

Datengetriebene Modelle stochastischer dynamischer Systeme

am Beispiel hydrodynamischer Turbulenz

Dissertation

zur Erlangung des Doktorgrades
der Fakultät der Naturwissenschaften (Dr. rer. nat.)
der Bergischen Universität Gesamthochschule Wuppertal

WUB-DIS 2001-13

vorgelegt von

Mario Ragwitz

geboren am 05.04.1972 in Cottbus

Oktober 2001

erstellt am

MAX-PLANCK-INSTITUT FÜR PHYSIK KOMPLEXER SYSTEME
DRESDEN

Abstract

This thesis deals with the development of data driven models for stochastic dynamical systems based on measured time sequences. Reconstruction of deterministic dynamical systems from time series is a well established technique, whereas only very few methods exist in the case of a nonlinear stochastic setting. The need for such algorithms arises in a variety of fields, such as physics, biology, economics, or meteorology. First we derive a method to construct Fokker-Planck- and Langevin-equations from time series, if all relevant dynamical variables are measured. The latter assumption, however, is unrealistic in many applications. Therefore we discuss, whether it is possible to embed a scalar time series in a stochastic setting and, if so, how to choose the correct parameters. This leads us to novel time series predictors for Markovian-processes. The performance of these algorithms is demonstrated by numerical examples.

In the second part of this work we apply the techniques developed in the first section to hydrodynamic turbulence and the problem of predicting the fluctuations in wind energy production. For a hydrodynamic flow with pronounced coherent structures, nonlinear phase space methods prove to have significantly higher predictive power within such structures than ordinary linear schemes. This knowledge is then applied to time series of atmospheric surface wind velocities and of the power output of a wind turbine. In these cases as well, strong turbulent fluctuations are clearly better predicted by nonlinear methods than by linear ones. This fact has relevant applications for the control of rotor blades of wind turbines as well as for the regulation of conventional power supplies.

Inhaltsverzeichnis

1	Einleitung	1
2	Datengetriebene Modelle dynamischer Systeme	5
2.1	Datenanalyse deterministischer dynamischer Systeme	5
2.1.1	Dynamische Systeme	6
2.1.2	Beispiele für dynamische Systeme	8
2.1.3	Einbettungsmethoden für dynamische Systeme	10
2.1.4	Vorhersage und Modellierung für Zeitreihen	13
2.2	Datenanalyse stochastischer dynamischer Systeme	17
2.2.1	Langevin- und Fokker-Planck-Gleichungen von Zeitreihen	18
2.2.2	Markov-Modelle von einer Zeitreihe	28
2.2.3	Ist eine skalare Beobachtung einer multivariaten Langevin-Gleichung Markovisch?	29
2.2.4	Einbettungsparameter	38
2.2.5	Nichtlineare Fluktuationen in stochastischen Systemen	41
2.2.6	Nichtlineare Meßfunktion	44
2.3	Diskussion	45
3	Datenanalyse voll entwickelter Turbulenz	47
3.1	Phänomenologie voll entwickelter Turbulenz	48
3.2	Experiment	54
3.3	Allgemeine statistische Eigenschaften	57
3.4	Deterministische Struktur	62
3.4.1	Abhängigkeit von Parametern	64
3.4.2	Skalenverhalten der PDF	67
3.4.3	Determinismus	69
3.4.4	Diskussion	71
3.5	Fokker-Planck-Methode	72
4	Vorhersagen des Oberflächenwindes	77
4.1	Möglichkeiten der Vorhersage	78

4.2	Zeitreihen und lineare Statistik	80
4.3	Nichtlineare Struktur in Bodenwindgeschwindigkeiten	87
4.4	Multivariate Analyse	92
4.5	Nichtlineare Struktur von Leistungszeitreihen	93
4.6	Diskussion	99
5	Zusammenfassung	101

Kapitel 1

Einleitung

Eine Vielzahl von Phänomenen in unserer Umwelt kann durch die Analyse einer zeitlichen Folge von Messungen oder Beobachtungen untersucht und klassifiziert werden. Diese Beobachtungsfolgen, welche auch Zeitreihen genannt werden, enthalten sehr oft einen wesentlichen Teil der zur Verfügung stehenden Information über das betreffende System. Die Analyse von Zeitreihen ist in einer Vielzahl von Disziplinen, wie z.B. der Meteorologie, der Medizin oder der Finanzmathematik, von Bedeutung. So zählen Elektrokardiogramme, die tägliche Durchschnittstemperatur an einem Ort oder der Umtauschkurs verschiedener Währungen zu den populärsten Vertretern untersuchter Zeitreihen.

Beobachtete Zeitreihen weisen häufig ein sehr unregelmäßiges und scheinbar unvorhersagbares Verhalten auf. Das Hauptanliegen der Zeitreihenanalyse ist es, die wesentlichen Eigenschaften dieser Unregelmäßigkeit zu erfassen und das Wissen über den verborgenen physikalischen Prozeß bzw. die Dynamik zu erweitern. Die Modellierung von Zeitreihen durch einen linearen stochastischen Prozeß ist in der statistischen Literatur wohletabliert, und es existiert eine Vielzahl von Verfahren zur Bestimmung von Modellparametern, zur Vorhersage und Kontrolle in linearen Systemen [1]. Insbesondere haben lineare Zeitreihenmodelle mit Gaußschen Inkrementen wegen ihrer einfachen analytischen Eigenschaften starke Beachtung erfahren.

Seit einiger Zeit, insbesondere seit den Arbeiten von Edward Lorenz, ist bekannt, daß auch deterministische dynamische Systeme sehr unregelmäßiges und scheinbar unvorhersagbares Verhalten erzeugen können. Für diese Systeme können Differentialgleichungen oder Abbildungen als Entwicklungsregeln angegeben werden, sie sind daher deterministisch. Dieses Phänomen, welches heute als Chaos bezeichnet wird und im Prinzip schon Poincaré zu Beginn des vergangenen Jahrhunderts bekannt war, erreichte starke Popularität durch die Entwicklung schneller Computer während der letzten Dekaden. Die Chaos-Theorie, auch Theorie dynamischer Systeme genannt, brachte eine Vielzahl leistungsfähiger Methoden zur Analyse deterministischer Zeitreihen hervor. Die Grundlage nahezu sämtlicher Methoden bildet das Takens-Theorem [2], welches er-

laubt, den Zustandsraum des Systems auf der Grundlage einer gemessenen Zeitreihe zu rekonstruieren. Darauf aufbauend, wurden Verfahren zur Abschätzung der Anzahl der aktiven Freiheitsgrade, zur Vorhersage chaotischer Zeitreihen oder zur Chaoskontrolle entwickelt. Außerdem existieren umfangreiche Arbeiten zur Rauschunterdrückung in chaotischen Systemen, welche sich dem Einfluß von Meßrauschen widmen, siehe [3]. Auch in der statistischen Literatur wurde die Bedeutung von Nichtlinearität in den vergangenen Jahren hervorgehoben, wobei besonders die Arbeiten Tongs [4] erwähnenswert sind.

Für viele realistische Zeitreihen von biologischen, meteorologischen und auch physikalischen Systemen ist jedoch offensichtlich, daß weder das Konzept des linearen stochastischen Prozesses noch das des deterministischen Chaos plus Beobachtungsrauschen eine zufriedenstellende Beschreibung darstellt. Entweder besitzen diese Systeme zu viele Freiheitsgrade, um durch skalare Zeitreihen aufgelöst zu werden, oder die deterministische Entwicklung einiger makroskopischer Freiheitsgrade wird durch Rauschen, welches von anderen Freiheitsgraden herrührt, angetrieben. Daher ist häufig ein nichtlinear stochastisches Verfahren für die Analyse solcher Systeme angebracht.

Wir wollen in dieser Arbeit Methoden entwickeln und anwenden, um die Dynamik eines stochastischen dynamischen Systems aus einer Zeitreihe zu rekonstruieren. Dabei werden wir zwei verschiedene Wege beschreiten, um mit modellfreien Ansätzen den deterministischen und den stochastischen Anteil der Bewegungsgleichungen zu erfassen. Zunächst werden wir eine Arbeit von Siegert *et al.* [5] aufgreifen, in welcher vorgeschlagen wurde, Fokker-Planck- und Langevin-Gleichungen von Daten zu bestimmen. Diese Verfahren sind streng nur im Grenzwert unendlicher zeitlicher Auflösung der gemessenen Daten gültig. Wir werden Korrekturterme herleiten, die die Anwendung derartiger Methoden auch in realistischen Situationen mit endlichen Abtastraten der verwendeten Zeitreihen ermöglichen. Außerdem werden wir einen alternativen Algorithmus zur Bestimmung von Fokker-Planck-Gleichungen aus Daten für eindimensionale Probleme vorschlagen, dessen Güte von der Abtastrate der Zeitreihen weitgehend unabhängig ist. Ein wesentlicher Nachteil dieser Methoden ist, daß für die Bestimmung der vollständigen Fokker-Planck-Gleichung eines Systems mit n Freiheitsgraden auch die entsprechenden n Zeitreihen gemessen werden müssen. Diese Forderung ist in vielen realistischen Situationen nicht erfüllt. Daher werden wir uns in Anlehnung an die Einbettungsverfahren in deterministischen Systemen mit der Frage beschäftigen, ob die Dynamik eines mehrdimensionalen stochastischen Prozesses, insbesondere eines Markov-Prozesses, aus einer einzelnen gemessenen Zeitreihe rekonstruiert werden kann. Im allgemeinen wird diese Annahme ungültig sein. Wir werden dennoch Situationen identifizieren, in denen die skalare Messung eines multidimensionalen Markov-Prozesses ein endliches Gedächtnis besitzt, und andere Konstellationen, in denen dies näherungsweise der Fall ist. Letztere Verfahren werden auf *lokale* Vorhersagealgorithmen im Phasenraum führen, wogegen die Fokker-Planck-Verfahren *globale* Ansätze darstellen.

Wir wollen diese Techniken auf Zeitreihen anwenden, deren Charakter offensichtlich

stochastischer Natur ist. Dies sind Daten voll entwickelter hydrodynamischer Turbulenz, atmosphärischer Bodenwindgeschwindigkeiten und der produzierten elektrischen Leistung einer Windturbine. In allen Fällen ist die Frage von Interesse, ob sich nicht-linear deterministische Anteile in diesen Zeitreihen mit starkem stochastischen Charakter feststellen lassen. In den Experimenten voll entwickelter Turbulenz werden wir die Frage untersuchen, ob sich kohärente Strukturen in einer turbulenten Flüssigkeit durch nichtlineare Phasenraummethoden identifizieren lassen und ob innerhalb dieser Strukturen Vorhersagen möglich sind, die über die Auswertung linearer Korrelationen hinausgehen. In dem betrachteten Experiment ist es durch Variation der Rotationsrate eines Flüssigkeitsbehälters möglich, zweidimensionale und dreidimensionale Turbulenz innerhalb ein und desselben physikalischen Systems zu erzeugen. Wir werden dort einen wesentlichen Unterschied zwischen den kohärenten Strukturen in zweidimensionaler und dreidimensionaler Turbulenz herausarbeiten. Auch die Anwendung der Fokker-Planck-Verfahren wird einen deutlichen Unterschied zwischen den beiden Flußtypen ergeben. Somit lassen sich aus den entwickelten Zeitreihen-Methoden für stochastische Systeme wesentliche Schlüsse über den Charakter der Fluktuationen in voll entwickelter Turbulenz ableiten.

Für die Zeitreihen des Oberflächenwindes und der elektrischen Leistung von Windturbinen hat das Vorhandensein von nichtlinearem Determinismus direkte Auswirkungen auf technisch relevante Vorhersageprobleme. Die Fluktuationen der Windgeschwindigkeit haben auf einem weiten Bereich verschiedener Zeitskalen wesentliche Konsequenzen für die Effizienz der Windenergieproduktion. Oft können Vorhersagen des Windes einen starken Nutzen haben, wie z.B. bei der Steuerung konventioneller Energieerzeuger. Wir werden uns auf Vorhersagen mit Horizonten von einigen Sekunden konzentrieren, deren Anwendung in der Steuerung der Rotorblattstellung von Windturbinen liegt. Windturbinen reagieren auf die Fluktuationen des Windes durch die Drehung der Rotorblätter entsprechend der momentan produzierten Leistung. Aufgrund der Trägheit des Rotors und der Verzögerungen in den Auswertungssystemen kommt die Reaktion auf plötzliche Änderungen der Windgeschwindigkeit immer einen Moment zu spät. Es wäre wünschenswert, von einer bevorstehenden Bö Bruchteile einer Sekunde im Voraus zu wissen. Dies definiert das von uns bearbeitete Vorhersageproblem. Wir werden zeigen, daß die Verbesserung nichtlinearer Methoden bei der Vorhersage von Windböen substantiell ist und eine intelligente Steuerung von Windturbinen ermöglicht.

Die Arbeit ist folgendermaßen gegliedert. Im Kapitel 2 werden grundlegende Konzepte dynamischer Systeme und der nichtlinearen Zeitreihenanalyse vorgestellt. Insbesondere werden wir auf modellfreie Methoden zur Bestimmung von Bewegungsgleichungen in dynamischen Systemen eingehen, da diese später im Zentrum der Arbeit stehen werden. Danach werden wir die Verfahren zur Bestimmung von Fokker-Planck- und Langevin-Gleichungen von Daten besprechen. Darauf folgend werden wir uns in Analogie zum Einbettungsverfahren in deterministischen Systemen mit der Rekonstruktion von vektorwertigen Markov-Prozessen aus skalaren Zeitreihen beschäftigen. Wir werden

Vorhersageverfahren für diese Situation vorstellen und die optimale Wahl der Parameter in stochastischen Systemen besprechen. Wir werden den Nutzen dieser Verfahren anhand numerischer Beispiele demonstrieren. Im Kapitel 3 werden wir die entwickelten Methoden auf die Analyse von Zeitreihen zweidimensionaler und dreidimensionaler hydrodynamischer Turbulenz anwenden. Wir werden zunächst die phänomenologische Sprache der Turbulenz einführen, danach das Experiment vorstellen und letztlich die Zeitreihen auf nichtlinearen Determinismus untersuchen. Im Kapitel 4 werden wir die Vorhersagbarkeit von Windzeitreihen und Zeitreihen der elektrischen Leistung von Windturbinen untersuchen, wobei wieder die in Kapitel 2 entwickelten Verfahren zur Anwendung kommen. Wir werden in diesem Kapitel besonderen Wert auf die technische Relevanz der Ergebnisse für die Windenergieproduktion legen. Im Kapitel 5 werden wir unsere Ergebnisse diskutieren und zusammenfassen.

Kapitel 2

Datengetriebene Modelle dynamischer Systeme

In diesem Kapitel sollen die Verfahren der datengetriebenen Modellbildung für dynamische Systeme vorgestellt werden. Zunächst werden die etablierten Methoden für deterministische Systeme, die im wesentlichen auf der Rekonstruktion von Zustandsräumen basieren, zusammengefaßt. Später wird die Anwendung dieser Verfahren zur Analyse stochastischer Daten besprochen. Insbesondere soll der Unterschied zwischen linearen Techniken der konventionellen spektralen Zeitreihenanalyse und den auf dem Konzept des deterministischen Chaos beruhenden nichtlinearen Phasenraummethoden herausgearbeitet werden. Weiterhin werden neuartige Verfahren, basierend auf der Extraktion von Fokker-Planck-Gleichungen aus Daten, eingeführt. Diese Fokker-Planck-Verfahren stellen eine *globale* Methode im Phasenraum des stochastischen Systems dar, während die Verallgemeinerung der aus der Chaostheorie stammenden Verfahren auf *lokale* Modelle im Zustandsraum führen wird.

2.1 Datenanalyse deterministischer dynamischer Systeme

In diesem Abschnitt wird die auf Zeitreihen basierende Modellbildung für deterministische dynamische Systeme eingeführt. Hierzu werden zunächst die zugrundeliegenden Konzepte, wie die Theorie dynamischer Systeme und des deterministischen Chaos, besprochen. Danach wird das Takensche Einbettungstheorem vorgestellt, welches erlaubt, den Phasenraum eines deterministischen Systems aus einer Zeitreihe zu rekonstruieren. Anschließend werden Vorhersagemodelle in einem solchen Phasenraum vorgestellt. Diese Methoden werden in den folgenden Kapiteln für stochastische dynamische Systeme mit Langevin-Dynamik verallgemeinert und nehmen einen zentralen Platz in dieser Arbeit ein.

2.1.1 Dynamische Systeme

Ein dynamisches System kann als eine mathematische Beschreibung für die Fortentwicklung des Zustandes eines Systems in der Zeit verstanden werden, wobei die Zeit entweder eine kontinuierliche oder diskrete Variable darstellt. Der Standardfall für ein dynamisches System mit kontinuierlicher Zeitvariable t ist ein System von n autonomen gekoppelten Differentialgleichungen erster Ordnung

$$\frac{d\vec{x}(t)}{dt} = \mathbf{f}(\vec{x}(t)), \quad (2.1)$$

wobei $\vec{x}(t) = (x_1(t), \dots, x_n(t))$ der Zustandsvektor zum Zeitpunkt t ist. Bei einer diskreten Zeitvariable wird die Dynamik durch eine Abbildung definiert:

$$\vec{x}_{n+1} = \mathbf{F}(\vec{x}_n). \quad (2.2)$$

Ist die Anfangsbedingung $\vec{x}(0)$ bzw. \vec{x}_0 bekannt, so sind, falls \mathbf{f} Lipschitz-stetig ist, durch Gleichung (2.1) bzw. (2.2) alle zukünftigen Zustände des Systems eindeutig festgelegt. Systeme dieser Art sind daher deterministisch. Mathematisch exakter läßt sich ein dynamisches System beschreiben durch ein Tripel (Ω, ϕ, T) , wobei Ω der Phasenraum oder Zustandsraum ist, der alle möglichen Zustände des Systems enthält. Wir werden hier annehmen, daß der Phasenraum ein endlichdimensionaler Raum \mathcal{R}^n ist. Zu allen möglichen Zeiten $t \in T$ befindet sich das System in einem der möglichen Zustände $\vec{x}(t) \in \Omega$. Das dynamische System besitzt außerdem einen Evolutionsoperator,

$$\phi : \Omega \times T \rightarrow \Omega. \quad (2.3)$$

Für jedes t definiert $\phi(\vec{x}, t)$ eine Abbildung

$$\phi^t : \Omega \rightarrow \Omega, \phi^t(\vec{x}) = \phi(\vec{x}, t), \quad (2.4)$$

welche als der Fluß des Systems in der Zeit t bezeichnet wird. Für festes $\vec{x} \in \Omega$, wird $\phi(\vec{x}, t)$ als Trajektorie durch \vec{x} bezeichnet. Damit durch ϕ^t ein dynamisches System definiert wird, muß der Evolutionsoperator noch die Halbgruppeneigenschaft erfüllen, d.h., es muß $\phi^t \circ \phi^s = \phi^{t+s}$ gelten. Typische Trajektorien divergieren entweder gegen unendlich oder bleiben beschränkt, der letztere Fall ist in dieser Arbeit von Interesse. Welche Form der Lösung tatsächlich auftritt, hängt sowohl von der Form von \mathbf{F} bzw. \mathbf{f} ab als auch von der jeweiligen Anfangsbedingung. Viele dynamische Systeme haben nicht nur beschränkte Lösungen, sondern besitzen weiterhin die Eigenschaft, daß ein Phasenraumvolumen, welches die Anfangsbedingungen enthält, im Mittel unter der Dynamik kontrahiert wird. Dann gilt im Mittel $|\det \mathbf{J}_{\mathbf{F}}| < 1$ oder $\text{div} \cdot \mathbf{f} < 0$, wobei $\mathbf{J}_{\mathbf{F}}$ die Jacobische von \mathbf{F} ist. Diese Systeme heißen **dissipativ**. Für solche Systeme wird eine Menge von Anfangsbedingungen mit von Null verschiedenem Maß auf eine Untermenge des Phasenraumes zusammengezogen, welche **Attraktor** genannt wird.

Attraktoren sind invariant unter der Dynamik des Systems. Wenn man im Fall der diskreten Dynamik mit einer Ausgangsverteilung $\rho_0(\vec{x}_k)$ von Anfangsbedingungen $\vec{x}_k \in \Omega$ beginnt, welche durch die Abbildung \mathbf{F} iteriert werden, so erhält man nach langer Zeit n eine Verteilung $\rho_n(\vec{x}_k)$ von Punkten im Phasenraum, welche von der Dynamik unabhängig ist. Damit bezeichnet man $\rho_n(\vec{x}_k)$ als invariante Menge, wenn gilt $\rho_n(\vec{x}_k) = \rho_{n+1}(\vec{x}_k)$, und die Verteilung erfüllt die Frobenius-Peron-Gleichung:

$$\rho(\vec{x}) = \int \rho(\vec{y}) \delta[\vec{x} - \mathbf{F}(\vec{y})] d\vec{y}. \quad (2.5)$$

Ein dynamisches System wird als **chaotisch** bezeichnet, wenn seine Zeitentwicklung exponentiell sensitiv von den Anfangsbedingungen $\vec{x}(0)$ abhängt. Man betrachte zwei Anfangsbedingungen $\vec{x}_1(0)$ und $\vec{x}_2(0) = \vec{x}_1(0) + \delta(0)$, welche durch ein in der Zeit kontinuierliches dynamisches System fortentwickelt werden und somit Lösungen $\vec{x}_1(t)$ und $\vec{x}_2(t)$ erzeugen. Falls die Lösungen im Grenzfall $\delta(0) \rightarrow 0$ und großer t beschränkt bleiben und deren Differenz exponentiell wächst, d.h. $|\delta(t)|/|\delta(0)| \propto \exp(ht)$ für $h > 0$ gilt, so spricht man von sensitiver Abhängigkeit von den Anfangsbedingungen, und das System ist chaotisch. In chaotischen dynamischen Systemen wächst somit die Differenz zwischen zwei anfänglich benachbarten Zuständen exponentiell an, so daß schon nach kurzer Zeit keine Ähnlichkeit mehr zwischen den beiden Zuständen besteht. Wird eine Komponente des Zustandsvektors $\vec{x}(t)$ in einer Zeitreihe dargestellt, so macht sich das Chaos typischerweise durch ein unregelmäßiges Verhalten der Zeitreihe bemerkbar, da die das asymptotische Auseinanderlaufen verursachende Streckung durch nichtlineare Faltungen kompensiert werden muß. Zeitreihen deterministischer chaotischer Systeme sind auf den ersten Blick nicht von solchen unterscheidbar, die durch Systeme mit stochastischem Antrieb erzeugt werden. Die strenge Definition von Chaos ist durch die Eigenschaft des Mischens gegeben, da diese die sensitive Abhängigkeit von den Anfangsbedingungen impliziert. Ist μ ein invariantes Wahrscheinlichkeitsmaß im Phasenraum und seien die Mengen A und B als Ereignisse interpretiert, so gelten diese als unabhängig, wenn gilt $\mu(A \cap B) = \mu(A) \cdot \mu(B)$. Das Mischen durch ein dynamisches System kann als eine Art asymptotische Unabhängigkeit der Ereignisse betrachtet werden. Ein invertierbares dynamisches System heißt mischend, wenn für beliebige A, B die Beziehung

$$\lim_{n \rightarrow \infty} \mu(\phi^n(A) \cap B) = \mu(A) \cdot \mu(B) \quad (2.6)$$

gilt. Die Mischeigenschaft impliziert also, daß sich $\phi^n(A)$ für große n mit einer wachsenden Verteilung ausbreitet, wobei der Anteil von $\phi^n(A)$, der für große n in einer beliebigen Menge $B \in \Omega$ liegt, proportional zum Maß $\mu(B)$ ist.

Eine natürliche Frage ist, welche deterministischen Systeme in der Lage sind, chaotisches Verhalten zu erzeugen. Zunächst ist hierfür notwendig, daß die Funktion $\mathbf{f}(\vec{x}(t))$ bzw. $\mathbf{F}(\vec{x}_n)$ nichtlinear ist. Für ein System von n gekoppelten gewöhnlichen Differentialgleichungen erster Ordnung gilt außerdem als Voraussetzung für Chaos $n \geq 3$, für

invertierbare Abbildungen gilt $n \geq 2$, und für nichtinvertierbare Abbildungen genügt bereits ein eindimensionaler Phasenraum. Eine genaue Analyse zeigt allerdings, daß die Notation “Nichtlinearität” weniger exakt ist als “nicht entkoppelbar”, da auch die Bewegungsgleichungen für einen harmonischen Oszillator durch eine geschickte Variablentransformation auf nichtlineare Form gebracht werden können.

Eine typische Eigenschaft nichtlinearer Systeme ist die fehlende Robustheit der Lösungen gegen leichte Änderungen von Systemparametern. Dynamische Systeme können von regulärem auf chaotisches Verhalten umschalten, wenn ein Parameter in den Bewegungsgleichungen variiert wird. In solchen Fällen beobachtet man einen abrupten Übergang von einer Attraktorgeometrie zu einer anderen, z.B. von einem Torus zu einem seltsamen Attraktor. Dieser Übergang bei einem bestimmten Wert eines Kontrollparameters wird als **Bifurkation** bezeichnet.

2.1.2 Beispiele für dynamische Systeme

Wir wollen nun einige Beispiele für deterministische dynamische Systeme geben, einerseits weil diese Systeme in ähnlicher Form im Verlauf der Arbeit zur Demonstration unserer Methoden dienen werden, andererseits vermitteln sie einen Eindruck von den zu erwartenden Phänomenen in nichtlinearen dynamischen Systemen und von der den später verwendeten Modellen zugrundeliegenden Physik.

Ein System mit nichtlinearen Bewegungsgleichungen, jedoch ohne chaotisches Verhalten, ist der Van-der-Pol-Oszillator. Das Van-der-Pol-Modell [6] beschreibt das Verhalten der elektrischen Ladung $q(t)$, welche durch eine Triodenröhre fließt. Die Bewegungsgleichung lautet:

$$\frac{d^2q}{dt^2} - (r - q^2)\frac{dq}{dt} + q = 0, \quad (2.7)$$

bzw. in der üblichen Schreibweise als System von Differentialgleichungen erster Ordnung:

$$\begin{aligned} \frac{dq}{dt} &= p \\ \frac{dp}{dt} &= (r - q^2)p - q. \end{aligned} \quad (2.8)$$

Ohne den nichtlinearen Term q^2 wäre die Lösung des Systems durch eine gedämpfte harmonische Oszillation gegeben. Van der Pols Leistung bestand darin, den Dämpfungsterm von der Ladungsmenge q abhängen zu lassen. Für kleine q würde die Triodenröhre die Amplitude des Signals verstärken, wogegen die Ladungsmenge für große Werte von q durch Sättigungseffekte begrenzt ist.

Das Van-der-Pol-System (2.8) hat einen instabilen Fixpunkt bei $(0, 0)$. Der Attraktor

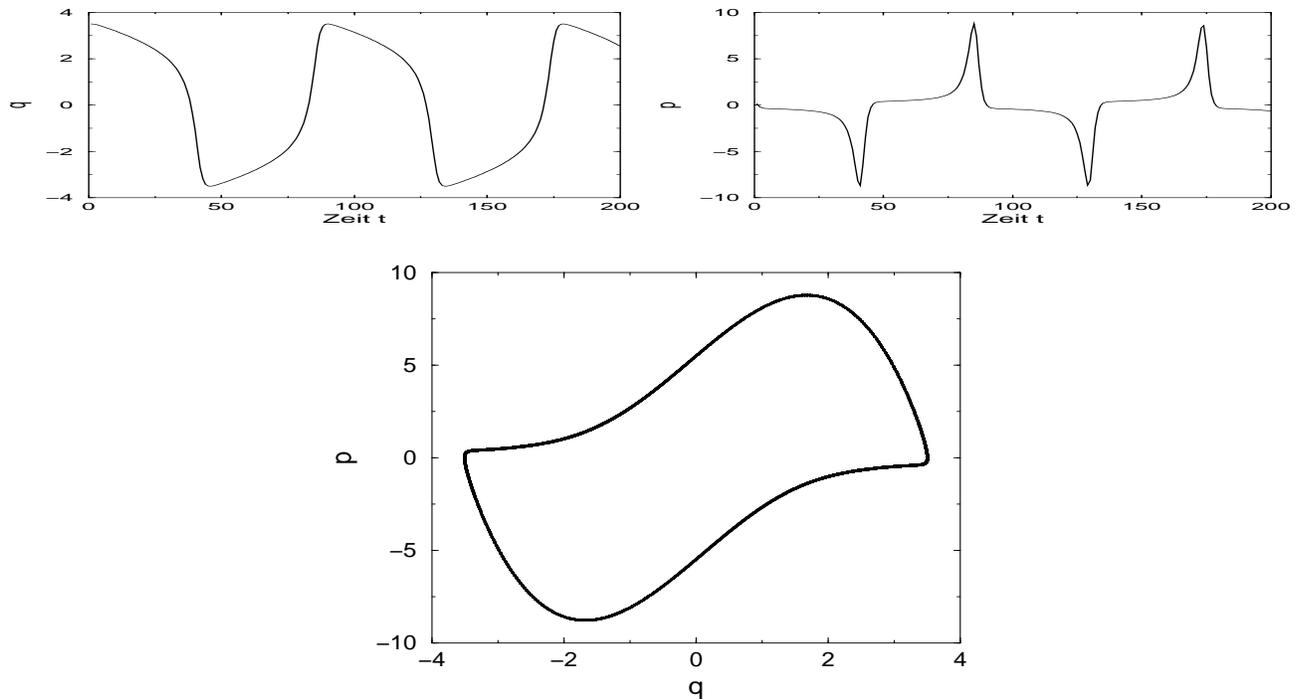


Abbildung 2.1: Verhalten der Variablen q und p des Van-der-Pol-Systems und der stabile Grenzzyklus im $q - p$ Zustandsraum.

des Systems ist ein stabiler Grenzzyklus, der für kleine r kreisförmig ist. Für größere Werte von r sind die Oszillationen in q nicht länger sinusförmig, und der Attraktor hat eine Gestalt wie in Abb. 2.1 dargestellt.

Eines der ersten dynamischen Systeme, in welchem chaotisches Verhalten festgestellt wurde, ist das Lorenz-Modell [7]. Es ist ein in der Zeit kontinuierliches dynamisches System, das durch drei gekoppelte gewöhnliche Differentialgleichungen beschrieben wird:

$$\begin{aligned} \frac{dx}{dt} &= \sigma(y - x) \\ \frac{dy}{dt} &= (r - z)x - y \\ \frac{dz}{dt} &= xy - bz, \end{aligned} \tag{2.9}$$

mit den kanonischen Werten für die Parameter $\sigma = 10$, $r = 28$ und $b = 8/3$, für welche das System chaotisch ist. Die Gleichungen stellen ein vereinfachtes Modell für eine getriebene dissipative hydrodynamische Strömung (Rayleigh-Bénard-Konvektion) dar,

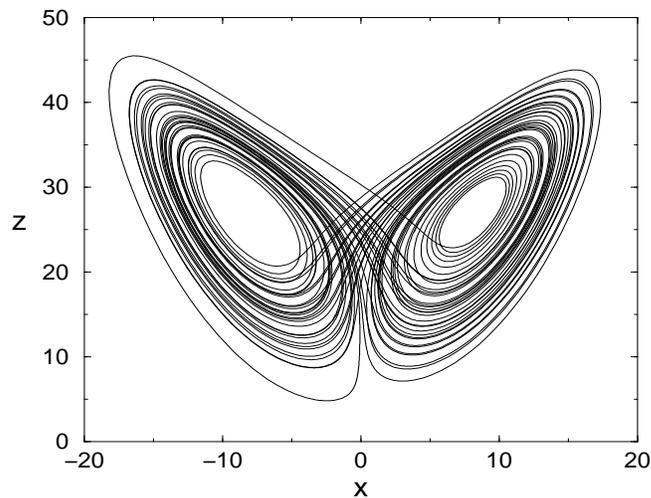


Abbildung 2.2: Projektion einer Trajektorie des Lorenz-Systems auf die x - z Ebene mit den Parametern $b = 8/3$, $\sigma = 10$ und $r = 28$.

in welcher x die Strömungsgeschwindigkeit beschreibt, y ist die Temperaturdifferenz zwischen aufsteigender und sinkender Flüssigkeit, und z stellt die Abweichung des Temperaturprofils vom linearen Verhalten dar. Weiterhin läßt sich σ mit der Prandtl-Zahl identifizieren, und r stellt einen externen Kontrollparameter dar, welcher proportional zur Reynolds-Zahl ist.

Das Lorenz-System besitzt drei instabile Fixpunkte: den Ursprung $(x, y, z) = (0, 0, 0)$ und die symmetrischen Punkte $(x, y, z) = (\pm\sqrt{b(r-1)}, \pm\sqrt{b(r-1)}, r-1)$. Die Projektionen dieser Fixpunkte auf die $x-z$ Ebene liegen in den Zentren der Flügel des Attraktors, dessen Projektion auf diese Ebene in Fig. 2.2 dargestellt ist. Die Zeitentwicklung des Systems ist chaotisch.

2.1.3 Einbettungsmethoden für dynamische Systeme

In Experimenten oder in Feldmessungen werden in der Regel nicht alle unabhängigen Variablen eines dynamischen Systems gemessen. Häufig steht nur eine skalare Zeitreihe zur Verfügung, um die Dynamik des Systems zu untersuchen. Die Grundlage nahezu sämtlicher Methoden der nichtlinearen Zeitreihenanalyse bildet die Rekonstruktion des Phasenraumes eines deterministischen dynamischen Systems aus einer gemessenen Zeitreihe. Das mathematische Fundament der Rekonstruktion eines zum originalen Phasenraum des dynamischen Systems äquivalenten Zustandsraumes bilden die Einbettungstheoreme von Takens und Sauer [2, 8]. Wir möchten an dieser Stelle die wesentlichen Ideen dieser Arbeiten darstellen.

Der Ausgangspunkt der Rekonstruktionstheoreme ist ein dynamisches System, welches durch die unbekannte Abbildung

$$\vec{x}_{n+1} = \mathbf{F}(\vec{x}_n) \quad (2.10)$$

im Zustandsraum Ω beschrieben wird. Im Beweis des Rekonstruktionstheorems wird angenommen, daß die Entwicklung des Systems auf einen endlichdimensionalen Unterraum M des Zustandsraumes Ω beschränkt ist. Die Meßfunktion h sei glatt und nur von dem momentanen Zustand des dynamischen Systems abhängig, das heißt, sie hängt nicht explizit von der Zeit ab. Außerdem nehmen wir an, h sei eine reellwertige skalare Funktion, d.h., h ist eine Abbildung von M auf die reelle Achse \mathbb{R} . Zunächst sei für jedes n die Beobachtung y_n das Bild von $\vec{x} \in M$ unter der Abbildung $h : M \rightarrow \mathbb{R}$ definiert durch $\vec{x} \rightarrow h(\vec{x})$. Nun läßt sich die Zeitversatzeinbettung durch die m -dimensionale Abbildung

$$\Phi_m : M \rightarrow \mathbb{R}^m, \vec{x} \rightarrow (h(\vec{x}), h(\mathbf{F}(\vec{x})), \dots, h(\mathbf{F}^{m-1}(\vec{x}))) =: (y_n, \dots, y_{n+m-1}) \quad (2.11)$$

definieren, die den Zustand \vec{x} auf (y_n, \dots, y_{n+m-1}) abbildet. Es ist intuitiv klar, daß die Kenntnis von y_{n_i} zu unterschiedlichen Zeiten n_i äquivalent ist zur Kenntnis einer bestimmten Zahl unterschiedlicher Variablen zur selben Zeit, wenn die Abbildung \mathbf{F} die unterschiedlichen Freiheitsgrade derart koppelt, daß die verschiedenen Komponenten y_{n_i} des Einbettungsvektors gegenseitig unabhängig sind. Das Einbettungstheorem von Takens garantiert nun für glatte Meßfunktionen h und für hinreichend große Werte von m , daß Φ_m generisch eine glatte invertierbare Abbildung von M nach \mathbb{R}^m mit glatter Umkehrabbildung ist. Es folgt daraus, daß die Punkte, welche durch den Einbettungsvektor (y_n, \dots, y_{n+m-1}) gebildet werden, eine eins zu eins Abbildung des originalen Attraktors darstellen. Das Bild dieser Abbildung bezeichnet man als den rekonstruierten Attraktor. Wenn die Dynamik des Systems auf eine kompakte Mannigfaltigkeit der Dimension D beschränkt ist, so garantiert das Takens-Theorem, daß die Abbildung Φ_m generisch eine Einbettung darstellt für $m = 2D + 1$ und Meßfunktionen h , die C^2 sind. Für fraktale Attraktoren wurde diese Aussage durch Sauer *et al.* [8] verallgemeinert. Man erhält hier generisch eine Einbettung, wenn gilt $m > 2D_F$, wobei D_F die fraktale Dimension des Attraktors ist. Als weitere Annahme ist hier vorausgesetzt, daß es keine periodischen Orbits der Periode δt bzw. $2\delta t$ gibt und nur endlich viele mit der Periode $p\delta t$, $p > 2$ existieren, wobei δt die Abtastrate der Zeitreihe ist. Das wesentliche Ergebnis dieses Einbettungstheorems ist, daß nicht die Dimension des eigentlichen Phasenraumes Ω für die Rekonstruktion der Dynamik nötig ist, sondern häufig genügt die wesentlich kleinere Dimension D_F des Trägers des invarianten Maßes, welches durch die Dynamik generiert wird.

Zeitversatz: Wir haben bisher der Einfachheit halber diskrete Systeme mit Zeitschritten n betrachtet. Für kontinuierliche Systeme mit der Zeitvariable t betrage der

Abstand zwischen zwei Elementen des Einbettungsvektors τ . Der Zeitversatz τ zwischen aufeinanderfolgenden Elementen des Einbettungsvektors wird durch die Einbettungstheoreme nicht diskutiert, abgesehen davon, daß einige spezielle Werte verboten sind. Im Grenzfall unendlicher Genauigkeit der Daten und für unendlich lange Zeitreihen sind alle anderen Werte von τ mathematisch äquivalent. In praktischen Anwendungen ist eine geeignete Wahl des Zeitversatzes jedoch wichtig. Für sehr große Werte von τ können aufeinanderfolgende Elemente des Einbettungsvektors nahezu unabhängig sein. Dies ist typisch für chaotische Systeme, bei denen die Autokorrelationsfunktion schnell zerfällt (exponentiell für mischende Systeme). Der rekonstruierte Attraktor besitzt dann eine komplizierte Struktur, die numerisch wie auch visuell kaum noch auswertbar ist.

Falls τ sehr klein ist, sind aufeinanderfolgende Elemente des Einbettungsvektors sehr stark korreliert, und die Punkte liegen im rekonstruierten Phasenraum nahe der Diagonale. Wenn die Daten durch *Meßrauschen* gestört sind, können verschiedene Regionen des Phasenraumes mit sehr unterschiedlicher Dynamik im Fall zu kleiner τ ununterscheidbar werden. Die Variation des Signals sollte im Zeitraum τ somit im Mittel größer sein als das Rauschniveau. Wenn σ_{rausch}^2 die Varianz des Rauschens und σ_{sig}^2 die Varianz des Signals sind, so sollte τ größer sein als die Zeit, in der die auf eins normierte Autokorrelationsfunktion auf $1 - \sigma_{rausch}^2 / \sigma_{sig}^2$ abfällt. Somit wächst die optimale Zeitverzögerung mit wachsendem Rauschniveau. Wir werden später in dieser Arbeit zeigen, daß für *dynamisches Rauschen* das Gegenteil richtig ist.

Es gibt eine Vielzahl von Methoden, um die optimale Zeitverzögerung in deterministischen dynamischen Systemen mit Beobachtungsrauschen abzuschätzen. Im wesentlichen basieren diese Verfahren auf dem geometrischen Argument, daß der rekonstruierte Attraktor in alle Raumrichtungen etwa dieselbe Ausdehnung haben sollte. Statistiken, die diese Anforderung quantitativ auswerten, sind z.B. der *Füllfaktor* [9] oder die *Entfernung von der Diagonalen* [10]. Die populärste Statistik ist jedoch die Autokorrelationsfunktion, einerseits weil man diese ohnehin berechnet, um ein Verständnis über die Zeitreihe zu bekommen. Andererseits steht sie in einem engen Verhältnis zur Geometrie des Attraktors - dieser ist etwa symmetrisch entfaltet, wenn die Zeitverzögerung dem ersten Nulldurchgang der Autokorrelationsfunktion entspricht. Dies wird deutlich, wenn man eine sinusförmige Oszillation betrachtet. Der erste Nulldurchgang der Autokorrelationsfunktion erfolgt bei einem Viertel der Periode. Genau bei dieser Zeitverzögerung wird der Sinus durch die Zeitversatzeinbettung auf einen Kreis abgebildet. Um alle nichtlinearen Korrelationen zu berücksichtigen, sollte korrekterweise bei nichtlinearen dynamischen Systemen das erste Minimum der zeitverzögerten *mutual information* als Optimierungskriterium anstelle der Autokorrelationsfunktion genutzt werden [11].

Die zweidimensionale Projektion des rekonstruierten Lorenzattraktors ist in Abbildung 2.3 dargestellt. Hierfür wurde die Zeitversatzeinbettung für die x -Koordinate durchgeführt.

Ableitungskoordinaten: Numerische Ableitungen von in diskreten Zeitschritten ge-

messenen Daten sind Linearkombinationen von aufeinanderfolgenden Datenpunkten. Daher sind Ableitungskordinaten, d.h. numerisch geschätzte Vektoren $(y_n, \dot{y}_n, \ddot{y}_n, \dots)$, Linearkombinationen der oben definierten Einbettungsvektoren. Bei derartig konstruierten Einbettungsvektoren wird so die Optimierung von τ überflüssig, und man kann häufig ein besseres physikalisches Verständnis der rekonstruierten Variablen gewinnen [12]. In vielen praktischen Anwendungen sind Ableitungskordinaten jedoch problematisch, da das Beobachtungsrauschen durch diese Methode verstärkt wird. Wir werden später zeigen, daß die Konstruktion von Ableitungskordinaten auch in bestimmten stochastischen dynamischen Systemen ermöglicht, eine Einbettung zu definieren. Neben

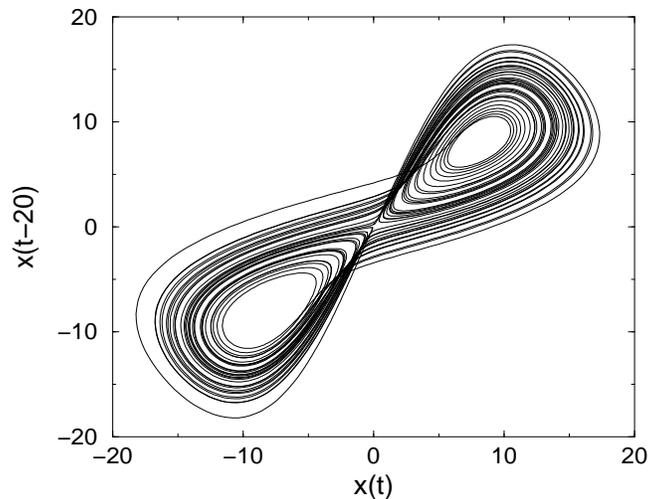


Abbildung 2.3: Rekonstruktion einer Trajektorie des Lorenz-Systems durch Zeitversatzeinbettung mit den Parametern $b = 8/3$, $\sigma = 10$ und $r = 28$.

den metrischen Eigenschaften des Attraktors, wie z.B. seiner fraktalen Dimension, wird auch die Dynamik des ursprünglichen Systems durch die Einbettungsmethoden rekonstruiert. Dies bedeutet, daß ein Zustand im rekonstruierten Phasenraum zu einer Zeit t den Zustand zu späteren Zeiten eindeutig festlegt. Somit kann man die dynamischen Entwicklungsregeln allein aus einer Zeitreihe bestimmen, ohne Kenntnis über die physikalischen Gesetze des zugrundeliegenden Systems zu benötigen. Diese Tatsache hat zu einer Reihe von Vorhersagealgorithmen für Zeitreihen geführt, die wir im folgenden Abschnitt genauer beschreiben möchten.

2.1.4 Vorhersage und Modellierung für Zeitreihen

Die auf dem Konzept des deterministischen Chaos basierenden Vorhersagealgorithmen stellen eine recht neue Klasse von Verfahren zur Beschreibung aperiodischer und unregelmäßiger Signale dar. Wir wollen zunächst diese Methoden vorstellen, bevor

wir AR-Prozesse einführen, welchen das Konzept eines linearen stochastischen Prozesses zugrundeliegt. Diese sind die gebräuchlichsten linearen Vorhersagealgorithmen und werden uns daher als Maßstab zur Beurteilung der Leistungsfähigkeit nichtlinearer Verfahren dienen.

