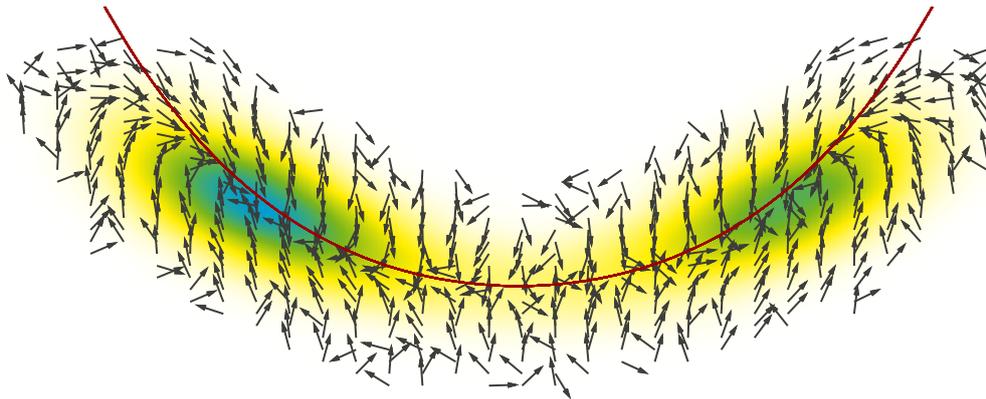


# Stochastische Analyse kollektiver Variabler in komplexen Systemen

als Diplomarbeit vorgelegt von Christoph Honisch



*Never express yourself more clearly  
than you are able to think.*

NIELS BOHR

# Inhaltsverzeichnis

<b>Einleitung</b>	<b>1</b>
<b>1. Stochastische Prozesse</b>	<b>5</b>
1.1. Einige grundlegende Begriffe der Stochastik . . . . .	5
1.2. Multivariate Wahrscheinlichkeitsdichten . . . . .	8
1.3. Schätzen von Wahrscheinlichkeitsdichten . . . . .	9
1.3.1. Der Histogrammschätzer . . . . .	10
1.3.2. Der Kernschätzer . . . . .	10
1.4. Zeitabhängige Zufallsvariablen . . . . .	12
1.5. Markov-Prozesse . . . . .	13
1.6. Langevin-Gleichungen . . . . .	14
1.7. Die Fokker-Planck-Gleichung . . . . .	15
1.7.1. Herleitung . . . . .	15
1.7.2. Stationäre Lösung . . . . .	17
1.8. Schätzen von Drift und Diffusion . . . . .	17
1.8.1. Die generelle Methode . . . . .	18
1.8.2. Stationäre, nicht stationäre und periodische Prozesse . . . . .	19
<b>2. Synergetik</b>	<b>21</b>
2.1. Einführung . . . . .	21
2.1.1. Komplexe Systeme und Selbstorganisation . . . . .	21
2.1.2. Nichtgleichgewichts-Phasenübergänge . . . . .	22
2.1.3. Was leistet die Theorie? . . . . .	22
2.1.4. Von unten nach oben oder andersherum . . . . .	22
2.2. Einige Grundbegriffe der nichtlinearen Dynamik . . . . .	23
2.2.1. Dynamische Systeme . . . . .	23
2.2.2. Lineare, autonome dynamische Systeme . . . . .	24
2.2.3. Nichtlineare autonome Systeme . . . . .	24
2.3. Das Versklavungsprinzip . . . . .	26
2.3.1. Anwendung auf autonome dynamische Systeme . . . . .	26
2.3.2. Ein Beispiel: Das Haken-Zwanzig-System . . . . .	29
2.3.3. Anwendung auf die Fokker-Planck-Gleichung . . . . .	30

2.4. Das Prinzip der unvoreingenommenen Schätzung . . . . .	33
2.4.1. Die Informationsentropie . . . . .	34
2.4.2. Anwendung auf Systeme in der Nähe von Instabilitäten . . . . .	35
2.5. Die Information von Systemen in der Nähe von Instabilitäten . . . . .	38
2.5.1. Aufspaltung der Information . . . . .	38
2.5.2. Die Ordnungsparameterinformation in der Nähe von Instabilitäten . . . . .	39
<b>3. Ordnungsparameterdynamik für ein System gekoppelter Oszillatoren</b>	<b>41</b>
3.1. Einführung . . . . .	41
3.2. Die gaußsche Approximation . . . . .	43
3.3. Numerische Simulation . . . . .	46
3.4. Stationarität . . . . .	47
3.5. Markov-Eigenschaft . . . . .	48
3.6. Schätzen der Kramers-Moyal-Koeffizienten . . . . .	51
3.6.1. Zeitabhängigkeit . . . . .	51
3.6.2. Der Grenzwert $\tau \rightarrow 0$ . . . . .	54
3.6.3. Diffusion und höhere Momente . . . . .	54
3.6.4. Die qualitative Form der Drift . . . . .	56
3.6.5. Der Gleichgewichtszustand . . . . .	57
3.7. Rekonstruktion und Vergleich . . . . .	59
3.7.1. Stationäre pdf . . . . .	59
3.7.2. Einpunkt-Korrelationsfunktion . . . . .	60
3.8. Zusammenfassung und Ausblick . . . . .	62
<b>4. Identifikation von Ordnungsparametern und invarianten Mannigfaltigkeiten</b>	<b>65</b>
4.1. Motivation . . . . .	65
4.2. Testsystem . . . . .	67
4.2.1. Erzeugung der Zeitreihe . . . . .	68
4.3. Der Driftansatz . . . . .	69
4.3.1. Motivation . . . . .	69
4.3.2. Numerische Umsetzung . . . . .	70
4.3.3. Ergebnis . . . . .	70
4.4. Der Informationsansatz . . . . .	71
4.4.1. Motivation . . . . .	71
4.4.2. Numerische Umsetzung . . . . .	73
4.4.3. Ergebnis . . . . .	76
4.5. Invariante Mannigfaltigkeit . . . . .	78
4.6. Schätzen der Ordnungsparametergleichung . . . . .	80
4.7. Zusammenfassung und Ausblick . . . . .	82
<b>A. Gradient der Information</b>	<b>85</b>
<b>Literaturverzeichnis</b>	<b>89</b>

# Einleitung

Komplexe Systeme sind von enormer Wichtigkeit und beeinflussen uns in nahezu allen Bereichen des Lebens, betrachtet man nur das Klima, die Finanzmärkte oder das Leben selbst als Beispiele. Daher erlangten komplexe Systeme in den letzten Jahrzehnten zunehmend mehr Aufmerksamkeit in der Wissenschaft.

