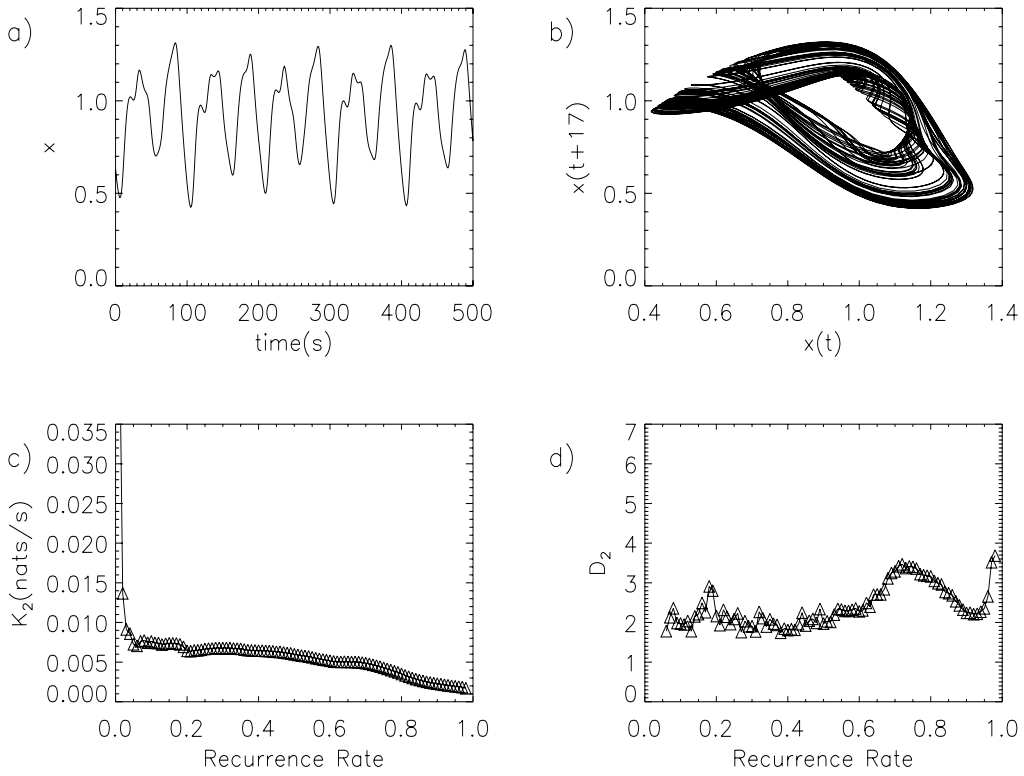

Figure 4.4: (a) Sample of the time series of the Mackey-Glass system for $\delta = 17$ (chaotic regime), (b) Reconstructed attractor, (c) Correlation entropy in dependence on the recurrence rate, (d) Correlation dimension in dependence on the recurrence rate.

usually have to embed the time series $x(t)$. Therefore, one has to determine the embedding dimension. The embedding dimension is not necessarily infinite. Mathematically speaking, an embedding is a smooth map $f : X \longrightarrow Y$ that is a diffeomorphism from a smooth manifold X to a smooth submanifold Y. The embedding dimension d is then defined as the minimum dimension of a subset

of Euclidean space into which a smooth manifold containing the attractor can be "embedded", i.e. d variables are sufficient to uniquely specify a point on the attractor.
As mentioned in the preceding section, the determination of the embedding dimension can be problematic for observed time series. However, since the RP based method for the estimation of K_2 and D_2 is independent of the embedding parameters, we do not use any embedding in this study.
Fig. 4.4 (c) shows the estimates of K_2 for different recurrence rate. We have used an algorithm that computes the threshold ε for a fixed RR (Sec. 5.1). This has the advantage that the values of RR are normalised. The plateau allows us to estimate the correlation entropy for the Mackey Glass system to be $K_2 = 6.66 \cdot 10^{-3} \pm 10^{-4}$nats/s. This result is in good accordance with values reported in the literature for its Lyapunov exponents $[0.007, 0, -0.071, -0.15, ...]$ [Farmer, 1982], since the correlation entropy is numerically close to the sum of the positive Lyapunov exponents. The metric entropy for non-hyperbolic systems is approximately the sum of the positive Lyapunov exponents (slightly less) and the metric entropy is an upper bound for the correlation entropy [Beck & Schlögl, 1993].
The RP based estimate for the correlation dimension is $D_2 = 2.13 \pm 0.03$. Farmer computes $D_F = 2.13 \pm 0.03$ for the fractal dimension and $D_{KY} = 2.10 \pm 0.02$ for the Kaplan-Yorke dimension, again in very good accordance with our results [Mackey & Glass, 1977].
Our estimates were obtained from the scalar time series without any embedding. However, if we had used embedding we would have obtained the same estimates for all choices of the embedding parameters.

4.7 Flow Data

In this section we estimate invariants from the RP obtained from an analysis of some fluid flow data. The experiment consists of a rotating, differentially-heated cylindrical annulus, in which a fluid (a water-glycerol mixture) is contained within the annular gap between two coaxial, brass cylinders and horizontal, thermally-insulating base and lid. The apparatus is rotated uniformly about its vertical axis of symmetry, and motion is driven by differential heating of the cylindrical sidewalls. Further details may be found in [Read et al., 1992]. The time series consisted of temperatures measured in the fluid at intervals of 2 s for periods of up to 8 $\times 10^4$ s, and were obtained from copper-constantan thermocouples on fine-wire probes located at mid-height and mid-radius in the convection chamber. The flows measured were in the baroclinically unstable regime, and took the form of azimuthally-propagating travelling waves with various quasi-periodic or chaotic modulations. The particular time series investigated here were taken from a single thermocouple probe for cases (ii) and (iii) of [Read et al., 1992]; case

(iii) was identified as a quasi-periodic amplitude-modulated wave number $m = 3$ flow, while case (ii) was identified as a low-dimensional chaotically-modulated wave with both $m = 3$ and $m = 2$ present. Fig. 4.5 (a) shows a section of the time series from case (ii) and Fig. 4.5 (b) the reconstructed attractor. To estimate K_2 we compute the distribution of diagonals $P_\varepsilon^c(l)$ and determine the slope for different thresholds ε (respectively versus the recurrence rate (Sec. 5.1)) in the second scaling region. Fig. 4.5 (c) shows the result. The plateau allows us to estimate K_2 reasonably well.
Based on Eq. (4.14) we then estimate D_2. Fig. 4.5 (d) represents the outcome for different recurrence rates. We obtain for the correlation entropy $K_2 =$

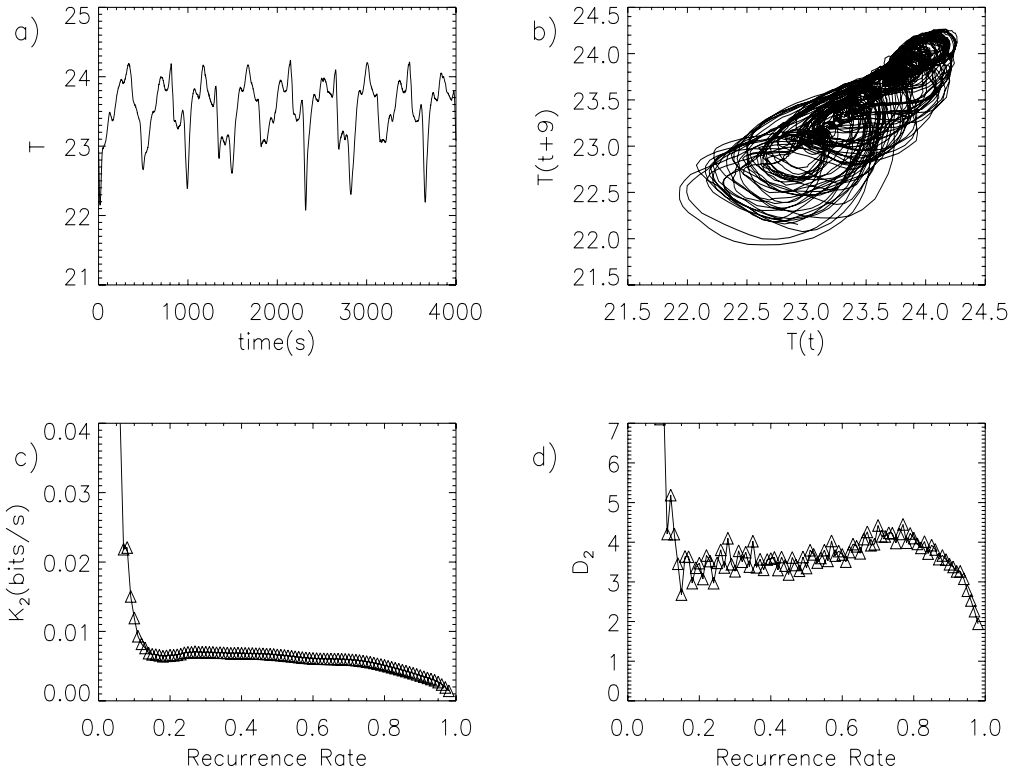

Figure 4.5: (a) Sample of the time series for the flow in chaotic regime, (b) Reconstructed attractor, (c) dependence of Rényi entropy of second order on the recurrence rate, (d) dependence of correlation dimension on the recurrence rate.

$6.7 \cdot 10^{-3} \frac{\text{bits}}{\text{s}} \pm 2.4 \cdot 10^{-4}$bits/s and for the correlation dimension $D_2 = 3.4 \pm 0.4$. Our results are in accordance with previous results [Read et al., 1992] and indicate low dimensional chaos. Note that our results were obtained from 40,000 data points without filtering and without embedding.
We have also performed the analogous computation based on the Grassberger

Procaccia algorithm (Fig.4.6) and found that the CPU time needed is about one or two orders of magnitude higher. The values we have obtained for the correlation entropy base on the G-P algorithm are $K_2 = 1.1 \cdot 10^{-3} \pm 5.0 \cdot 10^{-4}$bits/s and for the correlation dimension $D_2 = 4.1 \pm 0.4$. The estimation of the correlation

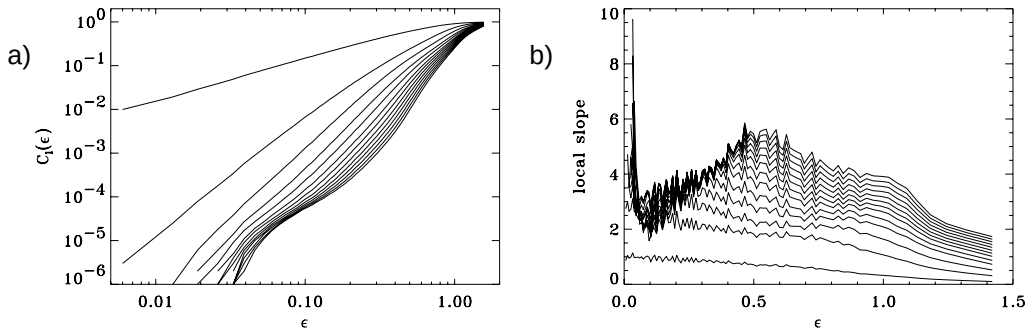

Figure 4.6: (a): Correlation integral for the flow data. The estimation of the invariants is problematic in this representation as there is no clear straight line for large embedding dimensions. (b): Local slope of the curves represented in (a).

dimension and the entropy is rather problematic as the scaling is not very well pronounced for large embedding dimensions [Kantz & Schreiber, 1997]. The estimation based on the RP method is more robust.
In the quasiperiodic case (iii) we obtain with the RP method for the correlation entropy $K_2 = 4.59 \cdot 10^{-4} \pm 2.2 \cdot 10^{-5} \frac{\text{bits}}{\text{s}}$ and for the correlation dimension $D_2 = 2.27 \pm 0.28$ also in very good agreement with previously reported results [Read et al., 1992].

Chapter 5

Automatisation of the Algorithm to Estimate K_2

5.1 Automatisation of the Algorithm

For many applications, e.g. if spacio-temporal data is to be analysed, it is desirable to automate the algorithm to estimate K_2 based on RPs. Such an automated algorithm is also more objective, as otherwise the choice of the proper scaling regions of $P_\varepsilon^c(l)$ depends to some extent on the choice of the data analyst.
For the practical application, one has first to compute the cumulative distribution of diagonals $P_\varepsilon^c(l)$ resp. $P_{\tau,\varepsilon}^d(l)$ for different thresholds ε. The question arises, which values of ε one should consider. As each system has its proper amplitude, which may differ from one system to another one, the choice will be different for each case and it is subjected to some arbitrariness. To overcome this problem, we can fix the value of RR, because it is normalised, and then calculate the corresponding ε. This can be done by the following algorithm:

1. We compute the distances between each pair of vectors $i = 1, \ldots, N$ and $j = 1, \ldots, i$. Then we obtain the series d_l with $l = 1, \ldots, N^2/2$ (because of the symmetry of the RP, we consider only the half of the matrix. Actually the length of the series of the distances is equal to $N^2/2 - N$, but for large N, we can write $N^2/2$.).

2. We sort the distances d_l in ascending way and denote the rank ordered distances by $\tilde{d}_l$, with $l = 1, \ldots, N^2/2$.

3. For a fixed RR (Eq. 2.4) the corresponding ε is then given by $\tilde{d}_m$, with $m = RR\frac{N^2}{2}$. For example, if $RR = 0.01$, then $\varepsilon = \tilde{d}_{0.01N^2/2}$. We then know that 1% of the distances are less or equal than ε, and hence $RR = 0.01$.

Like this, we avoid the arbitrariness of choosing appropriate values for ε and we can apply the same procedure for all systems. Our simulations have shown that

this way to choose the thresholds has the convenient by effect that the scaling regions become broader.
The next step is crucial for the automatisation: the scaling region of $\ln P_{RR}(l)$ vs. l and the plateau in $K_2(RR)$ vs. RR must be estimated automatically. In both cases we apply a cluster dissection algorithm [Späth, 1992]. The algorithm divides the set of points into distinct clusters. In each cluster a linear regression is performed. The algorithm minimises the sum of all square residuals in order to determine the scaling region and the plateau. To find both regions automatically, we used the following parameters:

- We consider only diagonal lines up to a fixed length l_{max}. Longer lines are excluded because of finite size effects. Reasonable values of l_{max} are at about 10% of the length of the time series.

- We consider only values of $P_{RR}(l)$ with $P_{RR}(l) > 500$ to obtain a reliable statistic.

- We use about 100 different values for ε, corresponding to 100 equally spaced recurrence rates RR between 1% and 99%, to have a good defined plateau in $K_2(RR)$ vs. RR.

- We in general use $5,000$ to $10,000$ data points of each simulated trajectory. The more data points one uses, the more pronounced the scaling regions. However, especially for maps about 500-1,000 data points are sufficient to obtain reliable estimates. Note, that the computation time increases approximately with N^2.

- We further have to specify the number of clusters when applying the cluster dissection algorithm: for the detection of the scaling region in $\ln P_{RR}(l)$ vs. l, 2 different clusters seem to be a rather good choice (see Sec. 3.3). Then, we use the slope of the largest cluster. For the detection of the plateau in $K_2(RR)$ vs. RR, we chose 3 clusters and use the value of the cluster with the minimum absolute slope.

These choices have proven to be the most appropriate ones for the estimation of the scaling regions. All these parameters are defaults of a computer program. The following two section present results we have obtained for two model systems.

5.2 Shrimps in ODEs

This automated algorithm to estimate K_2 can then be used to scan through the parameter or configuration space of any system. In Chap. 8 we present two examples of its application to evaluate the stability of extrasolar planetary systems on the one hand and the predictability of the earth's surface temperature on the

other hand.
We will start however with an investigation of the parameter space of the Rössler system (Eqs. (3.21)). After integrating this system with a Runge-Kutta integrator of fourth order, we compute, based on the automatisation of the algorithm K_2 in a cut of the parameter space varying $a = b$ on the one hand and c on the other. c is the standard bifurcation parameter, and a and b change mainly the topology of the attractor. We used about 150-250 oscillations of the chaotic oscillator and a total of 5,000 data points for each set of parameters. Fig. 5.1 shows the result. Grey areas mark a divergence of the trajectories. Blue regions are characterised

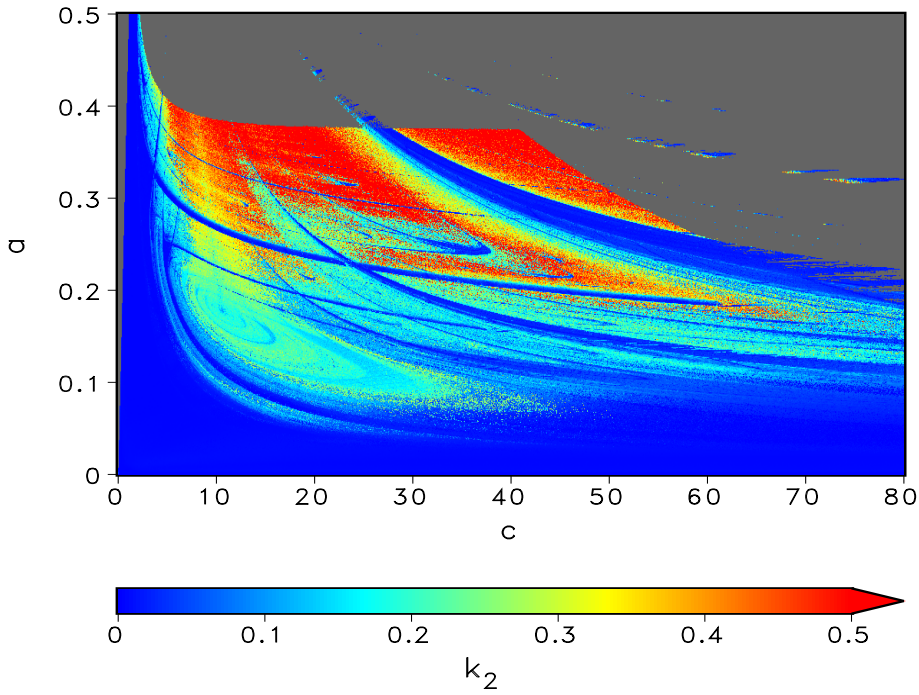

Figure 5.1: Diagram of K_2 for a section in the parameter plane of the Rössler system Eqs. 3.21. One recognises different features such as the vortex shaped structure in the lower left and blue (periodic) bands running through the chaotic regions. The grey region marks parameters which lead to diverging trajectories.

by $K_2 \approx 0$ indicating very regular or periodic behaviour. Red parts of the plot indicate highly chaotic behaviour.
One recognises a fairly complicated structure which seems to display fractal behaviour. In for $a = b \leq 0.25$ and $c \leq 40$ the structure seems to be roughly vortex shaped. Increasing a, b the chaotic regions are riddled with periodic bands. These bands occasionally intersect. Note also, that these bands correspond to periodic windows in the bifurcation diagram. Another structure is found in the region $c \in [20 : 45]$ and $a = b \in [0.2 : 0.3]$, which is represented in Fig. 5.2. In the centre of this box there is a blue swallow like structure, which is called shrimp. In this

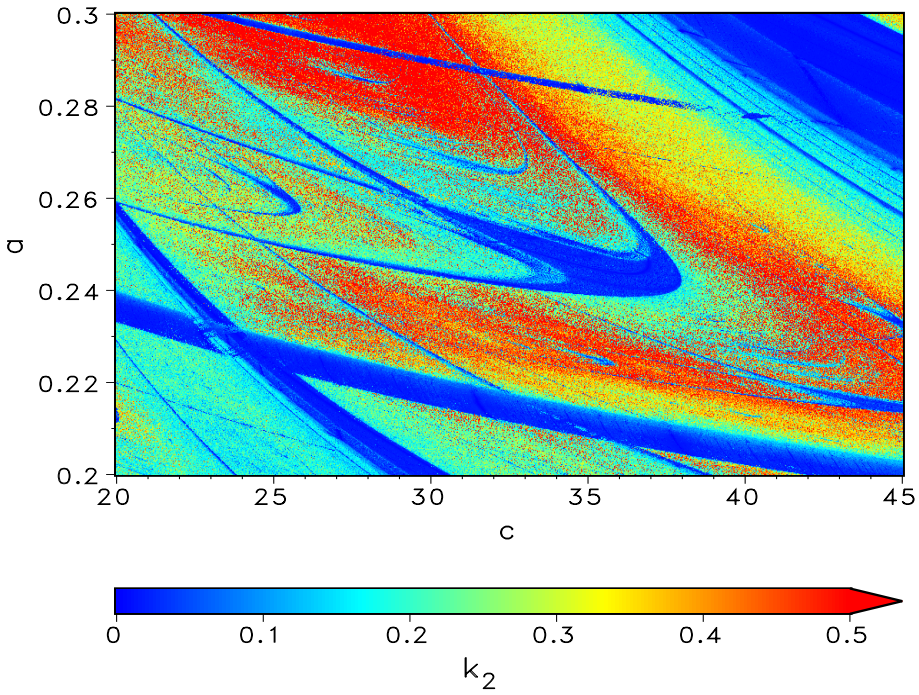

Figure 5.2: Zoom into the parameter plane presented in Fig. 5.1. The blue structure in the centre of the picture is called shrimp.

form it was so far only known for discrete systems [Gallas, 1994]. Its symmetry is determined by ten so called noble points. The shrimp is a unit cell of a family of self similar shrimps. Two shrimps can be recognised between the arms and the tail of the large one. We are still working on a mathematical description of this structure in ODEs.
However, this structure requires some specific characteristics of the underlying dynamical system. The Lorenz system

$$\begin{aligned} \dot{x} &= -\sigma x + \sigma y \\ \dot{y} &= rx - y - xz \\ \dot{z} &= -bz + xy \end{aligned} \tag{5.1}$$

for example does not seem to be endowed with shrimps in its parameter space. This systems describes convection cells in a Rayleigh-Bernard convection. Fig. 5.3 displays K_2 for a section in the parameter space with $b = 8/3$ (b is a measure for the geometry of the convection cell). The relative Rayleigh number r and the Prantl number σ are then varied. The resulting structure also seems to posses fractal characteristics.
Note, that there are similar diagrams ("Lyapunov diagrams") displaying the Lyapunov exponents, estimated from the maps and not from time series as in our case, which show rather similar structures [Markus, 1995]. There, the structures are known to be fat fractals. Due to the analogy, one can also suppose

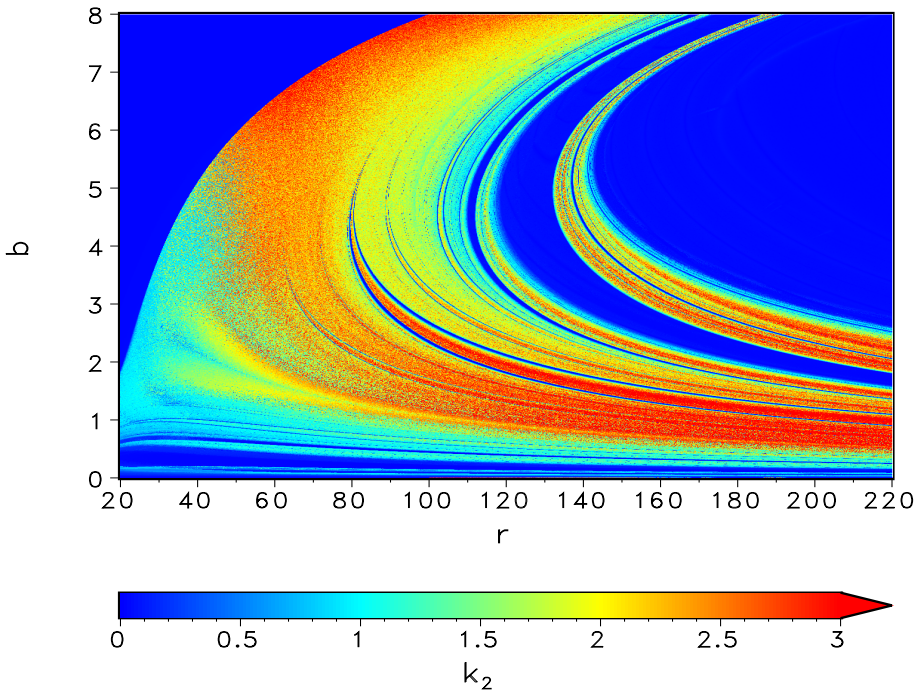

Figure 5.3: K_2 in the parameter space of the Lorenz system.

that the structures induced by K_2 in the parameter plane are also fat fractals [Markus & Tamames, 1996], [Grebogi et al., 1985].

