Klassifikation von Datenreihen mit Hilfe des Lyapunov-Exponenten

Dissertation

zur Erlangung des Grades einer
Doktorin der Naturwissenschaften
der Universität Dortmund

angefertigt am
Fachbereich Statistik
der Universität Dortmund

und im Institut für Neuroinformatik
der Ruhr-Universität Bochum

vorgelegt von

Anja M. Busse

geboren in Hattingen

eingereicht am Fachbereich Statistik
im Februar 2003

1. Gutachter: Prof. Dr. Claus Weihs

2. Gutachterin: Prof. Dr. Katja Ickstadt

Tag der mündlichen Prüfung: 1. April 2003

Danksagung

Auch wenn diese Arbeit nur meinen Namen trägt, so möchte ich doch einigen Menschen danken, die mir auf unterschiedlichste Art und Weise bei ihrer Entstehung geholfen haben.

Anfangen möchte ich mit Herrn Weihs, der mich aufgrund einer von mir unbedachten Bemerkung über den Lyapunov-Exponenten direkt mit einem „Sehr gut, Frau Reimetz, damit beschäftigen sie sich mal weiter“ zu dieser Arbeit getrieben und mich über die Zeit mit neuen Ideen unterstützt hat. Danken möchte ich auch Herrn von Seelen, der als 'sPapilie auf bemerkenswerte Weise die Seele des Instituts für Neuroinformatik ausmachte und mir als Statistikerin unter all den Physikern, Ingenieuren und Informatikern eine Nische freigehalten hat. In dem Zuge möchte ich auch dem Nachfolger der Institutsleitung, Herrn Schöner, danken. Er gab mir die Möglichkeit, die Interdisziplinarität meiner Arbeit weiter fortführen zu können. Des Weiteren bedanke ich mich bei Katja Ickstadt für die Erstellung des Zweitgutachtens und für die Erkenntnisse, die ich zur Heisenbergschen Unschärferelation gewinnen konnte. Für seine fachfremde Korrektur danke ich meinem Onkel, Paul Wiezoreck, der sich sogar die Mühe gab, diese Arbeit zu verstehen.

Mein besonderer Dank gilt meinen Kollegen, die auf erstaunliche Weise meine „Späße“ und Absonderlichkeiten ertragen haben. Da seien zuerst einmal die Neuros mit dem weltbesten Unix-Systemadministrator Micha Neef, insbesondere aber der SONN-Gruppe: Unser Heilpraktiker und Hypnotiseur Rudi Alberts, meinem Lieblingskapitalisten Axel Dietrich, Marc Toussaint und Stefan Wiegand. Michael Hüsken, mit dem ich das Interesse am VfL, aber nicht an Schlagern teile. Mein ganz persönlicher Informatikgott Martin Kreutz, der mir auf seine kaum auffallende Art Wesentliches über Informatik beigebracht hat. Danken möchte ich dem sich selbst zur Perfektion treibenden und weltgroßen (Statistik-)Informatiker Christian Igel, der bevor er - egal welche - statistische Frage stellt, sowieso schon die Antwort kennt. Meinem Lieblingsbernhardo, Dr. Sendhoff, der mal schnell alle entstandenen Physikfragen auf menschliche Weise zu erklären wusste. Meinem literarischen Freund 'sArminPeterle, Peter Stagge, der mit mir den ein oder anderen Humpen geleert und mir die Sinnhaftigkeit zweier Trajektorien erklärte.

Weiter zu danken habe ich meinen Stats: Meiner Kaffeezeitkollegin Martina Erdbrügge, meinem persönlichen Bücherausleiher Olaf Schoffer, Anne Christmann, meinen ehemaligen und jetzigen Lehrstuhlkollegen Karsten Lübke, Brigitta Voss, Frau Kürbis, Daniel Enache, Conny Pumplün und Gero Szepannek. Danken möchte ich weiter Winni Theis, der schnell ein mathematisches Problem zu lösen wusste, über das ich bereits Stunden brütete und Uwe Ligges, der egal welches Computerproblem auftrat, mal eben alles zum Besten richtete, zudem immer bereit war, nachmittags einen guten Mr. Tom mit mir zu teilen. Die angenehmste Art, Stücke dieser Arbeit zu diskutieren, war sicherlich mit meinem Lieblingkollegen von nebenan, Detlef Steuer, und 'sUschilie, Uschi Garczarek, gegeben. Beide füllten die Räume nicht nur mit Wissen, sondern vor allem mit Wärme. Der Uhli, Ulrike Held, möchte ich danken für die vielen Saftpausen und Plauschrunden, die dazu führten, dass wir nicht nur eine gute Ablenkung von der Arbeit hatten, sondern zudem gute Freunde werden konnten.

Meinen Freunden außerhalb des Unilebens sei dafür gedankt, dass sie mich trotz ständiger Abwesenheit nicht vergessen haben und bereit sind, mich wieder im sozialen Leben zu reintegrieren.

Zum Schluss möchte ich aber vor allem meiner Familie für so viele Dinge danken, jedoch reicht dazu der Platz nicht aus. An dieser Stelle sei nur kurz gesagt, dass ich froh bin, so tolle Eltern zu haben, die mir nicht nur durch die Kinderbetreuung Zeit zum Schreiben ließen, sondern vor allem dafür, dass ich immer noch Tochter sein kann. Auch meinen Schwiegereltern gilt dieser Dank. Meiner Tochter Johanna möchte ich dafür danken, dass sie mir klar machte, dass eigene Kinder die weltbesten Beispiele für chaotische Datenreihen liefern. Der am Ende kommende Dank gilt meinem Mann, dass er meine schlechten Launen ertragen hat und durch sein „Für-mich-da-sein" ein gutes Klassifikationskriterium zur Unterscheidung zwischen wichtigen und unwichtigen Dingen abgibt – Danke, Jens.

Inhaltsverzeichnis

Kapitel 1

Einleitung

In der Zeitreihenanalyse tritt oft das Problem auf, nur eine einzige Datenreihe zur Verfügung zu haben, um einen unbekannten zeitlich-funktionalen Zusammenhang zur Beschreibung des Verlaufs der Beobachtungen der Datenreihe, bzw. allgemeiner formuliert, eine geeignete Modellierung zu finden. Notwendig hierzu ist, zumindest eine Vorstellung von einer geeigneten Modellklasse zu haben.

Eine grobe Unterscheidung von Modellklassen in solche für gut oder schlecht prädizierbare Systeme und eine Zuordnung der Datenreihe zu diesen stellt einen wichtigen Schritt auf dem Weg zu einer angemessenen Modellierung dar. So können für so genannte *stabile* Prozesse häufig gute Prädiktoren angegeben werden, wohingegen bei *instabilen* Prozessen bereits eine Ein-Schritt-Prognose unzulänglich sein kann. Im Falle von chaotischen Datenreihen, einem Spezialfall von instabilen Prozessen, kann z. B. die Prädiktionsgenauigkeit schon nach wenigen Zeitschritten abnehmen.

Soll die Prädizierbarkeit für eine Entscheidung über geeignete Modellklassen genutzt werden, bedarf es eines geeigneten Kriteriums zur Schätzung der Prädizierbarkeit auf Grund einer Datenreihe. Dabei muss über die Prädizierbarkeit entschieden werden können, ohne auf irgendeine Kenntnis über den - eigentlich gesuchten - zeitlich-funktionalen Zusammenhang zurückzugreifen, da ansonsten eine Einteilung hinfällig wäre. Eben die Entscheidung über die Prädizierbarkeit soll ja gerade dazu dienen, eine geeignete Modellklasse für die Beschreibung eines solchen Zusammenhangs zu finden.

Der Lyapunov-Exponent leistet eine solche Klassifikation in gut und schlecht prädizierbare bzw. in stabile und instabile Prozesse. Ursprünglich wurde der Lyapunov-Exponent für deterministische Systeme hergeleitet. Gleichwohl wird in dieser Arbeit eine Übertragung auf stochastische Zeitreihen geleistet.

Um empirische Aussagen über die Güteeigenschaften des Lyapunov-Exponenten als Klassifikationskriterium für gut oder schlecht prädizierbare Systeme zu machen, bedarf es einer Stichprobe von Datenreihen.

In dieser Arbeit werden Datenreihen mit festgelegter Prädizierbarkeit ausgehend von einer gegebenen Datenreihe erzeugt, die so genannten Surrogate, ohne dass dabei auf einen bekannten zeitlich-funktionalen Zusammenhang zurück gegriffen wird. Die Stichprobe sollte dabei die Eigenschaft besitzen, ein bestmögliches Abbild der Ausgangsreihe darzustellen. Diese Methode der Simulation wird gewählt, weil sie es grundsätzlich ermöglicht, maßgeschneiderte Berechnungen durchzuführen genau in den Anwendungsfällen des Lyapunov-Exponenten, wenn ganz wenig Vorkenntnis über diesen funktionellen Zusammenhang vorliegt, wohl aber eine Datenreihe.

Ausgehend von solchen Simulationen wird in dieser Arbeit ein Test vorgelegt, mit der die Güte genau einer Entscheidung des Lyapunov-Exponenten als Klassifikationskriterium für eine bestimmte Datenreihe überprüft werden kann. So wird insgesamt eine Methodik bereitgestellt, eine Klassifikationsentscheidung in die Klasse der gut bzw. schlecht prädizierbaren Systeme durch ein statistisch abgesichertes Verfahren zu treffen.

Im Folgenden werden in Kapitel 2 die grundlegenden Begriffe und Notationen erklärt und festgelegt. Somit ist eine Grundlage geschaffen, den Lyapunov-Exponenten sowohl deterministisch als auch stochastisch herzuleiten (vgl. Kapitel 3). Da keine Kenntnisse über den zeitlich-funktionalen Zusammenhang vorausgesetzt werden, muss der Lyapunov-Exponent numerisch berechnet werden. In Kapitel 4 wird ein Verfahren vorgestellt, das dieses ermöglicht. Hierzu werden Simulationen mit stochastischen Prozessen und stabilen und instabilen Formen einer Modulo-1-Abbildung durchgeführt, um das Verfahren experimentell zu unterstützen (vgl. ebenfalls Kapitel 4). Die Grundlage für einen Test zur Überprüfung der Klassifikationsentscheidung, d.h. die Erzeugung einer Stichprobe ausgehend von einer Datenreihe, wird in Kapitel 5 gelegt. Hier werden wiederum Simulationen mit den bereits bekannten Prozessen durchgeführt. Kapitel 6 beschreibt dann die Herleitung der Testprozedur, die über mögliche Fehlklassifikationen entscheidet. Experimentelle Unterstützung wird wieder anhand der Prozesse aus den vorangegangenen Kapiteln geleistet. Um einen Eindruck zu gewinnen, wie gut eine Klassifikation mit Hilfe des Lyapunov-Exponenten auf einem realen Datensatz funktioniert, werden Beobachtungen aus dem so genannten *BTA*-Tiefbohren verwendet. Hierbei ist das Ziel, das während eines Bohrvorgangs möglicherweise auftretende, unerwünschte Rattern rechtzeitig zu identifizieren (vgl. Kapitel 7). Die Arbeit wird durch eine Zusammenfassung und einen Ausblick in Kapitel 8 abgerundet.

Kapitel 2

Deterministische und Stochastische Prozesse

In fast allen Feldern der empirischen Forschung, wie z. B. Ökonometrie, Technometrie oder Biometrie, werden Phänomene beobachtet, die zeitlich voneinander abhängige Komponenten aufweisen. Modelle für solche Phänomene werden gerne als „Prozesse“, die beobachteten Einheiten als „dynamische Systeme“ bezeichnet. In den verschiedenen Disziplinen der Wissenschaft werden deterministische von stochastischen Prozessen unterschieden. Eine intuitive Trennung zwischen deterministischen und stochastischen dynamischen Prozessen umfasst den Aspekt, einen dem Phänomen zu Grunde liegenden funktionalen Zusammenhang mit oder ohne zufälligem Fehler zu modellieren. Für beide Modellansätze sind Eigenschaften, wie die Stabilität oder die Ergodizität, beschreibbar.

2.1 Deterministische Prozesse

In den technischen Disziplinen werden dynamische Systeme vorrangig deterministisch modelliert, d. h. ohne die Annahme eines Zufallseinflusses. Es wird versucht, über das zeitliche Auftreten von funktional zusammenhängenden Systemzuständen zu einer Beschreibung der Dynamik zu gelangen. Ziel ist es, das zu

Grunde liegende System so gut wie möglich zu beschreiben. Die Basis der Systembeschreibung bilden die so genannten Beobachtungsreihen oder auch Flüsse.

Definition 2.1: Beobachtungsreihe

Die Folge von Zuständen oder Beobachtungen $(x_t)_{t \in T}$, deren Dynamik durch

$$x_{t+1} = f_t(x_0) = f(x_t) \tag{2.1}$$

beschrieben wird, heißt Fluss *oder auch* Beobachtungsreihe. *Der Anfangspunkt wird mit $x_0 \in I\!R^k$ bezeichnet, x_t beschreibt den Systemzustand zum Zeitpunkt t. Der funktionale Zusammenhang wird mit f bezeichnet, wobei f überall mindestens einmal differenzierbar sei.*

Soweit in dieser Arbeit nicht explizit darauf aufmerksam gemacht wird, wird $k = 1$ gesetzt.

Mit Hilfe der deterministischen Beobachtungsreihe lässt sich ein deterministischer Prozess definieren.

Definition 2.2: Deterministischer Prozess

Ein deterministischer Prozess $(X_t)_{t \in T}$ wird durch die Menge aller $(x_t)_{t \in T}$ beschrieben, die einer funktionalen Beziehung, Gleichung (2.1), mit einem bestimmten, aus der Menge aller differenzierbaren Funktionen stammenden f genügen.

Der deterministische Prozess ist die übergeordnete Form einer Beobachtungsreihe. Eine bestimmte Beobachtungsreihe $(x_t)_{t \in T}$ eines deterministischen Prozesses wird als Trajektorie oder Orbit bezeichnet.

Ist der funktionale Zusammenhang f des deterministischen Prozesses bekannt, wird er i. d. R. durch Differentialgleichungen beschrieben. Da in dieser Arbeit jedoch nur die Fälle betrachtet werden, in denen keine Informationen über f vorliegen, verweise ich für diesen Fall auf die bestehende Literatur, wie z. B. Arrowsmith und Place (1990), Wiggings (1990).

Einen Spezialfall deterministischer Prozesse bilden die chaotischen Prozesse. In Kapitel 3.3.3 wird er noch genauer beschrieben. An dieser Stelle sei nur darauf hingewiesen, dass sich chaotische Prozesse dadurch auszeichnen, dass kleine Änderungen im Anfangszustand x_0 zu großen Änderungen über die Zeit, also in x_t, führen können. Dadurch sind chaotische Beobachtungsreihen schlecht prädizierbar. Der grundlegende funktionale Zusammenhang ist nichtlinear (siehe z. B. Abarbanel (1996), Tong (1993) und Eckmann und Ruelle (1982)).

Eine oft untersuchte Eigenschaft von dynamischen Systemen ist ihre Stabilität. Anschaulich gesehen bedeutet Stabilität, dass unabhängig von dem Startpunkt immer wieder das gleiche asymptotische Verhalten eintritt. Eine formale Definition des in dieser Arbeit verwendeten Begriffs von Stabilität basiert auf stabilen Fixpunkten bzw. stabilen Orbits (Jetschke (1989), S. 59-60).

Definition 2.3: Stabiler Fixpunkt, stabiler Orbit

a) Ein Fixpunkt $x^0 = f(x^0)$ heißt asymptotisch stabil, wenn gilt:

$$\begin{aligned} &\exists\, \delta > 0, \text{ sodass } \forall x_0 \text{ mit } \| x_0 - x^0 \| < \delta \text{ gilt :} \\ &\lim_{N\to\infty} \| x_N - x^0 \| = 0. \end{aligned} \tag{2.2}$$

b) Eine Trajektorie $(x_i)_{i=0}^{N-1}$ eines Prozesses $(X_t)_{t\in T}$ heißt periodischer Orbit, wenn $f^N(x_0) = x_0$ und $f^i(x_0) \neq x_0$ für $i = 1, \ldots, (N-1)$, wobei $f^i = \underbrace{f \circ \cdots \circ f}_{i\ mal}$, mit $i = 1, \ldots, N-1$.

Ein periodischer Orbit C heißt asymptotisch stabil, wenn ein Punkt $x \in C$ ein asymptotisch stabiler Fixpunkt von f^N ist.

Ein Fixpunkt ist somit asymptotisch stabil, wenn der Abstand zwischen einem späteren Zustand x_N und dem Fixpunkt beliebig klein ist, wenn der Abstand zwischen Anfangszustand und Fixpunkt kleiner als δ war und die unendliche Folge von Zuständen in den Fixpunkt läuft.

Für die weitere Arbeit ist es wichtig, dass ab einem bestimmten Zeitpunkt alle Beobachtungen der Trajektorie in einer ϵ-Umgebung vom Fixpunkt liegen für beliebiges $\delta > \epsilon > 0$. Dieses ist eine direkte Folge aus der Folgerung, dass die Folge $(x_N)_{N \in T}$ gegen den Fixpunkt konvergiert (vgl. Blatter (1980), Kapitel 6.62).

Mit dieser Defintion ist es nun möglich, formale Definitionen für die Stabilität eines Prozesses und einer Beobachtungsreihe $(x_t)_{t \in T}$ zu geben.

Definition 2.4: Stabilität eines deterministischen Prozesses
Ein deterministischer Prozess $(X_t)_{t \in T}$ heißt stabil, wenn alle seine Fixpunkte und periodischen Orbits asymptotisch stabil sind.
Ein deterministischer Prozess heißt instabil, wenn es einen Fixpunkt oder periodischen Orbit gibt, der nicht asymptotisch stabil ist oder wenn überhaupt kein Fixpunkt existiert.

Definition 2.5: Stabilität einer Beobachtungsreihe
Eine Beobachtungsreihe $(x_t)_{t \in T}$ zu einem stabilen deterministischen Prozess $(X_t)_{t \in T}$ heißt asymptotisch stabil, wenn sie gegen einen asymptotisch stabilen Fixpunkt oder asymptotisch stabilen periodischen Orbit konvergiert.

Zur Verdeutlichung sind in Abbildung 2.1 die Möglichkeiten bzgl. der Stabilität dargestellt (vgl. Jetschke (1989), S. 42).

In Abbildung 2.1 Teil a) ist der stabile Zustand dargestellt. Hier reicht es, dass die einzelnen x_t in einem Schlauch um den Fixpunkt bleiben. Teil b) beschreibt das Verhalten bei asymptotischer Stabilität, Gleichung (2.2). Die Beobachtungen laufen asymptotisch auf den Fixpunkt zu. Die Instabilität zeigt Teil c) der Abbildung 2.1.

Eine weitere für diese Arbeit wichtige Eigenschaft deterministischer Prozesse ist die Ergodizität. Hierbei wird geprüft, wann Querschnittanalysen durch Längsschnittanalysen ersetzt werden können. Bei Querschnittanalysen („quer zur Trajektorie") werden verschiedene Beobachtungsreihen eines Prozesses zu einem Zeitpunkt $(x_{t_0})_n, n \in 1, \ldots, N \subset \mathbb{N}$ untersucht, im Gegensatz zu Längsschnittanalysen („längs der Trajektorie"). Hier wird eine Beobachtungsreihe zu jeweils verschiedenen Zeitpunkten $(x_t)_{t \in T^*}$ betrachtet (Schlittgen und Streitberg, 1994).

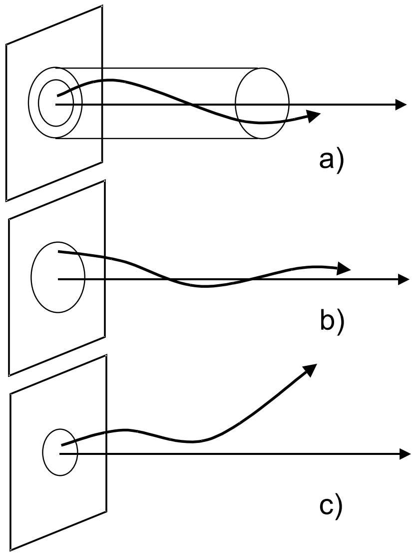

Abbildung 2.1: Stabilität eines Fixpunktes: Abstrakte Definition

Ein ergodischer Prozess soll genau dann gegeben sein, wenn die Mittel aus der Querschnittanalyse für $N \to \infty$ und der Längsschnittanalyse für $|T^*| \to \infty$ gegeneinander konvergieren.

Definition 2.6: Langzeitmittel

Ein Mittel aus der Längsschnittanalyse ist gegeben durch:

$$\bar{x}_{n_0} = \frac{1}{T^*} \sum_{t=0}^{T^*-1} (x_t)_{n_0}. \tag{2.3}$$

Definition 2.7: Ensemblemittel

Ein Mittel aus der Querschnittanalyse ist gegeben durch:

$$\bar{x}_{t_0} = \frac{1}{N} \sum_{n=1}^{N} (x_{t_0})_n, \tag{2.4}$$

wobei N die Anzahl der verschiedenen beobachteten Pfade eines deterministischen Prozesses angibt.

Definition 2.8: Ergodischer determinstischer Prozess

Ein deterministischer Prozess heißt ergodisch, wenn das zeitliche Mittel aus der Längsschnittanalyse und das Ensemblemittel aus der Querschnittanalyse gegen einen gemeinsamen Wert, ihren „Erwartungswert“ konvergieren.

Bemerkung 2.1:

Im Falle der Stabilität eines Prozesses existiert genau ein Wert, gegen den beide Mittel konvergieren. Das bedeutet weiter, dass beide Mittel den Erwartungswert schätzen.

Eine genaue Beschreibung ergodischer Systeme und unter welchen Bedingungen verschiedene deterministische Prozesse ergodisch sind, findet sich u. a. in Arnold und Avez (1968) und Friedman (1970).

2.2 Stochastische Prozesse

In realen Systemen ist es jedoch oft von Vorteil, einen stochastischen Einfluss in die Definition der modellierenden Prozesse einzubauen. So führen Störungen, die durch Messfehler oder ungenaue Sensoren verursacht werden, zu verrauschten Daten, so dass eine Beschreibung durch einen reinen Determinismus unzulänglich wird. Im Gegensatz zu der Modellierung mit einem deterministischen Prozess wird folglich dem funktionalen Zusammenhang des zu Grunde liegenden Systems in der stochastischen Modellierung ein zufälliger Fehler hinzugefügt. Hierbei wird die Definition eines stochastischen Prozesses in Anlehnung an die Definition deterministischer Prozesse wieder mit Hilfe der Folge von Systemzuständen geleistet. Das Zulassen zufälliger Einflüsse in dem funktionalen Zusammenhang der Zustände macht jedoch das Einführen von Zufallsvariablen notwendig. Die stochastischen Zeitreihen werden verbal von den deterministischen Beobachtungsreihen unterschieden.

Definition 2.9: Zeitreihe, stochastischer Prozess

Eine Folge von Beobachtungen $(x_t)_{t \in T}$, die als Realisationen von Zufallsvariablen $(X_t)_{t \in T}$ aufgefasst werden und deren Dynamik durch

$$x_{t+1} = f_t(x_0, \epsilon) = f(x_t, \epsilon) \tag{2.5}$$

beschrieben wird, heißt Zeitreihe. *Dabei ist ϵ eine Zufallsvariable mit Verteilung P_ϵ.*

Die Folge der Zufallsvariablen $(X_t)_{t \in T}$ heißt stochastischer Prozess. Dabei ist die Indexmenge T höchstens abzählbar.

Bemerkung 2.2:

Anders als sonst üblich in der Statistik wird auch eine unendliche Länge, d. h. gar nicht wirklich beobachtbare Zeitreihen als „Realisationen“ aufgefasst.

Bemerkung 2.3:

Ein stochastischer Prozess definiert sowohl Zufallsvariablen für Längsschnitt-, als auch für Querschnittanalysen (Rinne (1997), S. 437).

Ein stochastischer Prozess wird typischerweise durch eine Folge von Zufallsvariablen beschrieben. X_0 bezeichnet die Zufallsvariable zum Startzeitpunkt, in dem der Anfangspunkt x_0 realisiert wird. Die Zufallsvariable X_t mit der Realisation x_t beschreibt den Zustand zur Zeit $t, t \in T$. Der funktionale Zusammenhang f ist stochastisch gestört mit einem nicht notwendigerweise additiven Rauschen. Die Annahme eines additiven Rauschens, d. h. $f_t(x_0, \epsilon) = f(x_t) + \epsilon_t, t \in T$, ist in vielen Fällen sehr einschränkend. Zwar können Messfehler damit gut modelliert werden, in vielen Fällen unterliegen jedoch die Systeme selbst einer zufälligen Störung, was wiederum zu einem nichtadditiven, möglicherweise dynamischen Rauschen führt (Kreutz (2000), S. 9).

Eine bestimmte Zeitreihe $(x_t)_{t \in T}$ eines stochastischen Prozesses $(X_t)_{t \in T}$ wird genauso wie das deterministische Pendant als Trajektorie oder Orbit bezeichnet.

Können in der weiter führenden Arbeit sowohl die deterministischen Beobachtungsreihen als auch die stochastischen Zeitreihen gemeint sein, wird auch oft zusammenfassend der Begriff der *Datenreihe* verwendet.

Wenn der zu Grunde liegende funktionale Zusammenhang, Gleichung (2.5), bekannt ist, können stochastische Systeme durch stochastische Differentialgleichungen beschrieben werden. Diese Situation wird in dieser Arbeit jedoch nicht behandelt. Eine gute Einführung für eine Modellierung mit stochastischen Differentialgleichungen leisten u. a. Ash und Doleans-Dade (2000) und Arnold (1998). Insbesondere Arnold hat sich für bekanntes f mit Fragen des Lyapunov-Exponenten auseinandergesetzt (Arnold, 1998), (Arnold, Gundlach und Demetrius, 1994), (Arnold und Scheutzow, 1995).

Die wichtige Eigenschaft der Stabilität kann von den deterministischen auf stochastische Prozesse übertragen werden. Auch hier gilt, dass das asymptotische Verhalten idealerweise unabhängig von kleinen Schwankungen im Anfangszustand sein soll. Gemäß der stochastischen Modellierung wird jedoch keine 100%ige asymptotische Stabilität eines Fixpunktes oder periodischen Orbits gefordert, sondern nur eine asymptotische stochastische Stabilität. Ausgehend von Gleichung (2.2) wird die asymptotische stochastische Stabilität mit Hilfe „stochastisch stabiler“ Fixpunkte bzw. Orbits neu eingeführt.

Definition 2.10: Stochastisch stabiler Fixpunkt, stochastisch stabiler Orbit

a) *Ein Fixpunkt $x^0 = f(x^0)$ heißt asymptotisch stochastisch stabil, wenn gilt:*

$$\begin{aligned} &\exists\, \delta > 0, \text{ sodass } \forall x_0 \text{ mit } \| x_0 - x^0 \| < \delta \text{ gilt :} \\ &P\left(\lim_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N-1} X_t - x^0 = 0 \right) = 1. \end{aligned} \tag{2.6}$$

b) *Eine Trajektorie $(x_i)_{i=0}^{N-1}$ eines stochastischen Prozesses $(X_t)_{t\in T}$ heißt periodischer Orbit, wenn $f^N(x_0) = x_0$ und $f^i(x_0) \neq x_0$* für $i = 1, \ldots, (N-1)$, *wobei* $f^i = \underbrace{f \circ \cdots \circ f}_{i\ mal}$, *mit* $i = 1, \ldots, N-1$.

Ein periodischer Orbit C *heißt asymptotisch stochastisch stabil, wenn ein Punkt* $x \in C$ *ein asymptotisch stochastisch stabiler Fixpunkt von* f^N *ist.*

Bemerkung 2.4:

Wird für x^0 *in Gleichung (2.6) der Erwartungswert des stochastischen Prozesses* $E(X_t) = \mu\ \forall t \in T$ *eingesetzt, stimmt die Gleichung (2.6) mit dem „Starken Gesetz der großen Zahlen von Borel und Cantelli", Rinne (1997), S. 466, überein. Dieses Gesetz besagt die fast sichere Konsistenz des Folgemittels gegen den Erwartungswert des stochastischen Prozesses.*

Die Definition von stochastisch stabilen Fixpunkten und Orbits (Gleichung (2.6)) ist ein Spezialfall der deterministisch stabilen Fixpunkte (Gleichung (2.2)). Bei stochastischer Stabilität wird ein „Rauschen" um den Fixpunkt zugelassen, wohingegen der stabile Fixpunkt exakt getroffen werden muss.

Mit Hilfe der Definition stochastisch stabiler Fixpunkte und stochastisch stabiler Orbits können „stabile" stochastische Prozesse und „stabile" Zeitreihen festgelegt werden.

Definition 2.11: Stabilität eines stochastischen Prozesses

Ein stochastischer Prozess $(X_t)_{t \in T}$ *heißt stabil, wenn alle seine Fixpunkte und periodischen Orbits asymptotisch stabil sind.*

Ein stochastischer Prozess heißt instabil, wenn es einen Fixpunkt oder periodischen Orbit gibt, der nicht asymptotisch stochastisch stabil ist oder wenn überhaupt kein Fixpunkt existiert.

Definition 2.12: Stabilität einer Zeitreihe

Eine Zeitreihe $(x_t)_{t \in T}$ *zu einem stochastisch stabilen Prozess* $(X_t)_{t \in T}$ *heißt asymptotisch stabil, wenn sie gegen ihren Fixpunkt oder periodischen Orbit konvergiert.*

Bemerkung 2.5:

Im Gegensatz zur gewöhnlichen Stabilitätsaussage, bei der es sich um eine Vertei-

lungskonvergenz handelt (vgl. Brockwell und Davis (1996), S. 535), wird in dieser Arbeit eine fast-sichere Punktkonvergenz betrachtet.

Die Stabilität stochastischer Prozesse unterscheidet sich von der deterministischer Prozesse in dieser Arbeit nur durch die Bedingung, dass der Fixpunkt nicht exakt, sondern mit Wahrscheinlichkeit eins im Mittel erreicht werden muss.

Die Eigenschaft der Ergodizität spielt für stochastische Prozesse eine ebenso große Rolle wie für deterministische Systeme. Auch wenn der Ursprung der Ergodizität in der Mechanik liegt, so wurde doch dieser Begriff bereits 1931/1932 in die Stochastik übertragen (Schlittgen und Streitberg (1994), S. 230). Hier stellt sich die Frage, wann stochastische Prozesse hinsichtlich ihres Mittelwertes asymptotisch ihre Zeitabhängigkeit verlieren. Wie bei deterministischen Systemen soll das zeitliche Mittel gegen das Ensemblemittel konvergieren. Zusätzlich wird für die Definition von ergodischen stochastischen Prozessen die Mittelwertstationarität gefordert.

Definition 2.13: Stationarität

Ein stochastischer Prozess $(X_t)_{t \in T}$ heißt (mittelwert-)stationär, wenn der Erwartungswert invariant bleibt gegenüber Verschiebungen entlang der Zeitachse. Das bedeutet, dass

$$E(X_t) = E(X_{t'}) =: E_P(X) \ \forall t, t' \in T, \tag{2.7}$$

wobei $E_P(X)$ den Erwartungswert des stochastischen Prozesses bezeichnet.

Definition 2.14: Ergodischer stochastischer Prozess

Ein mittelwertstationärer stochastischer Prozess $(X_t)_{t \in T}$ mit Erwartungswert $E_P(X)$ heißt ergodisch (Schlittgen und Streitberg (1994), S. 230), wenn

$$P\left(\lim_{N \to \infty} \frac{1}{N} \sum_{t=0}^{N-1} X_t = E_P(X)\right) = 1. \tag{2.8}$$

Die Definition ergodischer stochastischer Prozesse stimmt mit der Definition der ergodischen deterministischen Systeme überein, wenn für das Ensemblemittel (vgl. Gleichung (2.4)) das Prozessmittel $E_P(X)$ eingesetzt wird.

Ergodische Prozesse besitzen die Eigenschaft, unabhängig vom Startwert immer wieder das gleiche asymptotische Verhalten aufzuweisen (Arnold und Avez (1968), S. 17).

Ein für die Arbeit nützlicher Satz ergodischer Prozesse ist, dass bei einer Transformation mit einer messbaren Funktion g die Ergodizitäteigenschaft eines (dann transformierten) Prozesses weiter bestehen bleibt (vgl. Stout (1974), S. 167ff, S. 182).

Satz 2.1:
Gegeben sei ein ergodischer stochastischer Prozess $(X_t)_{t\in T}$. Für alle messbaren Funktionen g gilt:

$$P\left(\lim_{N\to\infty}\frac{1}{N}\sum_{t=0}^{N-1} g(X_t) = E(g(X))\right) = 1. \tag{2.9}$$

Der Beweis findet sich bei Stout (1974), S. 183.

2.3 Einbettung deterministischer Prozesse in die Stochastik

In realen Systemen ist es oft nicht möglich zu unterscheiden, ob es sich eher um einen deterministischen oder um einen stochastischen Prozess handelt. So ist es von Vorteil, deterministische Beobachtungsreihen als stochastische Zeitreihen ausdrücken zu können. Es ist natürlich klar, dass dieses konzeptuelle Überführen auch eine philosophische Frage ist. So lehnen es die einen ab, beide Arten von Prozessen in einen „stochastischen Topf“ zu werfen. Die anderen vertreten den Standpunkt, dass in der Realität niemals reine deterministische Prozesse auftreten können, weil kein absolut exaktes - so genanntes optimales Messen - möglich sei. Für diese Arbeit ist es nicht wichtig, diese Frage zu beantworten. Es werden beide Standpunkte zugelassen.

Die konzeptuelle Einbettung von Beobachtungsreihen in den stochastischen Rahmen wird bewerkstelligt durch die Darstellung des funktionalen Zusammenhangs mit einem Fehler ϵ, der einer Einpunktverteilung folgt.

Satz 2.2:
Die Beobachtungsreihe $(x_t)_{t \in T}$ eines deterministischen Prozesses mit Funktion g kann durch

$$x_{t+1}^{determ} = g^t(x_0) = g(x_t) = f(x_t, e) = f^t(x_0, e) = f^t(x_0, \epsilon) = x_{t+1}^{stoch}, \qquad (2.10)$$

als Zeitreihe eines stochastischen Prozesses mit einer Funktion f dargestellt werden, wobei ϵ einer Einpunktverteilung, Ep(e), mit Parameter e folgt.

Beweis:
Ein stochastischen Prozess $(X_t)_{t \in T}$ mit

$$\begin{aligned} x_{t+1}^{stoch} &= f^t(x_0, \epsilon) && \text{funktionaler Zusammenhang (siehe Gleichung (2.5))} \\ &= f(x_t, e) && \text{mit } \epsilon \text{ einpunktverteilt (Ep(e)), Ep(e) beschreibt} \\ & && \text{„eigentlich" ein deterministisches Geschehen} \end{aligned}$$

Denn: Definiere $g(x) = f(x, e)\ \forall x \in X$. Beachte, dass kein stochastischer Anteil vorhanden ist. Definiere weiter, $(X_t)_{t \in T}$ ist eine Beobachtungsreihe eines deterministischen Prozesses mit Funktionsgleichung $g(x)$.
Also folgt weiter:

$$\begin{aligned} f(x_t, e) &= g(x_t) && \text{wegen der obigen Definition} \\ &= g^t(x_0) && \text{(siehe Gleichung (2.1))} \\ &= x_{t+1}^{determ}. \end{aligned}$$

q.e.d

Um also vom deterministischen Prozess zur stochastischen Darstellung überzugehen, bedarf es nur einer geschickten Wahl von e.

2.4 Spezielle Prozesse

Zwei in der Statistik häufig verwendete Zeitreihenmodelle sind lineare Modelle (wie z. B. die ARMA-Modelle) und Schwingungsmodelle. Da zur experimentellen Unterstützung in dieser Arbeit diese speziellen stochastischen Prozesse verwendet werden, sollen sie im Folgenden kurz mit den hier wichtigen Eigenschaften vorgestellt werden.

AR-Prozesse

Zuerst einmal soll ein Prozess vorgestellt werden, dessen Definitionsgleichung einer formalen multiplen Regression entspricht, jedoch die erklärenden Variablen nicht wie dort unabhängige Variablen sind, sondern die Vergangenheitswerte des Prozesses selbst (Schlittgen und Streitberg, 1994).

Definition 2.15: AR(p)-Prozess
Ein stochastischer Prozess $(X_t)_{t\in T}$ *heißt Autoregressiver Prozess der Ordnung* p, *kurz* $AR(p)$*-Prozess, wenn er der Beziehung*

$$X_t = \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} + \epsilon_t \tag{2.11}$$

genügt. Dabei ist ϵ_t *Weißes Rauschen* [1].

Damit dieser Prozess stationär ist, muss eine Stationaritätsbedingung, dass alle Lösungen der charakteristischen Gleichung außerhalb des Einheitskreises liegen, erfüllt sein (vgl. Schlittgen und Streitberg (1994), S. 122, 124). Die Eigenschaft der Ergodizität von AR-Prozessen, einschließlich der Verallgemeinerung zu nichtlinearen AR-Prozessen, werden in Schlittgen und Streitberg (1994), S. 232, 437 dargestellt. Eine für diese Arbeit wichtige Frage nach der Stabilität von AR-Prozessen beantworten Schlittgen und Streitberg (1994), S. 437. Für den speziellen AR(1)-Prozess ist bekannt, dass er explosiv ist, wenn der Parameter $|\phi_1|$

[1] Weißes Rauschen zeichnet sich dadurch aus, dass die Zufallsvariablen ϵ_t alle den gleichen Erwartungswert und die gleiche Varianz besitzen (Schlittgen und Streitberg (1994) S. 96).

größer als eins ist. Der Prozess läuft dann gegen $\pm\infty$. Somit ist der Prozess instabil. Ist hingegen $|\phi_1|$ echt kleiner als eins, so ist der Startwert auf lange Sicht irrelevant. Der Prozess ist normalverteilt. Die Stabilität ist in diesem Falle gegeben.