Unter diesem Schlagwort fasst man Systeme zusammen, die aus einer großen Anzahl von Subsystemen bestehen, deren gegenseitige Wechselwirkung als mehr oder weniger *komplex* angesehen werden kann [21]. Werden solche Systeme durch einen Strom von Energie, Materie, Geld oder Nahrung, je nachdem, was in dem betrachteten System relevant ist, aus dem Gleichgewicht getrieben, können sie die Fähigkeit besitzen, aus sich heraus makroskopische, geordnete Strukturen zu entwickeln, ohne dass diese dem System von außen aufgeprägt werden. Dieses Phänomen bezeichnet man als *Selbstorganisation*.

Selbstorganisationsprozesse entstehen durch meist nichtlineare Wechselwirkungen einer großen Anzahl von Subsystemen. Den Modellgleichungen, die diese Wechselwirkungen beschreiben, kann man jedoch im Allgemeinen nicht "ansehen", welche Strukturen sie produzieren. Ein besseres Verständnis der Mechanismen, nach denen Selbstorganisation funktioniert, wird im Rahmen der *Synergetik* [23, 24]<sup>1</sup>, einer in den 1970er Jahren entstandenen Theorie über Selbstorganisation, erlangt. Danach entstehen makroskopische Strukturen, indem einige wenige kollektive Größen, in der Theorie als *Ordnungsparameter* bezeichnet, die Dynamik des Systems beherrschen und die verbleibenden Freiheitsgrade unterwerfen.

Die Problemstellung, mit der sich diese Arbeit beschäftigt, besteht darin, aus der Analyse von Messdaten komplexer Systeme zu einem Modell in Form von Ordnungsparametergleichungen zu gelangen. Diese Problematik ist kein völliges Neuland. Um die Jahrtausendwende wurde eine Methode [12, 49] entwickelt, die es unter bestimmten Voraussetzungen erlaubt, deterministische (Drift) und stochastische (Diffusion) Anteile von Bewegungsgleichungen aus Zeitreihen von Messgrößen getrennt voneinander zu extrahieren (s. Abschnitt 1.8). Dadurch lässt sich ein Modell in Form langevinartiger stochastischer Differentialgleichungen oder dazu äquivalenter Fokker-Planck-Gleichungen gewinnen. Diese Methode wurde mittlerweile zur Analyse vieler komplexer Systeme

---

<sup>1</sup>Etwaigen Lesern, die sich selbst als naturwissenschaftliche Laien bezeichnen würden, sei das in weiten Teilen allgemein verständliche Buch [22] empfohlen.

aus verschiedenen Bereichen wie Physik [10], Medizin [15] oder Finanzwirtschaft [11] angewendet.

Im Rahmen dieser Arbeit wird obiges Verfahren verwendet, um das Verhalten einer kollektiven Variablen aus einem Modellsystem zu analysieren, wobei die Daten synthetisch, d. h. durch numerische Simulationen, erzeugt werden. Das Modellsystem besteht aus einer großen Anzahl identischer, miteinander gekoppelter, verrauschter Oszillatoren in einem bimodalen Potential im Grenzfall starker Dämpfung. Die kollektive Variable entspricht dabei dem Mittelwert aller Oszillatorzustände und kann als Ordnungsparameter betrachtet werden.

Dieses System zeigt ein für komplexe, selbstorganisierende Systeme typisches Phänomen, nämlich die Rückkopplung zwischen mikroskopischen Freiheitsgraden und Ordnungsparametern. Das Verhalten des Ordnungsparameters, hier die mittlere Oszillatorvariable, ergibt sich aus den Zuständen sämtlicher einzelner Oszillatoren, den mikroskopischen Freiheitsgraden. Jeder dieser Oszillatoren wird hingegen durch die Kopplung mit allen anderen Oszillatoren maßgeblich durch den Ordnungsparameter beeinflusst. Man spricht in diesem Zusammenhang auch von *zirkularer Kausalität*. Das gleiche Phänomen findet man beispielsweise auch im Laser, dem Musterbeispiel der Synergetik [23]. Ein weiterer Grund, weshalb dieses System von allgemeinem wissenschaftlichen Interesse ist, besteht in der überraschend starken stochastischen Resonanz<sup>2</sup>, die man bei der Variation verschiedener Systemparameter beobachtet [5, 38, 40]. Ein qualitatives Verständnis dieser Effekte würde durch die Kenntnis einer effektiven eindimensionalen stochastischen Bewegungsgleichung, die das Verhalten der kollektiven Variablen beschreibt, erheblich vereinfacht werden [17, 39]. Eine solche Gleichung kann jedoch analytisch nicht exakt abgeleitet werden. Daher wird das oben erwähnte Verfahren zur Schätzung von Drift und Diffusion dazu verwendet, auf numerischem Wege zu einem effektiven Modell zu gelangen.

Ein bisher ungelöstes Problem bei der Analyse von komplexen Systemen, deren Bewegungsgleichungen weder auf mikroskopischer noch auf makroskopischer Ebene bekannt sind, besteht in der Identifikation von Ordnungsparametern. Dies ist insbesondere bei der Analyse raumzeitlicher Daten von Bedeutung. In der Nähe von Instabilitäten lässt sich ein raumzeitliches Signal eines komplexen Systems durch eine Überlagerung stabiler und instabiler räumlicher Moden darstellen. Die Amplituden der instabilen Moden fungieren dabei als Ordnungsparameter und dominieren die Dynamik des Systems, die sich im Wesentlichen auf einer invarianten Mannigfaltigkeit abspielt. Zur Identifikation von Ordnungsparametern und invarianten Mannigfaltigkeiten werden in dieser Arbeit neue Methoden aus der Theorie der Synergetik heraus entwickelt und an einem einfachen Modellsystem getestet. Die Abbildung auf der Titelseite zeigt das Driftvektorfeld, die farbkodierte Verbundwahrscheinlichkeitsdichte sowie die invariante Mannigfaltigkeit dieses Modellsystems.

Die Arbeit gliedert sich in vier Kapitel. Das erste Kapitel beinhaltet eine Einführung

---

<sup>2</sup>s. [14] für einen Übersichtsartikel über stochastische Resonanz

in die Theorie der stochastischen Prozesse, welche bei der Beschreibung komplexer Systeme eine wichtige Rolle spielen. Im zweiten Kapitel werden einige Grundprinzipien der Synergetik eingeführt. Die oben erwähnte Analyse der Ordnungsparameterdynamik eines Modellsystems gekoppelter Oszillatoren ist Gegenstand des dritten Kapitels. Schließlich widmet sich das vierte Kapitel der Identifikation von Ordnungsparametern und invarianten Mannigfaltigkeiten.