5.3 The Standard Map

In this section we present results for a conservative chaotic system, the standard map

$$\begin{aligned} \varphi_{i+1} &= \varphi_i + p_i - \kappa \sin(\varphi_i) \\ p_{i+1} &= \varphi_{i+1} - \varphi_i \end{aligned} \tag{5.2}$$

This system has a "physical" interpretation [Beck & Schlögl, 1993]. Consider the following differential equation:

$$\frac{d^2\varphi}{dt^2} + \kappa \sin\varphi \sum_{n=0}^{\infty} \delta(t - n\tau) = 0 \tag{5.3}$$

This equation describes a "kicked oscillator", i.e. a plane mathematical pendulum under the influence of a force that is switched on at discrete time points $n\tau$ only. Integrating Eq. (5.3) one finds that for $n\tau < t < (n+1)\tau$ the acceleration $\ddot{\varphi}$ vanishes. Hence, the velocity $\dot{\varphi}$ is a constant, which we call p_{n+1}. It follows that

$$\varphi(t) = (t - n\tau)p_{n+1} + \varphi(m\tau). \tag{5.4}$$

Denoting $\varphi(n\tau) = \varphi_n$, we obtain

$$\varphi_{n+1} = \tau p_{n+1} + \varphi(n\tau) \tag{5.5}$$

To get a recurrence relation for p_n, we integrate Eq. (5.3) from $n\tau - \varepsilon$ to $n\tau + \varepsilon$, where $\pm\varepsilon$ again indicates respectively that the time is taken immediately after and before the kick:

$$\int_{n\tau-\varepsilon}^{n\tau+\varepsilon} \left(\frac{d^2\varphi}{dt^2} + \kappa \sin\varphi_n \right) dt = 0 \qquad (\varepsilon \to 0). \tag{5.6}$$

As $\dot{\varphi}(n\tau + \varepsilon) = p_{n+1}$ and $\dot{\varphi}(n\tau - \varepsilon) = p_n$, we obtain

$$p_{n+1} - p_n + \kappa \sin\varphi_n = 0. \tag{5.7}$$

Substituting Eq. (5.7) in Eq. (5.5) we end up with

$$\varphi_{n+1} = \varphi_n + (p_n - \kappa \sin\varphi_n)\tau. \tag{5.8}$$

Hence, Eq. (5.8) and Eq. (5.5) are equivalent to the standard map for the choice $\tau = 1$. Thus, we can interpret φ_n as the angle, and p_n as the momentum of a kicked rotator.

Conservative chaotic systems, like the kicked rotator, exhibit some features which are different from dissipative systems. The initial conditions (here φ_0 and p_0) determine to a large degree the chaoticity of the trajectories[1]. An interesting feature of this system is, that in contrast to the investigation of the Rössler system one does not have to integrate a trajectory starting at each point of the plot as we consider now the phase space and not the parameter space. In the case of the standard map, we start at a random initial condition and iterate 4,999 times. For this time series we compute then K_2. All the 5,000 points of the time series are characterised by the same K_2. Hence, each initial condition yields 5,000 points in the plot and not only one as in the case of the Rössler system. Note, that the distances in the recurrence metric are computed with respect to the torus the standard map lives on. Numerical results have shown however, that computing the distances on an Euclidian space, has a rather low effect on the picture.

Fig. 5.4 shows the phase space of the standard map for $\kappa = 1.4$. The colours at each point represent K_2 of the trajectory starting at that coordinates. The blue patches are periodic regions. They also have an striking substructure. It is caused by the coaction of several effects. One of which is that due to the finite sampling and the finite length ($N = 5,000$) even periodic time series may "seem" not to be purely periodic. At the transition from periodic to chaotic regions there is a layer of intermediate K_2.

[1] A further example of this class of chaotic processes will be given in Chap. 8, when the stability of extrasolar planetary systems is studied.

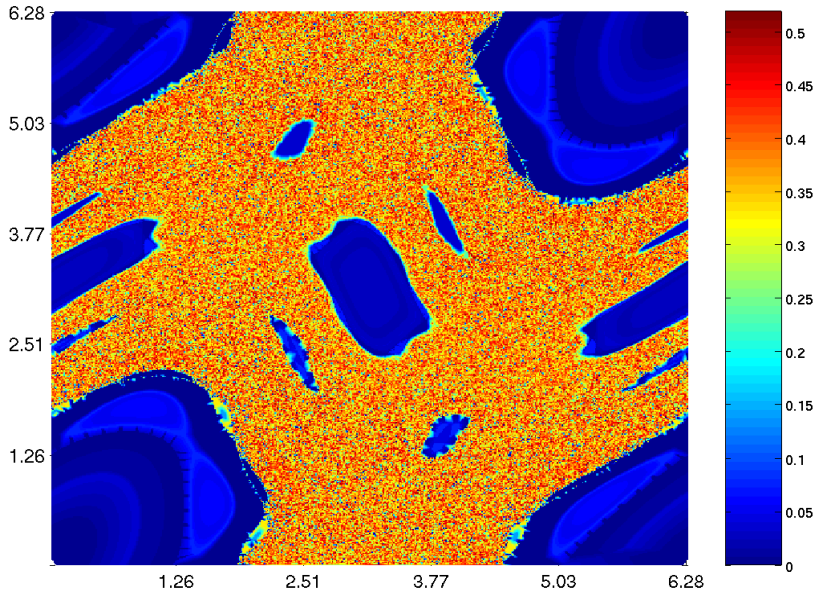

Figure 5.4: K_2 in the phase space of the standard map. The entropy depends on the initial conditions. Blue (periodic) islands are surrounded by chaotic basins.

Chapter 6

How Much Information is Contained in an RP?

6.1 Reconstruction of the Attractor from the Recurrence Plot

In Chap. 3 we have seen that at least some dynamical invariants can be estimated from an RP. The question arises how much information in principle can be extracted from an RP, as the complex behaviour of a system in a d dimensional phase space is mapped onto a binary matrix.
To answer this question we study in this section how to reconstruct a (univariate) time series $\{x_1, x_2, x_3, \ldots, x_n\}$ from a binary (i.e. black and white) RP given by the recurrence matrix $R_{i,j} = \Theta(\varepsilon - |x_i - x_j|)$. Let us further assume that the values of x_i are distributed rather continously between the minimum $x_{\text{Min.}}$ and the maximum $x_{\text{Max.}}$ of the time series with respect to the threshold ε. By this, we mean that there is no subinterval in $[x_{\text{Min.}}; x_{\text{Max.}}]$ of length ε which does not contain at least one point of the time series. In other words, we assume that the following condition is fulfilled:

Condition for the reconstructability
If it is impossible to divide the entries of the time series in two subsets, such that no point of one of the subsets has a neighbour (with respect to its ε-neighbourhood) in the second subset, then the time series is reconstructible.

We will show that then the following algorithm can be used to reconstruct the (univariate) time series from the RP. It is important to note that the so reconstructed time series has the same rank order and the same number of entries as the one underlying the RP. Additionally reconstructing by Takens theorem [Takens, 1980] the attractor from this time series yields a topologically identical object [Bandt et al., 2002].

The algorithm to reconstruct the time series from the RP consists of three main phases and a total of 10 steps.

1. Sorting

   (a) If n columns of the matrix $R_{i,j}$ are identical, store the indices of the columns and cancel $n-1$ of them, so that every column is unique.

   (b) Compute for all pairs of neighbouring points with coordinates i, j (i.e. pairs for which $R_{i,j} = 1$) the number of neighbours of x_j which are not neighbours of x_i. We will call this number $n_{i,j}$.

   (c) There exist exactly two points, say x_{j_1} and x_{j_2}, so that $n_{i,j_{1/2}} = 0 \quad \forall i$. These two points are the maximum and the minimum of the time series.

   (d) Choose one of these two indices as starting point. (We will call this index k.)

2. Iteration

   (a) Denote the last index which has been ranked in the reconstruction of the time series so far by k. If there is a unique minimum in the set of $\{n_{i,k}\}_i$, i.e. there is a $i_{\min}$ so that $n_{i_{\min},k} < n_{i,k} \quad \forall i \neq i_{\min}$, take the point with the index $i_{\min}$ as the second point.

   (b) If there is no unique minimum in the $\{n_{l,k}\}_l$, i.e. there is a set of m indices $\{\mu_1, \ldots, \mu_m\}$, so that $n_{\mu_1,k} = \ldots = n_{\mu_m,k} \leq n_{i,k} \quad \forall i$, choose the minimum $w = \min\{n_{k,\mu_i}\}_i$. The next point for the reconstruction then is x_w.

   (c) Iterate the two steps (a) and (b) of this phase, until all indices are ranked. Then, you have a rank order of the points of the time series $n_i^{\text{ord.}}$ which underlies the RP.

3. Final Reconstruction

   (a) Generate random numbers y_i so that for each entry in the ordered series there is one number. Then rank order these random numbers.

   (b) Generate a time series by putting the value y_i at the position $n_i^{\text{ord.}}$. Then one obtains $y_{n_i^{\text{ord.}}}$.

   (c) Reintroduce at the position of the "identical columns" obtained in step 1 the values of the points at the corresponding indices which remained in the RP.

In the next section we illustrate the application of this algorithm by an easy example.

6.2 Step by Step Reconstruction of a Time Series from an RP

To illustrate how the algorithm works, we apply it to the RP of the following toy time series $\{3, 2, 4, 8, 2, 5, 7, 6, 0, 2\}$. The threshold is chosen $\varepsilon = 2$. The corresponding RP is represented in Fig.6.1. Now we apply the algorithm described

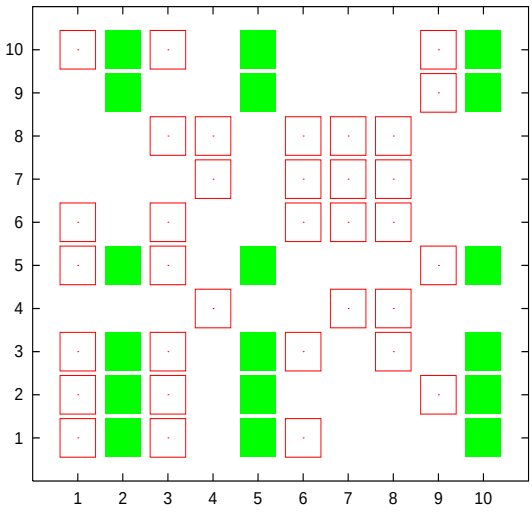

Figure 6.1: RP obtained from a toy time series of ten entries.

in the last section. First, we notice that the columns $2, 5, 10$ are identical (here marked in grey). Following the instruction in step "1" we cancel columns 5 and 10, i.e. we will ignore them carrying out the next steps of the algorithm.
Next, we compute the $n_{i,j}$ for all neighbours.

i	1	1	1	2	2	2	3	3	3	3	4	4	6	6	6	6	7	7	7	8	8	8	8	9
j	2	3	6	1	3	9	1	2	6	8	7	8	1	3	7	8	4	6	8	3	4	6	7	2
$n_{i,j}$	1	1	2	1	2	0	0	1	1	2	1	2	1	1	1	1	0	2	1	2	0	1	0	2

Following step "3" we search the two indices j_1 and j_2 for which $n_{i,j_{1,2}} = 0 \quad \forall i$. We find $j_1 = 9$ and $j_2 = 4$. This means that one of the values x_9 and x_4 is the largest and one is the smallest of the time series. We choose one of them, e.g. $j_1 = 9$ resp. x_9 (step 4).
Then (step 5), we search all neighbours of x_9. There is only one such value x_2. This is the next value in the rank order. Hence, we have $x_9 < x_2$ (or $x_9 > x_2$).
Next, we search all neighbours of x_2. These are x_1, x_3, x_9 and x_9 is already arranged. Hence, we consider $n_{2,1} = 1$ and $n_{2,3} = 2$. The minimum is $n_{2,1} = 1$, i.e. x_1 is next in the series. We have $x_9 < x_2 < x_1$ (or $x_9 > x_2 > x_1$) .

The neighbours of x_1 are x_2, x_3, x_6. x_2 is already arranged in the reconstructed time series. Then, we consider $n_{1,3} = 1$ and $n_{1,6} = 2$. This means that x_3 is the next value in the time series: $x_9 < x_2 < x_1 < x_3$ (or $x_9 > x_2 > x_1 > x_3$).
Proceeding with the neighbours of x_3, we find $n_{3,6} = 1$ and $n_{3,8} = 2$. Hence, we have $x_9 < x_2 < x_1 < x_3 < x_6$ (or $x_9 > x_2 > x_1 > x_3 > x_6$).
The next step is different from the last steps. The two relevant neighbours of x_6 are x_7 and x_8. As $n_{6,7} = 1$ and $n_{6,8} = 1$, we have to apply step "6" of the algorithm and consider $n_{7,6} = 2$ and $n_{8,6} = 1$. The minimum is $n_{8,6} = 1$. Hence, we get $x_9 < x_2 < x_1 < x_3 < x_6 < x_8$ (or $x_9 > x_2 > x_1 > x_3 > x_6 > x_8$).
In the next iteration we consider $n_{8,4} = 0$ and $n_{8,7} = 0$. Again, we must follow step "6". The minimum of $n_{4,8} = 2$ and $n_{7,8} = 1$ is $n_{7,8}$. Hence, the next point is x_7 and the last one x_4. We obtain the time series $x_9 < x_2 < x_1 < x_3 < x_6 < x_8 < x_7 < x_4$ (or $x_9 > x_2 > x_1 > x_3 > x_6 > x_8 > x_7 > x_4$).
Next, one generates 9 random numbers and orders them. Then x_9 is identified with the smallest number, x_2 with the smallest but one and so on.
Finally, one reintroduces x_5 and x_{10} which have been cancelled in the first step. Both are set equal to x_2.
Then, the algorithm is finished and one has a time series $\{x_i\}$ which is reconstructed from the RP.

6.3 Reconstruction of the Time Series for Three Prototypical Systems from Their Respective RPs

Now we show that the algorithm works well for very different systems, both dynamical and stochastic ones. The algorithm does not depend on the dynamics of the underlying system.
A) We start with the reconstruction of the time series of a logistic map $x_{n+1} = 4x_n(1-x_n)$ (Fig. 6.2 (a)). Given only the RP from a simulation of the logistic map (in this case we used $\varepsilon = 0.1$, length of the time series $N = 1,000$), we reconstruct the time series based on the upper algorithm. The result of the reconstruction is displayed in Fig. 6.2 (b). Plotting the original and the reconstructed time series one on top of the other, one obtains Fig. 6.2 (c). Note, that the reconstruction is much more precise than the error bounds given by the threshold ε. Actually, the precision increases with the length of the underlying time series. If the length of the time series increases and the distribution of the values is assumed to be known, the errors of the construction tend to zero. The main dynamical properties of the time series (e.g. the correlation entropy and dimension) are hence captured. This can also be seen in an x_n vs. x_{n+1} plot (Fig. 6.3). Plotting this diagram for the original time series yields a parabola. For the reconstructed time series one finds a graph which seems to be a continously deformed parabola. This effect is

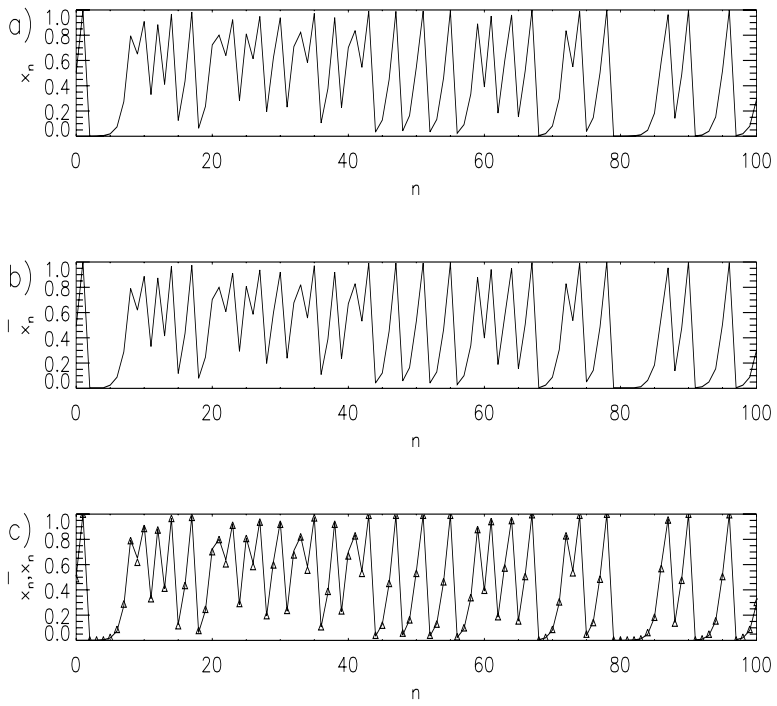

Figure 6.2: Reconstruction of a time series of a logistic map from an RP. (a) Original time series, (b) reconstructed time series, (c) both time series: original one (line) reconstructed one (points).

mainly due to the different distribution of the original and the reconstructed time series. However, it is still a one humped map and also its correlation entropy is unchanged.
B) The reconstruction also works for continuous systems, such as the Rössler system (Eqs. (3.21)). Fig. 6.4 (a) shows a section of the time series. The reconstruction coincides nearly perfectly with the original time series (see Fig. 6.4 (c)). If we assume the distribution of the values of the original time series is known, the mean square error of the reconstructed and the original time series vanishes in this case.
C) The next case we want to present is independent uniformly distributed noise. This system is not dynamical, but the reconstruction of the trajectory based on the algorithm still works. Fig. 6.5 (a) represents the original time series. Fig. 6.5 (b) and (c) represent the reconstructed time series. There is nearly perfect coincidence.
In all cases the algorithm succeeded in reconstructing the time series from the RP. In the next section we will discuss more of the characteristics and limitations of the algorithm.

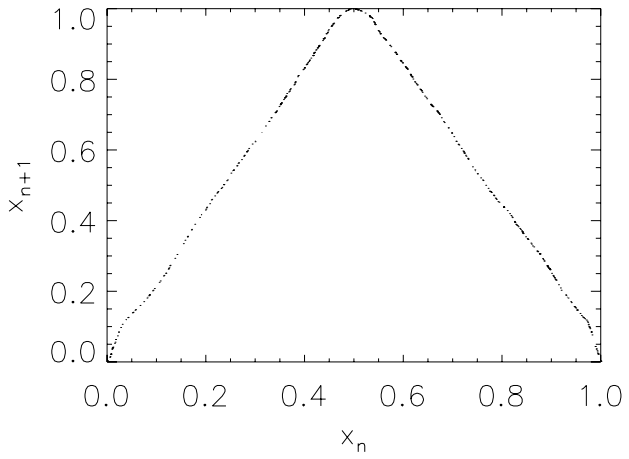

Figure 6.3: x_n vs. x_{n+1} diagram for the reconstructed logistic map. Due to the different distribution of the time series it is not a parabola.