MA-Prozesse

Ein linearer Prozess, der im Grunde nichts anderes ist als ein gewogenes Mittel aus unkorrelierten Zufallszahlen, wobei die Summe der Gewichte im Allgemeinen nicht gleich eins ist, ist durch den MA(q)-Prozess gegeben (Stier (2001), S. 52).

Definition 2.16: MA(q)-Prozess

Ein stochastischer Prozess $(X_t)_{t\in T}$ heißt Moving-Average Prozess der Ordnung q, kurz $MA(q)$-Prozess, wenn er der Beziehung

$$X_t = \theta_1\epsilon_{t-1} + \ldots + \theta_q\epsilon_{t-q} + \epsilon_t \tag{2.12}$$

genügt. Dabei ist ϵ_t Weißes Rauschen.

Offensichtlich sind MA(q)-Prozesse stationär (Stier (2001), S. 52, Schlittgen und Streitberg (1994)), S. 116. Die Ergodizität zeigt sich in Schlittgen und Streitberg (1994), S. 235.

ARMA-Prozesse

Eine Kombination der beiden in diesem Kapitel eingeführten Prozesse wird durch einen ARMA(p,q)-Prozess beschrieben.

Definition 2.17: ARMA(p,q)-Prozess

Ein stochastischer Prozess $(X_t)_{t\in T}$ heißt Autoregressiver-Moving-Average Prozess der Ordnung p, q, kurz $ARMA(p, q)$-Prozess, wenn er der Beziehung

$$X_t = \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} + \theta_1\epsilon_{t-1} + \ldots + \theta_q\epsilon_{t-q} + \epsilon_t \tag{2.13}$$

genügt. Dabei ist ϵ_t Weißes Rauschen.

ARMA-Prozesse sind genau dann stationär, wenn alle Wurzeln der charakteristischen Gleichung des AR-Teils, $1 - \phi_1 z - \phi_2 z^2 - \ldots - \phi_p z^p = 0$, außerhalb des Einheitskreises liegen (Rinne (1997), S. 454, Stier (2001), S. 55). Das ist gleichbedeutend, dass der AR-Teil als Potenzschreibweise mit absolut summierbarer Koeffizientenfolge beschreibbar ist (Schlittgen und Streitberg (1994), S. 134). Unter geeigneten Bedingungen kann ein ARMA-Prozess durch einen AR- bzw. MA-Prozess mit jeweils genügend großer Ordnung approximiert werden (Schlittgen und Streitberg (1994), S. 135). Weiter wird gezeigt, dass stationäre ARMA(p,q)-Prozesse ergodisch sind (Schlittgen und Streitberg (1994), S. 231). In diesem Falle sind unter weiteren Bedingungen (vgl. Brockwell und Davis (1996), S. 535, 536) ARMA-Prozesse auch stabil.

Random Walk

Ein in der Ökonometrie häufig verwendeter Prozesstyp ist der Random Walk. Er wird z. B. zur Modellierung von Aktienkursen verwendet.

Definition 2.18: Random Walk
Ein Random Walk Prozess $(X_t)_{t\in T}$ ist definiert durch die Rekursion:

$$X_t = \begin{cases} \epsilon_t & , \quad t = 1 \\ X_{t-1} + \epsilon_t & , \quad t = 2, 3, \ldots \end{cases} , \tag{2.14}$$

wobei ϵ_t Weißes Rauschen ist.

Brockwell und Davis (1996), S. 14, zeigen, dass ein Random Walk nicht varianzstationär ist. Die Mittelwertfunktion ist nur für $E[X_t] = 0$ konstant. Ist der Erwartungswert größer als null, ist sie linear steigend, andernfalls linear fallend. Die Varianzfunktion ist in jedem Fall eine wachsende Funktion der Zeit (Schlittgen und Streitberg (1994), S. 96). Das bedeutet anschaulich, dass sich die einzelnen Realisationen des Prozesses im Mittel immer weiter voneinander entfernen.

Schwingungsmodelle

Schwingungsmodelle eignen sich besonders gut zur Beschreibung von zyklischen Einflüssen. Vor allem die Überlagerung harmonischer Schwingungen (Sinus- und Kosinuswellen) spielen bei der Analyse zyklischer Schwankungen in Zeitreihen eine besondere Rolle. Dieses ist darauf zurück zu führen, dass sich auch sehr komplizierte und keineswegs „sinusförmige“ periodische Funktionen durch Überlagerung harmonischer Schwingungen verschiedener Frequenzen darstellen lassen (Schlittgen und Streitberg (1994), S. 51, 52). Ein Hauptanwendungsgebiet solcher Prozesse findet sich zudem bei der Fourieranalyse. Die Grundperiode oder auch Frequenz des Schwingungsmodells liegt bei 2π (Schlittgen und Streitberg (1994), S. 50-52). In dieser Arbeit wird nur ein Spezialfall der Schwingungsmodelle betrachtet:

Definition 2.19: sin-Prozess

Ein stochastischer Prozess $(X_t)_{t \in T}$ heißt sin-Prozess, wenn er der Beziehung

$$X_t = \alpha \sin x_{t-1} + \epsilon_t \tag{2.15}$$

genügt. Dabei ist ϵ_t Weißes Rauschen.

Für den so definierten sin-Prozess, ist für alle Parametereinstellungen Stationarität gegeben. Die allgemeinen Bedingungen für die Ergodizität dieses Prozesses finden sich in Schlittgen und Streitberg (1994), S. 231. Weil die Werte der Sinusfunktion absolut beschränkt sind auf dem Intervall $[-1, 1]$, sind insgesamt die Voraussetzungen für die Stabilität sehr günstig. Die Bedingungen an den Parameter α müssen jedoch dabei garantieren, dass die Schwankungen im Bereich $[-e-\alpha, \alpha+e]$ bleiben und nicht über die Zeit „explodieren“.

Kapitel 3

Lyapunov-Exponent

Zur Modellierung realer Systeme steht dem Anwender im Allgemeinen nur eine Datenreihe zur Verfügung. Selten sind Informationen über den zu Grunde liegenden Prozess und somit über die Modellklasse zur Beschreibung des Systems vorhanden. Eine Vorauswahl der vielen möglichen Modellklassen, die alleine durch die Zeit- bzw. Beobachtungsreihe getroffen werden kann, ist ein gutes Hilfsmittel für eine weitere detaillierte Analyse des Prozesses. Eine Möglichkeit zur Einteilung von Datenreihen bietet der größte Lyapunov-Exponent. Dieser unterscheidet zwischen stabilen, d. h. gut prädizierbaren, und instabilen Prozessen, also schlecht prädizierbaren. Da in dieser Arbeit nicht das Lyapunov-Spektrum (Abarbanel, Brown, Sidorowich und Tsimring, 1993) sondern nur der größte Exponent betrachtet wird, wird im Folgenden nur noch kurz vom Lyapunov-Exponenten oder auch nur Exponenten gesprochen.

Im Folgenden wird dieser zuerst für deterministische Beobachtungsreihen vorgestellt. Gleichwohl wird in Kapitel 3.2 gezeigt, dass eine direkte Übertragung der Definition des Lyapunov-Exponenten unter Annahme von additivem Rauschen oder unter der Annahme der Ergodizität des Prozesses in einem statistischen Kontext gewährleistet werden kann. Eine Übertragung des Lyapunov-Exponenten für Zeitreihen, wobei der funktionale Zusammenhang nicht eingeschränkt ist, wird in Kapitel 3.2.2 geleistet. Die Unterscheidung zwischen stabilen und instabilen

Datenreihen und die daraus resultierenden Modellklassen werden in Kapitel 3.3 beschrieben. Welche Bedeutung diese Einteilung auf die Prädizierbarkeit und auf den Informationsverlust über die Zeit hat, zeigt Kapitel 3.3.1.

3.1 Deterministische Herleitung

Seien $x_0 \in \mathbb{R}^k$ der Anfangspunkt oder Anfangszustand und x_t der Systemzustand zum Zeitpunkt t. Weiter sei ein deterministisches dynamisches System gegeben, dessen Dynamik definiert ist als (vgl. Gleichung 2.1)

$$x_{t+1} = f_t(x_0) = f(x_t) \ .$$

Der Lyapunov Exponent $\lambda(x_0)$ eines Flusses $(x_t)_{t \in T}$ wurde formal durch Eckmann und Ruelle (1982) definiert.

Definition 3.1: Lyapunov-Exponent
Der Lyapunov-Exponent eines Prozesses mit Startwert x_0 ist gegeben durch:

$$\lambda(x_0) := \lim_{N \to \infty} \frac{1}{N} \sum_{i=0}^{N-1} \ln |f'(x_i)| \ . \tag{3.1}$$

Das Langzeitverhalten der durchschnittlichen logarithmischen Ableitung des potenziellen funktionalen Zusammenhangs eines Flusses wird durch diese charakteristische Größe beschrieben. Es untersucht das Konvergenz- bzw. Divergenzverhalten zweier benachbarter Trajektorien. Somit erfolgt die Herleitung des Lyapunov-Exponenten über diese benachbarten Trajektorien, die zu einem zu Grunde liegenden Prozess gehören. Abbildung 3.1 zeigt, wenn N unendlich wird, schematisch das Auseinanderlaufen zweier benachbarter Pfade eines Flusses, wobei die Anfangszustände durch einen kleinen Abstand voneinander entfernt liegen.

Bezeichnen x_0 and x_0^* die zwei unterschiedlichen Anfangspunkte eines Flusses. Beide Pfade folgen dem gleichen funktionalen Zusammenhang, da sie aus dem gleichen Prozess mit unterschiedlichen Anfangszuständen kommen. Der Startpunkt x_0^* sei in der Nähe von x_0.

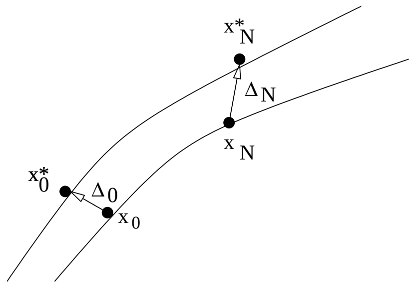

Abbildung 3.1: Über die Zeit werden zwei Trajektorien betrachtet, um das Divergenz- bzw. Konvergenzverhalten eines Prozesses zu beobachten.

Der Abstand zwischen x_0 und x_0^* sei gegeben durch:

$$\Delta_0 = |x_0^* - x_0|. \tag{3.2}$$

Nach einer Iteration beträgt die Distanz zwischen den Pfaden Δ_1. Diese nächste Distanz Δ_1 lässt sich durch Anwendung der ersten Taylorentwicklung in folgender Form approximieren:

$$\begin{aligned}
\Delta_1 &= |x_1^* - x_1| \\
&= |f(x_0^*) - f(x_0)| \\
&\approx |f(x_0) + f'(x_0) \cdot \Delta_0 - f(x_0)| \\
&= |f'(x_0)| \cdot \Delta_0 \\
&= |f'(x_0)| \cdot |x_0^* - x_0| \; .
\end{aligned}$$

Für die weitere Analyse des Langzeitverhaltens einer Beobachtungsreihe ist die Berechnung der Distanz zweier Trajektorien nach N Iterationen notwendig. Mit Hilfe der Gleichungen (2.1) und (3.2) und der Kettenregel folgt für die N-te

Distanz:

$$\begin{aligned}
\Delta_N &= |x_N^* - x_N| \\
&\approx \left|\frac{\mathrm{d}}{\mathrm{d}x} f^N(x_0)\right| \cdot |x_0^* - x_0| \\
&= \left|\frac{\mathrm{d}}{\mathrm{d}x} f^N(x_0)\right| \cdot \Delta_0 \\
&= |f'(x_0)| \cdot |f'(x_1)| \ldots |f'(x_{N-1})| \cdot \Delta_0 \\
&= \prod_{i=0}^{N-1} |f'(x_i)| \cdot \Delta_0 \ ,
\end{aligned}$$

wobei $f^N = \underbrace{f \circ \cdots \circ f}_{N\ Mal}$ ist.

Unter der Annahme eines exponentiellen Zuwachses kann die Expansionsrate der Pfade durch:

$$\frac{\Delta_N}{\Delta_0} \approx \prod_{i=0}^{N-1} |f'(x_i)| \tag{3.3}$$

$$=: e^{N \cdot \lambda_N(x_0)} \tag{3.4}$$

beschrieben werden.

Die charakteristische Größe λ_N hängt sowohl von den Iterationen N als auch von dem Startwert x_0 ab. Die Expansionsrate beschreibt das Verhalten der benachbarten Pfade nach N Iterationen in Abhängigkeit von Δ_0 und x_0.

Durch Umformung von Gleichung (3.4) und bei Betrachtung des asymptotischen Verhaltens für $N \to \infty$ ist es nun möglich, den Lyapunov-Exponenten zu definieren:

$$\lambda(x_0) := \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} \ln |f'(x_i)| \tag{3.5}$$

$$\approx \lim_{N\to\infty} \frac{1}{N} \ln\left(\frac{\Delta_N}{\Delta_0}\right) . \tag{3.6}$$

Der Lyapunov-Exponent misst die asymptotische, durchschnittliche logarithmische Ausdehnung infinitesimal kleiner Intervalle entlang zweier Trajektorien. Jedoch hängt $\lambda(x_0)$ von dem Anfangszustand x_0 ab. Ist λ kleiner als null, laufen die Pfade zusammen. Der zu Grunde liegende Prozess konvergiert. Auch die

Abhängigkeit vom Startwert ist nicht mehr gegeben. Ist hingegen λ größer als null, divergiert der Prozess mit einer Rate von $e^{\lambda N}$.

Die Ableitung f' des funktionalen Zusammenhangs f aus Gleichung (3.5) ist in realen Systemen jedoch oft nicht bekannt. Es muss anhand der gegebenen Beobachtungsreihe berechnet werden. Hierzu werden in der Literatur verschiedene Möglichkeiten bereit gestellt (vgl. Kapitel 4). An dieser Stelle sei schon einmal darauf hingewiesen, dass sich die Beobachtungsreihe noch in einem transienten, d. h. sich noch ändernden Zustand befinden kann. Der Lyapunov-Exponent ist jedoch eine Größe des Attraktors, d. h. des Grenzverhaltens eines Prozesses. Der Attraktor ist zeitunabhängig, so dass für die Berechnung von λ nur Beobachtungen in einem sich nicht mehr ändernden Zustand benutzt werden dürften. Bei zu kurzen Datenreihen kann das zu fehlerhaften Berechnungen führen (vgl. Kapitel 4). Bestenfalls sollte sich die Beobachtungsreihe zur Berechnung im asymptotischen Zustand befinden.

3.2 Stochastische Herleitung

Auch wenn der Lyapunov-Exponent ursprünglich für deterministische Systeme beschrieben wurde, ist es gleichwohl möglich, diesen auf stochastische Systeme zu übertragen. Wird ein additives Rauschen im funktionalen Zusammenhang angenommen, ist eine direkte Übertragung dieser Kenngröße gegeben (vgl. Kapitel 3.2.1). Sollen keine Einschränkungen im funktionalen Zusammenhang angenommen werden, zeigt sich der Lyapunov-Exponent unter gewissen Annahmen als asymptotischer Erwartungswert einer Transformation des gegebenen stochastischen Prozesses (vgl. Kapitel 3.2.2).

3.2.1 Additives Rauschen

Im Gegensatz zu den deterministischen Prozessen besitzt der funktionale Zusammenhang eines stochastischen Prozesses einen zufälligen Einfluss (vgl. Gleichung (3.7)). Sei hier der zufällige Einfluss durch additives Rauschen in der Dynamik gegeben, d. h.

$$X_{t+1} = f_t(X_0) + \epsilon = f(X_t) + \epsilon. \tag{3.7}$$

Für stochastische Prozesse mit additivem Rauschen ist es auf die gleiche Weise wie bei deterministischen Systemen möglich, den Lyapunov-Exponenten zu erklären.

Hierzu muss eine einfache Transformation $g(x) = f(x) + \epsilon$ vorgenommen werden, so dass:

$$X_{t+1} = g(X_t), \quad \text{mit} \quad g(X_t) = f(X_t) + \epsilon \quad \text{und} \quad g'(x) = f'(x). \tag{3.8}$$

Durch eine Verallgemeinerung der Gleichung (3.1) auf bekannte Weise ist der Lyapunov-Exponent für stochastische Prozesse mit additivem Rauschen gegeben durch:

$$\lambda(X_0) := \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} \ln |g'(X_i)| \tag{3.9}$$

$$= \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} \ln |f'(X_i)|. \tag{3.10}$$

Der additive Fehler verschwindet durch den Gebrauch der Ableitung, so dass die Übertragung ohne weitere Annahmen vollzogen werden kann. Die Definition des Lyapunov-Exponenten für stochastische Prozesse mit additivem Rauschen stimmt also formal mit der Definition für deterministische Prozesse überein (vgl. Gleichungen (3.1), (3.9) und (3.10)).

3.2.2 Ergodische Prozesse

Im Gegensatz zu Kapitel 3.2.1 muss nun der stochastische Einfluss nicht notwendigerweise durch additives Rauschen in den funktionalen Zusammenhang eingehen. Ganz allgemein wird hier von einer Dynamik, wie sie in der Definition von Zeitreihen (vgl. Gleichung (2.5)) beschrieben wurde, ausgegangen.

Auf natürliche Weise kann der Lyapunov-Exponent als asymptotischer Erwartungswert, sofern er existiert, einer Transformation des gegebenen stochastischen Prozesses verallgemeinert werden. Die Eigenschaft, dass der Lyapunov-Exponent ein Charakteristikum des Attraktors ist, d. h. des asymptotischen Verhaltens, ist direkt an der Verallgemeinerung ablesbar.

Definition 3.2: Stochastische Interpretation

Für einen transformierten stochastischen Prozess spiegelt der Lyapunov-Exponent den asymptotischen Erwartungswert wider in der Form:

$$\tilde{\lambda}(x_0) := \lim_{t\to\infty} E[\ln|f'(X_t(x_0))|]. \tag{3.11}$$

Jedoch ist dieser Erwartungswert in den meisten Fällen unbekannt und müsste daher geschätzt werden. Eine offensichtliche Möglichkeit ist die Berechnung des Langzeitmittels, welches für diskrete Prozesse in folgender Form definiert ist:

$$\bar{g}(x_0) := \lim_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N-1} g(X_t(x_0)). \tag{3.12}$$

Hierbei ist g eine willkürliche, messbare Funktion. In dem Fall der Schätzung des Lyapunov-Exponenten wird $g(x) = \ln|f'(x)|$ gesetzt. Es sei darauf hingewiesen, dass $\bar{g}(x_0)$ (vgl. Gleichung (3.12)) der Definition des Lyapunov-Exponenten für stochastische Prozesse mit additivem Rauschen entspricht (vgl. Gleichung (3.10)).

Das Langzeitmittel kann vom Anfangszustand x_0 des Prozesses abhängig sein. Bei der Annahme der Ergodizität ist aus Kapitel 2 bekannt, dass das Langzeitmittel gegen das Ensemblemittel konvergiert. Mit Hilfe der Gleichung (2.9) folgt unter der Annahme der Ergodizität (Stout (1974), S. 181), dass das Langzeitmittel fast sicher gegen das Ensemblemittel konvergiert (Birkhoff (1931), Arnold und Avez (1968), S. 16).

Satz 3.1: Ergodensatz von Birkhoff

Für einen ergodischen Prozess gilt:

$$\bar{g}(x_0) \stackrel{\text{a.s.}}{=} E(g(x)). \tag{3.13}$$

Der Beweis findet sich u. a. in Jetschke (1989), S. 142.

Nun kann das Langzeitmittel durch das Ensemblemittel ersetzt werden. Somit wird eine Unabhängigkeit vom Anfangszustand x_0 geschaffen.

Für ergodische Prozesse ist der Lyapunov-Exponent $\tilde{\lambda}(x_0)$ aus der Gleichung (3.11) unabhängig vom Anfangszustand. Zusätzlich entspricht er der Definition des Lyapunov-Exponenten für den stochastischen Fall mit additivem Rauschen in der Dynamik. Formal lässt er sich schreiben als:

$$\tilde{\lambda} = \int \ln|f'(x)| p(x) dx, \tag{3.14}$$

wobei $g(x) = \ln|f'(x)|$ und $p(x)$ eine Dichte des zu Grunde liegenden Prozesses beschreibt.

Da das Ensemblemittel offensichtlich durch:

$$\hat{\lambda} = \frac{1}{N} \sum_{i=0}^{N-1} \ln|f'(x_i)|, \tag{3.15}$$

geschätzt werden kann, stellt dieses ebenfalls einen Schätzer für den Lyapunov-Exponenten aus der Gleichung (3.1) dar.

Da in dieser Arbeit, bis auf die explizit erwähnten Experimente, die Dichte des zu Grunde liegenden Prozesses nicht bekannt ist, somit auch der wahre Lyapunov-Exponent nicht angegeben werden kann, wird der Einfachheit halber immer, wenn vom Lyapunov-Exponenten für stochastische Prozesse die Rede ist, implizit der Schätzer des Lyapunov-Exponenten betrachtet.

3.3 Lyapunov-Exponent als Klassifikationskriterium

Um eine erste grobe Einteilung von sowohl deterministischen als auch stochastischen Beobachtungs- bzw. Zeitreihen oder Prozessen machen zu können, ist ein geeignetes Klassifikationskriterium gesucht. Hierbei eignet sich die Güte der Prädizierbarkeit eines Prozesses gut als ein solches Kriterium. Da der Lyapunov-Exponent das Langzeitverhalten beschreibt und zu den Invarianten eines zu Grunde liegenden Systems gehört, d. h. er bleibt bei veränderten Initialbedingungen sowie unter einem Diffeomorphismus unverändert (Abarbanel, Brown, Sidorowich und Tsimring, 1993), bietet sich dieser als Kriterium sehr gut an. Mit Hilfe des Lyapunov-Exponenten ist es somit möglich, zwischen gut und schlecht prädizierbaren Zeitreihen zu unterscheiden (vgl. Kapitel 3.3.2). Durch Annahmen gewisser Eigenschaften an den Prozess, wie z. B. die Ergodizität, ist es ebenfalls möglich, stabile von instabilen Systemen zu unterscheiden (vgl. Kapitel 3.3.3).

Wie in Kapitel 2 dargestellt, lässt sich der Lyapunov-Exponent als Expansionsrate zweier Trajektorien auffassen. Hierbei wird das Auseinanderlaufen dieser Pfade, die zu einem Prozess gehören, von x_t auf x_{t+1} beobachtet. Wie sich zeigen wird, hängt die hier untersuchte Expansionsrate und der Informationsverlust über die Zeit zusammen. Ist die Expansionsrate positiv, d. h. ist der Lyapunov-Exponent positiv, wird der Wechsel von x_t auf x_{t+1} mit einem Informationsverlust durch das mit der Zeit immer größer werdende Unsicherheitsintervall einhergehen (vgl. Kapitel 3.3.1).

3.3.1 Informationsverlust

Das Wissen über die Güte der Prädizierbarkeit von Zeitreihen bzw. Prozessen ist eine wichtige Eigenschaft für die Interpretierbarkeit der prognostizierten Ergebnisse. Je größer der Informationsverlust bei einer Mehrschrittprognose sein wird, desto stärker wird auch die Güte der Prädizierbarkeit abnehmen. So ist es von Interesse, ein Maß des Informationsverlustes zu kennen, um mögliche Überinterpretationen bzw. Fehleinschätzungen zu vermeiden.

Der Lyapunov-Exponent kann als Expansionsrate aufgefasst werden, die in einem direkten Zusammenhang zum Informationsverlust über die Zeit betrachtet werden kann.

Nehmen wir o. B. d. A. an, der Startwert der gegebenen Beobachtungs- bzw. Zeitreihe x_0 (vgl. Gleichung (2.1) und Gleichung (2.5)) sei in einem Informationsgebiet I_0. Hierbei sei weiter angenommen, dass der Startwert x_0 um ein bestimmtes ϵ verschoben sein kann, was durch den Informationsball ausgedrückt wird. Das bedeutet, dass der exakte Startwert nicht bekannt ist, sondern nur seine Umgebung.

Durch die Annahme, dass sich der Startwert irgendwo in dem Informationsgebiet I_0 befindet, wird sich zur Zeit t die Datenreihe im Informationsgebiet I_t und zur Zeit $t+1$ im Informationsgebiet I_{t+1} befinden (vgl. Abbildung 3.2).

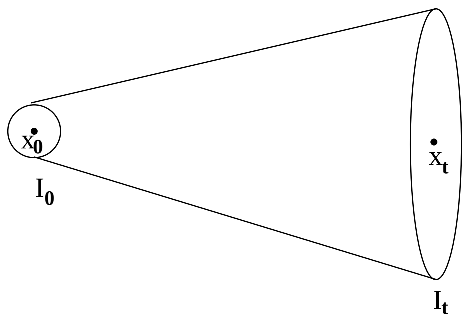

Abbildung 3.2: Informationszuwachs bzw. -verlust vom Informationsgebiet I_0 zum Informationsgebiet I_t

Je kleiner das Informationsgebiet ist, desto mehr Informationen liegen über die wahre Position des jeweiligen Iterationspunktes vor (Beck und Schlögl, 1993).

Als ein geeignetes Maß von Information bieten sowohl die Informationstheorie (Rohling und May, 1999) als auch die Thermodynamik (Beck und Schlögl, 1993) die Bitzahl an. Sie beschreibt die Anzahl an Stellen, die zur Bearbeitung eines Problems benötigt werden. Durch die Stellenangabe wird die Bitzahl in $\log 2$ als Einheit beschrieben (Rohling und May, 1999); in dieser Arbeit wird die normierte Form $\ln 2$, wie sie in der Thermodynamik vorkommt, verwendet (Ballard, 1997).

In dem hier benötigten Zusammenhang sei der Informationsgehalt b_n in folgender Form definiert (Beck und Schlögl, 1993):

Definition 3.3: Informationsgehalt

Der Informationsgehalt b_n, gemessen in einer Bitzahl, einer wahren Position eines Datenpunktes in einem Informationsgebiet I_n des Volumens Δ_n sei gegeben durch:

$$b_n := \ln \frac{1}{\Delta_n} = -\ln(\Delta_n). \tag{3.16}$$

Die Bitzahl ist genau dann positiv, wenn die Größe des Informationsgebietes kleiner als eins ist. Zudem spiegelt die Bitzahl offensichtlich das Verhältnis zwischen Größe und Informationsgehalt wider, da je größer das Gebiet ist, desto weniger Information vorhanden ist.

Die Größe Δ_n lässt sich durch $|x_n^* - x_n|$ charakterisieren und entspricht dem Abstand Δ_n, der zwischen zwei Trajektorien eines Prozesses zum Zeitpunkt n vorherrscht (vgl. Kapitel 3.1).

Die Länge bzw. Größe Δ_n des Informationsintervalls bzw. Informationsgebietes I_n lässt sich durch Umformung der Definitionsgleichung (3.16) ausrechnen.

Satz 3.2:

Die Größe eines Informationsgebietes, die die wahre Position eines Datenpunktes angibt, sei gegeben durch:

$$\Delta_n = \exp(-b_n). \tag{3.17}$$

Beweis:

Umformung der Gleichung (3.16)

q.e.d

Zur Berechnung der Prädiktionsgüte stellt sich nun die Frage, wie groß der Informationsverlust von einem Zeitpunkt zum nächsten Zeitpunkt ist. Hierzu wird die Differenz zweier Bitzahlen vor und nach einem Iterationsschritt bestimmt.

Satz 3.3:

Der Informationsverlust IV über die wahre Position eines Datenpunktes bei einem Iterationsschritt ist gegeben durch:

$$IV = b_n - b_{n+1} = \ln \Delta_{n+1} - \ln \Delta_n \approx \ln |f'(x_n)|. \tag{3.18}$$

Beweis:

$$\begin{aligned} b_n - b_{n+1} &= -\ln \Delta_n + \ln \Delta_{n+1} \\ &= \ln \left(\frac{\Delta_{n+1}}{\Delta_n} \right) \\ &\approx \ln \left(\frac{|f'(x_n)| \cdot \Delta_n}{\Delta_n} \right) \\ &= \ln |f'(x_n)|, \end{aligned}$$

wobei für Δ_1 und Δ_{n+1} gilt (vgl. Kapitel 3.1):

$$\begin{aligned} \Delta_1 &\approx |f'(x_0)| \cdot \Delta_0 \\ \Delta_{n+1} &\approx |f'(x_n)| \cdot \Delta_n. \end{aligned}$$

q.e.d

Ist die Differenz aus Gleichung (3.18) positiv, so beschreibt die Größe IV einen Informationszuwachs. Dies ist genau dann der Fall, wenn die Länge bzw. Größe Δ_n des Informationsintervalls bzw. -gebietes I_n größer ist als die Länge bzw. Größe von I_{n+1}. Ein Informationsverlust ist hingegen dann gegeben, wenn I_n kleiner ist als I_{n+1}.

Da der Informationsverlust die logarithmische erste Ableitung des funktionalen Zusammenhangs eines Prozesses an einer Stelle x_n darstellt, kann der Lyapunov-Exponent als Größe zur Beschreibung des durchschnittlichen Informationsverlusts bzw. auch Informationsgewinns aufgefasst werden.

Bemerkung 3.1:

Der Lyapunov-Exponenten beschreibt den durchschnittlichen Informationsverlust in folgender Form:

$$\lambda(X_0) = \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} \ln |f'(X_i)| \tag{3.19}$$

$$\approx \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} (b_i - b_{i+1}). \tag{3.20}$$

3.3.2 Unterscheidung zwischen prädizierbaren und nicht prädizierbaren Prozessen

Zur Analyse von Prozessen ist es oft hilfreich, ein geeignetes Kriterium zu finden, die gegebenen Beobachtungs- bzw. Zeitreihen sinnvoll zu klassifizieren. Durch die Güte der Prädizierbarkeit ist ein solches Kriterium gegeben.

Durch die Interpretation des Lyapunov-Exponenten als durchschnittlicher Informationsverlust (vgl. Kapitel 3.3.1) ergibt sich auf natürliche Weise eine Größe zur Klassifizierung zwischen gut prädizierbaren und schlecht prädizierbaren Zeit- und Beobachtungsreihen.

Im Gegensatz zu dem üblichen Vorgehen bei der Klassifizierung von gegebenen Datenreihen (Hastie, Tibshirani und Friedman, 2001) werden hier nicht die Datenpunkte selber, sondern ihre möglichen Positionsgebiete wie bei gewöhnlichem *k-means-Clustering* (Steinhausen und Langer, 1977), (Hastie, Tibshirani und Friedman, 2001) zur Analyse herangezogen.

Ausgehend von Kapitel 3.3.1 stellt die positive oder negative Ausdehnung der auseinander laufenden Trajektorien eine gute oder schlechte Prädizierbarkeit dar. Da der Lyapunov-Exponent als Größe zur Beschreibung des durchschnittlichen Informationsverlusts bzw. -gewinns aufgefasst werden kann, erfolgt die natürliche Klassifikation über die Betrachtung der Trajektorien. Laufen zwei benachbarte Trajektorien eines zu Grunde liegenden Prozesses zusammen, so ist der Lyapunov-Exponent kleiner als null und der Informationsgehalt nimmt in der Langzeitbetrachtung zu. Dehnt sich hingegen der Abstand zwischen den Trajektorien aus, liegt eine Divergenz der Pfade vor, so ist der Lyapunov-Exponent kleiner als null. In diesem Falle nimmt der Informationsgehalt, langzeitlich gesehen, durchschnittlich ab, da das Informationsgebiet, in dem der wahre Datenpunkt liegt, immer größer wird.

Zusammengefasst lässt sich die Klassifikation in gut und schlecht prädizierbare Datenreihen in folgender Form darstellen.

Definition 3.4: Gut versus schlecht prädizierbare Datenreihen

- *Sei* $\lambda(x_0) < 0 \Leftrightarrow \Delta_N < \Delta_0 \Rightarrow$ *Gute Prädizierbarkeit*

  Die Informationen über die wahre Positionen der Datenpunkte wachsen somit durch die Verkleinerung der Informationsgebiete. Folglich ist eine gute langzeitliche Prädiktion möglich und die Zeit- bzw. Beobachtungsreihe wird als gut prädizierbare Reihe klassifiziert. Die Trajektorien einer gegebenen Datenreihe mit unterschiedlichen Startwerten erreichen einen stabilen Grenzpunkt oder stabilen Grenzzyklus. In diesem Falle ist eine Konvergenz bzw. Konsistenz gegeben (vgl. Satz 3.4).

- *Sei* $\lambda(x_0) \approx 0 \Leftrightarrow \Delta_N \approx \Delta_0 \Rightarrow$ *Prädizierbarkeit wie ein Random Walk (vgl. Kapitel 2.4)*

  Der Informationsgehalt bleibt langzeitlich wie am Anfang des Prozesses. Da für einen Prozess wie dem Random Walk gilt: $\Delta_1 \approx 1 \cdot \Delta_0$ *und somit* $\Delta_N \approx \Delta_0$*, tritt weder ein Informationsverlust, noch ein Informationsgewinn ein. Folglich liegt eine langzeitliche Prädiktion wie bei einem Random Walk vor (vgl. Kapitel 3.4).*

- *Sei* $\lambda(x_0) > 0 \Leftrightarrow \Delta_N > \Delta_0 \Rightarrow$ *Schlechte Prädizierbarkeit*

  Die Trajektorien zweier unterschiedlicher, benachbarter Trajektorien divergieren exponentiell im Mittel mit einem Faktor $e^{N\lambda}$ *nach* N *Iterationen. In* *diesem Fall ist das Grenzverhalten nicht gleichförmig und der Informationsverlust über die wahre Position des Datenpunktes nimmt langfristig immer mehr zu. Folglich ist keine gute langzeitliche Prädiktion möglich und die Zeit- bzw. Beobachtungsreihe wird als schlecht prädizierbar klassifiziert.*

Dieses Klassifikationskriterium kann ohne weitere Annahmen über die Modellierung des Prozesses benutzt werden. Es ist möglich, zwischen gut und schlecht prädizierbaren Prozessen zu unterscheiden, auch wenn nur die Beobachtungs- bzw. Zeitreihe gegeben ist (Busse, Steuer und Weihs, 2001).

Daher eignet sich der Lyapunov-Exponent als Klassifikationskriterium sehr gut, da dadurch eine grobe Einteilung in prädizierbar oder nicht prädizierbar gemacht werden kann, so dass z. B. für eine weiterführende Modellierung des Prozesses die Klasse der möglichen Modelle eingeschränkt werden kann.

3.3.3 Unterscheidung zwischen stabilen und instabilen Prozessen

Ohne Einschränkung bzgl. des zu Grunde liegenden Prozesses bzw. des funktionalen Zusammenhangs ist eine Einteilung in gut und schlecht prädizierbare Datenreihen möglich. Nimmt man gewisse Einschränkungen des Prozesses in Kauf, d. h. werden bestimmte Eigenschaften vorausgesetzt, so ist es mit Hilfe des Lyapunov-Exponenten ebenfalls möglich, zwischen stabilen und instabilen Prozessen zu unterscheiden.

Da der Lyapunov-Exponent als Konvergenz oder Divergenzkriterium angesehen werden kann, deutet ein negativer Lyapunov-Exponent auf einen stabilen Prozess hin, da egal, bei welchem Startwert begonnen wird, immer das gleiche Grenzverhalten auftritt. Ist hingegen der Lyapunov-Exponent positiv, so hängt das Langzeitverhalten stark von dem Anfangspunkt ab. In diesem Fall liegt ein instabiler Prozess vor. Genauer gilt:

Satz 3.4: Stabile versus instabile Datenreihen

- *Gegeben sei ein stabiler Prozess* $\Rightarrow \lambda(x_0) < 0$
- *Gegeben sei ein Prozess wie ein Random Walk* $\Rightarrow \lambda(x_0) \approx 0$.
- *Gegeben sei ein instabiler Prozess* $\Rightarrow \lambda(x_0) > 0$

Beweis:

Im Beweis, dass der Lyapunov-Exponent eines stabilen Prozesses kleiner als null ist, wird zwischen vier Fällen unterschieden, deterministischen und stochastischen und Fixpunkt und periodischem Orbit.