Zuerst haben Farmer & Sidorowich festgestellt, daß es in rekonstruierten Zustandsräumen möglich ist, die Bewegungsgleichungen eines dynamischen Systems mit hoher Genauigkeit zu extrahieren [13]. Da die dynamischen Eigenschaften des Systems unter Messung und anschließender Einbettung erhalten bleiben, existieren in rekonstruierten Phasenräumen Bewegungsgleichungen der Form:

$$y_{n+1} = \mathbf{G}(\vec{y}_n). \quad (2.12)$$

Es gibt eine Vielzahl von Methoden, das Vektorfeld \mathbf{G} aus den Daten zu ermitteln. Zunächst wählt man einen Ansatz für die funktionale Gestalt von \mathbf{G} . Je nach dem, ob es günstiger ist \mathbf{G} für den gesamten Attraktor einmal zu bestimmen oder für den jeweiligen Bezugspunkt auf dem Attraktor neu, spricht man von einem *globalen* bzw. *lokalen* Modell. Im Falle eines globalen Modells würde man \mathbf{G} typischerweise als Entwicklung in einen geeigneten Satz von Basisfunktionen mit den Entwicklungsparametern p_i darstellen; $\mathbf{G}(\vec{y}) \approx \sum_{i=1}^n \mathbf{G}_{p_i}(\vec{y})$. Als ein solcher Satz von Funktionen können beispielsweise Polynome, Radiale Basis-Funktionen oder Neuronale Netze dienen. Die Parameter p_i werden dann durch die Minimierung einer geeigneten Kostenfunktion bestimmt. Die Wahl der Kostenfunktion hängt vom Zweck der Vorhersage bzw. Modellierung ab. Wenn beispielsweise nur der größte Fehler von einer Vielzahl von Vorhersagen von Bedeutung ist, wäre die Maximumsnorm die geeignete Kostenfunktion. Im Fall, daß positive Fehler andere Kosten verursachen als negative Fehler, müßte die Kostenfunktion asymmetrisch sein. Diese Situation wird im Verlauf der Arbeit bei der Vorhersage von Windböen zur Steuerung von Windkraftanlagen auftreten. Ein anderer Ansatz zur Bestimmung der optimalen Parameter bzw. des optimalen Modells besteht darin, die Wahrscheinlichkeit der Beobachtungen in Bezug auf das Modell zu maximieren (maximum-likelihood principle). Wenn die Vorhersagefehler einer Gaußverteilung genügen, ist die Gesamtwahrscheinlichkeit, eine Zeitreihe bei vorgegebenem Modell zu beobachten, durch das Produkt der einzelnen Gaußverteilungen der Abweichungen gegeben, und die Maximierung der Wahrscheinlichkeit führt zur Minimierung der quadratischen Fehlerfunktion

$$\epsilon^2 = \sum_{k=m}^{N-1} (y_{k+1} - \mathbf{G}_{p_i}(\vec{y}_k))^2 \stackrel{!}{=} \min. \quad (2.13)$$

Anstatt die globale Dynamik aus den Daten zu ermitteln, kann man lokale Modelle im Phasenraum konstruieren. Die Idee hierbei ist, daß man die Entwicklung des Vektorfeldes $\mathbf{G}(\vec{y}_n)$ nur bis zu wesentlich geringerer Ordnung durchzuführen braucht, um einen guten Fit an den Attraktor zu bekommen, als bei einem globalen Modell. Entwickelt man das Vektorfeld $\mathbf{G}(\vec{y}_n)$ als Taylorreihe um den Bezugspunkt \vec{y}_n , so sind lokal

konstante Modelle die nullte Ordnung dieser Entwicklung, und lokal lineare Modelle entsprechen der ersten Ordnung. Dieser Ansatz kann globalen Methoden besonders in Situationen überlegen sein, wenn die Daten nichtstationär sind oder die globale Dynamik zu kompliziert ist, um durch einfache Funktionen, wie Polynome, angepaßt werden zu können. Das Problem der Nichtstationarität kann durch lokale Modelle wirkungsvoll bearbeitet werden, da dynamisch unterschiedliche Abschnitte der Zeitreihe verschiedene Regionen des Phasenraumes ausfüllen. Der Nutzen lokaler Modelle steigt mit der Länge der verfügbaren Zeitreihen, da mit wachsender Datenbasis die Anzahl der Nachbarn in einem Volumenelement des Phasenraumes wächst und die Modelle somit statistisch robuster werden. Der Nachteil lokaler Modelle liegt zum einen gerade in dieser statistischen Unsicherheit und andererseits in dem hohen numerischen Aufwand, der mit der Konstruktion der Nachbarschaft im Phasenraum verbunden ist.

Lokal konstante Modelle: Der einfachste lokale Algorithmus ist Lorenz' "Methode der Analogen" [14]. Hier wird zur Vorhersage eines Punktes im Einbettungsraum einfach nach dem nächsten Nachbarn gesucht und das Bild dieses Nachbarn als Vorhersagewert verwendet. Bei diesem Verfahren wird kein bestimmtes Modell verwendet, sondern lediglich die Struktur im rekonstruierten Phasenraum ausgewertet. Diese Methode ist natürlich sehr anfällig gegen Meßrauschen in den Daten. Sie gewinnt an statistischer Robustheit, wenn man über mehrere Nachbarn innerhalb einer bestimmten Umgebung U_n im Phasenraum mittelt. Die Vorhersage ist dann der arithmetische Mittelwert über die "Zukünfte" der betrachteten k Nachbarn

$$y_{n+1} = \frac{1}{k} \sum_{\vec{y}_i \in U_n} y_{i+1}, \quad (2.14)$$

wobei $U_n = \{y_i : |\vec{y}_n - \vec{y}_i| < \epsilon\}$. Diese Methode kann verfeinerten Modellen überlegen sein, wenn der deterministische Anteil in den Daten schwach ist bzw. der zugrundeliegende Phasenraum zu hochdimensional ist, um auf genaue Weise durch die Einbettung rekonstruiert zu werden. Dieser Fall liegt bei allen in dieser Arbeit untersuchten Datensätzen vor, so daß wir bei den späteren Anwendungen jeweils lokal konstante Modelle verwenden, wenn von nichtlinearen Vorhersagen die Rede ist.

Lokal lineare Modelle: Lokal konstante Modelle können in gewissen Situationen zu ungenauen Vorhersagen führen, z.B. für Punkte die nah am Rand des Attraktors liegen. Nachbarn werden typischerweise asymmetrisch um diese Punkte verteilt sein (mit höherer Dichte zum Zentrum des Attraktors), so daß auch die Vorhersagen eine systematische Abweichung vom tatsächlichen Wert zeigen werden. Wenn die Daten hinreichend rauschfrei und niedrigdimensional sind, können lokal lineare Modelle Verbesserungen gegenüber lokal konstanten Modellen bringen. Die Vorhersage für \vec{y}_n lautet in lokal linearer Näherung:

$$y_{n+1} = \vec{A}_n \vec{y}_n + b_n,$$

wobei die Koeffizienten \vec{A}_n und \vec{b}_n im maximum likelihood - Sinne durch die Minimierung der Kostenfunktion:

$$e_n^2 = \sum_{\vec{y}_i \in \mathcal{U}_n} (y_{i+1} - \vec{A}_n \vec{y}_i - b_n)^2 \stackrel{!}{=} \min \quad (2.15)$$

bestimmt werden.

Global lineare Modelle: Die Qualität lokaler Phasenraummodelle wird im Allgemeinen durch den Vergleich mit entsprechenden global linearen Modellen getestet. Die populärsten linearen Verfahren mit Vorhersagekraft sind autoregressive Modelle der Ordnung m , sogenannte *AR*-Modelle, bei denen der Zukunftswert als Linearkombination der Zeitreihenwerte in der Vergangenheit angenommen wird:

$$x_{n+1} = \sum_{i=1}^m a_i x_{n-(i-1)\tau} + \xi_{n+1}. \quad (2.16)$$

Dies beschreibt ein *AR*-Modell der Ordnung m , wobei ξ_{n+1} gaußverteilte Zufallszahlen sind. Daher ist auch x_n eine gaußsche Zufallszahl. *AR*-Modelle können verrauschte harmonische Oszillationen erzeugen, wenn die Koeffizienten geeignet gewählt werden. In diesem Fall müssen die Wurzeln des Polynoms $P(z) = \sum_{i=1}^m a_i z^{m-i}$ im Innern des Einheitskreises liegen, d.h., die Amplitude folgt einem Diffusionsprozeß. Das Spektrum S_k des Prozesses ist durch die Koeffizienten a_i eindeutig festgelegt: $S_k = 1/|1 - \sum_{j=1}^m a_j e^{i2\pi kj/N}|^2$. Ein solches Spektrum hat im Limes $N \rightarrow \infty$ Pole bei bestimmten Frequenzen k , welche für endliches N ausgewaschen sind.

Die Überlegenheit lokaler Modelle gegenüber einem *AR*(m)-Modell wird üblicherweise als Indikation für die Existenz nichtlinear deterministischer Struktur in den Daten gewertet. Dies wurde z.B. durch den Casdagli-Test formal eingeführt [15]. Dabei wird in einem lokal linearen Modell die Umgebungsgröße für die Suche der Nachbarn variiert. Wenn die Umgebungsgröße die Ausdehnung des Attraktors erreicht, bekommt man ein global lineares Modell. Dagegen ergeben kleine Umgebungsgrößen lokal lineare Modelle. Somit erwartet man für die Zeitreihe eines niedrigdimensionalen nichtlinearen Systems ein monoton ansteigendes Vorhersagefehler mit der Umgebungsgröße, wogegen man für einen linearen stochastischen Prozeß einen monotonen Abfall beobachten würde. Für Daten eines nichtlinear stochastischen Systems ist ein Minimum des Fehlers in der Umgebungsgröße zu beobachten. Ein umfassender Überblick über lineare Verfahren ist z.B. in [1] zu finden. Der erste Schritt zur Einführung von Nichtlinearität in *AR*-Modellen wurde durch das "Zerschneiden" des Einbettungsraumes in separate Abschnitte erreicht [16]. Die sogenannten Threshold-*AR* (*TAR*)-Modelle bestehen aus einer Anzahl verschiedener *AR*-Modelle, wobei jedes einzelne nur in einem bestimmten Abschnitt des Phasenraumes gültig ist. Zur Bestimmung des Modells unterteilt man den Phasenraum in verschiedene Abschnitte und fittet die Parameter, indem man nur Datenpunkte aus dem entsprechenden Abschnitt nutzt. *TAR*-Modelle sind daher eine spezielle Form lokal linearer Modelle.

Eine weitere Klasse populärer linearer Modelle sind *moving average* (*MA*)-Modelle. Dabei ist der aktuelle Zeitreihenwert eine Linearkombination verschiedener gaußscher Zufallszahlen:

$$x_n = \sum_{i=1}^m b_i \xi_{n-i}. \quad (2.17)$$

Da diese Zufallszahlen aber unbekannte Eingangsdaten darstellen, haben *MA*-Modelle keine Vorhersagekraft und werden daher im Verlauf der Arbeit nicht in Betracht gezogen, da wir uns für Vorhersagen in dynamischen Systemen interessieren.

Modellierung versus Vorhersage: Bei der Beschreibung aperiodischer Signale wird man auf natürliche Weise mit der Frage konfrontiert, ob das Modell darauf optimiert sein sollte, die Statistik (z.B. die Verteilung) der Daten möglichst gut wiederzugeben bzw. bei einer Iteration eines bestimmten Anfangszustandes den zugrundeliegenden Attraktor möglichst gut wiederzugeben (Modellierung) oder die Vorhersagefehler der einzelnen Vorhersagen zu minimieren (Vorhersage). In deterministischen dynamischen Systemen ohne Rauschen sind beide Probleme äquivalent. Ausgehend von einem beliebig genau bekannten Anfangszustand wäre man durch Iteration der deterministischen Bewegungsgleichung in der Lage, eine gemessene Zeitreihe exakt zu reproduzieren. Dies würde sowohl die richtige Statistik als auch genaue Vorhersagen erzeugen. In einem stochastischen dynamischen System sind die Probleme der Vorhersage und Modellierung verschieden, da der stochastische Teil der Dynamik per Definition nicht vorhersagbar ist. Somit würde ein idealer Vorhersager nur den deterministischen Teil der Bewegungsgleichungen iterieren, womit jedoch in der Regel die Statistik der Daten falsch wiedergegeben würde. Wir werden später in dieser Arbeit zeigen, daß die oben eingeführten lokal konstanten Modelle in Markov-Prozessen generell in der Lage sind, beiden Anforderungen zu genügen, abhängig von der Zahl der Nachbarn, die zur Bestimmung des Vorhersagewertes genutzt werden.

2.2 Datenanalyse stochastischer dynamischer Systeme

In diesem Abschnitt wollen wir verschiedene Methoden vorstellen, die Dynamik stochastischer dynamischer Systeme aus einer Zeitreihe zu ermitteln. Dabei wird sowohl die Anwendbarkeit der für das deterministische Chaos entwickelten Konzepte auf stochastische Systeme untersucht als auch ein Verfahren vorgestellt, welches es ermöglicht, Langevin-Gleichungen bzw. Fokker-Planck-Gleichungen aus Daten abzuleiten [5]. Zunächst wollen wir letztere Methode besprechen.

2.2.1 Langevin- und Fokker-Planck-Gleichungen von Zeitreihen

Wir wollen nun eine Methode vorstellen, die es erlaubt, die Drift- und Diffusionskoeffizienten einer Fokker-Planck-Gleichung aus Daten zu bestimmen. Hierfür sei ein stochastischer Prozeß durch die Langevin-Gleichung für einen Satz von n Variablen $\vec{x} = \{x_k(t)\}, k = 1, \dots, n$ gegeben:

$$\frac{d}{dt}\vec{x}(t) = \mathbf{h}(\vec{x}(t), t) + \mathbf{g}(\vec{x}(t), t)\Gamma(t), \quad (2.18)$$

Der stochastische Antrieb $\Gamma_j(t)$ wird als m -dimensionales δ -korreliertes gaußsches Rauschen mit verschwindendem Mittelwert und auf eins normierter Varianz angenommen:

$$\begin{aligned} \langle \Gamma_i(t) \rangle &= 0 \quad \forall i, \\ \langle \Gamma_i(t)\Gamma_j(t') \rangle &= 2\delta_{ij}\delta(t-t') \quad \forall i, j, \end{aligned}$$

\mathbf{h} ist das Vektorfeld, welches die deterministische Kraft repräsentiert, und \mathbf{g} ist eine $n \times m$ tensorwertige Funktion, welche die Zufallskräfte darstellt. Es ist eine bekannte Eigenschaft [18], daß Variablen $\vec{x}(t)$, die der Langevin-Gleichung (2.18) genügen, die Markov-Eigenschaft aufweisen. Dies heißt, daß die Wahrscheinlichkeitsverteilung des multidimensionalen Zustandes $\vec{x}(t')$ zu einem späteren Zeitpunkt t' nur von dem Zustand zum momentanen Zeitpunkt t abhängt. Man sagt, das System besitzt kein Gedächtnis. Dieses Konzept wird in eine quantitative Form gebracht durch die Definition für die bedingten Wahrscheinlichkeitsverteilungen p [18]:

$$p(\vec{x}_n, t_n | \vec{x}_{n-1}, t_{n-1}; \dots; \vec{x}_1, t_1) \equiv \frac{w(\vec{x}_n, t_n; \dots; \vec{x}_1, t_1)}{w(\vec{x}_{n-1}, t_{n-1}; \dots; \vec{x}_1, t_1)}, \quad (2.19)$$

wobei $\int w(\vec{x}_n, t_n; \dots; \vec{x}_1, t_1) \prod_i d\vec{x}_i$ die Wahrscheinlichkeit dafür darstellt, das System im Zustand $\vec{x}_i \dots \vec{x}_i + d\vec{x}_i$ zur Zeit t_i für alle i zu finden. Die Markov-Eigenschaft wird nun ausgedrückt durch

$$p(\vec{x}_n, t_n | \vec{x}_{n-1}, t_{n-1}; \dots; \vec{x}_1, t_1) = p(\vec{x}_n, t_n | \vec{x}_{n-1}, t_{n-1}), \quad t_n > t_{n-1} > t_1, \quad (2.20)$$

wobei t_{n-1} die nächst frühere Zeit vor t_n ist. Im Fall, daß die Markov-Eigenschaft nicht erfüllt ist, kann man durch die Erhöhung der Dimension des Zustandsvektors immer versuchen, diese Eigenschaft herzustellen. Somit können nicht-Markovische Prozesse durch die Einführung zusätzlicher Variablen eventuell Markovisch gemacht werden [17].

Fokker-Planck-Gleichung: Für lineare Langevin-Gleichungen (2.18) gibt es etablierte Verfahren, Erwartungswerte der stochastischen Variablen $\vec{x}_k(t), k = 1, \dots, n$ zu berechnen. Für nichtlineare Langevin-Gleichungen ist es allerdings wesentlich schwieriger, eine formale Lösung zu bestimmen. Daher versucht man in diesem Fall zunächst eine

Gleichung für die Wahrscheinlichkeitsverteilung $w(\vec{x}_n, t_n)$ herzuleiten, welche vorteilhafterweise linear in $w(\vec{x}_n, t_n)$ ist. Wir wollen diese Herleitung für den eindimensionalen Fall hier kurz skizzieren. Die Verallgemeinerung auf mehrdimensionale Probleme ist offensichtlich, wird aber im Verlauf der Arbeit nicht benötigt. Hierbei startet man bei der Definition der Übergangswahrscheinlichkeiten (2.19), durch welche die Wahrscheinlichkeitsverteilungen $w(x, t + \tau)$ zur Zeit $t + \tau$ und $w(x, t)$ zur Zeit t wie folgt verknüpft werden:

$$w(x, t + \tau) = \int p(x, t + \tau | x', t) w(x', t) dx', \quad (2.21)$$

wobei der Zeitindex an der stochastischen Variablen $x(t)$ weggelassen wurde. Da man daran interessiert ist, einen Ausdruck für das Differential $\partial w(x, t) / \partial t$ zu finden, sollte man die Übergangswahrscheinlichkeiten $p(x, t + \tau | x', t)$ für kleine τ kennen. Zunächst nimmt man an, daß alle Momente der Verteilung bekannt seien:

$$M_n(x', t, \tau) = \langle [\xi(t + \tau) - \xi(t)]^n \rangle_{|\xi(t)=x'} = \int (x - x')^n p(x, t + \tau | x', t) dx, \quad (2.22)$$

wobei $|\xi(t)=x'$ bedeutet, daß die Zufallsvariable ξ zur Zeit t den Wert x' annimmt. Man kann nun eine Entwicklung der Übergangswahrscheinlichkeiten in den Momenten der Verteilung angeben [18]:

$$p(x, t + \tau | x', t) = \left[1 + \sum_{n=1}^{\infty} \frac{1}{n!} \left(-\frac{\partial}{\partial x} \right)^n M_n(x, t, \tau) \right] \delta(x - x'). \quad (2.23)$$

Setzt man (2.23) in Gleichung (2.21) ein so ergibt sich:

$$w(x, t + \tau) - w(x, t) = \frac{\partial w(x, t)}{\partial t} \tau + O(\tau^2) = \sum_{n=1}^{\infty} \left(-\frac{\partial}{\partial x} \right)^n [M_n(x, t, \tau) / n!] w(x, t) \quad (2.24)$$

Als letzter Schritt wird nun noch angenommen, die Momente M_n können in eine Taylor-Reihe bezüglich τ entwickelt werden ($n \geq 1$):

$$M_n(x, t, \tau) / n! = D^{(n)}(x, t) \tau + O(\tau^2). \quad (2.25)$$

Wenn man nur die linearen Terme in Betracht zieht, ergibt sich:

$$\frac{\partial w(x, t)}{\partial t} = \sum_{n=1}^{\infty} \left(-\frac{\partial}{\partial x} \right)^n D^{(n)}(x, t) w(x, t). \quad (2.26)$$

Gleichung (2.26) ist die Kramers-Moyal-Entwicklung. Für die Lösung dieser Gleichung ist es nun wichtig zu wissen, wieviele Terme in der Entwicklung zu berücksichtigen sind. Hierfür wird eine wichtige Information durch das Theorem von Pawula gegeben [18]. Dieses besagt, daß für eine positive Übergangswahrscheinlichkeit die Entwicklung

entweder nach dem ersten oder zweiten Term endet oder anderenfalls eine unendliche Anzahl von Termen enthält. Wenn die Reihe nach dem zweiten Term endet, bezeichnet man (2.26) als Fokker-Planck-Gleichung, wobei man $D^{(1)}$ als Driftkoeffizienten und $D^{(2)}$ als Diffusionskoeffizienten bezeichnet. Die Fokker-Planck-Gleichung schreibt sich somit:

$$\frac{\partial w(x, t)}{\partial t} = -\frac{\partial}{\partial x} D^{(1)}(x, t)w(x, t) + \frac{\partial^2}{\partial x^2} D^{(2)}(x, t)w(x, t). \quad (2.27)$$

Für die Langevin-Gleichung (2.18) kann gezeigt werden, daß nur die ersten beiden Summanden von Null verschieden sind und somit eine Fokker-Planck-Gleichung gilt. Im eindimensionalen Fall gilt dann die folgende Beziehung zwischen der Langevin-Gleichung und der entsprechenden Fokker-Planck-Gleichung in der Definition von Stratonovich:

$$\begin{aligned} D^{(1)}(x(t), t) &= f(x(t), t) + g(x(t), t) \frac{\partial}{\partial x} g(x(t), t) \\ D^{(2)}(x(t), t) &= g(x(t), t)^2. \end{aligned}$$

Gleichung (2.25) zeigt, daß die vollständige Information über ein durch eine Fokker-Planck-Gleichung beschriebenes stochastisches dynamisches System gegeben ist, wenn die ersten beiden Momente der Wahrscheinlichkeitsverteilung bekannt sind. Diese kann man sehr leicht aus den gemessenen Zeitreihen der unabhängigen Variablen bestimmen. Bevor wir für einige numerische Beispiele demonstrieren, wie Drift- und Diffusionskoeffizienten aus einer Zeitreihe bestimmt werden können, wollen wir ein wichtiges technisches Detail besprechen. Gleichung (2.25) ist nur gültig im Grenzübergang verschwindender τ . Für endliche Werte von τ besteht besonders bei der Schätzung des Diffusionskoeffizienten die Gefahr, künstliche Struktur zu "rekonstruieren". Wir wollen hierzu die Übergangswahrscheinlichkeit $p(x, t + \tau | x', t)$ nach τ entwickeln, wobei wir zwei Ergebnisse aus [18] zitieren. Für kleine τ gilt in linearer Näherung:

$$p(x, t + \tau | x', t) = [1 - \frac{\partial}{\partial x} D^{(1)}(x(t), t)\tau + \frac{\partial^2}{\partial x^2} D^{(2)}(x(t), t)\tau + O(\tau^2)]\delta(x - x').$$

Ersetzt man nun die δ -Funktion durch deren Darstellung als Fourier-Integral, so ergibt sich für kleine τ :

$$p(x, \tau | x_0, 0) = \frac{1}{\sqrt{4\pi D^{(2)}(x_0)\tau}} e^{-\frac{(x-x_0-D^{(1)}(x_0)\tau)^2}{4D^{(2)}(x_0)\tau}}. \quad (2.28)$$

Setzt man dies in Gleichung (2.22) ein, so gilt:

$$M_n(x, t, \tau) = [-i\sqrt{D^{(2)}(x(t), t)\tau}]^n H_n \left[\frac{1}{2}iD^{(1)}(x(t), t)\sqrt{\tau/D^{(2)}(x(t), t)} \right].$$

Dabei sind H_n die Hermite-Polynome ($H_0 = 1$, $H_1 = 2x$, $H_2 = 4x^2 - 2$, ...). Man erhält somit für die Momente folgende Ausdrücke:

$$M_1(x, t, \tau) = \tau D^{(1)}(x(t), t) \quad (2.29)$$

$$M_2(x, t, \tau) = 2\tau D^{(2)}(x(t), t) + \tau^2 D^{(1)}(x(t), t)^2. \quad (2.30)$$

Damit ist $D^{(1)}(x(t), t)^2$ der Beitrag zweiter Ordnung in τ zu $D^{(2)}(x(t), t)$. Nach $D^{(2)}(x(t), t)$ aufgelöst, ergibt sich aus der obigen Beziehung:

$$D^{(2)}(x(t), t) = \frac{1}{2\tau} \left(M_2(x, t, \tau) - \tau^2 D^{(1)}(x(t), t)^2 \right). \quad (2.31)$$

Abhängig vom Größenverhältnis von $D^{(1)}(x(t), t)$ und $D^{(2)}(x(t), t)$ kann die obige Abschätzung ziemlich schlecht werden, wenn man den zweiten Beitrag in Gleichung (2.31) vernachlässigt. Diese Korrektur berücksichtigt die Tatsache, daß ein Bündel von Trajektorien, welches bei $x(t=0) = x_0$ startet, sich entsprechend $D^{(2)}(x(t), t)$ verbreitert und gleichzeitig einen Drift entsprechend $D^{(1)}(x(t), t)$ erfährt. Diese Verschiebung des Bündels ist von der Ordnung τ , wogegen die Verbreiterung von der Ordnung $\sqrt{\tau}$ ist. Damit ist garantiert, daß im Grenzwert $\tau \rightarrow 0$ schon Gleichung (2.25) eine richtige Abschätzung für $D^{(2)}(x(t), t)$ liefert, was allerdings für endliche τ nicht korrekt ist. Es ist somit nicht das zweite Moment, sondern das zentrale zweite Moment (nach Abzug des Mittelwertes), welches die korrekte Abschätzung für $D^{(2)}(x(t), t)$ ergibt. Anderenfalls beobachtet man möglicherweise nur das Quadrat des Drifts.

Diese Entwicklung in τ ermöglicht es, einen kritischen Wert für τ abzuschätzen, ab welchem der quadratische Term vernachlässigbar ist. Es sollte gelten:

$$\tau \ll \frac{2D^{(2)}(x(t), t)}{D^{(1)}(x(t), t)^2}. \quad (2.32)$$

Somit ist der kritische Wert für τ abhängig von x .

Die numerischen Analysen, die wir weiter unten durchführen, zeigen, daß Gleichung (2.31) nicht die einzige Korrektur zu $D^{(2)}$ von der Ordnung τ ist, obwohl aus der Entwicklung der Fokker-Planck-Gleichung nichts Weiteres gewonnen werden kann. Der Grund ist, daß die zweite wesentliche Korrektur, die von $D^{(1)}$ stammt, ein nichtlokaler Effekt ist. Bevor wir einen mathematischen Ausdruck für diese Korrektur angeben, möchten wir die physikalische Ursache dieses Effektes beschreiben. Man starte mit dem oben eingeführten Trajektorienbündel exakt auf dem Mittelpunkt (lokalen Maximum) eines Doppelmuldenpotentials. Das Paket von Anfangsbedingungen wird sich nur aufgrund der Diffusion ausbreiten, da an diesem Ort der Drift Null ist. Bereits nach infinitesimaler Zeit werden die Teilchen jedoch einen Drift in Richtung eines der Potentialminima erfahren, der jedoch auf den beiden Seiten des lokalen Maximums ein unterschiedliches Vorzeichen hat. Somit wird die Diffusion erhöht. Der umgekehrte Effekt ergibt sich in den Potentialminima, wo Teilchen von beiden Seiten durch den Drift

zum Minimum zurückgetrieben werden. Intuitiv sollte die Krümmung des Potentials für diesen Effekt verantwortlich sein.

Um einen analytischen Ausdruck hierfür zu finden, müssen wir die Lösung der Fokker-Planck-Gleichung nach endlicher Zeit diskutieren, wenn man als Anfangsbedingung \vec{x}_0 eine δ -Verteilung ansetzt. Es ist einfach zu überprüfen, daß

$$w(x, \tau) = \frac{1}{\sqrt{4\pi D^{(2)}\tau(1 + \tau D^{(1)'})}} e^{-\frac{(x-x_0 - D^{(1)}(x_0)\tau)^2}{4D^{(2)}\tau(1 + \tau D^{(1)'})}} \quad (2.33)$$

die Anfangsbedingung erfüllt sowie die Gesamtwahrscheinlichkeit erhält (das Integral über x ist normiert). Hierbei ist $D^{(1)'} := d/dx D^{(1)}(x)$. Setzt man (2.33) in die Fokker-Planck-Gleichung (2.27) ein und vernachlässigt die Ortsabhängigkeit von $D^{(2)}$, so zeigt sich, daß der $D^{(1)'}\rho(x)$ -Term, welcher bei entsprechendem Einsetzen von Gleichung (2.28) auftritt und von der Ordnung τ ist, durch diese Form der Übergangswahrscheinlichkeit annulliert wird. Es ergibt sich nach Division durch $w(x, \tau)$ und Multiplikation mit $4\tau D^{(2)}(1 + \tau D^{(1)'})$:

$$\begin{aligned} & -2\mathbf{D}^{(2)}(1 + 2\tau\mathbf{D}^{(1)'}) + 2D^{(1)}(x - x_0 - \tau D^{(1)}(x_0)) \\ & + \frac{(x - x_0 - \tau D^{(1)}(x_0))^2 4D^{(2)}(1 + 2\tau D^{(1)'})}{4\tau D^{(2)}(1 + \tau D^{(1)'})} + \tau\mathbf{D}^{(1)'} 4\mathbf{D}^{(2)}(1 + \tau D^{(1)'}) \\ & - 2D^{(1)}(x - x_0 - \tau D^{(1)}(x_0)) + 2D^{(2)} - \frac{4D^{(2)}(x - x_0 - \tau D^{(1)}(x_0))^2}{4\tau D^{(2)}(1 + \tau D^{(1)'})} = 0. \end{aligned} \quad (2.34)$$

Die entscheidenden Terme, welche durch die Anwendung der Verteilung (2.33) annulliert werden, sind fett gedruckt. Man erkennt, daß sich nun alle Terme der Ordnung τ aufheben. Beachtet man, daß die Differenz $x - x_0$ von der Ordnung τ ist, so ergibt sich für den Ausdruck $[(x - x_0 - D^{(1)}(x_0)\tau)^2 4D^{(2)} 2\tau D^{(1)'}] / [4\tau D^{(2)}(1 + \tau D^{(1)'})]$ ein quadratisches Verhalten in τ . Damit ist dieser Term eine Ordnung höher in τ als die in Gleichung (2.34) fett gedruckten Terme. Es zeigt sich daher, daß Gleichung (2.33) die Fokker-Planck-Gleichung um eine Ordnung genauer löst als (2.28). Die Korrektur kann wie erwartet die Diffusion verstärken oder unterdrücken, abhängig vom Vorzeichen der Ableitung des Drift-Terms.

Mit dem numerischen Algorithmus zur Bestimmung von $D^{(2)}$ berechnet man Schätzwerte für \dot{x} bei festem x . Daher konstruiert man implizit die Zeitentwicklung der Wahrscheinlichkeitsentwicklung von Beispieltrajektorien mit der Anfangsverteilung $w(x, 0) = \delta(\vec{x} - \vec{x}_0)$. Daher erwartet man nach endlicher, aber kurzer Zeit τ die Gültigkeit von Gleichung (2.33) und somit die folgende Korrektur für den Diffusionsterm:

$$D_{ij}^{(2)} = \frac{\frac{1}{2\tau} \left(\langle (\vec{x}_{t+\tau} - \vec{x}_t)_i (\vec{x}_{t+\tau} - \vec{x}_t)_j \rangle_{\vec{x}} - \tau^2 \hat{D}_i^{(1)}(\vec{x}) \hat{D}_j^{(1)}(\vec{x}) \right)}{\left(1 + \tau \frac{d\hat{D}_i^{(1)}(\vec{x})}{dx_j} \right)} \quad (2.35)$$

Ein anderer Weg, sich der Gültigkeit der obigen Gleichung zu versichern, ist den effektiven Drift $D_{eff}^{(2)} = D^{(2)}(1 + \tau D^{(1)'})$ in Gleichung (2.30) einzusetzen und diese nach $D^{(2)}$ aufzulösen. Die räumlichen Ableitungen von $D^{(1)}$ können mit hoher statistischer Genauigkeit ermittelt werden, indem eine glatte Kurve durch den empirisch geschätzten Verlauf von $D^{(1)}$ gefittet wird.

Wir wollen die Schätzungen von $D^{(1)}$ und $D^{(2)}$ für verschiedene eindimensionale stochastische dynamische Systeme demonstrieren. Als erstes Beispiel betrachten wir Brownsche Bewegung:

$$\frac{dx}{dt} = \alpha x(t) + \beta \Gamma(t). \quad (2.36)$$

Abbildung 2.4 zeigt einen Teil der iterierten Zeitreihe des Systems (2.36) mit $\alpha = -0.1$ und $\beta = 0.05$. Zur Integration wurde ein Euler-Integrator mit der Schrittweite 0.01 verwendet. In Abbildung 2.5 sind die gefitteten Drift- und Diffusionskoeffizienten für verschiedene Werte von τ sowie die jeweilige exakte Funktion in Abhängigkeit von x dargestellt, wobei die Schätzung mit Gl. (2.25) durchgeführt wurde. Dabei sind die oben hergeleiteten Korrekturen noch nicht berücksichtigt. Man erkennt, daß die funktionale Form des Driftkoeffizienten recht genau geschätzt wird unabhängig von τ und x . Dagegen ergeben sich für den Diffusionskoeffizienten deutliche Abweichungen vom theoretischen Verhalten für große Werte von τ . $D^{(2)}(x(t), t)$ verhält sich in etwa wie x^2 wie man es nach Gleichung (2.30) erwarten würde, wenn $D^{(1)}(x(t), t)$ linear von x abhängt. Die zusätzlich zu beobachtende Verschiebung des Minimums von $D^{(2)}(x(t), t)$ rührt von der zweiten Korrektur her ($d/dx D^{(1)}(x(t), t)$ ist in diesem Falle konstant). Setzt man typische Werte von $D^{(1)}$ und $D^{(2)}$ in (2.32) ein, so ergibt sich ein kritischer Wert für τ von der Größenordnung 1. Daher ergeben ohne die Berücksichtigung der Korrekturen nur die Fits mit $\tau = 0.1$ sinnvolle Ergebnisse für $D^{(2)}$. In Abb. 2.6 sind für $\tau = 0.5$ die Diffusionskoeffizienten ohne Korrektur, nach Subtraktion des Quadrats des Drifts Gl.(2.31) sowie nach Multiplikation mit $(1 + \tau D^{(1)'})$ Gl.(2.35) gezeigt. Man erkennt, daß die zweite Korrektur von vergleichbarer Größenordnung zur ersten ist. Als zweites Beispiel betrachten wir einen nichtlinearen Oszillator:

$$\frac{dx}{dt} = -\alpha x(t) - x(t)^3 + \beta \Gamma(t). \quad (2.37)$$

Abbildung 2.7 zeigt einen Teil der iterierten Zeitreihe des Systems (2.37). Abbildung 2.8 zeigt die Drift- und Diffusionskoeffizienten für zwei verschiedene Werte von τ , welche wieder durch Anwendung der Gleichungen (2.22) und (2.25) (ohne Korrekturen) bestimmt wurden. Man erhält für den Diffusionskoeffizienten wieder deutliche Abweichungen vom theoretischen Verhalten für große τ . Es ist auch die Funktion $D^{(2)}(x(t), t) + (1/2)D^{(1)}(x(t), t)^2$ dargestellt, die einen wesentlichen Teil des qualitativen Verhaltens von $D^{(2)}(x(t), t)$ für $\tau = 1$ erklärt. Die Abschätzung (2.32) ergibt auch hier, daß $\tau \ll 1$ sein sollte, für nicht zu große Werte von x . Daher ergibt der unkorrigierte

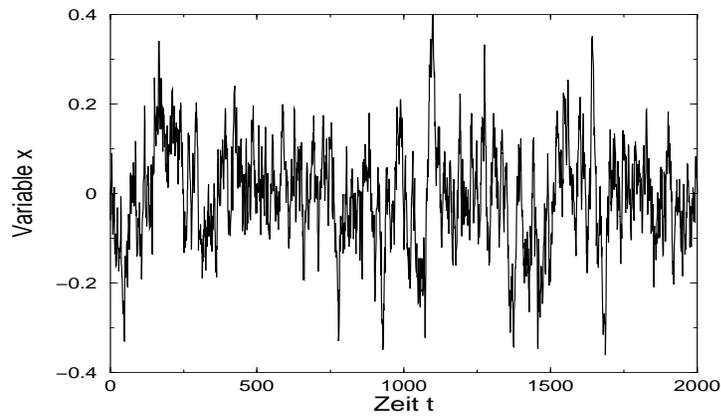


Abbildung 2.4: Zeitreihe für die Lösung $x(t)$ der Langevin-Gleichung (2.37) mit $\alpha = 0.1$ und $\beta = 0.05$.

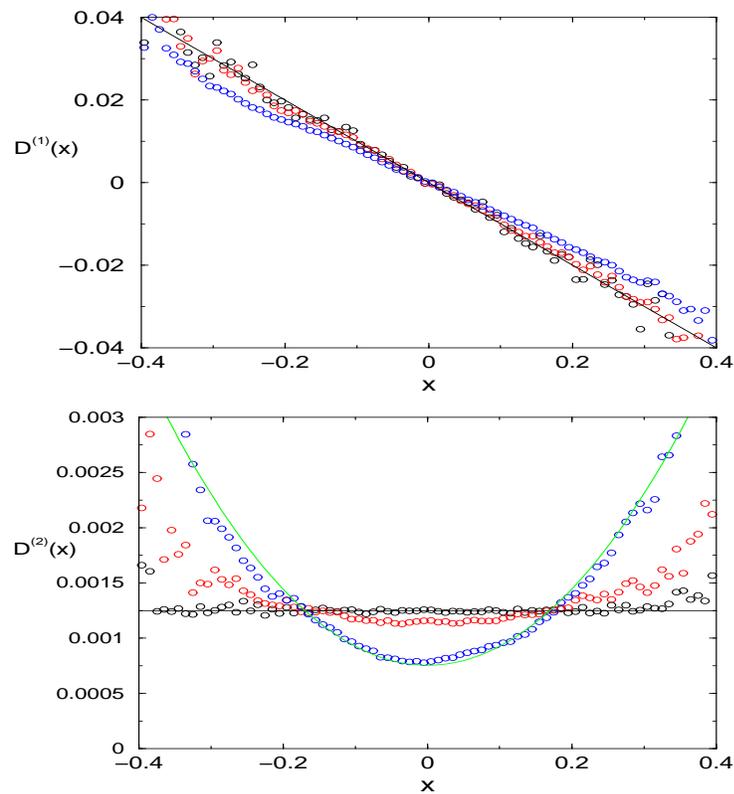


Abbildung 2.5: Driftkoeffizient und Diffusionskoeffizient für das System (2.36) aus den Daten ermittelt (Gl. (2.25)) mit $\tau = 0.1$ (schwarze Punkte), mit $\tau = 1.0$ (rote Punkte) und mit $\tau = 5.0$ (blaue Punkte) und die exakten Werte $D^{(1)} = 0.1x$ und $D^{(2)} = 0.00125$ (schwarze Linien) sowie ein Fit $D^{(2)} = a + bx^2$ (grüne Linie).

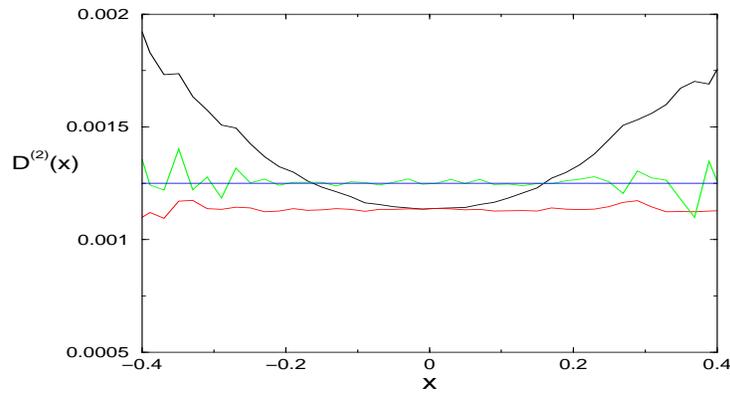


Abbildung 2.6: Diffusionskoeffizient für das System (2.36) aus den Daten ermittelt mit $\tau = 0.5$ nach Gleichung (2.25) (schwarz), nach Gleichung (2.31) (rot) und nach Gleichung (2.35) (grün) sowie exakter Wert (blau).

Fit mit $\tau = 0.1$ zufriedenstellende Ergebnisse für kleine x . An den Rändern (für große x) erkennt man jedoch, daß auch $\tau = 0.1$ nicht ausreichend klein ist. In Abb. 2.9 sind für $\tau = 1.0$ die Diffusionskoeffizienten ohne Korrektur, nach Subtraktion des Quadrats des Drifts sowie nach Multiplikation mit $(1 + \tau D^{(1)'})$ gezeigt. Man erkennt wieder, daß die zweite Korrektur von vergleichbarer Größenordnung zur ersten ist.

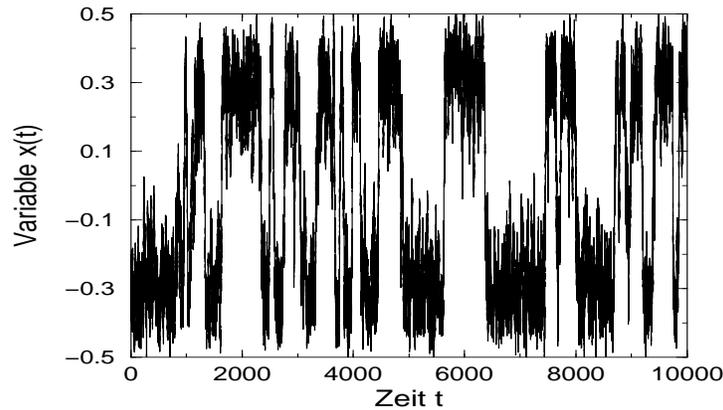


Abbildung 2.7: Zeitreihe für die Lösung $x(t)$ der Langevin-Gleichung (2.37) mit $\alpha = -0.1$ und $\beta = 0.05$.

Für eindimensionale Probleme empfehlen wir einen zweiten unabhängigen Schätzer für $D^{(2)}$, der auf einer direkten Auswertung der Fokker-Planck-Gleichung basiert. Wenn die Zeitreihe eine invariante Dichte repräsentiert, gilt $\dot{\rho} = 0$ und somit:

$$\int_{-\infty}^x \frac{d}{dx'} D^{(1)}(x') \rho(x') dx' = \int_{-\infty}^x \frac{d^2}{dx'^2} D^{(2)}(x') \rho(x') dx'. \quad (2.38)$$

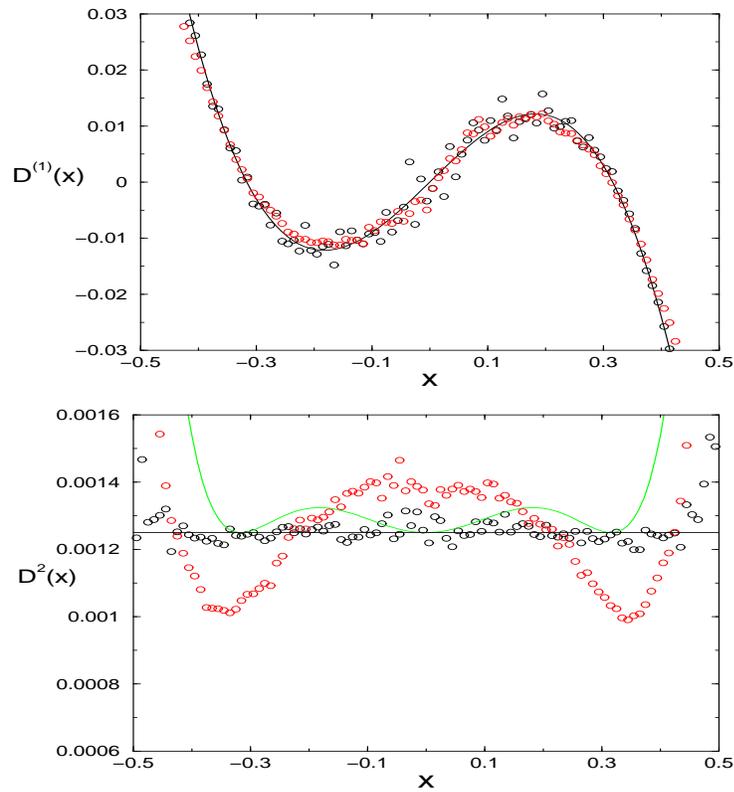


Abbildung 2.8: Driftkoeffizient und Diffusionskoeffizient für das System (2.37) aus den Daten ermittelt (Gl. (2.25)) mit $\tau = 0.1$ (schwarze Punkte) und mit $\tau = 1.0$ (rote Punkte) und die exakten Werte $D^{(1)} = 0.1x - x^3$ und $D^{(2)} = 0.00125$ (schwarze Linien) sowie $D^{(2)}(x(t), t) + (1/2)D^{(1)}(x(t), t)^2$ (grüne Linie).