6.4 Discussion of the algorithm

The reconstruction algorithm works by considering the neighbourhoods of the points of the time series. The **condition for the reconstructability** assures, that the neighbourhoods overlap sufficiently. This makes it possible to reconstruct the time series. It is equivalent to the condition that in the projection of the values there is no ε-interval void of points.
Assuming that the values are uniformly distributed, one can estimate the number of points which are needed to reconstruct the time series for a given threshold ε. The distance d of two neighbouring points in the interval of the values $[x_{\min}, x_{\max}]$ is then exponentially distributed

$$\rho(d) = N \cdot e^{-N*d}, \tag{6.1}$$

where N is the length of the time series. Let us, without loss of generality, assume that the interval in which the values are distributed is the unit interval. Then there are $N + 1$ intervals, which have to be all smaller than ε. One obtains the following relation between the number N of points in the time series, the threshold ε and the probability to find a void interval which is larger than the threshold p.

$$p = \left(1 - e^{-N\varepsilon}\right)^{N+1} \tag{6.2}$$

This formula allows to estimate that, in order to be able (with a probability of about 0.999) to reconstruct the time series (and the attractor) for an RP which has a recurrence rate of 1%, one should have more than about 1,400 points in

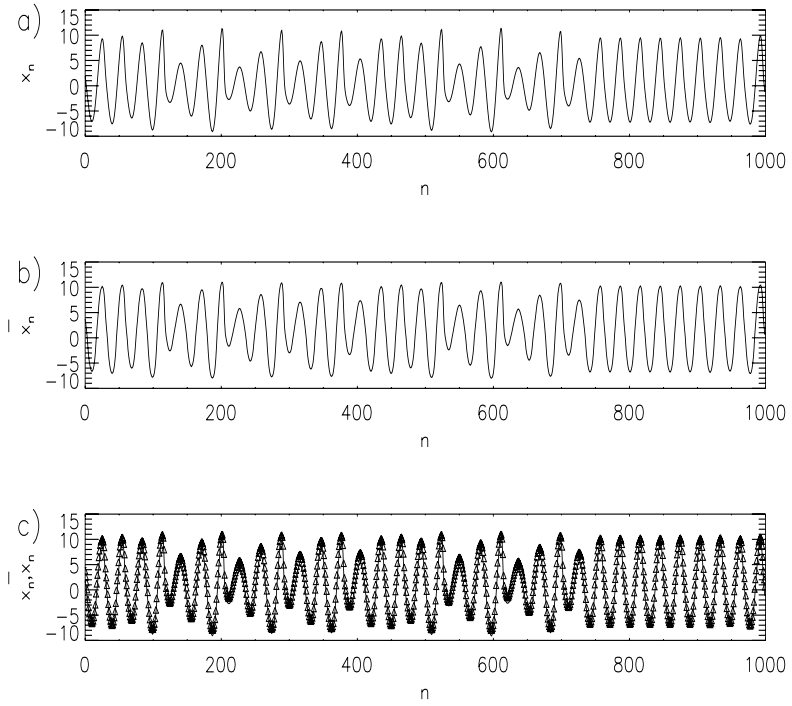

Figure 6.4: Reconstruction of a time series of the x-component of the Rössler system an RP. (a) Original time series, (b) reconstructed time series (for better comparability the standard deviation and the mean have been adapted), (c) both time series: original one (line) reconstructed one (points).

the time series. Using $\varepsilon = 0.1$ one only needs about 90 points to reconstruct the time series.
Hence, if ε is larger, one needs less points for the algorithm to reconstruct the time series. If, on the other hand, ε is too large, the reconstruction algorithm works but cannot distinguish different points properly. Let us take example C of values which are distributed uniformly in the unit interval. Then, if ε is $0.5 + \delta$, a band of width 2δ around 0.5 has all points of the time series as neighbours. All these points have equal columns in the RP and are not distinguishable. (They are "cancelled" in the first step of the algorithm.) Whenever $\varepsilon < 0.5$, i.e. half the interval width of the value of the time series, the time series can be reconstructed as accurately as one wishes by considering sufficiently long time series.
Note, that based on Takens theorem [Takens, 1980], it is possible to reconstruct the attractor from the reconstructed time series. Hence, the attractor can be recovered from the RP (of only one component) of the system, at least topologically [Bandt et al., 2002]. This means, that the RP contains all topological information of the underlying system, even though it is only computed from one

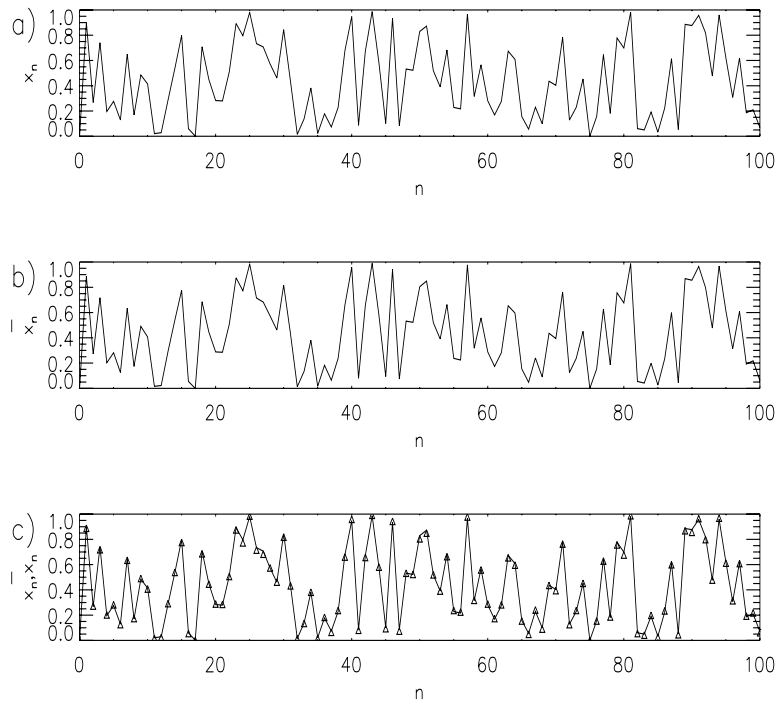

Figure 6.5: Reconstruction of a time series of independent uniformly distributed noise. (a) Original time series, (b) reconstructed time series, (c) both time series: original one (line) reconstructed one (points).

of its components. Fig. 6.6 shows the reconstruction of the Rössler attractor from the reconstructed time series. However, the reconstruction of the time series from the RP of more than one component of the system, e.g. the three dimensional vectors of the Rössler system, is not possible with this algorithm. Such a plot is the pointwise product of the RP of the single components, if one uses the maximum norm in Eq. (2.1). Hence, one loses information. The open question is, if it is possible to reconstruct the attractor from such a multidimensional RP. This is an important problem as the RP of only one component (i.e. the projection of the attractor onto one coordinate axis) contains seemingly less information than the whole n-dimensional phase space. It even does not represent real recurrences but due to the projection also false ones ("false nearest neighbours").
However, in the context of contact maps (see [Baldi & Pollastri, 2003] and references therein) where RP matrices in three dimensions are used, algorithms to reconstruct the "attractor" (or rather the three dimensional structure of a protein) at least approximately are known.

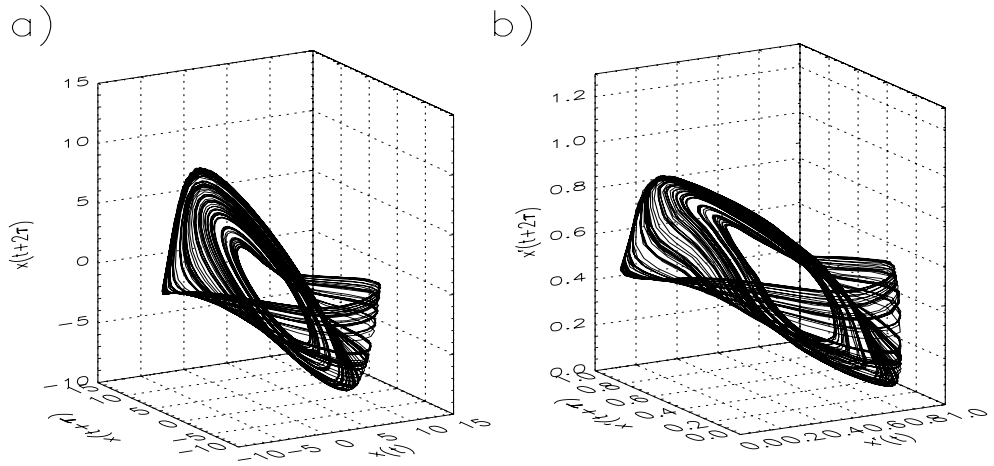

Figure 6.6: (a) Reconstruction of the phase space by Takens Theorem, based on the x-component of the Rössler system (b) Reconstruction of the attractor of the Rössler system from the RP of a time series of its x-component.

Chapter 7

Surrogates based on Recurrences

7.1 The Concept of Complex Synchronisation

The study of synchronisation goes back to the 17th century and begins with the analysis of synchronisation of nonlinear periodic systems. Well known examples are the synchronisation of two pendulum clocks that hang on the same wooden beam (it was through this system, that Huygens discovered synchronisation), the synchronised flashing of fireflies, or the peculiarities of adjacent organ pipes which can almost annihilate each other or speak in unison. But the research of chaotic synchronisation does not begin until the eighties [Fujisaka & Yamada, 1983], [Afraimovich et al., 1986], [Pikovsky, 1984], when it was shown that two chaotic systems can become completely synchronised, i.e. their time evolution becomes identical. This finding has had very important consequences for the design of secure communication devices [Sousa Veiera et al., 1992], [Yuan Zhao, 1983], [Itoh et al., 1999]. The synchronised chaotic trajectories can be used to mask messages and prevent their interception. In [Afraimovich et al., 1986] and later in [Rulkov et al., 1995] the notion of complete synchronisation of chaotic systems was generalised allowing the non identity between the coupled systems. And some time later, Rosenblum et al. considered a rather weak degree of synchronisation between chaotic oscillators, where their associated phases become locked, whereas their amplitudes remain almost uncorrelated. Hence, they called this kind of synchronisation, phase synchronisation [Rosenblum et al., 1996]. To study PS of chaotic signals, one has to identify a well defined phase variable in both coupled systems. If the flow of the chaotic oscillators has a proper rotation around a certain reference point, the phase can be defined in a straightforward way. For example, for the Rössler system with standard parameters (Eq. 3.21) the projection of the chaotic attractor on the (x, y) plane looks like a smeared limit cycle. Hence, the phase can be defined as

$$\Phi(t) = \arctan(y(t)/x(t)). \tag{7.1}$$

A more general approach to define the phase in chaotic oscillators, is the analytic signal approach introduced in [Gabor, 1946]. The analytic signal $\chi(t)$ is then given by

$$\chi(t) = s(t) + i\tilde{s}(t) = A(t)e^{i\Phi(t)}, \tag{7.2}$$

where $\tilde{s}(t)$ denotes the Hilbert transform of the observed scalar time series s(t)

$$\tilde{s}(t) = \frac{1}{\pi} P.V. \int_{-\infty}^{\infty} \frac{s(t')}{t - t'} dt', \tag{7.3}$$

where *P.V.* denotes the Cauchy principal value for the integral [Pikovsky et al., 2001]. The phase of a chaotic oscillators can also be defined based on an appropriate Poincaré section which the chaotic trajectory crosses once for each rotation. Each cross of the orbit with the Poincaré section corresponds to an increment of 2π of the phase, and the phase in between two crosses is linearly interpolated,

$$\Phi(t) = 2\pi k + 2\pi \frac{t - t_k}{t_{k+1} - t_k}, \quad (t_k < t < t_{k+1}), \tag{7.4}$$

where t_k is the time of the kth crossing of the flow with the Poincaré section.
For phase coherent chaotic oscillators, i.e. for flows which have a proper rotation around a certain reference point, the phases calculated by these different ways are in good agreement [Boccaletti et al., 2002]. If the so defined phases of two or more chaotic oscillators lock, one speaks of phase synchronisation.
Not only laboratory experiments have demonstrated phase synchronisation of chaotic oscillators, such as electronic circuits, lasers and electrochemical oscillators, but also natural systems can exhibit phase synchronisation. For example, the dynamics of the cardiorespiratory system, an extended ecological system, and the elecroencephalographic activity of Parkinsonian patients display synchronisation features [Schaefer et al., 1998], [Blasius et al., 1999], [Tass et al., 1998].
However, especially in "passive experiments", where one cannot change the parameters of the system or of the coupling, one is faced with certain problems when performing a synchronisation analysis. Examples of such systems are measurements of human ECG and the corresponding respiratory signal, or the putative synchronisation between a mother's and her fetus' heartbeats. Pikovsky et al. write [Pikovsky et al., 2001]:
"The general problem is, what kind of information can be obtained from a passive experiment? In particular, the natural question appears: can one detect synchronisation by analysing bivariate data? Generally, the answer to the above question is negative. As synchronisation is not a state, but a process of adjustment of phases and frequencies, its presence or absence cannot be established from a single observation."
In Sec. 7.6 we will introduce a kind of test which is designed do draw some conclusions about whether the process of synchronisation is in progress. Then, the analysis of passive experiments becomes feasible.
But first we introduce the idea of statistical hypothesis tests and surrogates.

7.2 Hypothesis Testing Based on Surrogate Data

The general procedure of surrogates data methods has been described in the mid 90's by Theiler [Theiler et al., 1992], [Theiler & Prichard, 1996], [Theiler, 1995], [Theiler & Rapp, 1996] and Takens [Takens, 1993]. The basic idea of a surrogate data test is the following. One first assumes that the data comes from some specific class of dynamical process, possibly by fitting a parametric or non-parametric model to the data. One then generates surrogate data (e.g. using different random increments for the realization of an ARMA process) from this hypothetical process and calculates various statistics of the surrogates and of the original data. The surrogate data will have some distribution of statistic values and one can check whether the statistic of the original data are typical. If the original data has a statistic which differs from the one of its surrogates, then we reject the hypothesis that the process that generated the original data is of the assumed class. One usually progresses from simple and specific assumptions to broader and more sophisticated models.
Let φ be a specific hypothesis and $\mathcal{F}_\varphi$ the set of all processes (or systems) consistent with that hypothesis. Let $x_i \in \mathcal{R}^N$ be a time series (consisting of N scalar measurements), and let $T : \mathcal{R}^N \to \mathcal{U}$ be a statistic which we will use to test the hypothesis φ that x_i was generated by some process $F \in \mathcal{F}_\varphi$ (generally $\mathcal{U}$ will be $\mathcal{R}$). One can discriminate between the data x_i and the surrogates x_i^{surr} consistent with the hypothesis on the basis of the approximate probability density $p_{T,F}$, i.e. the probability density of T given F.
In the next section we introduce Theiler's Fourier surrogates as an example for a surrogate technique.

7.3 Fourier Surrogates

The theory of nonlinear dynamical systems offers notions to characterise processes beyond linearity. Different quantities are used therefore, e.g. the correlation dimension, Lyapunov exponent, nonlinear forecasting errors and the mutual information [Kantz & Schreiber, 1997]. To investigate the reliability of the estimates of these characteristics, the method of Fourier surrogate data has been developed [Kurths & Herzel, 1987], [Theiler et al., 1992], [Theiler & Rapp, 1996],
[Theiler & Prichard, 1997], [Schreiber, 1998], [Schreiber & Schmitz, 2000]. The basic idea of generating Fourier surrogates is that the linear properties of the time series are specified by the mean and the squared amplitudes of the (discrete) Fourier transform. Surrogate time series are readily created by multiplying the Fourier transform of the data by random phases and then transforming back to the time domain. So, testing for a linear Gaussian process X_n one takes the

Fourier transform of the data $\{x_n\}_{n=1}^N$

$$\tilde{x}_k = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} x_n e^{-i2\pi nk/N} \tag{7.5}$$

Then the complex components $\tilde{x}_k$, $1 < k < N$ are multiplied (or equivalently substituted) by random independently and uniformly in $[0, 2\pi)$ distributed phases φ_k

$$\tilde{x_k^s} = \tilde{x}_k e^{i\varphi_k}, \tag{7.6}$$

with the constraint $\varphi_{N-k} = -\varphi_k$. Then one computes the inverse Fourier transform

$$x_n^s = \frac{1}{\sqrt{N}} \sum_{k=1}^{N} \tilde{x_k^s} e^{i2\pi nk/N} \tag{7.7}$$

and takes x_n^s as a surrogate of the data. Different realizations of the phases φ_k generate new surrogates. This process of *phase randomisation* preserves the periodogram and the Gaussian distribution (at least asymptotically for large N) [Schreiber & Schmitz, 2000].
If the time series $\{y_n\}$ is not Gaussian distributed one uses Amplitude Adjusted Fourier (AAFT) surrogates [Schreiber & Schmitz, 2000]. One assumes that $\{y_n\}$ comes from a linear Gaussian process with a nonlinear, monotonic (invertible) observation function S (e.g. $y_n = s(x_n) = \exp(x_n)$ where $\{x_n\}$ comes from a linear Gaussian process). Transforming it to Gaussian marginal distribution one recovers the original linear Gaussian process (i.e. $s^{-1}(y_n) = \ln(\exp(x_n)) = x_n$). Then, one generates surrogates x_n^s and transforms them using the function s, i.e. $y_n^s = s(x_n^s)$. These surrogates are constraint with the null hypothesis of the linear Gaussian process with a monotonic observation function. Alternatives to this approach are discussed in the literature [Schreiber & Schmitz, 1996, Kugiumtzis, 2000].

7.4 Twin Surrogates

As mentioned above, the method of Fourier surrogates has been applied to test for nonlinearity. They correspond to realizations of linear stochastic processes that have the same periodogram and mean value as the measured time series. After generating the surrogates some indicator of nonlinearity or chaotic behaviour, such as the largest Lyapunov exponent or the correlation dimension is calculated for the measured time series and for the surrogates. If the result computed for the measured time series does not differ significantly from the one calculated for the surrogates, the null hypothesis (that the measured time series is consistent with the assumption of a stochastic linear process) cannot be rejected.
But standard (linear) surrogate methods are only useful for time series exhibiting

no characteristic oscillations [Small et al., 2002]. If one wants for example to test for synchronisation between two time series, the Fourier surrogates are not longer applicable. In this case, we need surrogates that capture the nonlinear properties and characteristic oscillations of the underlying system. But they must not show the dynamical adjustment of their rhythms. Hence, the maintenance of the periodogram and mean value is not sufficient.
We consider the most general case of two systems $\vec{x}, \vec{y}$ which are bidirectionally coupled

$$\begin{aligned} \dot{\vec{x}} &= \vec{f}(\vec{x}, \vec{y}) \\ \dot{\vec{y}} &= \vec{g}(\vec{x}, \vec{y}). \end{aligned} \tag{7.8}$$

Hence, the surrogates we need, correspond to a copy of this system, i.e.

$$\begin{aligned} \dot{\tilde{\vec{x}}} &= \vec{f}(\tilde{\vec{x}}, \tilde{\vec{y}}) \\ \dot{\tilde{\vec{y}}} &= \vec{g}(\tilde{\vec{x}}, \tilde{\vec{y}}). \end{aligned} \tag{7.9}$$

Note, that the "copy of the system" corresponds to a trajectory of the whole system starting at different initial conditions. In this way, the trajectories of $\tilde{\vec{x}}, \tilde{\vec{y}}$ have the same nonlinear properties (in particular the same phase diffusion) and characteristic oscillations as $\vec{x}, \vec{y}$, but we know that $\tilde{\vec{x}}, \tilde{\vec{y}}$ might synchronise with each other but not with $\vec{x}, \vec{y}$.
However, the vector fields $\vec{f}$ and $\vec{g}$ are in general not known. Typically, we have only two measured (univariate) time series.
Note, that the generation of surrogates always implies a random element. An example for this is the random shuffling of the phases used for the construction of the FFT surrogates (Eq. (7.6)). The important point is that generating surrogates of a deterministic system we have to introduce a random element but the deterministic structure has to be maintained. This seems to be contradictory. However, any trajectory of a chaotic system is endowed with a certain "randomness" due to the random initial conditions. Slightly different initial conditions lead to diverging trajectories (and phases) and as a consequence, to unpredictability of the system's long term behaviour.
Closely linked to this randomness is the shadowing theorem. It is "the" accepted explanation of the validity of trajectories of chaotic systems generated by a digital computer [Ott, 1993]. Due to the limited computational precision one introduces in every iteration a small random error. The shadowing theorem assures that if these errors are smaller than some (small) threshold, the generated trajectories are shadowed by "real trajectory" of the system. I.e. roughly speaking, that there is a hypothetical round-off error free trajectory, which starts at slightly different initial conditions as the computer generated trajectory but always stays in its vicinity. Hence, small errors may only marginally influence the trajectory so that is stays in some sense valid [Grebogi et al., 1990].

We now tackle the problem, where to introduce the randomness in a deterministic system without changing its main characteristics. Especially, the recurrence properties are desired not to change because they seem to lay down the dynamics of the system, as the reconstruction of the attractor from the RP has shown (Chap. 6). However, the reconstruction was only possible for univariate time series. For the generation of surrogates we will use the RP computed from the vectors in phase space. If only a univariate time series is observed we will use e.g. delay embedding for the reconstruction of the phase space. Considering the phase space will allow us to also generate multivariate surrogates.
The algorithm that we propose to generate the surrogates is based on the recurrence matrix (RM) defined by $R_{i,j} = \Theta\left(\delta - ||\vec{x_i} - \vec{x_j}||_\infty\right)$. The idea to generate the surrogates is to change the structures in an RP consistently with the ones produced by the underlying dynamical system. In this way, one can reconstruct a new realization from the modified RM. However, the structures in RPs are not yet fully understood and one cannot simply interchange arbitrarily columns and lines in an RP to modify it, because in general one would change e.g. the distribution of diagonal lines and hence the entropy and predictability of the system (Sec. 3.2).
Therefore, this first approach must be modified. In Sec. 6.1 it has been reported that in general there are identical columns in an RP, i.e. $R_{i,k} = R_{j,k}\ \forall k$. This means that there are points which are not only neighbours (i.e. $||\vec{x}_i - \vec{x}_j||_\infty < \delta$), but which also share the same neighbourhood. Reconstructing the attractor from an RP, the respective neighbourhoods of these points cannot help to distinguish them, i.e. from this point of view they are identical. As neighbours cannot distinguish them, we will call them **twins**. The number of twins in such a system depends on δ and is typically of the order of 10-50% of the entries of the time series. Twins are special points of the time series as they are undistinguishable by their neighbours but still different [1] and hence have different pasts and -more important- different futures.
The key idea for introducing the randomness needed for the generation of surrogates of a deterministic system is that one can jump randomly to one of the two possible futures of the twins. A surrogate then is generated in the following way:

1. One identifies all pairs of twins.
2. Then one chooses an arbitrary point, say $\vec{x}_k$, as the starting point.
3. If it has no twin, the next point of the surrogate time series is $\vec{x}_{k+1}$.
4. If on the other hand, it has a twin, say $\vec{x}_m$, then one can go with equal probability to $\vec{x}_{k+1}$ or $\vec{x}_{m+1}$. If triplets occur one proceeds analogously.