1. Deterministischer Prozess, Fixpunkt:

   Gegeben sei ein deterministischer Prozess $x_{t+1} = f_t(x_0) = f(x_t)$ (vgl. Gleichung (2.1)). Der Startwert x_0 gehöre zum Anziehungsbereich eines asymptotisch stabilen Fixpunktes x^0. Somit wird impliziert:

$$
\begin{aligned}
&\lambda(x_0) = \lambda(x^0). \qquad \text{Also gilt:} \\
&\lambda(x^0) = \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} \ln|f'(x^0)| \qquad \text{(Lyapunov-Exponent,} \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{gestartet im Fixpunkt)} \\
&\quad = \lim_{N\to\infty} \frac{1}{N} N \ln|f'(x^0)| \qquad \text{(Summe ist unabhängig von } i) \\
&\quad = \lim_{N\to\infty} \ln|f'(x^0)| \\
&\quad = \ln|f'(x^0)| \qquad \text{(Unabhängigkeit von } N) \\
&\quad = \ln\left|\lim_{x\to x^0} \frac{f(x)-f(x^0)}{x-x^0}\right| \\
&\quad < \ln\frac{\epsilon}{\delta} \qquad \text{(vgl. Gleichung (2.2) und anschließende Bemerkung)} \\
&\quad \leq \ln 1 \qquad \text{(o. B. d. A. } \epsilon \text{ ist kleiner als } \delta \text{, vgl. Gleichung (2.2))} \\
&\quad = 0
\end{aligned}
\tag{3.21}
$$

Hierzu sei noch einmal auf Kapitel 2 und auf Jetschke (1989), S. 117, hingewiesen.

2. Stochastischer Prozess, Fixpunkt:

   Gegeben sei ein stochastischer Prozess $X_{t+1} = f_t(X_0, \epsilon) = f(X_t, \epsilon)$ (vgl. Gleichung (2.5)). Der Startwert x_0 gehöre zum ϵ-Anziehungsbereich eines asymptotisch stochastisch stabilen Fixpunktes x^0.

   Zuerst einmal ist zu zeigen, dass $\lambda(x_0) \leq \lambda(x^0)$.

$$
\begin{aligned}
\lambda(x_0) &= \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} \ln|f'(x_i)| \qquad \text{(Lyapunov-Exponent,} \\
&\qquad\qquad\qquad\qquad\qquad\qquad \text{gestartet im Anfangspunkt)} \\
&= \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} g(x_i) \qquad \text{(mit einer Transformation } g(x) = \ln|f'(x)|) \\
&= E[g(X)] \qquad \text{(vgl. Gleichung (2.9))} \\
&\leq g[E(X)] \qquad \text{(Jensensche Ungleichung} \\
&\qquad\qquad\qquad \text{mit der konkaven Funktion ln)} \\
&= g(x^0) \qquad \text{(Der Erwartungswert von } X \text{ ist auf } x^0 \text{ gesetzt,} \\
&\qquad\qquad\quad E(X) = x^0\text{).} \\
&= \ln|f'(x^0)| \qquad \text{(Rücktransformation)} \\
&= \lambda(x^0) \qquad \text{(vgl. deterministischer Fall).}
\end{aligned}
\tag{3.22}
$$

   Der weitere Beweis ist analog zum deterministischen Fall. Der Lyapunov-Exponent, wobei die Berechnung im Anfangszustand begann, ist kleiner als null.

3. Deterministischer Prozess, periodischer Orbit:

   Gegeben sei ein deterministischer Prozess $x_{t+1} = f_t(x_0) = f(x_t)$ (vgl. Gleichung (2.1)). Der Anfangspunkt x_0 gehöre zum Anziehungsbereich eines periodischen Orbits mit einer Periode K, $\overline{x^1}, ..., \overline{x^K}$. Für einen periodischen Orbit gilt dann:

   $\overline{x^k} = \overline{x^{(k+K)}}$ und $\overline{x^{(k+1)}} = \overline{x^{(k+1+K)}}$
   und $\overline{x^{(k+2)}} = \overline{x^{(k+2+K)}}$ u. s. w.
   Definiere also: $\overline{x^{K+k}} = \overline{x^k}, k = 1, \ldots, K$.
   Weiter gilt: $\overline{x^k} = f^K(\overline{x^k}), k = 1, \ldots, K$.

   Diese Gleichungen implizieren somit auch, dass:

   $f'(\overline{x^k}) = (f^{K+1})'(\overline{x^k})$ und $f^{K+1}(\overline{x^k}) = f^K(\overline{x^{k+1}})$, weil $f(\overline{x^k}) = \overline{x^{(k+1)}}$.

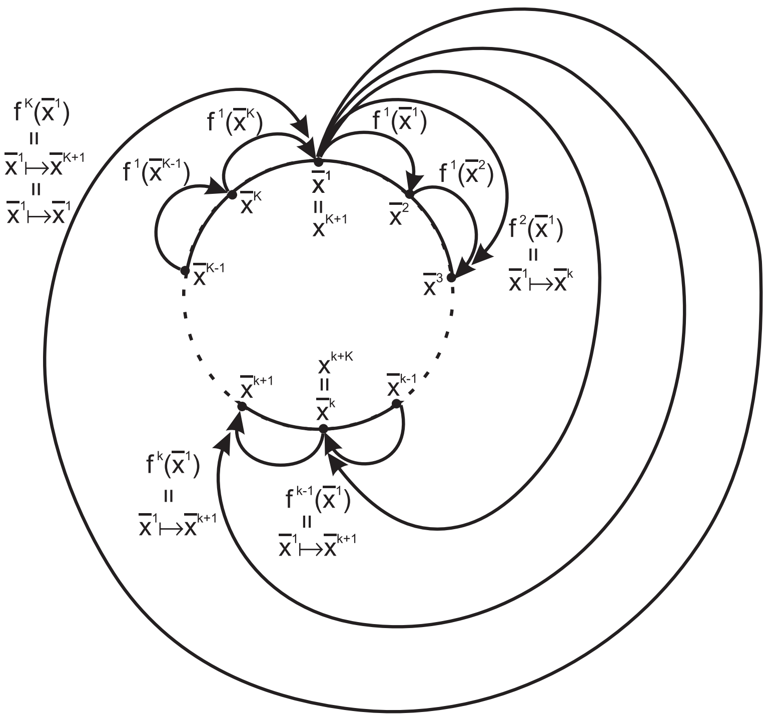

Abbildung 3.3: Darstellung der Zeitzustände und funktionalen Zusammenhänge in einem periodischen Orbit

Die Abbildung 3.3 verdeutlicht noch einmal die Zusammenhänge zwischen den unterschiedlichen Zeitpunkten und ihren funktionalen Zusammenhängen.

Somit gilt dann mit K Fixpunkten, wobei $k = 1, \ldots, K$:

$$
\begin{aligned}
|(f^K)'(\overline{x^{k+1}})| &= \left| \lim_{\Delta x \to 0} \frac{f^K(\overline{x^{k+1}} + \Delta x) - f^K(\overline{x^{k+1}})}{\Delta x} \right| < 1 \qquad \text{(wg. Definition,} \\
&\qquad \text{vgl. determ. Prozess mit Fixpunkt)} \\
|(f^{K+1})'(\overline{x^k})| &< 1 \qquad \text{(wegen } f^{K+1}(\overline{x^k}) = f^K(\overline{x^{k+1}})\text{)} \\
|f'(\overline{x^k})| &< 1 \qquad \text{(wegen } f'(\overline{x^k}) = (f^{K+1})'(\overline{x^k})\text{).}
\end{aligned}
\tag{3.23}
$$

Für den Lyapunov-Exponenten folgt nun:

$$
\begin{aligned}
\lambda(x_0) &= \frac{1}{K} \sum_{k=1}^{K} \ln |f'(\overline{x^k})| \qquad \text{(Anziehungsbereich eines period. Orbits)} \\
&< \frac{1}{K} \cdot K \cdot \ln 1 \qquad \text{(vgl. deterministischer Fall, Fixpunkt)} \\
&< 0.
\end{aligned}
\tag{3.24}
$$

4. Stochastischer Prozess, periodischer Orbit:

   Gegeben sei ein stochastischer Prozess $X_{t+1} = f_t(X_0, \epsilon) = f(X_t, \epsilon)$ (vgl. Gleichung (2.5)). Der Anfangszustand x_0 gehöre zu dem Anziehungsbereich eines periodischen Orbits der Länge K, $\overline{(x^1)}, ...\overline{(x^K)}$.

   Ohne Beschränkung der Allgemeinheit gilt:
   $\lambda(x_0) \leq \lambda(\overline{x^k})$ (vgl. stochastischer Fall, Fixpunkt im ϵ-Anziehungsbereich).

   Somit gilt für den Lyapunov-Exponent eines stochastischen Prozesses mit einem periodischen Orbit:

$$\begin{aligned}
\lambda(x_0) &\leq \frac{1}{K}\sum_{i=1}^{K} \ln|f'\overline{(x^i)}| && \text{(vgl. determ. Fall, periodischer Orbit)}\\
&< \frac{1}{K}\cdot K \cdot \ln 1 && \text{(vgl. deterministischer Fall, Fixpunkt)}\\
&< 0.
\end{aligned} \tag{3.25}$$

q.e.d

Bemerkung 3.2:

Modellierungsvorschläge:

- *Ist der Lyapunov-Exponent kleiner als null, ist also die Zeit- bzw. Beobachtungsreihe als stabil klassifiziert worden, so gibt es z. B. die Möglichkeit, eine Modellierung des Prozesses mit Hilfe von ARMA-Prozessen zu versuchen. Hierbei muss eine zum ARMA-Prozess benötigte Parametereinstellung verwendet werden, die einen ergodischen und mittelwertstationären und stabilen Prozess erzeugt (vgl. Kapitel 2.4 und 3.4).*

- *Ist der Lyapunov-Exponent null, ist eine erste Modellierung mit Hilfe eines Random Walk sicherlich eine gute Wahl (vgl. Kapitel 2.4 und 3.4).*

- *Ist der Lyapunov-Exponent größer als null, somit als ein instabiler Prozess klassifiziert worden, bieten sich chaotische Prozesse an. Hierzu sei auf die nachfolgende Definition verwiesen.*

Bleibt zu der Klassifikation zwischen stabilen und instabilen Prozessen also noch zu zeigen, dass chaotische Systeme einen Lyapunov-Exponenten größer als null haben. Dieses ist jedoch schon durch die Definition von Chaos festgesetzt (Tong (1993), S. 59).

Definition 3.5: Chaotische Datenreihe

Eine chaotische Datenreihe liegt genau dann vor, wenn der Lyapunov-Exponent $\lambda(x_0) > 0$ *ist.*

Hierbei ist festzuhalten, dass somit das asymptotische Verhalten lokal instabil ist.

In der Physik werden Datenreihen als chaotisch bezeichnet, wenn sie einen *strange attractor* aufweisen. Instabilität im Grenzverhalten und der Begriff der *Strange Attraktoren* fallen hierbei zusammen (Eckmann und Ruelle, 1982), (Grassberger und Procaccia, 1983a), (Grassberger und Procaccia, 1983b), so dass auch hier Chaos über den Lyapunov-Exponenten definiert wird. Für weiterführende Literatur verweise ich jedoch auf Jetschke (1989), Abarbanel (1996), Kantz und Schreiber (1997) und Buzug (1994).

Eine erste Möglichkeit, chaotische Systeme zu modellieren, kann mit Hilfe von rekurrenten neuronalen Netzen (Stagge, 2000) oder durch Mischdichtemodellen (Reimetz, 1998), (Kreutz, 2000) realisiert werden. Ist also eine Datenreihe als instabil, also chaotisch, identifiziert worden, bieten sich diese beiden nichtlinearen und nichtparametrischen Modellierungsansätze an. Beide Verfahren lassen Strukturoptimierung zu, so dass eine gewisse Flexibilität gegeben ist (Tsoi und Back, 1997), (Reimetz, 1998), (Kreutz, 2000).

Zusammenfassend kann gesagt werden, dass nach einer groben Einteilung mit Hilfe der gegebenen Datenreihen in stabile und instabile Prozesse Modelle für eine weitere Analyse zur Verfügung stehen. Der Lyapunov-Exponent ist somit ein nützliches Werkzeug für detailliertere Analysen der zu Grunde liegenden Systeme.

3.4 Spezielle Prozesse

In Kapitel 3.3.3 zeigte sich bereits, dass aufgrund der Definition chaotische und instabile Datenreihen einen Lyapunov-Exponenten größer als null haben. Weiter zeigte sich, dass die stabilen Prozesse hingegen einen Exponenten kleiner als null haben.

In dieser Arbeit werden spezielle Prozesse zur experimentellen Unterstützung heran gezogen (vgl. Kapitel 4.2, 5.2 und 6.2). Somit ist es hilfreich, die Lyapunov-Exponenten dieser Prozesse anzugeben, um eine Einordnung der Ergebnisse zu ermöglichen. Hierzu ist die Definition auf die speziellen Prozesse anzuwenden.

AR-Prozesse

Der Lyapunov-Exponent für einen AR(1)-Prozess hängt vom Parameter ϕ_1 ab, der ebenfalls über die Stationarität dieses Prozesses entscheidet. Ist $|\phi_1|$ kleiner als eins, so ergibt sich ein Lyapunov-Exponent kleiner als null.

Gegeben sei also die folgende Prozessgleichung (vgl. Gleichung (4.1)):

$$AR(1) : X_t = \phi_1 X_{t-1} + \epsilon_t, \tag{3.26}$$

Für den Lyapunov-Exponenten ergibt sich dann nach Einsetzen des funktionalen Zusammenhangs in die Definition:

$$\lambda(x_0)_{AR(1)} = \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} \ln|f'(x_i)| = \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} \ln|\phi_1|. \tag{3.27}$$

Ist also der Parameter ϕ_1 vom Betrag her kleiner als eins, ist der Lyapunov-Exponent kleiner als null. Ist er hingegen größer als eins gewählt, so ist $\lambda(x_0)_{AR(1)}$ größer als null.

Für den in den folgenden Kapiteln speziellen AR(1)-Prozess mit $\phi_1 = 0.9$ ergibt sich ein Lyapunov-Exponent von -0.105.

MA-Prozesse

Für den MA(1)-Prozess, der in folgender Form definiert ist (vgl. Gleichung (4.2)):

$$MA(1) : X_t = \theta_1 \epsilon_{t-1} + \theta_0,\ \theta_1 \neq 0, \theta_0 = 1 \tag{3.28}$$

ist aufgrund der Theorie zu vermuten, dass der Lyapunov-Exponent sich als stabil erweist. Um dieses jedoch zeigen zu können, muss der Prozess in Form einer autoregressiven Darstellung gewählt werden, da ansonsten kein von x abhängendes Signal in der Funktionsgleichung erscheint. Aus der Theorie ist bekannt, das ein MA(1)-Prozess als $AR(\infty)$-Prozess darstellbar ist. Es können also folgende Umformungen vorgenommen werden:

$$\begin{aligned}
X_{t-2} &= \theta_1\epsilon_{t-3} + \epsilon_{t-2} \Rightarrow \epsilon_{t-2} = X_{t-2} - \theta_1\epsilon_{t-3}, \\
X_{t-1} &= \theta_1\epsilon_{t-2} + \epsilon_{t-1} \Rightarrow \epsilon_{t-1} = X_{t-1} - \theta_1\epsilon_{t-2}, \\
X_t &= \theta_1\epsilon_{t-1} + \epsilon_t \Rightarrow \theta_1(X_{t-1} - \theta_1\epsilon_{t-2}) + \epsilon_t \\
&= \theta_1 X_{t-1} - \theta_1^2\epsilon_{t-2} + \epsilon_t \\
&= \theta_1 X_{t-1} - \theta_1^2(X_{t-2} - \theta_1\epsilon_{t-3}) + \epsilon_t \\
&= \theta_1 X_{t-1} - \theta_1^2 X_{t-2} + \theta_1^3\epsilon_{t-3} + \epsilon_t \quad \text{u. s. w.}
\end{aligned}$$

Hierbei tritt nun zum ersten Mal das Problem auf, mehrere Signale zu unterschiedlichen Zeitpunkten nutzen zu müssen. Da jedoch in dieser Arbeit nur der größte Lyapunov-Exponent und nicht das Lyapunov-Spektrum (Abarbanel, Brown, Sidorowich und Tsimring, 1993) betrachtet wird, ist die interessierende Größe nur die, wo der Vorfaktor der sich ergebenen X_i am größten ist. Nehmen wir an, $0 < \theta_1 < 1$, dann ist θ_1 und nicht (z. B.) die quadrierte Form ausschlaggebend. Für einen stabilen Prozess muss $|\theta_1|$, wie bei einem AR(1)-Prozess, kleiner als eins gewählt werden (vgl. AR-Prozesse).

ARMA-Prozesse

Ein ARMA(1,1)-Prozess (vgl. Gleichung (4.3)):

$$ARMA(1,1) : X_t = \phi_1 X_{t-1} + \theta_1\epsilon_{t-1} \tag{3.29}$$

sollte sich bzgl. des Lyapunov-Exponenten wegen des autoregressiven Anteils ähnlich wie ein AR(1)-Prozess verhalten.

Das Berechnen des Lyapunov-Exponenten bestätigt dieses:

$$\lambda(x_0)_{ARMA(1,1)} = \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} \ln|f'(x_i)| = \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} \ln|\phi_1|. \tag{3.30}$$

Wird also der Parameter, der zum autoregressiven Anteil gehört, vom Betrag her kleiner als eins gewählt, so zeigt sich ein stabiler Zustand. Andernfalls wird der Prozess im Sinne des Lyapunov-Exponenten als instabil charakterisiert.

Für den in den nachfolgenden Kapiteln gewählten speziellen ARMA(1,1)-Prozess mit $\phi_1 = 0.9$ und $\theta_1 = 0.1983$ ergibt sich ein Lyapunov-Exponent von -0.105.

Random Walk

Ein nächster wichtiger Prozess ist durch den Random Walk gegeben, der wie folgt definiert ist:

$$X_t = \begin{cases} \epsilon_t & , \quad t = 1 \\ X_{t-1} + \epsilon_t & , \quad t = 2, 3, \ldots \end{cases}, \tag{3.31}$$

wobei ϵ_t Weißes Rauschen ist.

Von Park und Whang (1999) wurde bereits bewiesen, dass der Lyapunov-Exponent gleich Null ist. Der Vollständigkeit halber sei es noch einmal aufgezeigt.

$$\lambda(x_0)_{RandomWalk} = \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} \ln|1| = 0. \tag{3.32}$$

Schwingungsmodelle

Sei zum Schluss noch das spezielle Schwingungsmodell (vgl. Gleichung (2.15))

$$X_t = \alpha \sin x_{t-1} + \epsilon_t \tag{3.33}$$

bezüglich seines Lyapunov-Exponenten untersucht.

$$\lambda(x_0)_{sin} = \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} \ln|f'(x_i)| = \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} \ln|\alpha \cdot \cos(x_i)|. \tag{3.34}$$

Da der Kosinus vom Betrag her immer kleiner oder gleich eins ist, muss der Gewichtungsparameter α für einen stabilen Prozess so gewählt werden, dass $|\alpha \cdot \cos x| < 1$ ist.

In den folgenden Kapiteln wird der Spezialfall $\alpha = 0.699$ gewählt, so dass ein stabiler Prozess generiert wird.

Kapitel 4

Untersuchungen zum Lyapunov-Exponenten

Ist der funktionale Zusammenhang des zu Grunde liegenden Prozesses bekannt, so ist es analytisch möglich, den Lyapunov-Exponenten direkt zu berechnen (Bougerol und Picard, 1992). Jedoch erübrigt sich dann eine Klassifikation, da ein direktes Modell vorhanden ist. In dieser Arbeit werden nur Fälle betrachtet, in denen der funktionale Zusammenhang nicht bekannt und nur eine gegebene Datenreihe zur Analyse zur Verfügung steht. Hieraus entsteht das Problem der Berechnung des Lyapunov-Exponenten. Anstatt sich den funktionalen Zusammenhang zu Nutze machen zu können, muss das Klassifikationskriterium aus der gegebenen Datenreihe geschätzt werden. Hierzu stehen in der Literatur einige Verfahren zur Verfügung, die in Kapitel 4.1 kurz dargestellt werden. Auch wird hier auf das in dieser Arbeit zur Berechnung des Lyapunov-Exponenten verwendete Verfahren eingegangen.

Zur Überprüfung des Klassifikationskriteriums werden einige Simulationen durchgeführt. Zum einen werden aus der Statistik bekannte Modelle verwendet (vgl. Kapitel 4.2.1), zum anderen wird die Modulo-1-Abbildung benutzt, da sie sowohl stabiles als auch instabiles Verhalten je nach Parameterwahl aufweisen kann. Somit ist es möglich, zuerst Simulationen für eine gewöhnliche Klassifikation zu machen (vgl. Kapitel 4.2.2), anschließend das Startwertproblem zu untersuchen (vgl. Kapitel 4.2.3), mögliche Fehlklassifikationen im kritischen Bereich, d. h. im Bereich vom stabilen zum instabilen Zustand, zu identifizieren (vgl. Kapitel 4.2.4) und Experimente mit kurzen Datenreihen auf ihre Klassifikationsfähigkeit zu überprüfen (vgl. Kapitel 4.2.5).

4.1 Berechnung des Lyapunov-Exponenten

Ist nur die Datenreihe für eine Analyse des Prozesses mit Hilfe des Lyapunov-Exponenten gegeben, so steht der funktionale Zusammenhang zur analytischen Berechnung nicht zur Verfügung. Die Ableitung f' (vgl. Gleichungen (2.1) und (2.5)), die zur Berechnung des Lyapunov-Exponenten bekannt sein muss, ist somit ebenfalls nicht gegeben. Folglich muss versucht werden, mit Hilfe der zur Verfügung stehenden Datenreihe das Klassifikationskriterium auf anderem Wege zu schätzen.

Auf der einen Seite kann die Ableitung f' numerisch berechnet werden, auf der anderen Seite wird das Divergenz- bzw. Konvergenzverhalten von zwei benachbarten Trajektorien graphisch analysiert.

In der Literatur sind bereits Verfahren aus der nichtlinearen Dynamik bekannt, den Lyapunov-Exponenten alleine durch die gegebene Zeit- bzw. Beobachtungsreihe zu berechnen, ohne dabei den zu Grunde liegenden Prozess explizit zu kennen.

Wird die Repräsentation einer Datenreihe als Trajektorie betrachtet, wie in Kapitel 3.1 bereits erläutert wurde, dann ist der Lyapunov-Exponent darstellbar als eine Anfangsdistanz zwischen den Startwerten x_0 und x_0^*, also Δ_0, multipliziert mit dem exponentiellen Wachstum der beiden benachbarten Trajektorien.

Diese Beziehung wird auch für viele numerische Verfahren zur Berechnung des Lyapunov-Exponenten genutzt. Das in dieser Arbeit verwendete Verfahren basiert ebenfalls auf der anschaulichen Darstellung und hat zudem den Vorteil, bereits in dem Programmpaket *TISEAN* (Hegger, Kantz und Schreiber, 1999), das sich mit der Analyse von nichtlinearen Dynamiken befasst, implementiert zu sein. Hierbei können natürlich nicht zwei Trajektorien des zu Grunde liegenden Prozesses verwendet werden. Hingegen werden in einer kleinen Umgebung (einem kleinen Zeitfenster) ϵ der Beobachtungs- bzw. Zeitreihe die Daten in einem Raum eingebettet und durch die entstehende Punktewolke eine Gerade angepasst. Die Steigungen der entstehenden Geraden über alle möglichen Umgebungen der Datenreihen geben dann als arithmetisches Mittel einen guten Schätzer für den Lyapunov-Exponenten wieder (Hegger, Kantz und Schreiber, 1999). Es kann jedoch passieren, dass bei einer zu starken Varianz des Rauschlevels die Steigungen der lokal entstandenen Geraden größer als eins werden. Die Folge davon ist eine falsche Schätzung des Lyapunov-Exponenten.

Weitere in der Literatur zur Verfügung stehenden Verfahren basieren ebenfalls auf dieser Vorgehensweise, jedoch wird z. B. keine Umgebung, sondern die direkt aufeinander folgenden Referenzpunkte betrachtet (Rosenstein, Collins und DeLuca, 1993).

Sano und Sawada (1985) liefern ein Verfahren, was nicht nur den in dieser Arbeit betrachteten (größten) Lyapunov-Exponenten, sondern das so genannte Lyapunov-Spektrum berechnet (Sano und Sawada, 1985), (Eckmann, Kamphorst, Ruelle und Ciliberto, 1986).

Ein anderer Weg, den Lyapunov-Exponenten zu schätzen, ist die Ableitung f' des zu Grunde liegenden, unbekannten funktionalen Zusammenhangs f numerisch zu bestimmen. Hierzu wurden mehrere Methoden vorgeschlagen. Für eine genauere Beschreibung sei z. B. auf Abarbanel (1996), Eckmann und Ruelle (1982), Gencay (1996) und Wolf, Swinney und Vastan (1985) verwiesen.

Es sei noch einmal angemerkt, dass der Lyapunov-Exponent ein Attribut des Attraktors, des Grenzverhaltens, des zu Grunde liegenden Systems ist. Dieser hängt somit nicht von der Zeit ab. Dieses wiederum impliziert, dass sich die Beobachtungen für die Berechnung des Lyapunov-Exponenten λ nicht mehr im *transienten*, d. h. sich ändernden Zustand befinden. Die Datenreihe sollte im günstigsten

Fall also den „asymptotischen Zustand“ erreicht haben. Ist die gegebene Reihe lang genug, so sind in der Regel die geschätzten Lyapunov-Exponenten, die mit Hilfe der in der Literatur vorgeschlagenen Verfahren berechnet werden, gute Vertreter, um die gewünschte Klassifikation richtig zu treffen.

4.2 Experimentelle Ergebnisse

In dieser Arbeit wird für alle Berechnungen des Lyapunov-Exponenten das von Kantz und Schreiber (1997) vorgeschlagene und in dem Programmpaket *TISEAN* implementierte Verfahren verwendet. Anhand von in der Statistik häufig benutzen und viel untersuchten linearen Prozessen, die hier durch einen AR(1), einen MA(1) und einen ARMA(1,1) repräsentiert werden, soll zunächst exemplarisch gezeigt werden, wie Zeitreihen - und nicht nur Beobachtungsreihen - klassifiziert werden (vgl. Kapitel 4.2.1). Dem anschließen sollen sich Untersuchungen mit Hilfe der Modulo-Abbildung. Durch die Wahl der Parameter ist es möglich, sowohl stabile als auch instabile Prozesse zu erzeugen (vgl. Kapitel 4.2.2). Zudem hat es den Vorteil, dass der wahre Lyapunov-Exponent analytisch berechenbar ist. Durch die Berechnung ist es dann weiter möglich, alle interessierenden Prozesse durch Variation der Parameter zu erreichen und auch im kritischen Bereich, d. h. an der Klassifikationsgrenze um null herum, Untersuchungen anzustellen (vgl. Kapitel 4.2.4). Da durch die Berechnung im günstigen Fall der asymptotische Zustand erreicht sein sollte, werden in Kapitel 4.2.5 Untersuchungen mit kurzen Zeitreihen von ausgewählten Modulo-Abbildungen vorgenommen. Hierdurch werden exemplarisch Fehlklassifikationen durch zu kurze Zeitreihen vorgestellt. Sowohl Kapitel 4.2.4 als auch Kapitel 4.2.5 beschäftigen sich mit der Frage, ob bei Verwendung des Lyapunov-Exponenten zur Klassifikation durch die numerischen Verfahren Fehlklassifikationen denkbar sein können.

4.2.1 Untersuchungen mit stochastischen Prozessen

In der Statistik sind die wohl bekanntesten linearen Prozesse die *Autoregressiven Prozesse*, kurz AR-Prozesse, die *Moving-Average-Prozesse*, kurz MA-Prozesse, und die *Autoregressiven-Moving-Average-Prozesse*, kurz ARMA-Prozesse. Ebenfalls oft in der Statistik in Erscheinung tretende Prozesse sind die Schwingungsmodelle (vgl. Kapitel 2.4). Da ursprünglich die Verwendung des Lyapunov-Exponenten nur für deterministische Beobachtungsreihen vorgesehen war, werden hier stellvertretend für bekannte Zeitreihen ein AR(1), ein MA(1), ein ARMA(1,1) und ein sin-Prozess untersucht.

Ein AR(1)-Prozess zeichnet sich dadurch aus, nur vom vorangehenden Wert abzuhängen und mit einem Weißen Rauschen gestört zu werden. Die hier verwendete Form ist gegeben durch:

$$AR(1) : X_t = \phi_1 X_{t-1} + \epsilon_t, \tag{4.1}$$

wobei ϵ_t Weißes Rauschen ist mit Varianz Eins (vgl. z. B. Fahrmeir, Kaufmann und Ost (1981), S. 215 und Kapitel 2.4).

Je nach Parameterwahl besitzt der AR(1)-Prozess die Eigenschaft, konvergentes bzw. divergentes Grenzverhalten aufzuweisen. Wird der Parameter $|\phi_1|$ kleiner als eins gewählt, so konvergiert die Zeitreihe. Wird hingegen der Parameter $|\phi_1|$ größer als eins gewählt, so tritt eine Divergenz in der asymptotischen Betrachtung auf. In dieser Simulation soll das Augenmerk auf einen AR(1)-Prozess gelegt werden, wo das asymptotische Verhalten stabil ist. Jedoch finden sich im Anhang ebenfalls Versuchsergebnisse zu der instabilen Form (Anhang A.1).

Es ist also der Parameter ϕ_1 so zu wählen, dass $|\phi_1| < 1$ ist. Hierzu wurden im Vorfeld Voruntersuchungen durchgeführt, um eine geeignete Wahl für ϕ_1 treffen zu können. Die Ergebnisse der Voruntersuchungen finden sich in Anhang A.1.

Ausgehend von diesen Untersuchungen wird der Parameter ϕ_1 auf $\phi_1 = 0.9$ gesetzt. Somit wird zum einen dem Anspruch genügt, ein stabiles asymptotisches Verhalten zu erzeugen, zum anderen liegt der Parameter nicht zu weit vom kritischen Bereich ($\phi_1 \approx 1$) entfernt.

Nun sollen 1024 Datenpunkte aus einem AR(1) mit $\phi_1 = 0.9$ erzeugt werden, um zu überprüfen, ob durch die Berechnung des Lyapunov-Exponenten die simulierte Zeitreihe richtig klassifiziert wird. Ist also der durch *TISEAN* berechnete Lyapunov-Exponent kleiner als null, so konnte der Prozess richtig klassifiziert werden. Wird hingegen der Lyapunov-Exponent größer als null sein, so würde es auf einen instabilen Prozess hindeuten und es läge eine Fehlklassifikation vor.

Voruntersuchungen zum Startwert ergaben, dass mit einem Erfolgsprozentsatz von 96% bei unterschiedlich gewählten Startwerten die richtige Klassifikationsentscheidung getroffen wurde. Als geeigneter Startwert erwies sich $x_0 = 0.699$, so dass dieser festgelegt wird. Zur Veranschaulichung der erzeugten Zeitreihe diene deren graphische Darstellung in Abbildung 4.1.

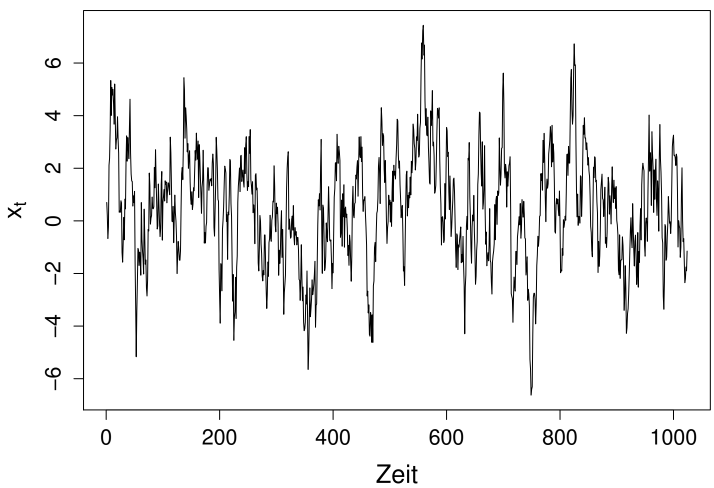

Abbildung 4.1: Reihendarstellung des durch Gleichung (4.1) erzeugten AR-Prozesses.

Um schon einmal einen Eindruck zu gewinnen, wie sich die Zustände x_t gegen x_{t-1} verhalten, wird ein Phasenbild erstellt. Ist dieses „regelmäßig“, deutet es auf einen eher konvergenten Prozess hin. Ein solches Phasenbild wird in Abbildung 4.2 dargestellt.

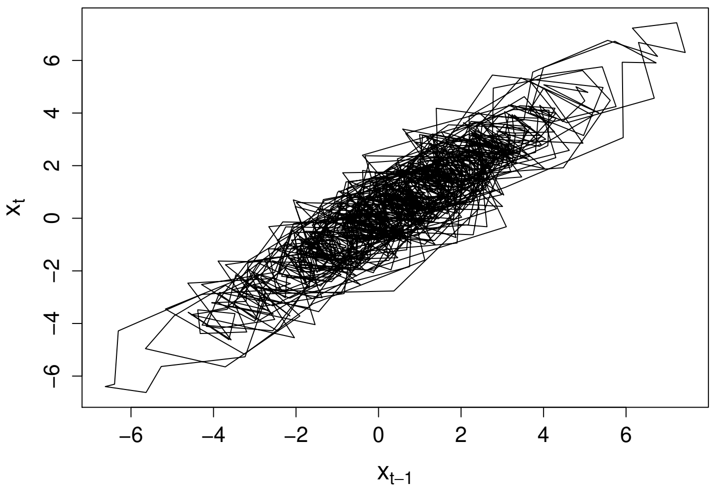

Abbildung 4.2: Phasendarstellung von x_t und x_{t-1} des durch Gleichung (4.1) erzeugten AR-Prozesses.

Es deutet wegen der regelmäßigen Struktur auf einen stabilen Prozess hin. Jedoch sind Phasenbilder nur veranschaulichend und sehr subjektiv, so dass die Berechnung des Lyapunov-Exponenten als Klassifikationskriterium mehr Objektivität leistet.

Einen Wert von $\hat{\lambda} = -0.0683$ ergab die Berechnung des Lyapunov-Exponenten, also einen Wert kleiner als null, so dass die Zeitreihe richtig klassifiziert wurde.

Eine weitere bedeutende Prozessklasse in der linearen Zeitreihenanalyse stellt der MA-Prozess dar (vgl. Stier (2001), S. 52):

$$MA(1) : X_t = \theta_1 \epsilon_{t-1} + \epsilon_t. \tag{4.2}$$

Ihre Haupteigenschaft liegt darin, nur vom Fehler des voran gegangenen Zeitpunktes abzuhängen und somit den Wert selbst nicht zu berücksichtigen.

Dieser Prozess ist also im Hinblick auf eine Klassifikation sehr schwierig, da nur die zufälligen Störungen in die Berechnungen eingehen. Da hier wiederum ein Parameter θ_1 frei wählbar ist, wurden Voruntersuchungen hierzu durchgeführt (vgl. Anhang A.2). Die Wahl fiel auf $\theta_1 = 0.1983$. So sollte aus der Theorie heraus ein stabiler Attraktor erzeugt werden.

Wie schon bei dem simulierten AR(1)-Prozess werden zur Überprüfung der Klassifikation in einen stabilen oder instabilen Prozess bzw. zur Überprüfung der Klassifikation in einen gut oder schlecht prädizierbaren Prozess 1024 Datenpunkte aus einem MA(1)-Prozess mit $\theta_1 = 0.1983$ erzeugt. Ist der berechnete Lyapunov-Exponent wieder kleiner als null, so ist der Prozess richtig klassifiziert. Andernfalls läge eine Fehlklassifikation vor.

Voruntersuchungen zum Startwert ergaben, dass mit einem Erfolgsprozentsatz von 95% bei unterschiedlich gewählten Startwerten die richtige, der Theorie entsprechenden Klassifikationsentscheidung getroffen wurde. In dieser Simulation wird damit der Startwert auf $x_0 = 2.4$ festgesetzt. Die erzeugte Zeitreihe aus dem MA(1)-Prozess ist in Abbildung 4.3 dargestellt.

Das Phasenbild von x_t auf x_{t-1} deutet auf ein „gleichförmiges“ asymptotisches Verhalten hin (vgl. Abbildung 4.4).

Durch die Berechnung des Lyapunov-Exponenten soll nun eine weniger subjektive Klassifikation vorgenommen werden.

Der Lyapunov-Exponent wurde auf $\hat{\lambda} = -0.0475$ berechnet, also kleiner als null. Da die richtige Entscheidung bei einem Wert kleiner als null läge, ist eine Aussage bzgl. einer richtigen Klassifikation anzunehmen.

Zum Schluss der als Stellvertreter der linearen Prozesse untersuchten Zeitreihe werden Datenpunkte aus einem ARMA(1,1)-Prozess erzeugt (vgl. z. B. Fahrmeir,

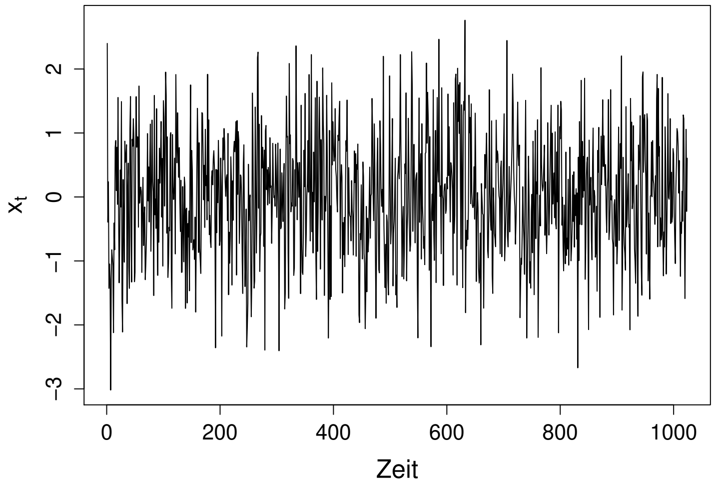

Abbildung 4.3: Reihendarstellung des durch Gleichung (4.2) erzeugten MA-Prozesses.