Da die Kapitel 3 und 4 weitgehend unabhängig voneinander entstanden sind, werden die erzielten Resultate jeweils am Ende der Kapitel zusammengefasst und diskutiert. Auf eine Zusammenfassung der Arbeit als Ganzes wird verzichtet.



# 1. Stochastische Prozesse

Durch das in der Einleitung angesprochene Ordnungsparameterkonzept gelangt man zu einer makroskopischen Beschreibung komplexer Systeme mit einer stark reduzierten Anzahl von Freiheitsgraden. Den nicht berücksichtigten mikroskopischen Freiheitsgraden wird in den Ordnungsparametergleichungen durch Rauschterme Rechnung getragen; folglich haben wir es mit stochastischen Prozessen zu tun. Daher werden in diesem Kapitel einige für das Weitere unverzichtbare Begriffe und Zusammenhänge aus diesem Bereich dargestellt. Dazu werden im ersten Abschnitt der Wahrscheinlichkeitsbegriff sowie die Wahrscheinlichkeitsdichte für univariate Zufallsvariablen über Grenzwerte von Häufigkeiten eingeführt und im darauf folgenden Abschnitt auf multivariate Zufallsvariablen verallgemeinert. Es wird dabei mehr Wert auf die Vermittlung eines intuitiven Verständnisses der Begriffe als auf mathematische Exaktheit gelegt. Der dritte Abschnitt beschäftigt sich mit der Schätzung von Wahrscheinlichkeitsdichten aus Stichproben. In den darauf folgenden Abschnitten geht es dann schließlich um zeitabhängige Zufallsvariablen, also um stochastische Prozesse. Der dabei eingeschlagene Weg führt uns von Markov-Prozessen über Langevin-Gleichungen zur Fokker-Planck-Gleichung. Dabei wird jeweils nur der eindimensionale Fall behandelt. Schließlich widmet sich der letzte Abschnitt einer Methode zur Schätzung von Kramers-Moyal-Koeffizienten aus Daten. Als hauptsächliche Quellen für dieses Kapitel dienen [45] und [42].

## 1.1. Einige grundlegende Begriffe der Stochastik

Ein *Zufallsexperiment* ist ein Experiment, dessen Ausgang nicht vorhersagbar ist. Den Ausgang eines Zufallsexperimentes bezeichnet man als *Ereignis*, die Menge aller möglichen Ereignisse als *Ergebnismenge*  $\Omega$ . Ordnet man jedem Ereignis eine Zahl zu, bezeichnet man diese Zahl als *Zufallsvariable* oder *stochastische Variable*  $\xi$ :

$$\xi : \Omega \rightarrow \Omega' \subseteq \mathbb{R} . \tag{1.1}$$

Beispielsweise gibt es beim klassischem Münzwurf die Ereignisse “Kopf” oder “Zahl”. Wir können nun eine stochastische Variable  $\xi$  definieren, die beim Ereignis “Kopf” die Zahl 1 annimmt und beim Ereignis “Zahl” die Zahl 42. Die Zahlen 1 und 42 sind dann *Realisationen* von  $\xi$ . In diesem Beispiel ist  $\Omega = \{\text{“Kopf”}, \text{“Zahl”}\}$  und  $\Omega' = \{1, 42\}$ .

Wirft man obige Münze  $N$  mal, werden die Realisationen 1 bzw. 42 mit den *Häufigkeiten*  $H_{\xi,N}(1)$  bzw.  $H_{\xi,N}(42)$  auftreten, wobei  $H_{\xi,N}(1) + H_{\xi,N}(42) = N$  ist. Teilt man die Häufigkeiten der jeweiligen Realisationen durch die Anzahl  $N$  der Münzwürfe, erhält man die *relativen Häufigkeiten*

$$h_{\xi,N}(i) = \frac{H_{\xi,N}(i)}{N}, \quad \forall i \in \Omega', \quad (1.2)$$

die offensichtlich der Beziehung

$$h_{\xi,N}(1) + h_{\xi,N}(42) = \frac{1}{N} [H_{\xi,N}(1) + H_{\xi,N}(42)] = 1 \quad (1.3)$$

genügen. Als *Wahrscheinlichkeit*, bei einem Münzwurf die Realisation  $i$  zu erhalten, definieren wir

$$P_{\xi}(i) = \lim_{N \rightarrow \infty} h_{\xi,N}(i). \quad (1.4)$$

In diesem Beispiel besteht  $\Omega'$  aus abzählbar vielen Elementen. Es gibt genau zwei Realisationen. In der Physik haben wir häufig mit Zufallsexperimenten zu tun, bei denen es nicht abzählbar viele Realisationen gibt. In diesem Fall kann man meist nicht mehr wie oben eine Wahrscheinlichkeit für eine einzelne Realisation definieren. Man kann jedoch eine Wahrscheinlichkeit für eine Teilmenge der Menge aller möglichen Realisationen definieren, indem wir beispielsweise fragen, wie groß die Wahrscheinlichkeit ist, dass die Realisation eines Experimentes in einem bestimmten Intervall liegt.

Unterteilen wir  $\Omega'$  in nichtüberlappende Intervalle  $I_j$ ,

$$\Omega' \subseteq \mathbb{R} = \bigcup_{j=1}^M I_j, \quad (1.5)$$

und wiederholen das Experiment  $N$  mal, können wir jedem Intervall eine relative Häufigkeit  $h_N(I_j)$  zuweisen, wobei analog zu Gl. (1.3)

$$\sum_{j=1}^M h_{\xi,N}(I_j) = 1 \quad (1.6)$$

gilt. Die Wahrscheinlichkeit, die Realisation eines Experimentes im Intervall  $I_j$  zu finden, kann dann definiert werden als

$$P_{\xi}(x \in I_j) = \lim_{N \rightarrow \infty} h_{\xi,N}(I_j). \quad (1.7)$$

Um aus den relativen Häufigkeiten den Begriff der Wahrscheinlichkeitsdichte abzuleiten, führen wir die *Häufigkeitsdichten*

$$\tilde{h}_{\xi,N}(I_j) = \frac{h_{\xi,N}(I_j)}{|I_j|} \quad (1.8)$$

ein, wobei  $|I_j|$  die Breite des Intervalls  $I_j$  bedeutet. Lässt man nun die Anzahl der Intervalle  $M$  gegen unendlich und gleichzeitig Länge eines jeden Intervalls gegen null gehen, wird aus der Häufigkeitsdichte eine Funktion, die sich aus einer Summe von  $\delta$ -Funktionen zusammensetzt:

$$\tilde{h}_{\xi,N}(I_j) \rightarrow \tilde{h}_{\xi,N}(x) = \frac{1}{N} \sum_{i=1}^N \delta(x - \xi_i). \quad (1.9)$$