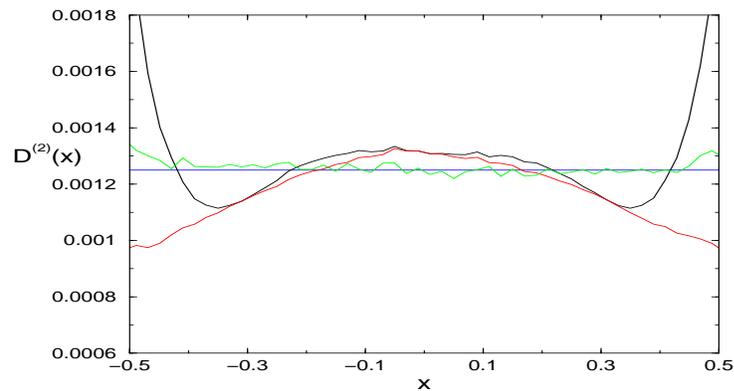


Abbildung 2.9: Diffusionskoeffizient für das System (2.37) aus den Daten ermittelt mit $\tau = 0.5$ nach Gleichung (2.25) (schwarz), nach Gleichung (2.31) (rot) und nach Gleichung (2.35) (grün) sowie exakter Wert (blau).

Wenn die Verteilung gaußförmig bzw. exponentiell abfällt und Drift- und Diffusionskoeffizienten sich algebraisch verhalten, verschwinden jeweils die unteren Grenzen des Integrals. Schätzt man den Drift-Term wie vorher und ermittelt die invariante Dichte ρ aus den Daten, so ergibt sich für die Diffusion:

$$D^{(2)} = \frac{\int_{-\infty}^x D^{(1)}(x')\rho(x')x'}{\rho(x)}, \tag{2.39}$$

wenn man wieder annimmt, daß $D^{(2)}(-\infty)\rho(-\infty) \approx 0$, was für gaußisch abfallendes ρ und algebraisches $D^{(2)}$ wieder gerechtfertigt ist. Die Abtastrate τ tritt in dieser Gleichung nur über die Schätzung von $D^{(1)}$ auf, daher ist diese Methode besonders gut geeignet für große τ . Für das System (2.37) wurde diese alternative Methode (2.39) zur Bestimmung von $D^{(2)}$ genutzt. Die Ergebnisse für $\tau = 0.5$ sind in Abbildung 2.10 dargestellt.

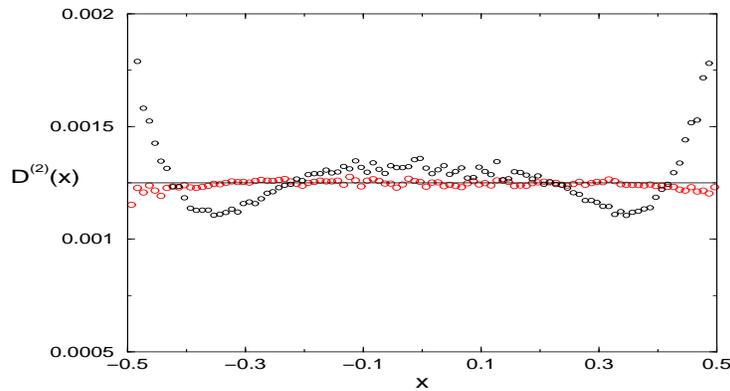


Abbildung 2.10: Diffusionskoeffizient für das System (2.37) aus den Daten ermittelt mit $\tau = 0.5$ nach Gleichung (2.25) (schwarze Punkte) und nach Gleichung (2.39) (rote Punkte) und der exakte Wert $D^{(2)} = 0.00125$ (schwarze Linie).

Es ist ersichtlich, daß für eindimensionale Systeme sehr effizient Fokker-Planck-Gleichungen aus Zeitreihen ermittelt werden können, wenn τ nicht deutlich größer ist als die typische Zeitskala des Systems. Dieses Verfahren ist prinzipiell auch für höherdimensionale Systeme anwendbar [5], allerdings wird die Statistik schlechter, da in Gleichung (2.22) bedingte Wahrscheinlichkeiten für mehrdimensionale Bedingungen auszuwerten sind. Ein weiteres Problem besteht darin, daß in Feldmessungen oder Laborexperimenten häufig lediglich eine skalare Zeitreihe gemessen wird. Daher wollen wir in den folgenden Abschnitten untersuchen, welche Information aus einer skalaren Messung eines multivariaten Langevin-Prozesses gewonnen werden kann.

2.2.2 Markov-Modelle von einer Zeitreihe

Wir haben in Abschnitt 2.1.3 gesehen, daß es in deterministischen dynamischen Systemen möglich ist, den Phasenraum eines Systems mit mehreren Freiheitsgraden aus einer skalaren Zeitreihe zu rekonstruieren. Aus aufeinanderfolgenden Messungen y_n werden Zeitverzögerungsvektoren $\vec{y}_n = (y_n, y_{(n-1)\tau}, \dots, y_{(n-m+1)\tau})$ gebildet, welche den ursprünglichen Bewegungsgleichungen genügen. Vorhersagen können auf der Basis der Nachbarschaft im Phasenraum durchgeführt werden.

Eine einfache Beobachtung zeigt, daß dieselbe Klasse von Vorhersagealgorithmen nicht nur im Fall m -dimensionaler Einbettungsräume anwendbar ist, sondern auch in der allgemeineren Situation eines Markov-Prozesses der Ordnung m . Ein skalarer Markov-Prozeß der Ordnung m in diskreter Zeit ist definiert durch die Tatsache, daß für alle aufeinanderfolgenden Zeiten die Übergangswahrscheinlichkeiten die folgende Bedingung erfüllen:

$$p(y_n, t_n | y_{n-1}, t_{n-1}; \dots; y_1, t_1) = p(y_n, t_n | y_{n-1}, t_{n-1}; \dots; y_{n-m+1}, t_{n-m+1}), \quad t_n > t_{n-1} > t_1, \quad (2.40)$$

d.h., die Übergangswahrscheinlichkeit hängt nur von den letzten m Beobachtungen ab. Da die Werte dieser Übergangswahrscheinlichkeiten beliebig sein können, ist solch ein Markov-Prozeß wesentlich allgemeiner als die oben eingeführten AR-Modelle.

Wir bemerken nun, daß die unabhängigen Übergangswahrscheinlichkeiten in Gleichung (2.40) genau die Elemente des Einbettungsvektors \vec{y}_n sind, wenn wir die Zeiten t_k mit den ganzzahligen Vielfachen der Zeitverzögerung τ identifizieren. Wir werden die Zeiten t_k daher im folgenden weglassen und die Übergangswahrscheinlichkeiten als $p(y_{n+1} | \vec{y}_n)$ schreiben. Um diese Übergangswahrscheinlichkeiten aus den Daten zu gewinnen, ist wieder die Annahme erforderlich, daß deren Abhängigkeit von \vec{y} glatt ist. Es ist dann sinnvoll die folgende Approximation zu schreiben:

$$p(y_{k+1} | \vec{y}_k) \approx \hat{p}(y_{n+1} | \vec{y}_n) \quad \forall \vec{y}_k \in \mathcal{U}_n, \quad (2.41)$$

d.h., wir benutzen wieder die lokal konstante Näherung. $\hat{p}(y_{n+1})$ kann von den beobachteten Werten von \vec{y}_{k+1} geschätzt werden, den "Zukünften" der Elemente $\vec{y}_k \in \mathcal{U}_n$. Im deterministischen Fall unterliegen diese Elemente der δ -Verteilung, $p(y_{n+1} | \vec{y}_n) = \delta(y_{n+1} - G(\vec{y}_n))$, und die numerische Abschätzung in Gleichung (2.41) liefert demnach eine schmale Verteilung. Dagegen kann die Verteilung im stochastischen Fall eine endliche Breite haben und auch multimodal sein. Auf welche Weise $\hat{p}(y_{n+1})$ ausgewertet wird, hängt von der jeweiligen Anwendung ab.

Wenn Vorhersagen das Ziel sind, ist typischerweise die quadratische Fehlerfunktion $e^2 = \sum (y_{n+1} - \hat{y}_{n+1})^2$ zu minimieren. Der ideale Vorhersager ist dann $\hat{y} = \int p(y'_{n+1} | \vec{y}_n) dy'_{n+1}$, d.h. genau der lokal konstante Vorhersager aus Gleichung (2.14). Abhängig von der Form der Verteilung $p(y_{n+1} | \vec{y}_n)$ kann der Mittelwert ein sehr unwahrscheinlicher Wert für y_{n+1} sein, und eine Iteration dieser Vorhersage kann eine sehr untypische Sequenz

von y 's ergeben, die sich deutlich vom Verhalten der eigentlichen Daten unterscheidet. Modellierung würde in diesem Fall bedeuten, einen zufälligen Wert aus der beobachteten Verteilung $p(y_{n+1}|\vec{y}_n)$ zu wählen [19]. Wenn der Mittelwert ein unwahrscheinlicher Vertreter der vollen Verteilung ist, sind Modellierung und Vorhersage zwei sehr verschiedene Aufgaben. In solchen Situationen kann es sinnvoll sein, auch andere Kostenfunktionen für Vorhersagen zu untersuchen. Wenn der Fehler genauso oft positiv wie negativ sein soll, ist der Median der optimale Vorhersager. Wenn man besonders große Fehler vermeiden will, wird die optimale Kostenfunktion das Maximum aller Fehler sein. In diesem Fall ist die optimale Vorhersage $\hat{y}_{n+1} = \frac{1}{2}(y_{max} - y_{min})$, wobei $y_{max/min}$ die größten/kleinsten Werte sind, für die $p(y_{n+1}|\vec{y}_n)$ von Null verschieden ist.

Unabhängig von der gewählten Kostenfunktion und von der Tatsache, ob der betrachtete Prozeß stochastisch oder deterministisch ist, ist die Breite der Verteilung $p(y_{n+1}|\vec{y}_n)$ ein direktes Maß für die zu erwartende Genauigkeit der Vorhersage. Je größer die Spannweite der y_{k+1} ist, desto stärker ist die lokale Instabilität und daher die mögliche Abweichung von y_{n+1} vom Mittelwert der Verteilung. Dies ist in Abb. 2.11 dargestellt für Daten generiert durch eine nichtlineare Langevin-Gleichung (2.43) und durch ein deterministisches chaotisches Modell (Lorenz-Modell in Gleichung (2.9)) mit Meßrauschen. Bei letzterem ist die Breite der Verteilung erzeugt durch die Divergenz benachbarter Trajektorien, was unter der Bezeichnung lokaler Lyapunov-Exponent bekannt ist. Daher ist es sinnvoll, die Standardabweichung der Verteilung $p(y_{n+1}|\vec{y}_n)$ als Kriterium für die Güte der jeweiligen individuellen Vorhersage zu verwenden. Diese Untersuchung bestätigt somit auch noch einmal, daß individuelle Vorhersagen in einem gegebenen System sehr unterschiedliche Zuverlässigkeit besitzen können.

2.2.3 Ist eine skalare Beobachtung einer multivariaten Langevin-Gleichung Markovisch?

Hat man die Hypothese, daß eine skalare Zeitreihe die Observable eines vektorwertigen deterministischen Systems darstellt, erlaubt das Einbettungstheorem, den multidimensionalen Zustandsraum zu rekonstruieren. In diesem sind die deterministischen Bewegungsgleichungen wiederhergestellt. Das entsprechende Problem in einem Markov-Prozeß ist das folgende: Es seien ein multidimensionaler Markov-Prozeß und eine skalare Beobachtung gegeben. Repräsentiert die Zeitreihe einer solchen skalaren Beobachtung einen skalaren Markov-Prozeß endlicher Ordnung m in der Zeit? Die allgemeine Antwort auf diese Frage ist negativ [17]. Um die Verbindung zur deterministischen Situation so deutlich wie möglich darzustellen, betrachten wir als Erzeuger der Markov-Dynamik eine Langevin-Gleichung vom Typ:

$$\dot{\vec{y}} = \vec{f}(\vec{y}(t)) + \mathbf{G}(\vec{y}(t))\vec{\Gamma}(t), \quad (2.42)$$

welche die Verallgemeinerung eines zeit-kontinuierlichen dynamischen Systems mit zusätzlicher Zufallskraft $\mathbf{G}(\vec{y}(t))\vec{\Gamma}(t)$ beschreibt. Die Zufallszahlen $\vec{\Gamma}(t) \in \mathcal{R}^l$ sind Gauß-

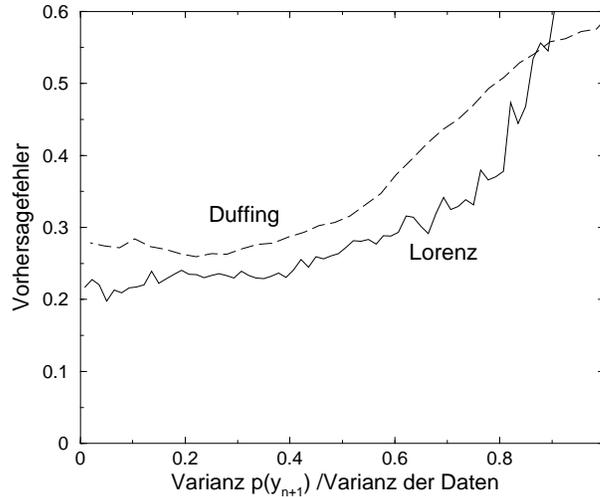


Abbildung 2.11: Mittlerer Vorhersagefehler $\sqrt{\langle (\hat{y}_{n+1} - y_{n+1})^2 \rangle}$ als Funktion der Standardabweichung der Verteilung von $p(y_{n+1}|\vec{y}_n)$. Durchgezogene Linie: Zeitreihe des Lorenz-Systems Eq.(2.9) mit Meßrauschen, gestrichelte Linie: Duffing-System mit dynamischem Rauschen Eq.(2.43).

verteilt weißes Rauschen mit $\langle \Gamma_k(t)\Gamma_{k'}(t') \rangle = \delta_{k,k'}\delta(t-t')$, und $\mathbf{G}(\vec{y})$ ist ein $(n \times l)$ -dimensionaler Tensor. In einem n -dimensionalen Zustandsraum mit Zustandsvektoren $\vec{y}(t) = (y_1(t), \dots, y_n(t))$ beschreibt Gleichung (2.42) einen Markov-Prozeß der Ordnung m , da die Übergangswahrscheinlichkeit $p(\vec{y}(t+\delta)|\vec{y}(t))$ durch die Angabe der Anfangsbedingung $\vec{y}(t)$ eindeutig festgelegt ist. Kann die Information, welche der n -dimensionale Zustandsvektor $\vec{y}(t)$ enthält, irgendwie durch die Messung von nur s Komponenten dieses Vektors mit $s < n$ zu aufeinanderfolgenden Zeiten rekonstruiert werden? Dies würde die Existenz eines endlichen Gedächtnisses in diesen s Komponenten voraussetzen. Im allgemeinen ist die Annahme, der reduzierte Phasenraum der Dimension s habe ein endliches Gedächtnis, falsch. Man kann nicht erwarten, daß die Kenntnis nur weniger Komponenten des Zustandsraumes ausreicht, um die zukünftige Wahrscheinlichkeitsverteilung dieser Komponenten festzulegen [17].

Wir werden jedoch zeigen, daß in bestimmten Situationen die Dynamik in einem Einbettungsraum der Markov-Eigenschaft genügt und in einer größeren Anzahl von Situationen die Markov-Eigenschaft in einem solchen Raum näherungsweise gültig ist, da das Gedächtnis hinreichend schnell abfällt. Ob eine skalare Beobachtung einer vektorwertigen Langevin-Gleichung die Markov-Eigenschaft aufweist, ist sowohl vom betrachteten System als auch von der Meßfunktion abhängig.

Wir demonstrieren unsere Argumentation zunächst anhand eines einfachen Bei-

spiels. Das Duffing-System

$$\begin{aligned}\frac{dx}{dt} &= v(t) \\ \frac{dv}{dt} &= av(t) - x^3(t) + x + b\Gamma(t)\end{aligned}\tag{2.43}$$

beschreibt die stochastische Bewegung eines gedämpften Teilchens in einem nichtlinearen Potential, wobei $v(t)$ die Geschwindigkeit des Teilchens und $x(t)$ dessen Ort bezeichnen. Diese Gleichung definiert einen vektorwertigen Markov-Prozeß in (x, v) . Das Teilchen bewegt sich aufgrund des stochastischen Antriebs $\Gamma(t)$. Die Geschwindigkeitsänderungen des Teilchens sind durch die stochastischen Kräfte sowie durch dessen jeweiligen Ort festgelegt. Wie in van Kampen [17] dargelegt wird, hängt der Ort des Teilchens über $x(t) = \int_{-\infty}^t v(t')dt'$ von dessen Geschwindigkeit zu allen früheren Zeiten ab. Daher besitzt die Geschwindigkeit ein unendliches Gedächtnis. Man muß die Geschwindigkeit für die gesamte Vergangenheit kennen, um aus der Evolutionsgleichung für die Wahrscheinlichkeitsverteilung von v die Variable x zu eliminieren. Im deterministischen Fall (ohne stochastischen Antrieb) basiert die Rekonstruktion der Dynamik auf der Tatsache, daß die zweite Gleichung in (2.43) nach x aufgelöst werden kann, wenn eine hinreichende Anzahl von Ableitungen in y gegeben ist (hier genügen die ersten beiden). Diese Invertierungseigenschaft ist verletzt, wenn eine stochastische Kraft hinzugefügt wird, und x muß daher aus der ersten Gleichung in (2.43) ermittelt werden. Im Gegensatz zur Geschwindigkeit des Teilchens ist dessen Ort noch immer ein Markov-Prozeß zweiter Ordnung, d.h., es genügt den Ort zu den Zeiten t und $t - \delta$ zu wissen, um die Wahrscheinlichkeitsverteilung für den folgenden Wert zu schätzen. Dies basiert auf der Tatsache, daß die Kenntnis von $x(t)$ und $x(t - \delta)$ eine Abschätzung für die Geschwindigkeit $v(t)$ gibt und die Kenntnis von $x(t - 2\delta)$ keine zusätzliche Information beiträgt. Dies kann auch folgendermaßen gesehen werden: Leitet man die erste Gleichung nach t ab und ersetzt dv/dt durch die zweite und v durch die erste Gleichung, so bekommt man eine stochastische Differentialgleichung zweiter Ordnung in t für x , die einen Markov-Prozeß erzeugt.

Wir wollen diese Argumentation nun auf ein allgemeineres Beispiel anwenden, dessen Bewegungsgleichungen gegeben sind durch:

$$\begin{aligned}\frac{dx}{dt} &= f(x, y) \\ \frac{dy}{dt} &= g(x, y, z) \\ \frac{dz}{dt} &= h(x, y, z) + \Gamma(t),\end{aligned}\tag{2.44}$$

mit $f(x, y)$ invertierbar in y und $g(x, y, z)$ invertierbar in z . Diese Gleichungen beschreiben einen multivariaten Markov-Prozeß in (x, y, z) . Diese Markov-Eigenschaft ist auch

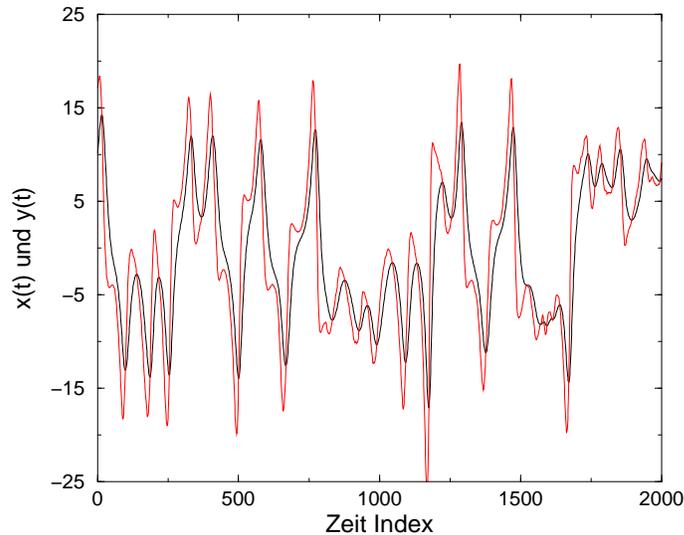


Abbildung 2.12: Die x -Koordinate (schwarz) und die y -Koordinate (rot) versus Zeit.

für die x -Koordinate allein gültig, d.h., die Kenntnis von $x(t)$, $x(t - \delta)$ und $x(t - 2\delta)$ ist ausreichend, um die Wahrscheinlichkeitsverteilung von $x(t + \delta)$ vorherzusagen. Dies kann man folgendermaßen sehen: Kennt man die x Komponente zu drei aufeinanderfolgenden Zeiten, hat man eine Abschätzung für die erste und zweite Ableitung dieser Größe nach t . Die Ableitung der ersten Gleichung von (2.44) ergibt:

$$\frac{d^2x}{dt^2} = \frac{\partial}{\partial x} f(x, y) \frac{dx}{dt} + \frac{\partial}{\partial y} f(x, y) \frac{dy}{dt}. \quad (2.45)$$

Da $x(t)$, dx/dt und d^2x/dt^2 näherungsweise bekannt sind und $f(x, y)$ implizit durch die Vergangenheit der Zeitreihe gegeben ist, legt diese Gleichung zusammen mit der ersten Gleichung von (2.44) $y(t)$ und dy/dt eindeutig fest. Berücksichtigt man zusätzlich die zweite Gleichung von (2.44), so erhält man auch z . Damit ist $x(t)$ ein Markov-Prozeß dritter Ordnung. Dies gilt nicht für die beiden anderen Koordinaten y und z . Beispiele für den Satz von Gleichungen in (2.44) sind das Lorenz- und das Rössler-System.

Abbildung 2.12 zeigt Ausschnitte der Zeitreihen der x - und der y -Koordinate des Lorenz-Systems in Gleichung (2.50). Man erkennt, daß sich die Zeitreihen sehr ähnlich verhalten. In Abb. 2.13 ist für das lokal konstante Modell der Vorhersagefehler gegen die Ordnung des Modells für x bzw. y als Einbettungskordinate aufgetragen. Das dargestellte Verhalten bestätigt die obige Argumentation. Für die x Koordinate werden optimale Vorhersagen bereits für 3-dimensionale Einbettungen erreicht, wogegen der Vorhersagefehler selbst bei einer 6-dimensionalen Einbettung noch nicht gesättigt ist, wenn y als Koordinate für die Einbettung verwendet wird.

Wir wollen nun ein einfaches nichtlineares System numerisch untersuchen, um zu zeigen,

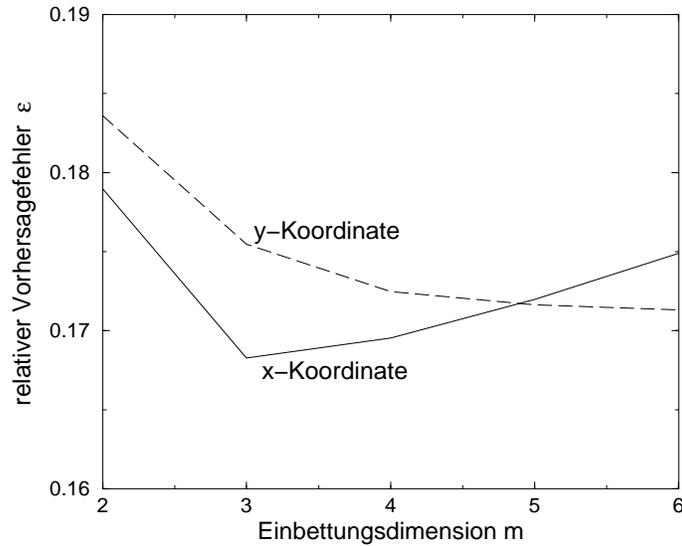


Abbildung 2.13: Mittlerer Vorhersagefehler $\sqrt{\langle(\hat{y}_{n+1} - y_{n+1})^2\rangle}$ als Funktion der Einbettungsdimension m mit x oder y als Einbettungscoordinate.

welche verschiedenen Situationen bei der Einbettung dynamisch verrauschter Daten zu erwarten sind. Wir betrachten den Van-der-Pol-Oszillator, da dieser einen stabilen Grenzzyklus als asymptotische Lösung besitzt und somit nicht chaotisch ist (damit läßt sich der Einfluß des Rauschens unabhängig von der chaotischen Divergenz der Trajektorien untersuchen):

$$\begin{aligned}\dot{x} &= y(t) + a\Gamma(t) \\ \dot{y} &= (r - x(t)^2)y(t) - x(t) + b\Gamma(t).\end{aligned}\tag{2.46}$$

Das System wurde integriert, indem entweder die erste oder die zweite Gleichung durch weißes Gaußsches Rauschen mit der Varianz normiert auf eins angetrieben wurde. Die Parameter sind entweder $(a, b) = (0.5, 0)$ oder $(a, b) = (0, 0.5)$ und $r = 3.0$. Für die numerische Integration wurde ein Euler-Integrator mit einer Schrittweite von 0.01 verwendet, in dem das Rauschen mit der Wurzel der Schrittweite normiert wird (Ito- und Stratonovich-Definition sind hier identisch, da a und b unabhängig von x und y sind.). Verwendet man dieselben Argumente wie zuvor für das Duffing-System, ergeben sich drei unterschiedliche Situationen, die wir im folgenden numerisch untersuchen wollen: (a) Das Rauschen wird zur unbeobachteten Koordinate addiert, und die beobachtete Variable unterliegt der Markov-Eigenschaft. (b) Die beobachtete Variable ist verrauscht, und die Markov-Eigenschaft ist daher ungültig. (c) Die unbeobachtete Variable ist verrauscht, und die beobachtete Koordinate ist nicht Markovisch.

Zunächst zeigen wir in Abb. 2.14 für diese drei Fälle die Phasenportraits. Im Fall, daß das Rauschen zur y -Variablen addiert ist, betrachten wir in Abb. 2.14 (a) das Phasen-

portrait, wobei x für die Einbettung benutzt wurde. In Abb. 2.14 (b) ist wieder x die Einbettungskordinate, aber das Rauschen ist jetzt in x . In 2.14 (c) ist derselbe Fall dargestellt, wenn y für die Einbettung benutzt wird. Im ersten Fall (a) sind nur schwache Abweichungen vom Grenzzzyklus erkennbar. Das zweite Portrait (b) wirkt wie ein Grenzzzyklus mit Beobachtungsrauschen, wobei im dritten Bild deutlich nichttriviale höherdimensionale Struktur erkennbar ist.

Im nächsten Schritt analysieren wir die drei Situationen mit Hilfe der oben eingeführten lokal konstanten Vorhersagealgorithmen. Wie oben dargestellt, werden wir Vorhersagen in Abhängigkeit von der Einbettungsdimension und der Zeitverzögerung durchführen. Normalerweise sind bei höheren Einbettungsdimensionen größere Nachbarschaftsradien ϵ notwendig, um eine bestimmte Anzahl von Nachbarn in der Umgebung $\mathcal{U}_n(\epsilon)$ zu sammeln. Dies kann höherdimensionale Einbettungen benachteiligen. Um einen solchen Effekt auszuschließen, nutzen wir ein festes ϵ und suchen eine feste Anzahl von Nachbarn für alle Werte der Zeitverzögerung τ und der Einbettungsdimension m . Dabei variieren wir die Länge der Zeitreihe, innerhalb derer Nachbarn gesucht werden (wir laufen solange rückwärts in der Zeit, bis die festgelegte Anzahl von Nachbarn gefunden wurde).

Das Ergebnis der Vorhersagen ist in Abb. 2.14 (d) - (f) dargestellt, wo der Vorhersagefehler (normiert durch die Standardabweichung der Daten) in Abhängigkeit von der Zeitverzögerung für verschiedene Einbettungsdimensionen gezeigt ist. Damit erhalten wir eine Aussage nicht nur über die Ordnung des optimalen Markov-Prozesses, sondern zugleich über den Zeitversatz τ , der, wie vorher diskutiert, in der Praxis eine wesentliche Rolle spielt. Wir beobachten verschiedene Formen des Gedächtnisses abhängig davon, wie das Rauschen in das System gekoppelt ist. Im ersten Fall sieht man, daß das Minimum des Vorhersagefehlers nicht von der Einbettungsdimension abhängt, falls $m \geq 2$. Außerdem erkennt man, daß die Länge des optimalen Einbettungsfensters, d.h., das Zeitintervall $(m - 1)\tau$, welches durch den Einbettungsvektor bei optimalem τ aufgespannt wird, unabhängig von m ist für $m \geq 2$. Dies zeigt ebenfalls, daß ein Modell zweiter Ordnung ausreichend ist. Im Gegensatz dazu findet man im Fall (b), daß der Vorhersagefehler mit steigender Einbettungsdimension m abfällt - jede höhere Dimension fügt Information zur Vorhersage der Wahrscheinlichkeitsverteilung hinzu. Die Verbesserung beträgt jedoch nur wenige Prozent des absoluten Fehlers. Wenn wir die x -Variable durch das Rauschen antreiben und y messen (c), so ist auch in diesem Fall die Markov-Eigenschaft verletzt. Die Verbesserung der Vorhersage durch höherdimensionale Einbettungen ist sogar deutlich stärker als im Fall (b). Hier wird deutlich, daß man tatsächlich in der Lage ist, höherdimensionale rauschinduzierte Struktur im Phasenraum aufzulösen. Der optimale Zeitversatz variiert auf nichttriviale Weise für verschiedene Einbettungsdimensionen.

Wir möchten die Ergebnisse für die betrachteten drei Fälle noch anhand einer etablierten Methode für stochastische dynamische Systeme untersuchen, dem Konzept grobgekörneter (coarse grained) dynamischer Entropie. Zunächst geben wir eine kur-

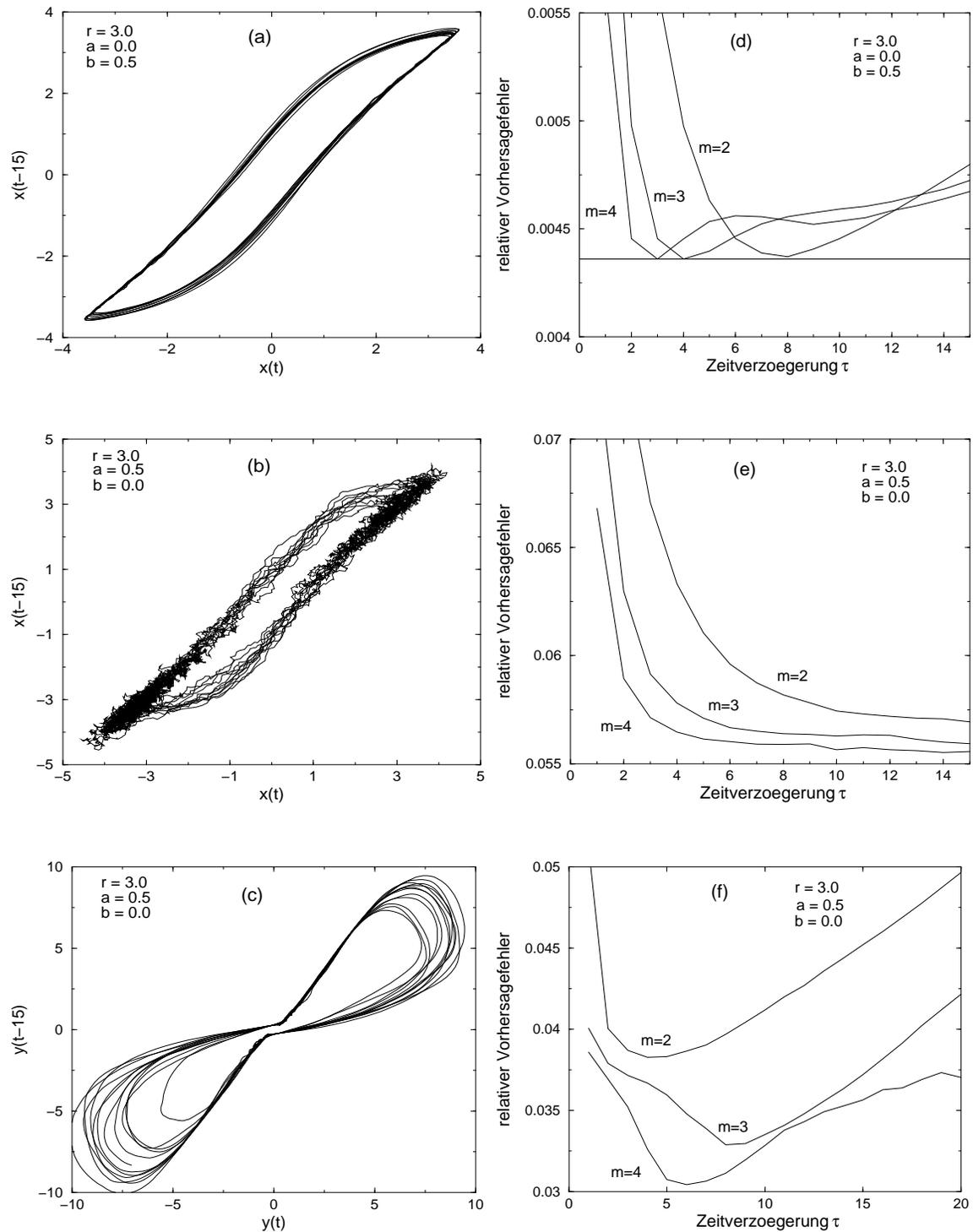


Abbildung 2.14: Bilder a) -c): Zweidimensionale Projektion des Phasenportraits des Van-der-Pol-Systems für verschiedene Konfigurationen des Rauschens und bei Nutzung verschiedener Variablen für die Einbettung. Bilder d) -f): Relative Vorhersagefehler für dieselben Rausch- und Einbettungskonfigurationen wie in den jeweils nebenstehenden Abbildungen a) -c).

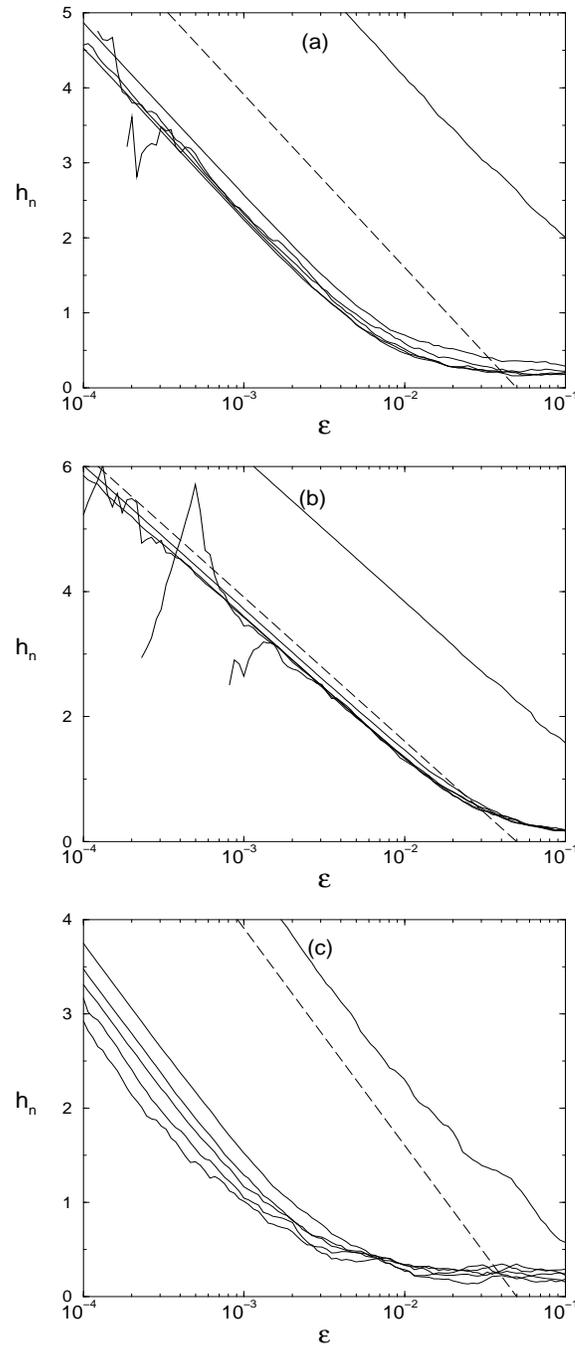


Abbildung 2.15: Korrelationsentropien h_0, \dots, h_5 für die Attraktoren des Van-der-Pol-Systems mit Zeitverzögerung $\tau = 35$. Konfigurationen des Rauschens und der Einbettungskordinaten entsprechen den Werten in Abb. 2.14 a) - c). Die gestrichelte Linie zeigt die Funktion $f(\epsilon) = -\ln(\epsilon) - 3$.

ze Darstellung dieser Methode. Die Entropie stellt in der Thermodynamik ein Maß für die Unordnung eines untersuchten Systems dar. Dieselbe Bedeutung kommt dieser Größe in der Informationstheorie zu. Man betrachte eine binäre Reihe, die durch eine beliebige Quelle generiert sei. Die Wahrscheinlichkeit p_{i_1, i_2, \dots, i_n} bezeichne die Verbundwahrscheinlichkeit zu einer beliebigen Zeit eine Sequenz der Länge n von der Form (i_1, i_2, \dots, i_n) zu finden, mit $i_k \in \{0, 1\}$, $k = 1, \dots, n$. Die Shannon-Entropien, welche man auf diesen Verbundwahrscheinlichkeiten berechnet

$$H_n = - \sum p_{i_1, i_2, \dots, i_n} \ln p_{i_1, i_2, \dots, i_n}, \quad (2.47)$$

werden auch Block-Entropien der Blocklänge m genannt. Deren Differenzen

$$h_n := H_{n+1} - H_n, \quad h_0 = H_1 \quad (2.48)$$

werden bedingte Entropien genannt, da diese interpretiert werden können als Shannon-Entropie basierend auf der bedingten Wahrscheinlichkeit $p(i_{n+1} | i_1, i_2, \dots, i_n)$. Diese Größe beschreibt den Betrag an Information, welcher benötigt wird, um den nächsten Wert vorherzusagen, wenn die letzten n Werte bekannt sind. Grob gesagt, sollte die bedingte Entropie konstant bleiben, sobald die Länge des maximalen Gedächtnisses des Prozesses erreicht ist. Damit eignen sich bedingte Entropien prinzipiell gut, um das Gedächtnis der oben vorgestellten Langevin-Gleichungen zu untersuchen.

Hierzu muß dieses Konzept jedoch auf Reihen verallgemeinert werden, die aus einer beliebigen Anzahl $|A|$ verschiedener Symbole zusammengesetzt sind. Hier bezeichnet A die Menge der Symbole und $|A|$ deren Anzahl. Um die Methoden auf eine Reihe reellwertiger Vektoren $\vec{x} \in \Gamma \subset \mathcal{R}^d$ anwenden zu können, muß diese in eine Symbolsequenz übersetzt werden. Dies geschieht formal durch die Einführung einer Partition \mathcal{P} auf Γ , welche eine Menge von Zellen c_i mit $c_i \subset \Gamma$, $\cup_i c_i = \Gamma$ und $c_i \cap c_j = \emptyset \quad \forall i \neq j$ ist. Dann kann jeder Vektor eindeutig durch die Bezeichnung der Zelle, welche diesen enthält, ersetzt werden, und Verbundwahrscheinlichkeiten können völlig analog zu der Situation einer binären Sequenz berechnet werden. Die bedingten Entropien h_n werden natürlich von der gewählten Partition abhängen, insbesondere von der räumlichen Auflösung in Γ . Der Grad der räumlichen Auflösung kann durch den maximalen Durchmesser der Zellen $\epsilon = \max_i \|c_i\|$ charakterisiert werden, wobei $\|\cdot\|$ eine passende Norm für die Menge c_i darstellt. Für numerische Berechnungen wird die Partitions-basierte Shannon-Entropie meist durch die durch Überdeckungen definierten Renyi-Entropien 2. Ordnung ersetzt:

$$\begin{aligned} \hat{H}_n(\epsilon) &= - \ln C(n, \frac{\epsilon}{2}) \\ &= - \ln \left(\frac{2}{(N-n)(N+1-n)} \sum_{i < j} \Theta(\frac{\epsilon}{2} - |\vec{s}_i - \vec{s}_j|) \right), \end{aligned} \quad (2.49)$$

wobei $\Theta(\cdot)$ die Heaviside-Funktion und $C(\cdot)$ die Korrelationssumme ist, \vec{s}_i sind n -dimensionale Einbettungsvektoren, und N ist die Länge des Datensatzes. Die bedingte

Entropie $\hat{h}_n(\epsilon) = \hat{H}_{n+1}(\epsilon) - \hat{H}_n(\epsilon)$ wird auch Korrelationsentropie genannt. Es gibt eine Vielzahl technischer Details zum Zusammenhang zwischen Überdeckungs- und Partitions-Definition der Entropie, siehe z.B. in [20],[3], auf die wir aber hier nicht eingehen wollen. Statt dessen wollen wir die Eigenschaften der bedingten Entropie für Markov-Prozesse zitieren. Ein bekanntes Ergebnis ist, daß sich die bedingten Entropien für einen stochastischen Prozeß verhalten wie $\hat{h}_n(\epsilon) = c_n - \ln \epsilon$, wobei die Konstanten c_n monoton fallend in n sind. Falls der Prozeß die Markov-Eigenschaft der Ordnung m hat, so gilt $c_n = c_\infty$ für alle $n \geq m$. Daher findet man in einer logarithmischen Darstellung der bedingten Entropien $\hat{h}_1 \dots \hat{h}_n$ für einen Markov-Prozeß m 'ter Ordnung m parallele, aber voneinander verschiedene Linien. Alle Kurven für $n > m$ fallen auf den Graph von h_m , weil es kein Gedächtnis der Ordnung $m+1$ oder höher in den Daten gibt, welches die Entropie weiter senken könnte.

Die bedingten Entropien $\hat{h}_0 \dots \hat{h}_5$ für den Prozeß in Gleichung (2.46) sind in Abb. 2.15 (a), (b), (c) für die jeweils zugehörigen Situationen (a), (b), (c) in Abb. 2.14 gezeigt. Im Fall, daß die Markov-Eigenschaft zweiter Ordnung vorliegt (das Rauschen ist in y , und x wird gemessen), sieht man, daß $\hat{h}_2 \dots \hat{h}_5$ für den Bereich von ϵ -Werten, die dem stochastischen Regime entsprechen, auf einer Linie liegen. Im Fall, daß das Rauschen zu x addiert ist und x gemessen wird, sieht man im Wesentlichen die Zufallsbewegung um den Grenzyklus. Gedächtnisse höherer Ordnung sind hier mittels bedingter Entropien schwierig festzustellen. Nur im letzten Fall sieht man, daß das Rauschen nichttriviale höherdimensionale Struktur erzeugt, die sich durch Phasenraummethoden auflösen läßt.

Zusammenfassend kann man sagen, daß nur in Ausnahmefällen eine beobachtete skalare Zeitreihe, die durch einen vektorwertigen Langevin-Prozeß erzeugt wurde, die Markov-Eigenschaft aufweist. Es zeigt sich jedoch, daß es in vielen Fällen möglich ist, den Prozeß als annähernd Markov zu betrachten (wie hier im Fall (b)), weil das Gedächtnis schnell abfällt. Wenn der Fehler, welcher durch diese Näherung verursacht wird, kleiner als jener ist, der z.B. von der endlichen Länge der Daten stammt, so gibt es keinen praktischen Unterschied zwischen der Markov-Näherung und einem formal unendlich-dimensionalen Modell.

2.2.4 Einbettungsparameter

Wir haben in Abschnitt 2.1.3 beschrieben, wie die Einbettungsparameter m und τ in deterministischen dynamischen Systemen zu wählen sind. Wir wollen nun untersuchen, wie sich die optimalen Werte der Parameter ändern, wenn das System dynamisch verwechselt ist.

Wie im vorangegangenen Abschnitt ausgeführt, gibt es hinsichtlich der Dimension zwei verschiedene Fälle. Eine Möglichkeit ist, daß die Zeitreihe einen Markov-Prozeß beschreibt, der die Ordnung des ursprünglichen multidimensionalen Prozesses hat. Dann

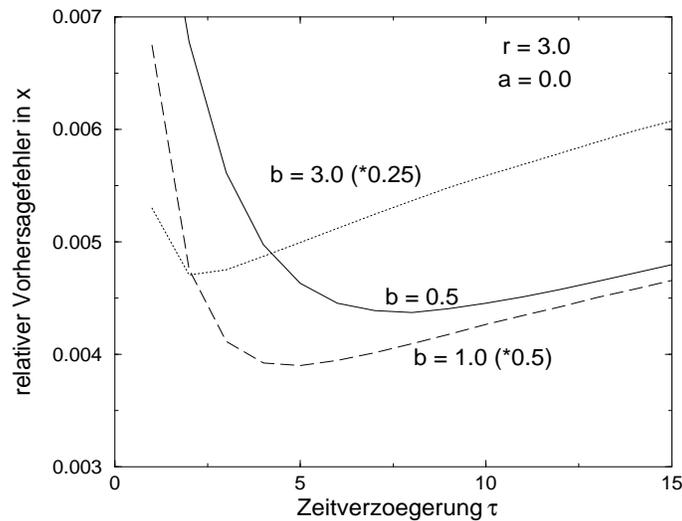


Abbildung 2.16: Relative Vorhersagefehler versus Zeitverzögerung für verschiedene Werte des Rauschniveaus für das Van-der-Pol-System $b = 0.5, 1.0, 3.0$ (x -Variable beobachtet und $m = 2$). Die beiden Kurven $b = 1.0$ und 3.0 sind mit einem Faktor $1/2$ bzw. $1/4$ reskaliert.

entspricht diese Ordnung natürlich auch der optimalen Einbettungsdimension m . Andernfalls kann das optimale m deutlich größer und formal sogar unendlich sein. Aufgrund der endlichen Länge der Zeitreihen wird man in praktischen Situationen jedoch stets einen endlichen Wert für m beobachten.