These steps are then repeated iteratively until the surrogate time series has the same length as the original one. If at some point in the algorithm one gets to the

[1] In a chaotic system there are not two equal points, otherwise it would be periodic.

last point in the time series which is a twin, one must jump back to its "brother". This algorithm yields twin surrogates which are shadows of a (typical) trajectory of the system [Ott, 1993], [Katok & Hasselblatt, 1995]. They are characterised by the same entropies (Lyapunov exponents), dimensions and the same attractor as the original system. Also their spectra and correlation functions are consistent with the ones of the original time series. But the key property of the twin surrogates, is that they do not adapt their rhythms to the original time series from which the surrogates were generated. That means, each twin surrogate corresponds to a copy of the underlying system in the sense of Eq. (7.9).
The surrogates do not only seem to give reasonable results for dynamical systems. The twin surrogates of e.g. an ARMA process, again show the typical behaviour of a linear Gaussian process. Even the parameters of the process can be estimated correctly from them.
In the next section we will compare twin surrogates to nearest neighbour surrogates, which may seem at first more direct.

7.4.1 Twin Surrogates and NN-Surrogates

The basic idea of generating twin surrogates is looking for a special class of nearest neighbours and interchanging their future evolutions. An important point is why it is necessary to choose twins and not simply nearest neighbours to alter the future evolutions. Note, that nearest neighbours are linked to a concept which Lorenz has termed "naturally occurring analogues" [Lorenz, 1969]. He proposes a weather prediction scheme based on analogue weather situations which already have been recorded. To predict the weather he uses the observed evolution of the past analogue (i.e. the weather of tomorrow will be similar to the weather one day after the preceeding analogue).
Twins are also "naturally occurring analogues", but a restricted class of them. They are even more similar than nearest neighbour analogues. Because we do not -as Lorenz- focus on prediction but rather on surrogates, we interchange the time evolutions of the twins. However, the question arises why we use twins and not simply nearest neighbours, as considered in [Small et al., 2002]. Fig. 7.1 opposes both procedures. The left panel illustrated the alternative method of nearest neighbours surrogates. In this case all ε-neighbours can be used to obtain surrogate futures. The right panel shows the same but using twins only. In this case the two twins are arranged in a way that they are indistinguishable from the point of view of their neighbours. Only the surrogates based on these twins do not change the recurrence structures in the RP qualitatively, because twins are characterised by the identity of their columns in the RP. On the other hand, they have different pasts and different future evolutions. Hence, they can be considered as a joint between two pasts and two future evolutions.
Due to these different characteristics, the two ways to generate surrogates have quite different performances. Fig. 7.2 shows examples of nearest neighbours sur-

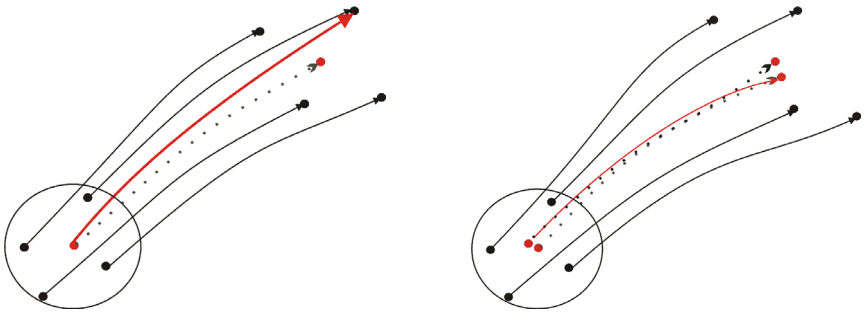

Figure 7.1: **Left panel**: Diagram representing the idea of Nearest-Neighbour-Surrogates. The future evolution is interchanged randomly between points in phase space which lie in an ε-vicinity. **Right panel**: Diagram illustrating the idea of the twin surrogates. The future evolution can only been altered between points which have the exactly identical neighbourhoods. They have the same recurrence characteristics with respect to the other points in phase space.

rogates of the logistic map ($r = 4$). For a rather small threshold $\varepsilon = 0.001$ the expected parabola of the phase portrait is obtained rather well (b), but the surrogate settles down to a fixed point (or sometimes a periodic orbit). These lead to a reduced complexity and entropy of the surrogate with respect to the original time series. For larger thresholds (e.g. $\varepsilon = 0.05$), the surrogates also have a rather large number of periodic windows (Fig. 7.2 (c)), and the phase portrait (d) yields a fuzzy parabola.
These cases are typical for the nearest neighbours surrogates. In general it is difficult to determine an appropriate ε for the generation of the surrogates. This is one reason why the use of these surrogates problematic, in spite of the low effort needed to compute them.
The twin surrogates overcome this problem (and other problems). They yield rather stable results for a broad interval of thresholds. Fig. 7.3 (a) shows a typical twin surrogate for the logistic map. There are no obvious and overemphasised periodic windows and also the phase portrait displays a sharp parabola (Fig. 7.3 (b)). The dynamics of the map seems to be reproduced by the surrogates rather well. The surrogates' autocorrelation function and K_2 are consistent with the ones of the original time series.
Fig. 7.3 (c) and (d) display the respective graphics for uniformly distributed independent noise. Also in this case the "dynamics seems" to be reproduced well. Generating surrogates of an ARMA process one can also estimate successfully

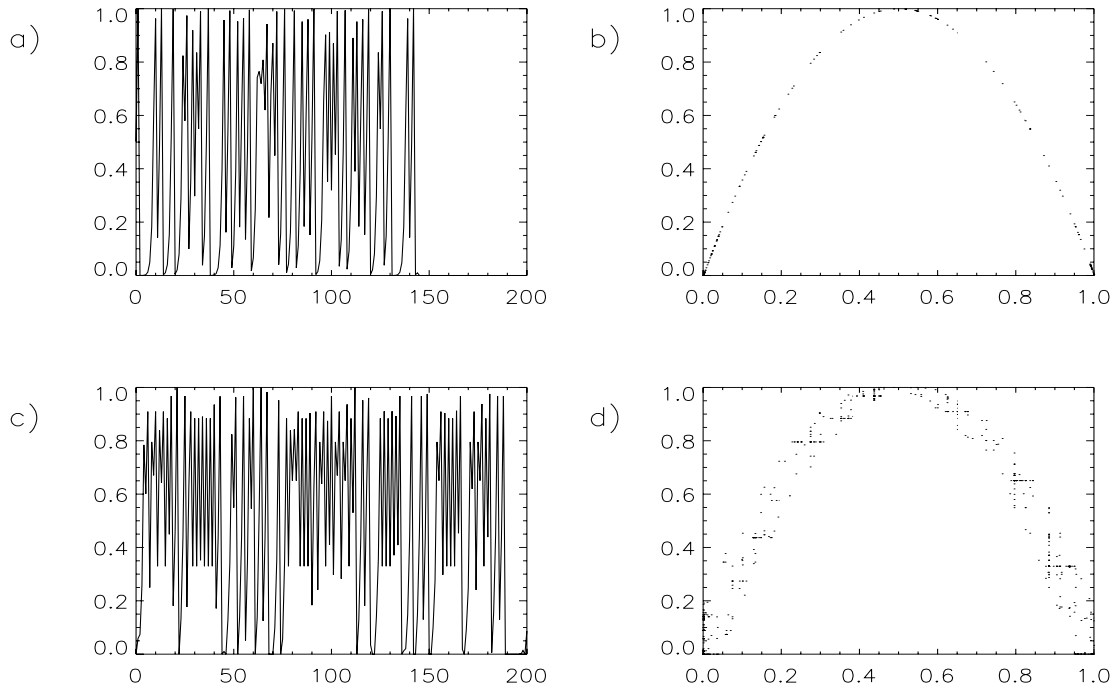

Figure 7.2: (a) Nearest-Neighbours-Surrogate of a time series ($N = 1,000$) the logistic map ($r = 4$) based on a small threshold $\varepsilon = 0.001$. After about 140 iterations it settles down to a fixed point. (b) Phase portrait of the same surrogate, which yields a rather good approximation of the expected parabola. (c) Nearest-Neighbours-Surrogate of a time series ($N = 1000$)the logistic map ($r = 4$) based on a rather large threshold $\varepsilon = 0.05$. The surrogate has (too) many periodic windows. (d) Phase portrait for the same surrogate. The expected parabola is smeared out.

the parameters of the original process from its surrogates.
Fig. 7.3 (e) represents the x-component of the surrogate of a trajectory from the chaotic Rössler system (Eq. (3.21)). The corresponding attractor (f) represents the typical attractor of the original time series. The linear quantities like correlation function, power spectrum but also nonlinear quantities like K_2 and D_2 are (up to fluctuations which also occur for trajectories of the Rössler system when starting at different initial conditions) the same as the one of the underlying system.
The obtained results are rather independent of the choice of the threshold ε. The generation of surrogates is not as fast as in the case of the nearest neighbours surrogates due to the fact that all the neighbours of each point of the time series have to be compared to all the neighbourhood to every other point. However, once all the neighbourhoods have been determined the computation time to gen-

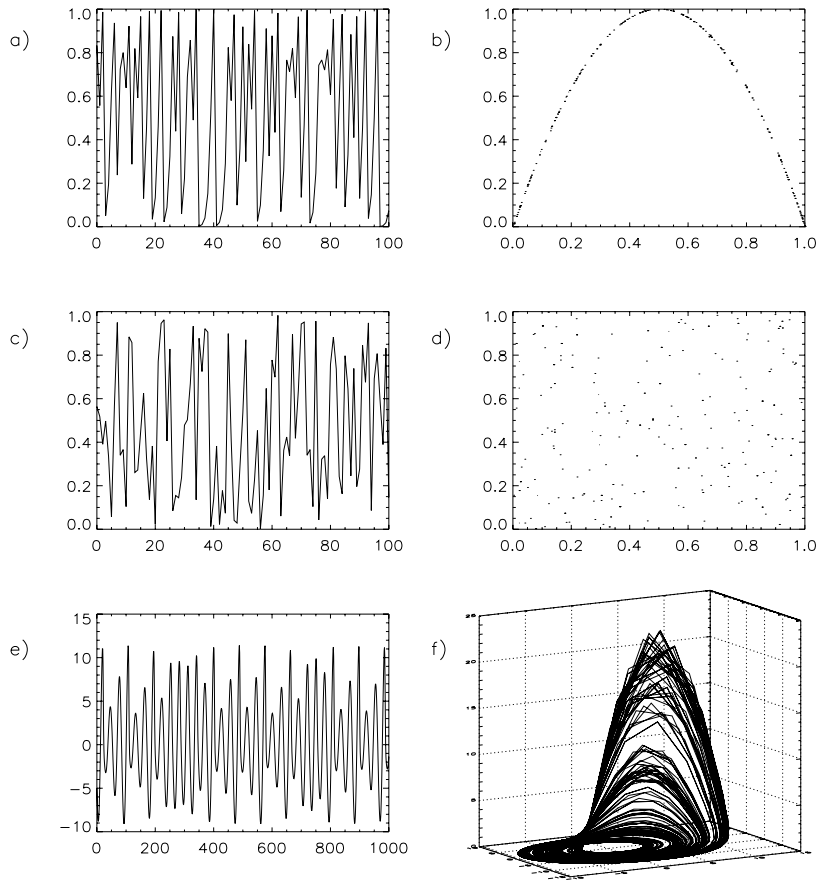

Figure 7.3: (a) Twin surrogate of a time series from a logistic map ($r = 4$). (b) Phase portrait for the surrogate of figure (a). (c) Surrogate for uniformly distributed independent noise. (d) Corresponding phase portrait. (e) x-component of a surrogate of the Rössler system. (f) Attractor of the surrogate of the Rössler attractor.

erate 1,000 surrogates instead of 10 is only marginally higher.
As mentioned above, twin surrogates exhibit the same oscillatory structures as the underlying time series. Another important point is that they also reflect the inherent phase diffusion.

7.5 Twin Surrogates and Phase Diffusion

Twin surrogates do maintain as mentioned in the last section most of the linear and nonlinear characteristics of a system, i.e. of the given time series. But what do they actually change?
For two close initial conditions the phases of a chaotic oscillator or a noisy limit cycle can be assumed to drift apart 'random walk like' [Pikovsky et al., 2001].

The diffusion constant depends on the system and its parameters. It decreases if the system is in PS with another system. This change is the basis of many measures which quantify PS.
The twin surrogates' phases drift apart, due to their construction, with a diffusion constant which corresponds to the one of the underlying system. Hence, they mimic the phase diffusion of the underlying time series.
Taking the Rössler system (Eqs. (3.21)) as an example, we generate 500 twin surrogates. Then we compute the instantaneous phase $\Phi(t)$ of the original time series by Eq. 7.1 and the respective ones $\Phi_1^s(t)$ of its surrogates. Here s takes all values from 1 to 500. Fig. 7.4 (a) shows the difference between the phase of the original time series and the mean of the respective phases of its surrogates, i.e.

$$\Delta_{\text{orig}}(t) = \Phi(t) - \langle \Phi^s \rangle_s \,. \tag{7.10}$$

As the mean frequency of the surrogates is by construction the same as the one of the original time series, the curve fluctuates about zero with no apparent drift. Fig. 7.4 (b) shows the respective curves for some surrogates, i.e.

$$\Delta_j(t) = \Phi^j(t) - \langle \Phi^s(t) \rangle_s \,, \tag{7.11}$$

where the index j denotes the j-th surrogates. The surrogates drift away from this mean due to their random walk nature. If one studies two systems, the

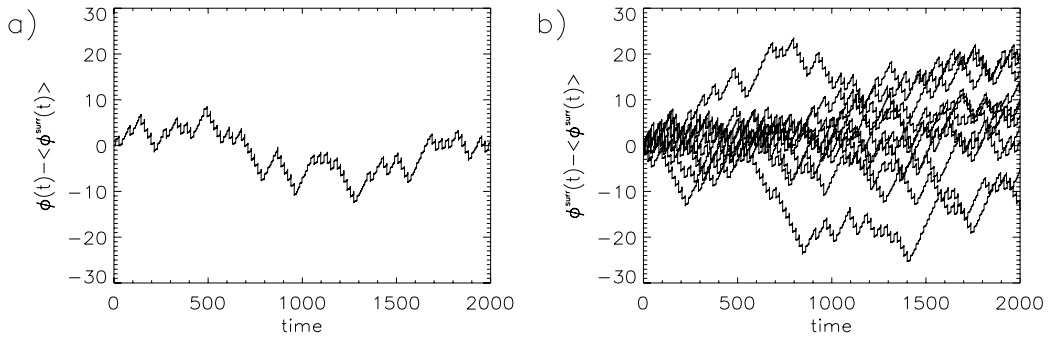

Figure 7.4: (a) Difference of the instantaneous phases of a time series of the Rössler system and the mean of 500 of its surrogates at each point in time. The curve fluctuates about zero with no obvious drift. (b) Some examples of the difference of the instantaneous phases of the surrogates with the mean of the surrogates. Some surrogates drift away from the mean.

procedure is similar. We now consider

$$\Delta_{\text{orig}}(t) = \Phi_1(t) - \Phi_2(t) \tag{7.12}$$

Fig. 7.5 (a) displays the phase difference $\Delta_{\text{orig}}(t)$ of two time series of identical and uncoupled Rössler systems (marked with an arrow)

$$\begin{aligned} \dot{x}_1 &= -x_2 - x_3, & \dot{y}_1 &= -y_2 - y_3, \\ \dot{x}_2 &= x_1 + 0.15x_2, & \dot{y}_2 &= y_1 + 0.15y_2, \\ \dot{x}_3 &= 0.2 + x_3(x_1 - 10), & \dot{y}_3 &= 0.2 + y_3(y_1 - 10). \end{aligned} \quad (7.13)$$

The further two curves show the phase difference of the first system with a surrogate of the second system

$$\Delta_j(t) = \Phi_1(t) - \Phi_2^j(t), \quad (7.14)$$

where the indices 1 and 2 mark the first and second system, and j again stands for the j-th surrogate. All three phase differences exhibit a random walk like structure. Fig. 7.5 (b) shows the distribution of the standard deviations of $\Delta_j(t)$, i.e.

$$\sigma_{\Delta_j} = \left\langle (\Delta_j(t) - < \Delta_j(t) >_t)^2 \right\rangle_t, \quad (7.15)$$

where $< \cdot >_t$ denotes the time average. The corresponding standard deviation for the two original time series, i.e. the standard deviation of $\Delta_{orig}(t)$, is marked by a dashed line. It lies well within the peak of the distribution, suggesting that

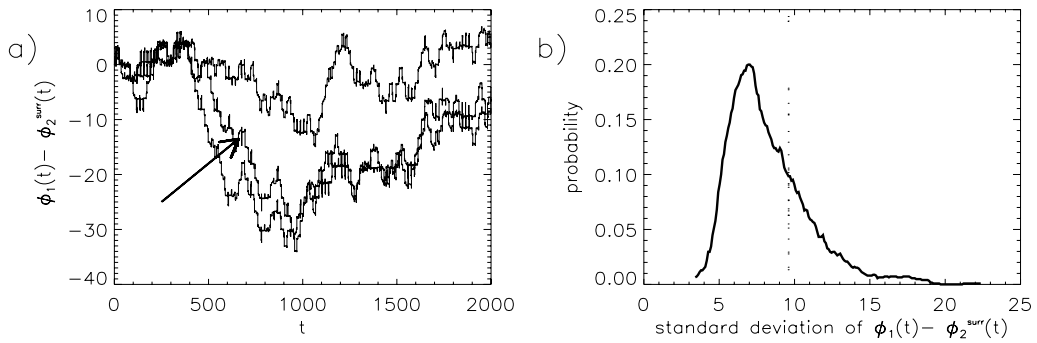

Figure 7.5: (a) Difference of the instantaneous phases of two time series of identical and uncoupled Rössler systems (marked with an arrow). The further two curves show the phase differences of the first system with two surrogates of the second system. All three phase differences exhibit a random walk like structure. (b) Distribution of the standard deviations of the phase differences of the first system with 500 surrogates of the second system. The dashed vertical line marks the value obtained for the two underlying time series from the Rössler system. This value is not an "outlier" assuming the given distribution obtained from the surrogates.

there is no significant difference between the second time series and its surrogates, with respect to the rhythm's adjustment to the first time series. This is expected,

because the two original time series were not coupled.
Fig. 7.6 shows the same graphics for two coupled Rössler systems

$$\begin{aligned} \dot{x}_1 &= -(1+\nu)x_2 - x_3 + \mu(y_1 - x_1), & \dot{y}_1 &= -(1-\nu)y_2 - y_3 + \mu(x_1 - y_1), \\ \dot{x}_2 &= (1+\nu)x_1 + 0.15x_2, & \dot{y}_2 &= (1-\nu)y_1 + 0.15y_2, \\ \dot{x}_3 &= 0.2 + x_3(x_1 - 10), & \dot{y}_3 &= 0.2 + y_3(y_1 - 10), \end{aligned} \tag{7.16}$$

with a frequency mismatch $\nu = 0.015$ and a coupling strength of $\mu = 0.015$ (with this choice of parameters, they are not in PS [Pikovsky et al., 2001]). Due to ν

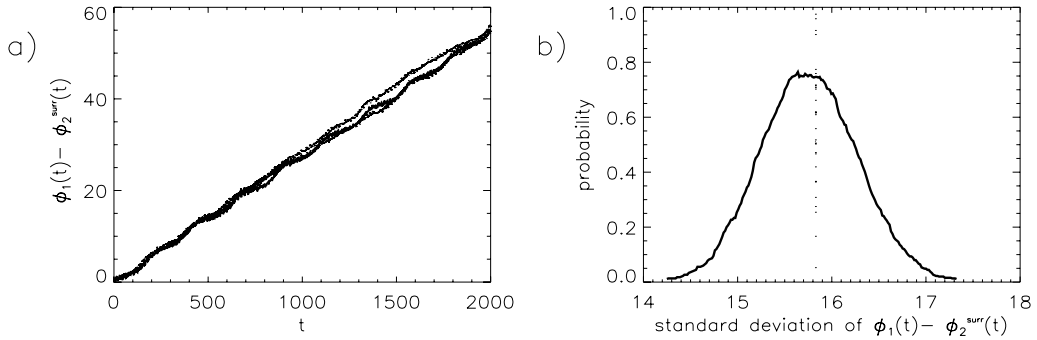

Figure 7.6: (a) Difference of the instantaneous phases of two time series of coupled Rössler systems with slightly different eigenfrequencies which are not in PS, and two curves showing the phase differences of the first system with two surrogates of the second system. All three phase differences exhibit a random walk like structure about the mean frequency (ωt). (b) Distribution of the standard deviations of the phase differences of the first system with 500 surrogates of the second system. The dashed vertical line marks the value obtained for the two underlying time series from the Rössler system. This value is not an "outlier" assuming the given distribution obtained from the surrogates.

there is a clear drift of all phase differences (Fig. 7.5 (a)). However, in spite of the drift, the distribution of the standard deviation $\Delta_{orig}(t)$ is not significantly different from the ones computed based on the surrogates $\Delta_j(t)$ - the dashed line lies inside of the peak.
The last case we consider are two coupled Rössler systems Eqs. (7.16) with a frequency mismatch $\nu = 0.015$ and a coupling strength of $\mu = 0.035$ (for these parameters, both oscillators are in PS) (Fig. 7.7). Also in this case the differences of the instantaneous phases show a behaviour which is comparable to the one of a random walk. However, the differences of the phases $\Delta_{orig}(t)$ for the original time series (Fig. 7.7 (a) middle time series) has a smaller standard deviation than the ones computed for $\Delta_j(t)$. This is due to the fact that the surrogates of the second time series do not "see" the first time series, and hence do not adjust their

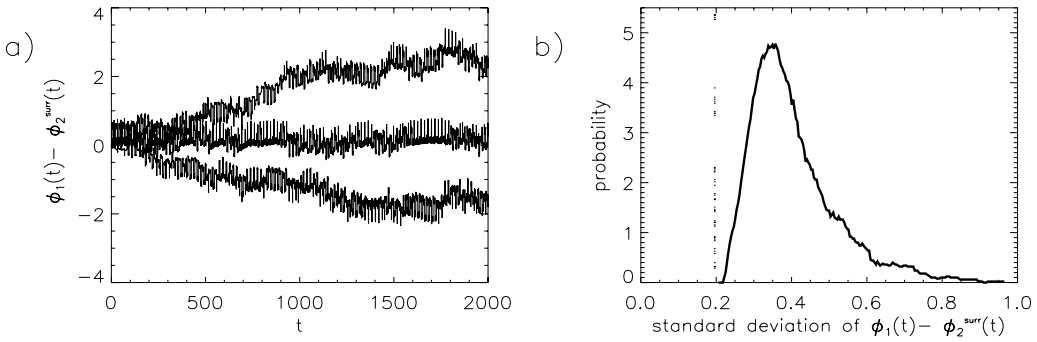

Figure 7.7: (a) Difference of the instantaneous phases of two time series of coupled Rössler systems with slightly different eigenfrequencies which are in PS (curve in the middle), and two curves showing the phase differences of the first system with two surrogates of the second system. All three phase differences exhibit a random walk like structure about the mean frequency (ωt), which become even more evident on longer time scales. (b) Distribution of the standard deviations of the phase differences of the first system with 500 surrogates of the second system. The dashed vertical line marks the value obtained for the two underlying time series from the Rössler system. This value is an "outlier" assuming the given distribution obtained from the surrogates.

rhythms to first oscillator. This is reflected in the distribution of the standard deviations of $\Delta_j(t)$ (Fig. 7.7 (b)) by the fact that the dashed line corresponding to the standard deviation of $\Delta_{orig}(t)$ lies outside of the peak.
An alternative interpretation of these results is the following. Chaotic oscillators and noisy limit cycles have, up to very artificial cases, always a certain phase diffusion. The phases are closely related to the zero Lyapunov exponent, i.e. perturbations in the phases will neither decrease nor increase but rather sum up over time. This leads to a "random walk" of the phases.
If both systems are not in PS, their respective instantaneous phases will perform a random walk uncorrelatedly, i.e. both will have "different random increments" [Pikovsky et al., 2001]. Also the difference of the phases will follow the mathematics of a random walk (Fig. 7.5 (a)).
If two oscillators synchronise, one of the two zero Lyapunov exponents becomes negative. This corresponds in the model of the two random walks to the fact that the random increments of both processes highly correlate. One can assume that they have nearly identical increments. The phase difference of the two oscillators becomes zero up to some very small fluctuations due to the remaining small phase diffusion. This means that even in the case of systems in PS, the phase difference will not be exactly zero (Fig. 7.7 (a), middle curve). This phase difference is however much smaller than in the case of random walks with increments which are independent (Fig. 7.7 (a), upper and lower curves). Fig. 7.7 (b) shows that

the standard deviation of the phase difference of the original time series $\Delta_{orig}(t)$, marked by a dashed vertical line, is significantly smaller than the ones obtained by the surrogates.
These considerations will now allow to construct a hypothesis test for the synchronisation of complex systems.