Dabei stellt  $\xi_i$  die Realisation der stochastischen Variable beim  $i$ ten Experiment dar. Nun können wir die Wahrscheinlichkeitsdichte (engl. *probability density function*, pdf) definieren:

$$\begin{aligned} f_{\xi}(x) &:= \lim_{N \rightarrow \infty} \tilde{h}_{\xi,N}(x) \\ &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \delta(x - \xi_i) \\ &\equiv \langle \delta(x - \xi) \rangle. \end{aligned} \quad (1.10)$$

In der letzten Umformung haben wir den *Erwartungswert*

$$\langle g(\xi) \rangle = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N g(\xi_i) \quad (1.11)$$

eingeführt. Die Wahrscheinlichkeit, dass die Realisation von  $\xi$  bei einem Experiment in einem beliebigen Intervall  $[a, b] \subseteq \Omega'$  liegt, lautet damit:

$$P_{\xi}(a \leq x \leq b) = \int_a^b f_{\xi}(x) dx. \quad (1.12)$$

Aus Gl. (1.10) ergibt sich die Normierungs-Bedingung

$$\int_{\Omega} f_{\xi}(x) dx = 1. \quad (1.13)$$

Ist die pdf einer stochastischen Variablen bekannt, lassen sich damit Erwartungswerte berechnen:

$$\langle g(\xi) \rangle = \int_{\Omega'} g(x) f_{\xi}(x) dx. \quad (1.14)$$

Dass diese Formel mit der Definition (1.11) übereinstimmt, zeigt folgende Rechnung:

$$\begin{aligned}
\langle g(\xi) \rangle &= \int_{\Omega'} g(x) f_{\xi}(x) dx \\
&= \int_{\Omega'} g(x) \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \delta(x - \xi_i) \\
&= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \int_{\Omega'} g(x) \delta(x - \xi_i) \\
&= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N g(\xi_i). \tag{1.15}
\end{aligned}$$

Die pdf, die in der Natur am häufigsten vorkommt, ist die Gauß-Funktion

$$f_{\xi}(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right). \tag{1.16}$$

Zufallsvariablen, deren pdf eine Gauß-Funktion ist, nennt man *gaußverteilt* oder *normalverteilt*. Der Grund, weshalb so viele Zufallsvariablen gaußverteilt sind, liegt im zentralen Grenzwertsatz (s. [42]), der hier jedoch nicht behandelt wird.

## 1.2. Multivariate Wahrscheinlichkeitsdichten

Wir betrachten nun mehrere kontinuierliche stochastische Variablen

$$(\xi_1, \dots, \xi_r) \in \Omega' \subseteq \mathbb{R}^r. \tag{1.17}$$

Jetzt steht  $\xi_i$  nicht mehr für den Wert, den eine stochastische Variable bei der  $i$ ten Durchführung eines Experimentes annimmt, sondern für die  $i$ te stochastische Variable. Würfelt man beispielsweise mit drei Würfeln, könnte man stochastische Variablen  $\xi_1$ ,  $\xi_2$  und  $\xi_3$  jeweils für die Augenzahl eines Würfels definieren. Führt man ein Experiment mit  $r$  stochastischen Variablen  $N$  mal durch, bezeichnet  $\xi_{ij}$  den Wert, den die  $i$ te stochastische Variable bei der  $j$ ten Durchführung annimmt. Die  $r$ -dimensionale pdf lässt sich dann als Verallgemeinerung zu Gl. (1.10) definieren:

$$f_{\xi_1, \dots, \xi_r}(x_1, \dots, x_r) := \lim_{N \rightarrow \infty} \frac{1}{N} \left[ \sum_{j=1}^N \prod_{i=1}^r \delta(x_i - \xi_{ij}) \right] \tag{1.18}$$

$$= \langle \delta(x_1 - \xi_1) \cdots \delta(x_r - \xi_r) \rangle. \tag{1.19}$$

Der Übersichtlichkeit halber benutzen wir im Folgenden die Abkürzung

$$f_r(x_1, \dots, x_r) := f_{\xi_1, \dots, \xi_r}(x_1, \dots, x_r). \tag{1.20}$$

Wir bezeichnen  $f_r$  als *Verbundwahrscheinlichkeitsdichte* (engl. *joint probability density function*, jpdf).

Die Wahrscheinlichkeit, dass die Realisation eines Experimentes in einem bestimmten Bereich  $Y \subseteq \Omega'$  liegt, ist dann

$$P((\xi_1, \dots, \xi_r) \in Y) = \int_Y f_r(x_1, \dots, x_r) dx_1 \cdots dx_r. \quad (1.21)$$

Der Erwartungswert einer Funktion  $g(\xi_1, \dots, \xi_r)$  ist definiert als

$$\langle g(\xi_1, \dots, \xi_r) \rangle = \int_{\Omega'} g(x_1, \dots, x_r) f_r(x_1, \dots, x_r) dx_1 \cdots dx_r. \quad (1.22)$$

Neben der jpdf ist eine weitere wichtige Größe die bedingte Wahrscheinlichkeitsdichte (engl. *conditional probability density function*, cpdf). Sie gibt die pdf einer stochastischen Variablen an unter der Bedingung, dass die anderen stochastischen Variablen feste Werte annehmen. Bezeichnet man mit  $p(x_1|x_2, \dots, x_r)$  die pdf von  $\xi_1$  an der Stelle  $x_1$  unter der Bedingung, dass die anderen Variablen  $\xi_2, \dots, \xi_r$  die festen Werte  $x_2, \dots, x_r$  annehmen, gilt offenbar der Zusammenhang

$$f_r(x_1, \dots, x_r) = p(x_1|x_2, \dots, x_r) f_{r-1}(x_2, \dots, x_r), \quad (1.23)$$

wobei sich  $f_{r-1}$  aus  $f_r$  über

$$f_{r-1}(x_2, \dots, x_r) = \int f_r(x_1, \dots, x_r) dx_1 \quad (1.24)$$

ergibt. Damit können wir die cpdf definieren als

$$p(x_1|x_2, \dots, x_r) = \frac{f_r(x_1, \dots, x_r)}{\int f_r(x_1, \dots, x_r) dx_1}. \quad (1.25)$$

### 1.3. Schätzen von Wahrscheinlichkeitsdichten

Viele der in dieser Arbeit verwendeten Algorithmen erfordern die Schätzung von Wahrscheinlichkeitsdichten aus Stichproben von Messdaten bzw. Daten aus numerischen Simulationen. Gegeben ist dabei eine Menge von Messpunkten  $\mathbf{X}_1, \dots, \mathbf{X}_N$ , von der angenommen wird, dass sie einer Wahrscheinlichkeitsdichte  $f(\mathbf{x})$  unterliegen. Dabei können die Datenpunkte  $\mathbf{X}_i$  aus einer (univariate Daten) oder mehreren (multivariate Daten) Komponenten bestehen. Ziel ist es, anhand der gegebenen Messpunkte die Wahrscheinlichkeitsdichte  $f_{\mathbf{X}}(\mathbf{x})$  zu schätzen.