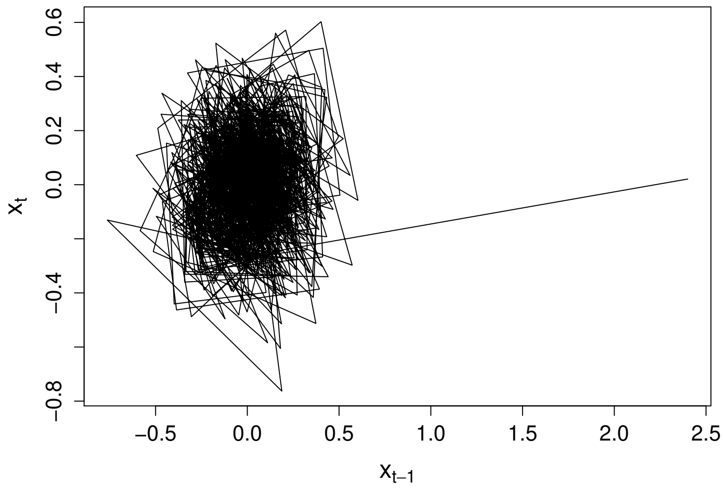

Abbildung 4.4: Phasendarstellung von x_t und x_{t-1} des durch Gleichung (4.2) erzeugten MA-Prozesses.

Kaufmann und Ost (1981), S. 215), d. h.:

$$ARMA(1,1): X_t = \phi_1 X_{t-1} + \theta_1 \epsilon_{t-1}. \tag{4.3}$$

Die Prozesszustände hängen zum einen vom Wert des vorangegangenen Zeitpunktes x_{t-1} ab und werden zusätzlich noch durch einen Fehlerterm zum ebenfalls letzten Zeitpunkt mit beeinflusst. Um theoretisch ein konvergentes Verhalten im asymptotischen Zustand zu sichern, werden die Parameter ϕ_1 und θ_1 aus den Untersuchungen des AR(1)- und des MA(1)-Prozesses übernommen. Der Parameter ϕ_1 wird wieder auf $\phi_1 = 0.9$ gesetzt, da der AR-Anteil somit sicher zur Konvergenz beiträgt. Der Parameter θ_1 wird auf $\theta_1 = 0.1983$ festgelegt, so dass insgesamt ein stabiler Attraktor klassifiziert werden sollte. Mit diesen Größen werden wieder 1024 Datenpunkte generiert, wobei der Startwert, wie bei dem MA(1)-Prozess, auf $x_0 = 2.4$ gesetzt wird. Auch hier gingen der Festsetzung Voruntersuchungen voraus, die mit unterschiedlichen Startwerten eine Erfolgsprozentzahl von 99% aufwiesen. Die erzeugte Zeitreihe aus diesem ARMA(1,1)-Prozess ist in Abbildung 4.5 wiedergegeben.

Das Phasenbild des jeweiligen und vorangegangenen Zustandes x_t und x_{t-1} weist graphisch eine mögliche Konvergenz auf (vgl. Abbildung 4.6).

Mit Hilfe der Simulationsreihe soll nun rechnerisch eine Einteilung in einen stabilen oder instabilen Prozess und somit auch in ein gut prädizierbares oder schlecht prädizierbares System durch die Berechnung des Lyapunov-Exponenten vorgenommen werden. Ist der berechnete Lyapunov-Exponent größer als null, läge eine Fehlklassifikation vor.

Das Ergebnis der Simulation jedoch ergab eine richtige Klassifikation in einen stabilen und gut prädizierbaren Prozess, da $\hat{\lambda} = -0.0331$, also $\hat{\lambda} < 0$ ist.

Als letzter stochastischer Prozess sei ein spezielles Schwingungsmodell betrachtet. Diese zeichnen sich dadurch aus, besonders gut zyklische Schwankungen

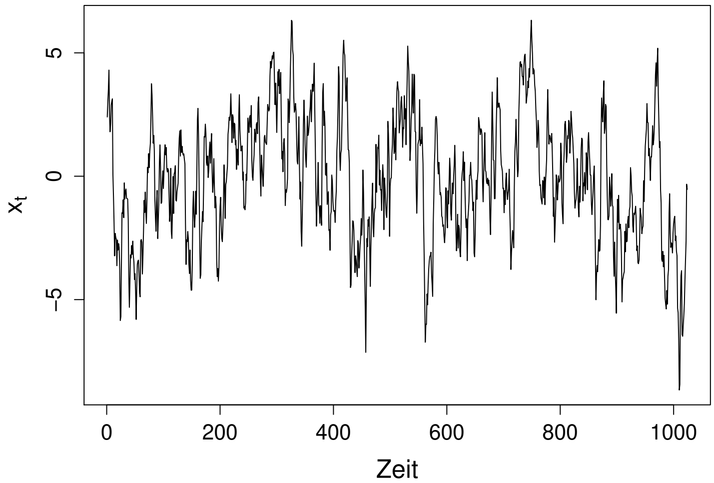

Abbildung 4.5: Reihendarstellung des durch Gleichung (4.3) erzeugten ARMA-Prozesses.

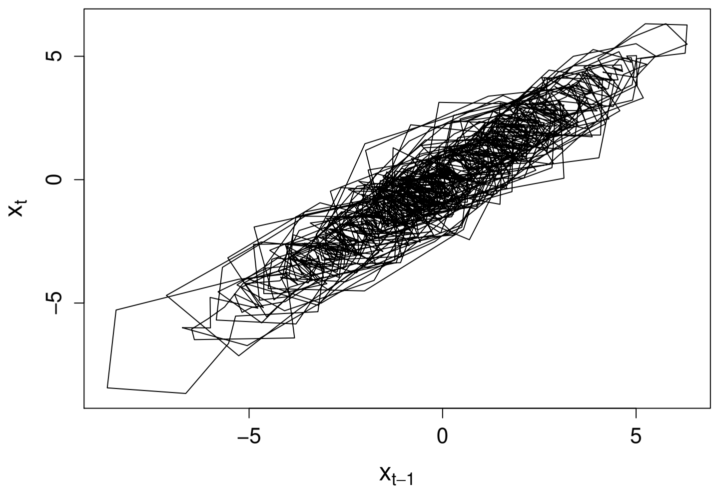

Abbildung 4.6: Phasendarstellung von x_t und x_{t-1} des durch Gleichung (4.3) erzeugten ARMA-Prozesses.

modellieren zu können. Die hier benutzte Form sei gegeben durch:

$$X_t = \alpha \sin x_{t-1} + \epsilon_t. \tag{4.4}$$

Die einzelnen Prozesszustände hängen mit einer harmonischen Schwingung zusammen, wobei eine natürliche Periode von 2π gegeben ist. Der Parameter α dient als Gewicht und kann über stabiles und instabiles Verhalten entscheiden.

Ausgehend von diesen Überlegungen wird der Parmeter auf $\alpha = 0.699$ festgelegt. Somit wird die Voraussetzung für ein stabiles asymptotisches Verhalten erfüllt.

Wie bereits in den anderen Untersuchungen werden wiederum 1024 Datenpunkte aus dem sin-Prozess erzeugt, um zu überprüfen, ob die simulierte Zeitreihe durch den berechneten Lyapunov-Exponenten richtig klassifiziert wird.

Voruntersuchungen mit 100 zufällig ausgewählten Startwerten ergaben eine richtige Klassifikationsentscheidung von 100%. Somit wird x_0 auf $x_0 = 2.4$ festgesetzt.

Die simulierten Daten sind in Abbildung 4.7 dargestellt. Es zeigt sich zuerst einmal, dass die Schwingungen durch den Gewichtsparameter in dieser Darstellung nicht sofort deutlich sichtbar sind.

Um einen Eindruck über das Verhalten von x_t zu x_{t-1} zu gewinnen, wird wiederum ein Phasenbild erstellt (vgl. Abbildung 4.8).

Es zeigt sich eine noch „regelmäßige" Struktur, obwohl eine graphische Einbettung in zwei Dimensionen keine optimale Wahl darstellt. Somit ist nur mit Vorsicht die Behauptung aufzustellen, dass schon anhand des Phasenbildes ein stabiler Prozess erkennbar ist.

Die Berechnung des Lyapunov-Exponenten hingegen bietet ein objektiveres Kriterium zur Entscheidung über einen stabilen oder instabilen Prozess. Sie ergab einen Wert kleiner als null, genauer $\hat{\lambda} = -0.0568131$, so dass die Zeitreihe richtig klassifiziert wurde.

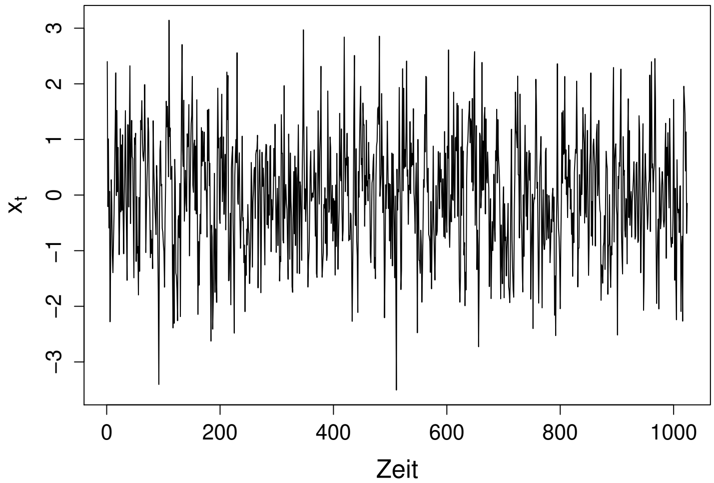

Abbildung 4.7: Reihendarstellung des durch Gleichung (2.15) erzeugten sin-Prozesses.

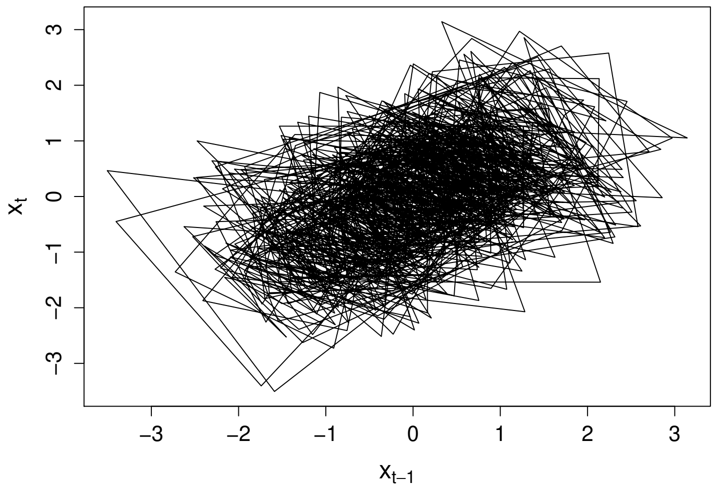

Abbildung 4.8: Phasendarstellung von x_t und x_{t-1} des durch Gleichung (2.15) erzeugten sin-Prozesses.

4.2.2 Untersuchungen mit stabilen und instabilen Prozessen

Um Aussagen über die Unterscheidung zwischen stabilen und instabilen Prozessen mit Hilfe des Lyapunov-Exonenten machen zu können, wäre es sinnvoll, eine Abbildung zu finden, die sowohl stabile als auch instabile Prozesse durch die Wahl der in der Abbildung vorhandenen Parameter erzeugen könnte. So wäre es in einem zweiten Schritt dann auch möglich, sich den kritischen Bereich, d. h. λ um null herum, von der stabilen als auch von der instabilen Parametrisierung zu nähern, um zu überprüfen, ob der Lyapunov-Exponent richtig klassifiziert. In einem dritten Schritt könnten dann ebenfalls Untersuchungen über die Länge der Datenreihen mit den gleichen Abbildungen vorgenommen werden. Das wäre wichtig, da der Lyapunov-Exponent eine Kenngröße des asymptotischen Verhaltens ist, so dass zu kurze Reihen zu Fehlklassifikationen führen könnten.

Eine solche Abbildung, die sowohl stabile als auch instabile Prozesse durch die Wahl des Parameters erzeugen kann, ist durch die Modulo-1-Abbildung, kurz mod1, gegeben. Sie bietet sich für die Untersuchungen zum einen deswegen so gut an, weil sie chaotisches Verhalten erzeugt, was mit einem klassischen instabilen Prozess gleichgesetzt wird. Zum anderen ist - zumindest im deterministischen Fall - der exakte Lyapunov-Exponent berechenbar, was für die Simulationen von Vorteil sein wird. Bei den stochastischen Zeitreihen ist die numerische Berechnung durch den Einfluss des Zufalls nicht immer exakt. Obwohl in dieser Arbeit nicht der wahre Wert des Lyapunov-Exponenten interessiert, sondern für eine Klassifikation nur seine Einteilung in kleiner oder größer als null, ist es möglich, den Grad an Instabilität durch den exakten Wert des Lyapunov-Exponenten der Modulo-Abbildung zu nutzen.

Allgemein bildet eine Modulo-Abbildung, kurz mod m, eine Zahl x auf den Rest ihrer Divisionen durch m ab. Sie läuft von den natürlichen Zahlen $\mathbb{N}$ (oder wahlweise auch von den ganzen Zahlen $\mathcal{Z}$) auf die Menge $\{0, \ldots, m-1\}$. Hierbei gilt (Padberg, 1996):

$$x \to i \qquad \text{mit } x \geq rm, \qquad \text{wobei } r \in N_0 \qquad \text{und } x < (r+1)m.$$

Formal ist die Modulo-1-Abbildung die Abbildung, die nur die Nachkommastellen zurück gibt, wenn der Wertebereich zwischen null und eins liegt.

Für die Simulationen werden Iterationen der Modulo-1-Abbildung betrachtet, d. h. allgemein:

$$x_t = (\lambda^* \cdot x_{t-1}) \mod (1). \tag{4.5}$$

Für diese Datenreihe ist der Lyapunov-Exponent λ direkt ausrechenbar, da der funktionale Zusammenhang bekannt ist. So ist $\lambda = \ln \lambda^*$. In Anhang A.3 ist eine spezielle Modulo-1-Abbildung, die sogenannte Bernoulli-Verschiebung, beschrieben, in der auch die Herleitung des Lyapunov-Exponenten für die Iterationen dieser Abbildung dargestellt ist.

Bei den stochastischen Zeitreihen wird eine kleine Störung zu der Abbildung hinzu addiert. Somit wird die Gleichung (4.5) modifiziert in:

$$x_t = (\lambda^* \cdot x_{t-1} + z \cdot \epsilon_t) \mod (1), \tag{4.6}$$

mit $\epsilon \sim Vtlg[\mu, \sigma^2]$, z=Gewichtungsparameter und $\lambda = ln\lambda^*$.

Dieses widerspricht nicht dem Simulationsziel, da der einzustellende Parameter der Modulo-1-Abbildung weiterhin stabile und instabile Prozesse bestimmt. An der Eigenschaft des Attraktors wird nichts verändert, wenn eine kleine Störung hinzu addiert wird (vgl. Kapitel 3.2.1). Was passieren kann, ist, dass der exakte Wert des Lyapunov-Exponenten nicht mehr übereinstimmt mit den Ergebnissen in den Simulationen. Jedoch gilt das Interesse in dieser Arbeit wegen der durchzuführenden Klassifikation auch nur dem Vorzeichen.

Nun soll anhand von simulierten Datenreihen aus einer gegebenen Modulo-1-Abbildung geprüft werden, ob der berechnete Lyapunov-Exponent die zu Grunde liegenden Prozesse richtig in stabile und instabile, also in gut prädizierbare und nicht gut prädizierbare Systeme klassifiziert.

Hierzu wird das gut prädizierbare System durch eine Modulo-1-Abbildung mit den Parametern $\lambda^* = 0.9$, $z = 0.05$ und ϵ gleichverteilt auf dem Intervall $[0, 1]$ simuliert, d. h.:

$$x_t = (0.9x_{t-1} + 0.05\epsilon_t) \mod 1, \qquad \text{mit } \epsilon_t \sim U[0, 1]. \tag{4.7}$$

Die schlecht prädizierbaren Beobachtungsreihen werden ebenfalls durch die Modulo-1-Abbildung erzeugt. Jedoch kann hier, wie oben beschrieben, auf den stochastischen Einfluss durch einen kleinen hinzu addierten Fehler verzichtet werden. Der einzustellende Parameter ist auf $\lambda^* = 2.5$ festgesetzt, d. h.:

$$x_t = (2.5x_{t-1}) \mod 1. \tag{4.8}$$

In beiden Fällen werden die Startwerte auf $x_0 = 0.699$ gesetzt (vgl. Kapitel 4.2.3) und durch Iteration der funktionalen Zusammenhänge Datenreihen der Länge 1024 erzeugt.

Hierbei ist die Wahl der frei wählbaren Parameter aus folgenden Gründen gesetzt: Für einen stabilen Prozess - im Sinne des Lyapunov-Exponenten λ, der kleiner als null sein soll - muss der Parameter λ^* kleiner als eins gewählt werden, da $\ln \lambda^* = \lambda$ ist. Mit einem $\lambda^* = 0.9$ zeigten sich schon in den Untersuchungen des AR(1)-Prozesses stabile Ergebnisse. Hingegen wird ein instabiler Prozess - generiert aus der Modulo-1-Abbildung - mit einem $\lambda^* = 2.5$, somit also größer als eins, sicher erzeugt. Die Wahl des Gewichtungsparameters z und der Gleichverteilung als Verteilung des hinzu addierten Fehlers sind aus Voruntersuchungen hervorgegangen (vgl. Anhang A.4).

Zur Veranschaulichung seien die Trajektorien der generierten Datenreihen in den Abbildungen 4.9 und 4.10 dargestellt.

In den Abbildungen sind keine Zyklen und Trends zu erkennen. Ebenfalls ist es nicht möglich, durch einfaches Hinschauen eine Unterscheidung zwischen den

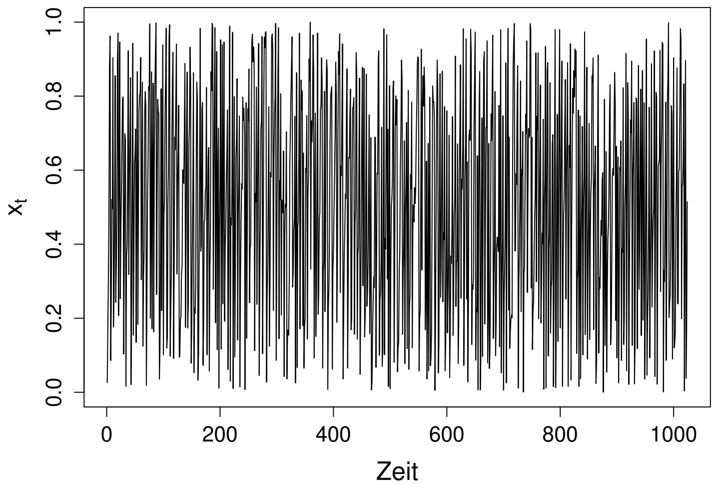

Abbildung 4.9: Die Modulo-1-Abbildung aus Gleichung (4.7) erzeugt eine stochastische Zeitreihe, ein Beispiel für eine gut prädizierbare, stabile Datenreihe.

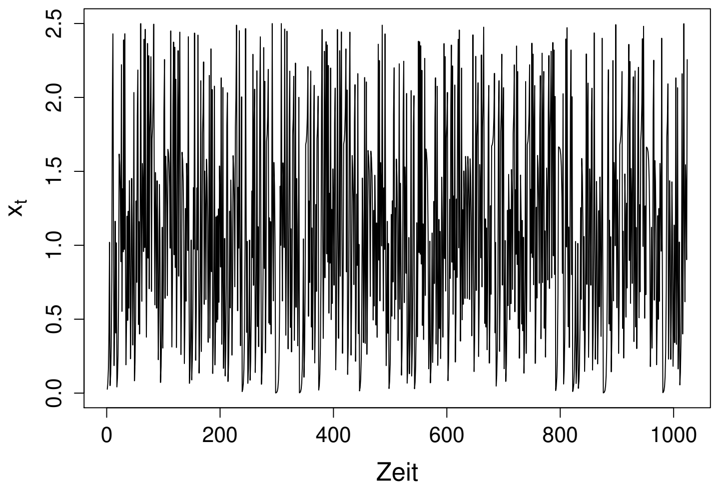

Abbildung 4.10: Die Modulo-1-Abbildung aus Gleichung (4.8) erzeugt eine chaotische Beobachtungsreihe, ein Beispiel für eine nicht gut prädizierbare, instabile Datenreihe.

gut prädizierbaren und nicht gut prädizierbaren Prozessen zu treffen. So wird der Lyapunov-Exponent numerisch berechnet bzw. geschätzt, um eine Klassifikation vorzunehmen.

Für den gut prädizierbaren, stabilen Prozess wurde $\hat{\lambda} = -0.0945$ geschätzt, wobei der wahre Lyapunov-Exponent bei $\lambda = -0.11$ liegt. Da in dieser Arbeit für die Klassifikation der exakte Wert des geschätzten Lyapunov-Exponenten keine Beachtung findet, ist nur wichtig festzuhalten, dass $\hat{\lambda} < 0$ ist und somit der Prozess richtig klassifiziert wurde.

Die Schätzung des Lyapunov-Exponenten führte in dem nicht gut prädizierbaren, instabilen, chaotischen Fall zu $\hat{\lambda} = 0.92$, wobei der wahre Lyapunov-Exponent mit $\lambda = 0.92$ exakt getroffen wurde. Jedoch auch hier ist nur wichtig festzuhalten, dass $\hat{\lambda} > 0$ berechnet wurde, somit auch bei dieser Simulation keine Fehlklassifikation auftrat.

Im Gegensatz zur reinen Betrachtung der Trajektorien der gegebenen Datenreihen ist es möglich, mit Hilfe der berechneten Lyapunov-Exponenten die simulierten Prozesse richtig zu klassifizieren.

4.2.3 Untersuchungen zum Startwert

Die Theorie zu stabilen und instabilen, hier also chaotischen, Prozessen besagt, dass die stabilen Systeme unempfindlich gegenüber dem Startwert sind, wohingegen sich eine Abhängigkeit von den Anfangswerten bei chaotischen, instabilen Prozessen herausgestellt hat (Jetschke, 1989). Jedoch ist der Lyapunov-Exponent eine Kenngröße des Attraktors, der diese Abhängigkeit mit berücksichtigt. Somit dürften bei den Berechnungen des Lyapunov-Exponenten keine Startwertprobleme auftauchen.

Um diese Frage zu klären, werden die beiden Modulo-1-Abbildungen aus Kapitel 4.2.2 verwendet, um mit unterschiedlichen Startwerten Datenreihen zu erzeugen. Für diese werden die Lyapunov-Exponenten berechnet und auf Fehlklas-

Tabelle 4.1: Klassifikationsergebnisse des stabilen und instabilen Prozesses in Abhängigkeit der Startwerte

Prozess	Startwert x_0	berechn. Lyap-Exp.	Klass.-entscheidung
stabiler Bereich	0.299	-0.09003	richtig
	0.699	-0.09450	richtig
	0.899	-0.11409	richtig
instabiler Bereich	0.299	0.9106	richtig
	0.699	0.9200	richtig
	0.899	0.9084	richtig

sifikationen untersucht. Für den stabilen Prozess werden mit Hilfe der Gleichung (4.7) und für den instabilen Prozess mit Hilfe der Gleichung (4.8) jeweils 1024 Datenpunkte mit unterschiedlichen Startwerten erzeugt. Da keine genauen Informationen über sinnvolle Startwerte zur Verfügung stehen, jedoch der Wertebereich der stabilen Abbildung (mit diesen Parametereinstellungen) nicht größer als $|1|$ wird, werden in einer ersten Voruntersuchung 100 Startwerte zufällig aus dem Intervall $(0, 1)$ gezogen. Hier zeigten sich bis auf natürliche Schwankungen keine Fehlklassifikationen. In einer zweiten Untersuchung werden aus diesem Intervall nochmals in 0.1-er Schritten, angefangen bei 0.099 und endend bei 0.999, zur möglichen Erkennung von „Struktur“ Simulationen für die stabile und instabile Form durchgeführt. Tabelle 4.1 zeigt exemplarisch unterschiedliche Startwerte, die berechneten Lyapunov-Exponenten und die richtigen und falschen Klassifikationen.

Es zeigt sich, dass alle Klassifikationen richtig waren und keine Struktur zu erkennen ist. Die berechneten Lyapunov-Exponenten schwanken lediglich auf natürliche Weise in ihren Werten. Da in dieser Arbeit jedoch nur die Unterscheidung zwischen stabilen und instabilen Prozessen und nicht das Maß an Chaos oder an Konvergenz von Interesse ist, bleibt für weitere Untersuchungen festzuhalten, dass das Startwertproblem hier keine große Rolle spielt.

4.2.4 Untersuchungen im kritischen Bereich

Nachdem in Kapitel 4.2.2 exemplarisch sowohl ein stabiler als auch ein instabiler Prozess mit Hilfe der Modulo-1-Abbildung simuliert und durch den Lyapunov-Exponenten richtig klassifiziert wurde, stellt sich die Frage, wie gut die Klassifikationen im kritischen Bereich funktionieren.

Zur Erinnerung sei noch einmal kurz erläutert, dass der kritische Bereich bei der Klassifikation mit dem Lyapunov-Exponenten um null herum liegt, da diese den Trennwert zwischen stabilen und instabilen Prozessen darstellt.

Um nun die Frage nach möglichen Fehlklassifikationen im kritischen Bereich beantworten zu können, werden die Parameter der Modulo-1-Abbildung in aufeinander folgenden Simulationen so variiert, dass sie sich von oben und unten, also vom instabilen und stabilen Zustand annähern. Für die Parameter der Modulo-1-Abbildung λ^* bedeutet das wegen des Zusammenhangs zum Lyapunov-Exponenten λ, d. h. $\ln \lambda^* = \lambda$, dass λ^* für stabile Prozesse immer näher von unten an eins, für instabile Prozesse immer näher von oben an eins gewählt wird.

Um keine Einstellungsänderungen zu den in Kapitel 4.2.3 zu den Startwerten zu den in Kapitel 4.2.2 gefundenen Ergebnissen in den nun folgenden Simulationen vorzunehmen, werden die Startwerte auf $x_0 = 0.699$ gesetzt. Da die Voruntersuchungen in Kapitel 4.2.2 und dem Anhang A.4 zeigten, dass eine gute Wahl des Gewichtungsparameters bei dem hinzu addierten Fehler im stochastischen Fall bei $z = 0.05$ liegt und die Gleichverteilung auf $[0, 1]$ keinen großen Einfluss auf die Berechnungen hat, werden diese beiden Einstellungen ebenfalls übernommen.

In jeder Simulation werden 1024 Datenpunkte mit jeweils den oben beschrie-

Tabelle 4.2: Klassifikationsergebnisse des stabilen und instabilen Prozesses im kritischen Bereich

λ^*	berechn. Lyap-Exp.	Klass.-entscheidung
0.9	-0.0945	richtig
0.95	-0.0907	richtig
0.97	-0.0939	richtig
0.99	-0.0676	richtig
0.99999	-0.0620	richtig
1.01	0.009	richtig
1.03	0.026	richtig
1.05	0.024	richtig
1.1	0.092	richtig

benen Parametereinstellungen und unterschiedlichem Stabilitätsparameter λ^* generiert. Tabelle 4.2 zeigt die unterschiedlichen λ^*, die berechneten Lyapunov-Exponenten und die richtigen und falschen Entscheidungen bzgl. der Frage, ob der Prozess stabil oder instabil, d. h. gut oder nicht gut prädizierbar ist.

Es zeigt sich, dass keine Fehlklassifikation auftritt. Wenn der kritische Bereich den Übergang vom stabilen zum instabilen ausdrückt, kann festgehalten werden, dass auch wenn sich eine Parametereinstellung zum stabilen Prozess den instabilen Parametereinstellungen nähert, trotzdem nur richtige Klassifikationen auftreten. Ebenso gilt das auch bei den Parametern zu instabilen, hier chaotischen Prozessen. Auch wenn das System immer weniger im Sinne des Lyapunov-Exponenten chaotisch ist, treten auch hier keine Fehlklassifikationen auf.

4.2.5 Untersuchungen mit kurzen Datenreihen

Obwohl der Lyapunov-Exponent eine Größe des asymptotischen Verhaltens eines Prozesses darstellt, muss diese aus der gegebenen Datenreihe geschätzt werden. Hierbei sind Fehlklassifikationen möglich, da sich die Reihen noch im transienten Zustand befinden könnten. Um also den Einfluss unterschiedlicher Datenreihenlängen zu untersuchen, werden wiederum die Modulo-1-Abbildung aus Kapitel 4.2.2 verwendet, um die Unterscheidung zwischen stabilen und instabilen Prozessen in Abhängigkeit der Längen zu verifizieren. Die Berechnung der Lyapunov-Exponenten kann schlechte Schätzer liefern und somit zu Fehlklassifikationen führen, wenn die Datenreihen zu kurz sind. In diesem Falle lägen die Datenreihen noch im transienten Zustand vor. Für kurze Reihen bleibt also zu entscheiden, ob eine Klassifikation überhaupt möglich ist.

Zur Klärung dieser Frage werden Zeitreihen mit unterschiedlichen Längen für den stabilen Prozess aus der schon bereits verwendeten Modulo-1-Abbildung (vgl. Gleichung (4.7)) erzeugt. Die Beobachtungsreihen unterschiedlicher Länge für den instabilen Bereich werden aus der ebenfalls schon verwendeten Modulo-1-Abbildung (vgl. Gleichung (4.8)) gewonnen. Die Startwerte werden wiederum auf $x_0 = 0.699$ festgesetzt, um keine weiteren frei wählbaren Parameter umzustellen.

Die Klassifikation in stabile, gut prädizierbare und instabile, schlecht prädizierbare Prozesse mit Hilfe des Lyapunov-Exponenten werden auf die Datenreihenlängen $10, 20, 30, 40, 50, 100, 150, 200, \ldots, 900$ angewendet. Abbildung 4.11 zeigt die berechneten Lyapunov-Exponenten in Abhängigkeit der Reihenlängen.

Die gestrichelte Linie zeigt die durch den Lyapunov-Exponenten bestimmte Grenze zwischen den stabilen und instabilen Systemen. Schätzungen, die oberhalb dieser Linie liegen, führen in diesem Fall zu Fehlklassifikationen, wohingegen die unterhalb Liegenden richtig klassifizieren. Die durchgezogene Linie bezeichnet den hier wahren Lyapunov-Exponenten, um einen Referenzbezug zu haben.

Es zeigt sich, dass jede Zeitreihenlänge zu einer korrekten Klassifikation führt,

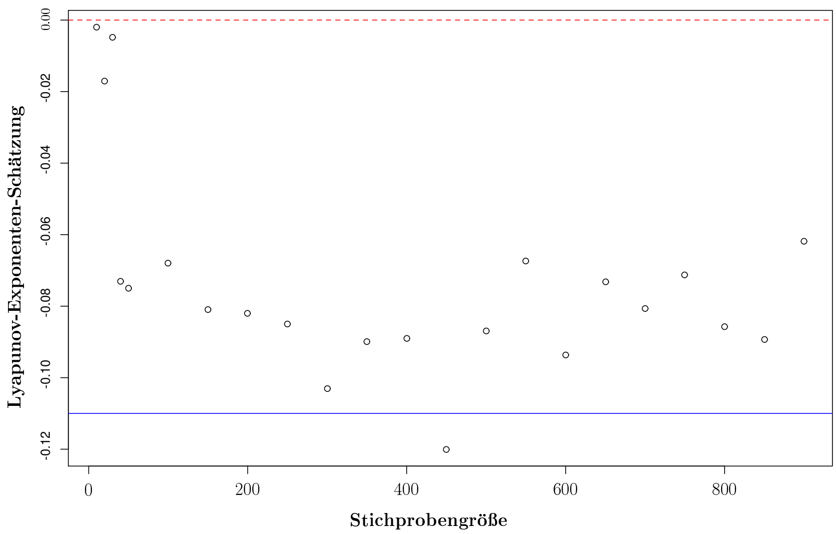

Abbildung 4.11: Die berechneten Lyapunov-Exponenten für die aus der Gleichung (4.7) generierten Zeitreihen in Abhängigkeit der Reihenlängen

auch wenn die Werte der berechneten Lyapunov-Exponenten vom wahren Lyapunov-Exponenten differieren. Jedoch sind die Längen von $10, 20$ und 30 zu klein für eine ernstzunehmende Klassifikation, da die Berechnungen wegen numerischer Schwierigkeiten nur unzureichend ausgeführt wurden.

Für den instabilen Prozess werden nun die gleichen Längen angewendet, wie in Abbildung 4.12 dargestellt.

Die gestrichelte Linie beschreibt wiederum die Klassifikationslinie. In diesem Fall weisen jedoch die Schätzungen, die unterhalb dieser Grenze liegen, auf die Fehlklassifikationen hin. Die oberhalb Liegenden zeigen die richtigen Entscheidung an. Die durchgezogene Linie indiziert den wahren Lyapunov-Exponenten des zu Grunde liegenden instabilen Systems.

Zuerst einmal bleibt festzuhalten, dass auch in diesem Fall die Beobachtungsreihenlängen von $10, 20$ und 30 schon bei den Berechnungen Schwierigkeiten aufwiesen, so dass sie für eine Klassifikation nicht genommen werden sollten. Es

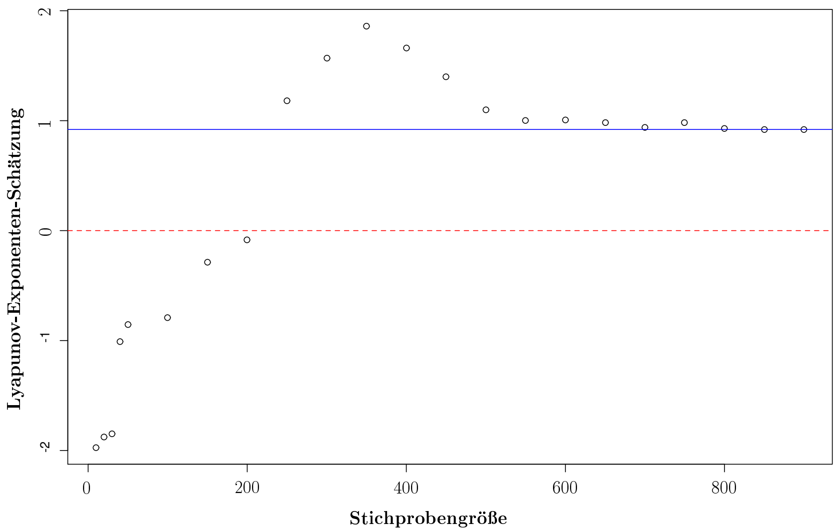

Abbildung 4.12: Die berechneten Lyapunov-Exponenten für die aus der Gleichung (4.8) generierten Beobachtungsreihen in Abhängigkeit der Reihenlängen

zeigt sich in Abbildung 4.12, dass kurze Beobachtungsreihen zu Fehlklassifikationen führen. Die Beobachtungsreihenlängen von 10 bis 200 liefern keine korrekten Klassifikationen. Jedoch Datenreihen mit den Längen zwischen 250 und 500 charakterisieren die Prozesse richtig. Gute bis hin zu fast exakten Schätzern werden durch Reihen der Länge 500 und mehr generiert.

Insgesamt zeigt sich, dass das aus der Theorie erwartete Ergebnis, zu kurze Datenreihen können zu Fehlklassifikationen führen, hier bestätigt wird.

Kapitel 5
Surrogate

Um empirische Aussagen über den Lyapunov-Exponenten einer Datenreihe machen zu können, ist es notwendig, eine Stichprobe ausgehend von der gegebenen Beobachtungs- bzw. Zeitreihe zu erzeugen. Wäre der funktionale Zusammenhang des zu Grunde liegenden Systems bekannt, könnten durch einfaches Ausführen der Prozessgleichung weitere Datenreihen generiert werden. Jedoch wird in dieser Arbeit nur der Fall betrachtet, in dem kein Wissen über das Zusammenwirken der einzelnen Beobachtungen bekannt ist. Somit muss ein Verfahren benutzt werden, das eine Stichprobe ohne das Wissen über den funktionalen Zusammenhang erzeugt (vgl. Kapitel 5.1). Um das Verfahren experimentell zu unterstützen, werden Untersuchungen mit den bereits verwendeten stochastischen Prozessen und den stabilen und instabilen Formen der Modulo-1-Abbildung durchgeführt (vgl. Kapitel 5.2).

5.1 Erzeugung von Surrogaten

Instabile von stabilen Prozessen zu trennen, ist eine nützliche Aufgabe zur Einschränkung der möglichen Modellklassen in der Datenreihenanalyse. Wie in Kapitel 4 gezeigt, ist es möglich, nur an der gegebenen Datenreihe mit Hilfe des Lyapunov-Exponenten den zu Grunde liegenden Prozess zu klassifizieren.