Interessanter ist die Frage nach dem optimalen Wert für die Zeitverzögerung τ_{opt} . In deterministischen dynamischen Systemen ist der erste Nulldurchgang der Autokorrelationsfunktion eine gute Näherung für τ_{opt} . Wir werden zeigen, daß τ_{opt} im stochastischen Fall deutlich kleiner ist und mit steigendem Rauschniveau abnimmt.

Einen Hinweis für dieses Verhalten bekommt man bereits, wenn man sich verdeutlicht, wie die unbeobachteten Variablen durch die Zeitversatzeinbettung rekonstruiert werden. Formal geschieht dies durch die Auswertung von Ableitungen der beobachteten Variable. Da diese Ableitungen aus Differenzen zweier Werte der Zeitreihe im Grenzübergang verschwindender Zeitdifferenz geschätzt werden, sollte der Zeitversatz im Prinzip so klein wie möglich sein, also der Abtastrate der Zeitreihe entsprechen. Wir werden für Situationen, in denen sich die Markov-Dynamik rekonstruieren läßt (Fall a) im obigen Van-der-Pol-System) sehen, daß dies für hohe Rauschamplituden tatsächlich der Fall ist. Für sehr kleines Rauschniveau beobachten wir den Wert, der im deterministischen Fall erwartet wird, z.B. der erste Nulldurchgang der Autokorrelation. Für mittlere Rauschamplituden beobachten wir einen Übergang zwischen beiden Extremfällen. Dies ist in Abb. 2.16 für das Van-der-Pol-System dargestellt, wo der relative Vorhersagefehler in Abhängigkeit von der Zeitverzögerung für verschiedene Rauschamplituden gezeigt ist. Das Minimum des Vorhersagefehlers verschiebt sich für

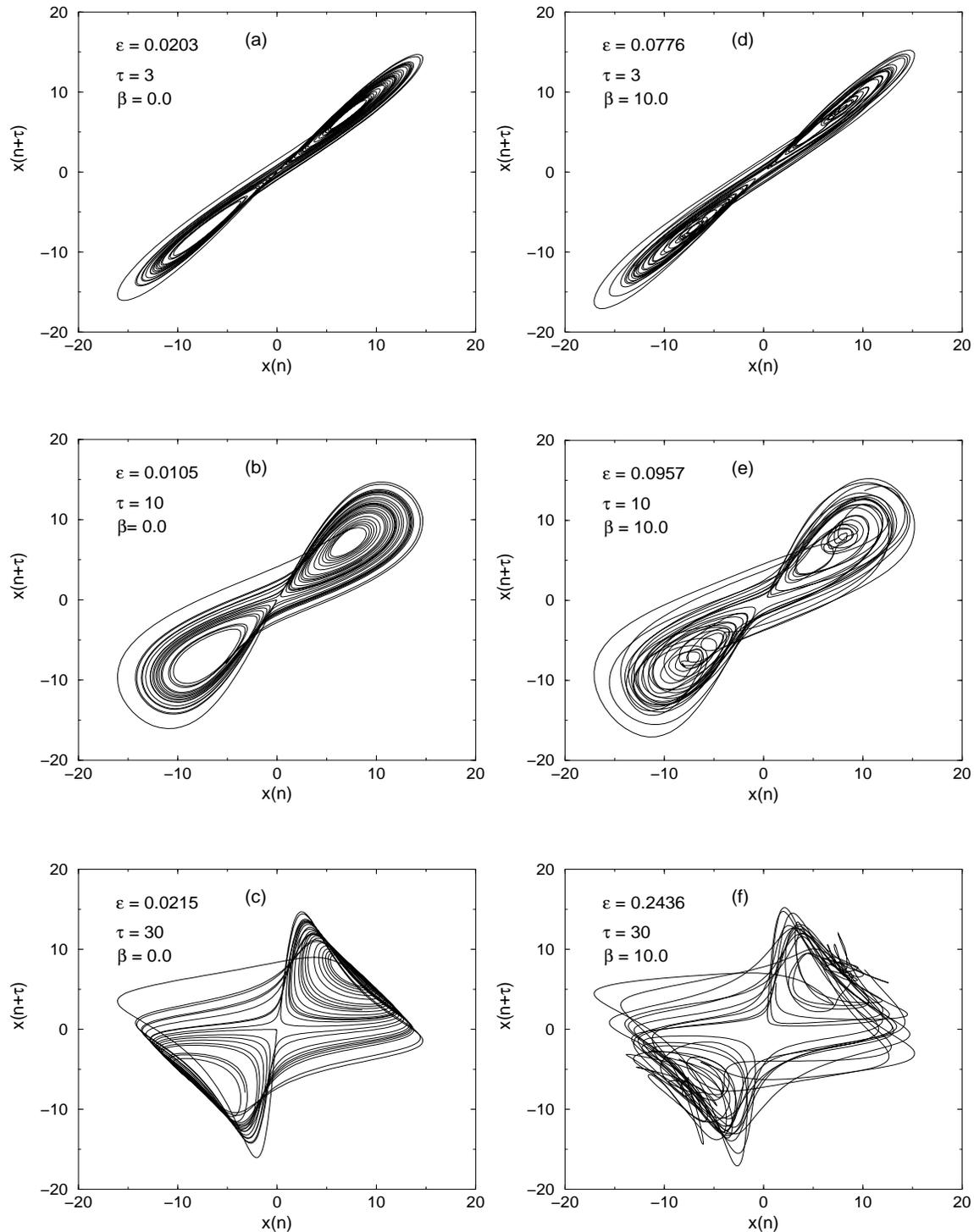


Abbildung 2.17: Zweidimensionale Projektion des Phasenportraits des Lorenz-Systems für verschiedene Werte der Zeitverzögerung $\tau = 3, 10, 30 \delta t$ und Rauschniveau $\beta = 0.0$ ((a)-(c)) und Rauschniveau $\beta = 10.0$ ((d)-(f)). Der relative Vorhersagefehler ϵ ist ebenfalls angegeben.

wachsendes Rauschniveau zu kleinerer Zeitverzögerung.

Als zweites Beispiel betrachten wir das durch Rauschen getriebene Lorenz-System. Die Bewegungsgleichungen sind:

$$\begin{aligned}\dot{x} &= 6(y - x) \\ \dot{y} &= 28x - y - xz \\ \dot{z} &= xy - (13/6)z + \beta\Gamma,\end{aligned}\tag{2.50}$$

Γ ist wieder weißes Gaußsches Rauschen mit Varianz eins und δ -Korrelation. Eine Zeitreihe der Länge $N = 1\,000\,000$ wurde durch die Integration dieser Gleichungen bei einer Abtastrate von $\delta t = 0.1$ Zeiteinheiten erzeugt. In Abbildung 2.17(a)-(c) zeigen wir für verschwindendes Rauschniveau und drei verschiedene Werte der Zeitverzögerung eine zwei-dimensionale Projektion des Phasenraumes für dieses dynamische System. Für kleines $\tau = 3\delta t$ ist das Phasenportrait um die Diagonale zentriert, da aufeinanderfolgende Werte der Zeitreihe sehr ähnlich sind. Für mittlere Werte von $\tau = 10\delta t$ ist der Attraktor gut entfaltet, wogegen große Werte von $\tau = 30\delta t$ zu komplizierten Graphen führen. In den Abbildungen ist neben der verwendeten Zeitverzögerung der relative Vorhersagefehler ϵ in der jeweiligen Situation angegeben. Für verschwindendes Rauschniveau entspricht das scheinbar einfachste Phasenportrait in Abb. 2.17(b) dem minimalen Vorhersagefehler ϵ , wobei der absolute Fehler wieder durch die Varianz der Daten normiert wurde. Der optimale Wert von τ fällt jedoch mit steigendem Rauschniveau. In den Abbildungen 2.17(d)-(f) zeigen wir die Phasenportraits des Lorenzsystems bei einem Rauschniveau von $\beta = 10.0$. Weiterhin ist der Attraktor bei $\tau = 10\delta t$ am besten entfaltet. In diesem Fall ist jedoch der Fehler bei einer Zeitverzögerung von $\tau = 3\delta t$ minimal. Daher sind bei einem stochastischen dynamischen System die visuelle Analyse des Attraktors sowie die Bestimmung des ersten Nulldurchgangs der Autokorrelation keine guten Methoden zur Ermittlung einer geeigneten Zeitverzögerung. Wir empfehlen daher eine explizite Optimierung der Vorhersagefehler abhängig von der gewählten Zeitverzögerung.

2.2.5 Nichtlineare Fluktuationen in stochastischen Systemen

Die Frage, ob ein stochastischer Prozeß linear oder nichtlinear ist, hat starke Implikationen hinsichtlich der Amplitude der auftretenden Fluktuationen in der Zeit. Nichtlineare Fluktuationen in Markov-Prozessen können mit dem vorgestellten Verfahren identifiziert werden. Wenn man ein lineares Modell nutzt, um einen intrinsisch nichtlinearen Prozeß vorherzusagen, sind die vorhergesagten Fluktuationen im Mittel zu klein. Daher analysieren wir Wahrscheinlichkeitsverteilungen (PDFs) der Inkremente des Signals $\Delta\hat{s}_T = \hat{s}_{n+T}^{model} - s_n$, vorhergesagt entweder durch das AR-Modell bzw. durch das lokal konstante Modell, wobei \hat{s}_{n+T}^{model} der durch eines der beiden Modelle vorhergesagte Wert des Signals T Zeitschritte in der Zukunft ist. Es ist eine typische Situation in stochastischen Systemen, daß die PDFs der beiden Modelle sich deutlich unterscheiden, obwohl

deren *mittlere* Vorhersagefehler sehr ähnlich sind. Dies beruht auf der Tatsache, daß die starken Fluktuationen eines nichtlinearen stochastischen Prozesses ihren Ursprung in dem stochastischen Antrieb sowie in nichtlinearen Korrelationen haben. Letztere können durch nichtlineare Verfahren, jedoch nicht durch lineare Algorithmen modelliert werden.

Für den dynamisch verrauschten Duffingoszillator in Gleichung (2.43) mit $a = -0.5$ und $b = 0.5$ ist die Bewegung im wesentlichen durch den stochastischen Antrieb generiert, da das Teilchen ohne Rauschen zur Ruhe kommen würde. Der deterministische Anteil der Bewegungsgleichungen ist nichtlinear. Eine Zeitreihe des Prozesses ist in Abb. 2.18 gezeigt, wobei die Abtastrate $\delta t = 0.1$ beträgt. In Abb. 2.19 sind die PDFs der linear und der nichtlinear vorhergesagten Inkremente dargestellt. Das lokal konstante Modell wurde in einem 2-dimensionalen Einbettungsraum bei einer Zeitverzögerung von $\tau = 10\delta t$ ausgewertet, und die linearen Vorhersagen ergeben sich aus einem entsprechenden $AR(2)$ -Modell mit gleichem τ . Der Vorhersagehorizont ist $T = 20\delta t$.

Ebenfalls gezeigt sind die PDFs der Inkremente der eigentlichen Zeitreihe sowie jene, die man erhält, wenn man den exakten deterministischen Anteil der Bewegungsgleichungen iteriert. Während das AR -Modell nicht in der Lage ist, die starken Inkremente des Signals zu modellieren, gibt das lokal konstante Modell eine gute Approximation der Inkremente, die durch den deterministischen Teil der Bewegungsgleichungen erzeugt sind. Abweichungen von der exakten deterministischen PDF beruhen auf der begrenzten Länge der verwendeten Zeitreihe.

Wenn das lokal konstante Modell in der Lage ist, den deterministischen Teil der Bewegungsgleichungen vorherzusagen, sollten alle Vorhersagen im Mittel über den stochastischen Anteil korrekt sein. Mittelt man über eine hinreichende Anzahl von Situationen, in denen ein bestimmter Wert vorhergesagt wird, so sollten sich die Zufallsfehler zu Null addieren. In demselben Sinne sollten die vorhergesagten Inkremente $\Delta \hat{s}_T = \hat{s}_{n+T}^{model} - s_n$ bei einem guten Modell im Mittel richtig sein. Diese in der Wettervorhersage gut bekannte Statistik wird dort als Zuverlässigkeitsfunktion (*reliability function*) bezeichnet und ist in Abb. 2.20 gezeigt. Wir beobachten dabei einen deutlichen Unterschied zwischen den beiden Vorhersagealgorithmen. Während das lokal konstante Modell im Mittel die richtigen Vorhersagen ergibt, unterschätzt das AR -Modell die Höhe der Fluktuationen systematisch. Das lineare Verfahren ist nicht in der Lage, den deterministischen Teil der Dynamik vorherzusagen. Es ist daher gezeigt, daß das lokal konstante Verfahren die nichtlinear deterministische Dynamik eines Markov-Prozesses 2. Ordnung aus einer einzelnen Zeitreihe extrahiert. Die vorgestellte Statistik kann genutzt werden, um nichtlineare Fluktuationen in einer Zeitreihe zu identifizieren. Wir werden dieses Verfahren im Verlauf der Arbeit nutzen, um voll entwickelte Turbulenz in einem Laborexperiment sowie aus Messungen in der atmosphärischen Grenzschicht zu analysieren.

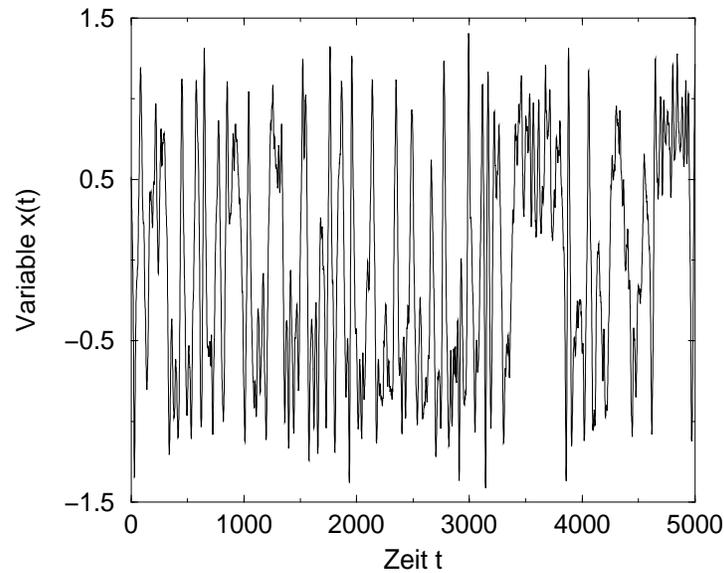
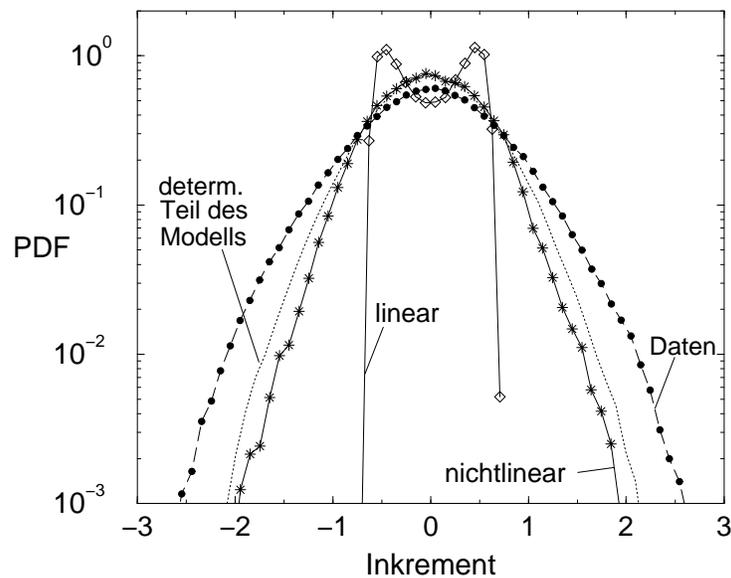
Abbildung 2.18: Variable x des Duffing-Systems versus Zeit t .

Abbildung 2.19: PDF der Inkremente der x -Variable des Duffing Systems. Ebenfalls gezeigt sind die durch das nichtlineare und das lineare Modell (AR-Modell) vorhergesagten Inkremente sowie die Inkremente, die man erhält, wenn der exakte deterministische Teil der Bewegungsgleichungen als bekannt angenommen wird.

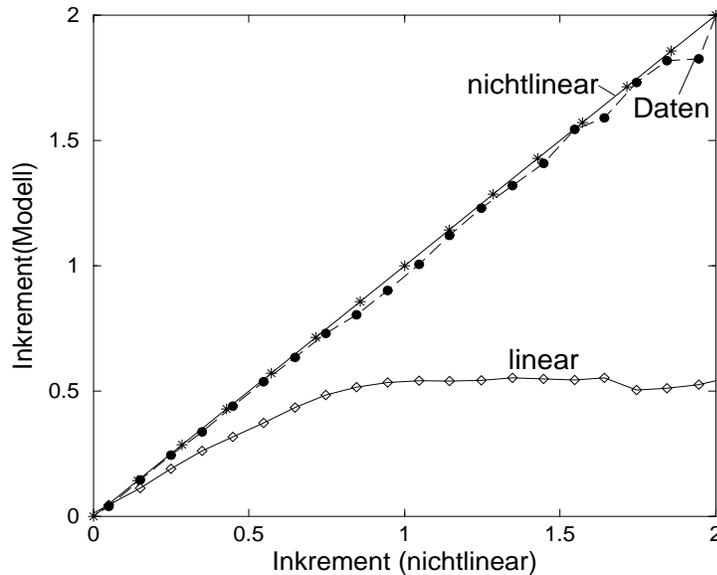


Abbildung 2.20: Mittelwert der gemessenen Inkremente Δs_T und durch das AR-Modell vorhergesagte Inkremente $\Delta \hat{s}_T^{AR}$ versus der nichtlinear vorhergesagten Inkremente $\Delta \hat{s}_T^{zero}$.

2.2.6 Nichtlineare Meßfunktion

Eine wichtige und typische Situation ist dadurch gegeben, daß die Observable eines linearen Systems durch eine nichtlineare Meßfunktion gemessen wird. Dies induziert nichtlineare Korrelationen in den Daten. Insbesondere wenn die Meßfunktion nicht bekannt oder nicht invertierbar ist, ist die Anwendung eines nichtlinearen Algorithmus notwendig. Um die Leistungsfähigkeit eines linearen und eines nichtlinearen Verfahrens an Daten, welche nichtlinear transformiert wurden, zu untersuchen, nutzen wir die Methode von Casdagli [15]. Dabei kann man zwischen linearem und nichtlinearem Verhalten des Modells stufenlos variieren, indem man den Vorhersagefehler als Funktion des Nachbarschaftsradius $\text{diam}(\mathcal{U})$ bestimmt. Für kleine Radien bekommt man ein lokales Modell, für Radien von der Ausdehnung des Attraktors hat man ein globales *AR*-Modell. Wir wollen nun diese Statistik am Beispiel eines *AR*(2)-Prozesses untersuchen, welches durch $x_{n+1} = a_1 x_n + a_2 x_{n-1} + \Gamma_n$ sowie $a_1 = 1.985$ und $a_2 = -0.995$ gegeben ist. Die Meßfunktion ist $s_n = \text{sign}(x_n) \sqrt{|x_n|}$ (gedämpfter harmonischer Oszillator). In Abb. 2.21 zeigen wir die mittleren Vorhersagefehler des lokal linearen Modells als Funktion der Anzahl der Nachbarn, die für den Fit genutzt wurden sowohl für die Zeitreihe von $\{s_n\}$ wie von $\{x_n\}$.

Bei der direkten Beobachtung des linearen Prozesses $\{x_n\}$ findet man, daß das *AR*-Modell geringfügig bessere Vorhersagen ergibt als das lokale Modell. Dies liegt an der höheren statistischen Robustheit globaler Modelle. Falls die Variable des linearen Pro-

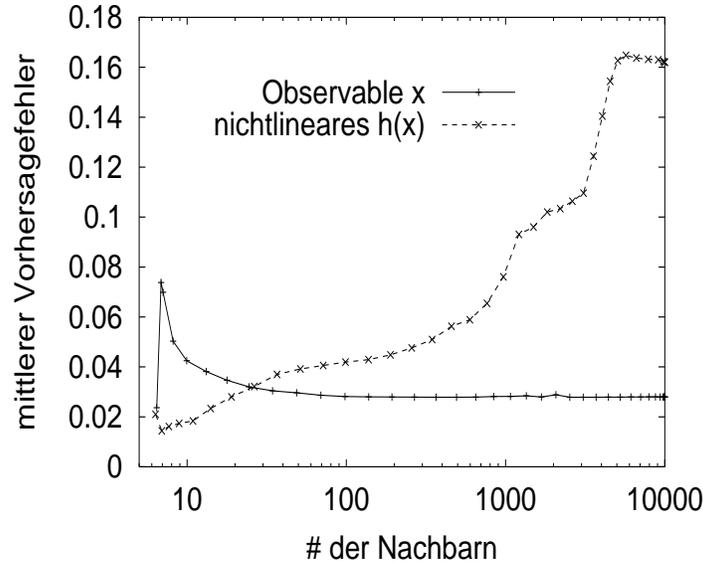


Abbildung 2.21: Gemittelter Vorhersagefehler lokal linearer Modelle in Abhängigkeit von der Anzahl der Nachbarn, die zur Vorhersage genutzt wurden, $\{x_n\}$ (durchgezogen) und $\{s_n\}$ (gestrichelt) für den AR(2)-Prozeß (siehe Text).

zesses nichtlinear transformiert wird, kann das nichtlineare Modell deutlich besser sein. Diese Eigenschaft werden wir im letzten Teil der Arbeit ausnutzen, wenn wir die Leistungsabgabe von Windturbinen vorhersagen werden.

2.3 Diskussion

In diesem Kapitel wurden Verfahren für die auf Zeitreihen basierende Modellbildung in stochastischen dynamischen Systemen vorgestellt. Dies sind Verfahren, mit denen im Zustandsraum des stochastischen Systems die deterministischen sowie die stochastischen Anteile der Bewegungsgleichungen ermittelt werden können. Werden alle relevanten Freiheitsgrade des Systems beobachtet, lassen sich die Bewegungsgesetze in Form von Langevin- und Fokker-Planck-Gleichungen aus den Daten rekonstruieren. Für solche aus der Literatur bekannten Methoden wurden in diesem Kapitel wesentliche Korrekturen hergeleitet, die aufgrund der endlichen Abtastrate gemessener Daten essentiell sind. Ist die Anzahl der gemessenen Zeitreihen kleiner als die Anzahl der aktiven Freiheitsgrade, so kann man in deterministischen dynamischen Systemen den ursprünglichen Zustandsraum durch die bekannten Einbettungsmethoden rekonstruieren. Wir konnten in diesem Kapitel zeigen, daß auch in stochastischen Systemen mit Langevin-Dynamik die Rekonstruktion in einigen Fällen durch die Einführung von Ableitungs koordinaten möglich ist. Obwohl es Fälle gibt, in denen dynamisches Rau-

schon nichttriviale Struktur in den rekonstruierten Zustandsräumen erzeugt, ist die Einbettung in einigen Situationen exakt und in einer Vielzahl von Fällen in sehr guter Näherung erlaubt. Diese Beobachtung erweitert den Gültigkeitsbereich der aus der Theorie deterministischer dynamischer Systeme stammenden Konzepte. Die optimalen Parameter, wie die Zeitverögerung zwischen aufeinanderfolgenden Datenpunkten, müssen jedoch bei der Anwendung auf stochastische Systeme modifiziert werden.

Kapitel 3

Datenanalyse voll entwickelter Turbulenz

Die Entdeckung von deterministischem Chaos in dynamischen Systemen hat die Hoffnung auf ein besseres Verständnis hydrodynamischer Turbulenz geweckt. Es zeigte sich jedoch schnell, daß voll entwickelte Turbulenz zu komplex ist, um auf geeignete Weise durch niedrigdimensionale Modelle mit wenigen Freiheitsgraden beschrieben zu werden. Das vorherrschende Bild ist dagegen, daß Turbulenz durch die nichtlineare Wechselwirkung einer sehr großen Anzahl von Freiheitsgraden entsteht, was zu Nichtvorhersagbarkeit auf einem weiten Bereich verschiedener Zeitskalen führt. Im Prinzip erfordert die physikalische Beschreibung einer Flüssigkeit die Lösung der Navier-Stokes-Gleichungen bei Kenntnis der Anfangsbedingungen des Geschwindigkeitsfeldes $\vec{u}(\vec{r}, t)$ und des Druckes $p(\vec{r}, t)$ zu einer Zeit t_0 und zusätzlicher geeigneter Randbedingungen. Eine Möglichkeit der Lösung der Navier-Stokes-Gleichungen besteht in der Entwicklung der Felder von Druck und Geschwindigkeit in einen vollständigen Satz von Funktionen. Die Amplituden dieser Funktionen gehorchen dann einer unendlichen Zahl gewöhnlicher Differentialgleichungen. Nahe dem Gleichgewicht mag es vorkommen, daß nur einige Freiheitsgrade angeregt sind und alle anderen irrelevanten Moden durch die wenigen relevanten Freiheitsgrade verklärt sind. In diesem Fall können einige nichtlineare gewöhnliche Differentialgleichungen ausreichen, um die Dynamik der Flüssigkeit zu beschreiben (Beispiel: Konvektionsrollen durch das Lorenz-System). Bei weiterer Entfernung vom Gleichgewicht werden jedoch mehr und mehr Moden instabil, und die einfache Beschreibung durch einige wenige Freiheitsgrade ist ungültig. In der Tat kann die Anzahl N der aktiven Freiheitsgrade in einer turbulenten Flüssigkeit abgeschätzt werden durch $N = (L/\eta)^3$, wobei L die Größe des Systems und η die Dissipationslänge ist. N kann auch zur Reynolds-Zahl in Beziehung gesetzt werden, mit $L/\eta = Re^{3/4}$ [21, 22]. Daher erwartet man eine sehr hohe Anzahl aktiver Freiheitsgrade bei hohen Reynolds-Zahlen. Ähnliche Argumente sollten auch für 2-dimensionale Turbulenz gültig sein, wobei jetzt gilt $N = (L/\eta)^2$.

Wir wollen in dem nun folgenden Kapitel verschiedene Methoden der Zeitreihenanalyse nutzen, um 2-dimensionale und 3-dimensionale Turbulenz zu untersuchen. Hierbei werden wir zunächst etablierte Methoden der Turbulenzforschung, wie die Analyse von Wahrscheinlichkeitsverteilungen (PDFs) sowie von Spektren und Strukturfunktionen, nutzen. Danach werden wir die im vorhergehenden Kapitel eingeführten Verfahren, Fokker-Planck-Gleichungen und Markov-Prozesse von Daten zu fitten, anwenden. Dabei wird sich zeigen, daß niedrigdimensionale Modelle geeignet sein können, kohärente Strukturen in turbulenten Flüssigkeiten zu untersuchen, obwohl es keine Indikationen für einen niedrigdimensionalen Attraktor in den Daten gibt. Zunächst wollen wir jedoch die Grundbegriffe der Turbulenz einführen und das Experiment erläutern.

3.1 Phänomenologie voll entwickelter Turbulenz

Wir wollen hier einige wesentliche Grundlagen voll entwickelter Turbulenz zusammentragen und die zum Verständnis der folgenden Abschnitte notwendige Terminologie einführen. Die grundlegenden Gesetze der modernen Hydrodynamik wurden zu Anfang des 19. Jahrhunderts in Form der Navier-Stokes-Gleichungen (NSG) formuliert. Die NSG beschreiben die zeitliche Evolution des Geschwindigkeitsfeldes $\vec{u}(\vec{x}, t)$ in Gasen und Flüssigkeiten. Sie hat die Form:

$$\partial_t \vec{u} + (\vec{u} \cdot \nabla) \vec{u} - \nu \nabla^2 \vec{u} = \frac{1}{\rho} (\vec{f} - \nabla p), \quad (3.1)$$

hierbei ist ν die Viskosität der Flüssigkeit ($\approx 1.5 \times 10^{-5} m^2/s$ in Luft, $\approx 10^{-6} m^2/s$ in Wasser), ρ ist die Dichte des Mediums, \vec{f} ist das externe Kraftfeld und p ist das Druckfeld. Der Druck ist dabei im Innern nicht unabhängig, sondern eine Konsequenz von \vec{u} . Die Euler-Gleichung ohne den Term $\nu \nabla^2 \vec{u}$ verkörpert das Newtonsche Gesetz $\vec{F} = m\vec{a}$ für ein Volumenelement der Flüssigkeit. Der Term $\nu \nabla^2 \vec{u}$ beschreibt die Reibungskräfte. Die NSG müssen noch durch die Kontinuitätsgleichung

$$\partial_t \rho + \nabla(\rho \vec{u}) \quad (3.2)$$

ergänzt werden. In den meisten Anwendungen wird die Flüssigkeit als inkompressibel betrachtet, d.h., ρ ist konstant, und \vec{u} ist divergenzfrei:

$$\nabla \cdot \vec{u} = 0. \quad (3.3)$$

Die NSG sind invariant unter den folgenden Umskalierungen:

$$\tilde{u}(\vec{x}, t) = \tau s^{-1} \vec{u}(s\vec{x}, \tau t)$$

$$\begin{aligned}\tilde{f}(\vec{x}, t) &= \tau s^{-1} \vec{f}(s\vec{x}, \tau t) \\ \tilde{p}(\vec{x}, t) &= \tau s^{-1} p(s\vec{x}, \tau t) \\ \tilde{\nu} &= \tau s^{-1} \nu.\end{aligned}$$

Sind \vec{u} und p Lösungen der NSG mit Viskosität ν und externer Kraft \vec{f} , so lösen \tilde{u} und \tilde{p} diese Gleichungen mit der Viskosität $\tilde{\nu}$ und Kraft \tilde{f} . Dies führt auf die Angabe einer dimensionslosen Größe (der inversen Viskosität), der **Reynolds-Zahl** - welche der wichtigste experimentelle Kontrollparameter der Turbulenz ist:

$$Re = \frac{L\delta_L u}{\nu}. \quad (3.4)$$

Hierbei ist $\delta_L u$ die charakteristische Größe der Geschwindigkeitsdifferenzen über eine Skala L , welche die Größe des Systems angibt. Für den Fluß durch ein Rohr mit Radius L kann man die Geschwindigkeit in der Rohrmitte für $\delta_L u$ annehmen, da die Geschwindigkeit an den Rohrwänden verschwindet. Die folgenden wesentlichen phänomenologischen Ergebnisse geben ein numerisches Verständnis von der Reynolds-Zahl. Für $Re \ll 1$ ist der Fluß **laminar**. Für $1 < Re < 100$ kann man komplexe Strukturen, deren Gestalt von der jeweiligen Geometrie abhängt, beobachten. Bei $Re \gg 100$ setzt **turbulentes** Verhalten mit sehr unregelmäßigen Mustern ein. Bei noch größeren Reynolds-Zahlen können diese Strukturen selbständig werden, und man spricht von **voll entwickelter Turbulenz**.

Im Bereich $Re \ll 100$ können einige wenige Moden ausreichen, um die Dynamik der Flüssigkeit durch gewöhnliche Differentialgleichungen mit niedrigdimensionalen Attraktoren zu beschreiben. Hier wurden in der Vergangenheit erfolgreich Verfahren aus der Theorie dynamischer Systeme angewandt, um spezifische Flußsituationen zu beschreiben, z.B. den Fluß zwischen zwei rotierenden Zylindern (Taylor-Couette-Fluß) [12]. Wir wollen im folgenden untersuchen, ob Konzepte aus der Theorie dynamischer Systeme auch zur Untersuchung entwickelter Turbulenz geeignet sind. Dabei möchten wir Eigenschaften 2-dimensionaler und 3-dimensionaler Turbulenz vergleichen.

2D-Turbulenz wird in dem zu untersuchenden Laborexperiment durch Rotation des Flüssigkeitsbehälters erreicht. Der wesentliche Kontrollparameter, der dabei die 2-Dimensionalität des Flusses quantitativ beschreibt, ist die **Rossby-Zahl**:

$$Ro = \frac{(U/D)}{2\Omega}, \quad (3.5)$$

wobei U die typische Geschwindigkeit des Mediums relativ zum Behälter ist, D gibt die charakteristische Längenskala des Systems an, und Ω ist der Betrag der Winkelgeschwindigkeit, deren Richtung antiparallel zur Gravitation gerichtet sei. Der Quotient U/D ergibt eine typische Frequenz der Flüssigkeitsbewegungen, somit ist die Rossby-Zahl der Quotient aus der typischen Frequenz der Flüssigkeit und der Rotation des

Containers. Für $Ro \gg 1$ sind Rotationseffekte vernachlässigbar, wogegen diese dominieren, wenn $Ro \ll 1$ gilt. Das Proudman-Taylor-Theorem [23] garantiert, daß die Bewegung einer Flüssigkeit mit uniformer Dichte 2-dimensional wird, im Grenzwert verschwindender Rossby-Zahl. Die Struktur der Flüssigkeit wird unabhängig von der Richtung des Rotationsvektors. Man kann dies sehr einfach sehen, wenn man die NSG in dimensionsloser Form schreibt als [24]:

$$Ro[\partial_t \vec{u} + (\vec{u} \cdot \nabla) \vec{u}] - \frac{Ro}{Re} \nabla^2 \vec{u} = -\vec{z} \times \vec{u} - \nabla p, \quad (3.6)$$

wobei \vec{z} die Rotationsachse ist. Wenn man nun den Viskositätsterm bei hohen Reynolds-Zahlen und kleinen Rossby-Zahlen vernachlässigt und Geschwindigkeit und Druck in dem kleinen Parameter Ro entwickelt, so bleibt in niedrigster Ordnung nur die rechte Seite von Gleichung (3.6). Wendet man den Rotationsoperator auf diese Seite an, so ergibt sich in nullter Ordnung in Ro :

$$\frac{\partial}{\partial z} \vec{u} = 0. \quad (3.7)$$

Wir wollen nun die wichtigsten experimentellen und theoretischen Ergebnisse bei der statistischen Beschreibung 2-dimensionaler und 3-dimensionaler Turbulenz zusammenfassen. Diese werden üblicherweise unter dem Titel “Phänomenologie voll entwickelter Turbulenz” behandelt und basieren im wesentlichen auf den Arbeiten Kolmogorovs [25, 26], welche in Frisch [22] zusammengefaßt sind. Bei den mathematischen Ausdrücken wird \sim “gleich bis auf Konstanten der Ordnung eins” bedeuten (Faktoren $1/2$ werden vernachlässigt).

Das vorherrschende Bild bei der phänomenologischen Beschreibung hydrodynamischer voll entwickelter Turbulenz ist eine Kaskade von Wirbeln, beginnend bei der sogenannten Injektionsskala bis zur Dissipationsskala. Die Flüssigkeit nimmt Energie auf in einer Längenskala l_0 , die für die geometrische Ausdehnung des Behälters typisch ist, z.B. dem Rohrdurchmesser bei Rohrströmungen. Die auf dieser Injektionsskala erzeugten Wirbel zerfallen unter der Erzeugung nächstkleinerer Wirbel, welche die Längenskalen $l_n = l_0 r^n$ haben. Die kleinsten Wirbel haben die Skala η , die Kolmogorov-Dissipationsskala. Dieses Bild ist die sogenannte Richardson-Kaskade. Die Anzahl der Wirbel je Volumeneinheit wächst in diesem Modell mit n wie r^{-3n} , dieses stellt sicher, daß kleine Wirbel dasselbe Volumen wie große ausfüllen. Energie, die am oberen Ende der Kaskade mit einer Rate ϵ in die Flüssigkeit injiziert wird, wird durch die Wirbel der verschiedenen Größen auf kleinere Längenskalen mit derselben Rate ϵ transferiert. Ebenfalls mit der Rate ϵ wird die Energie dann bei der Skala η aus der Flüssigkeit durch Dissipation entfernt. Ein Vorteil dieses Kaskaden-Bildes ist, daß es zwei wesentliche Annahmen der Kolmogorov-Theorie liefert, von denen die erste die *Skaleninvarianz* ist. Diese Annahme wäre verletzt, wenn z.B. kleinere Wirbel weniger und weniger raumfüllend wären. Solch eine Annahme führt zu Korrekturen der Kolmogorov-Theorie (Intermittenz).

Die zweite Annahme ist die *Lokalität* der Wechselwirkungen. Dies bedeutet, daß der Energiefluß im Inertialbereich bei einer Skala l überwiegend Wirbel von vergleichbarer Größe rl und $r^{-1}l$ involviert.

Wir wollen nun die wesentlichen quantitativen Relationen voll entwickelter Turbulenz angeben. Dabei ist v_l die mittlere Geschwindigkeitsdifferenz zweier Punkte mit Abstand l . Wir werden im folgenden wie vorher Flüssigkeitgeschwindigkeiten mit $u(x)$ bezeichnen und deren Differenzen $u(x+l) - u(x)$ an zwei Orten mit Abstand l mit v_l . Die Wirbel-turnover Zeit, in der sich eine Struktur auf der Skala l wesentlich verformt, ist durch

$$t_l \sim \frac{l}{v_l} \quad (3.8)$$

gegeben, wie in Abb. 3.1 illustriert. Aufgrund der Inkompressibilität müssen, wenn zwei Punkte eines Objektes wie 1 und 2 auseinanderlaufen, sich andere Punkte, wie 3 und 4, nähern. Daher entspricht t_l auch der typischen Zeit, in welcher eine Anregung von einer Skala l auf kleinere Skalen transferiert wird. Der Energie-Fluß Π_l von Skala

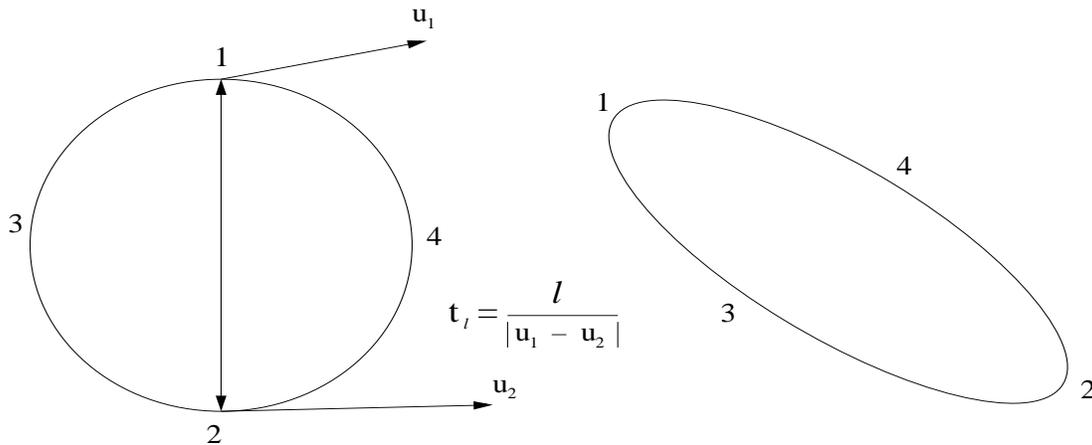


Abbildung 3.1: Querschnitt einer näherungsweise sphärischen Struktur, welche in ein Ellipsoid verformt wird. Während die Punkte 1 und 2 auseinanderrücken, müssen sich die Punkte 3 und 4 annähern.

zu Skala läßt sich damit wie folgt abschätzen:

$$\Pi_l \sim \frac{v_l^2}{t_l} \sim \frac{v_l^3}{l}. \quad (3.9)$$

Im Inertialbereich, wo es weder eine direkte Energiezufuhr noch einen direkten Energieverbrauch gibt, sollte der Energie-Fluß unabhängig von l sein und der mittleren Dissipationsrate ϵ entsprechen

$$\Pi_l \sim \frac{v_l^3}{l} \sim \epsilon. \quad (3.10)$$

Aus dieser einfachen Argumentation folgt eines der wesentlichen Ergebnisse von Kolmogorov, das Verhalten der Strukturfunktion erster Ordnung:

$$v_l \sim \epsilon^{1/3} l^{1/3}, \quad (3.11)$$

welches sagt, daß die longitudinalen Geschwindigkeitsinkremente skaleninvariant mit dem Exponenten $1/3$ sind. Dabei ist die Strukturfunktion der Ordnung n definiert durch

$$S_n(l) \equiv \langle |(\mathbf{v}(\mathbf{x}) - \mathbf{v}(\mathbf{x} + \mathbf{l})) * \frac{\mathbf{1}}{l}|^n \rangle \propto l^{n/3}. \quad (3.12)$$

Aus der Kombination von (3.8) und (3.11) ergibt sich:

$$t_l \sim \epsilon^{-1/3} l^{2/3}. \quad (3.13)$$

Eine Konsequenz dieser Beziehung ist, daß für Paare von Teilchen, die in einem Abstand l freigelassen werden, die Rate des gegenseitigen Abstandsquadrats l^2/t_l wie $l^{4/3}$ wächst. Die experimentelle Beobachtung dieses Gesetzes durch Richardson (1926) war eine der Grundlagen für die Kolmogorov-Theorie. Neben der Strukturfunktion dritter Ordnung in Gleichung (3.11) erfüllen nach Kolmogorov auch die anderen Strukturfunktionen ein entsprechendes Skalengesetz:

$$S_n = v_l^n \sim \epsilon^{n/3} l^{n/3}. \quad (3.14)$$

Der physikalische Inhalt der Kolmogorov-Theorie ist, daß mittlere Geschwindigkeitsdifferenzen von Wirbeln der Größe l wachsen wie $\epsilon^{1/3} l^{1/3}$.

Für das Energie-Spektrum [27]:

$$\varepsilon_K = \frac{1}{2} \frac{d}{dK} \int_{|\vec{k}| < K} \left(\int \langle \vec{u}(\vec{x}) \cdot \vec{u}(0) \rangle e^{-i\vec{k}\vec{x}} d\vec{x} \right) d\vec{k} \quad (3.15)$$

sagt die Kolmogorov-Theorie voraus:

$$\varepsilon_K \propto \epsilon^{2/3} K^{-5/3}. \quad (3.16)$$

Dieses Gesetz folgt direkt aus dem Verhalten der Strukturfunktion zweiter Ordnung. Das Skalenverhalten der Strukturfunktionen bestätigt sich sowohl experimentell als auch theoretisch sehr gut für $n < 4$. Für höhere Ordnungen bekommt man systematische Abweichungen zu kleineren Exponenten. Eine der entscheidenden offenen Fragen in der Theorie voll entwickelter Turbulenz ist es, den Unterschied zwischen Kolmogorov-Theorie und Experiment zu erklären, welcher sich in den anomalen Skalenexponenten äußert. Die abweichenden Exponenten deuten auf nichtgaußsches Verhalten der Wahrscheinlichkeitsverteilung hin und auf die Tatsache, daß in den kleinen Skalen größere Inkremente auftreten, als durch Kolmogorov vorhergesagt. Dieses Phänomen wird als

Intermittenz bezeichnet. Die intuitive Erklärung von Intermittenz ist, daß nur ein Teil der Moden an der turbulenten Kaskade beteiligt ist und daß der Anteil der aktiven Moden zu kleineren Skalen hin abnimmt. Das bedingt, daß die kurzwelligen Moden mehr Energie transferieren müssen und diese daher stärker angeregt sind. Diese Erklärung führte zu verschiedenen (multi)fraktalen Modellen, die in dem Buch von Frisch [22] zusammenfassend behandelt werden.