Zu diesem Zweck stellt die mathematische Statistik eine Vielzahl an Methoden, sogenannte *Schätzer*, bereit [26, 47, 50]. Neben dem am weitesten verbreiteten Histogrammschätzer, soll an dieser Stelle noch der Kernschätzer dargestellt werden, der gegenüber dem Histogrammschätzer einige Vorteile aufweist.

### 1.3.1. Der Histogrammschätzer

Die einfachste und verbreitetste Methode ist die Erstellung eines Histogramms. Bei einem univariaten Datensatz definiert man dazu Intervalle  $I_m = [x_0 + mh; x_0 + (m + 1)h)$ , sogenannte *Bins*, für ganzzahlige  $m$ . Dabei ist  $x_0$  der willkürlich festzulegende Ursprung und  $h$  bezeichnet die Breite der Bins. Der Histogrammschätzer ist folgendermaßen definiert:

$$\hat{f}(x) = \frac{1}{Nh} \cdot (\text{“Anzahl der Datenpunkte aus dem gleichen Bin wie } x\text{”}). \quad (1.26)$$

Die Breite  $h$  der Bins bestimmt den Grad der Glättung der resultierenden geschätzten Wahrscheinlichkeitsdichte  $\hat{f}(x)$ . Je größer  $h$  gewählt wird, desto “glatter” wird  $\hat{f}(x)$ . Wählt man  $h$  jedoch zu groß, wird möglicherweise eine vorhandene Struktur von  $\hat{f}(x)$  nicht mehr aufgelöst. Wie man  $h$  wählt, um ein optimales Ergebnis zu erzielen, hängt von der Form der realen Wahrscheinlichkeitsdichte ab, die natürlich a priori nicht bekannt ist. Nimmt man eine Gauß-Verteilung an, so kommt man nach [47] auf das Ergebnis

$$h_{\text{opt}} \approx 3.49\hat{\sigma}N^{-1/3} \quad (1.27)$$

für die optimale Binbreite. Dabei ist  $\hat{\sigma}$  ein Schätzer der Standardabweichung  $\sigma$ , wie z. B.

$$\hat{\sigma} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (X_i - \bar{X})^2}. \quad (1.28)$$

Obwohl Gl. (1.27) nur für Gauß-Verteilungen gültig ist, wird sie auch in anderen Fällen häufig als Daumenregel verwendet.

Bei der Verallgemeinerung auf multivariate Datensätze definiert man Bins der Form

$$I_{m_1, \dots, m_d} = [x_0^1 + m_1 h_1; x_0^1 + (m_1 + 1)h_1) \times \dots \times [x_0^d + m_d h_d; x_0^d + (m_d + 1)h_d). \quad (1.29)$$

In diesem Fall ist der Histogrammschätzer definiert als

$$\hat{f}(\mathbf{x}) = \frac{1}{N \prod_{i=1}^d h_i} \cdot (\text{“Anzahl der Datenpunkte aus dem gleichen Bin wie } \mathbf{x}\text{”}). \quad (1.30)$$

Wie man in diesem Fall die Binbreiten  $h_1, \dots, h_d$  optimal wählt, ist dem Verfasser dieser Arbeit nicht bekannt.

### 1.3.2. Der Kernschätzer

Kernschätzer werden im univariaten Fall definiert als

$$\hat{f}(x) = \frac{1}{Nh} \sum_{i=1}^N K\left(\frac{x - X_i}{h}\right). \quad (1.31)$$

Dabei ist die sogenannte *Kernfunktion*  $K(y)$  eine beliebige Funktion mit der Eigenschaft

$$\int_{-\infty}^{\infty} K(y) dy = 1. \quad (1.32)$$

Üblicherweise wird für  $K(y)$  eine symmetrische, nichtnegative Funktion mit genau einem Maximum bei  $y = 0$  gewählt. Stetigkeits- und Differenzierbarkeitseigenschaften der Kernfunktion übertragen sich auch auf  $\hat{f}(x)$ , was ein Vorteil gegenüber dem Histogrammschätzer ist, der als Stufenfunktion definiert ist.

Die am meisten verwendete Kernfunktion ist der *Epanechnikov-Kern*

$$K(y) = \begin{cases} \frac{3\sqrt{5}}{100}(5 - y^2) & \text{für } y^2 \leq 5 \\ 0 & \text{sonst} \end{cases}, \quad (1.33)$$

da er effizient zu berechnen ist und ein gutes Konvergenzverhalten zeigt [26]. Viel entscheidender als die Wahl der Kernfunktion ist nach [26] jedoch die Wahl der sogenannten *Bandbreite*  $h$ . Wie die Binbreite beim Histogrammschätzer, entscheidet die Bandbreite beim Kernschätzer über den Grad der Glättung. Es existiert eine Vielzahl an Methoden, die optimale Bandbreite für einen gegebenen Datensatz zu bestimmen. In dieser Arbeit wird die Daumenregel von Silverman [26]

$$h_{\text{opt}} \approx 1.06\hat{\sigma}N^{-1/5}, \quad (1.34)$$

mit  $\hat{\sigma}$  aus Gl. (1.28) verwendet.

Die natürliche Erweiterung des Kernschätzers auf multivariate Daten ist

$$\hat{f}(\mathbf{x}) = \frac{1}{N \prod_{k=1}^d h_k} \sum_{i=1}^N \prod_{j=1}^d K\left(\frac{x_j - X_i^j}{h_j}\right). \quad (1.35)$$

Auch für diesen Fall gibt es eine Daumenregel für die optimale Wahl der Bandbreiten  $h_j$  [26]:

$$h_{j,\text{opt}} \approx \hat{\sigma}_j N^{-1/(d+4)}. \quad (1.36)$$

Dabei ist  $\hat{\sigma}_j$  die Standardabweichung der  $j$ ten Komponente des Datensatzes.