Um jedoch empirische Eigenschaften des berechneten Lyapunov-Exponenten bestimmen zu können und einen Test über die möglichen Fehlklassifikationen zu erzeugen, muss eine Stichprobe der Zeit- bzw. Beobachtungsreihe, hier als Surrogate bezeichnet, erzeugt werden. Hierbei ist wieder die Nebenbedingung, dass der funktionale Zusammenhang des Systems bzw. der wohldefinierte Prozess nicht bekannt sind. Somit ist es notwendig, ein Verfahren zu verwenden, das unter gemachten Annahmen eine Stichprobe der gegebenen Datenreihe erzeugt.

In der Literatur werden in Verbindung mit Hypothesen, die einem statistischen Test zu Grunde liegen, Surrogatsätze erzeugt. Würde z. B. die Aufgabe sein zu testen, ob ein bestimmter Datensatz aus einem spezifizierten ARMA-Prozess stamme oder nicht, so würden zur Erzeugung einer Stichprobe aus diesem ARMA-Modell künstlich mehrere Zeitreihen generiert. Diese Zeitreihen können dann als Surrogate verwendet werden und die gewünschten empirischen Eigenschaften berechnet werden (Theiler, Galdrikian, Longtin, Eubank und Farmer, 1992). Schmitz und Schreiber (2000) versuchen, für nichtstationäre Prozesse Surrogate so zu erzeugen, indem sie Originaldatenreihen geschickt splitten, für jeden Teil einzeln Surrogate erzeugen, um sie anschließend wieder zusammen zu setzen. Für Daten aus linearen, korrelierten Gaußprozessen schlagen Prichard und Theiler (1994) einen modifizierten Algorithmus vor, der nicht, wie üblich, auf die Autokorrelationsfunktion, sondern ebenfalls auf die Kreuzkorrelation eingeht.

Als Ausgangssituation zur Erzeugung der Surrogate in dieser Arbeit dient das erstmals von Theiler, Galdrikian, Longtin, Eubank und Farmer (1992) bzw. Theiler, Eubank, Longtin, Galdrikian und Farmer (1992) vorgeschlagene *amplitude adjusted Fourier transform* (AAFT)-Verfahren, das später noch näher beschrieben wird (vgl. Anhang B). Erweitert wurde dieser Algorithmus von Schreiber und Schmitz (1996). Diese Version zur Erzeugung von Surrogaten ist im Programmpaket *TISEAN* (Hegger, Kantz und Schreiber, 1999) implementiert.

Um nun empirische Aussagen des zu berechnenden Lyapunov-Exponenten machen zu können, sollen Surrogate geschaffen werden. Die erzeugten Surrogate sollten ein bestmögliches Abbild der Originaldatenreihe sein. Da jedoch der wahre funktionale Zusammenhang des Prozesses unbekannt ist, können nicht einfach weitere Datenreihen aus dem zu Grunde liegenden System geschaffen werden. Die Surrogate müssen aus der gegebenen Zeit- bzw. Beobachtungsreihe erzeugt werden. Ohne weitere Annahmen ist das natürlich nicht möglich, so dass Kriterien festgesetzt werden müssen. Im Hinblick auf die Untersuchung bezüglich der Klassifikation zwischen stabilen, prädizierbaren und instabilen, schlecht prädizierbaren Prozessen ist es sinnvoll, die Kriterien an diesem Ziel zu orientieren. Da in Kapitel 3.2 vorgeschlagen wird, eine als stabilen Prozess klassifizierte Zeitreihe z. B. zuerst einmal mit einem gewöhnlichen linearen stochastischen Ansatz zu modellieren, ist eine Möglichkeit, für die Erzeugung der Surrogate die Autokorrelationsfunktion als Kriterium heranzuziehen. Die Eigenschaft der Autokorrelation wird bei durchgeführter Fouriertransformation auf Grund des Wiener-Khintchin-Wold Theorems (vgl. Schlittgen und Streitberg (1994), S. 161) beibehalten. Für die Erzeugung der Surrogate wird somit kein spezieller Prozess angenommen, jedoch wird gefordert, dass die Surrogate und die Originaldatenreihe die gleiche Autokorrelationsfunktion besitzen.

Der von Theiler, Eubank, Longtin, Galdrikian und Farmer (1992) vorgeschlagene AAFT-Algorithmus könnte für die Erzeugung der Surrogate genutzt werden. Schreiber und Schmitz (1996) zeigten, dass asymptotisch gute Ergebnisse mit diesem Verfahren erzielt werden. Jedoch für endliche Reihenlängen haben die Surrogate und die Originaldatenreihen zwar die gleiche Autokorrelationsfunktion, aber möglicherweise nicht die gleiche Verteilungsfunktion. Ein Grund hierfür liegt in der Phasenrandomisierung, die in der AAFT durchgeführt wird, so dass nur im Mittel die gleiche Verteilung erzeugt wird. Ein anderer Grund ist in den für das Verfahren notwendigen Transformationsfunktionen zu finden, die sowohl eine benötigte Funktion als auch ihre Inverse beschreiben sollten. Jedoch in der

Praxis sind diese nicht immer ganz exakt, so dass dieses zu kleineren Fehlern führen könnte (Schreiber und Schmitz, 1996).

Somit wird für die Surrogate gefordert, dass sie nicht nur die gleiche Autokorrelationsfunktion wie die Originaldatenreihe aufweisen, sondern zusätzlich auch der gleichen Verteilung folgen. Hierzu bieten Schreiber und Schmitz (1996) eine alternative Methode zur Erzeugung der Surrogate an, die in dieser Arbeit ihre Anwendung findet.

Der Algorithmus basiert auf einem zweigeteilten Iterationsschema. In einem ersten Schritt wird die Datenreihe so bearbeitet, dass die entstehenden Surrogate die gleiche Autokorrelationsfunktion aufweisen. In dem zweiten Schritt wird die Anpassung der Verteilung der Surrogate an die der Originaldatenreihe vorgenommen. Da jedoch durch den zweiten Schritt unglücklicherweise die Übereinstimmung der Autokorrelationsfunktion aufgehoben sein könnte, werden die beiden Schritte wiederholt, bis eine vorgegebene Genauigkeit erreicht ist. Somit wird zwar auch für endliche Datenreihenlängen keine exakte Konvergenz im strengen Sinne erreicht, aber zumindest eine bestimmbare Genauigkeit (Schreiber und Schmitz (1996), Schreiber und Schmitz (2000)). Im Folgenden werden zuerst die beiden Schritte einzeln erklärt, um dann das Zusammenfügen in den Iterationsstufen zu beschreiben.

Um das Kriterium der gleichen Autokorrelationsfunktion der Originaldatenreihe und der Surrogate zu erfüllen, werden zuerst die Werte der Originaldatenreihe so zurückskaliert, dass sie einem linearen stochastischen Prozess folgen. Hierzu weisen Theiler, Galdrikian, Longtin, Eubank und Farmer (1992) darauf hin, dass die Annahme getroffen werden muss, die beobachtbare Zeitreihe stamme mindestens aus einer nichtlinearen, monotonen Transformation eines linearen, stochastischen Prozesses. Danach wird mit der zurückskalierten Reihe eine Fouriertransformation durchgeführt, wobei die Amplitude erhalten bleibt, jedoch die Phase zufällig variiert wird. So wird erreicht, dass die Surrogate die gleiche Autokorrelationsfunktion beibehalten, jedoch durch eine kleine Zufallsänderung neue

Datenreihen entstehen können. Diese so entstandenen Surrogate müssen dann noch einmal zurückskaliert werden, damit sie auch Surrogate der Originaldatenreihe und nicht der in einem linearen stochastischen Prozess skalierten Reihe werden (Theiler, Galdrikian, Longtin, Eubank und Farmer (1992), S. 183).

Genauer beschrieben bedeutet das Folgendes: Zuerst muss die gegebene Datenreihe in einem Array abgespeichert und hierzu Ränge vergeben werden. Der nächste Schritt ist, einen linearen, stochastischen Prozess zu erzeugen, in dem jedes Element unabhängig durch einen normalverteilten Zufallsgenerator generiert wird. Danach wird diese Reihe ebenfalls so sortiert, dass die Ränge der beiden Zeitreihen übereinstimmen. Das ist genau dann gegeben, wenn der Wert des n-ten kleinsten Elements aller Werte aus der Datenreihe zum Zeitpunkt t ist, so muss dies auch für den Wert zum Zeitpunkt t der anderen Reihe gelten. So ist gewährleistet, dass die in Reihenfolge gebrachte künstlich erzeugte Datenreihe der gegebenen Originaldatenreihe „folgt“ und eine Amplitude eines linearen, stochastischen Prozesses besitzt. Nun werden hieraus die Surrogate erzeugt, indem eine Fouriertransformation so durchgeführt wird, dass die Datenreihe in einer komplexwertigen Darstellung erscheint. Das bedeutet, sie kann in Amplitude und Phase aufgesplittet werden. Danach wird die Phase randomisiert, indem zu jeder komplexen Amplitude $e^{i\phi}$ multipliziert wird, wobei ϕ jedes Mal unabhängig aus dem Intervall $[0, 2\pi]$ gezogen wird. Anschließend werden die Phasen symmetrisch gemacht, damit die nun folgende inverse Fouriertransformation wieder real ist. Diese inverse Fouriertransformation bildet ein Surrogat. Damit diese nun ein Surrogat der Originaldatenreihe wird, müssen wieder die Reihenfolgen der Surrogate und der Originaldatenreihe übereinstimmen. Das wird in dem letzten Schritt durchgeführt (Theiler, Galdrikian, Longtin, Eubank und Farmer (1992), S. 183 und Theiler, Eubank, Longtin, Galdrikian und Farmer (1992), S. 84-85).

Nachdem nun die Surrogate mit gleicher Autokorrelationsfunktion geschaffen wurden, soll das Kriterium der gleichen Verteilung in einem zweiten großen Schritt erfüllt werden. Hierzu wird das Surrogat in der Form in Ränge sortiert, dass sie die Werte der Originalreihenfolge exakt annehmen (Schreiber und Schmitz (1996), S. 636). Eine genaue Beschreibung des Algorithmus findet sich im Anhang B.

Dieser Schritt jedoch verschiebt das Spektrum wieder, so dass sowohl der erste als auch der zweite Schritt so oft wiederholt werden, bis das am Ende entstehende Spektrum eine vorgegebene Genauigkeit erreicht (Schreiber und Schmitz (2000), Kapitel 4.3).

5.2 Experimentelle Ergebnisse

Um das Verfahren experimentell zu unterstützen, ist es sinnvoll, für Prozesse, dessen zu Grunde liegender funktionaler Zusammenhang bekannt ist, Stichproben zu den gegebenen Datenreihen zu erzeugen und sie dann daraufhin zu prüfen, ob sie die Originaldatenreihe wirklich repräsentieren.

Als erste Versuche bieten sich die bereits in Kapitel 4.2.1 verwendeten stochastischen Prozesse an. Sie erfüllen die Kriterien, die für die Erzeugung der Surrogate festgelegt werden (vgl. Kapitel 5.2.1). In weiteren Simulationen sollen wiederum die stabilen und instabilen Formen der Modulo-1-Abbildung Anwendung finden, um auch hier das numerische Verfahren experimentell zu stützen.

5.2.1 Untersuchungen mit stochastischen Prozessen

Die aus der Statistik bekannten und in Kapitel 4.2.1 bereits für Simulationen verwendeten linearen Pozesse sollen hier ebenfalls zur Anwendung kommen. Die AR-, MA- und ARMA-Prozesse sind linear und besitzen als „Fehlerquelle“ jeweils gewöhnliches Weißes Rauschen bzw. sind direkt dadurch gekennzeichnet. Für solche Prozesse sollte das numerische Verfahren zur Erzeugung der Surrogaten sehr gut funktionieren, da der funktionale Zusammenhang alle Voraussetzungen erfüllt.

Um dieses auch in Simulationen experimentell zu unterstützen, werden jeweils aus einem AR(1), einem MA(1) und einem ARMA(1,1)-Prozess die Originaldatenreihen erzeugt. Diese dienen dann als Grundlage zur Erzeugung der Stichproben. Das im statistischen Programmpaket *R* (vgl. Ihaka und Gentleman (1996)) implementierte Verfahren *arima*, das auf Grundlage von Gardner, Harvey und Phillips (1980) eine Anpassung der Parameter an gegebene AR, MA, ARMA oder ARIMA-Zeitreihen vornimmt, dient hierzu als Prüfgröße.

Für jedes berechnete Surrogat wird mit Hilfe der *R*-Funktion eine Parameterschätzung vorgenommen. Die geschätzten Parameter werden dann mit denen in der simulierten Zeitreihe verglichen. Stimmen diese - mit einer zugelassenen numerischen Ungenauigkeit - überein, so bedeutet es, dass die berechneten Surrogate gute Repräsentanten der Originaldatenreihe darstellen.

Angefangen wird mit einem AR(1)-Prozess. Es werden, um mögliche Vergleichbarkeiten zu erhalten, wieder die Einstellungen aus Kapitel 4.2.1 verwendet. Somit werden 1024 Datenpunkte aus einem AR(1)-Prozess mit $\phi = 0.9$ erzeugt. Als Startwert dient $x_0 = 0.699$. Von dieser simulierten Ausgangsreihe werden dann mit Hilfe des in Kapitel 5.1 beschriebenen numerischen Verfahrens 20 Surrogate erzeugt. Tabelle 5.1 zeigt die geschätzten Parameter der berechneten Surrogate, den zugehörigen (in diesem Fall immer gleichen) Parameterwert der simulierten Originaldatenreihe und die Einteilung in gut oder schlechte Repräsentation der Ausgangsreihe. Hierbei sei als Kriterium eine Abweichung des Parameters aus der Originalzeitreihe und aus dem Surrogat kleiner als 0.025.

Es zeigt sich, dass kein schlechter Repräsentant für die Originaldatenreihe auftritt. Die durch das numerische Verfahren erzeugten Surrogate könnten also gut für weitere empirische Untersuchungen verwendet werden. Jedoch ist auffällig, dass alle geschätzten Parameter kleiner als 0.9 sind.

Tabelle 5.1: Repräsentationsergebnisse der erzeugten Surrogate bzgl. der Originalzeitreihe eines AR(1)-Prozesses

Surrogat	ϕ_1 aus Original	$\hat{\phi}_1$ aus Surrogat	Repräsentation
1	0.9	0.8805	gut
2	0.9	0.8808	gut
3	0.9	0.8816	gut
4	0.9	0.8824	gut
5	0.9	0.8812	gut
6	0.9	0.8802	gut
7	0.9	0.8834	gut
8	0.9	0.8801	gut
9	0.9	0.8808	gut
10	0.9	0.8810	gut
11	0.9	0.8805	gut
12	0.9	0.8817	gut
13	0.9	0.8808	gut
14	0.9	0.8878	gut
15	0.9	0.8808	gut
16	0.9	0.8811	gut
17	0.9	0.8818	gut
18	0.9	0.8805	gut
19	0.9	0.8807	gut
20	0.9	0.8808	gut

Der nächste betrachtete Prozess aus der linearen Zeitreihenanalyse ist ein MA(1)-Prozess. Auch wenn im Hinblick auf eine Modellierung mit Hilfe von linearen Prozessen bei einer Klassifikation in einen stabilen Prozess es sicher nicht sinnvoll wäre, einen MA-Prozess zu verwenden, so soll er dennoch auf das numerische Verfahren zur Erzeugung der Surrogate untersucht werden. Die Eigenschaft, dass nur der vorangegangene Fehler, nicht aber der eigentliche Wert in dem funktionalen Zusammenhang steckt, könnte bei den geschätzten Parametern aus den zu erzeugenden Surrogaten zu numerischen Ungenauigkeiten führen. Jedoch bietet dieser Prozess eine Möglichkeit, einen Eindruck über das Umgehen mit „Messfehlern" im numerischen Verfahren zu gewinnen.

Aus einem MA(1)-Prozess mit $\theta_1 = 0.1983$ werden zuerst einmal 1024 Datenpunkte erzeugt, die als Originaldatenreihe dienen. Der Startwert wird auf $x_0 = 2.4$ festgesetzt (vgl. Einstellungen in Kapitel 4.2.1). Wie bereits für den AR(1)-Prozess werden nun 20 Surrogate erzeugt, die auf Repräsentationsfähigkeit überprüft werden.

In Tabelle 5.2 sind wiederum der Parameter des simulierten MA(1)-Prozesses, die geschätzten Parameter der einzelnen Surrogate und die Entscheidung, ob das Surrogat als Stellvertreter der Originalreihe angesehen werden kann. Hierbei sei das Kriterium $|\theta_1^{Orig} - \theta_1^{Surr}| < 0.025$ festgelegt.

Wie in Tabelle 5.2 zu sehen ist, sind bis auf eins alle Surrogate gute Repräsentanten der simulierten Originalzeitreihe. Die mit R geschätzten Parameter der Surrogate stimmen - bis auf numerische Ungenauigkeiten - sehr gut mit dem Parameter des zu Grunde liegenden Prozesses überein. Jedoch fällt auch hier auf, dass die geschätzten Parameter alle kleiner als 0.1983 sind.

Als letzter aus der Statistik bekannter linearer Prozess soll ein ARMA(1,1)-Prozess verwendet werden, der ebenfalls in Kapitel 4.2.1 Anwendung gefunden

Tabelle 5.2: Repräsentationsergebnisse der erzeugten Surrogate bzgl. der Originalzeitreihe eines MA(1)-Prozesses

Surrogat	θ_1 aus Original	$\hat{\theta}_1$ aus Surrogat	Repräsentation
1	0.1983	0.1902	gut
2	0.1983	0.1909	gut
3	0.1983	0.1910	gut
4	0.1983	0.1706	schlecht
5	0.1983	0.1900	gut
6	0.1983	0.1919	gut
7	0.1983	0.1916	gut
8	0.1983	0.1950	gut
9	0.1983	0.1913	gut
10	0.1983	0.1912	gut
11	0.1983	0.1920	gut
12	0.1983	0.1911	gut
13	0.1983	0.1910	gut
14	0.1983	0.1919	gut
15	0.1983	0.1909	gut
16	0.1983	0.1920	gut
17	0.1983	0.1911	gut
18	0.1983	0.1923	gut
19	0.1983	0.1912	gut
20	0.1983	0.1918	gut

hat. Wäre eine Datenreihe mit Hilfe des Lyapunov-Exponenten als stabil klassifiziert worden, so wäre ein einfacher Modellierungsversuch sicherlich der ARMA-Ansatz. So ist es durchaus interessant zu wissen, wie gut es möglich ist, eine aus diesem zu Grunde liegenden Prozess generierten Datenreihe so zu guten Stichproben zu variieren, dass weiterführende empirische Aussagen gemacht werden können. Ein spezieller ARMA(1,1)-Prozess soll hierzu als Beispiel dienen.

Als Ausgangssituation wird aus einem ARMA(1,1)-Prozess mit $\phi_1 = 0.9$, $\theta_1 = 0.1983$ und Startwert $x_0 = 2.4$ wie in Kapitel 4.2.1 eine Zeitreihe generiert. Die als Originaldatenreihe dienenden 1024 Datenpunkte werden mit Hilfe des numerischen Verfahrens so verändert, dass 20 Surrogate entstehen.

Die Ergebnisse dieser Simulation werden wieder in tabellarischer Form (vgl. Tabelle 5.3) dargestellt. Die Spalten bestehen aus den Parametern der simulierten Datenreihe, aus den geschätzten Parametern der Surrogate und aus der Entscheidung über die Repräsentationsfähigkeit der Surrogate bzgl. der Originalzeitreihe. Hierbei stellen die Surrogate genau dann einen guten Stellvertreter dar, wenn $|\phi_1^{Orig} - \phi_1^{Surr}| < 0.025$ und $|\theta_1^{Orig} - \theta_1^{Surr}| < 0.025$ ist.

Aus Tabelle 5.3 geht eindeutig hervor, dass die geschätzten Parameter der Surrogate in der Nähe der Parameter der Originalzeitreihe liegen. Bei den geschätzten Parametern des AR-Anteils sind wieder alle kleiner als 0.9, wohingegen die geschätzten Parameter des MA-Teils um den wahren Parameter streuen. Eine Möglichkeit, warum sowohl bei den AR-Simulationen, als auch bei den MA-Simulationen die Parameter immer unterschätzt wurden, hier jedoch nicht, könnte daran liegen, dass sowohl für den autoregressiven Teil, als auch für den Fehlerteil ein eigener Prozess zu Grunde liegt, somit sich dieses auf die Surrogate auswirkt. Dieses ist allerdings nicht an den Simulationen nachweisbar.

Zusammenfassend kann aber gesagt werden, dass sich die Surrogate als gute Repräsentanten herausstellen, so dass auch - wie schon bei dem AR(1)- und dem MA(1)-Prozess - weitere empirische Untersuchungen oder weiterführende Tests möglich sind.

Tabelle 5.3: Repräsentationsergebnisse der erzeugten Surrogate bzgl. der Originalzeitreihe eines ARMA(1,1)-Prozesses

Surrogat	ϕ_1	$\hat{\phi}_1$	θ_1	$\hat{\theta}_1$	Repräsentation
1	0.9	0.8888	0.1983	0.1981	gut
2	0.9	0.8862	0.1983	0.1972	gut
3	0.9	0.8865	0.1983	0.1994	gut
4	0.9	0.8897	0.1983	0.1992	gut
5	0.9	0.8893	0.1983	0.1970	gut
6	0.9	0.8861	0.1983	0.1993	gut
7	0.9	0.8869	0.1983	0.2003	gut
8	0.9	0.8861	0.1983	0.1999	gut
9	0.9	0.8861	0.1983	0.1999	gut
10	0.9	0.8873	0.1983	0.2007	gut
11	0.9	0.8862	0.1983	0.2014	gut
12	0.9	0.8859	0.1983	0.1997	gut
13	0.9	0.8864	0.1983	0.1994	gut
14	0.9	0.8864	0.1983	0.1983	gut
15	0.9	0.8863	0.1983	0.1987	gut
16	0.9	0.8869	0.1983	0.1997	gut
17	0.9	0.8869	0.1983	0.2007	gut
18	0.9	0.8872	0.1983	0.1977	gut
19	0.9	0.8865	0.1983	0.1990	gut
20	0.9	0.8860	0.1983	0.1980	gut

Zum Schluss der Untersuchungen stochastischer Prozesse sei wieder das Schwingungsmodell, das bereits in Kapitel 4.2.1 Anwendung gefunden hat, verwendet. Die frei wählbaren Einstellungen seien wieder auf $\alpha = 0.699$ als Gewichtungsparameter und der Startwert auf $x_0 = 2.4$ festgesetzt. Die 20 Surrogate werden wieder mit Hilfe des numerischen Verfahrens erzeugt, indem die 1024 Datenpunkte als Originaldatenreihe dienen.

Da für das Schwingungsmodell nicht die in R implementierte Funktion genutzt werden kann, muss auf eine andere Überprüfung zurück gegriffen werden. Hierzu eignet sich die Berechnung der Konfidenzbereiche, Mittelwerte etc. für jede Reihe. Stimmen die Kenngrößen überein, so können die Surrogate als Stellvertreter der Originaldatenreihe verwendet werden.

Konkret wurden das Minimum, das 1. Quartil, der Median, das 3. Quartil und das Maximum einer jeden Reihe berechnet. Bis auf minimale numerische Ungenauigkeiten stimmen die Kenngrößen der Originalreihe und der einzelnen Surrgate überein, so dass auf eine tabellarische Aufführung der Ergebnisse verzichtet wird. Die Surrogate können mit diesem Resultat als gute Repräsentanten angesehen werden, so dass eine Grundlage für weitere empirische Untersuchungen geschaffen ist.

5.2.2 Untersuchungen mit stabilen und instabilen Prozessen

Da das Hauptziel der Arbeit ist, eine Klassifikation von gegebenen Datenreihen in stabile und instabile Prozesse vorzunehmen, bietet sich für experimentelle Untersuchungen die schon in Kapitel 4.2.2 vorgestellte Modulo-1-Abbildung an. Durch die Wahl der Parameter können sowohl stabile als auch instabile Systeme generiert werden.

Um nun zu prüfen, ob die Surrogate, die aus einer generierten Modulo-1-Abbildung erzeugt wurden, gute Repräsentanten des zu Grunde liegenden Prozesses sind, kann leider auch hier nicht auf ein Verfahren wie bei den linearen Prozessen in Kapitel 5.2.1 zurückgegriffen werden. So wird bei der Überprüfung der Repräsentationsfähigkeit der Surrogate die Verteilung mit ihren Konfidenzbereichen, Mittelwerten, Medianen etc. für jede Reihe berechnet, wie es bei dem sin-Prozess Anwendung fand. Stimmen diese Kenngrößen überein, so wird der Schluss gezogen, dass die Surrogate als Stellvertreter der Originaldatenreihe verwendet werden dürfen.

Die Überprüfung wird mit Hilfe der in Kapitel 4.2.2 beschriebenen stabilen und instabilen Form der Modulo-1-Abbildung exemplarisch durchgeführt. Die Parametereinstellungen bleiben zu Kapitel 4.2.2 unverändert. Es werden somit aus den Gleichungen (4.7) und (4.8) jeweils 1024 Datenpunkte als Originaldatenreihen erzeugt. Anschließend werden sowohl für die stabilen Zeitreihen als auch für die instabilen Beobachtungsreihen jeweils 20 Surrogate erzeugt und anhand der verschiedenen Kenngrößen verglichen.

Es stellt sich heraus, dass die Ergebnisse, bis auf numerische Instabilitäten in der 5. Nachkommastelle, exakt übereinstimmen, so dass auf eine tabellarische Darstellung der Ergebnisse verzichtete wird.

Wie schon für den stabilen Prozess zeigen sich für den instabilen Prozess ebenfalls deutliche Übereinstimmungen in den Kenngrößen. Auch hier wird wegen der Exaktheit auf eine tabellarische Darstellung verzichtet. In beiden Fällen sind die Surrogate hinsichtlich ihrer Verteilung sehr gute Repräsentanten für ihre Originaldatenreihe. Dieses Ergebnis verwundert jedoch nicht. Da das numerische Verfahren so aufgebaut ist, dass die Autokorrelationsfunktion und die Verteilung der Originaldatenreihen und der Surrogate gleich sein sollen, musste ein vergleichbares Ergebnis gefunden werden.

Kapitel 6

Test zur Unsicherheit der Klassifikationsentscheidung

Wie Kapitel 5 zeigt, ist es möglich, mit der Annahme von bestimmten funktionalen Eigenschaften Surrogate nur mit Hilfe einer gegebenen Datenreihe zu erzeugen. Hier schließt sich jedoch die Frage an, ob auch ein Test konstruiert werden kann, der auf Basis dieser Stichprobe Aussagen über die Wahrscheinlichkeit gemachter Fehlklassifikationen trifft. So ist es durchaus denkbar, dass z. B. durch Zufallseinflüsse eine Entscheidung über stabile und instabile Prozesse getroffen wurde, obwohl bei einer nächsten Datenreihe aus dem gleichen zu Grunde liegenden Prozess eine andere Entscheidung fallen würde. In dieser Arbeit ist es jedoch nicht möglich, auf das zu Grunde liegende System zurück zu greifen, so dass der zu konstruierende Test diesen Sachverhalt berücksichtigen muss. Wie sich zeigen wird, sind in der Literatur schon mögliche Testprozeduren bekannt, die aber einige Mängel aufweisen. Daher wird eine Alternative aufgestellt, die zum einen eine Testentscheidung hinsichtlich des Lyapunov-Exponenten als Klassifikationskriteriums erfüllt, zum anderen jedoch auch die Grundlagen des statistischen Testens berücksichtigt (vgl. Kapitel 6.1). Die bereits in den Kapiteln 4.2 und 5.2 verwendeten Prozesse sollen wiederum der experimentellen Unterstützung dienen (vgl. Kapitel 6.2).

6.1 Testverfahren

Um empirische Eigenschaften des berechneten Lyapunov-Exponenten bestimmen und einen Test über die Unsicherheit der Klassifikationsentscheidung erzeugen zu können, muss auf eine Stichprobe der gegebenen Datenreihe zurück gegriffen werden, da der funktionale Zusammenhang der Daten über die Zeit nicht bekannt ist. Damit die direkte Verbindung der Klassifikationsentscheidung zum Test gegeben ist, ist es sinnvoll, den Lyapunov-Exponenten in die Teststatistik einzubauen.

Anfangen möchte ich mit einer kurzen Literaturzusammenfassung (vgl. Kapitel 6.1.1). Anschließend werde ich den Ausgangstest zu dem in dieser Arbeit entwickelten Test näher beschreiben, um am Ende die neue Testprozedur vorzustellen (vgl. Kapitel 6.1.2).

6.1.1 Darstellung bestehender Testprozeduren

In der Literatur sind einige Verfahren bekannt, die auf Grund von Hypothesentests Klassifikationen von gegebenen Datenreihen mit Hilfe geeigneter Teststatistiken und Testgrößen vornehmen. An dieser Stelle werden exemplarisch einige Arbeiten vorgestellt. Park und Whang (1999) entwickelten auf Basis eines Nadaraya-Watson-Regressionsschätzers einen formalen Test, mit Hilfe des Lyapunov-Exponenten eine Testentscheidung zwischen einem Random Walk und Chaos zu treffen. Ihre experimentellen Anwendungen basierten hierbei auf typische Makro- und Finanzdatenreihen. Sie beschreiben ein Verfahren, bei dem nur alleine mit der gegebenen Datenreihe, also ohne Surrogate, gearbeitet wird.

Eine ebenfalls nicht auf der Basis von Surrogaten entstandene Untersuchung zur Unterscheidung zwischen chaotisch deterministischer und stochastisch nichtlinearer Modellierung ist erstmals von Casdagli (1991) eingeführt worden. Hierbei handelt es sich nicht um einen formal exakten statistischen Test, sondern vielmehr um einen Algorithmusvorschlag, Datenreihen an stückweise lineare Reihen anzupassen und zu entscheiden, ob Chaos oder stochastische Nichtlinearität in den Daten vorliegt. Der Hauptgedanke ist dabei, dass Modelle, die nur richtige Kurzzeitprognosen zeigen, eher dem deterministischen Chaos als den nichtlinearen, stochastischen Systemen zuzuschreiben sind. Der Lyapunov-Exponent findet hierbei keine Anwendung.

Für das Hypothesentesten auf Basis von Surrogaten kann die Arbeit von Theiler, Galdrikian, Longtin, Eubank und Farmer (1992), die in Theiler, Eubank, Longtin, Galdrikian und Farmer (1992) nochmals abgedruckt wurde, als Ausgangssituation angesehen werden. Sie behandelt die Identifikation von Nichtlinearitäten in Datenreihen. In dem Paper werden verschiedene Teststatistiken vorgeschlagen, wie z. B. die Korrelationsdimension, der Vorhersagefehler oder auch der Lyapunov-Exponent. Zu den einzelnen Testprozeduren, die jedoch formal unbenannt bleiben, werden je nach Teststatistik die verschiedenen Nullhypothesen formuliert. Den Schwerpunkt legen die Autoren hauptsächlich auf die Erzeugung der Surrogate, die grundlegend für die sich anschließenden Tests sind. In Kapitel 5.1 wurde dieses Verfahren bereits vorgestellt.

Seitdem wurde an dem Verfahren zur Erzeugung der Stichprobe gearbeitet (vgl. ebenfalls Kapitel 5.1) und weitere Teststatistiken vorgeschlagen. Paluš (1994) schlägt das informationstheoretische Funktional der Redundanz als Teststatistik vor, um auf Nichtlinearitäten in den Daten zu testen. Er verbindet ein sogenanntes qualitatives Testen mit einem quantitativen Testen. Hierbei beruht das qualitative Testen auf einer Berechnung der informationstheoretischen Maße an nur einer Datenreihe, wohingegen das quantitative Testen auf der Basis von Surrogaten erfolgt.

Auch Schreiber und Schmitz (1996) versuchen, Nichtlinearitäten in den Datenreihen zu identifizieren, indem sie die von Theiler, Galdrikian, Longtin, Eubank und Farmer (1992) vorgeschlagene Surrogatmethode weiter entwickeln (vgl. Kapitel 5.1). Bei der Formulierung der Hypothesen und Teststatistiken wird an der Ausgangsform, wie sie Theiler, Galdrikian, Longtin, Eubank und Farmer (1992) vorschlugen, festgehalten.

Jedoch führen Schmitz und Schreiber (1998) in einem nachfolgenden Paper Untersuchungen durch bzgl. anderer Teststatistiken, die auf dem so genannten *Lomb-Periodogramm* [1] beruhen. Es wird dabei versucht, ungleich gesampelte Datenreihen auf Nichtlinearitäten zu überprüfen.

Eine Unterscheidung zwischen nichtlinearen und nichtstationären Systemen wurde von Schmitz und Schreiber (2000) ebenfalls auf Basis von Surrogaten getroffen.

6.1.2 Herleitung der Testprozedur

Die in der Literatur beschriebenen Testverfahren eignen sich aus verschiedenen Gründen nicht zur direkten Anwendung auf das hier beschriebene Klassifikationsproblem, zwischen stabilen, gut prädizierbaren und instabilen, nicht gut prädizierbaren Zeit- und Beobachtungsreihen zu unterscheiden. Daher wird ein modifiziertes Testverfahren vorgestellt, das diese Aufgabe löst. In Anlehnung an die Testmethoden, die auf der Basis von Surrogaten getroffen werden, wird eine Verbindung zu einem Test aus der nichtparametrischen Statistik gezogen.

Zuerst wird eine Arbeit vorgestellt, die für den hier neu erzeugten Test als Ausgangssituation dient, da er auf Basis von Surrogaten arbeitet. Dem anschließen wird sich eine Darstellung über eine Verbesserung der Testsituation in der Ausgangsarbeit. Im Laufe des Kapitels wird dann eine Testprozedur entwickelt, die das oben beschriebene Klassifikationsproblem testet.

Theiler und Prichard (1996) schlagen u. a. eine Methode vor, auf der Basis von Surrogaten die Datenreihen auf die Nullhypothese zu testen, ob sie aus einem Gaußprozess stammen oder nicht. Hierbei nutzen sie das Mittel über das vierte Moment als Teststatistik, verweisen jedoch direkt auch auf andere denkbare Teststatistiken. Der Trick in dieser Testprozedur ist, auf Basis von Rangstatistiken

[1] Das *Lomb-Periodogramm* wird in der Astrophysik zur Schätzung des Powerspektrums verwendet.

zu einer Entscheidung zu kommen.

Die Testprozedur in dem von Theiler und Prichard (1996) vorgeschlagenen Verfahren beginnt mit der Erzeugung von Surrogaten und der Berechnung der Teststatistik zu jedem Surrogat und zur Originalreihe. Anschließend werden die berechneten Teststatistiken der Größe nach sortiert. Der Rang der Teststatistik, die zur Originalreihe gehört, entscheidet über Annahme oder Ablehnung der Nullhypothese. Das folgende Schaubild gibt kurz das Schema der Testprozedur wider.

Testprozedur nach Theiler und Prichard (1996)

1. Berechne Teststatistik T für Originalreihe: T_0,
2. Berechne N_S Surrogate auf Basis der Originaldatenreihe,
3. Berechne Teststatistiken zu den Surrogaten $T_1, \ldots, T_{N_S}$,
4. Bilde Ränge über die Menge $T_0, T_1, \ldots, T_{N_S}$ mit kleinstem Rang 1 ($r_i = 1$) und größtem Rang $(N_S + 1)$, d. h. $r_j = N_S + 1$ mit $i, j \in 0, 1, \ldots, N_S$,
5. Überprüfe Rang r_0 der Originalreihe,
6. Lehne H_0 ab, wenn r_0 größer als $(N_S + 1) - (N_S + 1) \cdot \frac{\alpha}{2}$ oder kleiner als $(N_S + 1) \cdot \frac{\alpha}{2}$ in der geordneten Liste ist.

Wie der Ablehnbereich berechnet wurde, wird in dem Artikel nicht erläutert. In dem Fall von $N_S = 19$ Surrogaten und einem Signifikanzniveau von $\alpha = 0.05$ wird die Hypothese H_0 immer dann verworfen, wenn die Originaldatenreihe eine Teststatistik T_0 hat, die entweder kleiner oder größer ist als alle Teststatistiken der Surrogate. Die Hypothese auf diese Weise zu verwerfen, scheint natürlich willkürlich, so dass hinsichtlich einer Testprozedur, die auf das hier verwendete Klassifikationsproblem zugeschnitten sein soll, dieses noch überprüft werden muss.

Da die Testentscheidung auf Basis von Rangstatistiken getroffen wird, ist es sinnvoll, die Berechnung von Toleranzgrenzen vorzunehmen. Denn egal welche Nullhypothese durch die Teststatistik überprüft wird, bei einer Testprozedur mit Überprüfung des Rangs der Originaldatenreihe zu den Rängen der Surrogate wird immer die Frage verfolgt, ob die Originaldatenreihe einen Ausreißer bzgl. der Surrogate darstellt. Die Frage anders herum formuliert bedeutet, ob die Surrogate gute Stellvertreter für die Originaldatenreihe sind hinsichtlich der Prüfgröße, die in die Teststatistik eingeht.