7.6 Tests for Synchronisation of Complex Systems

We now use the twin surrogates to test for PS. The test works based on similar assumptions as in the case of independent patients to test for synchronisation in the cardio-respiratory system or the (independent) heart beat time series of the surrogate mothers in the case of the mother-fetus synchronisation [Toledo et al., 1999], [van Leeuwen, 2003]. After generating the surrogates we use the following statistics for the analysis:

1. The Synchronisation Index defined as

$$SI = \frac{1}{N} \sum_{t=1}^{N} \left(\exp(i \cdot \Delta_{1,2}(t))\right)^2 , \tag{7.17}$$

   where $\Delta_{1,2}(t) = \Phi_1(t) - \Phi_2(t)$ are the differences of the instantaneous phases of the two signals. It is one for phase synchronised systems and asymptotically zero for independent systems [Boccaletti et al., 2002].

2. The standard deviation of the difference of the instantaneous phases $\sigma\left(\Delta\varphi_{1,2}(t)\right)$ which is small for systems in PS and otherwise large. The index $\sigma\left(\Delta_{1,2}(t)\right)$ reflects the broadness of the histogram of $\Delta_{1,2}(t)$. If the systems are in PS the histogram of $\Delta_{1,2}(t)$ has a predominant peak, whereas otherwise the histogram is rather broad). The surrogates help us to decide what large and small mean in this case.

To exemplify our procedure we consider two non-identical, mutually coupled Rössler oscillators with a frequency mismatch of $\nu = 0.015$ (Eqs. (7.16)). We vary the coupling strength ε from 0 to 0.12 and compute the two PS indices for the original time series and for each coupling strength. The phases are computed by Eq. (7.1).
We describe three approaches to detect PS in the following subsections. In each case the analysis is based on 200 surrogates. If SI $\left(\sigma\left(\Delta(t)\right)\right)$ for the original time series is higher (lower) then 99% of the values of SI $\left(\sigma\left(\Delta(t)\right)\right)$ computed based on the surrogates, we reject the null hypothesis that the original time series are not in PS.

7.6.1 First Approach

The first approach is rather appropriate if the time series come from two similar systems, e.g. two Rössler systems with a frequency mismatch. We first generate for the first time series (T1) 200 twin surrogates (S1). Then, the test statistic is computed for $\Phi_1(t) - \Phi_1^j(t)$, where $\Phi_1^j(t)$ denotes the phase of the j-th surrogate of the first time series. One obtains a distribution for the test statistic. This distribution is then compared with the respective test statistic computed for $\Phi_1(t) - \Phi_2(t)$. If the latter differs significantly from the distribution, the null-hypothesis is rejected.
This approach is based on the idea that we generate surrogates which are very similar but not in PS with T1. The assumption is that T2 can only be more similar to T1 than the surrogates S1, if T1 and T2 are synchronised.
Figs. 7.8 a) and b) show the results for both test statistics (bold line represents the test statistic computed for $\Phi_1(t) - \Phi_2(t)$; the solid line marks the significance threshold). As expected, SI increases when the transition to PS occurs. As the phases become more coherent then, also the significance limit increases, because it is more difficult then to decide if the regularities in the phases are due to an interaction, or accidental [2]. The zoom shows that even if the synchronisation index exceeds the value of **0.98** this does not give sufficient evidence for PS, i.e. such a value could also be obtained by independent but otherwise similar oscillators. Only at a bit larger coupling strength of about 0.038 there is sufficient evidence to reject the null hypothesis that the systems are not in PS. **This result shows that the knowledge of the synchronisation index alone is not sufficient evidence for PS.**
In the case of the standard deviation of the phase difference, low values indicate PS. Also in this case the test statistic seems to indicate a transition to PS. The surrogates show that only from a coupling strength of about 0.038 one has sufficient evidence to reject the null hypothesis. So both measures give consistent results.
Fig. 7.9 compares the results of the test for different couplings with the Lyapunov exponents. Fig. 7.9 (a) shows the rejection (at a 1% level) of the test based on SI. 1 means no rejection, -1 means rejection. Fig. 7.9 (b) displays the same for $\sigma(\Delta(t))$. Comparing theses results with Fig. 7.9 (c), namely the Lyapunov exponents, one finds that at a coupling strength of about 0.03 the zero Lyapunov exponent becomes negative (i.e. the random increments of the random walks of the phases become highly correlated), indicating the transition to PS. The tests indicate PS at about 0.038 hence at slightly higher coupling strength. This is expected, as due to the limited length of the time series (5,000 points), in a region

[2] For two sine functions with the same frequency, which do not have any phase diffusion, it is not possible to tell if this "synchronisation" is accidental or due to coupling. However, for systems which have a certain phase diffusion but the same frequency, such as two identical Rössler oscillators, the surrogates help to detect PS.

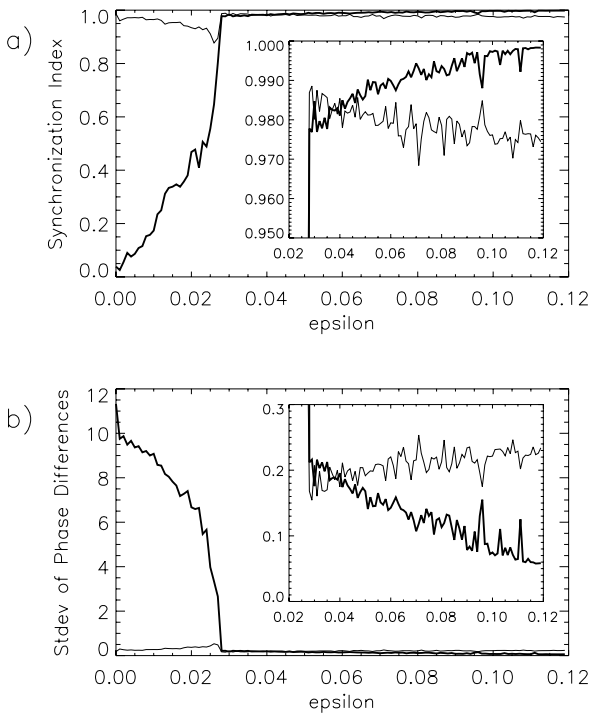

Figure 7.8: Synchronisation index (bold line) and significance level of 1% (solid line) computed based on the twin surrogates. At the transition to PS both curves increase due to the more coherent phases. The smaller windows display a zoom of curves. Note, that in a) the null-hypothesis is rejected when the value for the measure is larger then the threshold; and in b) when it is smaller.

from coupling strength between 0.03 and 0.038 there is not enough evidence to reject the null hypotheses at a 1% level. This is mainly due to the frequent phase slips which occur in this region. Using 5 or 10% levels this region decreases in size. Note also, that even though the surrogates correspond to independent realizations (i.e. copies of the original system, just as in the case of the natural surrogates for the mother fetus system) the test detects only PS. That means if two time series are coupled but not yet in PS, the null hypothesis is not rejected. Using an appropriate test statistic which is sensitive to couplings, one could use the same surrogates to test for coupling and eventually for the coupling direction. We next perform an analysis of the power of tests based on the two PS indices presented above. In the case of $\varepsilon = 0$ and no frequency mismatch ($\nu = 0$) we perform the test 100 times for 100 random initial conditions of the Rössler system and use a significance level of 5%. The null hypothesis was erroneously rejected only in 4 of the 100 cases. This is a rather auspicious result, as due to the same frequencies, it is extremely difficult to detect, that in this case there is no PS. In

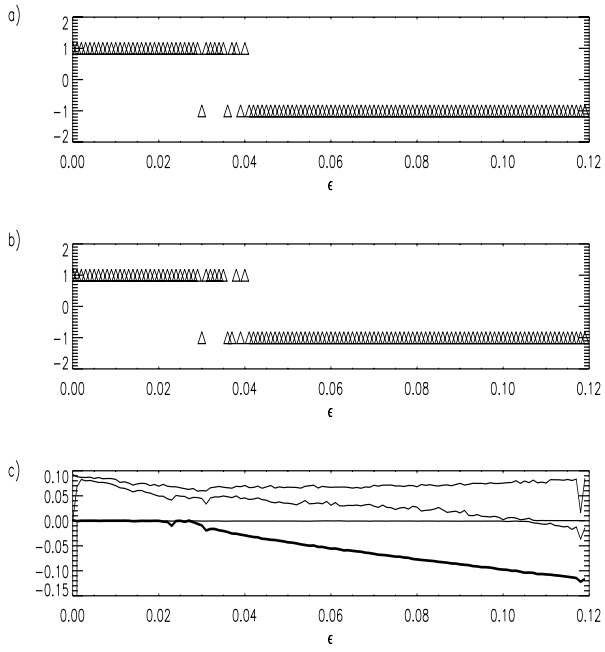

Figure 7.9: Results of the hypothesis tests (at a significance level of 1%) (a) SI and (b) $\sigma(\Delta(t))$ compared to the Lyapunov exponents (LEs) of the system (c). The LEs indicate a transition to PS at a coupling strength of $\varepsilon \approx 0.03$. The tests do not reject for coupling strength marked by "1" and do reject at "-1". In a region between 0.03 and 0.038 multiple phase slips occur in the TS which make it in some cases impossible for the test to reject the null hypothesis.

the case of a coupling strength of $\varepsilon = 0.02$ (e.g. no PS) and a frequency mismatch $\nu = 0.015$, there were no erroneous rejections of the null hypothesis. Finally, for PS ($\varepsilon = 0.45$ and $\nu = 0.015$), in all 100 test runs the null hypothesis was correctly rejected. These results indicate that the power of the test is rather good.

7.6.2 Second Approach

The second approach is slightly different to the first one. In this case the 200 surrogates are generated from T2. Then one computes the above introduced test statistics for $\Phi_1(t) - \Phi_2(t)$, and then for $\Phi_1(t) - \Phi_2^j(t)$. The basic idea is in this case that if T1 and T2 are originally synchronised the surrogates of T2 do not adapt their rhythm to T1. If the T1 and T2 were synchronised then they would have a rather narrow distribution of phase differences and their phase diffusion was rather small. As the surrogates generated from T2 correspond to trajectories from the second subsystem of the copy of the whole system (Eqs. (7.9)). They also have a small phase diffusion, because they are in PS with the first subsys-

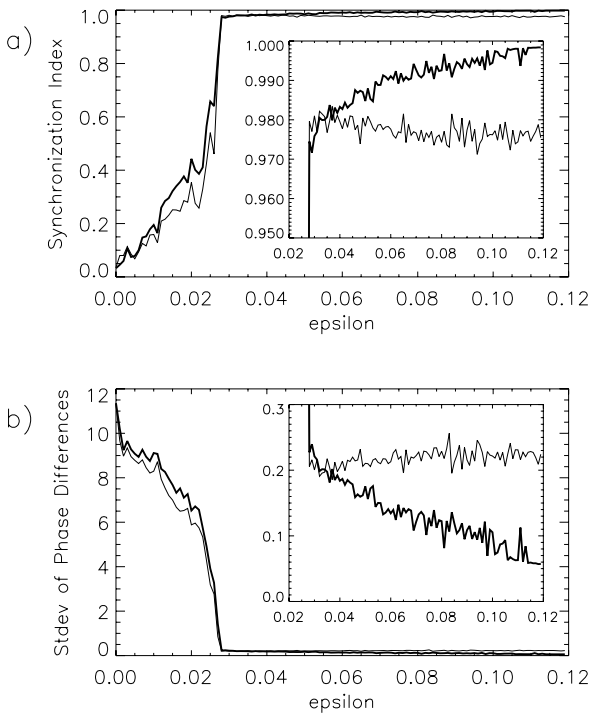

Figure 7.10: Synchronisation index (bold line) and significance level of 1% (solid line) computed based on the twin surrogates. At the transition to PS both curves increase due to the more coherent phases. The smaller windows display a zoom of curves. Note, that in a) the null-hypothesis is rejected when the value for the measure is larger then the threshold; and in b) when it is smaller.

tem of the copy (Eqs. (7.9)). However, they are not in PS with neither T1 nor T2. The resulting significance threshold (solid line) is especially for low coupling strength much lower than observed in the first approach (Fig. 7.10). However, the test only rejects correctly the thesis that the original time series are not in PS for coupling strengths larger than 0.036 for the standard deviation of the phase differences (i.e. $\sigma(\Delta(t))$ of the original time series is lower than the 1% level of the one computed with the surrogates (Fig. 7.10 (a))).
SI it already rejects the null hypothesis for coupling strengths larger than 0.003 (i.e. the measure for the original time series is always larger than the significance threshold (Fig. 7.10 (b))). This is expected as the surrogates actually are independent. The point is that SI is sensitive already to coupling, whereas $\sigma(\Delta(t))$ is sensitive to phase synchronisation. Hence, this second approach can only be applied with a properly chosen test statistic.

7.6.3 Third Approach

The third approach uses 200 surrogates of both T1 and T2. Then, $\Phi_1(t) - \Phi_2(t)$ is compared with $\Phi_1^j(t) - \Phi_2^j(t)$ for $j = 1, \ldots, 200$. Φ_1^j denotes the surrogates of T1 and Φ_2^j the surrogates of T2. This approach, takes fluctuations of the phases

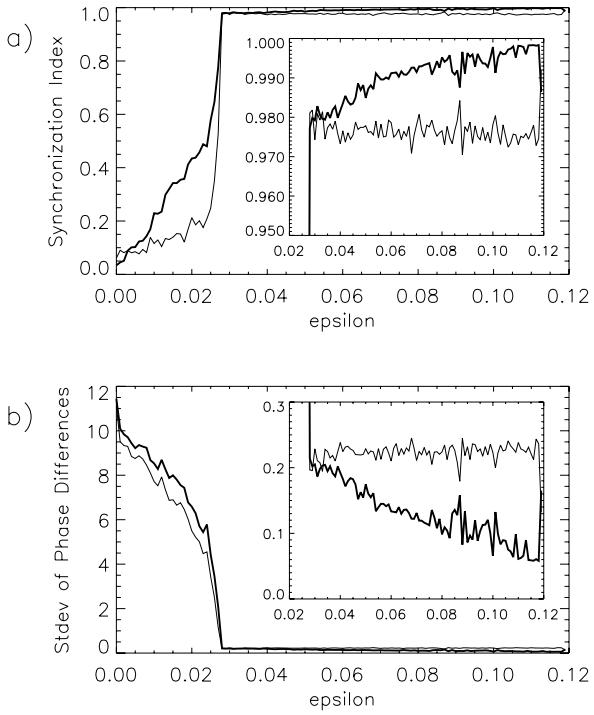

Figure 7.11: Synchronisation index (bold line) and significance level of 1% (solid line) computed based on the twin surrogates. At the transition to PS both curves increase due to the more coherent phases. The smaller windows display a zoom of curves. Note, that in a) the null-hypothesis is rejected when the value for the measure is larger then the threshold; and in b) when it is smaller.

of T1 and T2 into account. The results are similar to the ones of the second approach. Again SI already rejects the null-hypothesis for $\varepsilon > 0.003$ (Fig. 7.11 (a)), whereas $\sigma(\Delta(t))$ detects PS (Fig. 7.11 (b)). Note, that already at coupling strength of about 0.033 the hypothesis is rejected using $\sigma(\Delta(t))$. This is the best result which we have obtained.
Our results show that even though the surrogates correspond to trajectories of an independent copy of the whole system, one can construct a test which helps to detect PS. In the next section we will apply the test to experimental data from an electrochemical oscillator.

7.6.4 Electrochemical Oscillator

Next, we applied our algorithm to data from electrochemical oscillators [Kiss et al., 2004]. The experiment was performed in a standard three-compartment electrochemical cell consisting of an array of two iron working electrodes (1-mm diameter each with 2 mm spacing), a $Hg/Hg_2SO_4/K_2SO_4$ reference electrode and a Pt mesh counter electrode was used (a schematic of this experimental setup can be found in [Kiss et al., 2004].) The applied potentials (V) of the two electrodes were held at the same value with a potentiostat. Experiments were carried out in 0.5 mol/dm^3 H_2SO_4 solution at room temperature in stagnant solution. The working electrodes are embedded in epoxy and the ends of the electrodes, where reaction takes place, are exposed to the electrolyte. The experiments are carried out by sweeping the circuit potential (V) from the rest potential (-0.95V) slowly, at a rate of 1mV/s, to the target potential where the currents of the electrodes are measured independently at a sampling rate of 2kHz.
For the test we base our study on the second approach (Sec. 7.6.2) and use four different test statistics.

1. Fig. 7.12 (a) shows the results for the entropy S of the histogram of the phase differences. S expresses the sharpness of the maximum in the cyclic phase difference distribution and is obtained as $S = \frac{S_{max}-S_{real}}{S_{max}}$ where S_{real} is the Shannon entropy of the cyclic phase difference distribution ($S_{real} = -\sum_{i=1}^{M} p_i \ln(p_i)$, M is the number of bins in the histogram), and S_{max} is the maximum entropy (flat distribution). S takes on values from 0 to 1 as the distribution changes from flat to a delta function. The measure S already has been used to investigate these electrochemical oscillators [Kiss et al., 2004].

2. Fig. 7.12 (b) shows the test for the synchronisation index SI.

3. Fig. 7.12 (c) shows the result for $\sigma(\Delta(t))$.

4. Fig. 7.12 (d) shows the result for a measure based on recurrences in a reconstructed phase space [Romano et al., submitted].