Wir wollen nun noch zusammenfassen, welche Änderungen das gezeigte Verhalten erfährt, wenn die Turbulenz zweidimensional ist. Die hierbei wesentliche Größe ist die Vortizität $\omega = \nabla \times \vec{u}$, welche die Stärke und die Orientierung der Wirbel mißt. Die Vortizität ist in drei Dimensionen ein (Pseudo-)Vektorfeld und erfüllt die folgende Gleichung für inkompressible Flüssigkeiten:

$$\partial_t \vec{\omega} + (\vec{u} \cdot \nabla) \vec{\omega} - (\vec{\omega} \cdot \nabla) \vec{u} - \nu \nabla^2 \vec{\omega} = \nabla \times \vec{f}. \quad (3.17)$$

Bei verschwindenden ν und \vec{f} enthält die Gleichung (3.17) Terme der Form $(\partial_i u^j + \partial_j u^i)$, welche eine Dehnung von ω bewirken. Dieser Mechanismus erlaubt die Dehnung und Stauchung von Wirbeln in 3D-Turbulenz und ist ein wichtiger Effekt beim Energietransfer zwischen unterschiedlichen Skalen. In 2D-Turbulenz ist die Vortizität nur ein (Pseudo-)Skalar $\omega = \varepsilon^{ij} \partial_i u^j$, und Gleichung (3.17) reduziert sich zu:

$$\partial_t \omega + (\vec{u} \cdot \nabla) \omega - \nu \nabla^2 \omega = \varepsilon^{ij} \partial_i f^j. \quad (3.18)$$

Bei verschwindenden ν und \vec{f} wird die Vortizität hier nur entlang Lagrange-Trajektorien $\vec{x}(t)$ von (imaginären) Teilchen im Fluß transportiert, mit $\frac{d\vec{x}}{dt} = \vec{u}(\vec{x}, t)$. Die totale Zeitableitung von $\omega(t)$ verschwindet. Daraus folgt die wesentliche Beobachtung, daß in 2D-Turbulenz neben der Energie auch die Enstrophie $\Phi = \frac{1}{2} \langle \omega^2 \rangle$ erhalten bleibt. In seiner Arbeit von 1967 erkannte Kraichnan [28], daß dieses Erhaltungsgesetz ein völlig anderes Kaskadenbild für 2D-Turbulenz impliziert. Zunächst sei hierzu die Beziehung zwischen Energie- und Enstrophiespektrum hergestellt:

$$\Phi_K = \frac{1}{2} \frac{d}{dK} \int_{|\vec{k}| < K} \left(\int \langle \vec{\omega}(\vec{x}) \cdot \vec{\omega}(0) \rangle e^{-i\vec{k}\vec{x}} d\vec{x} \right) d\vec{k} = K^2 \varepsilon_K. \quad (3.19)$$

Damit müssen auch die Flüsse von Energie und Enstrophie aus einer Wellenzahl K in die benachbarten Wellenzahlen durch den Faktor K^2 zueinander in Beziehung stehen, wenn diese lokal im Fourierraum sind. Daher können nicht beide Flüsse gleichzeitig konstant in K sein. Kraichnan argumentierte, daß auf Skalen kleiner als der Injektionskala der Enstrophiefluß konstant sei, wobei die gesamte Enstrophie auf kleinere Skalen transferiert wird. Der Energiefluß wird dadurch gedämpft, und resultierend wird die Energie auf größere Skalen als die Injektionskala transferiert. Diesen Prozeß nennt man **inverse Energie-Kaskade**. Die Kolmogorov-Gesetze gelten in 2D-Turbulenz nur im

Bereich der inversen Energie-Kaskade. Dies bedeutet für das Energiespektrum das Verhalten (3.16) und für die Strukturfunktionen die Gültigkeit von (3.14). Darüber hinaus wurde in experimentellen [29] und numerischen [30] Analysen festgestellt, daß die inverse Kaskade nicht-intermittent ist. Dieses Verhalten werden wir in unserem Experiment bestätigen. Wir werden uns in den folgenden Analysen im wesentlichen auf den Bereich der inversen Kaskade beschränken. Für das Regime der direkten Enstrophie-Kaskade gilt für das Energie-Spektrum

$$\varepsilon_K \propto K^{-3}. \quad (3.20)$$

3.2 Experiment

Das nun beschriebene Experiment wurde am *Center for Nonlinear Dynamics* an der *University of Texas at Austin* durch die Gruppe von Prof. Harry Swinney durchgeführt und ist z.B. in [31] detailliert beschrieben. Der experimentelle Aufbau besteht aus einem rotierenden rotations-symmetrischen Tank, der mit Wasser gefüllt ist. Der Tank besitzt eine feste obere Abdeckung, und der Boden ist konisch, um den Beta-Effekt der Atmosphäre (die Variation der Coriolis-Kraft mit dem Breitengrad) zu simulieren. Wasser wird in einem geschlossenen Kreislauf durch in konzentrischen Ringen angeordnete Löcher in den Tank gepumpt und wieder abgesaugt, siehe Abb. 3.2. Der somit erzeugte radiale Fluß koppelt mit der Coriolis-Kraft, wobei ein azimuthaler Strom (Jet) erzeugt wird. Wir analysieren die Strömungsgeschwindigkeit in diesem stark turbulenten gegenrotierenden (entgegen der Rotationsrichtung des Tanks) Jet, der durch das radial nach außen gepumpte Wasser erzeugt wird. Die Messungen wurden für zwei Parametersätze durchgeführt: (a) eine Rotationsrate von $\Omega = 1.57 \text{ rad/s}$ und eine Pumprate von $Q = 450 \text{ cm}^3/\text{s}$, was zu einem 3-dimensionalen Fluß führt, (b) eine Rotationsrate von $\Omega = 11.0 \text{ rad/s}$ und eine Pumprate von $Q = 150 \text{ cm}^3/\text{s}$, was einen 2-dimensionalen Fluß ergibt.

Die Pumprate Q wurde so hoch ($450 \text{ cm}^3/\text{s}$) eingestellt, um voll entwickelte Turbulenz auch bei der kleineren der beiden Rotationsraten zu erhalten. Im Fall (a) war die Reynolds-Zahl 26 000, basierend auf der mittleren azimuthalen Geschwindigkeit U und der Längenskala L , die dem Abstand der beiden Pumpringe (16.2 cm) entspricht. Die Pumprate bei der höheren Rotationsrate wurde verringert, um die Reynolds-Zahl (35 000) für den resultierenden 2D-Fluß vergleichbar zu der im entsprechenden 3-D Fluß zu bekommen. Der Einfluß der Rotation wird, wie im vorigen Abschnitt ausgeführt, durch die Rossby-Zahl Ro quantitativ erfaßt. Kleine Rossby-Zahlen beschreiben Flüsse, in denen der Einfluß der Rotation dominiert. Für unsere Daten ist $Ro = 0.06$ im Fall hoher Rotationsrate und $Ro = 0.33$ im Fall geringer Rotation. Letztere Rossby-Zahl ist bereits ausreichend hoch, um einen deutlich drei-dimensionalen Fluß zu erzeugen. Die Flüssigkeitgeschwindigkeit wurde mittels heißer Filmsonden (hot film probes) bei einer Abtastrate von 150 Hz gemessen. Für jeden Satz von Kontrollparametern wurde

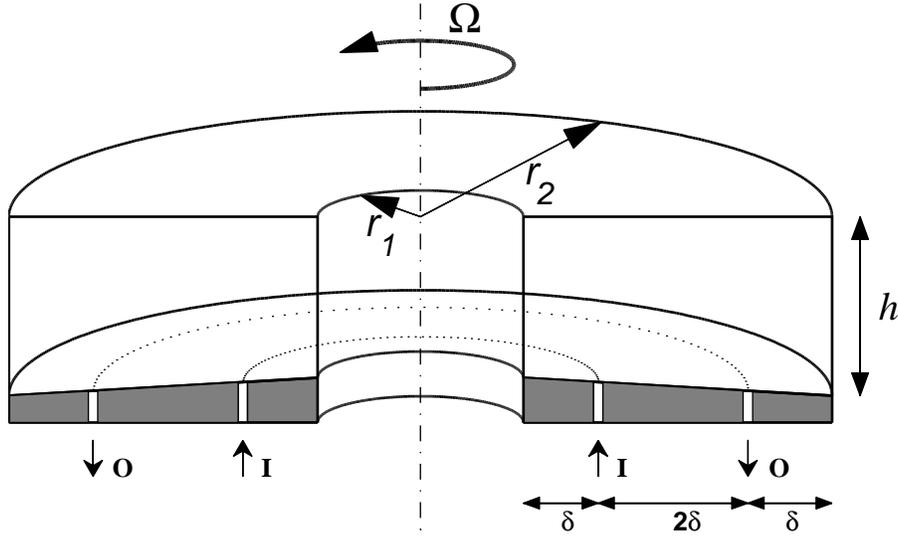


Abbildung 3.2: Querschnitt des rotierenden Tanks. $r_1 = 10.8$ cm, $r_2 = 43.2$ cm, $\delta = 8.1$ cm, $h = 17.1$ cm bei r_1 , und $h = 20.3$ cm bei r_2 . I ist der Zufluß und O der Abfluß.

eine Zeitreihe von 10^6 Datenpunkten gemessen.

Die Geschwindigkeitssonden wurden an der Deckelplatte des Tanks, 1 cm in die Flüssigkeit reichend, mittig zwischen innerem und äußerem Rand des Behälters angebracht. Die Orientierung wurde derart gewählt, daß die azimuthale Geschwindigkeit gemessen werden konnte, siehe Abb. 3.3. Die Abtastrate von $f = 150$ Hz entspricht einer räumlichen Skala von $U/f \approx 0.1$ cm, was nahe der geschätzten Komogorov-Dissipationslänge von $\eta = 0.07$ cm liegt. Diese Länge ist definiert durch $\eta = (\nu^3/\epsilon)^{1/4}$, wobei ϵ wieder die Rate des Energie-Flusses ist.

Die Geschwindigkeitsdaten wurden im Sinne der Taylor-Hypothese der eingefrorenen Turbulenz interpretiert, welche besagt, daß Fluktuationen des Geschwindigkeitsfeldes in der Zeit, die durch räumlich feste Sonden aufgenommen werden, als räumliche Fluktuationen der Strömung betrachtet werden. So werden beispielsweise Geschwindigkeitssinkremente Δu_S auf einer räumlichen Skala $S = \langle U \rangle t$ ausgewertet, wobei $\langle U \rangle$ die mittlere Geschwindigkeit ist.

Der im Experiment untersuchte gegenrotierende Jet ist sehr instabil und wird schon bei geringen Pumpraten Q turbulent [32]. Im zweidimensionalen Fall beobachtet man jedoch langlebige kohärente Wirbel, die mit dem mittleren Fluß bewegt werden. Diese Wirbel sind sehr gut in Abbildung 3.3 zu beobachten, wo die Stromfunktion $\phi(x, y)$, mit $v_x = \partial\phi/\partial y$ und $v_y = -\partial\phi/\partial x$, dargestellt ist. Die Inkompressibilität und die Zweidimensionalität der Flüssigkeit erlaubt es, eine Stromfunktion $\phi(x, y)$ zu definieren. Im

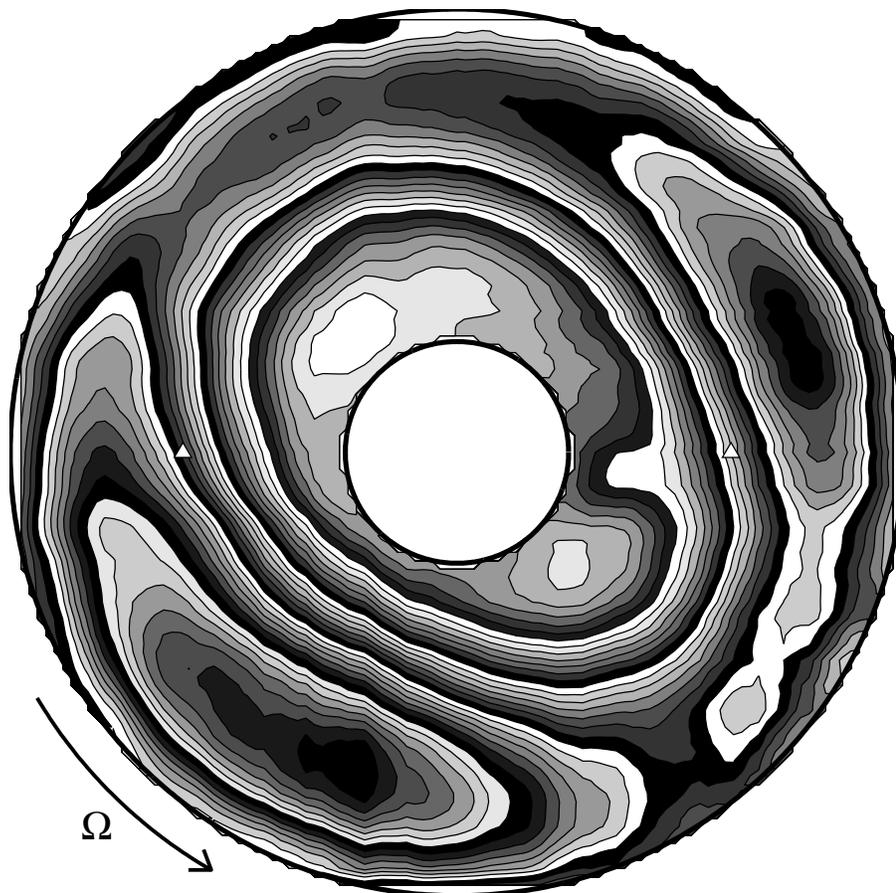


Abbildung 3.3: Höhenlinien der Stromfunktion des 2D-Flusses ($\Omega = 11.0 \text{ rad/s}$, $Q = 150 \text{ cm}^3/\text{s}$). Der Ort der Filmsonden ist durch \triangle angezeigt. Die Rotationsrate von 0.37 rad/s wurden von der Rotationsrate des Tanks 11.0 rad/s abgezogen, um den Beobachter den Blick aus dem mit der mittleren Flußgeschwindigkeit rotierenden Bezugssystem zu gestatten. Im Bezugssystem des Tanks werden die zyklonischen (dunkel) und anti-zyklonischen (hell) Wirbel entgegen des Uhrzeigersinns mit der Strömung bewegt. Die engen Stromlinien im Zentrum des Kanals (des Jets) deuten auf hohe Geschwindigkeiten, die weiteren Abstände im Zentrum der Wirbel auf geringe Geschwindigkeiten.

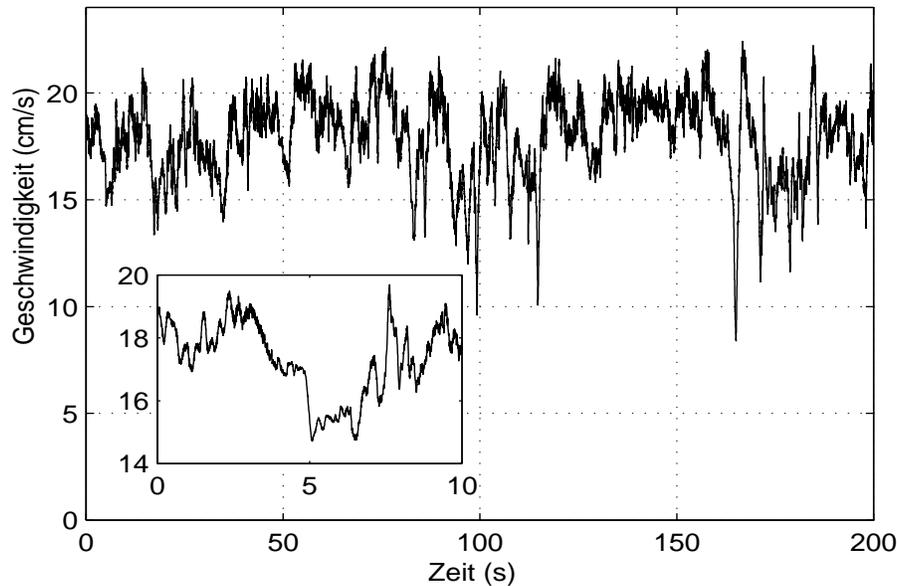


Abbildung 3.4: Typische Zeitreihe der Geschwindigkeit gemessen bei $\Omega = 11.0$ rad/s und $Q = 150$ cm³/s. Das 10 s Segment, welches in dem inneren Fenster gezeigt ist, illustriert die selbstähnliche turbulente Feinstruktur.

Bezugssystem des Tanks werden die zyklonischen (dunkel) und anti-zyklonischen (hell) Wirbel entgegen des Uhrzeigersinns mit der Strömung bewegt. Die engen Stromlinien im Zentrum des Kanals (des Jets) deuten auf hohe Geschwindigkeiten, die weiteren Abstände im Zentrum der Wirbel auf geringe Geschwindigkeiten. Die Geschwindigkeit, mit welcher sich die Wirbel bewegen, variiert durch deren Wechselwirkung mit dem Jet und mit anderen Wirbeln. Wirbel mit der selben Rotationsrichtung verschmelzen, wogegen sich solche mit unterschiedlicher Richtung abstoßen. Während sich der rotierende Jet über die Sonden hinwegbewegt, beobachtet man ein Umschalten zwischen Regionen mit hoher azimuthaler Geschwindigkeit innerhalb des Jets und Regionen, in denen der Fluß stark radial ist. In letzteren Abschnitten sieht man in der Zeitreihe ein deutliches Absinken der Geschwindigkeit. Eine typische Zeitreihe der Flüssigkeitsgeschwindigkeit ist in Abb. 3.4 dargestellt.

3.3 Allgemeine statistische Eigenschaften

In diesem Abschnitt möchten wir sowohl für 2D- als auch 3D-Turbulenz die wesentlichen statistischen Größen, wie Wahrscheinlichkeitsverteilungen, Strukturfunktionen und Energiespektren angeben. Diese stimmen mit dem in Abschnitt 3.1 vorhergesagten Verhalten in guter Näherung überein. Zunächst zeigen wir in Abb. 3.5 a) die Wahrscheinlichkeitsverteilungen für den dreidimensionalen Fall. Man erkennt deutlich die

unterschiedliche Form der Graphen für verschiedene Abstände. Während das Verhalten auf kleinen Skalen näherungsweise exponentiell ist, beobachtet man nahezu Gaußverteilungen auf großen Skalen. Dies ist die in Abschnitt 3.1 beschriebene Intermittenz, eine stärkere Aktivität auf kleinen Skalen als durch die Kolmogorov-Theorie vorhergesagt. In Abb. 3.5 b) zeigen wir die Wahrscheinlichkeitsverteilungen für den zweidimensionalen Fall. Man erkennt, daß alle Kurven eine ähnliche Gestalt zeigen, welche einer gestreckten exponentiellen Verteilung entspricht. Obwohl die Wahrscheinlichkeitsverteilungen nicht gaußförmig sind, zeigt das selbstähnliche Verhalten auf unterschiedlichen Skalen das Fehlen von Intermittenz.

Es ist wesentlich, den Bereich der Längenskalen, in welchem wir die Geschwindigkeitsinkremente berechnen, sorgfältig zu betrachten. Die Skalenargumente sollten im Inertialbereich gültig sein. In der 2D-Turbulenz ist der Inertialbereich in die inverse Energie-Kaskade (für Abstände größer als die Injektionsskala l_0) und die direkte Enstrophie-Kaskade (für Abstände kleiner als die Injektionsskala l_0) unterteilt. Wir wissen aus [32], daß die Injektionsskala für unser Experiment $l_0 \sim 3\text{cm}$ ist. Wir werden nun für den Bereich der inversen Energie-Kaskade die Strukturfunktionen analysieren.

Eine weitere Statistik, um den Unterschied zwischen 2D- und 3D-Turbulenz zu untersuchen sind die Strukturfunktionen S_n der Ordnung n in Gleichung (3.14). Wir analysieren diese mit einer Methode, welche durch Benzi *et al.* [33] vorgeschlagen wurde und Extended Self Similarity (EES) genannt wird. Dabei werden alle Strukturfunktionen S_n gegen S_3 aufgetragen anstatt gegen die Geschwindigkeitsdifferenz v_l . Dies macht im Idealfall keinen Unterschied zur konventionellen Methode, S_n versus v_l darzustellen, da $S_3 \propto v_l$ gilt. Für Daten mit begrenztem Inertialbereich und geringer Statistik wird das Skalenverhalten mit dieser Methode jedoch typischerweise besser dargestellt, da man systematische Fehler ausschalten kann, die in den Strukturfunktionen aller Ordnungen auftreten. Die Strukturfunktionen sind in Abb. 3.6 gezeigt. Man erkennt, daß die Graphen für 3D-Turbulenz bei höheren Ordnungen $n = 4, 5$ deutlich vom theoretisch vorhergesagten Verhalten abweichen. Die tatsächlich beobachteten Skalenexponenten sind kleiner als durch die Kolmogorov-Theorie vorhergesagt. Dies ist der Effekt der Intermittenz. Für 2D-Turbulenz sieht man eine sehr gute Übereinstimmung zwischen theoretischem und experimentell beobachtetem Verhalten. Die Daten sind selbstähnlich und nichtintermittent.

Als letzte Methode der “konventionellen” Statistik wollen wir noch die Energie-Spektren in beiden Flußsituationen berechnen, siehe Abb. 3.7. Man erkennt sowohl für 2D- als auch für 3D-Turbulenz deutliche Abweichungen vom vorhergesagten Verhalten. Das experimentell beobachtete Spektrum ist steiler, d.h., der Skalenexponent ist betragsmäßig größer. Dies bedeutet, daß große Strukturen häufiger vorkommen, als durch die Theorie vorhergesagt. Als Grund hierfür werden die in dem betrachteten Experiment beobachteten ausgeprägten kohärenten Strukturen angenommen.

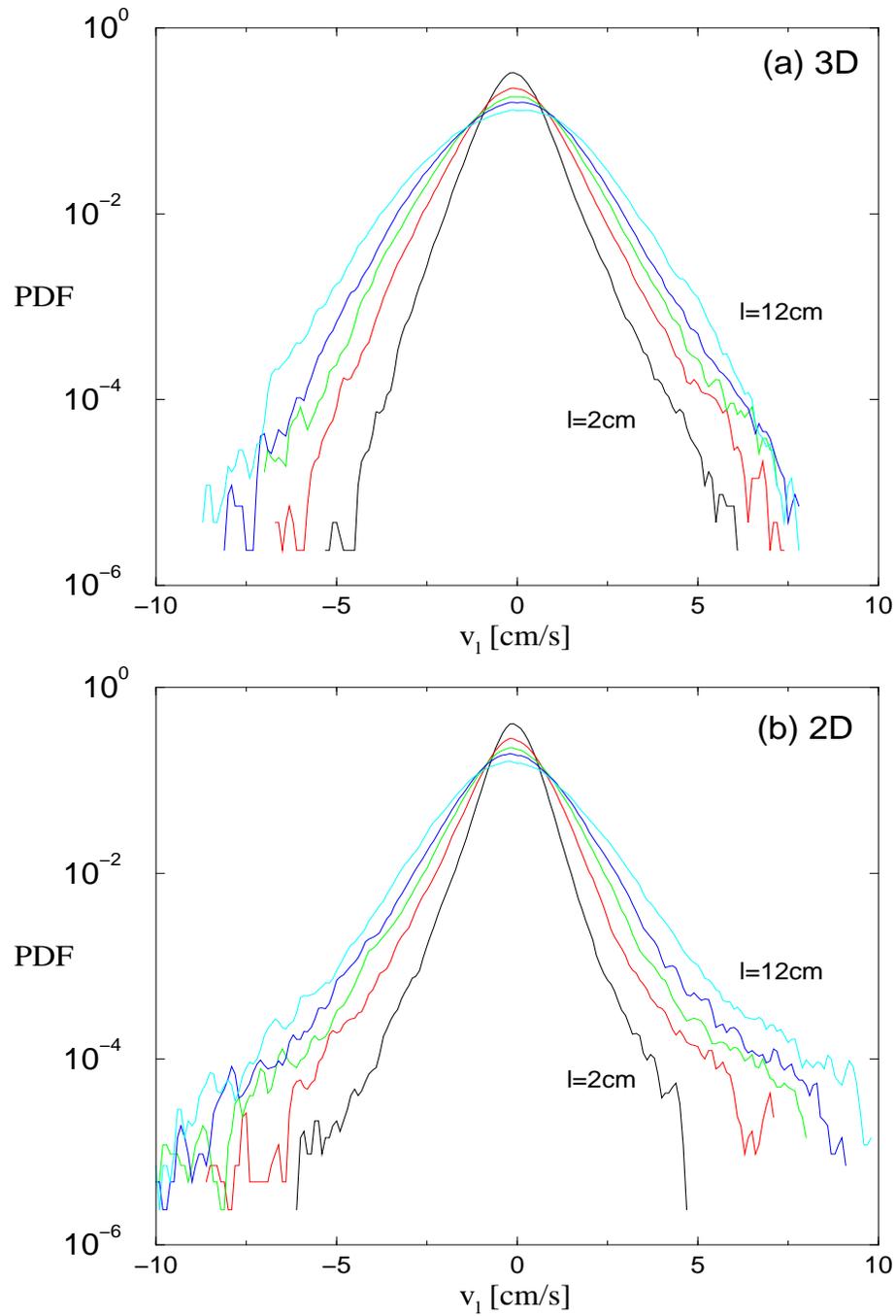


Abbildung 3.5: Wahrscheinlichkeitsverteilung (PDF) versus Geschwindigkeitsdifferenz v_l bei (a) $\Omega = 1.57$ rad/s und $Q = 450$ cm³/s (b) $\Omega = 11.0$ rad/s und $Q = 150$ cm³/s. Während 2D-Turbulenz selbstähnliches Verhalten zeigt, ist bei dem 3D-Fluß der Übergang von exponentiellem zu gaußischem Verhalten sichtbar.

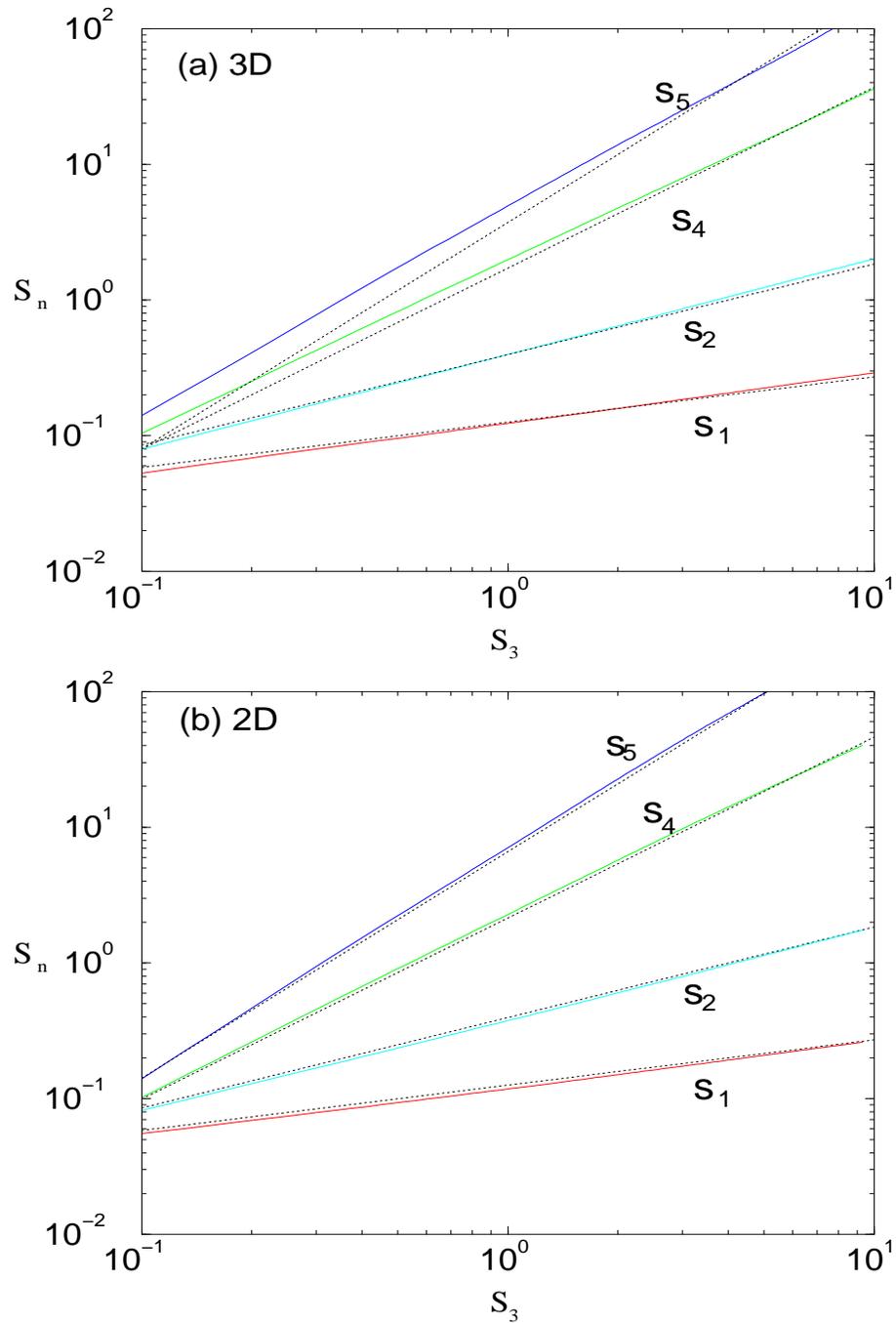


Abbildung 3.6: Strukturfunktionen S_n versus S_3 (EES) bei (a) $\Omega = 1.57$ rad/s und $Q = 450$ cm³/s (b) $\Omega = 11.0$ rad/s und $Q = 150$ cm³/s. Während 2D-Turbulenz das Kolmogorov-Verhalten zeigt, sind beim 3D-Fluß deutliche Abweichungen von der Theorie erkennbar.

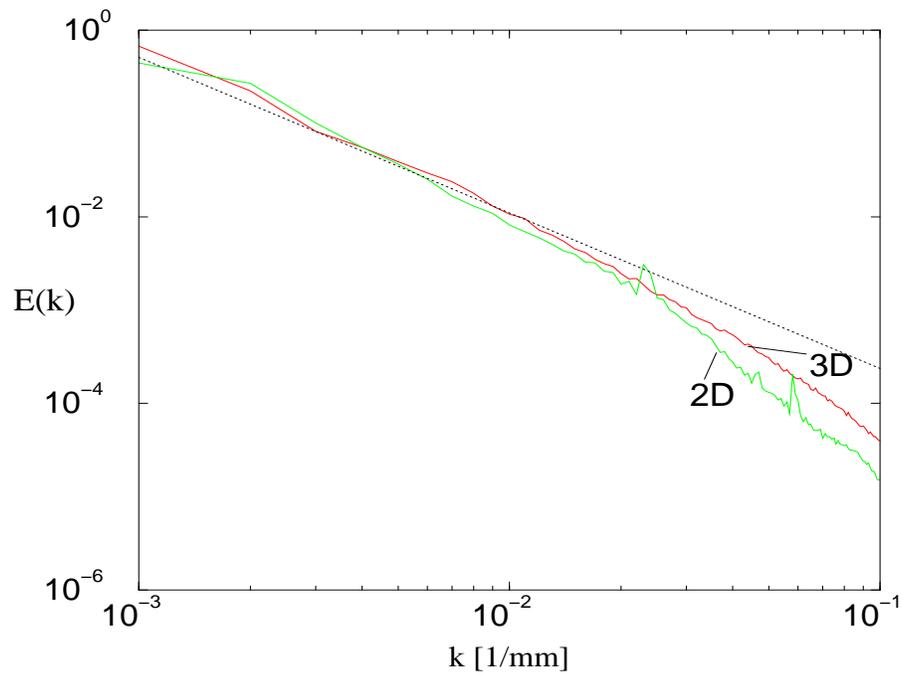


Abbildung 3.7: Energie-Spektren für den Fluß bei $\Omega = 1.57$ rad/s und $Q = 450$ cm³/s (rot) sowie $\Omega = 11.0$ rad/s und $Q = 150$ cm³/s (grün). Man erkennt, daß in beiden Fällen deutliche Abweichungen vom Kolmogorov-Verhalten (schwarz) auftreten. Als Ursache werden die ausgeprägten kohärenten Strukturen angenommen.

3.4 Deterministische Struktur

Wir werden nun die in Abschnitt 2.2.3 behandelten Modelle nutzen, um den deterministischen Anteil der Dynamik in voll entwickelter Turbulenz zu untersuchen. Auf der Basis der vergangenen m Meßwerte der Flüssigkeitsgeschwindigkeit werden wir einen zukünftigen Wert, einen Zeitschritt S in der Zukunft, vorhersagen. Wie im Abschnitt 2.2.5 werden wir dann die Statistik der Inkremente $v_S = \hat{u}(t+S) - u(t)$ zwischen dem momentanen Wert $u(t)$ und dem vorhergesagten Wert $\hat{u}(t+S)$ auswerten. Dieses Vorgehen ist in Abb. 3.8 schematisch dargestellt. Mathematisch wertet man wieder bedingte Wahrscheinlichkeiten aus

$$p(u_{k+S}|\vec{u}_k) \approx \hat{p}(u_{n+S}|\vec{u}_n) \quad \forall \vec{u}_k \in \mathcal{U}_n, \quad (3.21)$$

wobei $\vec{u}_n = (u_n, u_{n-1}, \dots, u_{n-m+1})$ der Einbettungsvektor ist, welcher aus den letzten m Meßwerten der Geschwindigkeit gebildet wird. Dabei erhalten wir eine Vorhersage \hat{u}_{n+S} wieder als arithmetischen Mittelwert über die Zukünfte der Nachbarn u_{k+S} .

Wir haben in Abschnitt 2.2.4 die optimale Wahl der Einbettungsparameter m und

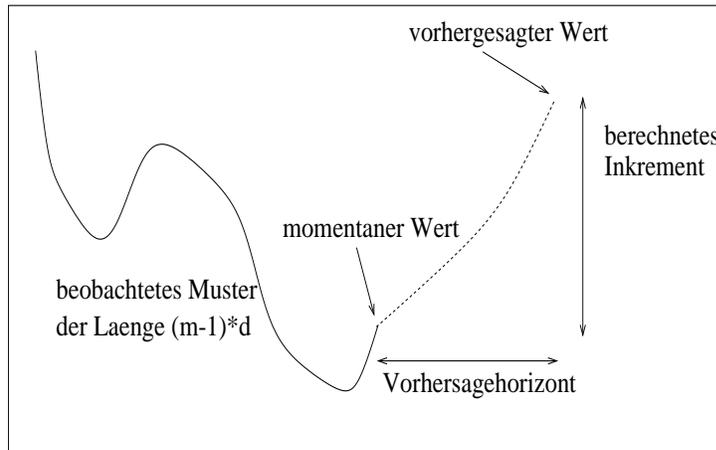


Abbildung 3.8: Schematische Darstellung des lokal konstanten Vorhersagers und Berechnung der Inkremente.

τ besprochen. Die Dimension m sollte in einem Markov-Prozeß der Ordnung des Prozesses entsprechen. Eine rigorose Ordnung kann für die Daten voll entwickelter Turbulenz natürlich nicht angegeben werden, und es ist zweifelhaft, ob diese endlich wäre. Wir nehmen daher an, daß die letzten m Meßwerte die wesentliche Information über die Wahrscheinlichkeitsverteilung des Zukunftswertes $\hat{p}(u_{n+S})$ enthalten und höhere Ordnungen Korrekturen hierzu darstellen. Wir werden für unsere Analysen den Wert $m = 10$ verwenden, was groß genug ist, um sinnvolle Vorhersagen zu erhalten und klein genug, um den numerischen Aufwand machbar zu halten. Wir werden in Abschnitt 3.4.1

zeigen, daß die qualitative Aussage unserer Ergebnisse nicht von der speziellen Wahl von m abhängt.

Die Zeitverzögerung τ sollte in einem Markov-Prozeß mit Langevin-Dynamik und starkem stochastischen Anteil wesentlich kleiner sein als z.B. dem ersten Nulldurchgang der Autokorrelationsfunktion entspricht. Wir haben in Abschnitt 2.2.4 gezeigt, daß bei dynamisch verrauschten Systemen mit hinreichend großer Rauschamplitude der optimale Wert von τ der Abtastrate der Zeitreihe entspricht. Auch für die Turbulenzdaten haben wir überprüft, daß der mittlere Vorhersagefehler minimal wird, wenn τ der zeitlichen Auflösung der Meßsonde entspricht. Daher wählen wir $\tau = 2 \text{ mm}$ für die folgenden Analysen, werden jedoch in Abschnitt 3.4.1 zeigen, daß die qualitative Aussage unserer Ergebnisse auch nicht von der speziellen Wahl von τ abhängt. Daß die Einheit des Zeitversatzes in mm angegeben wird, liegt an der oben dargestellten Interpretation der Daten im Sinne der Taylorhypothese, zeitliche Fluktuationen werden als solche im Raum betrachtet.

Wir nutzen mindestens 30 Nachbarn zum Fit der verwendeten Phasenraummodelle. Die Wahrscheinlichkeitsverteilungen (PDFs) $p(v_S)$ der gemessenen Inkremente, der li-

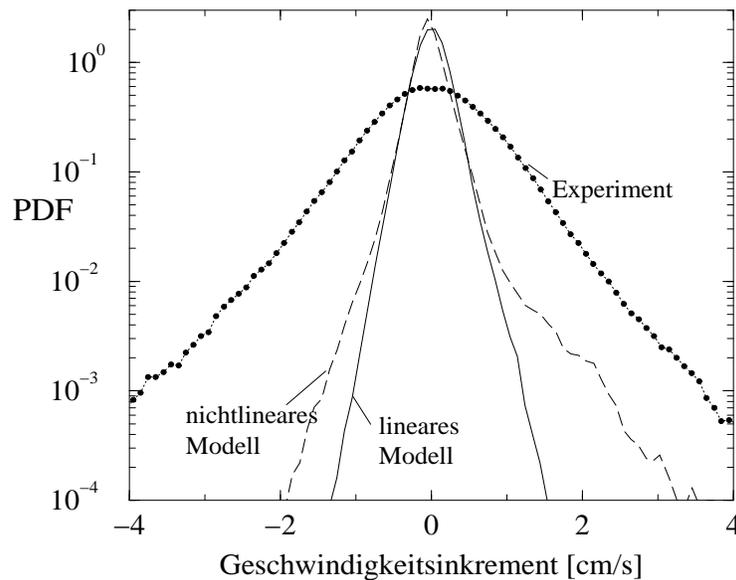


Abbildung 3.9: PDF der gemessenen Geschwindigkeitsinkremente v_S , der linear und nichtlinear vorhergesagten Inkremente, für den 2D-Fluß mit $m = 10$, $\tau = 2 \text{ mm}$, und $S = 20 \text{ mm}$.

near vorhergesagten Inkremente sowie der nichtlinear vorhergesagten Inkremente sind in Abb. 3.9 für den zweidimensionalen Fluß dargestellt. Lineare Vorhersagen wurden wieder mit einem $AR(m)$ -Modell durchgeführt, dessen Ordnung m der des lokal konstanten Modells entspricht. Man beobachtet eine deutliche Asymmetrie der PDF für

die nichtlinear vorhergesagten Inkremente, die negativen vorhergesagten Inkremente sind fast so stark unterdrückt wie die linear vorhergesagten. Dagegen erkennt man für positive v_S einen deutlichen Unterschied zwischen linearem und nichtlinearem Modell. Wir werden im Verlauf dieses Abschnitts zeigen, daß der Ursprung dieses Unterschieds kohärente bzw. deterministische Strukturen in der Flüssigkeitsbewegung sind. Wir werden uns im folgenden auf den positiven Teil der PDF beschränken und werden den Grund der Asymmetrie am Ende dieses Abschnitts erläutern.

Der positive Bereich der PDFs für den 2D-Fall im Vergleich zum 3D-Fluß ist in Abb. 3.10 für dieselben Einbettungsparameter wie in Abb. 3.9 gezeigt. Die PDFs der eigentlichen Daten hängen vom Abstand S ab, wie in Abschnitt 3.3 erläutert. Für den hier betrachteten Wert von S zeigen die PDFs sowohl im 2D- als auch im 3D-Fall exponentielles Verhalten.

Das Verhalten der vorhergesagten Inkremente unterscheidet sich deutlich in den beiden Situationen. Im 3D-Fluß fallen sowohl die linear als auch die nichtlinear vorhergesagte Verteilung exponentiell ab. Folglich enthält die Zeitreihe nur lineare Korrelationen, alle anderen Fluktuationen werden durch beide Modelle als Rauschen gedeutet. Wie wir auch schon in Abb. 3.9 gesehen haben, unterscheiden sich die Vorhersagen der beiden Modelle im 2D-Fall deutlich voneinander. Die Inkremente, die durch das lineare Modell vorhergesagt werden, fallen wie im 3D-Fall exponentiell ab. Dagegen ist der Zerfall der nichtlinear vorhergesagten Inkremente deutlich langsamer und legt ein algebraisches Verhalten nahe, siehe Abschnitt 3.4.2.

Die Diskrepanz zwischen experimenteller und nichtlinear vorhergesagter PDF rührt von dem stochastischen Anteil der Dynamik her. Dieser kann durch keinen Vorhersager genutzt werden, wie in Abschnitt 2.2.3 ausführlich diskutiert. Dieser Unterschied zeigt wieder die verschiedene Bedeutung von Modellierung und Vorhersage in stochastischen dynamischen Systemen. Hier interessieren wir uns für Vorhersagen und nicht dafür, ein Modell zu konstruieren, welches die experimentell beobachtete PDF möglichst genau wiedergibt. Der Unterschied zwischen linear und nichtlinear vorhergesagter PDF legt jedoch das Vorhandensein nichtlinearer Korrelationen in den Daten nahe und ist daher für uns von Interesse. Diese Differenz wird uns auf nichtlinearen Determinismus in den Daten führen und ist in Abschnitt 3.4.3 detailliert diskutiert.

3.4.1 Abhängigkeit von Parametern

Wir wollen nun zeigen, daß die qualitative Struktur der PDFs im 2D-Fall nicht von den Modellparametern m und τ abhängt. Zunächst variieren wir die Einbettungsdimension $m = 5, 10, 15$ und halten $\tau = 4$ mm fest, ändern jedoch auch den Vorhersagehorizont $S = 20, 40, 60$ mm. Wie in Abb. 3.11 a) gezeigt, ergibt sich in allen Fällen der beschriebene Unterschied zwischen dem linearen und dem nichtlinearen Modell. Im nächsten Schritt untersuchen wir das Verhalten für verschiedene Zeitverzögerungen $\tau = 2, 3, 4, 5$

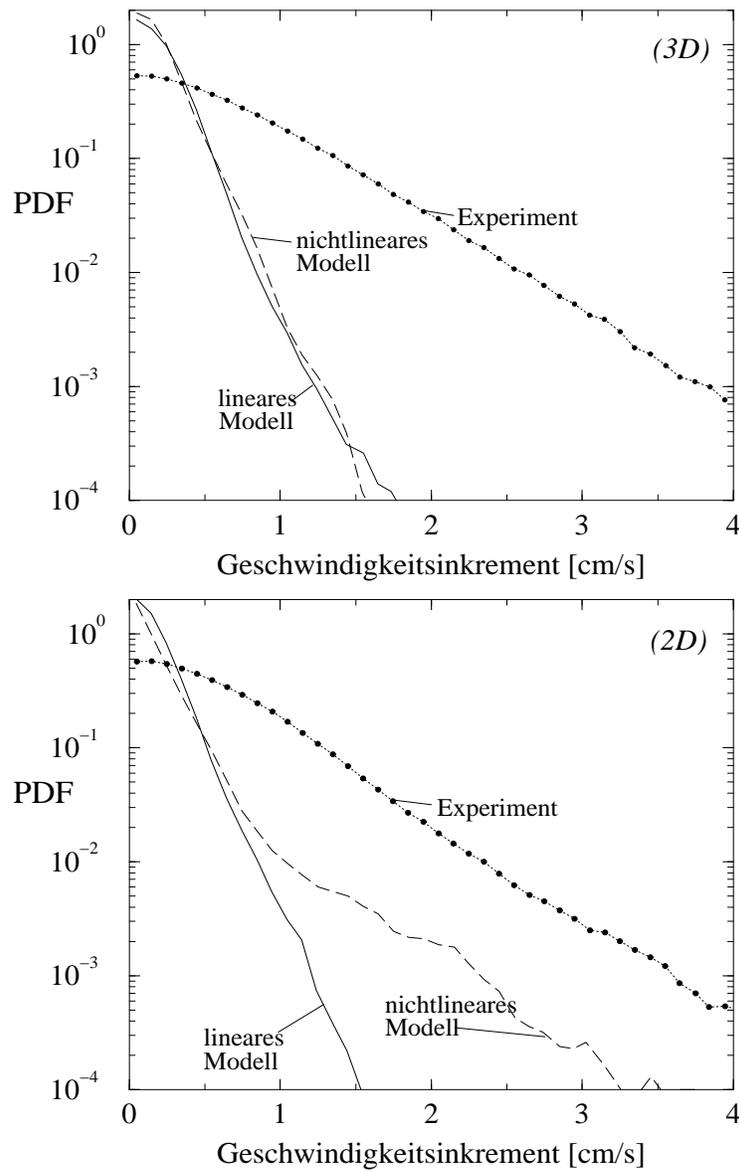


Abbildung 3.10: PDFs der positiven Geschwindigkeitsinkremente für das Experiment, das lineare Modell und das nichtlineare Modell für (a) den 3D-Fluß und (b) den 2D-Fluß mit $m = 10$, $\tau = 2$ mm und $S = 20$ mm.