Zu suchen ist also der Bereich $[TG_1, TG_2]$, in dem sich die Zufallsvariable T_0 befinden sollte, die auf Grund der Originaldatenreihe berechnet wurde. TG_1 und TG_2 sind ebenfalls Zufallsvariablen und werden Toleranzgrenzen genannt. Gesucht ist also die Wahrscheinlichkeit dafür, dass zwischen TG_1 und TG_2 der Anteil der Grundgesamtheitsverteilung größer oder gleich einem vorgegebenen β ist, $1 - \alpha$ beträgt, d. h.:

$$Q(P(TG_1 < T_0 < TG_2) \geq \beta) = 1 - \alpha. \tag{6.1}$$

Hierbei sei Q das zu der Zufallsvariablen $P(TG_1 < T_0 < TG_2)$ gehörende Wahrscheinlichkeitsmaß. Gesucht sind also die Statistiken TG_1 und TG_2, die die Gleichung (6.1) erfüllen. Nun wird das Problem so transformiert (vgl. Büning und Trenkler (1978), S. 80-81), dass zu gegebenen β und $1 - \alpha$ die Rangzahlen k und l mit $k < l$ so bestimmt werden, dass

$$Q(Y_{(l)} - Y_{(k)} \geq \beta) = 1 - \alpha \tag{6.2}$$

ist. Nach Umformungen und weiteren Transformationen (vgl. Büning und Trenkler (1978), S. 80-81) zeigt sich, dass

$$Q(Y_{(l)} - Y_{(k)} \geq \beta) = F(l - k - 1) \tag{6.3}$$

beträgt, wobei F die Verteilungsfunktion einer binomialverteilten Zufallsvariablen mit Parametern n und β ist. Es muss nun nur noch

$$F(l - k - 1) = 1 - \alpha \tag{6.4}$$

bzgl. l und k gelöst werden (vgl. geeignetes Tabellenwerk wie z. B. in Büning und Trenkler (1978), S. 342).

Mit Hilfe dieser Herleitung kann nun die Testprozedur von Theiler und Prichard (1996) auf eine statistische Grundlage gestellt werden.

Nehmen wir die Einstellungen zu den gemachten Experimenten aus dem Artikel von Schreiber und Schmitz (2000), d. h. $N_S = 19$ und $\alpha = 0.05$, an, so würde die berechnete Ablehngrenze zu einer Power des Tests von 0.2 führen. Dass dieses nicht akzeptiert werden kann, liegt auf der Hand. Schon um mindestens eine Power von 0.5 zu erreichen, müßte H_0 abgelehnt werden, wenn $TG_1 = T_{(3)}$ und $TG_2 = T_{(17)}$ bzw. $TG_1 = T_{(3)}$ und $TG_1 = T_{(16)}$ oder $TG_1 = T_{(2)}$ und $TG_1 = T_{(17)}$.

Natürlich ist auch eine Power von 0.5 nicht das Ziel eines guten Tests, da es gleichbedeutend mit gutem Raten oder einer zufälligen Entscheidung wäre. Vorgeschlagen wird daher, die Power des Tests auf 0.8 anzuheben. Durch Erhöhung der Stichprobe oder durch Verkleinerung des Toleranzbereichs ist dieses in jedem Fall zu erreichen. Für das in dem Artikel von Schreiber und Schmitz (2000) verwendete Beispiel von $N_S = 19$ und $\alpha = 0.05$ würde ein Toleranzbereich von $[T_{(7)}, T_{(14)}]$ bzw. $[T_{(7)}, T_{(15)}]$ oder $[T_{(6)}, T_{(14)}]$ eine Power von 0.8 erreichen.

Angewendet auf das Klassifikationsproblem, einen stabilen von einem instabilen Prozess mit Hilfe des Lyapunov-Exponenten, in der Teststatistik kurz mit LE abgekürzt, unterscheiden zu können, könnte folgende Prozedur vorgeschlagen werden: Das Testverfahren ist ein zweigeteilter Test. Zuerst wird auf gute Stellvertreter getestet und anschließend die Klassifikationsentscheidung überprüft. Als erstes läge ein Entscheidungsproblem der folgenden Art vor:

$$H_{01}: \quad LE_0 = LE_s \qquad \forall s \in 1, \ldots, N_S$$

vs.

$$H_{11}: \quad LE_0 \neq LE_s \qquad \text{für mindestens ein } s \in 1, \ldots, N_S.$$

Die Hypothese, die auch als $H_0 : \lambda_0 = \lambda_s$ geschrieben werden könnte, würde bedeuten, dass alle erzeugten Surrogate gute Stellvertreter der Originalreihe hinsichtlich ihres Lyapunov-Exponenten darstellen. Im nächsten Schaubild ist die dazu gehörige Testprozedur abgebildet.

Testprozedur 1 (auf gute Stellvertreter):

1. Berechne Teststatistik LE für Originalreihe: λ_0,
2. Berechne N_S Surrogate auf Basis der Originaldatenreihe,
3. Berechne Teststatistiken zu den Surrogaten $\lambda_1, \ldots, \lambda_{N_S}$,
4. Bilde Ränge über die Menge $\lambda_0, \lambda_1, \ldots, \lambda_{N_S}$ mit kleinstem Rang 1 ($r_i = 1$) und größtem Rang $(N_S + 1)$, d. h. $r_j = N_S + 1$ mit $i, j \in 0, 1, \ldots, N_S$,
5. Überprüfe Rang r_0 der Originalreihe,
6. Lehne H_{01} ab, wenn r_0 nicht im Toleranzbereich liegt (hierbei wird der Toleranzbereich durch das oben beschriebene Verfahren berechnet (wobei das Signifikanzniveau mit α_1 bezeichnet sei)).

Um nun nicht nur auf gute Stellvertreter zu testen, sondern ebenfalls die Klassifikationsentscheidung zu treffen, wird eine zweite Testprozedur durchgeführt. Diese berücksichtigt die Wahrscheinlichkeit p für eine Fehlklassifikation bzw. beschreibt die Unsicherheit der Klassifikationsentscheidung:

$$H_{02} : p \leq 0.05 \qquad \text{vs. } H_{12} : p > 0.05.$$

Hinter diesem Testproblem steckt also die einfache Frage, ob wirklich höchstens 5% der Surrogate falsch klassifiziert haben. Als Testprozedur wird das Vorzeichen der berechneten Lyapunov-Exponenten der Surrogate in einer Tabelle abgetragen und ein Binomialtest (Büning und Trenkler, 1978) darauf angewendet. Die dazugehörige Teststatistik ist die Anzahl B der Beobachtungen, die zu Klasse 1 gehören. Sie ist binomialverteilt mit den Parametern n und p. H_{02} wird abgelehnt, wenn die Teststatistik größer ist als das $(1-\alpha)$-Fraktil der Binomialverteilung mit den Parametern n und p. Für das Klassifikationsproblem sähe die zweite Testprozedur so aus, wie sie im nachfolgenden Schaubild dargestellt ist.

Testprozedur 2 (auf Klassifikation):

1. Schreibe Vorzeichen der berechneten Lyapunov-Exponenten aus der Testprozedur 1 in eine Tabelle,
2. Berechne Anzahl, die zur Klasse 1 gehört (z. B. alle negativen Vorzeichen),
3. Lehne H_{02} ab, wenn $P(B \geq t_{1-\alpha_2}) = \alpha_2$ mit dem Signifikanzniveau α_2.

Die allgemeine Hypothese, dass die Originalreihe hinsichtlich der Surrogate und des Lyapunov-Exponenten richtig klassifiziert ist, wird genau dann abgelehnt, wenn entweder Testprozedur 1 oder Testprozedur 2 zur Ablehnung führten. Dieses ist ein multipler Test, so dass das Signifikanzniveau beachtet werden muss. Unter der Annahme, dass $\alpha_1 = \alpha_2 = \alpha$ sein soll, folgt, dass das allgemeine Signifikanzniveau bei $c = \mathrm{Min}[1, \alpha \cdot \frac{n}{k}] = \mathrm{Min}[1, \alpha \cdot \frac{2}{1}]$ liegt. Hierzu sei auf Rüger (1978) verwiesen, der als Erster das maximale Signifikanzniveau berechnete zu der Art „Lehne H_0 ab, wenn k von n gegebenen Tests zur Ablehnung führen." Anschließend kam Morgenstern (1980) zu dem gleichen Ergebnis durch Verwendung der Markov'schen Ungleichung.

Bei diesem Test fällt jedoch auf, dass für das hier gestellte Klassifikationsproblem zu viel gefordert ist: Das Verfahren zur Erzeugung der Surrogate nimmt die gleiche Autokorrelationsfunktion und die gleiche Verteilung an, fordert jedoch in keinem Fall, dass die erzeugten Surrogate hinsichtlich des Lyapunov-Exponenten „gleichmäßig“ um die Originaldatenreihe streut. So scheitert der Test zur Klassifikationsentscheidung über stabile und instabile Prozesse bereits an der ersten Hypothesenüberprüfung, wenn nicht die Zusatzannahme bei der Erstellung der Surrogate gemacht wird, dass diese bzgl. des Lyapunov-Exponenten um die Originaldatenreihe streuen sollen.

Auf Grund dieser Erkenntnisse wird ein neuer Test vorgeschlagen, der zum einen auf das hier gestellte Klassifikationsproblem der Entscheidung über stabile und instabile Systeme mit Hilfe des Lyapunov-Exponenten eingeht, zum anderen jedoch auch keine weiteren Annahmen bei der Generierung der Surrogate benötigt.

Dieser Test beruht auf einer Teststatistik, die die Vorzeichen der berechneten Lyapunov-Exponenten berücksichtigt und an einen nichtparametrischen Binomialtest angelehnt ist.

Gegeben sei nun die Hypothese:

$$\begin{aligned} H_0 &: p \leq \text{vorgegebene Wahrscheinlichkeit für Fehlklassifikationen} \\ \text{vs. } H_1 &: p > \text{vorgegebene Wahrscheinlichkeit für Fehlklassifikationen} \end{aligned} \tag{6.5}$$

Dieses ist gleichbedeutend, dass höchstens eine vorgegebene Prozentzahl der Surrogate die falsche Klassifikationsentscheidung getroffen hat. Hierbei ist auch jede wünschenswerte Prozentzahl einsetzbar, vorzugsweise bietet sich eine 5%-ige Fehlerrate an. Nun werden, statt einfach das Vorzeichen der berechneten Lyapunov-Exponenten zu betrachten und die Anzahl der zu einer Klasse Gehörenden zu messen, wie es bei dem vorangegangenen Test der Fall war, die Lyapunov-Exponenten der Originaldatenreihe und jeweils eines Surrogats miteinander multipliziert. Als Teststatistik dient dann die Anzahl aller $\lambda_0 \cdot \lambda_i$, mit $i = 1, \ldots N_S$, die größer als null sind. So bedeutet eine gleiche Klassifikation der Originaldatenreihe und des Surrogats, dass die Größe $\lambda_0 \cdot \lambda_i$, mit $i = 1, \ldots N_S$, immer positiv ist. Unterscheiden sich hingegen die beiden Lyapunov-Exponenten, so wird $\lambda_0 \cdot \lambda_i < 0$ sein. Das Vorzeichen wird negativ. Die Teststatistik ist binomialverteilt mit dem Parameter p =Wahrscheinlichkeit der Fehlentscheidungen (bei einer 5% Fehlklassifikationsrate, z. B. $p = 0.05$) und $n = N_S$ (Anzahl der Surrogate). Auch dieser Test ist wieder im folgenden Schaubild dargestellt.

Testprozedur zur Unsicherheit der Klassifikationsentscheidung:

1. Berechne Lyapunov-Exponent für Originaldatenreihe: λ_0,
2. Berechne N_S Surrogate auf Basis der Originaldatenreihe,
3. Berechne Lyapunov-Exponenten der Surrogate $\lambda_1, \ldots, \lambda_{N_S}$
4. Bestimme $\lambda_0 \cdot \lambda_i$ mit $i = 1, \ldots, N_S$ für jedes Surrogat,
5. Berechne $LE = \#[(\lambda_0 \cdot \lambda_i) < 0]$,
6. Lehne H_0 ab, wenn $P(LE \geq \text{bin}_{1-\alpha}) = \alpha$ mit einem Signifikanzniveau α, wobei LE binomialverteilt ist mit $n = N_S$ und $p = \frac{\text{Fehlerrate (in \%)}}{100}$ (z. B. $p = 0.05$).

Zur Lösung der Probleme, dass das Signifikanzniveau bei einer diskreten Teststatistik nicht ganz ausgeschöpft wird, dem Auftreten von Bindungen sowie der Beschreibung, was bei großen Stichproben zu tun ist, sei auf Büning und Trenkler (1978) verwiesen.

Sei nun konkret eine Fehlklassifikationsrate von 5% zugelassen, das Signifikanzniveau auf $\alpha = 0.1$ festgelegt und werden $N_S = 20$ Surrogate erzeugt, so wird die Hypothese abgelehnt, wenn $LE \geq 3$ ist. Das bedeutet, dass maximal 2 unterschiedliche Klassifikationen auftreten dürfen.

Mit Hilfe dieses Tests ist es also möglich, mit einem vorgegebenen Signifikanzniveau eine Aussage darüber zu machen, ob einer Klassifikationsentscheidung zu einer originalen Datenreihe zu trauen ist oder nicht.

6.2 Experimentelle Ergebnisse

Um eine getroffene Klassifikationsentscheidung statistisch abzusichern, wurde in Kapitel 6.1.2 ein Test entwickelt, der Aussagen über eine mögliche Fehlentscheidung macht. In diesem Kapitel soll nun der Test mit den bereits aus den Kapiteln 4.2 und 5.2 bekannten Prozessen experimentell unterstützt werden. Hierzu werden mit Hilfe der Surrogate jeweils Tests durchgeführt und mit einem vorgegebenen Signifikanzniveau überprüft, ob die Hypothese, dass wirklich höchstens eine 5% Fehlerrate aufgetreten ist, angenommen oder abgelehnt werden muss.

Neben dem Test werden zusätzlich Dichtefunktionen, die über die Lyapunov-Exponenten zu der Originaldatenreihe und zu den Surrogaten berechnet werden, erstellt. Hierzu wird die bereits im Programmpaket *R* implementierte Funktion *density* verwendet, die auf einer Kerndichteschätzung beruht (Ihaka und Gentleman, 1996), (Silverman, 1986).

Die Dichtefunktionen sollen die Verteilung graphisch verdeutlichen, so dass auch ein Eindruck über das Streuen der Surrogate um den originalen Lyapunov-Exponenten vermittelt werden kann. Zudem ergibt sich zusätzlich eine Möglichkeit, die Testentscheidung anschaulich nachzuvollziehen.

6.2.1 Untersuchungen mit stochastischen Prozessen

Die aus den Kapiteln 4.2.1 und 5.2.1 bekannten stochastischen Prozessen sollen den aus Kapitel 6.1.2 entwickelten Test experimentell unterstützen. Für den AR(1)-, den MA(1)- und den ARMA(1,1)-Prozess werden die Lyapunov-Exponenten der Originalzeitreihen und der erstellten Surrogate berechnet. Mit Hilfe der neu entwickelten Teststatistik wird dann entschieden, ob der Klassifikationsentscheidung bzgl. der Originalzeitreihe zu einem vorgegebenen Signifikanzniveau von $\alpha = 0.01$ und bei einer Fehlerrate von höchstens 5% zu trauen ist.

Sei nun konkret das Testproblem der folgenden Art gegeben:

$$H_{02} : p \leq 0.05 \qquad \text{vs. } H_{12} : p > 0.05. \tag{6.6}$$

Dieses ist gleichbedeutend, dass höchstens 5% der Surrogate anders entscheiden als bei der Originaldatenreihe.

Zuerst werden nun für den AR(1)-Prozess aus Gleichung (4.1), für den MA(1)-Prozess aus Gleichung (4.2) und für den ARMA(1,1)-Prozess aus Gleichung (4.3) Originalzeitreihen generiert. Hierbei bleiben die weiteren Einstellungen wie in den Kapiteln 4.2.1 und 5.2.1 unverändert. Für diese Zeitreihen werden die originalen Lyapunov-Exponenten mit Hilfe des in Kapitel 4 dargestellten Verfahrens berechnet. Anschließend werden mit Hilfe des Verfahrens aus Kapitel 5 jeweils $N_S = 20$ Surrogate zu den einzelnen Prozessen erstellt und die zugehörigen Lyapunov-Exponenten ebenfalls berechnet. Hieraus ergeben sich dann die Werte der Teststatistiken, die über die Annahme oder über das Ablehnen der Hypothese entscheiden.

Da die Prüfgröße binomialverteilt ist mit den Parametern $n = N_S = 20$ und $p = 0.05$ ergibt sich (vgl. Büning und Trenkler (1978), S. 352, 353):

$$P(LE \geq 2) = 0.2642 \quad \text{d. h. } 1 - P(LE \leq 1) = 1 - 0.7358 = 0.2642, \tag{6.7}$$

$$P(LE \geq 3) = 0.0755 \quad \text{d. h. } 1 - P(LE \leq 2) = 1 - 0.9245 = 0.0755. \tag{6.8}$$

Somit wird die Nullhypothese abgelehnt, sobald mehr als zwei Surrogate anders entschieden haben als die Originaldatenreihe.

Für den AR(1)-Prozess zeigt Tabelle 6.1 die berechneten Lyapunov-Exponenten der Originaldatenreihe und der Surrogate und die Werte der Multiplikation

Tabelle 6.1: Bestimmung der Werte der Teststatistik für den AR(1)-Prozess

Surrogat	λ_0	λ_i	$\lambda_0 \cdot \lambda_i$
1	-0.0684	-0.0170	> 0
2	-0.0684	-0.0420	> 0
3	-0.0684	-0.0626	> 0
4	-0.0684	-0.0315	> 0
5	-0.0684	-0.0343	> 0
6	-0.0684	-0.0030	> 0
7	-0.0684	-0.0530	> 0
8	-0.0684	-0.0060	> 0
9	-0.0684	-0.0816	> 0
10	-0.0684	-0.0364	> 0
11	-0.0684	-0.0290	> 0
12	-0.0684	-0.0176	> 0
13	-0.0684	-0.0814	> 0
14	-0.0684	-0.0033	> 0
15	-0.0684	0.0109	$< \mathbf{0}$
16	-0.0684	-0.0317	> 0
17	-0.0684	-0.0742	> 0
18	-0.0684	-0.0297	> 0
19	-0.0684	0.0088	$< \mathbf{0}$
20	-0.0684	-0.0287	> 0

der verschiedenen Exponenten, die die Grundlage der Teststatistik bilden.

Da zwei Surrogate einen negativen Wert aufweisen, somit die Prüfgröße $LE = 2$ beträgt, wird H_0 auf dem Testniveau $\alpha = 0.1$ nicht abgelehnt. Der generierte Prozess ist also als stabil anzusehen. Dieses stimmt mit der Ausgangssituation überein, so dass der Test in diesem Fall eine gute, statistische Unterstützung darstellt. Es bleibt aber noch festzuhalten, dass bei der Überprüfung der Surrogate (vgl. Tabelle 5.1) keine Auffälligkeiten festzustellen waren, die einen Hinweis auf eine Fehlklassifikation gaben.

Eine graphische Unterstützung, die mit Hilfe der empirischen Dichtefunktion über die berechneten Lyapunov-Exponenten erstellt wurde, ist in Abbildung 6.1 wiedergegeben.

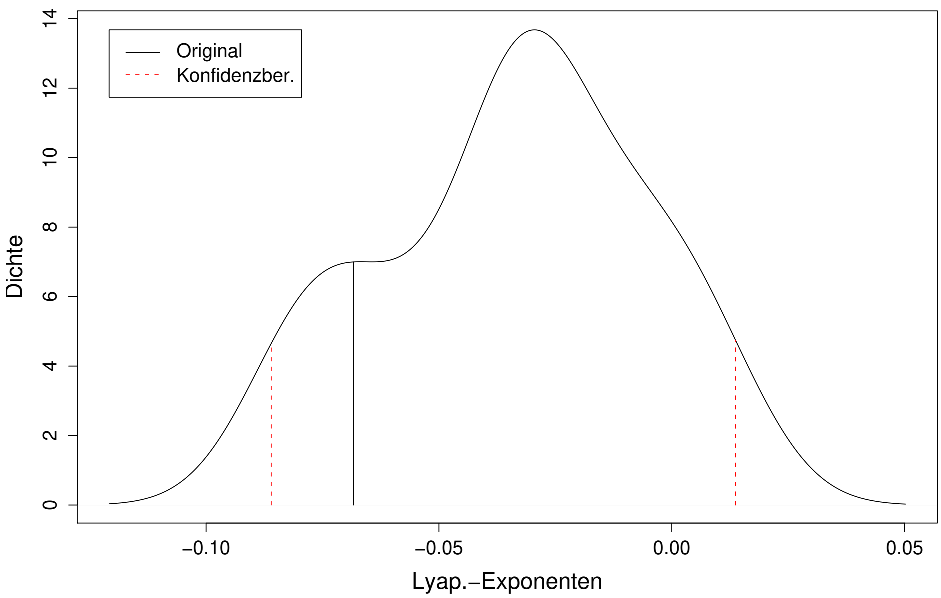

Abbildung 6.1: Geschätzte Dichtefunktion der Lyapunov-Exponenten zu dem AR-Prozess

In Abbildung 6.1 liegt viel Wahrscheinlichkeitsmasse im Bereich unter der kritischen Klassifikationsgrenze von null. Die Lyapunov-Exponenten streuen um den Lyapunov-Exponenten der Originalzeitreihe, auch wenn dieser nicht mit dem Median zusammenfällt.

Tabelle 6.2: Bestimmung der Werte der Teststatistik für den MA(1)-Prozess

Surrogat	λ_0	λ_i	$\lambda_0 \cdot \lambda_i$
1	-0.047	-0.067	> 0
2	-0.047	-0.048	> 0
3	-0.047	0.010	$< \mathbf{0}$
4	-0.047	0.053	$< \mathbf{0}$
5	-0.047	-0.022	> 0
6	-0.047	-0.013	> 0
7	-0.047	-0.011	> 0
8	-0.047	0.012	$< \mathbf{0}$
9	-0.047	0.050	$< \mathbf{0}$
10	-0.047	0.081	$< \mathbf{0}$
11	-0.047	0.018	$< \mathbf{0}$
12	-0.047	-0.049	> 0
13	-0.047	0.003	$< \mathbf{0}$
14	-0.047	-0.042	> 0
15	-0.047	0.028	$< \mathbf{0}$
16	-0.047	0.016	$< \mathbf{0}$
17	-0.047	-0.009	> 0
18	-0.047	0.037	$< \mathbf{0}$
19	-0.047	0.072	$< \mathbf{0}$
20	-0.047	0.026	$< \mathbf{0}$

Die Ergebnisse bzgl. des Tests für den MA(1)-Prozess sind in Tabelle 6.2 zusammengestellt.

Es zeigt sich deutlich (vgl. Tabelle 6.2), dass die Hypothese zu einem Signifikanzniveau von $\alpha = 0.1$ abgelehnt wird, da die Teststatistik deutlich größer als zwei ist ($LE = 12$). Es ist also nicht ratsam, der Klassifikationsentscheidung, die auf Basis des Lyapunov-Exponenten der Originalzeitreihe getroffen wurde, zu diesem Testniveau zu trauen. Dieses stimmt mit der Theorie nicht überein, da der wahre Lyapunov-Exponent kleiner als null sein sollte. So ist zu vermuten, dass das numerische Verfahren an dieser Stelle versagt. Eine Erklärung wäre, dass die Varianz des in dem MA(1)-Prozess einhergehenden Weißen Rauschens mit $\sigma^2 = 1$ zu groß gewählt ist, so dass ein stabiles Verhalten überdeckt wird (vgl. hierzu auch Anhang A.4). Um dieses zu überprüfen, wurden nochmals Simulationen mit einer kleinen Varianz ($\sigma^2 = 1/16$) mit anschließender Testentscheidung durchgeführt. Es zeigte sich, dass nur zwei Fehlklassifikationen auftraten, so dass der Prozess als stabil anzusehen ist. An dieser Stelle wird also zum ersten Mal richtig sichtbar, dass das Verhältnis vom Rauschen zum eigentlichen Signal bei der numerischen Berechnung eine große Rolle spielt. Überlagert das Rauschen mit großer Variabilität das Signal, kommt es vermehrt zu Fehlklassifikationen.

Wie schon bei dem AR-Prozess waren bei der Überprüfung der Surrogate zum MA-Prozess in Tabelle 5.2 keine Auffälligkeiten zu finden, die eine Fehlklassifikation, wie sie in Tabelle 6.2 sichtbar wird, vermuten lassen.

Verdeutlicht wird das Ergebnis durch die Betrachtung der empirischen Dichtefunktion über die Lyapunov-Exponenten.

Hier zeigt sich viel Wahrscheinlichkeitsmasse auch über den kritischen Klassifikationswert hinaus. Zudem zeigt sich in Abbildung 6.2, dass der Original-Lyapunov-Exponent am Rande der Verteilung liegt, so dass vermutet werden kann, dass die Surrogate hinsichtlich der Verteilung des Lyapunov-Exponenten in diesem Falle keine guten Stellvertreter abgeben.

Der letzte Test im Bereich der linearen Prozesse wird für den ARMA(1,1)-Prozess durchgeführt. Die Ergebnisse finden sich in Tabelle 6.3.

Da aus Tabelle 6.3 hervorgeht, dass die Prüfgröße $LE = 2$ beträgt, somit

Tabelle 6.3: Bestimmung der Werte der Teststatistik für den ARMA(1,1)-Prozess

Surrogat	λ_0	λ_i	$\lambda_0 \cdot \lambda_i$
1	-0.033	-0.047	> 0
2	-0.033	-0.004	> 0
3	-0.033	-0.014	> 0
4	-0.033	-0.037	> 0
5	-0.033	-0.016	> 0
6	-0.033	0.005	$< \mathbf{0}$
7	-0.033	-0.035	> 0
8	-0.033	-0.066	> 0
9	-0.033	-0.010	> 0
10	-0.033	-0.038	> 0
11	-0.033	-0.012	> 0
12	-0.033	-0.030	> 0
13	-0.033	-0.083	> 0
14	-0.033	0.052	$< \mathbf{0}$
15	-0.033	-0.014	> 0
16	-0.033	-0.048	> 0
17	-0.033	-0.068	> 0
18	-0.033	-0.062	> 0
19	-0.033	-0.013	> 0
20	-0.033	-0.066	> 0

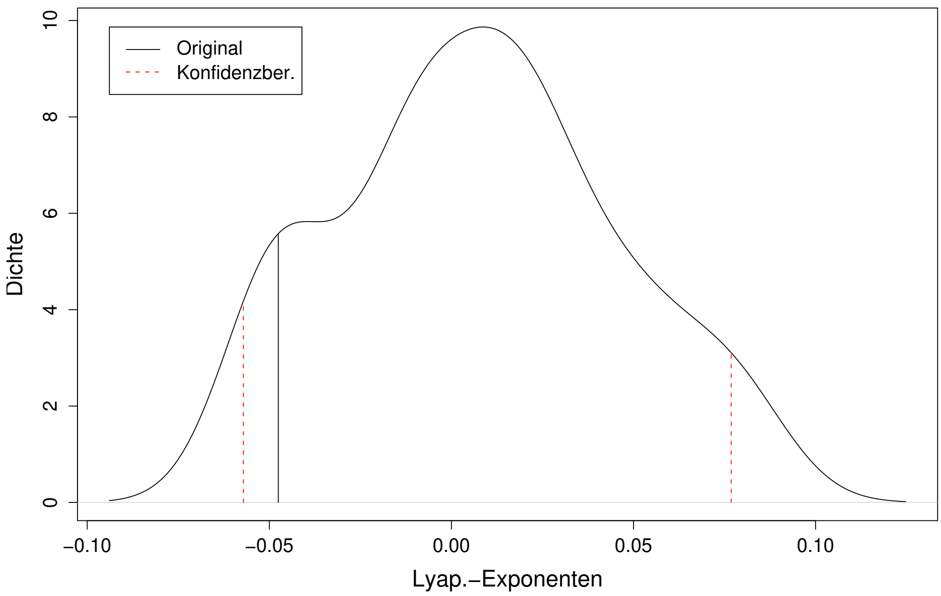

Abbildung 6.2: Geschätzte Dichtefunktion der Lyapunov-Exponenten zu dem MA-Prozess

also kleiner als drei ist, wird H_0 nicht abgelehnt. Die Klassifikationsentscheidung wird zu diesem Signifikanzniveau als richtig angesehen. Auch hier waren in den Simulationen zu den Surrogaten (vgl. Tabelle 5.3) keine Auffälligkeiten, die eine Fehlklassifikation ankündigen, zu vermerken.

Die grafische Darstellung der empirischen Verteilungsfunktion belegt dieses Testergebnis (vgl. Abbildung 6.3), da viel Wahrscheinlichkeitsmasse unterhalb der Null liegt, so dass wenig auf mögliche Fehlklassifikationen durch die Surrogate hinweist. Auch liegt der Lyapunov-Exponent, der zur Originalzeitreihe gehört, in der Nähe des Medians, so dass vieles darauf hinweist, dass die Surrogate hinsichtlich des Lyapunov-Exponeten gute Stellvertreter der Originalreihe darstellen.

Zum Schluss der Untersuchungen stochastischer Prozesse sei wieder das Schwingungsmodell, das bereits in Kapitel 4.2.1 Anwendung gefunden hat, verwendet.

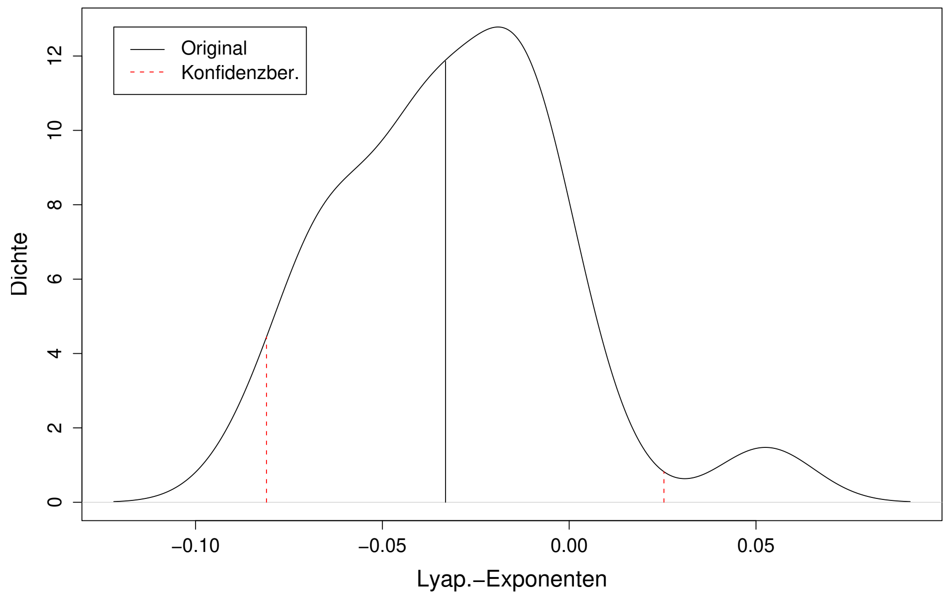

Abbildung 6.3: Geschätzte Dichtefunktion der Lyapunov-Exponenten zu dem ARMA-Prozess

Da für den MA-Prozess keine Testentscheidung getroffen werden konnte, dass weniger als 5% Fehlklassifikationen auftreten, könnte der Eindruck entstehen, dass die Berechnungen nur Prozesse mit einem autoregressiven Anteil richtig klassifizieren. So ist es interessant zu wissen, ob Fehlklassifikationen bei dem sin-Prozess auftreten (Tabelle 6.4).

Da kein Surrogat einen negativen Wert aufweist, somit die Prüfgröße $LE = 0$ beträgt, wird die Hypothese auf dem Testniveau $\alpha = 0.1$ nicht abgelehnt. Der erzeugte Prozess ist als stabil anzusehen.

Eine graphische Unterstützung, die mit Hilfe der Kerndichteschätzung über die berechneten Lyapunov-Exponenten erstellt wurde, ist in Abbildung 6.4 wiedergegeben.

Es zeigt sich, dass der Lyapunov-Exponent in der Nähe des Mittelwertes liegt und die gesamte Wahrscheinlichkeitsmasse unter der kritischen Klassifikations-

Tabelle 6.4: Bestimmung der Werte der Teststatistik für den sin-Prozess

Surrogat	λ_0	λ_i	$\lambda_0 \cdot \lambda_i$
1	-0.057	-0.05993	> 0
2	-0.057	-0.05591	> 0
3	-0.057	-0.05374	> 0
4	-0.057	-0.06148	> 0
5	-0.057	-0.05899	> 0
6	-0.057	-0.06447	> 0
7	-0.057	-0.05431	> 0
8	-0.057	-0.05335	> 0
9	-0.057	-0.06643	> 0
10	-0.057	-0.05596	> 0
11	-0.057	-0.05354	> 0
12	-0.057	-0.06866	> 0
13	-0.057	-0.04917	> 0
14	-0.057	-0.05195	> 0
15	-0.057	-0.05899	> 0
16	-0.057	-0.06498	> 0
17	-0.057	-0.04542	> 0
18	-0.057	-0.05405	> 0
19	-0.057	-0.05837	> 0
20	-0.057	-0.06264	> 0

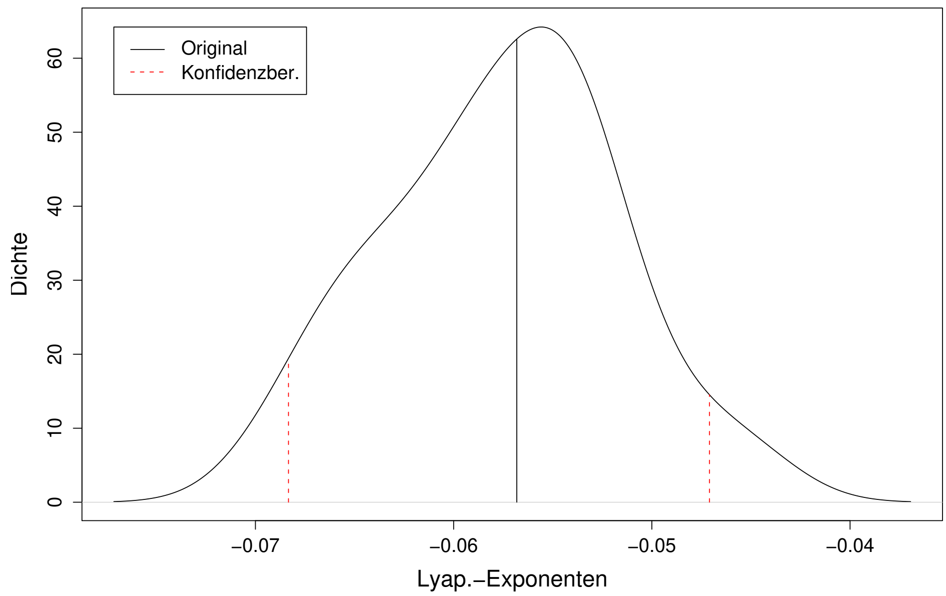

Abbildung 6.4: Geschätzte Dichtefunktion der Lyapunov-Exponenten zu dem sin-Prozess

grenze von null liegt. Die Lyapunov-Exponenten der Surrogate streuen um den Lyapunov-Exponenten der Originalreihe.

6.2.2 Untersuchungen mit stabilen und instabilen Prozessen

Da die in dieser Arbeit behandelte Klassifikation zwischen stabilen und instabilen Prozessen entscheidet, bieten sich wieder die beiden Modulo-1-Abbildungen aus dem Kapitel 5.2.2 an, um die neu entwickelte Testprozedur experimentell zu unterstützen.

Wie schon bei den Untersuchungen zu den linearen Prozessen, werden wieder die Lyapunov-Exponenten der Originaldatenreihen und der zugehörigen $n = N_S = 20$ Surrogate berechnet und eine Testprozedur zu der Hypothese:

$$H_{02} : p \leq 0.05 \qquad \text{vs. } H_{12} : p > 0.05 \tag{6.9}$$

durchgeführt. Hierbei sei das Signifikanzniveau wieder auf $\alpha = 0.1$ festgesetzt. Die Fehlerrate der Klassifikationsentscheidungen soll nochmals bei höchstens 5% liegen.

Die Hypothese wird abgelehnt, wenn die Prüfgröße abermals mehr als zwei beträgt (vgl. Kapitel 6.2.1).

In Tabelle 6.5 sind die Ergebnisse des durchgeführten Tests zusammen getragen.

Nur ein von 20 Surrogaten klassifizierte anders als die Originalzeitreihe. Somit liegt auch die Prüfgröße mit $LE = 1$ so, dass H_0 zu dem vorgegebenen Signifikanzniveau nicht abgelehnt wird. Der Entscheidung, die auf Grund der Originaldatenreihe getroffen wurde, kann getraut werden.

Wie schon in Kapitel 6.2.1 soll auch hier wieder die empirische Dichtefunktion anschaulich das Testergebnis bestätigen (vgl. Abbildung 6.5).