All pictures show that for the lowest three coupling strengths the systems are not in PS, and for the highest three couplings we find PS. Only in the case of $\sigma(\Delta(t))$ and $\varepsilon = 1.0$, the null hypothesis cannot be rejected. These results are in accordance with recent results reported in [Kiss et al., 2004]. Note, that already $S = 0.2$ indicates PS, although S is normalised, and hence only values of $S \sim 1$ should indicate PS.
In conclusion, we have introduced a method to generate (uni- or multivariate) surrogates, which is based on recurrences. The twin surrogates mimic all relevant dynamical properties, such as entropies, dimensions, correlations, spectra and also the attractor in phase space of the underlying system. They correspond

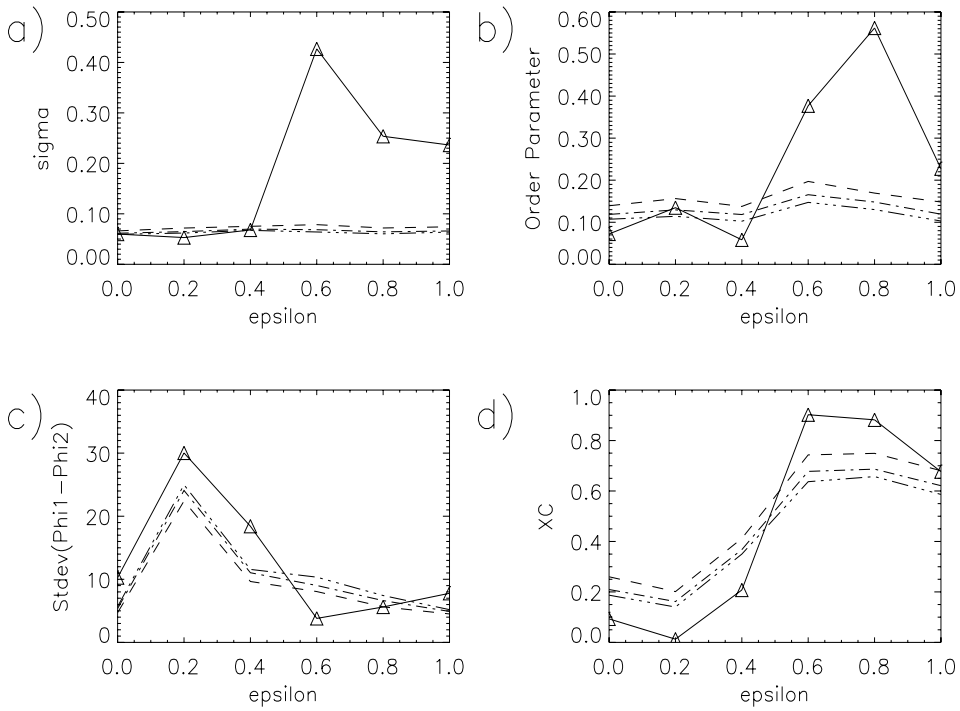

Figure 7.12: Results of the analysis of the synchronisation of the electrochemical oscillators in dependence of the coupling strength ε: (a) Entropy S, (b) SI (c) $\sigma(\Delta(t))$ (d) measure based on recurrences. The dashed lines indicate the 1%, 5% and 10% significance levels.

to trajectories of a copy of the whole underlying system, i.e. starting at different initial conditions. Hence, the twin surrogates do not synchronise with the underlying system, but they have the same phase diffusion as the original system. Using an appropriate test statistic it also is possible to test for PS due to the special character of the transition from non-PS to PS. These surrogates can in principle be used for tests in different fields especially testing for coupling directionality.

Chapter 8

Application to Data

8.1 Stability of Planetary Systems

The results presented in this section were obtained during the 3^{rd} Helmholtz Summer School in Potsdam (2003). The results summarised here are an almost word by word reproduction of parts of a common paper [Asghari et al., 2004]. The part to be related with this thesis is the data analysis by means of RPs which was carried out with the participants of the summer school and especially with the help of M.C. Romano and W. von Bloh. The findings are introduced in this work to show how illustrative results in various fields can be obtained by the method of RPs. However, detailed results are discussed hereafter only for two out of five extrasolar planetary systems (namely Gl 777 A and HD 72659) in this work. The results for all systems are summarised at the end of this section.
During the summer school we carried a thorough dynamical investigation of five extrasolar planetary systems using extensive numerical experiments. The systems Gl 777 A, HD 72659, Gl 614, 47 Uma and HD 4208 were examined concerning the question of whether they could host terrestrial-like planets in their habitable zones (HZ). First we investigated the mean motion resonances between fictitious terrestrial planets and the existing gas giants in these five extrasolar systems. Then a fine grid of initial conditions for a potential terrestrial planet within the HZ was chosen for each system, from which the stability of orbits was then assessed by direct integrations over a time interval of 1 million years. For each of the five systems the 2-dimensional grid of initial conditions contained 80 eccentricity points for the Jovian planet and up to 160 semimajor axis points for the fictitious planet. The equations of motion were integrated using a Lie-series integration method with an adaptive step size control. This integration method achieves machine precision accuracy in a highly efficient and robust way, requiring no special adjustments when the orbits have large eccentricities.
The stability of orbits was examined with a determination of the Rényi entropy of second order (Sec. 3.2), estimated from recurrence plots, and with a more straight-

forward method based on the maximum eccentricity achieved by the planet over the 1 million year integration. The eccentricity is an indication of the habitability of a terrestrial planet in the HZ; any value of $e > 0.2$ produces a significant temperature difference on a planet's surface between apoapse and periapse.
The results for possible stable orbits for terrestrial planets in habitable zones for the five systems are: for Gl 777 A nearly the entire HZ is stable, for 47 Uma, HD 72659 and HD 4208 terrestrial planets can survive for a sufficiently long time, while for Gl 614 our results exclude terrestrial planets moving in stable orbits within the HZ.
Studies such as this one are of primary interest to future space missions dedicated to finding habitable terrestrial planets in other stellar systems. Assessing the likelihood of other habitable planets, and more generally the possibility of extraterrestrial life, is the central question of astrobiology today. Our investigation indicates that, from the dynamical point of view, habitable terrestrial planets seem to be compatible with many of the currently discovered extrasolar systems.

8.1.1 Simulation method and stability analysis

The availability of a supercomputer with 128 processors[1] for this investigation enabled the direct computation of orbits to assess stability. Furthermore, the extent of the computational resources favoured the use of a very precise numerical integration scheme, the Lie-integration method, which is free from numerical difficulties experienced by other (lower order) techniques, particularly in the case of highly eccentric orbits. The Lie-integration method uses an adaptive stepsize and is quite precise and fast, as has been shown in many comparative test computations with other integrators such as Runge-Kutta, Bulirsch-Stoer or symplectic integrators. Although symplectic integrators are very effective when eccentricities remain small, the Lie integrator is a better choice in studies such as this one, where very large eccentricity orbits are explored. Details about this integration method can be found in Hanslmeier & Dvorak [Hanslmeier & Dvorak, 1984] as well as Lichtenegger [Lichtenegger, 1984].
For the analysis of the stability we used a straightforward check based on the eccentricities. For this we examined the behaviour of the eccentricity of the terrestrial planets along their orbit and used the *largest* value as a stability criterion; in the following we call it the maximum eccentricity method (MEM). This simple check has already been used in other studies of this kind and was found to be quite a good indicator of the stability character of an orbit (Dvorak et al. [Dvorak et al., 2003a]). An orbit was deemed unstable when the eccentricity exceeded a value of $e = 0.5$, after which we stopped further computation. In all former studies this stability limit turned out to be an appropriate tool because

[1]The PEYOTE cluster at the Max Planck Institute for Gravitational Physics (Albert Einstein Institute): www.aei-potsdam.mpg.de/facilities/public/computers.html

all the terrestrial planet orbits with $e = 0.5$ turned out to suffer, sooner or later, from a close encounter with the large planet, causing the terrestrial planet to escape (Dvorak et al. [Dvorak et al., 2003a]). Although some orbits in multiple exosolar systems have eccentricities larger than 0.5 their special configuration allows non-crossing orbits. By placing additional fictitious planets between their orbits with eccentricities larger than 0.5 they always suffer in the long run from close encounters leading to unstable orbits.
For the habitability of a planet we also used an additional criterion based directly on the eccentricity of the orbit within the HZ. This was done in order to take the variations in the "solar" insolation on the surface of the terrestrial planet into account. A good approximation (Lammer 2004; private communication), requiring $e < 0.2$, is sufficient to keep this variation in insolation small enough during an orbit.
On the other hand we computed K_2 by means of RPs to determine how predictable an orbit is. These values are comparable to the Fast Lyapunov Indicators (FLI) introduced by Froeschlé et al. [Froeschlé et al., 1997]. The RP based method is slower than determining the FLIs but it has the advantage that it can be used a posteriori.
The two methods used are complementary because on one side the MEM is the right tool to assure the stability of an orbit in the sense of being habitable (the eccentricities stay small) whereas on the other side the entropy method gives a direct measure of chaos and unveils the resonance structure of phase space in more detail (compare e.g. figures 8.8 and 8.9 for HD 72659).

8.1.2 The stability within resonances

We know from our planetary system about the importance of resonances between the mean motions of two bodies. Because all our planets have only small orbital eccentricities there are different simplified models available which can provide interesting results concerning the structure of phase space inside these resonance (e.g. Engels and Henrard [Engels & Henrard, 1994], Malhotra [Malhotra, 1998]). Most of these models can be used for the natural satellites (e.g. Malhotra and Dermott [Malhotra, 1990]) but in the case of extrasolar planets with large eccentricities theoretical considerations are not yet obtained. Therefore we choose to use a numerical approach for each system separately which we describe in the following.
For the investigation of the resonances, we choose initial conditions placed in the most relevant mean–motion resonances (MMRs) of the fictitious planet with the Jovian planet inside but also outside the HZ. These resonances were checked for stability in 8 different positions of the terrestrial planet (corresponding to $M = 0°, 45°, 90°, 135°, 180°, 225°, 270°, 315°$). Additionally the computations were carried out with the Jovian planet initially placed at the apoastron and periastron. For a detailed list of the resonant positions that were investigated for each system

Table 8.1: Stability of orbits in mean motion resonances. The numbers give the stable orbits according to the 8 different initial conditions

	Gl 777 A		47 Uma[a]		HD 72659		Gl 614[b]		HD 4208[c]	
	P	A	Mode I	Mode II	P	A	P	A	P	A
5:1	1	0	8	6	8	8	0	0	8	8
4:1	0	0	3	0	8	8	0	0	8	8
3:1	0	0	0	0	4	3	0	0	8	7
5:2	0	0	6	7	2	2	0	0	3	2
7:3	0	0	1	1	1	0	0	0	8	8
2:1	0	0	1	2	3	3	0	0	8	8
5:3	0	0	–	–	2	1	–	–	7	5
3:2	2	1	–	–	0	2	0	0	7	8
4:3	0	0	4	0	–	–	–	–	3	1
Sum [%]	4.7	1.6	28.6	69.6	42.2	43.8	0.0	0.0	83.3	76.4
Total Sum [%]	3.1		34.9		43.0		0.0		79.9	

[a]Note that in the case of the 47 Uma system, where two Jovian planets are known, we did not use peri- and apoastron position as initial conditions, but 2 different modes corresponding to an aligned or anti-aligned configuration of the two major bodies.

[b]Besides the 7 given resonances, we calculated the motion inside the 7:2, 9:2 and 8:3 resonance – again, we only found unstable motion.

[c]For this system, we calculated all resonances up to the 4^{th} order (see section 4); with the exceptions of the 15:11 and the 13:9 MMRs the other resonant positions showed the same amount of predominantly stable motion.

see table 8.1.

As an example we discuss the results of the investigation of the MMRs for the system HD 4208. We studied the following mean motion resonances up to the fourth order: 2:1, 3:2 and 4:3 (first order); 3:1, 5:3, 7:5 (second order); 4:1, 5:2, 7:4, 8:5, 10:7 and 11:8 (third order); 5:1, 7:3, 11:7, 13:9, 15:11 (fourth order). As shown in Fig. 8.1 for the first set of resonances (1^{st} and 2^{nd} order) the orbits close to the central star, which move well inside the HZ, are all stable. For the MMRs close to the Jovian planet we can see a preference for stable orbits for the initial conditions $M = 0°$, and $180°$. For the 3^{rd} order resonance (Fig. 8.2) the picture is very inhomogeneous; for the Jovian planet in the apoastron position the orbits are stable even for the resonances close to the giant planet. The 4^{th} order resonances are not destabilising an orbit as we can see from figure 8.3; most of them are stable! The percentage of stable orbits in resonances is very large for HD 4208.

Details of the results from the investigation of resonances for all systems can be found in table 8.1. For the resonances acting in three of the systems far from the

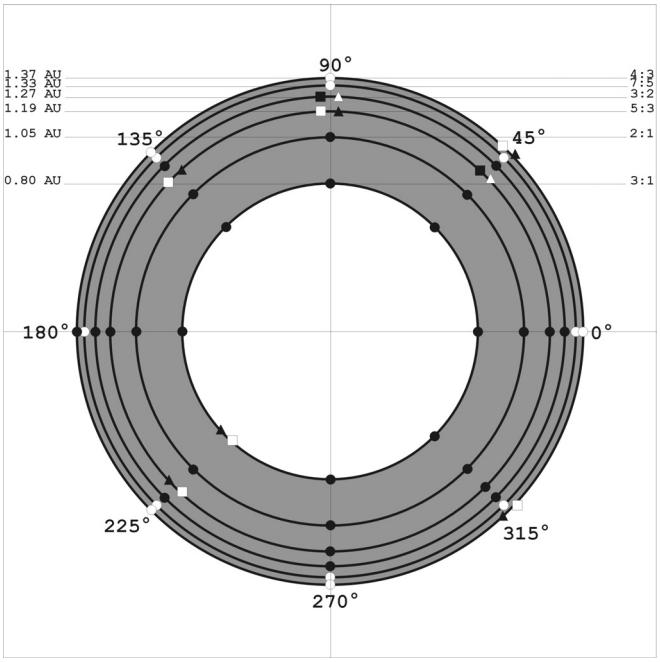

Figure 8.1: Schematic view of the stability of orbits in the resonances of 1st and 2nd order in HD 4208. Full (empty) circles stand for stable (unstable) orbits in apoastron and periastron position. When the stability is different we mark the apoastron by a triangle, the periastron with a square.

perturbing planet almost all of them are stable; closer to the perturbing planet, they are more and more unstable (47 Uma, HD 72659 and HD 4208). Two systems are very much dominated by unstable motion in resonances: Gl 777 A and Gl 614.

The results of the orbital computations which we started exactly in the resonances (with the properly chosen semimajor axis which corresponds to the MMR we were investigating) are in a certain sense redundant. For some resonances (e.g. the 3:1 MMR) the starting positions of the giant planet in the pericenter or the apocenter and the starting position of the terrestrial planet on the connecting line in between – on a circular orbit – should give the same results. Also there is a symmetry for the initial mean anomalies M and $M + 180°$ for the fictitious planet for some of the resonances (e.g. the 2:1 MMR). The reason that we have undertaken the computations for all positions is the following: when the results for the stability analysis were not the same – even when they should be equal – we took it as a sign of unstable motion. What we observed in our computations, which is in fact a known property, is that the unavoidable numerical errors can reveal the stability character of an orbit. Thus the results presented in table 3 can be regarded as a good estimate for the stability of motions in a MMR.

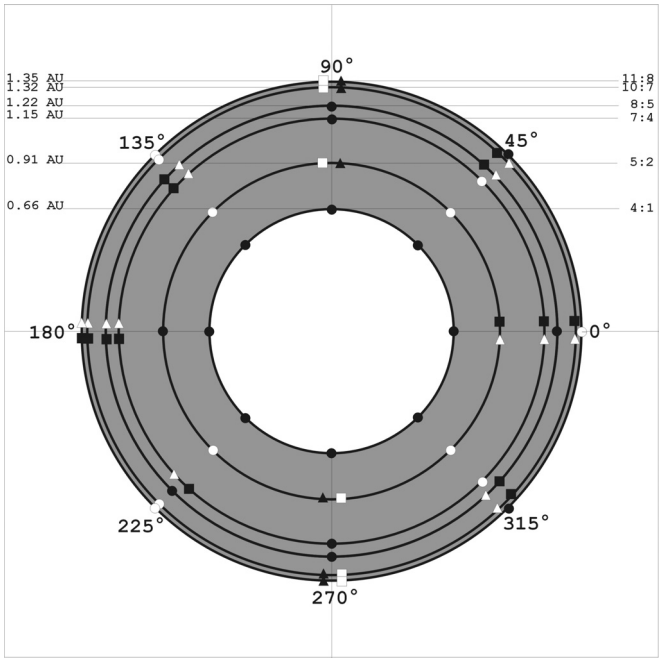

Figure 8.2: Schematic view of the stability of orbits in the resonances of 3rd order in HD 4208. Description like in figure 8.1.

We emphasise that a detailed theoretical study for larger eccentricities is highly desirable.

8.1.3 Gl 777 A

The first discovery of a planet in Gl 777 A (=HD 190360) was reported by Naef et al. ([Naef et al., 2003]) from the Geneva group of observers. This extrasolar planetary system is a wide binary with a very large separation (3000 AU); for our dynamical investigations of motions close to one star there was no need to take into account the perturbations of the very far companion. The central star is of spectral type G6 IV with $0.9 M_{\odot}$ and has a planet of minimum mass 1.33 $M_{\rm J}$ with a semimajor-axis of 4.8 AU. Because of the large eccentricity (e=0.48) the possible region of motion for additional planets is confined to $a < 2.4$ AU (= periastron). Nevertheless, to have a global stability picture of possible additional planets, we investigated the stability in the region of the MMR from the 4:3 to the 5:1 resonance located at a=1.64 AU. From table 8.1 one can see that only a few percent of the orbits started in the MMRs are stable.
The interesting region of habitability (see figure 8.4), where planets could have temperature conditions to allow liquid water on the surface, corresponds roughly to $0.7 < a < 1.3$ AU, where we ignore the eccentricity of the terrestrial planet. We have started our computations in a larger region ($0.5 < a < 1.3$) with a grid

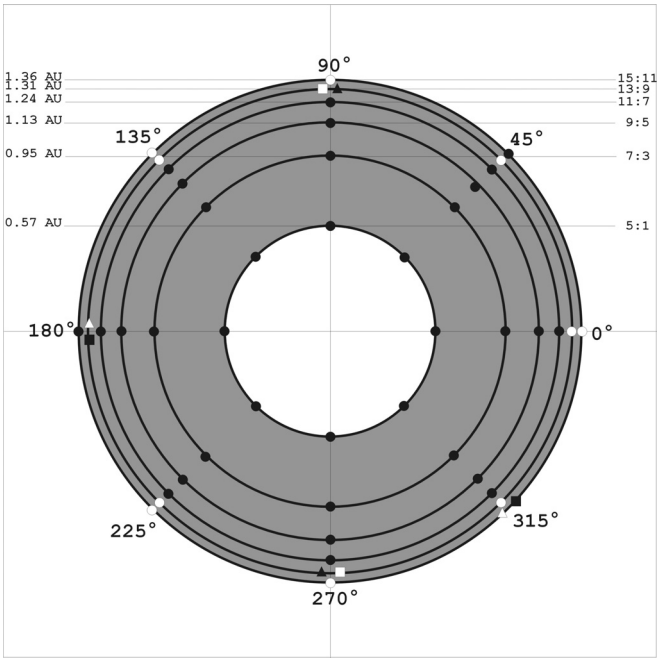

Figure 8.3: Schematic view of the stability of orbits in the resonances of 4^{th} order in HD 4208. Description like in figure 8.1.

spacing of $\Delta a = 0.01$ AU and changed also the eccentricity of the known planet between $0.4 < e < 0.5$ with a gridsize of $\Delta e = 0.01$. The results of the two methods of analysis of the orbital behaviour are shown in figures 8.5 and 8.6. In the first plot we show the results of the MEM, where two features are immediately visible: 1) strong vertical lines due to high order resonances, and 2) unstable orbits due to high eccentricity and high semimajor axes values (red or yellow colours). The latter feature is easy to understand because closer to the existing planet the perturbations are larger. The two methods complement each other in the information they convey; the MEM tells us about the variable distance to the central star and consequently it is a direct measure of the differential energy flux (insolation) on the planet. We can therefore determine where the variation of this distance does not exceed 50 percent, corresponding to an eccentricity of $e = 0.2$. The Rényi entropy of second order is a more sensitive probe of the dynamical character of the orbit, giving us a measure of the degree of chaos. In particular high order resonance features are made very clear using this second method, and we can even see the resonances acting when the eccentricity of the planet is as low as $e = 0.4$ (the bottom of Fig. 8.6).

As a result for habitability of a terrestrial planet inside the orbit of the Jovian planet, we find that for the system Gl 777 A there is quite a good chance that planets will last long enough in the HZ to acquire the necessary conditions for

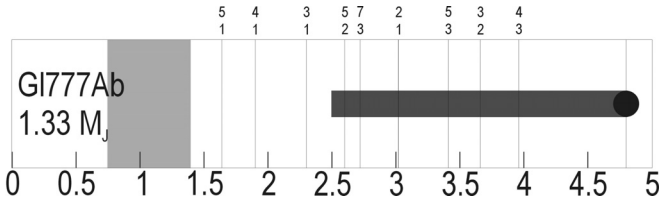

Figure 8.4: Main characteristics of the extrasolar system Gl 777 A. The light grey region shows approximately the position of the HZ; the dark grey bar indicates how closely the planet approaches the central star in its orbit.

life in the region with $a < 1$ AU.

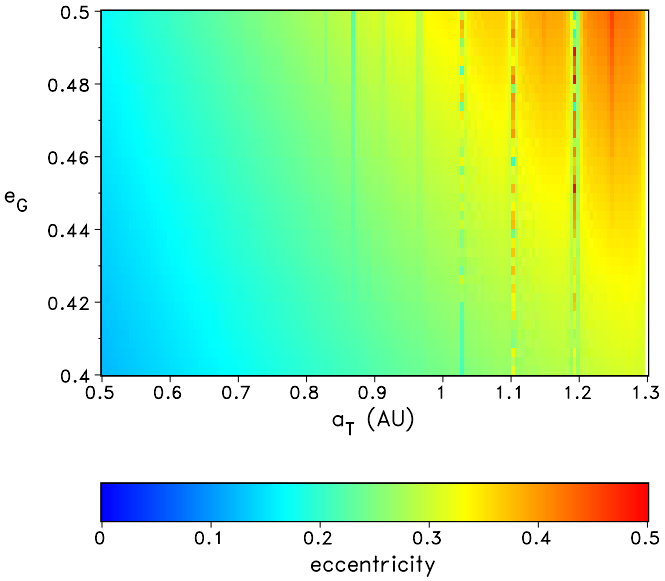

Figure 8.5: Initial condition diagram for fictitious planets in the system Gl 777 A: initial semimajor axes of the planet versus the eccentricity of the Jovian planet. The maximum eccentricity of an orbit during its dynamical evolution is marked with different colours.