Eine Verbesserung des Schätzers (1.35) ist

$$\hat{f}(\mathbf{x}) = \frac{1}{N \det \mathbf{H}} \sum_{i=1}^N \mathcal{K}\{\mathbf{H}^{-1}(\mathbf{x} - \mathbf{X}_i)\} \quad (1.37)$$

mit

$$\mathcal{K}\{\mathbf{H}^{-1}(\mathbf{x} - \mathbf{X})\} = \prod_{j=1}^d K\{[\mathbf{H}^{-1}(\mathbf{x} - \mathbf{X})]_j\}, \quad (1.38)$$

wobei  $\mathbf{H}$  die *Bandbreitematrix* ist. Für die optimale Bandbreitematrix existiert die Daumenregel [26]

$$\mathbf{H}_{\text{opt}} \approx \hat{\Sigma}^{1/2} N^{-1/(d+4)} \quad (1.39)$$

mit

$$\{\hat{\Sigma}^{1/2}\}_{ij} = \langle (X_i - \bar{X}_i)(X_j - \bar{X}_j) \rangle^{1/2}. \quad (1.40)$$

Die Verwendung dieser Bandbreitematrix führt auf das gleiche Ergebnis, das man erreichen würde, würde man den Datensatz derart transformieren, dass die Matrix (1.39) Diagonalgestalt hätte, und anschließend den Schätzer (1.35) benutzte und dabei als Bandbreiten  $h_j$  die Diagonalelemente der transformierten Matrix (1.39) verwendete. Der Schätzer (1.37) berücksichtigt daher im Gegensatz zu (1.35) Korrelationen zwischen den Komponenten von  $\mathbf{X}$ . Letzterer entspricht einem Spezialfall von (1.37), wenn man als Bandbreitematrix

$$\mathbf{H} = \text{diag}\{h_1, \dots, h_d\} \quad (1.41)$$

verwendet.

#### 1.4. Zeitabhängige Zufallsvariablen

Wir betrachten nun stochastische Variablen, deren Realisationen von der Zeit abhängen, d. h.  $\xi = \xi(t)$ . Eine Realisation der Variablen entspricht dann einer Kurve, wie in Abb. 1.1 dargestellt wird. Die zeitliche Entwicklung einer stochastischen Variable bezeichnet man als stochastischen Prozess. Analog zu (1.10) können wir die zeitabhängige pdf definieren als

$$f_{\xi}(x, t) = \langle \delta(x - \xi(t)) \rangle \quad (1.42)$$

und die  $n$ -Punkt jpdf als

$$f_n(x_1, t_1; \dots; x_n, t_n) = \langle \delta(x_1 - \xi(t_1)) \cdots \delta(x_n - \xi(t_n)) \rangle. \quad (1.43)$$

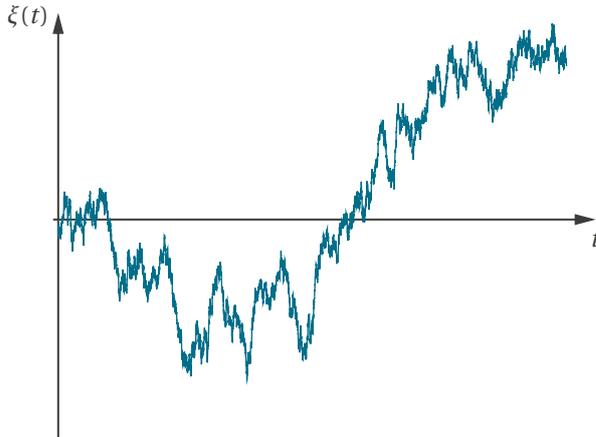
Die cpdf

$$p(x_1, t_1 | x_2, t_2) = \frac{f_2(x_1, t_1; x_2, t_2)}{f_1(x_2, t_2)} \quad (1.44)$$

wird für  $t_1 > t_2$  auch *Übergangswahrscheinlichkeitsdichte* genannt, da

$$p(x_1, t_1 | x_2, t_2) dx_1$$

die Wahrscheinlichkeit angibt, die Realisation von  $\xi$  zur Zeit  $t_1$  im Intervall  $[x_1, x_1 + dx_1]$  zu finden unter der Voraussetzung, dass zur Zeit  $t_2$  der Wert  $x_2$  realisiert wurde.



**Abb. 1.1** Beispiel einer Realisation eines stochastischen Prozesses.

Ein stochastischer Prozess ist durch die Angabe aller möglichen  $n$ -Punkt jpdfs vollständig charakterisiert. Es lassen sich dann alle zeitabhängigen Momente und Korrelationsfunktionen berechnen:

$$\langle \xi(t_1) \cdots \xi(t_n) \rangle = \int x_1(t_1) \cdots x_n(t_n) f_n(x_1, t_1; \dots; x_n, t_n) dx_1 \cdots dx_n . \quad (1.45)$$

Ein stochastischer Prozess ist *stationär*, falls alle  $n$ -Punkt jpdfs der Relation

$$f_n(x_1, t_1; \dots; x_n, t_n) = f_n(x_1, t_1 + \Delta t; \dots; x_n, t_n + \Delta t) \quad (1.46)$$

für beliebige  $\Delta t$  genügen.

## 1.5. Markov-Prozesse

Eine wichtige Unterklasse stochastischer Prozesse stellen sogenannte *Markov-Prozesse* dar. Ein Markov-Prozess ist ein stochastischer Prozess mit der Eigenschaft, dass für beliebige  $n$  und beliebige  $t_1 > t_2 > \dots > t_n$  gilt:

$$p_{1|n-1}(x_1, t_1 | x_2, t_2; \dots; x_n, t_n) = p_{1|1}(x_1, t_1 | x_2, t_2) . \quad (1.47)$$

Das heißt, die Wahrscheinlichkeit, dass der Zustand  $x_1$  zur Zeit  $t_1$  realisiert wird, unter der Bedingung, dass zur früheren Zeit  $t_2$  der Zustand  $x_2$  realisiert wurde, hängt nicht davon ab, welche Zustände zu Zeiten  $t_i < t_2$  realisiert wurden. In diesem Sinne bezeichnet man einen Markov-Prozess als einen Prozess ohne "Gedächtnis", da er sich an die Zeiten  $t_i < t_2$  nicht "erinnern" kann.