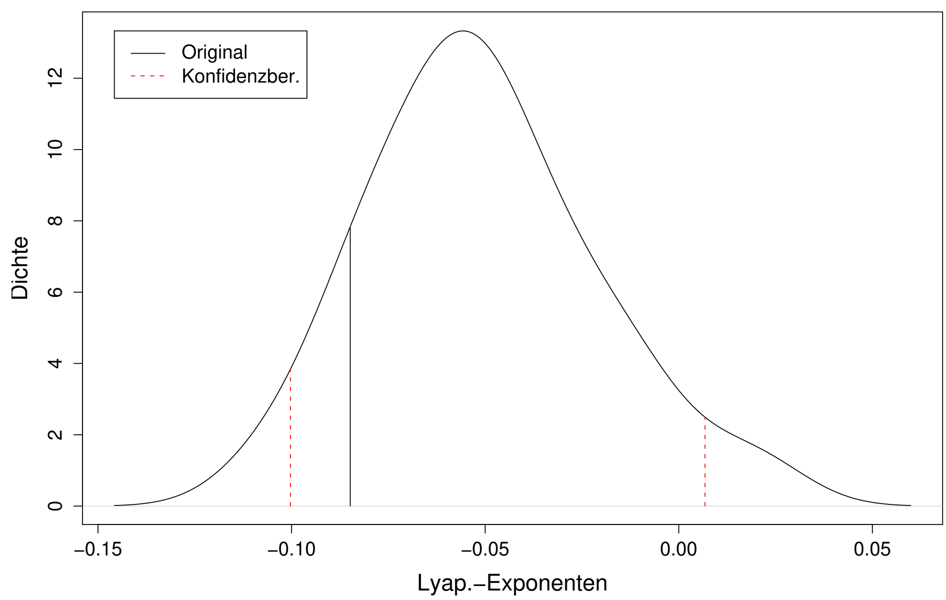

Abbildung 6.5: Geschätzte Dichtefunktion der Lyapunov-Exponenten zum stabilen Prozess

Tabelle 6.5: Bestimmung der Werte der Teststatistik für die stabile Form der Modulo-1-Abbildung

Surrogat	λ_0	λ_i	$\lambda_0 \cdot \lambda_i$
1	-0.094	-0.0848	> 0
2	-0.094	-0.0587	> 0
3	-0.094	-0.006	> 0
4	-0.094	-0.059	> 0
5	-0.094	-0.071	> 0
6	-0.094	-0.085	> 0
7	-0.094	-0.046	> 0
8	-0.094	-0.075	> 0
9	-0.094	-0.047	> 0
10	-0.094	-0.104	> 0
11	-0.094	-0.056	> 0
12	-0.094	-0.08	> 0
13	-0.094	-0.063	> 0
14	-0.094	-0.033	> 0
15	-0.094	-0.037	> 0
16	-0.094	-0.034	> 0
17	-0.094	-0.02	> 0
18	-0.094	-0.056	> 0
19	-0.094	0.018	$< \mathbf{0}$
20	-0.094	-0.014	> 0

Es fällt auf, dass viel Wahrscheinlichkeitsmasse in dem Bereich liegt, wo die Zeitreihe richtig klassifiziert würde. So untermauert die Darstellung der empirischen Dichtefunktion bzgl. der Lyapunov-Exponenten auch das Testergebnis. Jedoch wird ebenfalls deutlich, dass der Lyapunov-Exponent, der zur generierten Originaldatenreihe gehört, nicht in der Nähe des Medians liegt. Die Lyapunov-Exponenten der Surrogate streuen nicht gleichmäßig um den der Originalzeitreihe herum.

Die Ergebnisse der Testuntersuchungen eines instabilen Prozesses werden erneut in tabellarischer Form dargestellt (vgl. Tabelle 6.6).

Die Teststatistik beträgt $LE = 1$, so dass die Nullhypothese zu diesem Signifikanzniveau nicht abgelehnt wird. Die Klassifikationsentscheidung über einen instabilen Prozess auf Basis der Originalbeobachtungsreihe wird somit statistisch abgesichert.

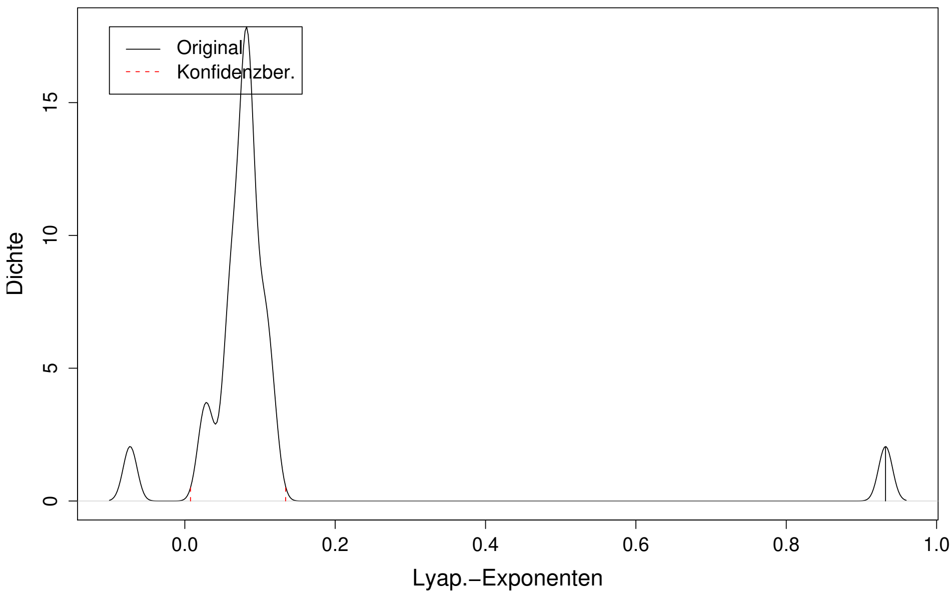

Abbildung 6.6: Geschätzte Dichtefunktion der Lyapunov-Exponenten zum instabilen Prozess

Tabelle 6.6: Bestimmung der Werte der Teststatistik für die instabile Form der Modulo-1-Abbildung

Surrogat	λ_0	λ_i	$\lambda_0 \cdot \lambda_i$
1	0.922	0.843	> 0
2	0.922	0.886	> 0
3	0.922	0.875	> 0
4	0.922	0.061	> 0
5	0.922	0.084	> 0
6	0.922	0.077	> 0
7	0.922	0.105	> 0
8	0.922	0.081	> 0
9	0.922	0.053	> 0
10	0.922	-0.072	$< \mathbf{0}$
11	0.922	0.109	> 0
12	0.922	0.118	> 0
13	0.922	0.102	> 0
14	0.922	0.064	> 0
15	0.922	0.024	> 0
16	0.922	0.032	> 0
17	0.922	0.081	> 0
18	0.922	0.063	> 0
19	0.922	0.083	> 0
20	0.922	0.075	> 0

Wie bereits in den Beispielen zuvor unterstützt die empirische Dichtefunktion das Testergebnis (vgl. Abbildung 6.6). Jedoch fällt auf, dass der Lyapunov-Exponent der Originaldatenreihe immer zu klein durch die Surrogate geschätzt wird. Hinsichtlich der Testprozedur ist dies nicht von Bedeutung. Es kann aber die Aussage getroffen werden, dass die Verteilung der Surrogate bzgl. des Lyapunov-Exponenten nicht mit der Originalbeobachtungsreihe übereinstimmen wird. Das ist allerdings auch an keiner Stelle gefordert.

Kapitel 7
Identifikation von Rattern

Die Klassifikation von stabilen und instabilen Datenreihen mit Hilfe des Lyapunov-Exponenten wird nun auf ein reales Datenbeispiel angewendet. Hierbei stammen die Reihen aus einem sogenannten *BTA*-Tiefbohren. Die Charakterisierung dieses Prozesses ist von besonderer Relevanz, da vereinzelt unerwünschte Zustände, wie z. B. Rattern, auftreten können. Ziel ist es, den Übergang vom nichtratternden in den ratternden Zustand zu identifizieren. Dies wird mit Hilfe der in dieser Arbeit vorgestellten Klassifikation versucht, indem die gegebene Datenreihe in Zeitfenstern untersucht wird. Zum Vergleich werden andere Methoden, wie z. B. die Berechnung der Autokorrelationsfunktionen, diesem Verfahren gegenübergestellt.

7.1 Darstellung des Tiefbohrproblems

Um lange Metallbohrungen mit hoher Oberflächengüte herstellen zu können, wurde im Maschinenbau das so genannte *Boring and Trespaning Association (BTA)-Tiefbohren* entwickelt (VDI-Richtlinie 3210, 1974). Die hierzu verwendeten Werkzeuge sind sehr schlank, so dass es aufgrund geringer Steifigkeit zu *Rattern* führt. Dieser Zustand verursacht Oberflächenfehler am Werkstück, den so genannten Rattermarken. Diese Rattermarken führen zu hohem Qualitätsverlust am Werkstück. Neben den Riefen im Werkstück geht Rattern mit einer starken Lärmbelästigung einher. Als weitere Ratterfolge ist ein hoher Verschleiß an Schneiden und Führungsleisten, die sich in der Tiefbohrmaschine befinden, zu nennen. Auch eine Gefährdung des Werkstücks bis hin zum Werkzeugbruch ist nicht selten eine Folge von Rattern. Zum anderen wird die Ausnutzung der Antriebsleistung der Tiefbohrmaschine durch ein ansteigendes Bohrmoment reduziert (Weinert, Webber, Busse, Hüsken, Mehnen und Stagge, 2001). Um diesen unerwünschten Zustand vermeiden zu können, sind Kenntnisse über den Bohrprozess und über die Ursachen des Ratterns notwendig.

7.2 Beschreibung der Tiefbohrdaten

Die Daten, die der Analyse zu Grunde liegen, werden in zwei unterschiedliche Bereiche unterteilt (Busse, Hüsken und Stagge, 2001). Einem nichtratternden Zustand folgt ein hochfrequenter Ratterzustand. Bei einer reinen Betrachtung der Zeitreihe lassen sich einige Charakteristiken ausmachen. Zuerst lassen sich zwei Abschnitte mit stark unterschiedlicher Amplitude identifizieren. Die Amplitude des Beschleunigungssensors ist in der ersten Phase um etwa zwei Größenordnungen geringer als in einer zweiten Phase. Mit Hilfe der unterschiedlichen Größenordnungen der Amplituden wird eine Unterteilung des Prozesses in zwei Zustände vorgenommen, einen ratterfreien Zustand und einen Zustand des Ratterns.

Der Mittelwert der Beschleunigungszeitreihe liegt in den Abschnitten nahe bei null. Das passt zu der Tatsache, dass mit der Beschleunigung ein differentieller Wert aufgetragen wird. In der ersten Phase des Bohrprozesses erscheint die Zeitreihe des Beschleunigungssensors als stark verrauscht mit einem sehr geringen periodischen Anteil, wohingegen die Phase des Ratterns ein stark periodisches Verhalten aufweisen. Wird der Übergang vom Nichtrattern zum Rattern genauer betrachtet, so wirkt dieser „wohlgeordnet“ und bezüglich der Amplitudenzunahme „trichterförmig“ (vgl. Abbildung 7.1).

Da etwa 1700 Messdaten eine Bohrkopfumdrehung entsprechen, setzt Rattern

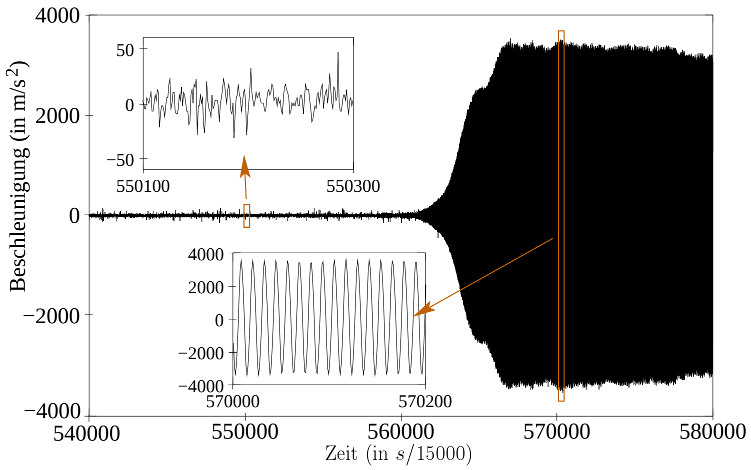

Abbildung 7.1: Beschleunigung des Bohrkopfs im ratterfreien (links) und ratternden (rechts) Zustand.

innerhalb von drei Bohrkopfumdrehungen ein. Während des Ratterns herrscht eine einzige Frequenz der Beschleunigung eindeutig vor, während im ratterfreien Fall keine dominante Schwingung sichtbar ist.

Abbildung 7.2 stellt das Leistungsdichtespektrum in der ratterfreien Phase dar. Schwach erkennbar ergibt sich eine charakteristische Frequenz, die bei etwa 1.2 kHz liegt. Aus Voruntersuchungen (Webber, 1999) ist bekannt, dass diese Frequenz eine der ersten sieben Torsionseigenfrequenzen der Bohrstange ist. Eine niederfrequentere Grundschwingung ist im ratterfreien Bereich nicht zu entdecken.

In dieser Zeitreihe gibt es einen Bereich des Ratterns, für diesen Bereich ist das Leistungsspektrum in Abbildung 7.3 gezeigt. Die Abbildung zeigt hochfrequentes Rattern und bezieht sich auf den zweiten charakteristischen Bereich der Messreihe. Die Grundfrequenz der Abbildung spiegelt deutlich die Torsionseigenfrequenz (1.26 kHz) der Bohrstange wider, wie sie ebenfalls in Abbildung 7.2 sichtbar ist. Zusätzlich sind Oberschwingungen deutlich erkennbar. Mit dem Anstieg der Frequenzen fällt die Stärke der Oberschwingungen ab.

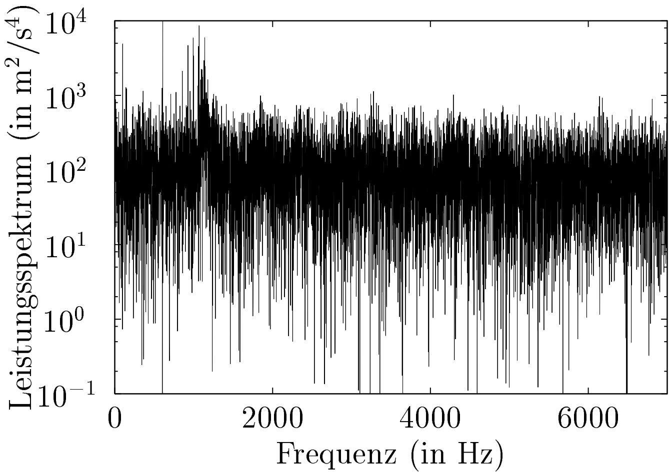

Abbildung 7.2: Frequenzspektrum über den Beschleunigungssensor über 16384 Daten aus dem ratterfreien Bereich (Datenpunkte 100001-116384).

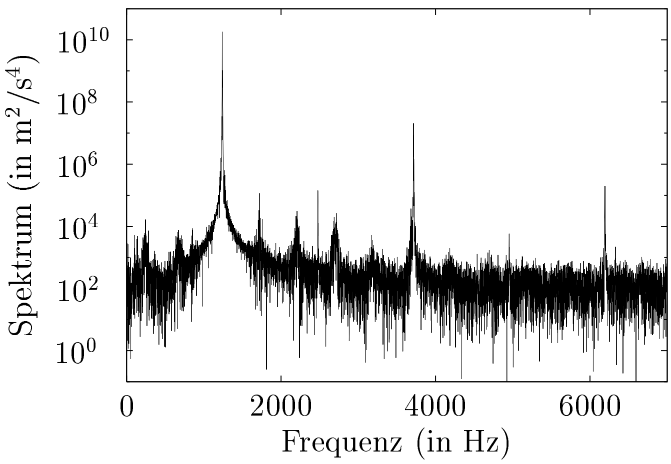

Abbildung 7.3: Frequenzspektrum der Beschleunigung beim hochfrequente Rattern (Datenpunkte 700001-716384)

Aus der Frequenzanalyse des Bohrprozesses lässt sich der Schluss ziehen, dass die hier aufgetretenen charakteristischen Frequenzen die unterschiedlichen Phasen, d. h. Nichtrattern und hochfrequentes Rattern, kennzeichnen. Diese charakteristischen Frequenzen entsprechen den Eigenfrequenzen der Bohrstange.

7.3 Datenanalyse

Die deskriptive Analyse des Bohrprozesses identifizierte zwei unterschiedliche Phasen. Da jedoch das Ziel die frühzeitige Identifikation von Rattern ist, wird der Übergang vom nichtratternden zum ratternden Zustand genauer betrachtet. Hierzu wird in Zeitfenstern die in dieser Arbeit behandelte Klassifikation von stabilen und nichtstabilen Prozessen mit Hilfe des Lyapunov-Exponenten durchgeführt. In Busse, Hüsken und Stagge (2001) wurde bereits eine Identifikation des Übergangs mit Hilfe der Berechnung von Spektren, von Autokorrelationsfunktionen und von Dichten, jeweils auch über gleitende Zeitfenster im Übergangsbereich, versucht. Diese Ergebnisse werden der Analyse der Klassifikationsergebnisse gegenübergestellt (vgl. Kapitel 7.3.2).

7.3.1 Identifikation des Übergangs

Da anhand der Fourieranalyse eine Unterscheidung der charakteristischen Phasen des Bohrprozesses möglich ist, schließt sich die Frage an, wie der Übergangsbereich vom Nichtrattern zum Rattern (vgl. Abbildung 7.1) frühzeitig erkannt werden kann. Hier wird zuerst eine Fourieranalyse in gleitenden Fenstern benutzt (1. Abschnitt). Dem anschließen wird sich eine ebenfalls in gleitenden Fenstern betrachtete Analyse mit der Autokorrelationsfunktion (2. Abschnitt). Dichteschätzungen in den gleitenden Zeitfenstern aus dem Übergangsbereich werden die Analyse zur Beschreibung des Wechsels vom Nichtrattern zum Rattern fortsetzen (3. Abschnitt). Die Berechnung der empirischen Dichtefunktion zu den Lyapunov-Exponenten wird die Charakterisierung des Wechsels der beiden unterschiedlichen Bohrzustände abrunden (4. Abschnitt).

Fouriertransformationen in Fenstern

Um die zeitliche Veränderungen der Fundamentalfrequenzen im Übergangsbereich vom Nichtrattern zum Rattern bestimmen zu können, werden Spektren zu gegebenen Zeitfenstern im Übergangsbereich gebildet. Mit diesem Vorgehen werden zusätzliche Informationen geliefert, wie z. B. die Analyse der Frequenz, die hauptsächlich für das Erregen zum Rattern verantwortlich ist.

Es werden Spektren über jeweils 1024 Messpunkte ausgehend von Messpunkt 550096 berechnet (siehe Abbildung 7.4). Deutlich sind die Veränderungen der Spektren im Übergang zum Rattern zu erkennen. In der noch ratterfreien Phase ist kaum eine charakteristische Frequenz zu erkennen, ähnlich zu Abbildung 7.2. Nach einiger Zeit (spätestens ab dem Fenster mit den Datenpunkten 558288-559312) wird das Aufschwingen einer Frequenz deutlicher, die sich zum Ende, also im ratternden Zustand, wesentlich heraushebt. Offenbar stellt sich Resonanz zwischen Bohrer und Anregung ein. In den letzten Bildern zeigt sich das Erscheinungsbild des Ratterns (vgl. Abbildung 7.3). Diese Fundamentalfrequenz ist die typische Torsionseigenfrequenz, die bei etwa 1.2 kHz liegt. Es ist somit möglich, die Spektren zur Charakterisierung des Übergangs zum Rattern heran zu ziehen.

Autokorrelationsfunktionen in Fenstern

Um weitere Einsichten über die Charakterisierung des Übergangsbereichs zu gewinnen, werden die empirischen Autokorrelationsfunktionen, kurz ACF, im Zeitfenster herangezogen. Hierbei beschreibt die ACF die Korrelation zwischen je zwei Zeitpunkten eines stochastischen Prozesses, jeweils in Abhängigkeit von unterschiedlichen zeitlichen Abständen (Kendall und Ord, 1990), (Brockwell und Davis, 1996).

Um den Wechsel vom nichtratternden zum ratternden Zustand möglicherweise mit Hilfe der ACF beschreiben zu können, werden die verschiedenen ACFs wie bei der Fourieranalyse im Zeitfenster wieder über jeweils 1024 Messpunkte ausgehend von Messpunkt 550096 berechnet (siehe Abbildung 7.5). Sowohl im ratterfreien als auch im ratternden Zustand klingen die ACF nicht ab. Eine Veränderung in der Abfolge der verschiedenen Autokorrelationsfunktionen ist sehr schnell zu

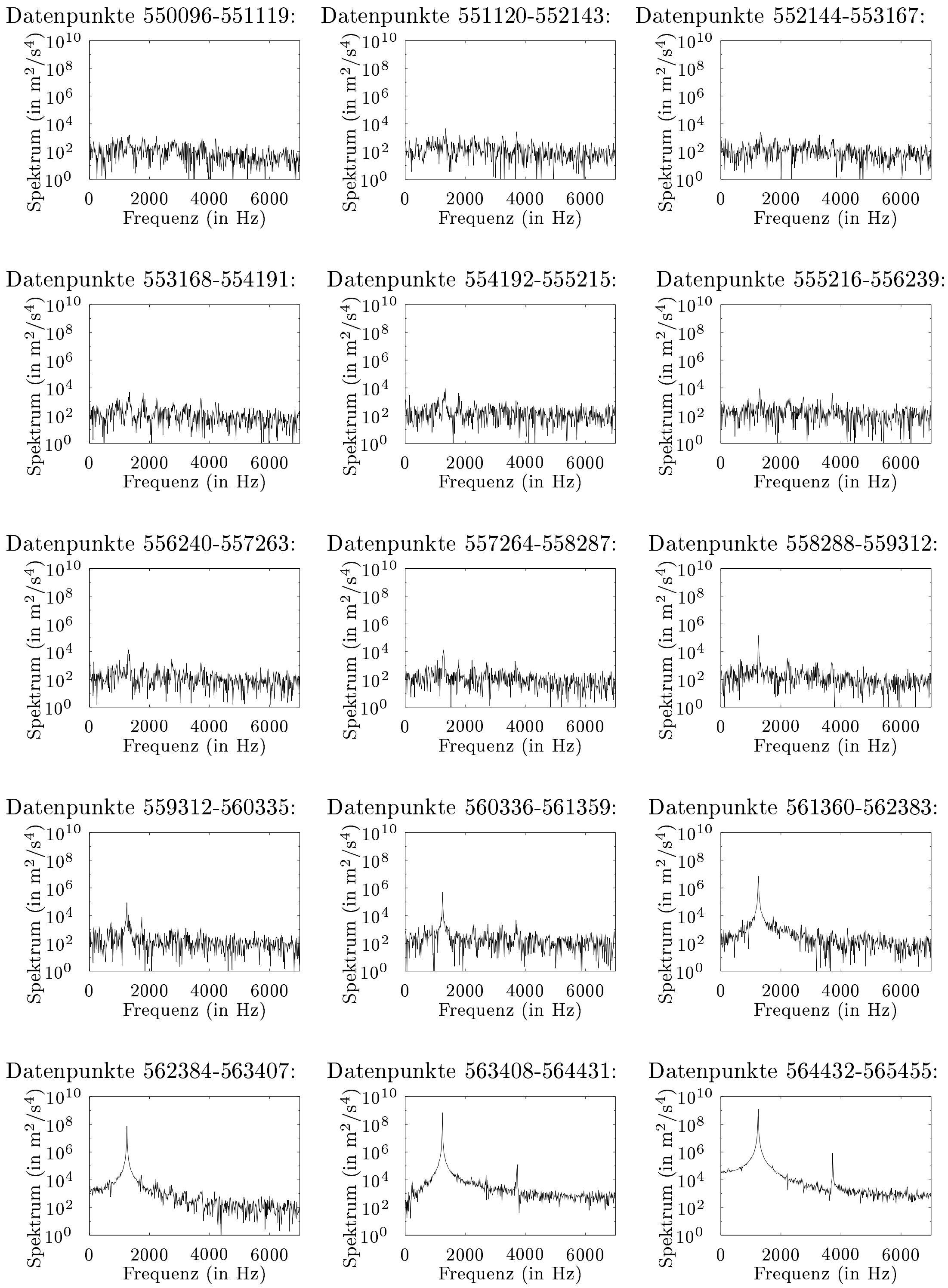

Abbildung 7.4: Spektren über je 1024 Datenpunkten im Bereich des Übergangs zum Rattern.

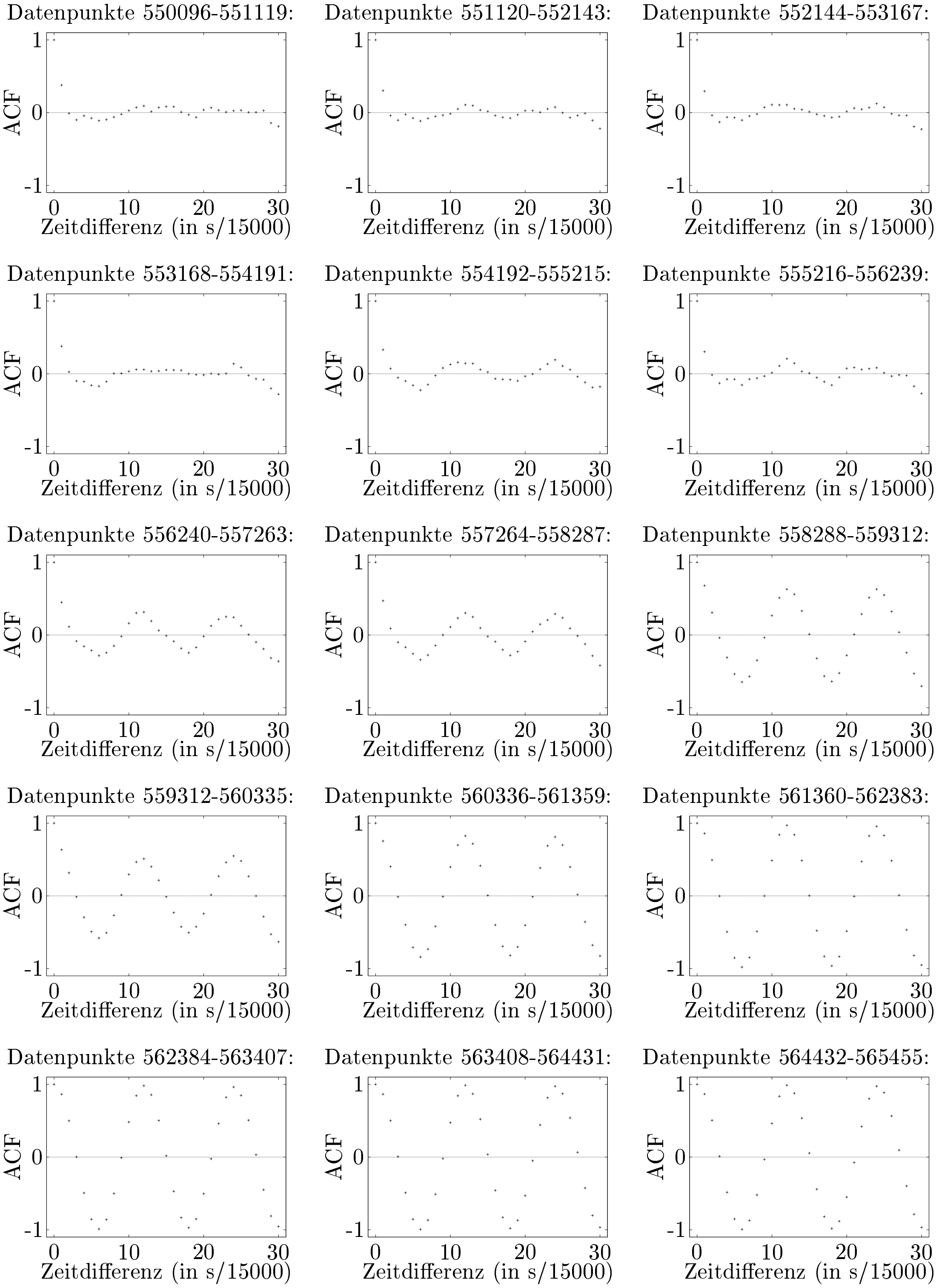

Abbildung 7.5: Autokorrelationsfunktionen über je 1024 Datenpunkten im Bereich des Übergangs zum Rattern.

erkennen. Klare Schwingungen werden deutlich sichtbar. Im Gegensatz zu den vorher untersuchten Spektren (vgl. Kapitel 7.3.1) ist das Rattern anhand der ACF früher erkennbar (Datenfenster 556240-557263).

Dichtefunktionen in Fenstern

Wird der zeitliche Aspekt vernachlässigt, ist eine weitere Untersuchungsmethode des Übergangs vom ratterfreien zum ratternden Zustand in dem untersuchten Bohrprozess die Dichteschätzung. Wieder wird der Übergang ins Rattern in der Zeitreihe des Beschleunigungssensors betrachtet, wie in Abbildung 7.1 gezeigt. Die durch Kerndichteschätzung (Silverman, 1986) ermittelten Dichten wurden aus jeweils 1024 Messpunkten gebildet, ausgehend von Messpunkt 550096 (siehe Abbildung 7.6). Es sei noch darauf hingewiesen, dass die Achsenskalierung aus Betrachtungsgründen unterschiedlich gewählt ist.

Es zeigt sich im Übergang vom Nichtrattern zum Rattern, dass sich die Verteilungen von eingipfligen, symmetrischen Funktionen zu zweigipfligen Dichten verändern. Eine mögliche Ursache ist, dass die Zweigipfligkeit durch eine sinusartige Schwingung zustande kommt. Die Eingipfligkeit ist ein Hinweis darauf, dass eine spätere Modellierung mit linearen Modellen im ratterfreien Zustand möglich sein kann. Eine Art „Bauchigkeit“ zwischen den Modi ist in den letzten Bildern der Abbildung 7.6 zu erkennen. Eine Symmetrie bleibt in allen Bildern erhalten.

Auch mit Hilfe von Dichten in einem Zeitfenster ist der Übergang also charakterisierbar (Datenfenster 560336-561359). Es bleibt festzuhalten, dass die Dichtefunktion im Gegensatz zur Spektrenbetrachtung und zur Betrachtung der ACF erst später zur Identifikation von Rattern neigt.

Lyapunov-Exponenten in Fenstern

Zum Schluss nun soll der Wechsel vom nichtratternden zum ratternden Zustand mit Hilfe des in dieser Arbeit vorgestellten Klassifikationsverfahrens identifiziert

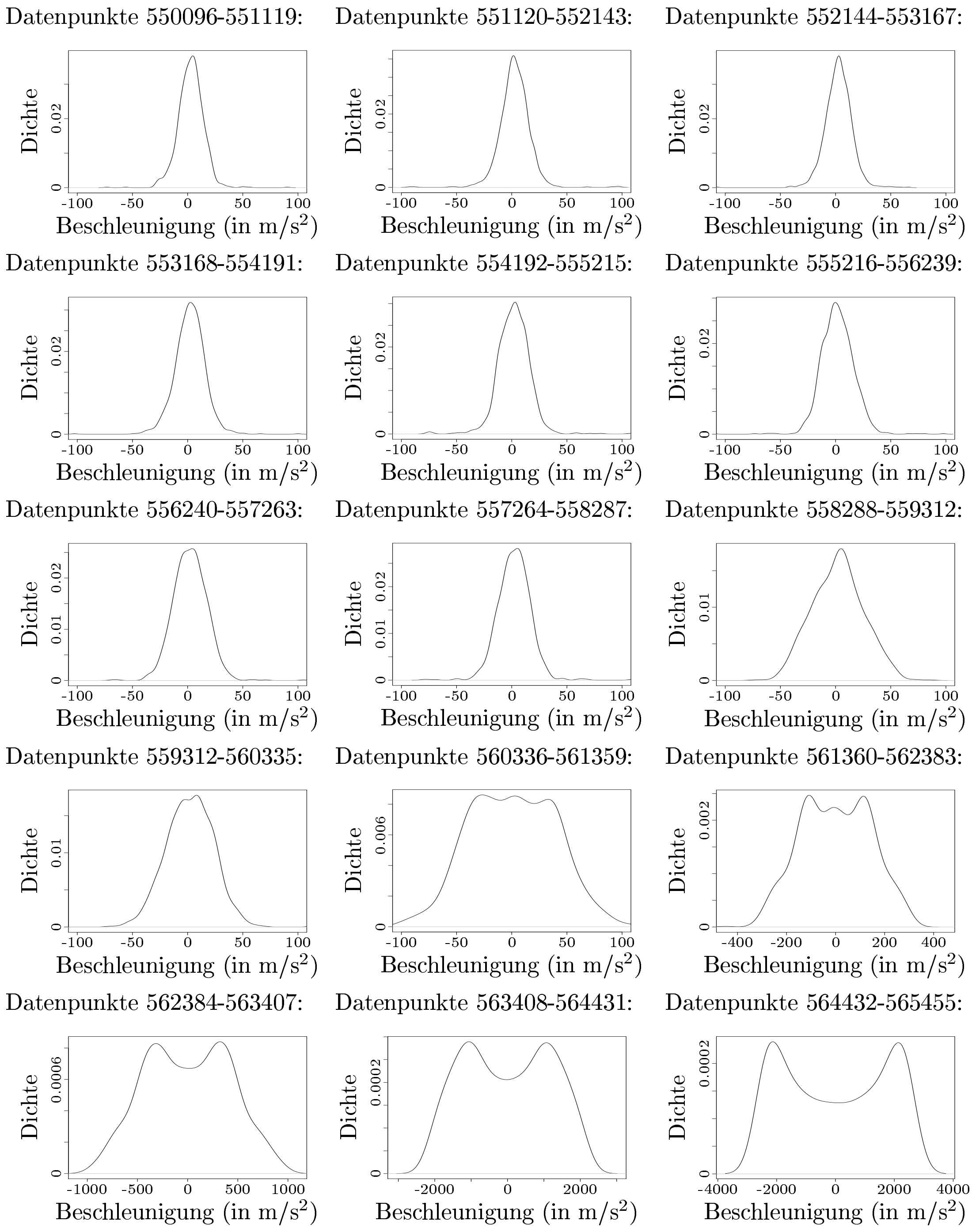

Abbildung 7.6: Geschätzte Dichtefunktionen über je 1024 Datenpunkten im Bereich des Übergangs zum Rattern.

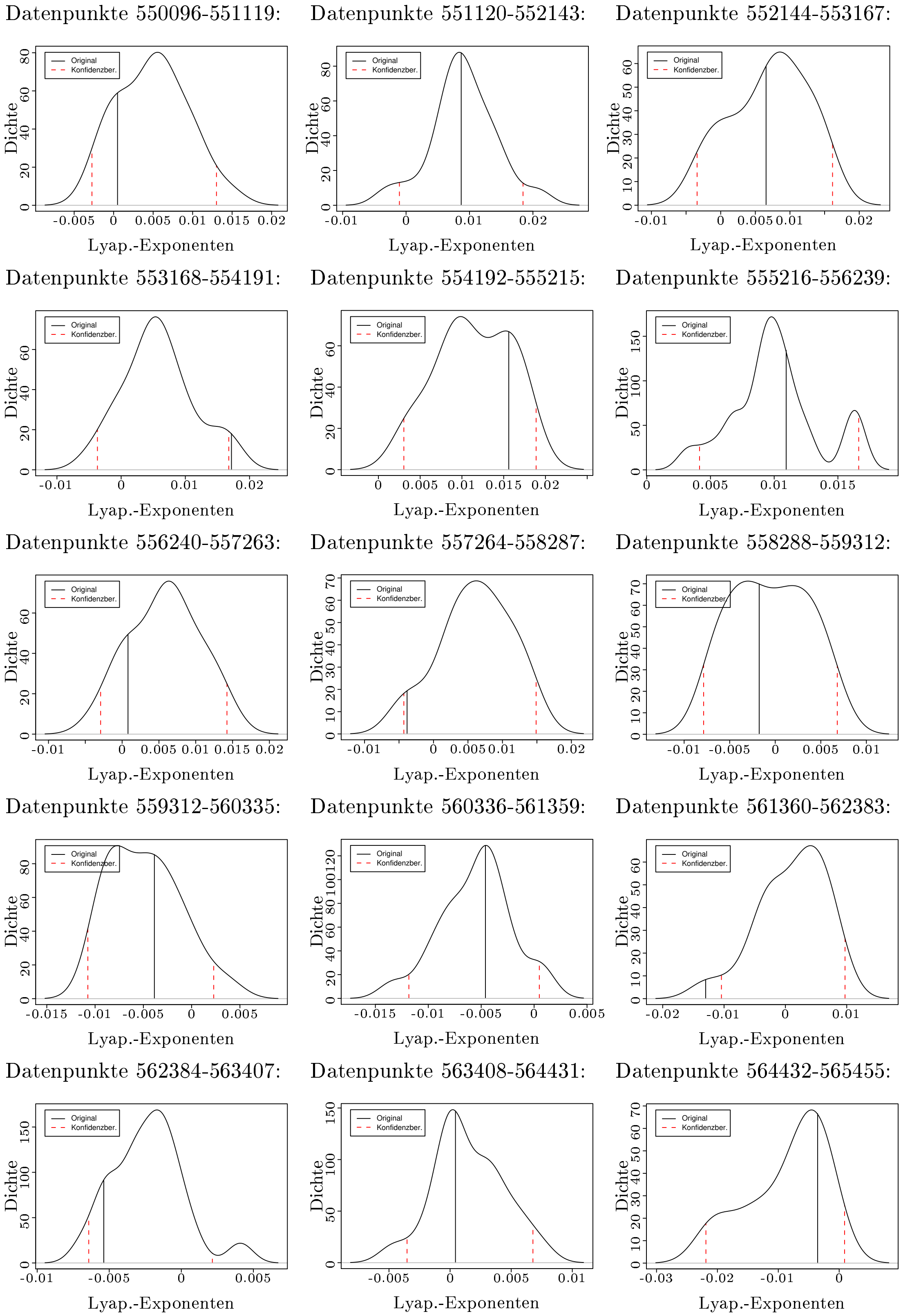

Abbildung 7.7: Empirische Dichtefunktion der Lyapunov-Exponenten über je 1024 Datenpunkten im Bereich des Übergangs zum Rattern.

werden. Hierzu werden wieder Zeitfenster über jeweils 1024 Messpunkten ausgehend vom Messpunkt 550096 zur Analyse herangezogen. Von diesen 1024 Datenpunkten wird zuerst der Lyapunov-Exponent bzgl. der Originalreihe berechnet. Anschließend werden jeweils 20 Surrogate erstellt und die dazu gehörigen Lyapunov-Exponenten berechnet (vgl. Kapitel 6.1). Zur Veranschaulichung wird danach die empirische Dichtefunktion der berechneten Lyapunov-Exponenten für jedes Zeitfenster abgetragen. Die Entscheidung wird dann durch den Original-Lyapunov-Exponenten getroffen, der entweder größer oder kleiner als die Klassifikationsentscheidungsgrenze, d. h. also größer oder kleiner als null, ist.