8.1.4 HD 72659

The G05 star HD 72659 was found to have a companion from the Keck Precision Doppler survey ([Butler et al., 2002]). The Jovian planet (2.55 M_J) has an orbit with a semimajor axis $a = 3.24$ AU and an eccentricity of $e = 0.18$. The MMRs are located from 2.47 AU (3:2) to 1.1081 AU (5:1); the 5:1, 4:1 and 3:1 are well inside the periastron position of 2.657 AU and lie in the HZ (around 1 AU, see Fig. 8.7). The resonances turned out to be stable in more than 40% of the orbits;

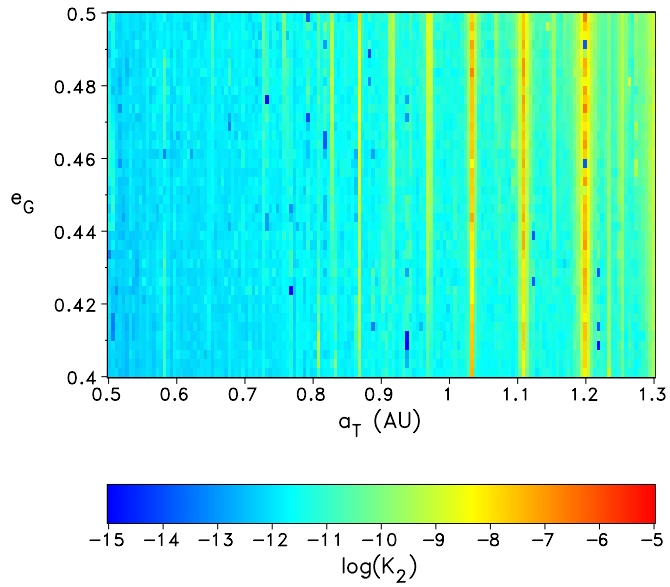

Figure 8.6: Initial condition diagram for fictitious planets in the system Gl 777 A: initial semimajor axes of the planet versus the eccentricity of the Jovian planet. The value of the entropy (=entropy plot) of an orbit is marked in different colours.

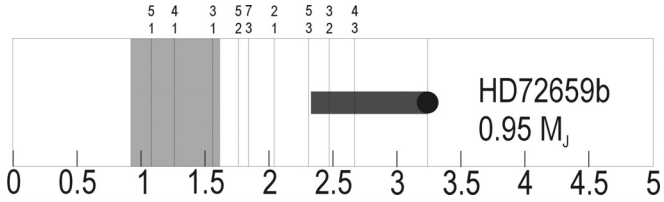

Figure 8.7: Main characteristics of the extrasolar planetary system HD 72659. Specifications like in figure 8.4.

especially the high order resonances close to the HZ are stable in both initial conditions (periastron and apoastron position). As a consequence we expected that these planetary systems may host additional terrestrial planets in stable orbits. Because of the uncertainties of the observed Jovian planet's eccentricity we varied it from 0.08 to 0.30 with a stepsize of $0.22/80 = 0.00275$ and chose the initial semimajor axis of the fictitious terrestrial planet to satisfy 0.4 AU $< a <$ 1.2 AU. The results are shown in Fig. 8.8 (MEM) and Fig. 8.9 (entropy plot). We can identify quite well in these plots the resonances up to the 7:1 resonance (only in the entropy plot). Again one can see that the dynamics of a single orbit can be determined quite well with this method; it not only confirms what is depicted in figure 8.8, it also shows many more details especially for the motions in resonances. On the contrary the MEM is the appropriate tool for

determining the eccentricity, which is – together with the semimajor axes – the crucial parameter for our research of determining planets in habitable zones.

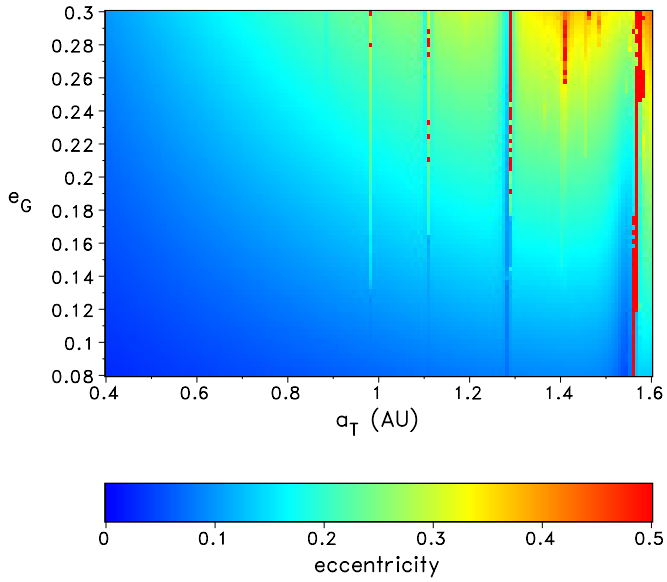

Figure 8.8: Results of the MEM for HD 72659.

Globally we can see a quite stable HZ in this extrasolar system which allows planets on orbits with small eccentricities. The strong unstable line close to 1.6 AU corresponds to the 3:1 resonance, while the other resonances, although giving rise to large perturbations in the eccentricities, are confined to the centre of the resonance. In figure 8.8 we can also see that for the most probable eccentricity value of the Jovian planet ($e = 0.18$) all orbits up to the 3:1 resonance are stable with low eccentricities ($e < 0.2$) and as a consequence the HZ could be populated by a terrestrial planet (or even more planets depending on their masses).

8.1.5 Conclusions

We have carried out a dynamical study of five extrasolar planetary systems using extensive numerical experiments to answer the question of whether they could host terrestrial-like planets in habitable zones. For the single–planet systems we used the elliptic restricted three body problem, and for the system 47 Uma the mutual perturbations of the two Jovian planets were taken into account and therefore the restricted four-body problem served as a dynamical model. Because of the dependence of the stability of an orbit in a mean motion resonance on angular position we have selected 8 different positions for the mean anomaly of the planet with a step of 45° in the apoastron and the periastron position of the Jovian planet.

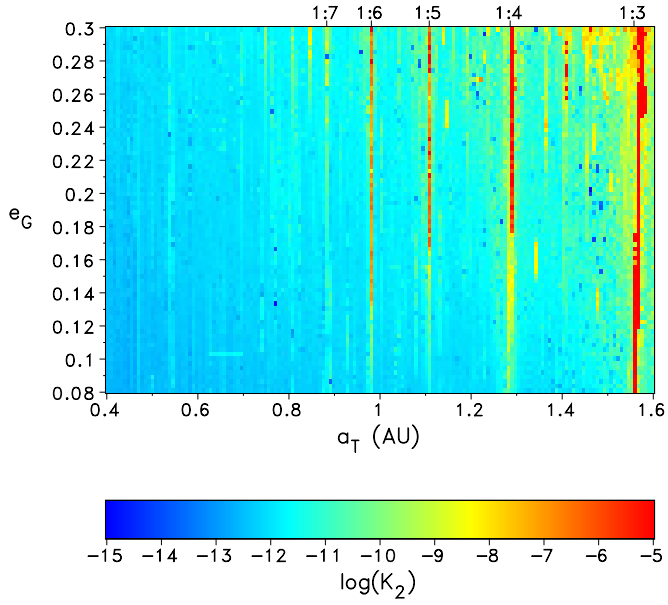

Figure 8.9: Entropy plot for HD 72659.

The characteristics of each system dictated the initial conditions, chosen in a fine 80 by 80 grid within the habitable zone, from which the orbits were computed using a robust numerical method (Lie–series integration) for 1 million years. The grid of the initial conditions of the fictitious terrestrial planets was chosen to cover the whole habitable zone of the system, and also to model the uncertainties in the elements of the observed planet(s). The stability of orbits was assessed with two methods, namely the computation of the Rényi entropy of second order as measure of the chaoticity of an orbit and the determination of the maximum eccentricity of the orbit of a fictitious planet during its orbital evolution of 1 million years.

We can say that our computations for such a fine grid, taking into account also the essential role of the MMRs, lead to a deeper insight concerning the dynamics of the five systems which we studied. We also give the percentage of orbits which survived in the paper (=MT) of Menou & Tabachnik ([Menou & Tabachnik, 2003]) where they investigated all known extrasolar planetary systems with respect to possible additional terrestrial planets. We note that a direct comparison of MT with the percentages of 'our' survivors is not useful here because of the different approaches used; we have emphasised the role of the MMRs and neglected possible inclinations. However, we know that terrestrial planets will form within a protoplanetary disk thus staying with small orbital inclinations (Richardson et al. [Richardson et al., 2000]; Lissauer [Lissauer, 1993]); additionally in a recent publication (Dvorak et al. [Dvorak et al., 2003a]) it was shown that the inclinations of the fictitious planets up to 15 degrees do not change the stability of orbits

in the HZ. The results for possible stable orbits may be summarised as follows:

- In the system **Gl 777 A** the stability zone for the motion of terrestrial planets is well inside of the HZ and suggests any planets residing there will survive for a sufficiently long time (in MT, 86.8 % of the orbits were found to be stable).
- In the system **47 Uma** we can also say that there is a good chance for planets to move inside the HZ with small eccentricities between the main resonances; this is a result which is consistent with others but not with MT where only 28 % remained.
- **HD 72659** turned out to be a very good candidate for hosting planets in the HZ; again this does not confirm the results of MT – they found that only 40.2 % of the orbits were stable.
- The results of the computations for **Gl 614** show that it is very unlikely that there is an additional planet moving in the HZ – these results are consistent with MT (9.2 % stable orbits)
- For **HD 4208** In the HZ there is enough room left for terrestrial planets and that they could survive for a sufficiently long time; these results are more or less consistent with those of MT where 50.2 % of the orbits were stable.

New observational possibilities provided by missions like COROT, DARWIN or the TPF make the first search for terrestrial exoplanets seem possible in the next decades. In an ESA study the goal of the missions is summarised as follows: "To detect and study Earth–type planets and characterise them as possible abodes of life". In this sense dynamical studies like the one we present here should help to define promising targets for observations.

8.2 Variability of the Earth Surface Temperature

The next application we present regards the earth's climate. Since the discovery of chaos in a conceptual climate model by Lorenz [Lorenz, 1963] the predictibility of weather (and climate) is still an open and not fully understood topic. The predictability of weather is strongly related to the so-called persistence of, i.e. how , e.g., the temperature of tomorrow depends on the temperature of today. [Koscielny-Bunde et al., 1998] found a long-term correlation between the daily temperature following a power law by analysing the data of 14 meteorological stations around the world. The correlation function decays with a universal exponent of ≈ 0.65. Later on, the detrended fluctuation analysis (DFA) method

has been applied in order to compare measured temperature data with state-of-the-art global circulation models of the climate [Bunde et al., 2003]. In this study we use observed as well as modelled gridded data on a global scale. We quantify the long-term predictability of mean daily temperature data from the General Circulation Model ECHAM (calculated at DKRZ) by means of the Rényi entropy of second order K_2. We are interested in the amplitude fluctuations of the temperature from year to year. Hence, the data are low-pass filtered. The obtained oscillatory signal has a constant frequency (earth rotation around the sun) but its amplitude fluctuates irregularly. We compare the results obtained with K_2 with the linear standard deviation analysis and also with the results obtained in the CRU data set (interpolated measured temperature in the year 1901-2003 with 0.5 degree resolution).

8.2.1 Description of the Data

For the measured data the Climate Research Unit (CRU) dataset [Mitschell et al., submitted] has been used. These monthly data on a 0.5 degree resolution have been constructed from available observed data all over the world. The primary purpose was to create an input data set for environmental modelling.
The global climate model consists of the spectral atmospheric model ECHAM4 [Roeckner et al., 1996] and the ocean model HOPE-G [Wolff et al., 1997], both developed at the Max-Planck-Institute of Meteorology in Hamburg. In this simulation the model ECHAM4 has a horizontal resolution of T30 (approx. 3.75° x 3.75°) and 19 vertical levels, five of them located above 200 hPa. The horizontal resolution of the ocean model HOPE-G is about 2.8° x 2.8° with a grid refinement in the tropical regions, where the meridional grid-point separation decreases progressively to the equator, reaching a value of 0.5°. This increased resolution allows for instance for a more realistic representation of ENSO events. The ocean model has 20 vertical levels.
In this simulation, the model has been driven by estimations of three past external forcing factors: solar variability, greenhouse gas concentrations in the atmosphere and an estimation of the radiative effects of stratospheric volcanic aerosols. No changes in the anthropogenic atmospheric aerosol concentrations have been considered. Changes in vegetation cover or land-use have also been neglected.
The atmospheric concentrations of two greenhouse gases, carbon dioxide and methane, have been estimated from analysis of air bubbles trapped in Antarctica ice cores ([Etheridge et al., 1996],[Blunier et al., 1995]). The past variations of solar output have been derived from the values used by Crowley [Crowley, 2000]. For the period after 1610 A.D., past solar variations are empirically estimated from observations of sun spots (Lean et al., 1995) and between 1500 and 1610 A.D. they are based on concentrations of the cosmogenic isotope ^{10}Be.
The third external factor is the stratospheric loading of volcanic aerosols.

After a volcanic eruption, these are washed out by precipitation, that influence the acidity of the ice layers in ice cores. Changes in optical densities of the stratosphere can be thus estimated from ice acidity through a semi-empirical model [Crowley, 2000].
The idea of how to estimate K_2 is linked to a weather prediction scheme proposed by Lorenz in 1963 [Lorenz, 1963]. He proposed to use naturally occurring analogues. The idea is to record long series of data describing the state of the atmosphere (or any other system under consideration). To predict the weather one then has to compare the actual state with all the states in the data bank and to identify a former state which is extremely close to the current one, so that the mismatch could also be attributed to a measurement error. Such state than is called an analogue. The prediction of the future is than given by the time evolution of the former state. Here we do not focus on the prediction but rather quantify the predictability of a system, i.e. the possibility to predict the system. Therefore, we identify close analogues of a given state and consider all their respective evolutions. Then we quantify the time that these possible evolutions stay similar. This time is directly linked to K_2.

8.2.2 Interpretation of Recurrence Plots with respect to Climate

Suppose we have a dynamical system represented by the trajectory $\{\vec{x}_i\}$ for $i = 1, \ldots, N$ in a d-dimensional phase space. Then we compute the recurrence matrix (Eq. (2.1)). As mentioned in Sec. 7.4.1 recurrences correspond to Lorenz's naturally occurring analogues. This means that all black points in the RP represent analogues. The analogues of a point $\vec{x}_i$ are visualised in the i-th column of the RP as black points. If $\vec{x}_i$ and $\vec{x}_j$ are analogues we have $R_{i,\,j} = 1$. If the evolution of $\vec{x}_i$ and $\vec{x}_j$ stays similar for n time step we have also $R_{i+1,\,j+1} = 1, \ldots, R_{i+n,\,j+n} = 1$, i.e. we have a diagonal line in the RP. The better the predictability of a system the longer are these diagonals. We will now use these diagonals to estimate K_2.

8.2.3 Results for the Two Data Sets

For the analysis of the data the automatic algorithm for the K_2 estimation (Chap. 5) had to be slightly modified. The problem is that due to the high dimensionality of the models and/or the random elements involved, the slope of $P_\varepsilon^c(l)$ is not independent of ε. This is an expected behaviour for systems with a rather high dimension and/or noise (Sec. 3.1). The entropy K_2 is also well defined in such cases but one has to estimate the limit $\varepsilon \to 0$. This limit can estimated by linear interpolation of the largest scaling region, in the "slope vs. RR" diagram (similarly as in Fig. 3.5). However, in this case we chose a different approach. As we considered temperature signals, we set $\varepsilon = 0.5$, which corresponds to half a

degree centigrade prediction error. The results of the application of the modified automatic algorithm are represented in Fig. 8.10. High variances seem to corre-

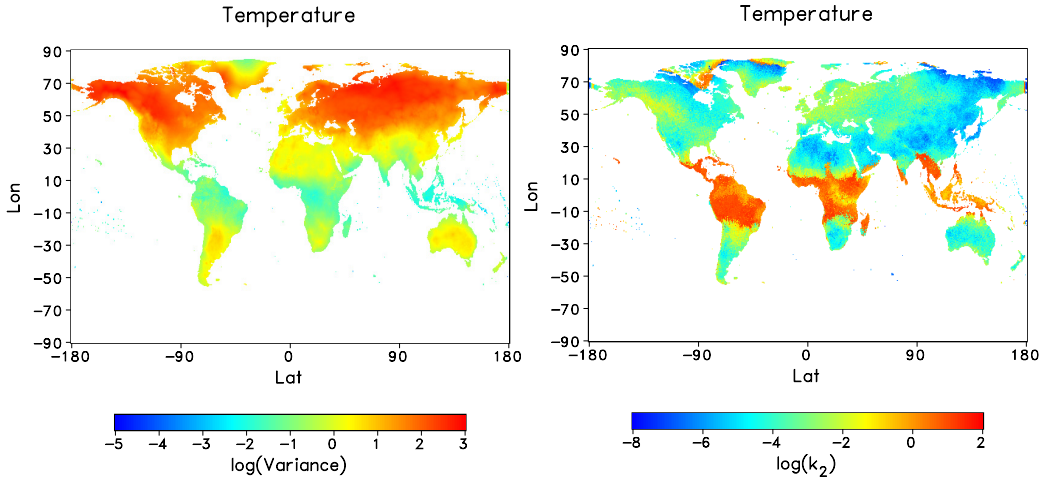

Figure 8.10: **Left panel**: Variances of the annual temperature fluctuations of the CRU data set. **Right panel**: K_2 estimates for the CRU data set.

spond to low entropies and vice versa. However, there are exceptions especially in Alaska but also in central Africa. Comparing this result with the one obtained for the GCM data (Fig. 8.11) one finds that there is a qualitative correspondence for both sets. The GCM data also allows to compute the corresponding values for oceanic regions. The GCM data gives an interesting result for the relationship between entropies and variances. In continental regions the variances and the entropies seem to be anticorrelated, whereas in oceanic regions they seem to be rather correlated.

Even though this study is still rudimentary, it shows that modified versions of the RP based K_2 estimation algorithm can be applied in various fields. Studies about the earth's climate are in progress in cooperation with the Potsdam Institute of Climate Impact Research (PIK).

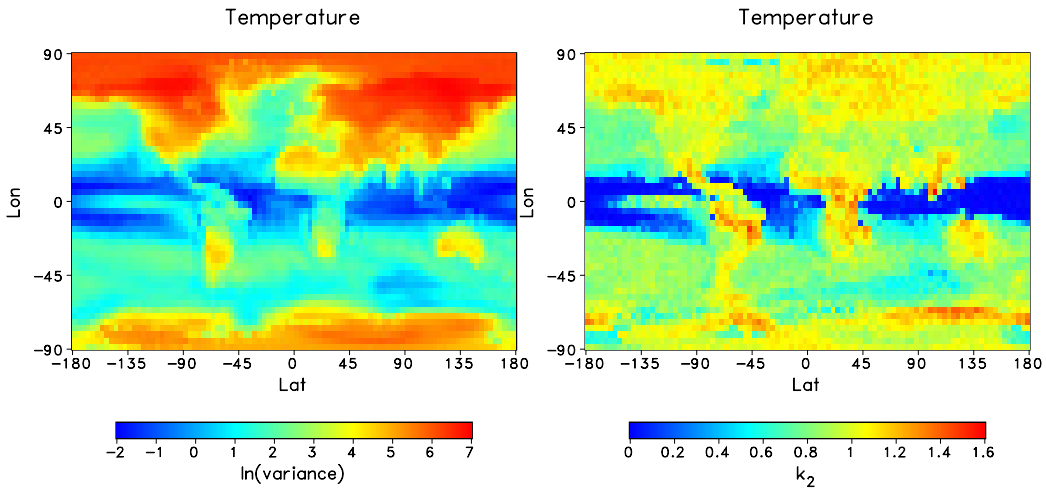

Figure 8.11: **Left panel**: Variances of the annual temperature fluctuations for the GCM data set. **Rigth panel**: K_2 estimates for the GCM data set.

Chapter 9

Conclusions

In this work we have exploited recurrences, which were termed "naturally occurring analogues" by Lorenz , for the analysis of time series. The three key results of this thesis are the following:

1. The diagonal structures, which dominate recurrence plots, are linked to generalised Rényi entropies and dimensions. These invariants can be estimated from recurrence plots by a rather robust algorithm, which can then be automated. Furthermore, their estimation by RPs is independent from the embedding parameters use for the attractor's reconstruction. It is also possible to evaluate off-diagonal structures, by estimating e.g. the mutual information.

2. The structure in recurrence plots determines the topology of the underlying attractor to a very high degree. If the recurrence matrix is computed from one component only, the underlying time series can be topologically reconstructed from it.

3. Recurrences allow not only to predict the evolution of a system, but also to generate alternative evolutions consistent with the system. They can be applied to test the reliability of phase synchronisation analysis. These surrogates have interesting dynamical properties. They seem to mimic the dynamical behaviour to a very high degree. They shadow real trajectories of the underlying system.

The algorithm for the estimation of K_2 has been applied to various systems such as the Rössler system but also to more complicated systems, e.g. a stability analysis of extrasolar planetary systems by means of K_2 has complemented results which were obtained with standard methods.
This work has led to further studies which are momentarily carried out. One of these studies is related to climate research. The task is to study the anthropogenic effect on the climate. K_2 estimations are applied to analyse climate data

and climate simulations.
At the JUMP cluster at the Jülich Rechenzentrum further extrasolar planetary systems are analysed. The RP method is permanently being refined and adapted to answer open questions in astrophysics.
A further field of interest is the analysis of physiological data, e.g. EEG and ECG data. Especially the surrogate test will help to analyse the so called "passive experiments", i.e. situations in which the system evades an experimental manipulation.
Furthermore, the construction of a hardware based RP evaluation tool is planned. This would increase the computation speed of RPs by a large amount and is highly desirable for the analysis of spacio-temporal systems.
Furthermore, a study about roundoff induced periodicity in chaotic systems is planned [Grebogi et al., 1988].