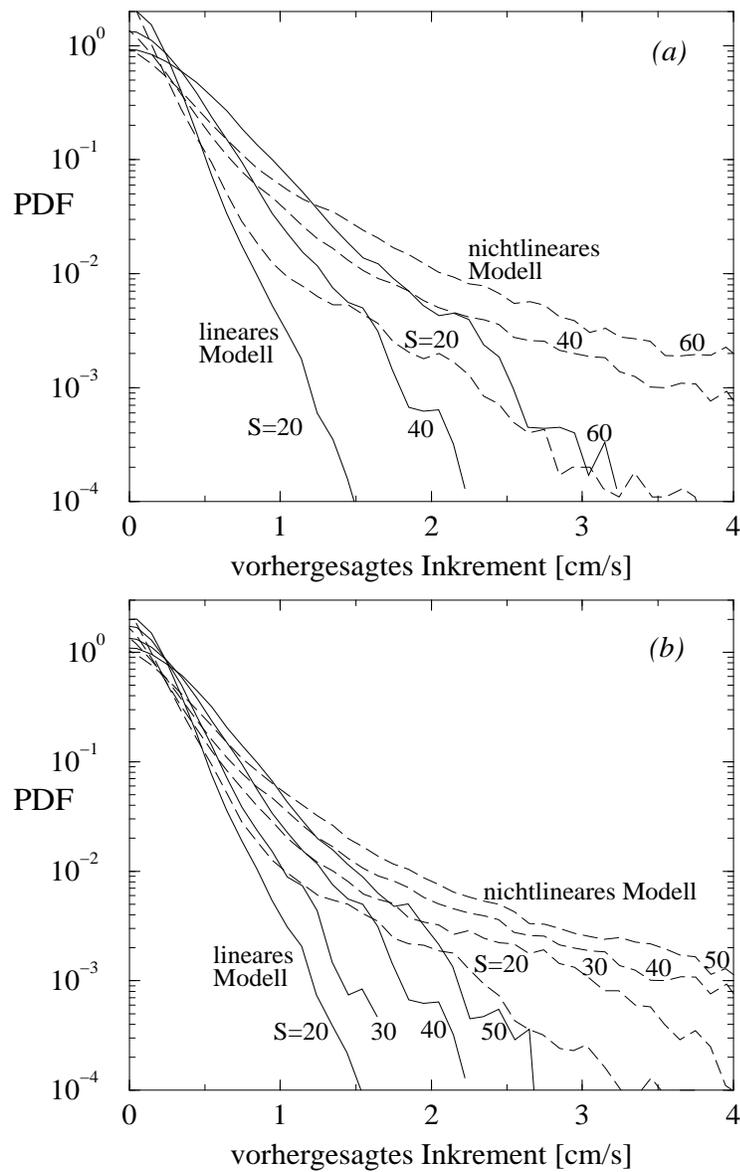


Abbildung 3.11: PDFs der positiven Geschwindigkeitsinkremente für 2D-Turbulenz für das lineare Modell (durchgezogen) und das nichtlineare Modell (gestrichelt) für (a) $m = 5, 10, 15$ mit $d = 4$ mm und $S = 20, 40, 60$ mm sowie (b) $\tau = 2, 3, 4, 5$ mm mit $m = 10$ und $S = 20, 30, 40, 50$ mm, wobei für $\tau = 2$ $S = 20$ gilt usw.

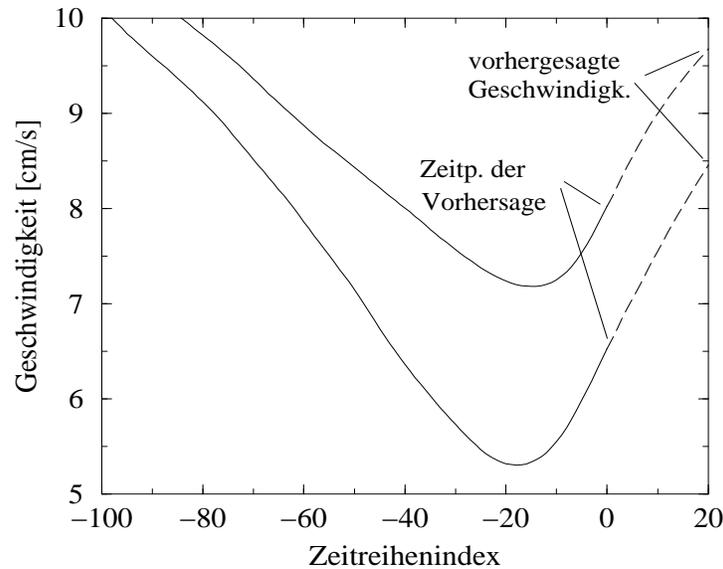


Abbildung 3.12: Gemittelttes Muster des Geschwindigkeitssignals, wenn ein positives Inkrement von 1.5 cm/s (obere Kurve) und 2 cm/s (untere Kurve) durch das nichtlineare Modell vorhergesagt wird mit $m = 10$, $\tau = 2$ mm und $S = 20$ mm.

mm bei konstantem $m = 10$ und $S = 20, 30, 40, 50$ mm, wie in Abb. 3.11 b) gezeigt. Für alle Parameterwerte fällt die PDF der nichtlinear vorhergesagten Inkremente deutlich langsamer ab.

3.4.2 Skalenverhalten der PDF

Wir wollen nun untersuchen, ob man das funktionale Verhalten der PDF des nichtlinearen Modells aus dem Wissen über die Skalentheorie voll entwickelter Turbulenz erklären kann. Wir haben festgestellt, daß der Graph der PDF für große Inkremente ein Potenzgesetz nahelegt, siehe Abb. 3.10. Dies kann man verstehen, wenn man die Wirkung des lokal konstanten Modells genauer analysiert. Dazu fragen wir zunächst, *wann* ein großes Inkrement durch das nichtlineare Modell vorhergesagt wird. In Abb. 3.12 ist das gemittelte Muster des Geschwindigkeitssignals gezeigt, wenn eine Erhöhung des Signals um 1.5 bzw. 2 cm/s vorhergesagt wird. Es ist ersichtlich, daß die Zeitreihe ein tiefes Minimum aufweist, bevor ein starkes Ansteigen der Geschwindigkeit vorhergesagt wird. Aus Abschnitt 3.2 wissen wir, daß diese Minima den Zentren der kohärenten Wirbel entsprechen. Wir nehmen also an, daß eine bestimmte vorhergesagte Erhöhung des Signals einer typischen Wirbelgröße entspricht und daß die Erhöhungen nur in einem kleinen Abschnitt, dem Zentrum, des Wirbels vorhergesagt werden, dessen Größe unabhängig von der tatsächlichen Länge des Wirbels ist.

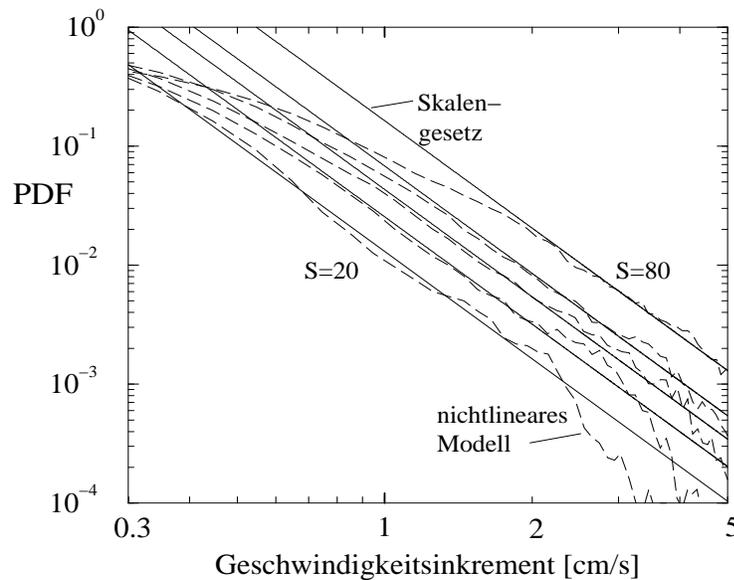


Abbildung 3.13: Vergleich der PDF der nichtlinear vorhergesagten Inkremente (gestrichelte Linien) mit dem aus der Kolmogorov-Theorie hergeleiteten Skalenverhalten (durchgezogene Linien). Die verschiedenen Kurven entsprechen unterschiedlichen Einbettungsparametern von unten nach oben: $\tau = 2, 3, 4, 5, 8$ mm und $S = 20, 30, 40, 50, 80$ mm.

Das lokal konstante Modell mittelt nun über alle ähnlichen Strukturen in der Vergangenheit der Zeitreihe, also über alle Wirbel derselben Größe. Zu jedem Wirbel der Größe r_0 wird damit das zugehörige mittlere Geschwindigkeitsinkrement $\langle v_r \rangle_{r=r_0}$ ausgerechnet. Damit wird die Anzahl der Wirbel mit der Größe r_0 direkt in die Anzahl von Wirbeln mit mittlerem Geschwindigkeitsinkrement der Größe $\langle v_r \rangle_{r=r_0}$ übersetzt. Wir interpretieren das lokal konstante Modell also hier nicht als Vorhersager, sondern als Mittelungsmethode im Phasenraum.

Wir wollen nun zeigen, welche Aussage die Kolmogorov-Theorie über die Verteilung der Wirbel versus deren mittlerem Geschwindigkeitsinkrement macht. Wie in Abschnitt 3.1 besprochen, liegt der Richardson-Kaskade die Annahme der Skaleninvarianz zugrunde, d.h., Wirbel verschiedener Größe sind gleichermaßen raumfüllend. Das bedeutet eine Verteilung der Wirbel wie r^{-3} in 3D-Turbulenz und wie r^{-2} im 2D-Fall. Da unsere Messungen entlang einer eindimensionalen Linie durchgeführt sind, müssen wir den Wirkungsquerschnitt der Wirbel berücksichtigen und erhalten in beiden Fällen eine Verteilung der Wirbel wie $N(r) = r^{-1}$. Die Beziehung zwischen der Wirbelgröße und deren mittlerem Inkrement wird durch die Strukturfunktionen hergestellt, siehe Gleichung (3.14). Deren physikalischer Inhalt ist, daß mittlere Geschwindigkeitsänderungen v_r in Wirbeln der Größe r sich wie $\epsilon r^{1/3}$ verhalten. Invertiert man diese Beziehung und setzt sie in den Ausdruck für $N(r)$ ein, so erhält man die gewünschte Abhängigkeit

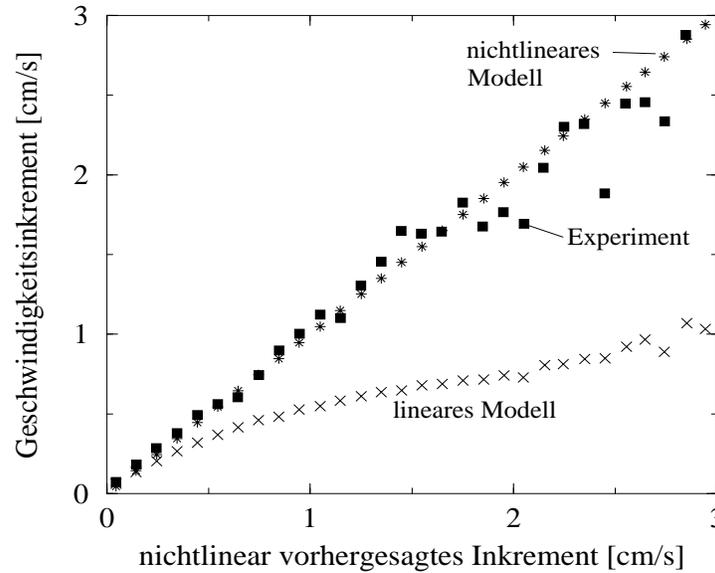


Abbildung 3.14: Die nichtlinear vorhergesagten Inkremente (Sterne) stimmen im Mittel gut mit den gemessenen Inkrementen (Quadrate), während das lineare Modell (x) die Fluktuationen systematisch unterschätzt (2D-Fluß mit $S = 20$ mm).

der Anzahl der Wirbel versus ihrem mittleren Inkrement v_r , diese solle skalieren wie $N(v_r) \propto v_r^{-3}$.

In Abb. 3.13 ist die Verteilung der nichtlinear vorhergesagten Inkremente für verschiedene Werte von τ und S gegen v_r in doppellogarithmischer Darstellung aufgetragen. Ebenfalls gezeigt sind Geraden, die das Kolmogorov-Verhalten $N(r) \propto v_r^{-3}$ repräsentieren. Die PDFs zeigen das erwartete Skalenverhalten für einen Bereich von etwa einer Dekade. Das Einsetzen des Potenzverhalten wird für höheren Vorhersagehorizont S zu größeren v_r verschoben. Das ist erwartet, da nur Wirbel durch das lokal konstante Modell aufgelöst werden können, die größer als die Summe aus der Länge des Einbettungsvektors und dem Vorhersagehorizont sind. Für die größten gezeigten Werte von S erreicht diese Summe etwa 16 cm, was etwa der Länge der größten kohärenten Wirbel in unserem System entspricht. Daher beobachtet man hier kein klares Skalenverhalten mehr.

3.4.3 Determinismus

In den vergangenen Abschnitten haben wir gesehen, daß das nichtlineare Modell für den 2D-Fluß mehr große Fluktuationen vorhersagt und daß die Verteilung der nichtlinear vorhergesagten Inkremente näher an der tatsächlichen Verteilung ist. Diese Statistik sagt uns allerdings noch nicht, ob die großen Inkremente wirklich in Situationen vorher-

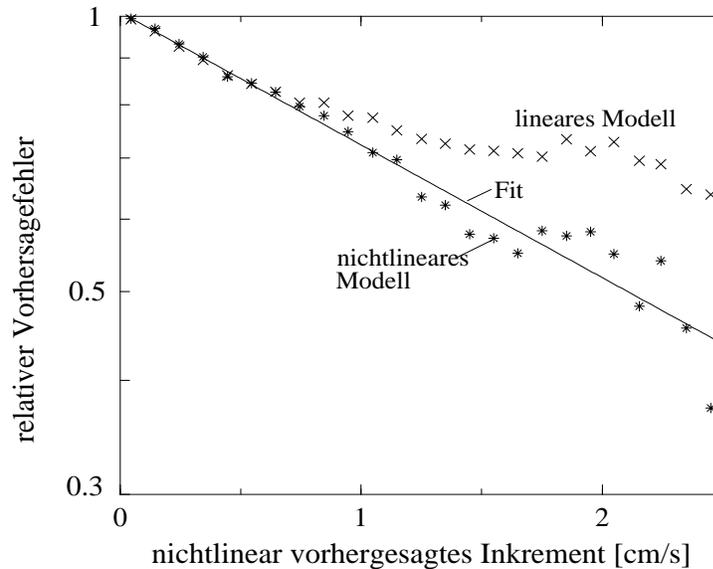


Abbildung 3.15: Relativer Vorhersagefehler des linearen und des nichtlinearen Modells versus der vorhergesagten Fluktuationen v_S (2D-Fluß mit $S = 20$ mm).

gesagt werden, wenn starke Fluktuationen tatsächlich auftreten. Die Frage lautet: Sind die großen nichtlinear vorhergesagten Inkremente zeitlich korreliert mit solchen im Signal, und führen diese zu verbesserten Vorhersagen im Vergleich zum linearen Modell? Welches der Modelle gibt die besseren Vorhersagen in dem Bereich, wo deren PDFs stark voneinander abweichen? In Abb. 3.14 sind die mittleren Inkremente in den Daten und durch das lineare Modell vorhergesagt versus dem nichtlinear vorhergesagten Inkrement gezeigt (*reliability function*). Während die gemessenen Inkremente im Mittel mit denen durch das nichtlineare Modell vorhergesagten übereinstimmen, unterschätzt das lineare Modell systematisch die auftretenden Fluktuationen. Nach Abschnitt 2.2.5 ist dies erneut ein Zeichen für nichtlineare Korrelationen in den Daten.

Dieses Verhalten ist in Abb. 3.15 in einer etwas anderen Form dargestellt. Dort werden die relativen mittleren Vorhersagefehler des AR-Modells ϵ_{lin}/v_S und des lokal konstanten Modells ϵ_{nlin}/v_S wieder gegen die vorhergesagten Inkremente \hat{v}_S^{nlin} des nichtlinearen Modells aufgetragen. Man erkennt die deutlich besseren Vorhersagen des lokal konstanten Modells im 2D-Fall. Im 3D-Fluß beobachtet man in entsprechenden Situationen das Gegenteil, dort ist das nichtlineare Modell im Mittel schlechter als das lineare, wenn es starke positive Fluktuationen vorhersagt.

Dies ist in Abb. 3.16 gezeigt, wo die relative Verbesserung $I = (\epsilon_l - \epsilon_{nlin})/v_r$ gegen das nichtlinear vorhergesagte Inkrement \hat{v}_S^{nlin} aufgetragen ist. Während die Verbesserung im 2D-Fluß positiv ist, erhält man für den 3D-Fluß eine Verschlechterung der Vorhersage mit dem lokalen Modell. Das lineare Modell ist hier aufgrund seiner höheren

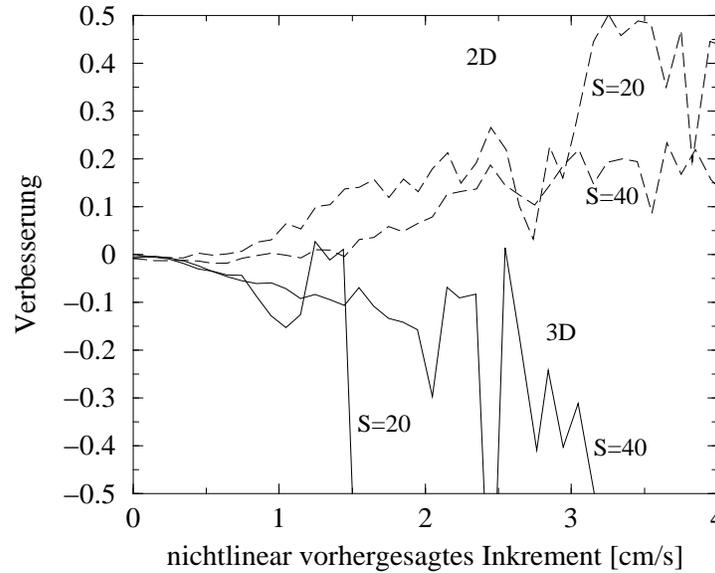


Abbildung 3.16: Relative Verbesserung I des nichtlinearen Modells gegenüber dem linearen Modell für 2D- (gestrichelt) und 3D-Turbulenz (durchgezogen) als Funktion des nichtlinear vorhergesagten Inkrements ($m = 10$; $\tau = 2, 4$ mm; $S = 20, 40$ mm).

statischen Robustheit überlegen.

Schließlich möchten wir noch die anfangs erwähnte Asymmetrie zwischen negativen und positiven Inkrementen für den 2D-Fall erklären. Weder das Skalenverhalten noch die deutliche Verbesserung der Vorhersage erhält man für negative Inkremente. Man kann dies folgendermaßen verstehen: Wenn ein Wirbel die Sonde passiert, beobachtet man zunächst am Übergang zwischen Jet und Wirbel ein negatives Inkrement. Dies ist jedoch nicht vorhersagbar, solange der Einbettungsvektor noch innerhalb des Jets liegt. Wenn der Wirbel später wieder durch den Jet abgelöst wird, erfolgt ein Wiederanstieg der Geschwindigkeit. Dieses positive Inkrement ist vorhersagbar, da nach dem Durchzug eines Wirbels immer der Wiederanstieg der Geschwindigkeit erfolgt.

3.4.4 Diskussion

Wir haben die Dynamik zweidimensionaler und dreidimensionaler Turbulenz mittels linearer und nichtlinearer Vorhersagen der Geschwindigkeitszeitreihen analysiert. Für 2D-Turbulenz fanden wir verbesserte Vorhersagen mit lokalen Phasenraummodellen verglichen zu konventionellen linearen Verfahren (AR-Modlle), wenn die vorhergesagten Fluktuationen positiv und groß waren. Die Wahrscheinlichkeitsverteilung der nichtlinear vorhergesagten Inkremente \hat{v}_r^{nlin} skaliert wie $(\hat{v}_r^{nlin})^{-3}$, wie durch die Kolmogorov-Theorie vorhergesagt. Im 3D-Fluß konnte keine Verbesserung des nichtlinearen Modells

gefunden werden. Eine intuitive Erklärung für den Unterschied zwischen dem betrachteten 2D- und 3D-Fluß ist das Fehlen von Wirbeldehnung im 2D-Fall. Die Dehnung bewirkt eine wesentlich kompliziertere Dynamik der Wirbel in 3D-Turbulenz.

3.5 Fokker-Planck-Methode

Wir wollen nun die in Abschnitt 2.2.1 eingeführten Methoden nutzen, um das Auftreten kohärenter Strukturen in 2-dimensionaler Turbulenz aus einer anderen Perspektive zu betrachten. Anstatt die Dynamik des Systems in der Zeit t zu untersuchen, werden wir das Verhalten der Geschwindigkeitsinkremente in Bezug auf die räumliche Skala r analysieren, um so Einblick in die Energiekaskaden und den Prozeß des Transfers von Skala zu Skala zu bekommen. Dies bedeutet, wir werden nicht wie im vorangegangenen Abschnitt im Sinne der bedingten Wahrscheinlichkeit $p(u_{t_2}, t_2 | \vec{u}_{t_1}, t_1)$ die Geschwindigkeit zur Zeit t_2 von der Geschwindigkeit zur Zeit t_1 schätzen, sondern die Evolution der Geschwindigkeitsinkremente in Abhängigkeit des räumlichen Abstands r untersuchen. Dies bedeutet, wir betrachten longitudinale Geschwindigkeitsinkremente $v_r(x) := u(x-r) - u(x)$ auf einer räumlichen Skala r und analysieren Ausdrücke der Form $P(v_{r_1}, r_1 | v_{r_2}, r_2)$. Die räumlichen Daten u_r werden wieder mittels der Taylor-Hypothese aus den gemessenen Zeitreihen “erzeugt”. Eine verbreitete Methode, die Komplexität voll entwickelter Turbulenz zu untersuchen, besteht darin, das skalenabhängige Verhalten der Wahrscheinlichkeitsverteilung (pdf) $P_r(v_r)$ entweder direkt oder im Sinne der Momente $\langle v_r^n \rangle = \int v_r^n P(v_r) dv_r$ zu untersuchen. Wie bereits ausgeführt, werden die skalenabhängigen Momente Strukturfunktionen genannt. In verschiedenen Arbeiten von Friedrich und Peinke wurde kürzlich gezeigt, daß die r -Entwicklung der Verteilungen in voll entwickelter Turbulenz als Markov-Prozeß betrachtet werden kann [34, 35, 36, 37]. Die Markov-Eigenschaft wird ausgewertet durch die Untersuchung bedingter Wahrscheinlichkeiten. Betrachtet man Geschwindigkeitsinkremente $v_{r_1}, v_{r_2}, v_{r_3}$ mit $r_1 > r_2 > r_3$, so sollte gelten:

$$P(v_{r_3}, r_3 | v_{r_2}, r_2; v_{r_1}, r_1) = P(v_{r_3}, r_3 | v_{r_2}, r_2).$$

Neben der Untersuchung der Markov-Eigenschaft können wie in Abschnitt 2.2.1 entsprechende Fokker-Planck-Gleichungen aus den Daten geschätzt werden. Als wesentliche notwendige Bedingung für die Gültigkeit der Markov-Eigenschaft ist die Chapman-Kolmogorov-Gleichung zu überprüfen:

$$P(v_{r_2}, r_2 | v_{r_1}, r_1) = \int_{-\infty}^{\infty} P(v_{r_2}, r_2 | v_{r'}, r') \times P(v_{r'}, r' | v_{r_1}, r_1) dv_{r'}.$$

Dies ist in Abbildung 3.17 dargestellt, wo die rechte und die linke Seite der Chapman-Kolmogorov-Gleichung für verschiedene Werte von v_{r_1} und feste r_1, r_2, r' gezeigt werden. Die Radien liegen innerhalb des Inertialbereichs und innerhalb des Bereichs,

für welchen kohärente Strukturen erwartet werden. Man erkennt, daß die Chapman-Kolmogorov-Gleichung im wesentlichen erfüllt ist. Die Abweichungen von der Gültigkeit der Chapman-Kolmogorov-Gleichung für 2D- und 3D-Turbulenz sind von derselben Größenordnung, und systematische Unterschiede sind zwischen den beiden Flußtypen nicht erkennbar.

Wie in Abschnitt 2.2.1 wird eine Fokker-Planck-Gleichung für die Dynamik von Skala zu Skala aufgestellt

$$-\frac{\partial}{\partial r}P(v_r, r) = \left(-\frac{\partial}{\partial v_r}D^{(1)}(v_r, r) + \frac{\partial^2}{\partial v_r^2}D^{(2)}(v_r, r) \right) P(v_r, r), \quad (3.22)$$

wobei die Koeffizienten wie in Gleichung (2.25) definiert sind durch:

$$\begin{aligned} D^{(k)}(v_r, r) &:= \lim_{\delta \rightarrow 0} M^{(k)}(v_r, r, \delta) \\ M^{(k)}(v_r, r, \delta) &:= \frac{1}{k!} \frac{1}{\delta} \int (v_{r'} - v_r)^k P(v_{r'}, r' | v_r, r) dv_{r'}, \end{aligned} \quad (3.23)$$

mit $\delta = r - r'$. Wir werden diesem Verfahren folgen sowie zum Vergleich die durch Gleichung (2.35) motivierten Korrekturen für den Diffusionsterm berücksichtigen. In der Sprache der Geschwindigkeitsinkremente schreibt sich der korrigierte Diffusionsterm als:

$$D^{(2)} = \frac{\frac{1}{2\delta} \left(\langle (v_{r+\delta} - v_r)(v_{r+\delta} - v_r) \rangle_v - \delta^2 D^{(1)}(v) D^{(1)}(v) \right)}{\left(1 + \delta \frac{dD^{(1)}(v)}{dv} \right)}. \quad (3.24)$$

Die aus den Daten der 2- und 3-dimensionalen Turbulenz gefitteten Drift- und Diffusionskoeffizienten sind in Abbildung 3.18 als Funktion von v_r dargestellt. Der Driftkoeffizient ist im Wesentlichen eine lineare Funktion von v_r mit schwachen Nichtlinearitäten an den Rändern. Dieses Verhalten wurde auch von Friedrich *et al.* [34] für voll entwickelte 3D-Turbulenz gefunden. Dieser Koeffizient, der den deterministischen Teil der Dynamik beschreibt, unterscheidet sich für den 2D- und den 3D-Fall nur unwesentlich. Ein anderes Verhalten ergibt sich für den Diffusionsterm. Für diesen findet man einen deutlichen Unterschied zwischen den beiden Flüssen. Man erkennt, daß dieser Koeffizient im 2D-Fall deutlich kleiner ist und auch einen geringeren Anstieg aufweist. Es ist sowohl der naive Schätzer für $D^{(2)}$ als auch der korrigierte Verlauf von $D^{(2)}$ unter Nutzung von Gleichung (3.24) dargestellt. Man erkennt sehr deutliche Korrekturen, der Anstieg der Diffusion wird mit dem einfachen Verfahren von Siegert *et al.* [5] deutlich überschätzt. Nimmt man ein quadratisches Verhalten des Diffusionsterms $D^{(2)}(v_r) = \alpha + \beta v_r + \gamma v_r^2$ an, wie in Renner *et al.* [37] getan, bekommt man mit Gleichung (3.24) wesentliche Korrekturen für β und γ . Die Überschätzung dieser Parameter bei Nutzung des unkorrigierten Verfahrens wurde in [37] auch bereits festgestellt, jedoch konnte dort die Ursache nicht korrekt identifiziert werden. Die Tatsache, daß $D^{(2)}$ im 2D-Fall deutlich kleiner ist als in 3D-Turbulenz weist erneut auf die Tatsache hin, daß das Verhalten des zweidimensionalen Flusses stärker deterministisch ist.

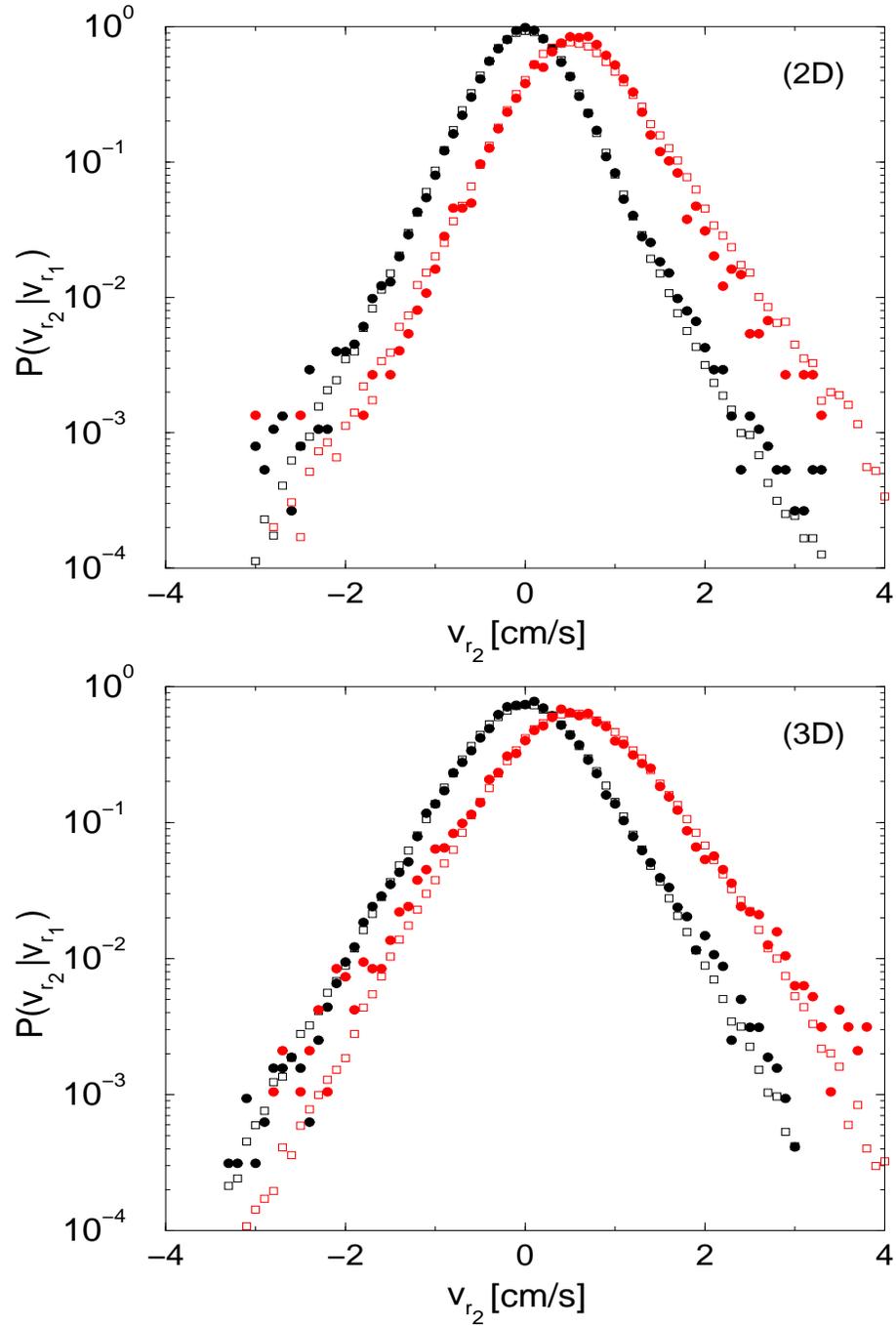


Abbildung 3.17: Überprüfung der Chapman-Kolmogorov-Gleichung (linke Seite gefüllt; rechte Seite offen) im Inertialbereich mit $r_1 = 8\text{cm}$, $r_{r'} = 6\text{cm}$, $r_2 = 2,5\text{cm}$ für zwei verschiedene Werte von $v_{r_1} = 0\text{cm/s}$ (schwarz) und $v_{r_1} = 2\text{cm/s}$ (rot) für 2D- und 3D-Turbulenz.

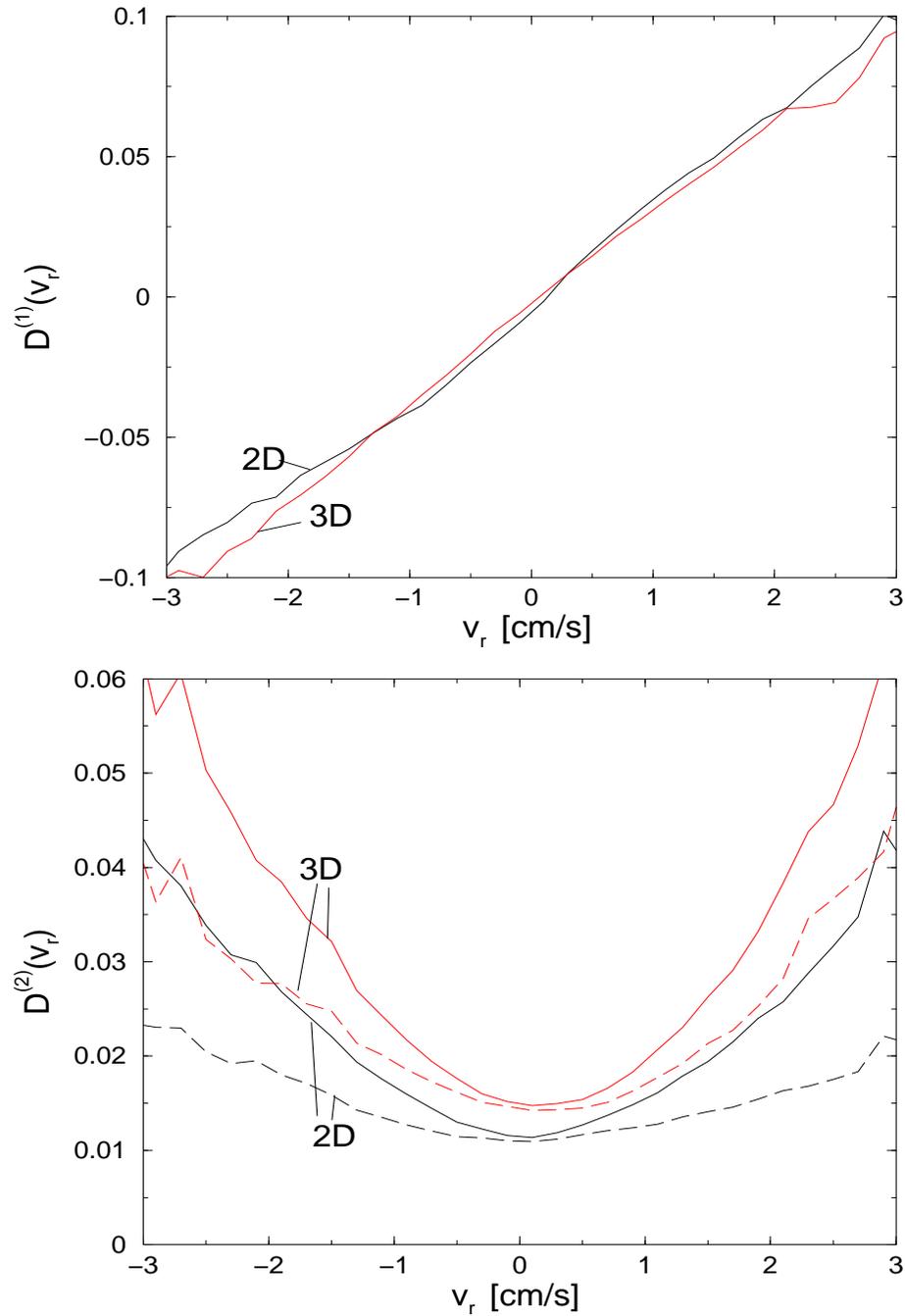


Abbildung 3.18: Drift- und Diffusionskoeffizienten für 2D- (schwarz) und 3D-Turbulenz (rot) für $r = 4\text{cm}$, für $D^{(2)}$ naive Methode in Gleichung (3.23) (durchgezogen) und korrekte Methode Gleichung (3.24) (gestrichelt).

Kapitel 4

Vorhersagen des Oberflächenwindes

In diesem Kapitel wollen wir die in den vorangegangenen Abschnitten entwickelten Verfahren auf die Vorhersage von atmosphärischen Bodenwindgeschwindigkeiten anwenden. Wir werden dabei besonders die technischen Anwendungen von Vorhersagen bei der “Windenergieproduktion” mittels Windturbinen betrachten. Daher nutzen wir neben den Winddaten auch Zeitreihen der von einer Windturbine produzierten elektrischen Leistung. Vorhersagen sind für die Konvertierung von Windenergie auf verschiedenen Zeitskalen von Bedeutung. Im Sekundenbereich lassen sich Vorhersagen für die direkte Steuerung der Rotorblätter einer Turbine nutzen. Im Minuten- bis Stundenbereich können Prognosen zur Steuerung konventioneller Energieerzeuger dienen. Auf noch längeren Vorhersagehorizonten sind Vorhersagen des Windes für die Kraftwerkeinsatzplanung nützlich. In diesem Kapitel werden wir uns auf die erste der genannten Anwendungen konzentrieren. Starke Änderungen der Windgeschwindigkeit im Sekundenbereich stellen erhebliche Schwierigkeiten bei der Windenergieproduktion dar. Im wesentlichen lassen sich drei Problemfelder nennen. Zunächst gefährden Windböen die mechanische Stabilität der Rotoren und können zu einer erheblichen Reduktion der Lebenserwartung der Turbinen, insbesondere in Mittelgebirgslagen, führen. Sind die Anlagen in ein Stromversorgungsnetz integriert, können die Böen zu starken Fluktuationen der in das Netz eingespeisten Leistung führen. Die Verminderung der Netzqualität ist die Folge. Dezentral genutzte isolierte Windkraftanlagen werden durch Dieselaggregate ergänzt, um bei plötzlicher Flaute die elektrische Leistung konventionell produzieren zu können. In dieser Situation ist Versorgungsunsicherheit die Folge der Fluktuationen, da die Dieselmotoren nicht hinreichend schnell reagieren können.

Im Prinzip können Windturbinen Fluktuationen des Windes bis zu einem bestimmten Grad durch die Drehung der Rotorflügel ausgleichen. Bei wachsender Windgeschwindigkeit werden die Blätter einfach um einige Grad gegen die Hauptwindrichtung geneigt. Leider kommt diese Justierung immer einen Moment zu spät. Um rechtzeitig reagieren zu können, wäre es daher wünschenswert, von der bevorstehenden Bö Sekunden im voraus zu wissen.

Wir wollen mit unseren Analysen in diesem Kapitel so nah wie möglich an einer realistischen technischen Aufgabenstellung arbeiten. Wir verwenden daher Zeitreihen der gemessenen elektrischen Leistung, einerseits weil die Leistung die räumlich gemittelte Windgeschwindigkeit direkt reflektiert und andererseits weil sie die übliche Steuerungsgröße von Windkraftanlagen ist. Neben der Leistungsmessung sind Windkraftanlagen typischerweise mit Schalenkreuzanemometern ausgestattet, welche Zeitreihen der Windgeschwindigkeit aufzeichnen. Diese Messungen werden üblicherweise zur Abschaltung der Turbinen genutzt, sobald z.B. der 10-Sekunden-Mittelwert der Windgeschwindigkeit eine obere Schranke überschreitet.

Im Sinne der vorangegangenen Kapitel werden wir zunächst die Windgeschwindigkeit auf nichtlineare Struktur untersuchen und dann den Fokus auf die Analyse der Leistungsdaten legen. Vorher werden wir einen kurzen Überblick über verwendete Methoden der Windvorhersage geben und die wesentlichen Eigenschaften der verwendeten Zeitreihen nennen.

4.1 Möglichkeiten der Vorhersage

Generell lassen sich bei der Vorhersage von Windgeschwindigkeiten meteorologische und datengetriebene Verfahren unterscheiden. Die meteorologischen Methoden beruhen auf globalen Vorhersagen des geostrophischen Windes durch die Wetterämter und dem anschließenden Herunterskalieren dieser Windgeschwindigkeiten mittels des logarithmischen Windprofils zur Bestimmung des Oberflächenwindes. Lokale Modelle dienen danach dazu, die jeweiligen Geländebeschaffenheiten in die Vorhersage zu integrieren. Obwohl diese Modelle gute Langzeitvorhersagen im Bereich von einigen Tagen liefern, sind sie für die Kurzzeitprognose recht ungeeignet. Bis zu etwa drei Stunden Prognosezeitraum ist sogar das einfache Modell der Persistenz (man verwendet den letzten Meßwert als Vorhersage) diesen aufwendigen Verfahren überlegen [38].

Verschiedene lineare Verfahren wurden in der Vergangenheit genutzt, um Windzeitreihen vorherzusagen. Dies waren insbesondere Autoregressivmodelle [39, 40] und Kalman-Filter [41], die den Wind als stochastisches Phänomen betrachten und lineare Korrelationen berücksichtigen. Der wesentliche Unterschied des Kalman-Filters zu herkömmlichen AR -Modellen besteht in den veröffentlichten Anwendungen auf Winddaten darin, daß die Parameter des linearen Modells von der Zeit abhängen. Dies wird in den Anwendungen durch die Variation der Länge des Gedächtnisses erreicht, auf welchem die Parameter des AR -Modells bestimmt werden. Dadurch wird der Nichtstationarität der Zeitreihen in begrenztem Maße Rechnung getragen. Man bezeichnet dieses Verfahren daher auch als adaptives AR -Modell. Wir werden diese Technik bei der Analyse der verwendeten Zeitreihen nutzen, wobei die Länge des Gedächtnisses typischerweise einige Stunden betragen wird. In Abb. 4.1 ist für eine Windzeitreihe die Abhängigkeit des Vorhersagefehlers von der Länge des Gedächtnisses gezeigt.

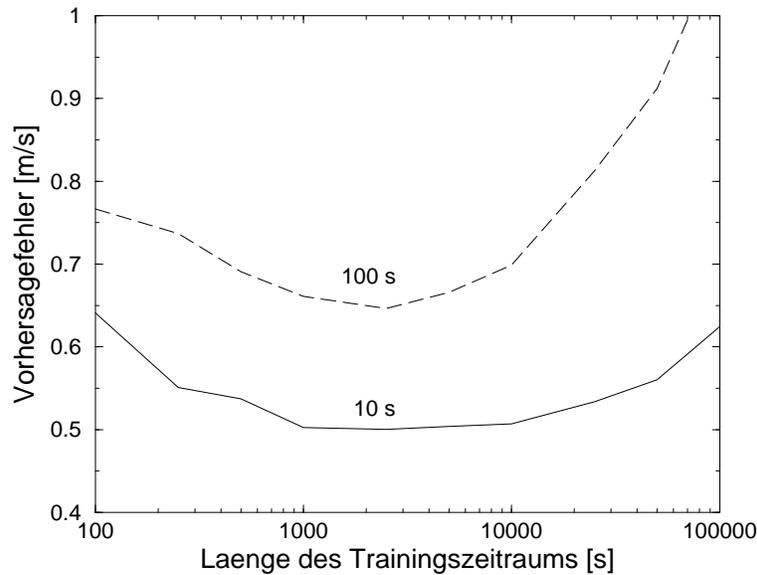


Abbildung 4.1: Vorhersagefehler in [m/s] für zwei verschiedene Werte des Vorhersagehorizonts in Abhängigkeit des Gedächtnisses, auf welchem das *AR*-Modell trainiert wird. Die Ordnung des Modells war $m = 5$ und die Zeitverzögerung $\tau = 2s$ für den Vorhersagehorizont von $10s$ bzw. $\tau = 20s$ anderenfalls.

In den drei zitierten Arbeiten [39, 40, 41] wurden Vorhersagen mit Horizonten im Minutenbereich durchgeführt. Die Ergebnisse zeigen übereinstimmend, daß die Verbesserungen der linearen Verfahren gegenüber der Persistenz im Bereich von etwa 10% des Fehlers der Persistenz liegen. Es zeigte sich jedoch in [41], daß die Vorhersagbarkeit stark von den orographischen Beschaffenheiten am Vorhersageort abhängen kann. Eine Abhängigkeit der Ergebnisse von der mittleren Windgeschwindigkeit konnte indessen nicht festgestellt werden. Schließlich sei noch auf eine Arbeit von Hausmann und Hoffmann [42] eingegangen, in der neuronale Netze zur ein- bzw. zehnmütigen Vorhersage genutzt werden. Bei Feed-Forward-Netzen wird mit einer Eingabedimension von 30 bzw. 60 (für verschiedene Datensätze) eine Verbesserung gegenüber der Persistenz von etwa 10% erreicht. Generell war die Abhängigkeit der Resultate von den verwendeten Netzparametern recht gering. Einzig die Variation der Eingabedimension bewirkte deutliche Unterschiede. Daher wurde für deren Arbeit jeweils ein Perzeptron mit variabler Ausgangsamplitude verwendet. Die Ergebnisse der Arbeit zeigten keine systematischen Verbesserungen gegenüber *AR*-Modellen.