Mit der Eigenschaft (1.47) lässt sich die  $n$ -Punkt jpdf schreiben als

$$f_n(x_1, t_1; \dots; x_n, t_n) = \left[ \prod_{i=1}^{n-1} p_{1|1}(x_i, t_i | x_{i+1}, t_{i+1}) \right] f_1(x_n, t_n) . \quad (1.48)$$

Sind also die pdf  $f_1$  sowie die Übergangswahrscheinlichkeitsdichten  $p_{1|1}$  bekannt, ergeben sich daraus alle möglichen  $n$ -Punkt pdfs, was die Beschreibung eines Markov-Prozesses erheblich vereinfacht.

Als Konsequenz aus der Markov-Eigenschaft ergibt sich die Chapman-Kolmogorov-Gleichung (CKG)

$$p(x_1, t_1 | x_3, t_3) = \int p(x_1, t_1 | x_2, t_2) p(x_2, t_2 | x_3, t_3) dx_2, \quad (1.49)$$

auf deren Beweis hier verzichtet wird (s. dazu [45]).

## 1.6. Langevin-Gleichungen

Das berühmteste Beispiel der Natur für einen Markov-Prozess ist die Brownsche Bewegung [8], deren Bewegungsgleichung

$$\dot{v}(t) = -\gamma v(t) + \sqrt{2q}\xi(t) \quad (1.50)$$

lautet. Dabei ist  $v(t)$  eine Geschwindigkeitskomponente eines in einer ruhenden Flüssigkeit suspendierten Teilchens zur Zeit  $t$ . Der erste Term der rechten Handseite steht für die Reibung nach dem Gesetz von Stokes. Der zweite Term modelliert die durch die unregelmäßigen Stöße mit den Molekülen der Flüssigkeit hervorgerufene Kraft.  $\xi(t)$  ist eine zeitabhängige, gaußverteilte stochastische Variable mit den Eigenschaften

$$\langle \xi(t) \rangle = 0 \quad (1.51a)$$

$$\langle \xi(t)\xi(t') \rangle = \delta(t - t'). \quad (1.51b)$$

Der Parameter  $q$  in Gl. (1.50) ist proportional zur Temperatur der Flüssigkeit, die die Stärke der Fluktuationen bestimmt. Man bezeichnet  $\xi(t)$  als *Langevin-Kraft* und Bewegungsgleichungen, in denen Langevin-Kräfte vorkommen, als *Langevin-Gleichungen*.

Da Langevin-Kräfte  $\delta$ -korreliert sind, beschreiben Lösungen von Langevin-Gleichungen Markov-Prozesse. Bei realen physikalischen Prozessen gibt es jedoch in der Regel eine endliche Korrelationszeit. Im Falle der Brownschen Bewegung ist diese durch die recht kurze Dauer eines Stoßes gegeben. Die Markov-Eigenschaft (1.47) gilt dann nur auf Zeitskalen, die groß im Vergleich zur Korrelationszeit sind.

Die allgemeine Form einer einkomponentigen Langevin-Gleichung lautet<sup>1</sup> [45]

$$\dot{X} = h(X, t) + \sqrt{2}g(X, t)\xi(t), \quad (1.52)$$

<sup>1</sup>Selbstverständlich kann der Faktor  $\sqrt{2}$  auch in die Funktion  $g$  hineingezogen werden. Dann ergibt sich ein entsprechend anderer Zusammenhang zu den Kramers-Moyal-Koeffizienten, s. Gln. (1.60). In anderen Darstellungen, wie auch in [45], wird der Faktor in die Normierung der Langevin-Kraft geschoben, so dass anstelle von (1.51b)  $\langle \xi(t)\xi(t') \rangle = 2\delta(t - t')$  gilt.

wobei  $\xi(t)$  die oben erwähnten Eigenschaften besitzt. Hängt die Funktion  $g$  nicht von  $X$  ab, spricht man von einer Gleichung mit additivem Rauschen, sonst von multiplikativem Rauschen.

Natürlich kann man stochastische Differentialgleichungen der Form (1.52) nicht lösen in dem Sinne, dass man zu einer bestimmten Anfangsbedingung die Lösung  $X(t)$  berechnet, da diese von der Realisation der Zufallsvariablen  $\xi(t)$  abhängt, die per definitionem nicht vorhersagbar ist. Wir können also anhand einer Langevin-Gleichung nicht vorhersagen, welchen Wert die Variable  $X$  zu einer bestimmten Zeit  $t$  haben wird. Es ist jedoch möglich, Wahrscheinlichkeiten für bestimmte Ereignisse oder Erwartungswerte bestimmter Größen anzugeben. Erstere wie Letztere lassen sich aus den  $n$ -Punkt pdfs  $f_n(x_1, t_1; \dots; x_n, t_n)$  berechnen. Da diese, wie im letzten Abschnitt gezeigt, bei Markov-Prozessen durch die cpdfs  $p(x_i, t_i | x_{i+1}, t_{i+1})$  sowie eine gegebene Anfangsverteilung  $f_1(x_n, t_n)$  ausgedrückt werden können, lautet das Ziel, die cpdfs zu berechnen. Diese sind Lösungen der *Fokker-Planck-Gleichung*, die im folgenden Abschnitt behandelt wird.

## 1.7. Die Fokker-Planck-Gleichung

Die Fokker-Planck-Gleichung (FPG) ist eine partielle Differentialgleichung für die Zeitentwicklung von Wahrscheinlichkeitsdichten. Ziel dieses Abschnittes ist es, die Herleitung der FPG zu skizzieren und dabei den Zusammenhang zur Langevin-Gleichung herzustellen. Außerdem wird die stationäre Lösung der FPG berechnet. Eine ausführlichere Darstellung sowie Methoden zur Lösung von FPGen finden sich in [45].