Acknowledgements

I would like to thank all the people, who have helped me to make this work possible.
First of all, I thank Prof. Dr. Kurths not only for giving me the opportunity to work in his group, but also for many and extremely helpful discussions. He has demonstrated a lot of patience and created a very stimulating scientific environment, which was very important for my studies.
Maria Carmen Romano has, of course, been the person who was a special counsellor: both scientifically and privately. She was literally always there when I needed someone.
I want to give my thanks to Dr. Udo Schwarz. He helped me when I started the studies in Potsdam. I had no experience in nonlinear dynamics at all and Udo discussed with me weekly, and proved to be a much better teacher, than he himself is aware of.
I am very grateful to Vadim Anishchenko, Dieter Armbruster, Jörg Assmuß, Werner von Bloh, Rudolf Dvorak, Fred Feudel, Celso Grebogi, Jack Hudson, Norbert Marwan, Eun-Hyoung Park, Michael Rosenblum, Norbert Seehafer, Annette Witt, Joseph Zbilut and Chang-Song Zhou and many others for fruitful discussions.
I want to thank Birgit Nader and Jörg Tessmer, both of which have supported me so much during the last years that I can hardly express my gratitude.
I also want to express my gratitude to the "Corellanos", who not only are good friends but also have confied her daughter to me.
Last but not least I want to thank my parents and my brother who have always been cheering me up. Besides writing my PhD, I had to "grow up" during the last years and they have helped me with that sometimes difficult task.
I know that I must have forgotten to mention many friends and colleagues here. But it is 3 am now and in 6 hours I have to have to be in the copy shop to bind this work. Hence, I am a bit stressed...

Bibliography

[Afraimovich et al., 1986] V. S. Afraimovich, N. N. Verichev, M. I. Rabinovich, Izvestiya Vysshikh Uchebnykh Zavedenii Radiofizika **29 (9)**, 1050 (1986).

[Alligood et al., 1996] K. T. Alligood, T. D. Sauer and J. A. Yorke, Chaos an introduction to dynamical systems, Springer, (1996).

[Anishchenko et al., 2004] V. S. Anishchenko, T. E. Vadivasova, J. Kurths, G. Okrokvertskhov, and G. I. Strelkova, Phys. Rev. E **69**, 036215 (2004).

[Arecchi et al., 1982] F. T. Arecchi, R. Meucci, G. Puccioni, and J. Tredicce, Phys. Rev. Lett. **49**, 1217 (1982).

[Asghari et al., 2004] N. Asghari, C. Broeg, L. Carone, R. Casas-Miranda, J.C. Castro Palacio, I. Csillik, R. Dvorak, F. Freistetter, G. Hadjivantsides, H. Hussmann, A. Khramova, M. Khristoforova, I. Khromova, I. Kitiashivilli, S. Kozlowski, T. Laakso, T. Laczkowski, D. Lytvinenko, O. Miloni, R. Morishima, A. Moro-Martin, V. Paksyutov, A. Pal, V. Patidar, B. Pečnik, O. Peles, J. Pyo, T. Quinn, A. Rodriguez, C. Romano, E. Saikia, J. Stadel, M. Thiel, N. Todorovic, D. Veras, E. Vieira Neto, J. Vilagi, W. von Bloh, R. Zechnervand E. Zhuchkova, A&A (accepted for publication)

[Baldi & Pollastri, 2003] P. Baldi, and G. Pollastri, J. Machine Learning Research **4**, 575 (2003).

[Bandt et al., 2002] C. Bandt, G. Keller, and B. Pompe, Nonlinearity **15**, 1595 (2002).

[Beck & Schlögl, 1993] C. Beck, F. Schlögl, Thermodynamics of Chaotic Systems, Cambridge University Press (1993).

[Blasius et al., 1999] B. Blasius, A. Huppert, and L. Stone, Nature **399**, 354 (1999).

[Blunier et al., 1995] T. Blunier, J. A. Chapellaz, J. Scwander, B. Stauffer, D. Raynaud, Nature **374**, 46 (1995).

[Boccaletti et al., 2002] S. Boccaletti, J. Kurths, G. Osipov, D. L. Valladares, and C. S. Zhou, Physics Reports **366**, 1 (2002).

[Bunde et al., 2003] A. Bunde, J. Eichner, R. Govindan, S. Havlin, E. Koscielny-Bunde, D. Rybski, and D. Vjushin, Power-Law persistence in the atmosphere: Analysis and Applications models in: Nonextensive entropy – Interdisciplinary applications, edited by M. Gell-Mann and C. Tsallis , Oxford University Press, New York (2003).

[Butler & Marcy, 1996] R. Butler, G. Marcy, AJ **464**, L153 (1996).

[Butler et al., 2002] R. Butler, G. Marcy, S. Vogt, et al., AJ **582**, 455 (2002).

[Crowley, 2000] T. J. Crowley, Science **289**, 270 (2000).

[Dechert & Gençay, 2000] W. D. Dechert, and R. Gençay, Physics Letters A **276**, 59 (2000).

[Ding et al., 1993] M. Ding, C. Grebogi, E. Ott, T. Sauer,and J. A. Yorke, Phys. Rev. Lett. **70**, 3872 (1993).

[van den Dool, 1994] H. M. van den Dool, Tellus, **46 A(3)**, 314 (1994).

[Dvorak et al., 2003a] R. Dvorak, E. Pilat-Lohinger, B. Funk, and F. Freistetter, A&A **398**, L1 (2003).

[Eckmann et al., 1987] J.-P. Eckmann, S. O. Kamphorst, and D. Ruelle, Europhysics Letters **4**, 973 (1987).

[Engels & Henrard, 1994] J. Engels, and J. Henrard, CMDA **58**, 215 (1994).

[Etheridge et al., 1996] D. Etheridge, L. P. Steele, R. L. Langenfelds, R. J. Francey, J. M. Barnola, and V. I. Morgan, J. Geophys. Res. **101**, 4115 (1996).

[Farmer, 1982] J. D. Farmer, Physica D **4**, 366 (1982).

[Faure & Korn, 1998] P. Faure, and H. Korn, Physica D **122**, 265 (1998).

[Froeschlé et al., 1997] C. Froeschlé, E. Lega, R. Gonczi, CMDA **67**, 41 (1997).

[Fujisaka & Yamada, 1983] H. Fujisaka, and T. Yamada, Prog. Theor. Phys. **69 (1)**, 32 (1983).

[Gabor, 1946] D. Gabor, J. IEE London **93**, 429 (1946).

[Gallas, 1994] J. Gallas, Physica A **202**, 196 (1994).

[Gao & Cai, 1999] J. Gao, and H. Cai, Physics Letters A **270**, 75 (1999).

[Gao, 1999] J. Gao, Phys. Rev. Lett. **83**, 3178 (1999).

[Gonzalez-Rouco et al., 2003] F. Gonzalez-Rouco, H. von Storch, and E. Zorita, Geophysical Research Letters **30**, 2116 (2003). thousand years.

[Grassberger & Procaccia, 1983a] P. Grassberger, and I. Procaccia, Phys. Rev. Lett. **50 (5)**, 346 (1983).

[Grassberger & Procaccia, 1983b] P. Grassberger, and I. Procaccia, Physica D **9** 189 (1983). of strange attractors,

[Grassberger & Procaccia, 1983c] P. Grassberger, and I. Procaccia, Physical Review A, 2591 (1983).

[Grassberger, 1983] P. Grassberger, Physics Letters A **97 (6)**, 227 (1983).

[Grassberger & Procaccia, 1984] P. Grassberger, and I. Procaccia, Physica D **13** 34 (1984).

[Grassberger et al., 1991] P. Grassberger, T. Schreiber, and C. Schaffrath, Int. J. Bif. Chaos **1 (3)**, 521 (1991).

[Grebogi et al., 1985] C. Grebogi, S. W. McDonald, E. Ott, and J. A. Yorke, Phys. Lett. A **110**, 1 (1985).

[Grebogi et al., 1988] C. Grebogi, E. Ott, and J.A. Yorke, Phys. Rev. A **38**, 463 (1988).

[Grebogi et al., 1990] C. Grebogi, S. M. Hammel, J. A. Yorke, and T. Sauer, Phys. Rev. Lett. **65**, 87 (1990).

[Hanslmeier & Dvorak, 1984] A. Hanslmeier, and R. Dvorak, A&A, **132**, 203 (1984).

[Holyst et al., 2001] J. A. Holyst, M. Zebrowska, and K. Urbanowicz, Eur. Phys. J. B **20**, 531 (2001).

[Hübner et al., 1993] U. Hübner, C-O. Weiss, N. Abraham, and D. Tang, Time Series Prediction: Forecasting the Future and Understanding the Past., Wigend and Gershenfeld Eds., Addison-Wesley, Reading, MA, (1993).

[Itoh et al., 1999] M. Itoh, T. Yang, and L. O. Chua, Int. J. Bifurc. Chaos **9** 1393 (1999).

[Iwanski, 1998 **] J. S. Iwanski, and E. Bradley, Chaos **8 (4)**, 861 (1998).

[Kantz & Schreiber, 1997] H. Kantz, and T. Schreiber, University Press, Cambridge, (1997).

[Katok & Hasselblatt, 1995] A. Katok, and B. Hasselblatt, Introduction to the Modern Theory of Dynamical Systems, Cambridge University Press (1995).

[Kennel et al., 1992] M. B. Kennel, R. Brown, and H.D.I. Abarbanel, Phys. Rev. A **45 (6)**, 3403 (1992).

[Khinchin, 1957] A. I. Khinchin, Mathematical Foundations of Information Theory, Dover Publ., New York (1957).

[Kiss et al., 2004] I. Z. Kiss, Q. Lv, and J. L. Hudson, submitted to PRL.

[Koebbe & Mayer-Kress, 1992] M. Koebbe, and G. Mayer-Kress, Use of Recurrence Plots in the Analysis of Time-Series Data, Nonlinear Modeling and Forecasting, in Proceedings of SFI Studies in the Science of Complexity, Vol. XXI, 361–378, Addison-Wesley, Santa Fe, (1992).

[Koscielny-Bunde et al., 1998] E. Koscielny-Bunde, A. Bunde, S. Havlin, H.E. Roman, Y. Goldreich, and H.J. Schellnhuber, Phys. Rev. Lett. **81**, 729 (1998).

[Kugiumtzis, 2000] D. Kugiumtzis, Phys. Rev. E **62**, R25 (2000).

[Kurths & Herzel, 1987] J. Kurths, and H. Herzel, Physica D **25**, 165 (1987).

[Kurths et al., 1994] J. Kurths, U. Schwarz, C.P. Sonett, and U. Parlitz, Nonlinear Processes in Geophysics **1**, 72 (1994).

[van Leeuwen, 2003] P. Van Leeuwen, D. Geue, S. Lange, D. Cysarz, H. Bettermann, and D. Grñemeyer2, BMC Physiology **3**, (2003).

[Leven et al., 1994] R. W. Leven, B. P. Koch, and B. Pompe, Chaos in dissipativen Systemen, Akademie Verlag, 1994.

[Lichtenegger, 1984] H. Lichtenegger, CMDA **34**, 357 (1984).

[Lissauer, 1993] J. J. Lissauer, ARA&A, **31**, 129 (1993).

[Lorenz, 1963] E. N. Lorenz, J. atmos. Sci. **20**, 130 (1963).

[Lorenz, 1969] E. N. Lorenz, J. atmos. Sci. **26**, 636 (1969).

[Mackey & Glass, 1977] M. C. Mackey, and L. Glass, Science **197**, 287 (1977).

[Malhotra, 1998] R. Malhotra, Orbital Resonances and Chaos in the Solar System. In Solar System Formation and Evolution: ASP Conference Series 149 (eds. D. Lazzaro, R. Vieira Martins, S. Ferraz-Mello, J. Fernandez, C. Beauge), 37 (1998).

[Malhotra, 1990] R. Malhotra, and S.F. Dermott, Icarus **85**, 444 (1990).

[Manetti et al., 1999] C. Manetti, M.-A. Ceruso, A. Giuliani, C. L. Webber Jr., and J. P. Zbilut, Physical Review E **59 (1)**, 992 (1999).

[Markus, 1995] M. Markus, Lyapunow-Diagramme, Spektrum der Wissenschaft, April, 66 (1995)/ Los diagramas de Lyapunov, Investigación y Ciencia, Sept., 70 (1995).

[Markus & Tamames, 1996] M. Markus, and J. Tamames, Fat Fractals in Lyapunov Space, in: Fractals in the Future (ed. by C. A. Pickover), St. Martin's Press, N. Y., 333 (1996).

[Marwan, 1999] N. Marwan, Untersuchung der Klimavariabilität in NW Argentinien mit Hilfe der quantitativen Analyse von Recurrence Plots, Diploma thesis, Dresden University of Technology (1999).

[Marwan et al., 2001] N. Marwan, M. Thiel, and N. R. Nowaczyk, Nonlinear Processes in Geophysics **9 (3-4)**, 325 (2002).

[Marwan et al., 2002] N. Marwan, N. Wessel, U. Meyerfeldt, A. Schirdewan, and J. Kurths, Phys. Rev. E **66**, 026702 (2002).

[Marwan & Kurths, 2002] N. Marwan, and J. Kurths, Phys. Lett. A **302 (5-6)**, 299 (2002).

[Marwan, 2003] N. Marwan, Encounters With Neighbours - Current Developments Of Concepts Based On Recurrence Plots And Their Applications.- Ph.D. Thesis, Institute of Physics, University of Potsdam, ISBN 3-00-012347-4, (2003).

[Menou & Tabachnik, 2003] K. Menou, and S. Tabachnik, AJ **583**, 473 (2003).

[Mitschell et al., submitted] T. Mitchell, T. Carter, P. Jones, M. Hulme, M. New, Journal of Climate, submitted.

[Naef et al., 2003] D. Naef, M. Mayor, S. G. Korzennik, et al., A&A **410**, 1051 (2003).

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

[Pikovsky, 1984] A. Pikovsky, Z.Physik B **55 (2)**, 149 (1984).

[Pikovsky et al., 2001] A. Pikovsky, M. Rosenblum, and J. Kurths, Synchronization, Cambridge University Press, (2001).

[Poincaré, 1890] H. Poincaré, Acta Math. **13**, 1 (1890).

[Pompe, 1993] B. Pompe, J. Stat. Phys. **73**, 587 (1993).

[Raab & Kurths, 2001] C. Raab, and J. Kurths, Phys. Rev. E **64**, 0162161 (2001).

[Read et al., 1992] P. L. Read, M. J. Bell, D. Johnson, and R. M. Small, J. Fluid Mech. **238**, 599 (1992).

[Rényi, 1970] A. Rényi, *Probability theory*, (North-Holland, 1970 (appendix)).

[Richardson et al., 2000] D. C. Richardson, T. Quinn, J. Stadel, and G. Lake, Icarus **143**, 45 (2000).

[Roeckner et al., 1996] E. Roeckner, L. Arpe, L. Bengtsson, M. Christoph, M. Claussen, L. Dümenil, M. Esch, M. Giorgetta, U. Schlese, and U. Schulzweida, Atmospheric general circulation model ECHAM-4: Model description and simulation of present-day climate. MPI Report No. 218, Max-Planck-Institute for Meteorology, Hamburg, Germany (1996).

[Romano et al., 2004] M. C. Romano, M. Thiel, J. Kurths, and W. von Bloh, Phys. Lett. A (in press).

[Romano et al., submitted] M. C. Romano, M. Thiel, J. Kurths, I. Z. Kiss, and J. L. Hudson, submitted to Phys. Rev. Lett.

[Rössler, 1976] O. E. Rössler, Phys. Lett. 57A, 397 (1976).

[Rosenblum et al., 1996] M. G. Rosenblum, A. S. Pikovsky, and J. Kurths, Phys. Rev. Lett. **76**, 1804 (1996).

[Ruelle, 1990] D. Ruelle, Deterministic chaos: the science and the fiction, Proc. R. Soc. Lond. A, **427** 241 (1990).

[Rulkov et al., 1995] N. F. Rulkov, M. M. Sushchik, L. S. Tsimring, and H. D. I. D. Abarbanel, Phys. Rev. E, **51 (2)**, 980 (1995).

[Sauer & Yorke, 1993] T. Sauer, and J. A. Yorke, Int. J. Bifurcation and Chaos **3**, 373 (1993).

[Schaefer et al., 1998] C. Schaefer, M. G. Rosenblum, J. Kurths, and H. H. Abel, Nature **392**, 239 (1998).

[Schreiber & Schmitz, 1996] T. Schreiber, and A. Schmitz, Phys. Rev. Lett. **77**, 635 (1996).

[Schreiber, 1998] T. Schreiber, Phys. Rev. Lett. **80 (10)**, 2105 (1998).

[Schreiber & Schmitz, 2000] T. Schreiber, and A. Schmitz, Physica D **142**, 346 (2000).

[Schuster, 1984] H. G. Schuster, Deterministic Chaos, Wiley VCH, (1984).

[Small et al., 2002] M. Small, R. G. Harrison, and C. K. Tse, A surrogate test for pseudo-periodic Time Series Data., In: Experimental Chaos, American Institute of Physics, ed. S. Boccaletti et al., 190-195 (2002).

[So et al., 1996] P. So, S. J. Schiff, D. T. Kaplan, E. Ott, T. Sauer, and C. Grebogi, Phys. Rev. Lett **76**, 4705 (1996).

[Sousa Veiera et al., 1992] M. de Sousa Vieira, P. Khoury, A. J. Lichtenberg, Int. J. Bifurc. Chaos **2** 645 (1992).

[Späth, 1992] H. Späth, Cluster Dissection and Analysis, Horwood, Chichester (1992).

[Sprott, 2001 ***] J. C. Sprott, and G. Rowlands, Int. J. Bif. Chaos **11 (7)**, 1865 (2001).

[Steeb, 1994] W. H. Steeb, Chaos und Quantenchaos in Dynamischen Systemen (BI Wissenschaftsverlag, 1994).

[Takens, 1980] F. Takens, Detecting strange attractors in turbulence, in: Dynamical Systems and Turbulence, eds. D. A. Rand and L.-S. Young, Lecture Notes in Mathematics, Vol. 898, Springer, Berlin, (1980).

[Takens, 1993] F. Takens, Int. J. Bif. Chaos **3**, 241 (1993).

[Tass et al., 1998] P. Tass, M. G. Rosenblum, M. G. Weule, J. Kurths, A. Pikovsky, J. Volkmann, A. Schnitzler, H. J. Freund, Phys. Rev. Lett. **81**, 3291 (1998).

[Theiler et al., 1992] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J. D. Farmer, Physica D **58**, 77 (1992).

[Theiler & Prichard, 1996] J. Theiler, and D. Prichard, Physica D **94**, 221 (1996).

[Theiler & Prichard, 1997] J. Theiler, and D. Prichard, Fields Inst. Commun. **11**, 99 (1997).

[Theiler, 1995] J. Theiler, Phys. Lett. A (196), 335 (1995).

[Theiler & Rapp, 1996] J. Theiler, and P. Rapp, Electroencephalogr. Glin. Neurophysiol. **98**, 213 (1996).

[Thiel et al., 2002] M. Thiel, M. C. Romano, J. Kurths, R. Meucci, E. Allaria, and T. Arecchi, Physica D **171 (3)** 138 (2002).

[Thiel et al., 2003] M. Thiel, M. C. Romano, and J. Kurths, Applied Nonlinear Dynamics **11 (3)**, 20 (2003).

[Thiel et al. 2004a] M. Thiel, M. C. Romano, P. Read, and J. Kurths, Chaos **14 (2)**, 234 (2004).

[Thiel et al., 2004b] M. Thiel, M. C. Romano, and J. Kurths, Phys. Lett. A (in press).

[Thiel et al., in prep.] M. Thiel, M. C. Romano, J. Kurths, I. Z. Kiss, J. L. Hudson, in preparation for PRL.

[Toledo et al., 1999] E. Toledo, M. Rosenblum, J. Kurths, and S. Akselrod, Computers in Cardiology **26**, 237 (1999).

[Trulla et al., 1996] L. L. Trulla, A. Giuliani, J. P. Zbilut, and C. L. Webber Jr., Physics Letters A , **223(4)**, 255 (1996).

[Webber & Zbilut, 1994] C. L. Webber Jr., and J. P. Zbilut, J. Appl. Physiol. **76** 965 (1994).

[Wolff et al., 1997] J.-O. Wolff, E. Maier-Reimer, and S. Legutke: The Hamburg Ocean Primitive Equation Model, Technical report, No. 13, German Climate Computer Center (DKRZ), Hamburg, 98 pp (1997).

[Yamada & Fujisaka, 1983] T. Yamada, and H. Fujisaka, Prog. Theor. Phys. **70 (5)**, 1240 (1983).

[Yuan Zhao, 1983] Y. Yuan Zhao, Int. J. Bifurc. Chaos **7**, 1401 (1992).

[Zbilut et al., 1998a] J. P. Zbilut, A. Giuliani, and C. L. Webber Jr., Physics Letters A **237** 131–135 (1998).

[Zbilut et al., 1998b] J. P. Zbilut, A. Giuliani, and C. L. Webber Jr., Physics Letters A **246** 122 (1998).

[Zbilut & Webber, 1992] J. P. Zbilut, and C. L. Webber Jr., Physics Letters A **171 (3-4)**, 199 (1992).

[1] We do not analyze a RP as we would have to consider spurious correlations due to the embedding. The CRP considers the same underlying process but different realizations of the noise in the two axis of the plot.

[2] *In practice we fix r as small as possible (typically not greater than* 10% *of the normalized mean distance of the first embedding) relative to the noise level.* [Zbilut & Webber, 1992]

[3] *..., we define the diameter,* $d \in \mathcal{R}^+$*, of a reconstruction to be* $d = max\{||\vec{v}_i - \vec{v}_j|\,|,\ i,j \in \{1,\ldots,N_v\}$ *and* $i \neq j\}$*. Ten percent of the diameter is a typical cutoff value, c.* [Koebbe & Mayer-Kress, 1992]

[4] *For the Radius* ε *one uses usually* 10% *of the maximum diameter* φ_{max} *of the phase space... If there are some outliers in the time series, this ansatz can yield values of* ε*, that are not reasonable any more. The values should be of the order of magnitude of the standard deviation.* [Marwan, 1999]

[5] *Besides being critically important, the selection of threshold corridor is also difficult to systematize in any sensible way. Solutions in the literature are unsatisfying: Webber and Zbilut, without comment, prescribe a threshold corridor, corresponding to the lower* 10% *of the entire distance range present in the corresponding UTRP [Unthreshold Recurrence plot]* [Iwanski, 1998 **]