Erwähnenswert für die Vorhersagbarkeit von Wetterdaten ist die gegen Ende der 80er Jahre vorgetragene Hypothese, es könne sich bei Wetter- und Klimaphänomenen um dynamische Systeme mit einem niedrigdimensionalen Attraktor handeln. Dabei wurde mittels des Grassberger-Procaccia-Algorithmus [43] die Anzahl der aktiven Frei-

Autoren	Art der Daten	Daten- menge	geschätztes D_2
Nicolis & Nicolis (1984)	$\delta^{18}\text{O}$ Isotop	500	3.1
Fraedrich (1986)	lokaler Druck	5500	3-4
	relative Sonnenscheindauer	11000	3-4
	$\delta^{18}\text{O}$ Isotop	200	3-4
Tsonis & Elsner (1989)	vertikale Windgeschwindigkeit	4000	7.3
Maasch (1989)	$\delta^{18}\text{O}$ Isotop	200	4-6
Sharifi <i>et al.</i> (1990)	Regenmenge	4000	4
Poveda-Jaramillo <i>et al.</i> (1993)	vertikale Windgeschwindigkeit	40000	4-7

Tabelle 4.1: Geschätzte Korrelationsdimensionen für Wetterdaten.

heitsgrade des Systems (in Form der Korrelationsdimension) bestimmt. Eine Zusammenstellung dieser Arbeiten [44, 45, 46, 47, 48, 49] mit den verwendeten Datensätzen und erzielten Resultaten ist in Tabelle 4.1 gezeigt. Diese Ergebnisse wurden jedoch bald aufgrund ihrer geringen Datenbasis angezweifelt [50]. Später wurden grundlegende Voraussetzungen bezüglich der notwendigen Datenmenge für die Anwendung des Grassberger-Procaccia-Algorithmus formuliert [51], welche von den in der Tabelle 4.1 angegebenen Arbeiten nicht erfüllt werden (alle verwendeten Zeitreihen sind zu kurz). Somit konnte die Hypothese eines niedrigdimensionalen Attraktors für Wetter- bzw. Klimaphänomene nicht bestätigt werden. Insbesondere sind mit den in dieser Arbeit verwendeten Datensätzen, siehe Abschnitt 4.2, bei korrekter Anwendung des Grassberger-Procaccia-Algorithmus keine endlichen Dimensionen identifizierbar.

4.2 Zeitreihen und lineare Statistik

Für die durchzuführende Arbeit stehen Datensätze der Windgeschwindigkeit gemessen an Standorten in Dänemark und Spanien zur Verfügung. Die Messungen in Dänemark wurden von Juli 1987 bis Juni 1988 im Lammefjord auf der Insel Seeland vorgenommen. Das Gelände um die Meßmasten ist extrem flach, da es sich um einen zurückgewonnenen Fjord handelt. Es befinden sich keine Gebäude im Umkreis von einem Kilometer um den Meßstandort. Gemessen wurde an vier unterschiedlichen Meßmasten deren Höhe 10, 30 und 45 m beträgt und die sich auf einem Gelände von 400 m^2 befinden, wie in Abb. 4.2 gezeigt. Dabei sind die drei höheren Türme derart angeordnet, daß sie die fiktive Ebene des Rotors einer Windturbine orthogonal zur mittleren Windrichtung (Süd-West) aufspannen. Der kleinere Mast ist rechtwinklig zu dieser Ebene und 15 Meter von dieser entfernt in die Richtung der mittleren Windrichtung angebracht. Dies wird später bei den multivariaten Analysen wichtig sein, wenn wir diesen vorgerückten Meßstandort nutzen werden, um Böen auf der "Rotorebene" vorherzusagen. Zur

Messung wurden Schalenkreuz- bzw. Ultraschallanemometer verwendet, für welche die Abtastfrequenz 8 bzw. 16 Hz beträgt. Insgesamt gibt es an den vier Masten zehn Meßorte. Typische Werte sind Windgeschwindigkeiten zwischen $3m/s$ und $15m/s$, deren Standardabweichungen zwischen $1m/s$ und $3m/s$ liegen.

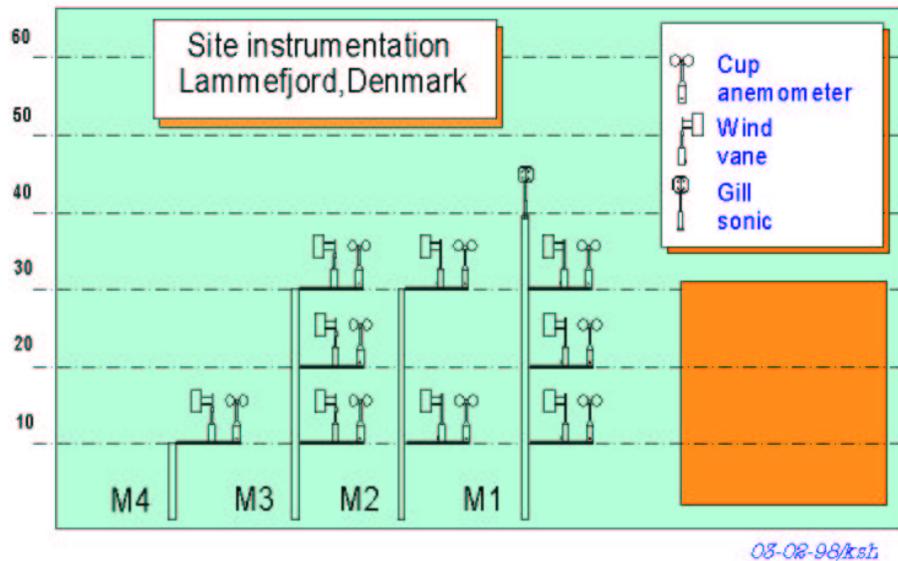


Abbildung 4.2: Meßanordnung am Standort Lammefjord.

Die Messungen in Spanien wurden in einer Mittelgebirgslage durchgeführt. Dabei wurde eine Windturbine insbesondere auf deren Belastungen unter stark turbulenten Windverhältnissen vermessen. So wurden neben der Windgeschwindigkeit $u(t)$ die von der Turbine erzeugte elektrische Leistung $P(t)$ und die Dehnungsmomente $M_D(t)$ an den Rotorblättern aufgezeichnet. Die Windgeschwindigkeit wurde in der Richtung orthogonal zur Rotorebene auf der Höhe der Rotornabe von 40 m gemessen. Die Abtastrate betrug für alle gemessenen Größen 36 Hz . Bei der Turbine handelt es sich um das Modell G74 des Herstellers Gamesa mit einem Rotordurchmesser von 47 m , welche ab einer Geschwindigkeit von 3.0 m/s und bis zu 25.0 m/s elektrische Leistung produziert, die bei 660 kW maximal ist. In Abb. 4.3 zeigen wir typische Zeitreihen der Windgeschwindigkeit und der Leistung.

Die Häufigkeitsverteilung des Betrags der Windgeschwindigkeit ist in Abb. 4.4 gezeigt. Man erkennt, daß die Verteilung asymmetrisch ist. Verschiedene nichtgaußsche Verteilungen wurden vorgeschlagen, um Windgeschwindigkeiten zu modellieren. Die in der

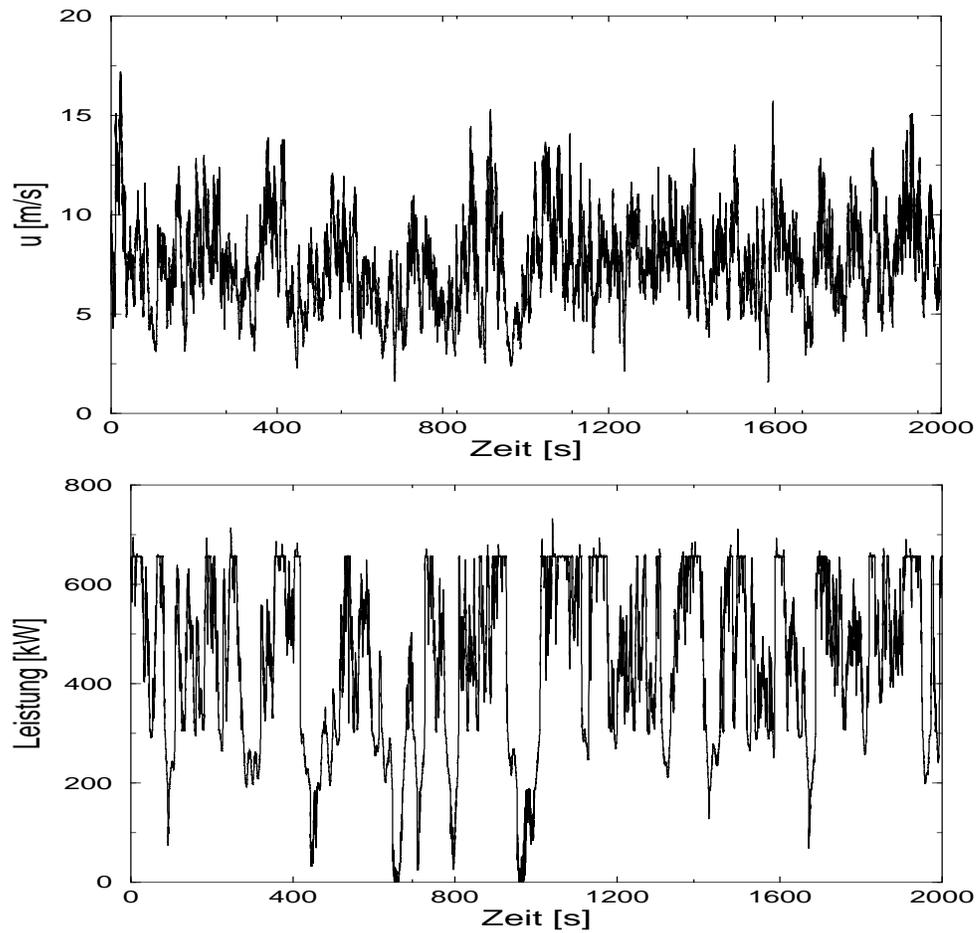


Abbildung 4.3: Typische Zeitreihen der Windgeschwindigkeit und der elektrischen Leistung.

Literatur am häufigsten verwendete Verteilung ist die sogenannte Weibull-Verteilung:

$$f(u) = \frac{g}{b} \left(\frac{u}{b}\right)^{g-1} \exp[-(u/b)^g], \quad (4.1)$$

wobei g und b Form- und Skalenparameter genannt werden. Es wurde eine solche Verteilung an die Daten angepaßt, wie ebenfalls in Abb. 4.4 dargestellt ist. Wir wollen hier nicht weiter auf die Details dieser Verteilung eingehen, da dies für die weiteren Abschnitte nicht von Bedeutung ist, sondern verweisen für eine umfassende Diskussion auf Brown *et al.* [39].

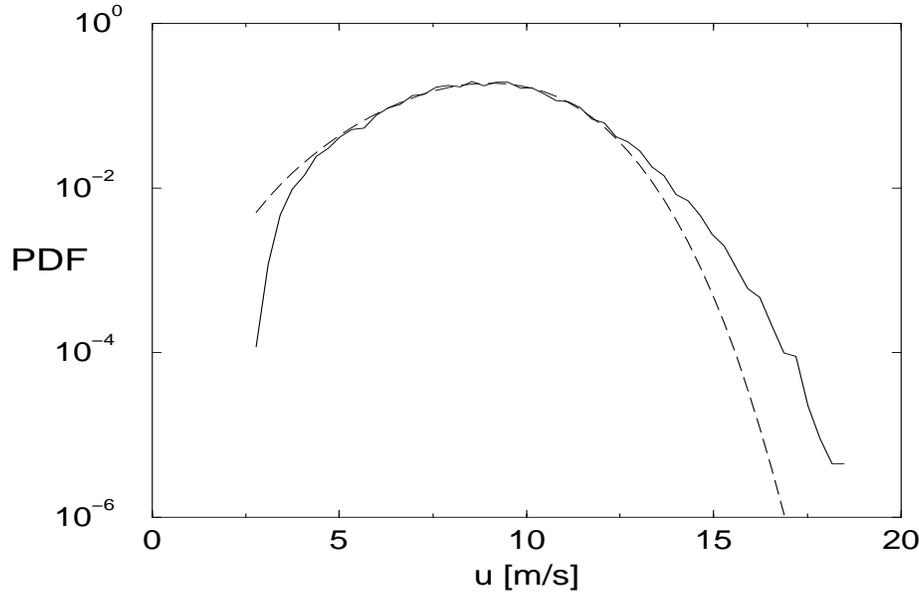


Abbildung 4.4: Häufigkeitsverteilung des Betrags der Windgeschwindigkeit (durchgezogen) und angepaßte Weibull-Verteilung (gestrichelt) in Gleichung (4.1) mit $b = 9.5$ und $g = 4.7$.

Wir wollen nun noch das Leistungsspektrum der Geschwindigkeitsdaten angeben. Dieses ist sowohl für die mit dem Ultraschall- als auch mit dem Schalenkreuzanemometer gemessenen Daten in Abb. 4.5 gezeigt. Man erkennt, daß beide Datensätze für Frequenzen kleiner als 1 Hz recht gut mit dem Kolmogorov-Gesetz ($k^{-5/3}$) übereinstimmen. Für höhere Frequenzen ergeben sich jedoch für die mit dem Schalenkreuzanemometer (8 Hz -Daten) gemessenen Zeitreihen deutliche Abweichungen. Dieses Meßgerät wirkt durch seine mechanische Trägheit wie ein Tiefpaßfilter und glättet somit die kurzwelligen Fluktuationen. Dadurch ähnelt das Spektrum in diesem Fall dem der 2-dimensionalen Turbulenz, vgl. Abschnitt 3.3. Wir werden später zeigen, daß für diese Daten auch wie im 2D-Fall nichtlineare deterministische Fluktuationen beobachtet werden können, obwohl atmosphärische Turbulenz sicherlich dreidimensional ist. Wir

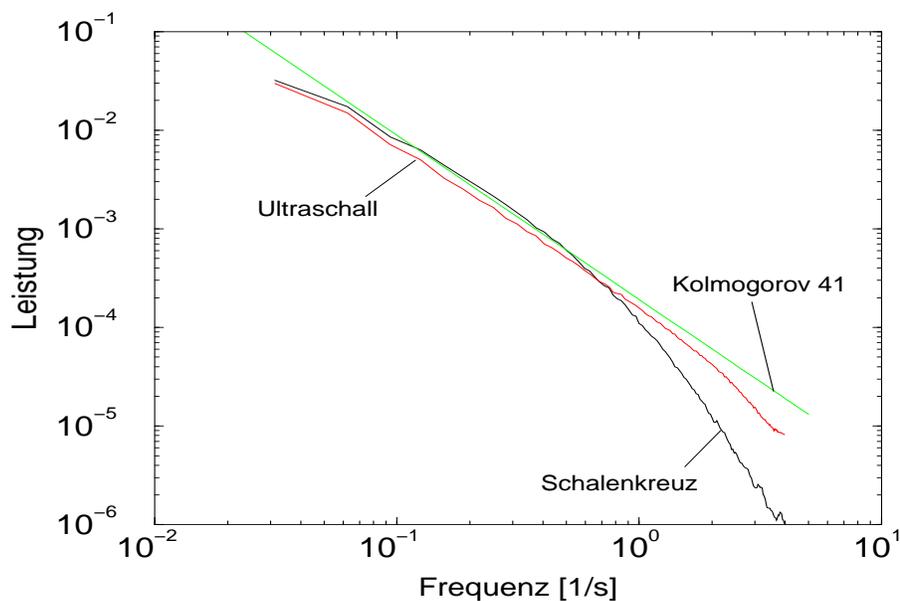


Abbildung 4.5: Leistungsspektrum der Windgeschwindigkeit für die mit dem Schalenkreuzanemometer und mit dem Ultraschallanemometer gemessenen Winddaten. Ebenfalls gezeigt ist das Kolmogorov 41 Skalenverhalten mit dem Exponenten $-5/3$.

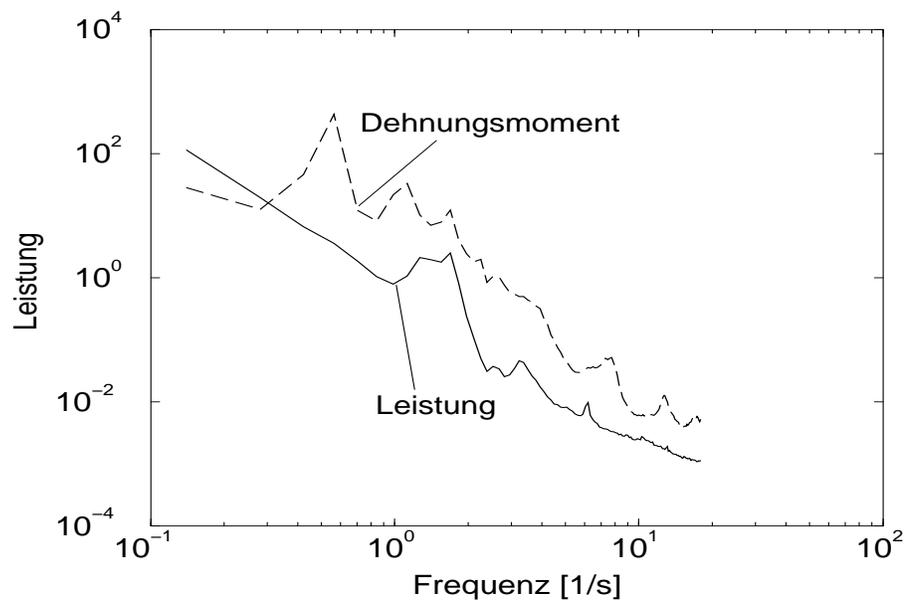


Abbildung 4.6: Leistungsspektrum des Signals der elektrischen Leistung sowie des Dehnungsmoments an den Rotoren.

werden daher für die folgenden Analysen im wesentlichen die 8 Hz -Daten nutzen, auch weil Windturbinen typischerweise mit Schalenkreuzanemometern ausgestattet sind (damit sind diese Daten praktisch relevanter). Außerdem stehen nur für die 8 Hz -Daten multivariate Zeitreihen zur Verfügung, um somit vektorwertige Analysen durchführen zu können.

In Abb. 4.6 ist das Leistungsspektrum für eine Zeitreihe des Dehnungsmoments an den Rotorblättern sowie für die Daten der elektrischen Leistung jeweils aus dem Spanien-Datensatz gezeigt. Es zeigen sich deutliche Peaks in den beiden Spektren, wobei das Hauptmaximum im Spektrum der Leistung bei etwa 1.5 Hz und im Spektrum des Dehnungsmoments bei etwa 0.5 Hz liegt, wobei letzterer Wert der Umlauffrequenz des Rotors entspricht. Einmal pro Umlauf erreicht die Belastung der Rotorblätter ein Maximum, dies ist immer dann, wenn das entsprechende Blatt senkrecht nach oben gerichtet ist. Entsprechend des logarithmischen Windprofils steigt die Windgeschwindigkeit u mit der Höhe h wie $u \propto \log h$. Das Spektrum der Leistung hat ein Maximum bei etwa dem dreifachen der Umlauffrequenz, da eine bestimmte Konfiguration der Rotorflügel einer maximalen Energieausbeute entspricht. Eine solche Konfiguration tritt aufgrund der Symmetrie des Rotors dreimal pro Umlauf auf.

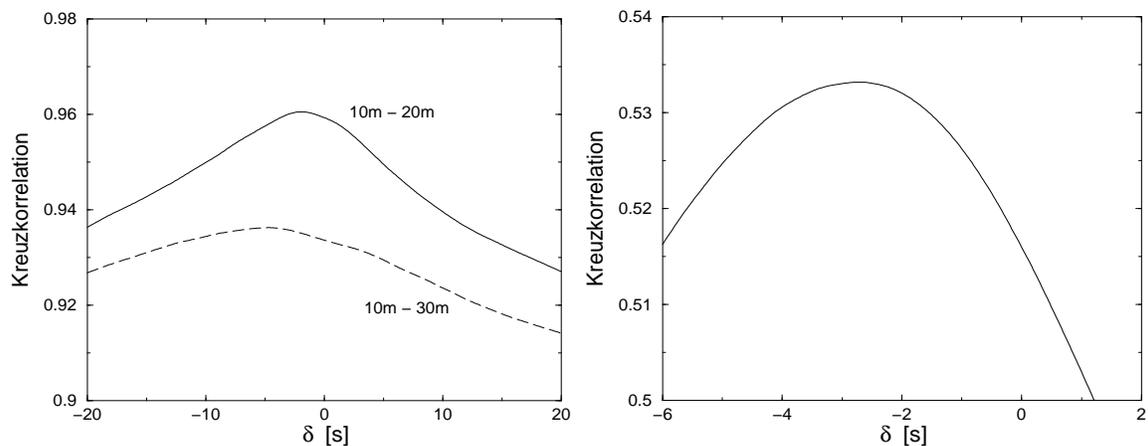


Abbildung 4.7: Lineare Kreuzkorrelation der Windgeschwindigkeit auf 10, 20 und 30 Metern über dem Boden (links) und zwischen Leistungssignal und Windgeschwindigkeit an einer Windturbine.

Die Reibung der Luft an der Grenzfläche Erde erzeugt einen Scherfluß, welcher kohärente Strukturen, sogenannte Streifen, enthält, die spitzwinklig zur Oberfläche in die Hauptflußrichtung zeigen [24]. Stellt man eine vertikale Fläche in den Fluß, so erreichen die Strukturen diese Fläche zunächst am obigen Ende. Somit treten gewisse Ereignisse, wie Böen, weiter entfernt von der Oberfläche früher auf als nahe am Boden. Dies wird

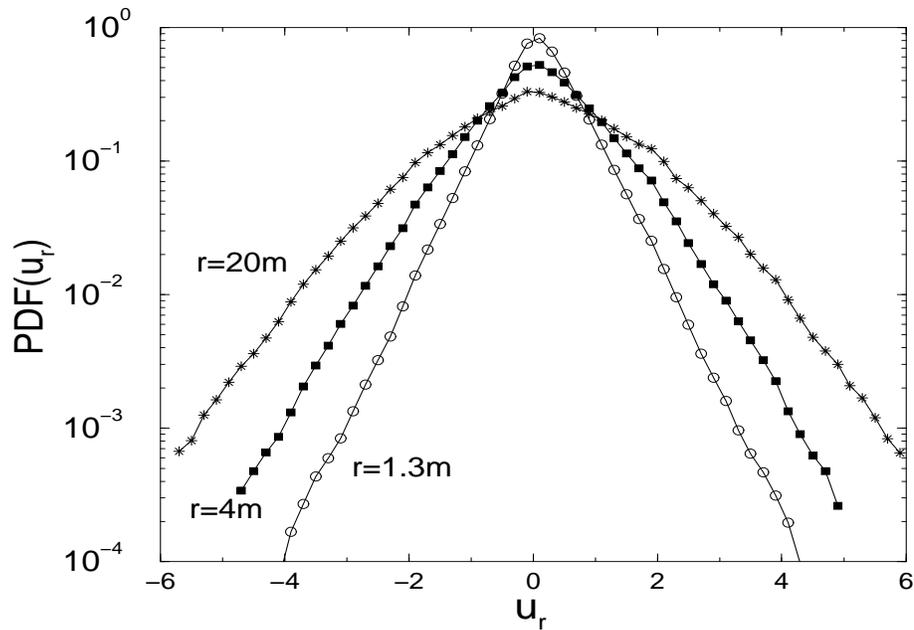


Abbildung 4.8: Wahrscheinlichkeitsverteilung der Leistungsinkremente auf unterschiedlichen Skalen 1.3... 20 m.

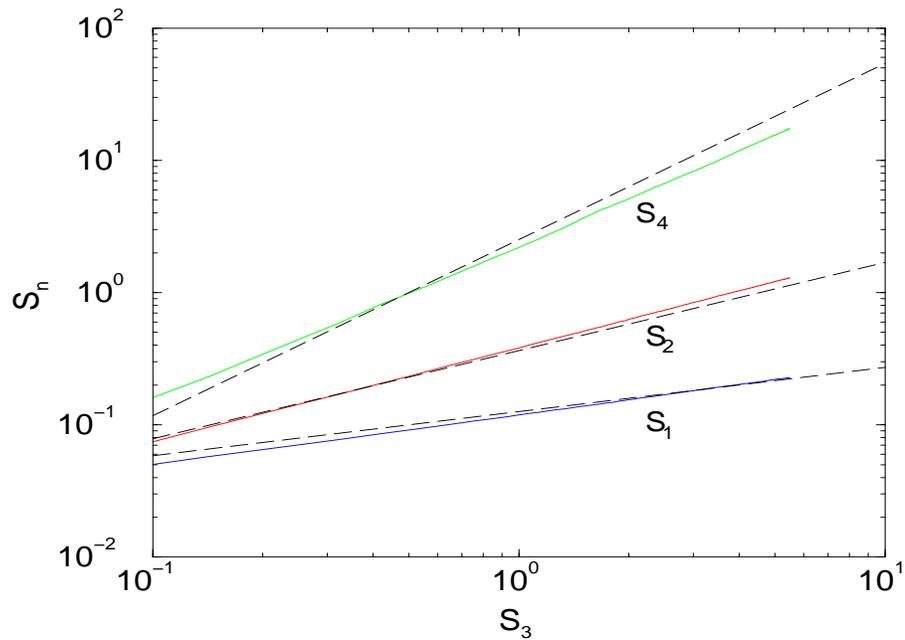


Abbildung 4.9: Strukturfunktionen S_1 , S_2 und S_4 nach der EES-Technik (farbige Linien) und das theoretische Verhalten nach Kolmogorov 41 (gestrichelt).

deutlich, wenn man die Kreuzkorrelationsfunktion $\langle u_t^{h_1} u_{t+d}^{h_2} \rangle$ der Geschwindigkeiten auf unterschiedlichen Höhen berechnet, wie in Abb. 4.7 mit $h_1 = 10 \text{ m}$ und $h_2 = 20, 30 \text{ m}$ gezeigt. Die Zeitreihen an höheren Meßpunkten sind denen an tieferen Meßpunkten voraus. Diese Tatsache hat Bedeutung für die Kreuzkorrelation von Windgeschwindigkeit und Leistung an einer Windturbine. Der Hauptbeitrag der produzierten elektrischen Leistung stammt von der Atmosphärenschicht über der Rotorachse, da dort die Windgeschwindigkeit höher ist. Die Windgeschwindigkeit hingegen wird auf der Höhe der Rotorachse gemessen. Somit ist die Leistung der Windgeschwindigkeit voraus, bei der typischen Anordnung des Anemometers an der Windturbine. Dies ist im rechten Bild von Abb. 4.7 gezeigt. Hier liegt das Maximum der Kreuzkorrelation bei etwa 3 Sekunden Zeitdifferenz. Dieses Verhalten ist wider der Intuition, da man naiv erwartet, daß die Leistung aufgrund der Trägheit des Rotors der Windgeschwindigkeit nachfolgt. Zum Ende dieses Abschnittes möchten wir uns wieder der Statistik der Geschwindigkeitsinkremente in voll entwickelter Turbulenz zuwenden. Wir berechnen die PDF der Inkremente und die Strukturfunktionen mittels EES wie schon in Abschnitt 3.3, wobei wir wieder die Taylor-Hypothese der eingefrorenen Turbulenz nutzen (Fluktuationen in der Zeit werden als solche im Raum interpretiert), siehe Abschnitt 3.2. Die PDFs zeigen wieder das für 3D-Turbulenz bekannte intermittente Verhalten. Auf kleinen Skalen beobachtet man exponentielle Verteilungen, welche zu größeren Skalen hin zunehmend gaußförmig werden. Auf Intermittenz deuten auch die Strukturfunktionen hin, welche in Abb. 4.9 dargestellt sind. Die beobachteten Skalenexponenten weichen, wie für 3D-Turbulenz erwartet, systematisch von dem durch die Kolmogorov-Theorie vorhergesagten Verhalten ab.

4.3 Nichtlineare Struktur in Bodenwindgeschwindigkeiten

Wie in den Abschnitten 2.2.5 und 3.4.3 möchten wir die lokal konstanten Vorhersageverfahren nutzen, um nun nichtlineare Strukturen in den Zeitreihen der Windgeschwindigkeit zu suchen und zu analysieren. Die ursprüngliche Hoffnung zu Beginn dieser Arbeit bestand darin, die Vorhersagen der Bodenwindgeschwindigkeiten **im Mittel** durch nichtlineare Methoden zu verbessern, d.h., der gemittelte Vorhersagefehler über z.B. alle 10-Sekunden-Vorhersagen eines ganzen Tages sollte durch die Anwendung nichtlinearer Verfahren gesenkt werden. Dieses Ziel konnte bei der skalaren Analyse nicht erreicht werden. Die im Abschnitt 4.1 zitierte Verbesserung linearer Verfahren gegenüber der Persistenz von etwa 10% konnte durch die Anwendung nichtlinearer Modelle nicht systematisch gesteigert werden. Daher beschlossen wir, die verschiedenen Modelle zeitaufgelöst zu analysieren, um somit Situationen zu identifizieren, in denen sich die Modelle in ihrer Vorhersagekraft unterscheiden. Dies ist in Abb. 4.10 gezeigt, wo für ein Intervall von 400 Sekunden die Fehler des lokal konstanten Modells

sowie des AR -Modells zeitaufgelöst dargestellt sind. Wir verwendeten als nichtlineares Modell wieder einen lokal konstanten Vorhersager mit einer Einbettungsdimension von $m = 10$, einer Zeitverzögerung von $\tau = 0.25$ s und einer Mindestzahl von 50 Nachbarn. Zunächst erkennt man, daß die Modelle im zeitlichen Mittel tatsächlich ähnlich genaue bzw. ungenaue Vorhersagen liefern. Bei sorgfältigerem Betrachten des zeitlichen Verlaufs der Fehler erkennt man jedoch, daß das nichtlineare Modell immer dann überlegen ist, wenn der Vorhersagefehler maximal wird. Für kleine Fehler sieht man dagegen, daß entweder beide Modelle etwa gleich sind oder daß das lineare Modell überlegen ist. Die maximalen Vorhersagefehler treten in Situationen auf, in denen sich das Signal stark ändert, also im Fall der Windgeschwindigkeit in Böensituationen. Diese Tatsache ist in Abb. 4.11 gezeigt, wo die Verbesserung des nichtlinearen Modells gegenüber dem linearen gegen die Änderung der Windgeschwindigkeit in der entsprechenden Situation aufgetragen ist. Man erkennt, daß in Situationen, in denen das Signal nahezu persistent ist, das lineare Modell überlegen ist. Dagegen beobachtet man in Böensituationen eine deutliche Verbesserung durch das nichtlineare Verfahren.

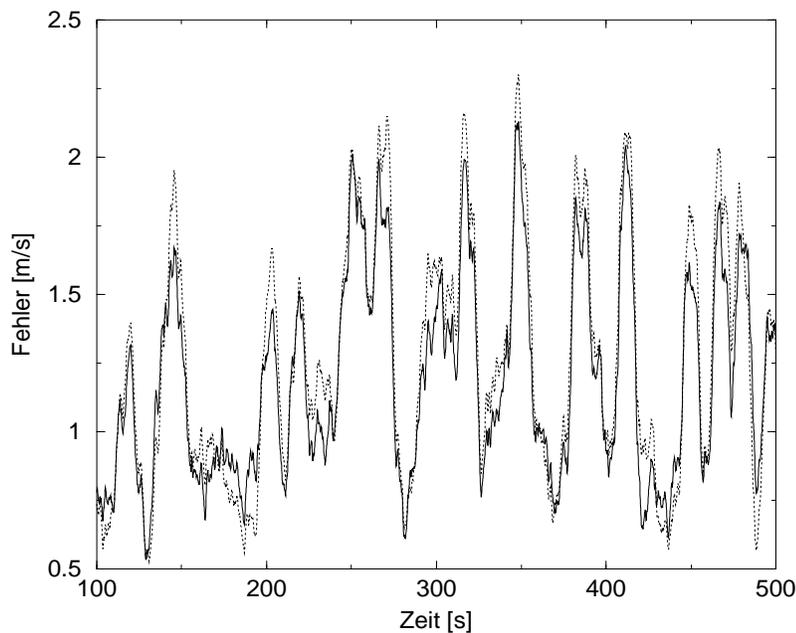


Abbildung 4.10: Zeitaufgelöster Vorhersagefehler von linearem (gepunktet) und nichtlinearem Modell (durchgezogen) für eine Windzeitreihe (gemessen mit einem Schalenkreuzanemometer). Das nichtlineare Modell ist im Bereich der Maxima des Fehlers überlegen.

Die Wahl des Modells nach dem jeweiligen Inkrement in der Zeitreihe entspricht einer Diskriminierung im Phasenraum, d.h. einer Einteilung in Regionen, in denen eines der jeweiligen Modelle angebrachter ist. Andere Ansätze, den Phasenraum nach ei-

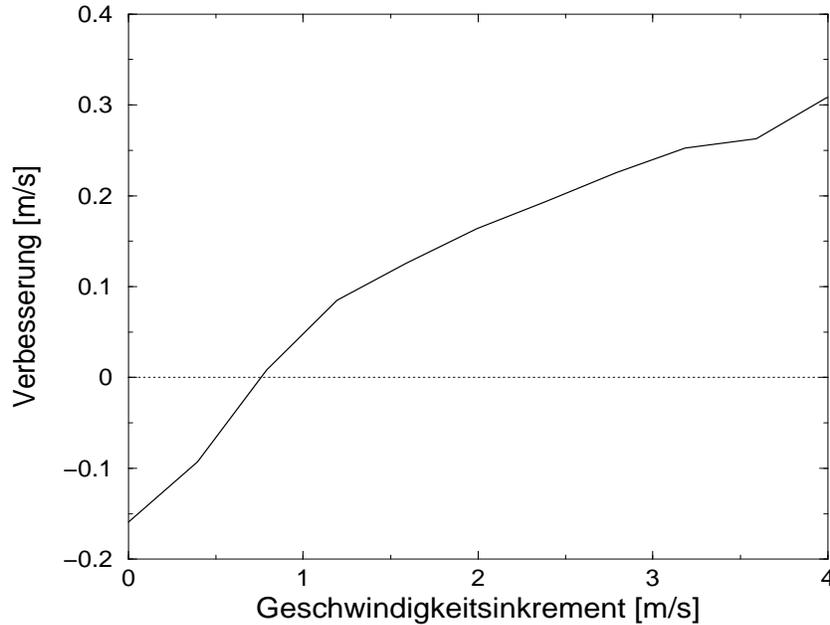


Abbildung 4.11: Verbesserung des nichtlinearen Modells gegenüber dem linearen als Funktion des Betrags der Änderung des Signals im Vorhersageintervall.

nem solchen Kriterium zu unterteilen, haben sich nicht als erfolgreich herausgestellt. So haben wir die Varianz des Einbettungsvektors oder die Varianz der Verteilung der Vorhersagen des lokal konstanten Modells als Kriterium für die Wahl des Modells in verschiedenen Situationen genutzt. Diese Kriterien haben sich aber nicht als nützlich erwiesen, d.h., unabhängig von der Varianz des Einbettungsvektors und unabhängig von der Verteilung der Vorhersagen des lokal konstanten Modells war die Vorhersagekraft beider Modelle in etwa gleich. Ebenfalls haben wir die einzelnen Felder in der Partition des Phasenraums nach dem Kriterium gekennzeichnet, ob das lineare oder das nichtlineare Modell überlegen ist. Auf unserer Testzeitreihe haben wir dann bei Vorhersagen zwischen beiden Modellen hin und her geschaltet, abhängig davon, in welchem Feld des Phasenraums sich der Einbettungsvektor gerade befand. Auch dieses Vorgehen führte jedoch nicht zu einer systematischen Steigerung der mittleren Vorhersagbarkeit. Daher kehren wir zurück zur Betrachtung der Geschwindigkeitsinkremente. Die Nutzung des zukünftigen Geschwindigkeitsinkrements zur Diskriminierung der Modelle hat einen entscheidenden Nachteil: man kennt das zukünftige Geschwindigkeitsinkrement zum Zeitpunkt der Vorhersage noch nicht. Daher liegt die Frage nahe, ob der Unterschied zwischen den Modellen auch auftritt, wenn man statt des tatsächlichen das vorhergesagte Geschwindigkeitsinkrement benutzt. Diese Statistik ist in Abb. 4.12 gezeigt, wo die relative Verbesserung $I = (\epsilon_{lin} - \epsilon_{nlin})/v_S$ des nichtlinearen Modells gegen das vom linearen bzw. nichtlinearen Modell vorhergesagte Inkrement aufgetragen

ist. Dabei sind wieder ϵ_{lin} und ϵ_{nlin} die absoluten Fehler des linearen bzw. nichtlinearen Modells, und v_S ist das Geschwindigkeitsinkrement im Vorhersageintervall. Zunächst erkennt man eine interessante Asymmetrie, positive vorhergesagte Inkremente eignen sich im Gegensatz zu negativen als Unterscheidungskriterium. Im positiven Fall erkennt man, daß unabhängig davon, ob das lineare oder das nichtlineare Modell eine Bö vorher sagt, die Überlegenheit des nichtlinearen Algorithmus klar hervortritt. Dies wird umso deutlicher, je stärker die vorhergesagte Änderung des Signals ist. Die Verbesserung des nichtlinearen Modells beträgt bis zu 40% des tatsächlichen Inkrements in dieser Situation.

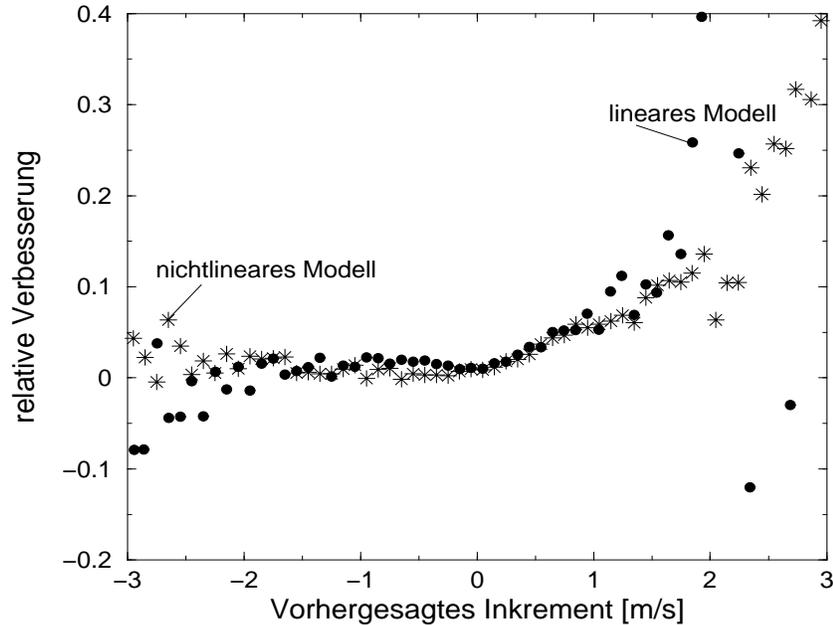


Abbildung 4.12: Relative Verbesserung des nichtlinearen Modells versus der Größe des entweder durch das lineare oder das nichtlineare Modell vorhergesagten Inkrements.

In Abb. 4.13 ist für die positiven vorhergesagten Inkremente wie schon in den Abschnitten 2.2.5 und 3.4.3 noch die Zuverlässigkeitsfunktion gezeigt. Es zeigt sich erneut, daß das nichtlineare Modell im Mittel die richtigen Vorhersagen macht, wogegen das lineare Modell die starken Fluktuationen der Zeitreihe systematisch unterschätzt. In Abb. 4.14 zeigen wir noch die positiven Bereiche der PDFs der durch die beiden Modelle vorhergesagten Inkremente sowie der Inkremente der eigentlichen Zeitreihe. Wieder beobachtet man, daß die Verteilung für das nichtlineare Modell deutlich breiter ist, im Gegensatz zur 2D-Turbulenz scheinen aber die PDFs für beide Modelle exponentiell abzufallen.

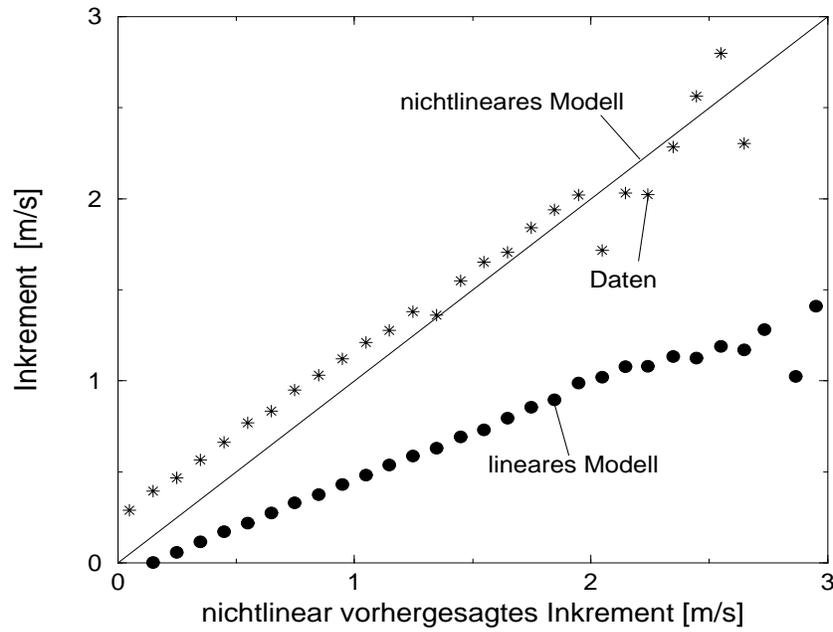


Abbildung 4.13: Zuverlässigkeitsfunktion für die nichtlinear vorhergesagten Geschwindigkeitsinkremente. Zum Vergleich sind die durch das lineare Modell vorhergesagten Inkremente in diesen Situationen gezeigt.

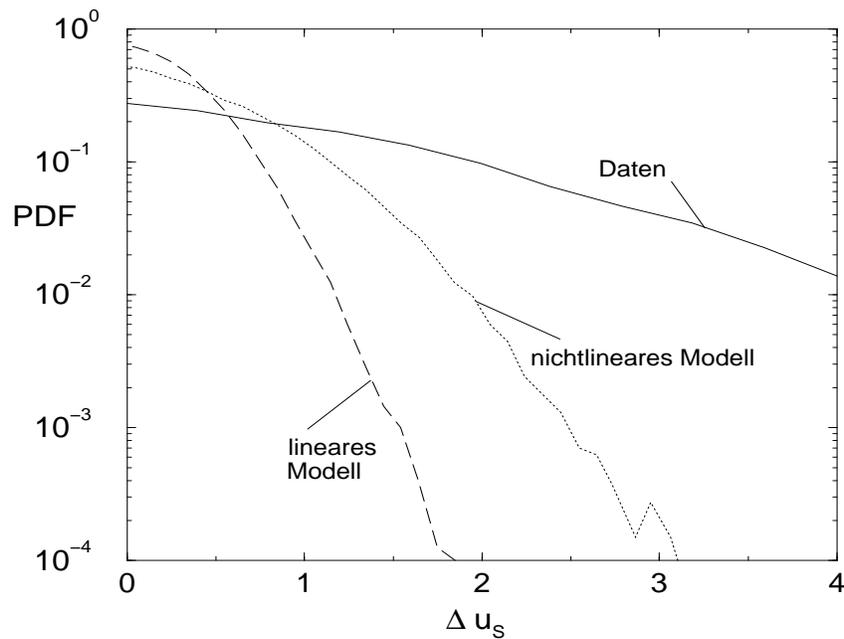


Abbildung 4.14: Positiver Teil der PDFs der tatsächlichen und der vorhergesagten Geschwindigkeitsinkremente.