Es zeigt sich in Abbildung 7.7, dass ab dem Fenster mit den Datenpunkten 557264-558287 der Prozess als stabil, d. h. $\lambda < 0$, klassifiziert wird. Zuvor weisen die Fensterbetrachtungen alle auf jeweils einen instabilen Prozess hin. Lediglich in dem Fenster der Datenpunkte 563408-554431, welches einen instabilen Prozess zeigt, scheint eine Fehlklassifikation aufzutreten, da die umliegenden Fensterprozesse immer als stabil eingestuft werden. Auch wenn die Entscheidung nahe der Klassifikationsgrenze getroffen wurde, muss dieses als falsche Klassifikation angesehen werden.

Es scheint logisch, dass die Prozesse in den Anfangsfenstern als instabil charakterisiert werden, da ein Wechsel der Prozessart zwischen den Zuständen des Nichtratterns und des Ratterns auftritt. Somit ist ebenfalls erklärbar, dass nach dem als Zustandänderung interpretierten Wechsel stabile Prozesse klassifiziert werden. Abbildung 7.1 bestätigt diese Entscheidung, da der ratternde Zustand wie harmonische Schwingungen wirkt.

Es bleibt festzuhalten, dass der Übergang vom Nichtrattern zum Rattern auch durch die in dieser Arbeit vorgestellten Klassifikation zwischen stabilen und instabilen Prozessen identifiziert werden kann. Das Erkennen des Strukturbruchs ist mehr als eine Bohrumdrehung vor dem eigentlichen Amplitudenanstieg sichtbar.

7.3.2 Vergleich der Ergebnisse

Die in den vorhergehenden Kapiteln dargestellten Analysen geben einen Überblick über die Veränderungen beim Übergang zum Rattern. Die durchaus unterschiedlichen Methoden zeigen im Wesentlichen ein einheitliches Bild bezüglich des Zeitpunktes erster Ratteranzeichen. Alle Verfahren weisen den Übergang durch signifikante Veränderungen aus. Besonders bemerkenswert ist, dass diese Modifikationen schon deutlich vor dem eigentlichen Amplitudenanstieg (etwa ab Datenpunkt 560000, vgl. Abbildung 7.1) sichtbar werden. Die ACF kündigt den Übergang etwa 4000 Datenpunkte eher, also mehr als zwei Bohrumdrehungen vor dem eigentlichen Amplitudenanstieg an (vgl. Abbildung 7.5, 7. Bild). Das entspricht einer Eingriffszeit von etwa 0.21 Sekunden. Die Identifikation des Wechsels vom nichtratternden zum ratternden Zustand mit Hilfe des Lyapunov-Exponenten erfolgt ein Zeitfenster später (vgl. Abbildung 7.7, 8. Bild). Das sind etwa eineinhalb Bohrumdrehungen und etwa 0.15 Sekunden eher als der eigentliche Amplitudenanstieg. Diese Identifikation ist ein Zeitfenster schneller als die Fourierspektren (vgl. Abbildung 7.4, 9. Bild). Etwas später sind dann auch die ersten Anzeichen für einen Strukturwechsel den Dichten zu entnehmen (vgl. Abbildung 7.6, 11. Bild). Der Lyapunov-Exponent im Zeitfenster ist also durchaus in der Lage, den Übergang der verschiedenen Bohrzustände zu identifizieren.

Kapitel 8

Zusammenfassung und Ausblick

In der Zeitreihenanalyse tritt oft das Problem auf, nur eine einzige Datenreihe zur Verfügung zu haben, um einen unbekannten zeitlich-funktionalen Zusammenhang zur Beschreibung des Verlaufs der Beobachtungen der Datenreihe, bzw. allgemeiner formuliert, eine geeignete Modellierung zu finden. Notwendig ist, zumindest eine Vorstellung von einer geeigneten Modellklasse zu haben. So ist eine grobe Unterscheidung von Modellklassen in solche für gut oder schlecht prädizierbare Systeme und eine Zuordnung der Datenreihe zu diesen ein wichtiger Schritt auf dem Weg zu einer angemessenen Modellierung.

Es zeigte sich, dass für eine solche Zuordnung bzw. Klassifikation der Lyapunov-Exponent herangezogen werden kann. Ursprünglich war dieser nur für deterministische Beobachtungsreihen definiert. Hier misst der Lyapunov-Exponent die asymptotische, durchschnittliche logarithmische Ausdehnung infinitesimal kleiner Intervalle entlang zweier Trajektorien. Anschaulich bedeutet das, dass zwei Pfade eines Prozesses zusammen laufen, wenn der Lyapunov-Exponent einen Wert kleiner als null hat. Hingegen divergiert der Prozess, wenn dieser größer als null ist.

Nun war ein weiteres Ziel, den Lyapunov-Exponenten auch auf stochastische Zeitreihen zu übertragen. Es konnte gezeigt werden, dass wenn ein additiver Fehler als Zufallseinfluss in der Prozessgleichung angenommen werden kann, ein direktes Überführen der deterministischen Herleitung in die Stochastik möglich ist. Jedoch ist die Annahme des additiven Fehlers nicht immer gerechtfertigt, da sie zu einschränkend bzgl. möglicher Modellklassen ist. So wurde in dieser Arbeit zusätzlich der allgemeinere Fall betrachtet, keine Annahmen zu treffen, wie der Zufall einen funktionalen Zusammenhang zwischen den Beobachtungen stört. Auch hier war es möglich, den Lyapunov-Exponenten für Zeitreihen zu definieren. Der Lyapunov-Exponent konnte auf natürliche Weise als asymptotischer Erwartungswert einer Transformation des gegebenen stochastischen Prozesses verallgemeinert werden. Die zusätzliche Annahme der Ergodizität des Prozesses diente als Grundlage für die spätere Klassifikation und zur Herleitung eines weiteren Schätzers für den Lyapunov-Exponenten.

Mit dieser Grundlage und mit der in dieser Arbeit ebenfalls entwickelten und begründeten Interpretation des Lyapunov-Exponenten als durchschnittlichen Informationsverlust, konnte nun gezeigt werden, dass der Lyapunov-Exponent als Klassifikationskriterium für die Prädizierbarkeit eines Prozesses, der eine gegebene Datenreihe angemessen modelliert, geeignet ist. Ist der Lyapunov-Exponent kleiner als null, verkleinert sich das Informationsgebiet, wo die wahre Position eines zukünftigen Datenpunktes liegt, so dass eine gute Prädiktion möglich ist. Ist der Lyapunov-Exponent hingegen größer als null, geht dieses mit einem Informationsverlust einher und der Prozess ist schlecht zu prädizieren.

Um einen Schritt weiter zu gehen und den Lyapunov-Exponenten als Klassifikationskriterium zwischen stabilen und instabilen Prozessen zu verwenden, war es im Vorfeld notwendig, die Stabilität eines deterministischen Prozesses auf stochastiche Prozesse zu übertragen. Dieses wurde mit der Definition einer asymptotisch stochastischen Stabilität umgesetzt. Hiermit konnte sowohl für deterministische als auch für stochastische Zeitreihen gezeigt werden, dass der Lyapunov-Exponent kleiner als null ist, wenn ein stabiler Prozess vorliegt, wohingegen er größer als null ist, wenn ein instabiler Prozess zu Grunde liegt. So ist es möglich, eine Klassifikation zwischen stabilen und instabilen Prozessen vorzunehmen.

Da bei der Anwendung des Lyapunov-Exponenten der funktionale Zusammenhang der Beobachtungen unbekannt ist, muss der Lyapunov-Exponent aus der gegebenen Datenreihe geschätzt werden, ohne einen solchen Zusammenhang zu modellieren. Hierzu wurde auf ein bereits frei verfügbares Programmpaket zurück gegriffen. Somit war es möglich, durch gezielte Simulationen zum einen spezielle stochastische Prozesse zu klassifizieren, zum anderen Untersuchungen an stabilen und instabilen Prozessen vorzunehmen. Die Überlegungen aus der Theorie wurden in den Simulationen bestätigt. So wurden Prozesse mit autoregressivem Anteil und mit stabilem Grenzverhalten auch als stabile Prozesse identifiziert. Auch die Untersuchungen im kritischen Bereich, d. h. an der Klassifikationsgrenze zwischen stabilem und instabilem Verhalten, zeigten keine Fehlklassifikationen. Hierzu wurde eine Prozessklasse gewählt, die zum einen stabiles als auch instabiles Verhalten je nach Parameterwahl zeigt.

Die Annahme, dass der Lyapunov-Exponent als eine Kenngröße des asymptotischen Verhaltens keine Startwertprobleme bei den numerischen Berechnungen aufweist, wurde ebenfalls in Simulationen bestätigt. Es traten keine zusätzlichen Fehlklassifikationen bei unterschiedlich gewählten Startwerten auf. Auch Untersuchungen mit kurzen Zeitreihen, bei denen sich die Datenreihe möglicherweise noch nicht im asymptotischen Bereich befinden könnte, wurden durchgeführt. Hier zeigte sich, dass bei zu kurzen Datenreihen falsche Einteilungen möglich sind.

Da bei der Verwendung des Lyapunov-Exponenten durchaus Fehlklassifikationen, wie z. B. bei den kurzen Datenreihen, auftreten können, war es ein weiteres Ziel dieser Arbeit, die Klassifikationsentscheidung statistisch abzusichern. Hierzu wurde ein statistischer Test entwickelt, der auf Surrogatreihen beruht. Surrogatreihen sind simulierte Datenreihen, die ein bestmögliches Abbild der Ausgangsreihe darstellen. Hierzu wurde ein bereits implementiertes auf Fouriertransformation beruhendes Verfahren anhand der bereits verwendeten Prozesse mit Hilfe von Simulationen untersucht. Es zeigte sich, dass die erzeugten Surrogate gute

Stellvertreter für die Ursprungsreihen darstellten.

Nun wurde ein nichtparametrischer Test konstruiert, der Aussagen über die Unsicherheit der Klassifikationsentscheidung zwischen gut und schlecht pädizierbaren bzw. stabilen und instabilen Prozessen macht. Er sollte berücksichtigen, dass die Entscheidung mit Hilfe des Lyapunov-Exponenten getroffen wird. Somit gehen die Vorzeichen der berechneten Lyapunov-Exponenten zu der Originaldatenreihe und zu den Surrogaten in die Teststatistik ein. Die Nullhypothese besteht darin, unterhalb einer maximalen Wahrscheinlichkeit von Fehlklassifikationen zu liegen. Der Test wurde mit Hilfe der bereits verwendeten Prozesse experimentell unterstützt. Hierbei zeigte sich, dass der Test auch numerische Mängel aufdeckt. So stellte sich heraus, dass bei zu starkem Rauschen das eigentliche Signal überdeckt wird und vermehrt Fehlklassifikationen auftraten.

Neben der Konstruktion eines Tests ist es mit den Surrogaten auch möglich, weiter gehende empirische Aussagen über den berechneten Lyapunov-Exponenten zu machen.

Zum Schluss der Arbeit wurde die Klassifikation mit Hilfe des Lyapunov-Exponenten auf einen realen Datensatz angewendet. Die Beobachtungen stammen aus dem sogenannten *BTA*-Tiefbohren. Das Ziel war es, das während eines Bohrvorgangs auftretende unerwünschte Rattern, das wie ein chaotisches, instabiles Phänomen erscheint, rechtzeitig zu identifizieren. Hierzu wurden die Klassifikationsentscheidungen durch Lyapunov-Exponenten in Zeitfenstern, die einen Bereich vom Nichtrattern zum Rattern abdecken, betrachtet. Es zeigte sich, dass eine Identifikation durch den Lyapunov-Exponenten ebenso gut funktioniert wie eine Identifikation mit Hilfe von Fourierspektren oder Autokorrelationsbetrachtungen in den Zeitfenstern, die zum Vergleich herangezogen wurden.

Da in dieser Arbeit Simulationen nur zur experimentellen Unterstützung angewendet wurden, ist es durchaus lohnenswert, eine groß angelegte Simulationsstudie für unterschiedliche Prozesstypen durchzuführen. So wäre es auch möglich, nicht nur die weite Klasse der chaotischen Prozesse, die typischerweise instabiles Verhalten zeigen, unter dem Gesichtspunkt des Lyapunov-Exponenten zu charakterisieren, sondern auch die speziellen stochastischen Verfahren. Mit einer einhergehenden Verbesserung der numerischen Verfahren wäre es dann möglich, nicht nur eine grobe Klassifikation durchzuführen, sondern Modellklassen anhand des Lyapunov-Exponenten einzuteilen.

Anhang A

Simulationsprozesse

A.1 Voruntersuchungen zum AR(1)-Prozess

Wird bei einem AR(1)-Prozess der Parameter ϕ_1, der den Zusammenhang zwischen dem zurück liegenden und dem gegenwärtigen Zeitpunkt gewichtet, kleiner oder größer als eins gewählt, so tritt entweder stabiles oder instabiles Verhalten des Prozesses auf.

Um zu überprüfen, ob durch die Berechnung des Lyapunov-Exponenten der Prozess richtig klassifiziert wird, werden aus einem AR(1)-Prozess mit unterschiedlichem ϕ_1 jeweils 1024 Datenpunkte erzeugt. Hierbei liegt der Startwert immer bei $x_0 = 0.699$. In Tabelle A.1 sind die unterschiedlichen ϕ_1 und die dazu gehörigen geschätzten Lyapunov-Exponenten abgetragen.

Bei einer Parameterwahl von $|\phi_1| < 1$ wird aus der Theorie heraus ein stabiler Prozess erzeugt. Damit dieser auch durch den Lyapunov-Exponenten richtig klassifiziert wird, müsste der durch *TISEAN* berechnete Lyapunov-Exponent kleiner als null sein. Wird hingegen das $|\phi_1| > 1$ gewählt, so werden instabile Systeme generiert. Tabelle A.1 zeigt hinsichtlich der Klassifikation von stabilen und instabilen Prozessen durch den Lyapunov-Exponenten nur richtige Ergebnisse. Zudem lässt sich hier eine Systematik vermuten. Je größer der Parameter ϕ_1, der über Stabilität und Instabilität entscheidet, ist, desto größer wird auch der berechnete Lyapunov-Exponent. Der Einfluss des additiven Zufallsfehlers wird also, wie es

Tabelle A.1: Geschätzte Lyapunov-Exponenten zu ausgewählten AR(1)-Prozessen

ϕ_1	Lyap.-Exp.
0.75	-0.0065
0.8	-0.0222
0.9	-0.0683
1.2	0.1987
1.3	0.2622
1.5	0.4052
1.7	0.5321

in Kapitel 3.3 beschrieben ist, durch die Berechnung des Lyapunov-Exponenten als Klassifikationskriteriums heraus gerechnet.

A.2 Voruntersuchungen zum MA(1)-Prozess

Der MA(1)-Prozess wird durch die Wahl seines freien Parameters θ_1 bestimmt, so dass je nach Festlegung unterschiedliche Prozesse generiert werden können. Aus der Theorie heraus sollten die hier untersuchten Parameter immer einen stabilen Prozess erzeugen. Um dieses mit Hilfe des Lyapunov-Exponenten zu überprüfen, werden zu unterschiedlicher Parameterwahl jeweils 1024 Datenpunkte erzeugt, wobei der Startwert jeweils bei $x_0 = 2.4$ liegt. Werden die Prozesse richtig klassifiziert, müssten die berechneten Lyapunov-Exponenten kleiner als null sein. In Tabelle A.2 sind die Lyapunov-Exponenten zu den unterschiedlichen Parametern dargestellt.

Es zeigt sich, dass alle Systeme richtig klassifiziert wurden, da alle Lyapunov-Exponenten auf einen Wert kleiner als null berechnet wurden. Zudem zeigt sich

Tabelle A.2: Geschätzte Lyapunov-Exponenten zu ausgewählten MA(1)-Prozessen

ϕ_1	Lyap.-Exp.
0.09983	-0.0119
0.099983	-0.0060
0.1011	-0.0393
0.1983	-0.0475
0.249	-0.0232
0.301	-0.0376
0.349	-0.0265

im Gegensatz zu den AR-Ergebnissen aus A.1 keine Systematik. Diese liegt vermutlich daran, dass sich ein MA-Prozess hauptsächlich dadurch auszeichnet, dass der voran gegangene zufällige Fehler in den Prozess eingeht.

A.3 Bernoulli-Shift-Abbildung

Ein Beispiel für eine besondere Modulo-1-Abbildung stellt die so genannte Bernoulli-Shift-Abbildung dar (vgl. Jetschke (1989), S. 113, 114 und 117). Sie gehört zu den Prozessen, die den zu deterministischem Chaos führenden Mechanismus beschreiben. Somit weist diese Abbildung instabiles Verhalten auf.

Gegeben sei also die Iteration der Modulo-1-Abbildung:

$$x_{t+1} = f(x_k) \qquad \text{mit } f(x) = \begin{cases} 2x & , \quad x \in [0, \frac{1}{2}) \\ 2x - 1 & , \quad x \in [\frac{1}{2}, 1]. \end{cases} \tag{A.1}$$

Die Wirkung der Abbildung lässt sich besser verstehen, wenn x in Dualdarstellung geschrieben wird,

$$x_0 = 0.a_1a_2a_3\ldots := \sum_{i=1}^{\infty} a_i 2^{-i}, a_i \in \{0,1\}. \tag{A.2}$$

Die Iterationen der Gleichung (A.1) bewirkt dann wegen

$$f(x_0) = 0.a_2a_3a_4 \tag{A.3}$$

ein Streichen der ersten Ziffer und anschließende Linksverschiebung um eine Stelle (Bernoulli-Verschiebung). Anschaulich wird bei der Iteration das Intervall $[0,1]$ um den Faktor 2 gestreckt, zerschnitten und übereinander gelegt. Daraus resultiert eine sensitive Abhängigkeit der Trajektorien von den Anfangswerten: Sind also zwei Startwerte x_{01} und x_{02} nach der $(n-1)$-ten dualen Ziffer verschieden, so wird der Unterschied durch die Iteration vergrößert. Die nächsten Zustände $f^n(x_{01})$ und $f^n(x_{02})$ haben nichts mehr gemeinsam. Da reale Messungen mit nur endlicher Genauigkeit durchgeführt werden, wird eine Langzeitprognose unmöglich, was eine Eigenschaft von Chaos darstellt. Dieses spiegelt sich auch in der Berechnung des Lyapunov-Exponenten wider. Der Betrag der ersten Ableitung dieser Modulo-1-Abbildung (vgl. Gleichung (A.1)) ist gegeben durch $f'(x) = 2$ für alle $x \in [0,1]$ (für $x = \frac{1}{2}$ nur links- bzw. rechtsseitig). Daher beträgt der Lyapunov-Exponent (vgl. Gleichung (3.1)):

$$\lambda(x) = \ln 2 \qquad \forall x \in [0,1]. \tag{A.4}$$

A.4 Bestimmung der Parameter für die Modulo-1-Abbildung

Die Modulo-1-Abbildung eignet sich als Simulationsbeispiel von stabilen und instabilen Prozessen sehr gut (vgl. Kapitel 4.2.2). Bei der Untersuchung der stabilen Varianten, die durch die Modulo-1-Abbildung generiert werden, wird ein kleiner Zufallsfehler zur Modellgleichung hinzu addiert. Somit kommen zwei weitere frei

Tabelle A.3: Geschätzte Lyapunov-Exponenten zu unterschiedlichen Gewichtungsparametern des Zufallseinflusses bei der Modulo-1-Abbildung

z	berechneter Lyap.-Exp.	wahrer Lyap.-Exp.
0.03	-0.04042	-0.11
0.04	-0.03060	-0.11
0.05	-0.09450	-0.11
0.1	-0.03770	-0.11
0.5	-0.02024	-0.11
0.9	-0.00220	-0.11

wählbare Parameter hinzu, zum einen der Gewichtsparameter, der den Einfluss der Zufallszahl bestimmt, zum anderen die Verteilung, aus der die Zufallszahl stammt.

Um für das weitere Vorgehen diese Parameter festzusetzen, werden Voruntersuchungen bezüglich ihres Einflusses auf das Klassifikationsergebnis durchgeführt.

Hierzu werden jeweils zu unterschiedlichen Parametereinstellungen immer 1024 Datenpunkte aus der Modulo-1-Abbildung generiert. Exemplarisch seien in Tabelle A.3 einige Ergebnisse bei der Variation des Gewichtsparameters z des Zufallseinflusses in der Modulo-1-Abbildung dargestellt.

Es zeigt sich, dass $z = 0.05$ in diesem Beispiel eine gute Wahl darstellt.

Bei der Wahl der Verteilung der hinzu addierten Zufallszahl erscheint zunächst die Gleichverteilung auf $[0, 1]$ ein guter Vertreter zu sein, da die in dieser Arbeit gewählte Modulo-1-Abbildung eine Abbildung des Einheitsintervalls $[0, 1]$ auf sich selbst ist. Jedoch werden noch weitere Verteilungen aus der Gleichverteilungsfamilie und aus der Familie der Normalverteilung ausprobiert. Einige Ergebnisse sind in Tabelle A.4 dargestellt.

Der erste Eindruck, die Gleichverteilung auf $[0, 1]$ zu verwenden, wird in den

Tabelle A.4: Geschätzte Lyapunov-Exponenten zu unterschiedlichen Verteilungen des Zufallseinflusses bei der Modulo-1-Abbildung

Verteilung	berechneter Lyap.-Exp.	wahrer Lyap.-Exp.
U[0,1]	-0.0945	-0.11
U[-1,1]	-1.2764	-0.11
N(0.5,0.5)	-0.8473	-0.11
N(0,1)	-1.1152	-0.11
N(0,10)	-0.0013	-0.11

Experimenten bestätigt. Somit wird für die in dieser Arbeit verwendeten Zufallszahlen bei der Modulo-1-Abbildung diese Verteilung verwendet. Zudem wird aus Tabelle A.4 deutlich, dass wenn die Varianz des Fehlers zu hoch gewählt wird, das numerische Verfahren zur Berechnung des Lyapunov-Exponenten diesen nicht mehr gut schätzt.

Anhang B

AAFT-Algorithmus

Bei der Erzeugung der in dieser Arbeit verwendeten Surrogate sollen die Kriterien erfüllt sein, dass die Originaldatenreihe und die Ersatzdatenreihen zum einen eine gleiche Autokorrelationsfunktion besitzen, zum anderen der gleichen Verteilung folgen. Die Erzeugung von Surrogaten mit gleicher Autokorrelationsfunktion wie die Originaldatenreihe wird mit Hilfe des von Theiler, Galdrikian, Longtin, Eubank und Farmer (1992) vorgestellten AAFT-Algorithmus durchgeführt, der im Folgenden näher beschrieben wird. Vergleiche hierzu Theiler, Galdrikian, Longtin, Eubank und Farmer (1992), S. 183, 184 und Theiler, Eubank, Longtin, Galdrikian und Farmer (1992).

Beschreibung zur Generierung von Surrogaten mit gleicher Autokorrelationsfunktion:

1. Eingabe der Datenpunkte in ein Array, $x[t]$, mit $t = 1, \ldots, N$.

2. Sortierung des Arrays der Reihe nach, $S_x[k]$, mit $k = 0, \ldots, N-1$.

3. Bilde Ränge über die Datenreihen, $R_x[t]$, so dass sie der Anforderung $S_x[R_x[t]] = x[t]$ genügt.
   Bemerke, dass $S_x[k]$ eine monotone Funktion mit einer wohldefinierten Inversen ist. Somit ist $R_x[t] = S_x^{-1}[t]$ eine statische Zurückskalierung von $x[t]$.

4. Erzeuge einen zufällig gaußverteilten Datensatz, $g[t]$, mit $t = 1, \ldots, N$.

5. Sortiere die gaußverteilten Zufallszahlen $S_g[k]$, mit $k = 1, \ldots, N$.

6. Definiere ein neue Zeitreihe: $y[t] = S_g[R_x[t]]$.
   Die neue Zeitreihe ist eine statische Zurückskalierung von $x[t]$ mit der Eigenschaft, dass die Amplitudenverteilung gaußverteilt ist.

7. Berechne die diskrete Fouriertransformation (DFT): $z[t] = DFTy[t]$.
   Bemerke, dass $z[t]$ eine komplexwertige Darstellung mit Real- und Imaginärteilen ist. Die komplexwertige Fouriertransformation splittet Amplitude und Phase auf.

8. Randomisiere die Phase: $\tilde{z}[t] = z[t]e^{i\phi[t]}$, mit $\phi[t] \sim U[0, 2\pi]$. Hierbei wird die Phase zufällig verschoben, wobei die Amplitude beibehalten wird (um die gleiche Autokorrelationsfunktion zu behalten).

9. Symmetrisiere die Phase, damit die inverse Fouriertransformation wieder real ist:
   $\mathrm{Re}\,\tilde{\tilde{z}}[t] = \mathrm{Re}\,(\tilde{z}[t] + \tilde{z}[N+1-t])/2,$
   $\mathrm{Im}\,\tilde{\tilde{z}}[t] = \mathrm{Im}\,(\tilde{z}[t] - \tilde{z}[N+1-t])/2.$

10. Invertiere die diskrete Fouriertransformation: $\tilde{y} = DFT^{-1}\tilde{\tilde{z}}[t]$.

11. Erstelle eine Rangzeitreihe für $\tilde{y}[t]$, $R_{\tilde{y}}[t]$.

12. Die Ersatzdatenreihe, Surrogat, ist dann gegeben durch:
    $\tilde{x}[t] = S_x[R_{\tilde{y}}[t]]$.

Literatur

Abarbanel, H. D. I. (1996). *Analysis of Observed Chaotic Data.* Institute for Nonlinear Science. New York: Springer Verlag.

Abarbanel, H. D. I., R. Brown, J. J. Sidorowich und L. S. Tsimring (1993). The analysis of observed chaotic data in physical systems. *Reviews of Modern Physics 65*, 1331–1392.

Arnold, L. (1998). *Random Dynamical Sytems.* Berlin Heidelberg New York: Springer Verlag.

Arnold, L., V. M. Gundlach und L. Demetrius (1994). Evolutionary formalism for products of positive random matrices. *Annals of Applied Probability 4*, 859–901.

Arnold, L. und M. Scheutzow (1995). Perfect cocycles through stochastic differential equations. *Probability Theory and Related Fields 101*, 65–88.

Arnold, V. I. und A. Avez (1968). *Ergodic problems of classical mechanics.* New York: W. A. Benjamin.

Arrowsmith, D. K. und C. M. Place (1990). *An Introduction to Dynamical Systems.* Cambridge: Cambridge University Press.

Ash, R. B. und C. A. Doleans-Dade (2000). *Probability and Measure Theory.* New York: Academic Press.

Ballard, D. H. (1997). *An Introduction to Natural Computation.* MIT Press.

Beck, C. und F. Schlögl (1993). *Thermodynamics of chaotic systems.* Cambridge: Cambridge University Press.

Birkhoff, G. D. (1931). Proof of the ergodic theorem. *Proc. Nat. Acad. Sci. 17*, 656–660.

Blatter, C. (1980). *Analysis I* (3 Aufl.). Berlin: Springer Verlag.

Bougerol, P. und N. Picard (1992). Stationarity of GARCH processes and of some nonnegative time series. *Journal of Econometrics 52*, 115–127.

Brockwell, P. J. und R. A. Davis (1996). *Introduction to Time Series and Forecasting.* New York: Springer.

Büning, H. und G. Trenkler (1978). *Nichtparametrische statistische Methoden.* Berlin, New York: de Gruyter.

Busse, A. M., M. Hüsken und P. Stagge (2001). Offline-Analyse eines BTA-Tiefbohrprozesses. Technischer Bericht 16, SFB 475, Fachbereich Statistik, Universität Dortmund, 44221 Dortmund, Germany.

Busse, A. M., D. Steuer und C. Weihs (2001). An Approach for the Determination of Predictable Time Series. Technischer Bericht 12, SFB 475, Fachbereich Statistik, Universität Dortmund, 44221 Dortmund, Germany.

Buzug, T. (1994). *Analyse chaotischer Systeme.* Mannheim, Leibzig, Wien: Wissenschaftsverlag BI.

Casdagli, M. (1991). Chaos and deterministic versus stochastic non-linear modeling. *J. R. Statist. Soc. B 54*, 303–328.

Eckmann, J. P., S. O. Kamphorst, D. Ruelle und S. Ciliberto (1986). Lyapunov exponents from a time series. *Physical Review A 34*(6), 4971–4979.

Eckmann, J. P. und D. Ruelle (1982). Ergodic theory of chaos and strange attractors. *Reviews of Modern Physics 57*, 617–657.

Fahrmeir, L., H. L. Kaufmann und F. Ost (1981). *Stochastische Prozesse. Eine Einführung in Theorie und Anwendungen.* Wien: Carl Hanser Verlag.

Friedman, N. A. (1970). *Introduction to ergodic theory.* New York: Van Nostrand Reinhold Company.

Gardner, G., A. Harvey und G. Phillips (1980). Algorithm AS 154. An algorithm for exact maximum likelihood estimation of autoregressive-moving average models by means of kalman filtering. *Applied Statistics 29*, 311–322.

Gencay, R. (1996). A statistical framework for testing chaotic dynamics via lyapunov exponents. *Physica D 89*, 261–266.

Grassberger, P. und I. Procaccia (1983a). Characterization of strange attractors. *Physical Review Letters 50*, 346–349.

Grassberger, P. und I. Procaccia (1983b). Measuring the strangeness of strange attractors. *Physica D 9*, 189–208.

Hastie, T., R. Tibshirani und J. Friedman (2001). *The Elements of Statistical Learning. Data Mining, Inference, and Prediction.* New York, Berlin, Heidelberg: Springer.

Hegger, R., H. Kantz und T. Schreiber (1999). Practical implementation of nonlinear time series methods: The TISEAN package. *CHAOS 9*, 413–435.

Ihaka, R. und R. Gentleman (1996). R: A Language for Data Analysis and Graphics. *JCGS-Journal of Computational and Graphical Statistics 5*(3), 299–314.

Jetschke, G. (1989). *Mathematik der Selbstorganisation.* Thun und Frankfurt (Main): Verlag Harri Deutsch.

Kantz, H. und T. Schreiber (1997). *Nonlinear time series analysis.* Cambridge: Cambridge University Press.

Kendall, M. und J. K. Ord (1990). *Time Series, Third Edition.* Suffolk: Edward Arnold.

Kreutz, M. (2000). *Modellierung von unvollständig beschriebenen Systemen.* Stuttgart: ibidem-Verlag.

Morgenstern, D. (1980). Berechnung des maximalen Signifikanzniveaus des Tests: Lehne H_0 ab, wenn k unter n gegebenen Tests zur Ablehnung führen. *Metrika 27*, 285–286.

Padberg, F. (1996). *Elementare Zahlentheorie* (2 Aufl.). Spektrum Akademischer Verlag GmbH.

Paluš, M. (1994). Identifying and quantifying chaos by using information-theoretic functionals. In A. S. Weigend und N. A. Gershenfield (Hrsg.), *Time Series Prediction: Forecasting the Future and Understanding the Past, Proc. of the NATO Advanced Research Workshop on Comparative Time Series Analysis held in Santa Fe, New Mexico, May 14-17, 1992*, Band Proceedings Volume XV aus *Santa Fe Institute Studies in the Sciences of Complexity*, Reading, MA, S. 387–413. Addison Wesley.

Park, J. Y. und Y.-J. Whang (1999). Random walk or chaos: A formal test on the Lyapunov exponent. Working Paper Series 9, Institute of Economic Research, Seoul National University.

Prichard, D. und J. Theiler (1994). Generating surrogate data for time series with several simultaneously measured variables. *Physical Review Letters 73*, 951–954.

Reimetz, A. M. (1998, January). Strukturbestimmung von probabilistischen neuronalen Netzen mit Hilfe von Evolutionären Algorithmen. Master's thesis (Diplomarbeit), Fachbereich Statistik, Universität Dortmund.

Rinne, H. (1997). *Taschenbuch der Statistik*, Band 2. Thun und Frankfurt am Main: Verlag Harri Deutsch.

Rohling, H. und T. May (1999). Informations- und Codierungstheorie. In P. Rechenberg und G. Pomberger (Hrsg.), *Informatik-Handbuch* (2 Aufl.), Kapitel B2, S. 191–215. München; Wien: Hanser.

Rosenstein, M., J. Collins und C. DeLuca (1993). A practical method for calculating largest lyapunov exponents from small data sets. *Physica D 65*, 117–125.

Rüger, B. (1978). Das maximale Signifikanzniveaus des Tests: Lehne H_0 ab, wenn k unter n gegebenen Tests zur Ablehnung führen. *Metrika 25*, 171–178.

Sano, M. und Y. Sawada (1985). Measurements of the lyapunov spectrum from a chaotic time series. *Physical Review Letters 55*, 1082–1085.

Schlittgen, R. und B. Streitberg (1994). *Zeitreihenanalyse* (5 Aufl.). Lehr– und Handbücher der Statistik. München, Germany: R. Oldenbourg GmbH.

Schmitz, A. und T. Schreiber (1998). Testing for nonlinearity in unevenly sampled time series. *Physical Review E 59*, 4044–4048.

Schmitz, A. und T. Schreiber (2000). Surrogate data for non-stationary signals. In K. Lehnertz, C. Elger, J. Arnhold und P. Grassberger (Hrsg.), *Proceedings "Chaos in brain?"*, Singapore. World Scientific.

Schreiber, T. und A. Schmitz (1996). Improved surrogate data for nonlinearity tests. *Physical Review Letters 4*, 635–638.

Schreiber, T. und A. Schmitz (2000). Surrogate time series. *Physica D 142*, 346–373.

Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis.* Monographs on Statistics and Applied Probability. London, New York: Chapman and Hall.

Stagge, P. (2000). *Strukturoptimierung rückgekoppelter neuronaler Netze.* Stuttgart: ibidem-Verlag.

Steinhausen, D. und K. Langer (1977). *Clusteranalyse. Einführung in Methoden* *und Verfahren der automatischen Klassifikation.* Berlin, New York: Walter de Gryter.

Stier, W. (2001). *Methoden der Zeitreihenanalyse.* Berlin, Heidelberg, New York: Springer.

Stout, W. F. (1974). *Almost sure convergence.* New York: Academic Press.

Theiler, J., S. Eubank, A. Longtin, B. Galdrikian und J. D. Farmer (1992, March). Testing for nonlinearity in time series: the method of surrogate data. *Physica D 58*, 77–94.

Theiler, J., B. Galdrikian, A. Longtin, S. Eubank und J. D. Farmer (1992). Using surrogate data to detect nonlinearity in time series. In M. Casdagli und S. Eubank (Hrsg.), *Nonlinear Modeling and Forecasting*, New York, S. 163–188. Adison-Wesley.

Theiler, J. und D. Prichard (1996). Constrained-realization Monte-Carlo method for hypothesis testing. *Physica D 94*, 221–240.

Tong, H. (1993). *Non-linear Time Series, A Dynamical System Approach.* Oxford Statistical Science Series. New York: Oxford University Press.

Tsoi, A. C. und A. Back (1997). Discrete time recurrent neural network architectures: A unifying review. *Neurocomputing 15*(4), 184–223.

VDI-Richtlinie 3210 (1974, Juni). Tiefbohrverfahren. VDI Düsseldorf.

Webber, O. (1999). Meßtechnische Erfassung und Analyse dynamischer Instabilitäten des BTA-Tiefbohrprozesses. Diplomarbeit, Institut für Spanende Fertigung, Fakultät für Maschinenbau, Universität Dortmund, 44227 Dortmund, Deutschland.

Weinert, K., O. Webber, A. Busse, M. Hüsken, J. Mehnen und P. Stagge (2001). In die Tiefe: Koordinierter Einsatz von Sensorik und Statistik zur Analyse und Modellierung von BTA-Tiefbohrprozessen. *Zeitschrift für wirtschaftlichen Fabrikbetrieb 96*(5), 262–265.

Wiggings, S. (1990). *Introduction to Applied Nonlinear Dynamical Systems and Chaos.* New York: Springer-Verlag.

Wolf, A., J. B. Swinney und H. L. Vastan (1985). Determining Lyapunov exponents from a time series. *Physika D 16*, 285–317.