4.4 Multivariate Analyse

Wegen der raumzeitlichen Natur des betrachteten Vorhersageproblems ist es natürlich naheliegend, räumliche, also multivariate Daten anstatt skalarer Zeitreihen zu nutzen. Wir haben uns jedoch so ausführlich mit skalaren Daten beschäftigt, weil in praktischen Situationen meist nur univariate Windzeitreihen zur Verfügung stehen und weil das Vorhandensein von nichtlinearem Determinismus in diesen Daten von wissenschaftlichem Interesse bei der Analyse kohärenter räumlicher Strukturen ausgehend von einzelnen zeitlichen Messungen ist. Es besteht jedoch kein Zweifel, daß die Nutzung räumlicher Informationen, sofern diese zur Verfügung stehen, die Vorhersagbarkeit verbessern muß. Insbesondere Daten, welche stromaufwärts, also entgegen der Windrichtung gemessen sind, werden Informationen über Strukturen enthalten, welche zu einer späteren Zeit unseren Referenzpunkt erreichen und somit eine deutliche Verbesserung der Vorhersagen bewirken. Die Analyse raumzeitlicher Winddaten wird insbesondere bei der Vorhersage von Windzeitreihen auf Horizonten von einigen Stunden in der Zukunft wichtig sein, da man aufgrund der steigenden Anzahl von Windkraftanlagen über eine gute räumliche Datenbasis verfügen wird und Windkraftanlagen zunehmend mit Online-Datenerfassungssystemen ausgestattet werden. Im Gegensatz dazu ist nicht erkennbar, ob die Nutzung räumlicher Daten auf den hier betrachteten kurzen Vorhersagezeiten in naher Zukunft stattfinden wird, da die Installation eines räumlichen Meßfeldes um eine Windturbine mit erheblichen finanziellen Aufwendungen verbunden wäre. Daher ist die nun folgende Analyse eher als eine Skizze für den Einsatz multivariater Verfahren zu verstehen. Sie soll einen Eindruck von den zu erwartenden Verbesserungen der Vorhersage geben, ohne jedoch Vollständigkeit zu beanspruchen. Die multivariate Anwendung der lokal konstanten, lokal linearen bzw. global linearen Verfahren erfolgt völlig analog zum skalaren Fall, mit dem Unterschied, daß Einbettungsvektoren nun entweder allein aus den Komponenten der vektorwertigen Zeitreihe gebildet werden oder eine Kombination aus der Nutzung der multivariaten Daten und einer Zeitversatzeinbettung angewendet wird. Wir werden im folgenden jedoch keine zusätzliche Zeitversatzeinbettung anwenden und mindestens 50 Nachbarn nutzen. Wie in Abschnitt 4.2 beschrieben, besteht der Dänemark-Datensatz aus neun Messungen, die die Ebene einer fiktiven Windturbine aufspannen, und einer Messung in etwa 15 Metern Entfernung entgegen der Hauptwindrichtung. Wir wollen eine Zeitreihe auf der Rotorebene unter Ausnutzung der gesamten übrigen Information (9 Messungen des Betrags der Windgeschwindigkeit und 9 Messungen der Windrichtung) vorhersagen. Das Resultat, welches in Abb. 4.15 gezeigt ist, läßt sich folgendermaßen zusammenfassen.

- Die Verbesserung bei der Anwendung multivariater Verfahren kann verglichen zur univariaten Analyse erheblich sein und beträgt hier abhängig vom Vorhersagezeitraum 10-20%.
- Abhängig vom Vorhersagehorizont können nichtlineare Phasenraummethoden li-

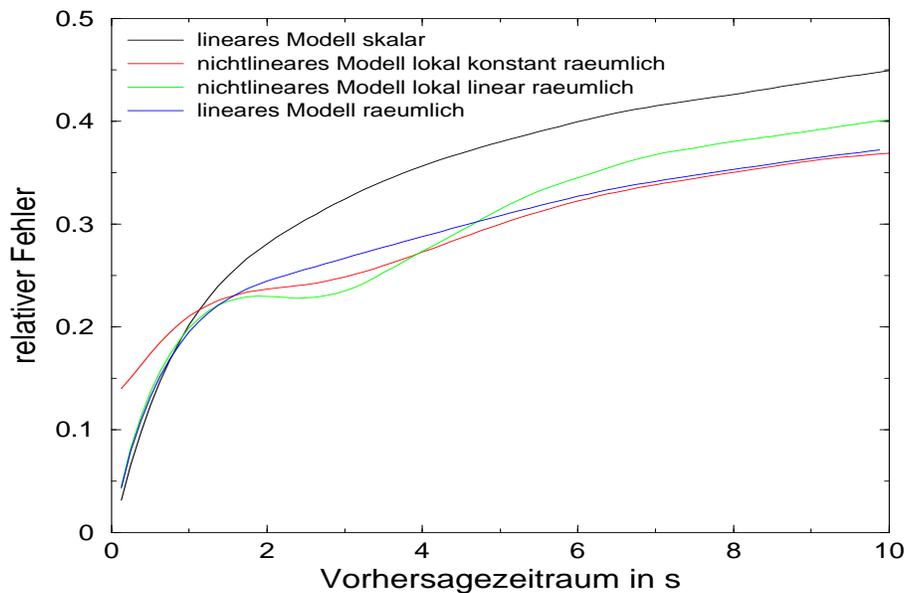


Abbildung 4.15: Relativer Vorhersagefehler skalarer und multivariater - linearer und nichtlinearer Modelle vs Vorhersagehorizont.

nearen Methoden auch **im Mittel** überlegen sein, da z.B. Wettersituationen mit wechselnden Windrichtungen sehr effizient durch Phasenraumverfahren getrennt werden können.

- Die Lage des lokalen Minimums des Vorhersagefehlers in Abhängigkeit vom Horizont entspricht dem Abstand von $l = 15 \text{ m}$ des Referenzpunktes zum nächsten Meßpunkt entgegen der Windrichtung bei einer mittleren Windgeschwindigkeit von etwa $U = 5 \text{ m/s}$. Eine Struktur, die sich zur Zeit t_0 am ersten Meßmast befand, erreicht zur Zeit $t_0 + l/U$ den Referenzpunkt. Die räumliche Meßanordnung kann somit abhängig vom gewünschten Vorhersagehorizont optimiert werden.

Es läßt sich festhalten, daß multivariate Verfahren skalaren Methoden bei der Vorhersage von Windzeitreihen in der Regel überlegen sein werden. Nichtlineare Methoden können im multivariaten Fall Verbesserungen erreichen, die stärker als im skalaren Fall sind.

4.5 Nichtlineare Struktur von Leistungszeitreihen

Im letzten Abschnitt dieses Kapitels wollen wir die Phasenraummethoden auf die Vorhersage von Zeitreihen der von einer Windturbine produzierten elektrischen Leistung anwenden. Zuerst möchten wir jedoch die Relevanz dieses Vorhersageproblems aufzei-

gen. Wie in Abschnitt 4.2 beschrieben, enthält der Spanien-Datensatz u.a. simultane Messungen der produzierten elektrischen Leistung sowie des Dehnungsmoments an den Rotorflügeln der Windturbine, welches die mechanische Belastung widerspiegelt. In Abb. 4.16 ist die PDF des Dehnungsmoments gezeigt sowie das durchschnittliche Dehnungsmoment gegen die Inkremente der Zeitreihe der elektrischen Leistung. Es zeigt sich, daß die Belastung an den Flügeln mit den Inkrementen deutlich zunimmt und das durchschnittliche Dehnungsmoment bei sehr großen Inkrementen bereits in den Ausläufern der gezeigten PDF liegt. Somit ist gezeigt, daß starke Leistungsinkremente Momenten hoher Belastung der Windturbine entsprechen. Eine Vorhersage der Leistungszeitreihe ist daher wünschenswert, so daß die Inkremente durch rechtzeitiges Regeln der Rotorblatteinstellung verringert werden können. Zur Vorhersage der Lei-

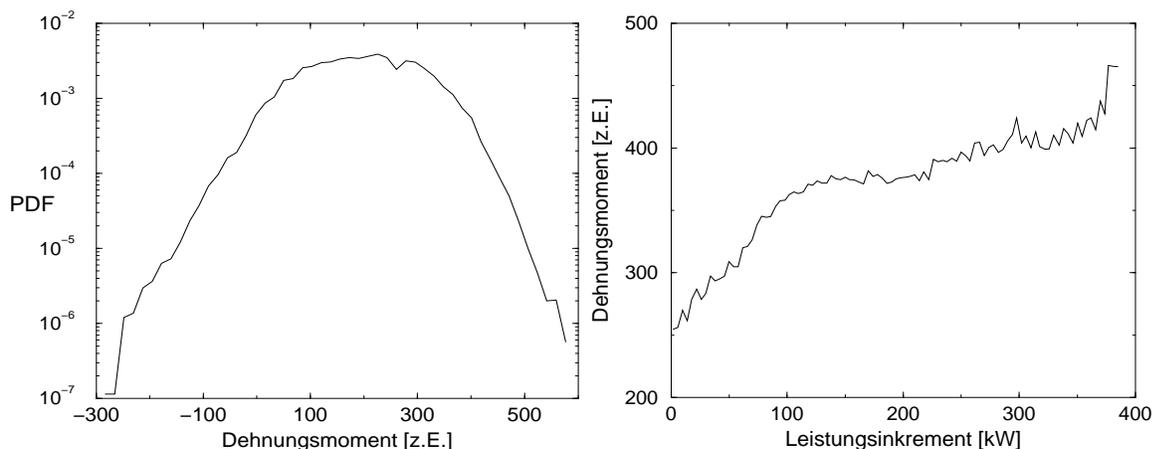


Abbildung 4.16: Wahrscheinlichkeitsverteilung der Dehnungsmomente am Rotorblatt (links) und das gemittelte Dehnungsmoment in zufälligen Einheiten [z.E.] versus des Leistungsincrements innerhalb 1 s Vorhersagezeitraum (rechts).

stung wurden wieder der lokal konstante Vorhersager sowie ein AR -Modell verwendet. Außerdem werden diese mit der Persistenz verglichen, die hier das arithmetische Mittel über die m Komponenten des Einbettungsvektors darstellt. Als Einbettungsdimension wurde wieder $m = 10$ gewählt, und als Zeitverzögerung wurde $\tau = 5$ benutzt. Die wesentlichen Ergebnisse der Vorhersagen können folgendermaßen zusammengefaßt werden:

1. Der gemittelte Vorhersagefehler bei der Leistungsvorhersage von Windturbinen ist bis zu 10 % geringer bei der Verwendung von nichtlinearen Phasenraummodellen verglichen zu herkömmlichen linearen Verfahren (AR -Modell). Die Verbesserung beträgt bis zu 20 % verglichen zur Persistenz, wo der momentane Wert der Leistung als Vorhersagewert verwendet wird. Dies gilt für den hier relevanten Vorhersagezeitraum von einigen Sekunden (siehe Abb. 4.17).

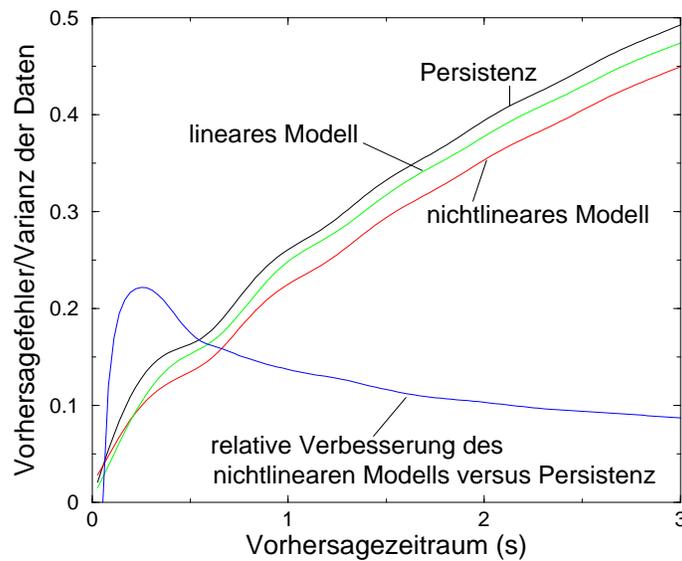


Abbildung 4.17: Relativer Vorhersagefehler der Leistung für das nichtlineare Modell (rot), lineare Modell (grün) und die Persistenz (schwarz). Außerdem relative Verbesserung des nichtlinearen Modells gegenüber der Persistenz (blau).

2. Die durch das nichtlineare Modell vorhergesagte PDF der Fluktuationen der Leistung beschreibt die tatsächliche Verteilung wesentlich genauer als lineare Modelle. Relativ häufig auftretende starke Schwankungen der Leistung werden durch das nichtlineare Modell besser vorhergesagt (siehe Abb. 4.18).

3. Besonders in Situationen, in denen sehr starke Schwankungen der Leistung auftreten oder durch das nichtlineare Modell vorhergesagt werden, sind die Vorhersagen des nichtlinearen Modells deutlich genauer als die des entsprechenden linearen Modells. In Abb. 4.19 ist das mittlere vorhergesagte Leistungsinkrement gegen die tatsächlichen Leistungsinkremente für beide Modelle dargestellt. Für die Vorhersagen des nichtlinearen Modells ist weiterhin die Streuung der Vorhersagen in Höhe der Standardabweichung um die mittleren Vorhersagen dargestellt. Es zeigt sich, daß das nichtlineare Modell besonders bei starken Inkrementen die tatsächlichen Fluktuationen im Mittel richtig beschreibt, wogegen ein lineares Modell zur Beschreibung der Daten ungeeignet ist. Die tatsächlich beobachteten Inkremente in Situationen, in denen das nichtlineare Modell starke Schwankungen vorhersagt, d.h. die oben definierte Zuverlässigkeitsfunktion, sind in Abb. 4.20 aufgetragen. Hier wurde wieder die Standardabweichung der Streuung der Vorhersagen abgezogen bzw. addiert. Diese Abbildung demonstriert die Sicherheit der Vorhersage. Es zeigt sich, daß starke vorhergesagte Schwankungen nicht nur im Mittel hohe tatsächliche Inkremente implizieren, sondern daß dies für jede einzelne Situation zutrifft. Man beobachtet eine deutliche Asymmetrie der Vorhersage, positive Inkremente können deutlich besser vorhergesagt werden als negative. Der Grund

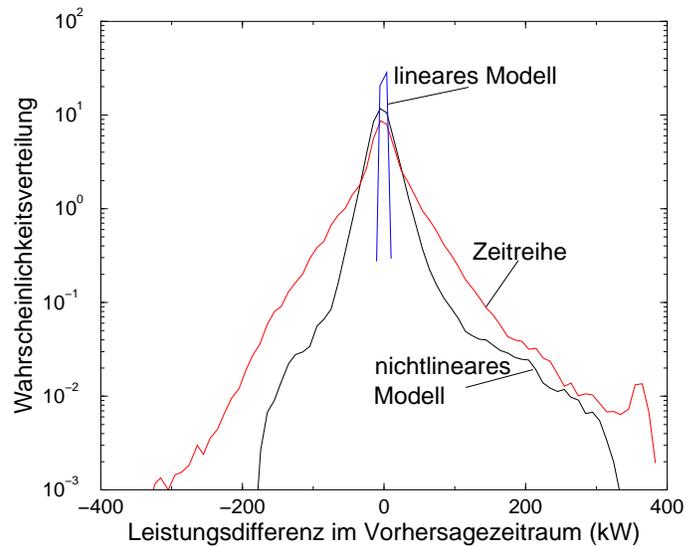


Abbildung 4.18: Wahrscheinlichkeitsverteilung der Leistungsdifferenzen in einem Vorhersagezeitraum von 1 s.

hierfür ist das Abschneiden der Leistung bei 660 kW. Negative Inkremente können sich nicht durch ein typisches Muster in der Leistungszeitreihe ankündigen.

Zusammenfassend wurde festgestellt, daß mit dem nichtlinearen Verfahren turbulente Schwankungen der Leistung verbessert vorhergesagt werden können. Die Verbesserung ist sogar wesentlich größer als für die Zeitreihen der Windgeschwindigkeit. Dies mag im ersten Moment überraschend erscheinen, da man die Geschwindigkeit durch eine weitere Meßfunktion transformiert. Es ist jedoch gerade diese Meßfunktion, die den nichtlinearen Charakter der Daten verstärkt. Diese Meßfunktion ist null unterhalb einer cut-in-Geschwindigkeit von 3 m/s und konstant überhalb einer cut-off-Geschwindigkeit von 25 m/s und wächst dazwischen kubisch an. Nach den Überlegungen aus Abschnitt 2.2.6 kann eine nichtlineare Meßfunktion die Anwendung von Phasenraummethoden gegenüber linearen Verfahren begünstigen, selbst wenn der eigentliche stochastische Prozeß linear ist.

Zuletzt wollen wir noch eine einzelne Böensituation herausgreifen, in welcher die Leistung deutlich ansteigt, um für diese das Verhalten der beiden Modelle zu demonstrieren. In Abb. 4.21 ist ein Ausschnitt aus der Leistungszeitreihe in einer solchen Situation gezeigt. Weiterhin ist in den farbkodierten Abbildungen der Vorhersagefehler gegen die Zeitachse und den Vorhersagezeitraum dargestellt (dunkle Farben bedeuten geringe Fehler). Der Fehler ist dabei für den Zeitpunkt des vorhergesagten Wertes berechnet. Daher sieht man für den Moment, in dem die Leistung stark ansteigt, eine deutliche vertikale Front im Vorhersagefehler. Werte, die in dem hellen Bereich dieser Front liegen, können nur noch schlecht vorhergesagt werden. Hier wird der Unterschied

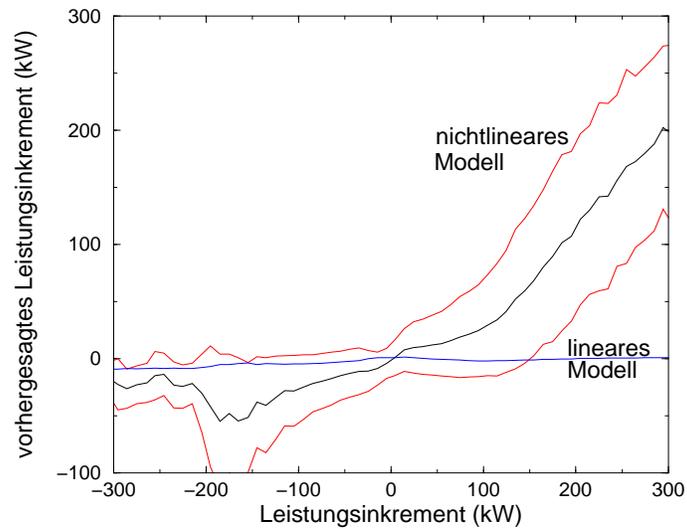


Abbildung 4.19: Stärke der vorhergesagten Fluktuationen gegen die tatsächlichen Leistungszunahmen für das lineare Modell (blau) und das nichtlineare Modell (schwarz) plus/minus der Standardabweichung der Streuung der Vorhersagen (rot). (Vorhersagehorizont 1 s; 660 kW-Anlage)

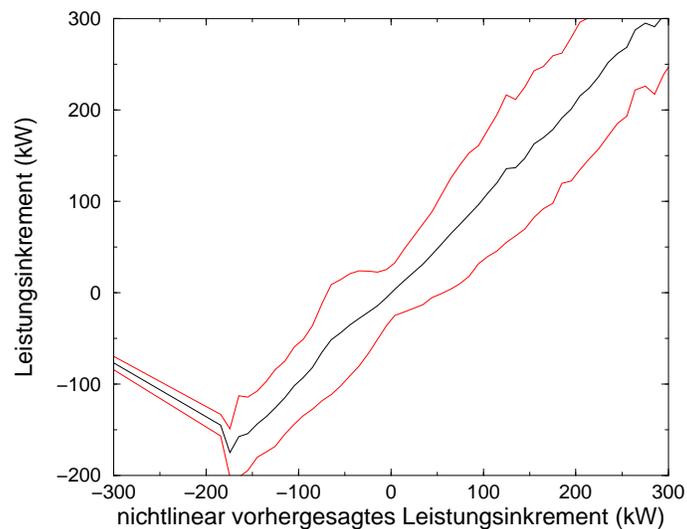


Abbildung 4.20: Stärke der tatsächlichen Fluktuationen gegen die vorhergesagten Leistungszunahmen für das nichtlineare Modell (schwarz) plus/minus der Standardabweichung der Streuung der Vorhersagen (rot). (Vorhersagehorizont 1 s; 660 kW-Anlage)

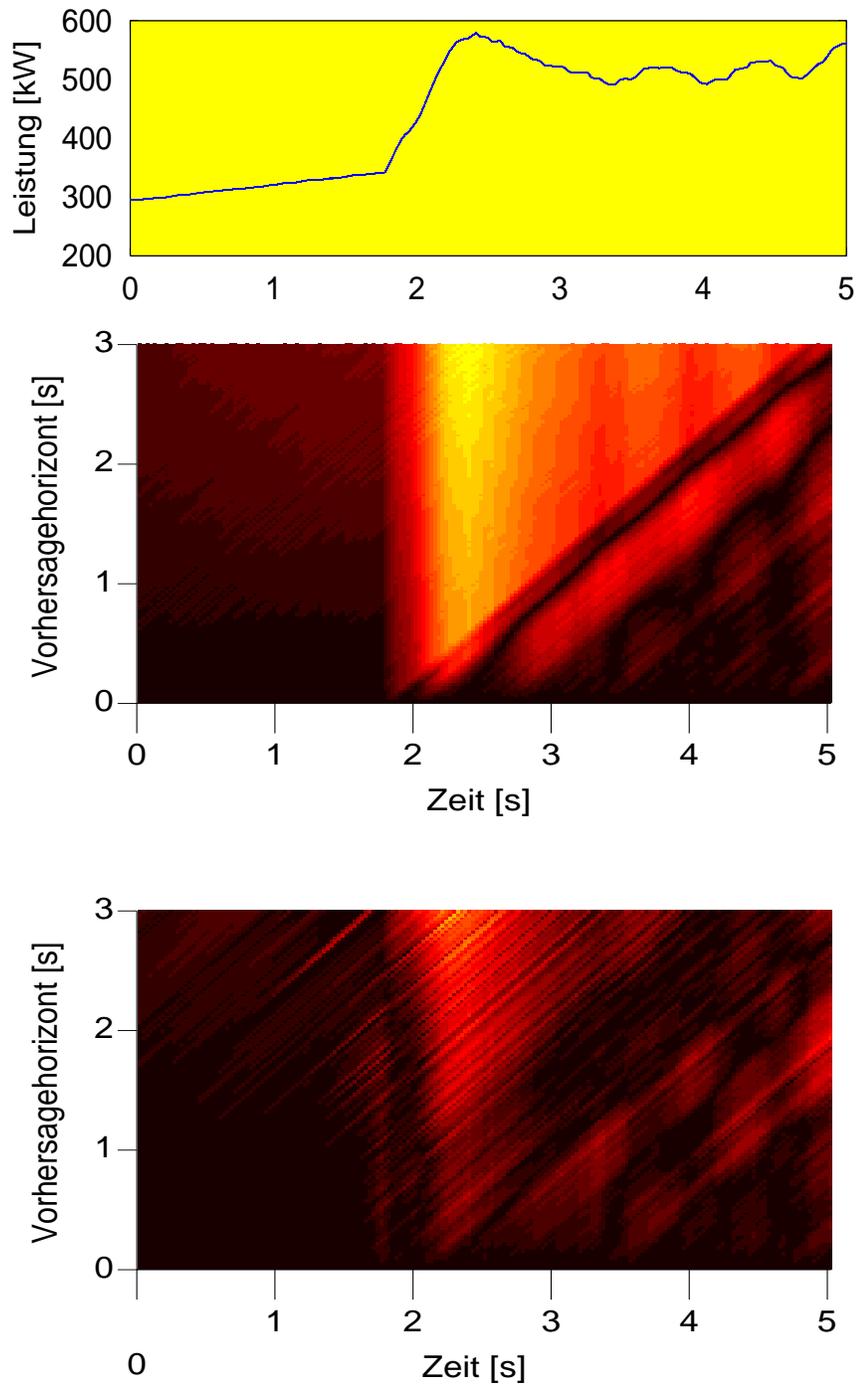


Abbildung 4.21: Leistungssignal in einer Böensituation (oben) und Vorhersagefehler des linearen (Mitte) und des nichtlinearen Modells (unten) in dieser Situation. Die Skala der Farbkodierung ist für beide Modelle gleich, helle Farben bedeuten einen großen Fehler.

zwischen linearem und nichtlinearem Modell deutlich. Während für das *AR*-Modell der Wert beim Zeitpunkt 2.5 *s* schon bei einem Vorhersagehorizont von unter 1 *s* schlecht vorhersagbar ist, tritt eine ähnlich schwache Vorhersagbarkeit für das nichtlineare Modell erst nach etwa 3 *s* auf. Die ausgeprägten diagonalen Linien, die insbesondere für das lineare Modell auftreten, deuten auf Momente in der Zeitreihe, in denen die Vorhersage unabhängig vom Horizont gleichbleibend schlecht oder gut funktioniert. Dies zeigt die Bedeutung des Einbettungsvektors für die Qualität der Vorhersage.

4.6 Diskussion

In diesem Kapitel wurden nichtlineare Phasenraummethoden auf praktisch relevante Vorhersageprobleme bei der Windenergieproduktion angewendet. Insbesondere wurden Zeitreihen der von einer Windturbine produzierten elektrischen Leistung analysiert. Vorhersagen dieser Größe können genutzt werden, um die Rotorblätter von Windkraftanlagen zu steuern, um somit kurzzeitigen Fluktuationen der Windgeschwindigkeit zu begegnen. Es konnte gezeigt werden, daß Phasenraumverfahren gemittelt über alle Situationen linearen Verfahren um etwa 10 % überlegen sind. Eine wesentlich stärkere Verbesserung ergibt sich jedoch in Böensituationen. Die turbulenten “Ausreißer” des Leistungssignals werden durch nichtlineare Methoden mit hoher Genauigkeit beschrieben, siehe Abb. 4.19, während lineare Methoden wesentlich persistenteres Verhalten vorhersagen. Die Überlegenheit nichtlinearer Methoden in Böensituationen ergibt sich nach Vergleich mit dem in Kapitel 3 besprochenen Verhalten aus kohärenten Strukturen in dem betrachteten Fluß. Das Ergebnis dieses Kapitels zeigt, daß auch in atmosphärischer Turbulenz ausgeprägte kohärente Strukturen existieren, die durch die Beobachtung einer Geschwindigkeitszeitreihe aufgelöst werden können.

Kapitel 5

Zusammenfassung

In dieser Arbeit beschäftigten wir uns mit der zeitreihengetriebenen Modellbildung in stochastischen dynamischen Systemen, wobei sich unsere Ausführungen in einen Theorie- und einen Praxisteil gliedern. Dabei haben wir zunächst in Kapitel 2 neuartige bzw. weiterentwickelte Methoden zur Modellbildung in stochastischen Systemen vorgestellt. In den folgenden Kapiteln 3 & 4 wendeten wir diese Verfahren zur Vorhersage und Analyse von Geschwindigkeitsdaten in voll entwickelter Turbulenz sowie von Zeitreihen atmosphärischer Bodenwindgeschwindigkeiten und der Ausgangsleistung einer Windturbine mit nichttrivialen Ergebnissen an.

Im Bereich “Methodenentwicklung” haben wir Verfahren für die Bestimmung von Langevin- bzw. Fokker-Planck-Gleichungen aus Daten hergeleitet. Wie bereits in Risken [18] dargestellt und von Siegert *et al.* [5] explizit für die Datenanalyse angewendet, können der Drift- und der Diffusionskoeffizient einer Fokker-Planck-Gleichung über das erste und zweite Moment einer Zeitreihe bestimmt werden, wenn man infinitesimale Abstraten der Daten annimmt. Diese Annahme ist in realistischen Situationen jedoch im allgemeinen falsch, so daß systematische Korrekturen für Zeitreihen mit endlichen Abstraten insbesondere für den Diffusionsterm notwendig sind. Wir haben in Abschnitt 2.2.1 die Korrekturterme für die Diffusion, die vom Betrag des Driftterms und dessen räumlicher Ableitung stammen, hergeleitet und für die Brownsche Bewegung sowie für ein Teilchen in einem nichtlinearen Doppelmuldenpotential die Wirkung der einzelnen Ausdrücke besprochen. Der Nachteil dieser Fokker-Planck-Methode liegt ganz generell in der Tatsache begründet, daß für die Ableitung einer n -dimensionalen Fokker-Planck-Gleichung aus Daten n Zeitreihen für die entsprechenden n Freiheitsgrade gemessen werden müssen. Diese Forderung ist für typische Anwendungen unrealistisch. Daher gingen wir im zweiten Teil des Kapitels 2 der Frage nach, ob in einem multidimensionalen Markov-Prozeß, z.B. generiert durch eine vektorwertige Langevin-Gleichung, die Zeitreihe für eine einzelne Variable ein endliches Gedächtnis haben kann. Obwohl die Antwort auf diese Frage im allgemeinen negativ ist, zeigt sich, daß die approximative Rekonstruktion des mehrdimensionalen Markov-Prozesses

in vielen Fällen möglich ist. Ob eine niedrigdimensionale Einbettung der stochastischen Zeitreihe sinnvoll ist, hängt sowohl von dem beobachteten Freiheitsgrad ab als auch davon, wie die stochastische Kraft in das System gekoppelt ist. Wir haben gezeigt, daß die optimalen Einbettungsparameter in stochastischen dynamischen Systemen anders als in deterministischen Systemen zu wählen sind. Der optimale Wert der Zeitverzögerung zwischen aufeinanderfolgenden Elementen des Einbettungsvektors ist monoton fallend mit der Amplitude des dynamischen Rauschens. Wir konnten außerdem zeigen, daß durch nichtlineare Phasenraummethoden besonders die großen Inkremente eines stochastischen Signals, welche eine Kombination aus Rauschen und nichtlinearen Korrelationen darstellen, besser vorhergesagt werden können. Eine weitere Anwendung der entwickelten nichtlinearen Verfahren ist in Fällen gegeben, in denen das Signal eines linearen stochastischen Systems durch eine nichtlineare Meßfunktion transformiert wird. Im Kapitel 3 wurden Phasenraummethoden zur Vorhersage von Geschwindigkeitszeitreihen in voll entwickelter Turbulenz angewendet. Dabei wurde die Turbulenz in einem rotierenden Wassertank erzeugt, dessen Winkelgeschwindigkeit variabel war. Bei hohen Winkelgeschwindigkeiten konnte man ausgeprägte kohärente Strukturen in der Flüssigkeit erzeugen, und der Fluß war annähernd zweidimensional. Niedrige Winkelgeschwindigkeiten führten indessen zu einem eher dreidimensionalen Fluß ohne deutliche kohärente Strukturen. Die Frage war, ob die im zweidimensionalen Fall vorhandenen kohärenten Strukturen in einer skalaren Zeitreihe feststellbar sind bzw. ob diese einen Einfluß auf die Vorhersagbarkeit des Signals haben. Es zeigte sich, daß nichtlineare Phasenraumverfahren verglichen mit linearen *AR*-Modellen deutlich verbesserte Vorhersagen innerhalb solcher Strukturen ergeben. Die Wahrscheinlichkeitsverteilung der nichtlinear vorhergesagten Inkremente entspricht einem Potenzgesetz mit dem Skalensexponent -3 , welcher durch einfache Argumente der Kolmogorov 41 - Theorie erklärt werden kann. Auch die Fokker-Planck-Methode, in diesem Fall auf die Dynamik innerhalb der Energieskalen angewandt, ergab einen deutlichen Unterschied zwischen dem zweidimensionalen und dem dreidimensionalen Fluß. Der Diffusionskoeffizient, welcher die Stärke der stochastischen Fluktuationen der Dynamik in den Skalen beschreibt, ist im zweidimensionalen Fall deutlich kleiner.

Im Kapitel 4 wurden die in Kapitel 2 entwickelten Vorhersageverfahren unter dem Aspekt der in Kapitel 3 ausgearbeiteten Fähigkeit, kohärente Strukturen vorherzusagen, auf technisch relevante Vorhersageprobleme angewendet. Diese Anwendung liegt bei der Produktion elektrischer Leistung durch Windturbinen, wo die turbulenten Fluktuationen die mechanische Stabilität der Rotoren gefährden und die Qualität der produzierten elektrischen Leistung reduzieren. Die Vorhersage der Fluktuationen ermöglicht eine rechtzeitige Reaktion durch die Steuerung der Rotorblätter der Turbinen. Die Analysen der Windgeschwindigkeit haben gezeigt, daß starke Fluktuationen, wie im Fall der zweidimensionalen Turbulenz, verbessert durch nichtlineare Phasenraumverfahren vorhergesagt werden können. Dies ist im ersten Moment überraschend, da atmosphärische Turbulenz als dreidimensional betrachtet wird. Die Beobachtung legt jedoch den

Schluß nahe, daß bodennahe atmosphärische Turbulenz ausgeprägte kohärente Strukturen enthält, wie dies im betrachteten Laborexperiment bei starker Rotation der Fall war. Die Verbesserung der Vorhersagbarkeit des Geschwindigkeitssignals wird noch deutlich erhöht, wenn man die Leistungsabgabe einer Turbine anstelle der Geschwindigkeit direkt betrachtet. Der wesentliche Grund hierfür ist die stark nichtlineare Transformation der Geschwindigkeitszeitreihe, wenn die Leistung gemessen wird. Außerdem stellt die Leistung ein räumliches Mittel der Windgeschwindigkeit dar. Dieses Ergebnis ist von großem technischen Interesse, da die elektrische Leistung der eigentliche Kontrollparameter von Windturbinen ist. Zuletzt wurden noch räumliche multivariate Geschwindigkeitszeitreihen betrachtet. Es zeigte sich hier, daß Verbesserungen der räumlichen Analyse im Vergleich zu skalaren Verfahren wesentlich sein können. So kann der gemittelte Vorhersagefehler durch ein vektorwertiges Phasenraumverfahren um etwa 20% geringer sein als bei Verwendung einer skalaren linearen Methode. Diese Analyse führte zu dem Schluß, daß vektorwertige Verfahren genutzt werden sollten, falls entsprechende Daten zur Verfügung stehen.

Zusammenfassend konnte in dieser Arbeit festgestellt werden, daß Phasenraummethoden, die ursprünglich aus dem Bereich der nichtlinearen deterministischen Dynamik stammen, sowohl konzeptionell zu neuen Wegen bei der Analyse stochastischer Systeme führen als auch leistungsstarke Verfahren für praktisch relevante Vorhersageprobleme darstellen.

Ein wesentlicher Teil der Ergebnisse wurde in einzelnen wissenschaftlichen Artikeln veröffentlicht, welche unter [52, 53, 54, 55] aufgeführt sind.

Literaturverzeichnis

- [1] G.E. Box und G.M. Jenkins, *TIME SERIES ANALYSIS*, Holden-Day, San Francisco, 1976.
- [2] F. Takens, *DETECTING STRANGE ATTRACTORS IN TURBULENCE*, Lecture Notes in Math. Vol. 898, (Springer, New York 1981).
- [3] H.I. Abarbanel, *ANALYSIS OF OBSERVED CHAOTIC DATA*, (Springer, New York 1996); H. Kantz and T. Schreiber, *NONLINEAR TIME SERIES ANALYSIS*, (Cambridge University Press, Cambridge, 1997).
- [4] H. Tong, *NONLINEAR TIME SERIES ANALYSIS*, (Oxford University Press, Oxford, 1990).
- [5] S. Siegert, R. Friedrich und J. Peinke, *Analysis of data sets of stochastic systems*, Physics Letters A **243**, 275 (1998).
- [6] B. Van der Pol, *On Relaxation Oscillations*, Phil. Mag.(7) **2**, 978 (1926).
- [7] E.N. Lorenz, *Deterministic non-periodic flow*, J. Atmos. Sci. **20**, 130, (1963).
- [8] T. Sauer, J. Yorke, und M. Casdagli, *Embedology*, J. Stat. Phys. **65**, 579 (1991).
- [9] M.T. Rosenstein, J.J. Collins und C.J. De Luca, *Reconstruction expansion as a geometry-based framework for choosing proper delay times*, Physics D **73**, 82 (1994).
- [10] Th. Buzug und G. Pfister, *Comparison of algorithms calculating optimal parameters from delay time coordinates*, Physics D **58**, 127 (1992).
- [11] A.M. Fraser und H.L. Swinney, *Independant coordinates for strange attractors from mutual information*, Phys. Rev. A **33**, 1134 (1986).
- [12] A. Brandstätter, J. Swift, H.L. Swinney und A. Wolf, *Low-dimensional chaos in a hydrodynamic system*, Phys. Rev. Lett. **51**, 1442 (1983).

- [13] J.D. Farmer und J.J. Sidorowich, *Predicting chaotic time series*, Phys. Rev. Lett. **59**, 845 (1987). Siehe auch J.P. Crutchfield und B.S. McNamara, *Equations of motion from data series*, Complex Systems **1**, 417 (1987) und E.J. Kostelich und J.A. Yorke, *Noise reduction in dynamical systems*, Phys. Rev. A **38**, 1649 (1988).
- [14] E.N. Lorenz, *Atmospheric predictability as revealed by naturally occurring analogues*, J. Atmos. Sci. **26**, 636 (1969).
- [15] M. Casdagli, *Chaos and deterministic versus stochastic nonlinear modelling*, J. Roy. Stat. Soc. B **54**, 303-328 (1991).
- [16] H. Tong, THRESHOLD MODELS IN NONLINEAR TIME SERIES ANALYSIS, Lecture notes in Statistics Vol. 21, (Springer, New York, 1983).
- [17] N. van Kampen, STOCHASTIC PROCESSES IN PHYSICS AND CHEMISTRY, (Elsevier Science, Amsterdam, 1992).
- [18] H. Risken, THE FOKKER PLANCK EQUATION, (Springer, Berlin, 1989).
- [19] F. Paparella, A. Provenzale, L.A. Smith, C. Taricco und R. Vio, *Local random analogue prediction of nonlinear processes*, Physics Letters A **235**, 233-240 (1997).
- [20] P. Gaspard und X.-J. Wang, *Noise, chaos and (ϵ, τ) -entropy per unit time*, Physics Reports **235**, 291 (1993). H. Kantz und E. Olbrich, *Coarse grained dynamical entropies - investigation of high-entropic dynamical systems*, Physica A **280**, 34-48 (2000).
- [21] L.D. Landau und E.M. Lifshitz, FLUID MECHANICS, (Pergamon, Oxford, 1959).
- [22] U. Frisch, TURBULENCE, (Cambridge University Press, Cambridge, 1995).
- [23] J. Pedlosky, GEOPHYSICAL FLUID DYNAMICS, (Springer Verlag, Heidelberg, 1979).
- [24] M. Lesieur, TURBULENCE IN FLUIDS, (Kluwer Academic Publishers, Dordrecht, 1990).
- [25] A.N. Kolmogorov, *The local structure in incompressible viscous fluid for very large Reynolds numbers*, Dokl. Akad. Nauk SSSR **30**, 301-305 (1941).
- [26] A.N. Kolmogorov, *On degeneration of isotropic turbulence in an incompressible viscous fluid*, Dokl. Akad. Nauk SSSR **31**, 538-541 (1941).

- [27] K. Gawedzki, *Easy Turbulence*, chao-dyn/9907024 v2.
- [28] R.H. Kraichnan, *Inertial ranges in two-dimensional turbulence*, Phys. Fluids **10**, 1417-1423 (1967).
- [29] J. Paret und P. Tabeling, *Intermittency in 2D inverse cascade of energy: experimental observations*, Phys. Fluids **10**, 3126-3136 (1998).
- [30] G. Boffetta, A. Celani und M. Vergassola, *Inverce cascade in two-dimensional turbulence: deviations from Gaussianity*, chao-dyn/9906016.
- [31] T. H. Solomon, W. J. Holloway und H. L. Swinney, *Shear Flow Instabilities and Rossby Waves in Barotropic Flow in a Rotating Annulus*, Phys. Fluids **5**, 1971 (1993).
- [32] J. Sommeria, S. D. Meyers und H. L. Swinney, *Laboratory model of a planetary eastward jet*, Nature **337**, 58 (1989). Auch in J. Sommeria, S.D. Meyers und H. L. Swinney, *Experiments on Vortices and Rossby Waves in Eastward and Westward Jets*, Nonlinear Topics in Ocean Physics, 227-269, editiert durch A. Osborne, (1988).
- [33] R. Benzi, S. Ciliberto, R. Tripicciono, C. Baudet, F. Massaioli, und S. Succi, *Extended self-similarity in turbulent flows*, Phys. Rev. E **48**, R29 (1993).
- [34] R. Friedrich und J. Peinke, *Description of a Turbulent Cascade by a Fokker-Planck Equation*, Phys. Rev. Lett. **78**, 863 (1997).
- [35] S. Lück, J. Peinke und R. Friedrich, *Uniform Statistical Description of the Transition between Near and Far Field Turbulence in a Wake Flow*, Phys. Rev. Lett. **83**, 5495 (1999).
- [36] R. Friedrich, J. Peinke und Ch. Renner, *How to Quantify Deterministic and Stochastic Influences on the Statistics of the Foreign Exchange Market*, Phys. Rev. Lett. **84**, 5224 (2000).
- [37] Ch. Renner, R. Friedrich und J. Peinke, *Markov properties of small scale turbulence*, Journal of Fluid Mechanics **433**, 383 (2001).
- [38] Landberg, L., *Predicting the Power Output from Windfarms*, European Wind Energy Conference, October 1997, Dublin, Ireland.
- [39] B.G. Brown, R.W. Katz und A.H. Murphy, *Time Series Models to forecast Wind Speed and Wind Power*, Amer. Meteor. Soc. **23**, 1184-1195 (1984).
- [40] U. Schlink und G. Tetzlaff, *Wind speed forecasting from 1 to 30 minutes*, Theoretical and Applied Climatology **60**, 191 (1998).

- [41] Bossanyi, E.A., *Short Term Prediction Using Kalman Filters*, Wind Engineering **9**, No.1, (1985).
- [42] Hausmann, J., Hoffmann, M., *Zeitreihenvorhersage und Steuerung autonomer Wind-Diesel-Systeme mit neuronalen Netzwerken*, Diplomarbeit, Uni Oldenburg (1995).
- [43] P. Grassberger und I. Procaccia, *Characterization of strange attractors*, Phys. Rev. Lett. **50**, 346 (1983).
- [44] K. Freadrich, *Estimating the dimension of of weather and climate attractors*, J. Atmos. Sci. **43**, 419 (1986).
- [45] C.L. Keepenne und C. Nicolis, *Global properties and local structure of the weather attractor over western Europe*, J. Atmos. Sci. **46**, 2356 (1989).
- [46] C. Nicolis und G. Nicolis, *Is there a climatic attractor ?*, Nature **311**, 529 (1984).
- [47] G. Poveda-Jaramillo und C.E. Puente, *Strange Attractors in Atmospheric Boundary-Layer Turbulence*, Boundary-Layer Meteorology **64**, 175 (1993).
- [48] A.A. Tsonis und J.B. Elsner, *The Weather Attractor over very short Timescales*, Nature **333**, 545 (1988).
- [49] M.B. Sharifi, K.P. Georgakakos und I. Rodriguez-Iturbe, *Evidence of deterministic chaos in the pulse of storm rainfall*, J. Atmos. Sci. **47**, 888 (1990).
- [50] P. Grassberger, *Do climatic attractors exist?*, Nature **323**, 609 (1986).
- [51] A. Wolf, J. B. Swift, H. L. Swinney, und J. A. Vastano, *Determining Lyapunov exponents from a time series*, Physica D **16**, 285 (1985). J.P. Eckmann and D. Ruelle, *Fundamental limitations for estimating dimensions and lyapunov exponents in dynamical systems*, Physica D **56**, 185 (1992).
- [52] M. Ragwitz und H. Kantz, *Detecting non-linear structure and predicting turbulent gusts in surface wind velocities*, Europhys. Lett. **51**, 595 (2000).
- [53] M. Ragwitz und H. Kantz, *Indispensable finite time corrections for Fokker-Planck equations from time series data*, akzeptiert zur Publikation bei Phys. Rev. Lett. (2001).
- [54] M. Ragwitz, C. N. Baroud, B. B. Plapp und H. L. Swinney, *Nonlinear determinism in time series measurements of 2-dimensional turbulence*, eingereicht bei Physica D (2001).

- [55] M. Ragwitz und H. Kantz, *Markov models from data by simple nonlinear time series predictors in delay embedding spaces*, eingereicht bei Phys. Rev. E (2001).

Danksagung

Ich möchte mich bei allen bedanken, die zum Entstehen dieser Arbeit beigetragen haben.

Mein besonderer Dank gilt Priv. Doz. Dr. Holger Kantz, der jederzeit in anregenden und kritischen Diskussionen den Fortschritt der Arbeiten begleitete. Seine tolerante Betreuung, die mir immer auch viele Freiheiten ließ, kam mir sehr entgegen.

Weiterhin danke ich Prof. Dr. Harry Swinney, der mir den überaus interessanten Aufenthalt am *Center for Nonlinear Dynamics* an der *University of Texas at Austin* ermöglichte. In intensiven Diskussionen lernte ich viel über experimentelle Aspekte der Nichtlinearen Dynamik.

Es war insgesamt sehr angenehm, in der Arbeitsgruppe *Nichtlineare Zeitreihenanalyse* arbeiten zu dürfen. Dafür danke ich allen Mitgliedern der Gruppe.

Dem *Max-Planck-Institut für Physik komplexer Systeme* danke ich für die sehr angenehme Arbeitsatmosphäre und dem *Bundesministerium für Wirtschaft* für die finanzielle Unterstützung.

Herrn Prof. Dr. Karl-Heinz Mütter danke ich für seine Bereitschaft, das Zweitgutachten zu übernehmen.

Insbesondere danke ich Ulrike Nitschke für Ihre Geduld während des gesamten Verlaufs der Arbeit und speziell beim Korrekturlesen.