### 1.7.1. Herleitung

Um eine Gleichung für die Zeitentwicklung der pdf  $f_X(x, t)$  der Form

$$\frac{\partial}{\partial t} f_X(x, t) = \text{“Funktion bekannter Größen”} \quad (1.53)$$

zu entwickeln, starten wir mit der Relation

$$f_X(x, t + \tau) = \int p(x, t + \tau | x', t) f_X(x', t) dx', \quad (1.54)$$

deren Beweis direkt durch Einsetzen von (1.44) folgt. Unter Verwendung der bedingten Momente

$$\begin{aligned} M_n(x', t, \tau) &= \left\langle [X(t + \tau) - X(t)]^n \right\rangle_{X(t)=x'} \\ &= \int (x - x')^n p(x, t + \tau | x', t) dx \end{aligned} \quad (1.55)$$

lässt sich nach [45] folgende Entwicklung für die Übergangswahrscheinlichkeitsdichte in (1.54) ableiten:

$$p(x, t + \tau | x', t) = \left[ 1 + \sum_{n=1}^{\infty} \frac{1}{n!} \left( -\frac{\partial}{\partial x} \right)^n M_n(x, t, \tau) \right] \delta(x - x'). \quad (1.56)$$

Setzt man diese Entwicklung in (1.54) ein, so ergibt sich

$$f_X(x, t + \tau) = f_X(x, t) + \sum_{n=1}^{\infty} \frac{1}{n!} \left( -\frac{\partial}{\partial x} \right)^n M_n(x, t, \tau) f_X(x, t) . \quad (1.57)$$

Damit folgt unmittelbar die *Kramers-Moyal-Entwicklung*

$$\begin{aligned} \frac{\partial}{\partial t} f_X(x, t) &= \lim_{\tau \rightarrow 0} \frac{f_X(x, t + \tau) - f_X(x, t)}{\tau} \\ &= \sum_{n=1}^{\infty} \left( -\frac{\partial}{\partial x} \right)^n D^{(n)}(x, t) f_X(x, t) , \end{aligned} \quad (1.58)$$

wobei mit  $D^{(n)}(x, t)$  die *Kramers-Moyal-Koeffizienten*

$$\begin{aligned} D^{(n)}(x, t) &= \frac{1}{n!} \lim_{\tau \rightarrow 0} \frac{1}{\tau} M_n(x, t, \tau) \\ &= \frac{1}{n!} \lim_{\tau \rightarrow 0} \frac{1}{\tau} \left\langle [X(t + \tau) - X(t)]^n \right\rangle_{X(t)=x} \end{aligned} \quad (1.59)$$

bezeichnet werden.

Falls der stochastische Prozess  $X(t)$  einer Langevin-Gleichung der Form (1.52) gehorcht, lassen sich die Kramers-Moyal-Koeffizienten explizit berechnen. Nach [45] ergibt sich folgender Zusammenhang<sup>2</sup>:

$$D^{(1)}(x, t) = h(x, t) \quad (1.60a)$$

$$D^{(2)}(x, t) = g^2(x, t) \quad (1.60b)$$

$$D^{(n)}(x, t) = 0 \quad \text{für} \quad n \geq 3 . \quad (1.60c)$$

Demzufolge kann der erste Kramers-Moyal-Koeffizient mit dem deterministischen Teil der Langevin-Gleichung (1.52) identifiziert werden und der zweite mit der Amplitude des Rauschterms. Daher bezeichnet man  $D^{(1)}(x, t)$  als *Drift* und  $D^{(2)}(x, t)$  als *Diffusion*.

Aufgrund von (1.60c) bricht bei einem Langevin-Prozess die Kramers-Moyal-Entwicklung nach dem zweiten Glied ab und es ergibt sich die Fokker-Planck-Gleichung

$$\frac{\partial}{\partial t} f_X(x, t) = \left[ -\frac{\partial}{\partial x} D^{(1)}(x, t) + \frac{\partial^2}{\partial x^2} D^{(2)}(x, t) \right] f_X(x, t) . \quad (1.61)$$

Es sei betont, dass die Differentialoperatoren sowohl auf die Kramers-Moyal-Koeffizienten als auch auf die pdf wirken.

Unter der Berücksichtigung spezifischer Randbedingungen gibt die FPG an, wie sich eine gegebene Anfangs-pdf mit der Zeit entwickelt. Die cpdf  $p(x, t | x_0, t_0)$  ist eine spezielle Lösung der FPG zu der Anfangsbedingung  $f_X(x, t_0) = \delta(x - x_0)$ .

<sup>2</sup>Dabei wird die Definition stochastischer Integrale nach Itô verwendet (s. [45]). Mit der Definition nach Stratonovich erhält der erste Koeffizient einen zusätzlichen Term:  $D^{(1)}(x, t) = h(x, t) + \frac{\partial g(x, t)}{\partial x} g(x, t)$ . Dieser wird als *rausch-induzierte Drift* bezeichnet und verschwindet im Falle von additivem Rauschen.

### 1.7.2. Stationäre Lösung

Sind die Kramers-Moyal-Koeffizienten zeitunabhängig, lässt sich eine einfache stationäre Lösung der FPG im eindimensionalen Fall angeben. Diese Lösung lässt sich leicht ableiten, indem man die FPG als Kontinuitätsgleichung formuliert. Führt man dazu den Strom

$$S(x, t) = D^{(1)}(x) f_X(x, t) - \frac{\partial}{\partial x} D^{(2)}(x) f_X(x, t) \quad (1.62)$$

ein, so lässt sich die FPG formulieren als

$$\frac{\partial}{\partial t} f_X(x, t) + \frac{\partial}{\partial x} S(x, t) = 0. \quad (1.63)$$

Für die stationäre Lösung soll gelten:

$$\frac{\partial}{\partial t} f_X^{\text{stat}}(x, t) = 0. \quad (1.64)$$

Damit folgt aus Gl. (1.63)

$$S(x) = \text{const}. \quad (1.65)$$

Ist der Strom an irgendeiner Stelle null, was insbesondere bei natürlichen Randbedingungen gegeben ist, folgt weiter  $S(x) = 0$ . Dann ergibt sich mit Gl. (1.62)

$$\frac{\partial}{\partial x} D^{(2)}(x) f_X^{\text{stat}}(x) = D^{(1)}(x) f_X^{\text{stat}}(x) = \frac{D^{(1)}(x)}{D^{(2)}(x)} D^{(2)}(x) f_X^{\text{stat}}(x). \quad (1.66)$$

Dies lässt sich leicht integrieren mit dem Ergebnis

$$f_X^{\text{stat}}(x) = \frac{\mathcal{N}}{D^{(2)}(x)} \exp \left\{ \int^x \frac{D^{(1)}(x')}{D^{(2)}(x')} dx' \right\}. \quad (1.67)$$

Dabei ist  $\mathcal{N}$  eine Normierungskonstante, in der bereits die Integrationskonstante des Exponenten enthalten ist.

## 1.8. Schätzen von Drift und Diffusion

Die Kramers-Moyal-Koeffizienten

$$D^{(n)}(x, t) = \frac{1}{n!} \lim_{\tau \rightarrow 0} \frac{1}{\tau} \int (x - x')^n p(x, t + \tau | x', t) dx \quad (1.68a)$$

$$= \frac{1}{n!} \lim_{\tau \rightarrow 0} \frac{1}{\tau} \left\langle [X(t + \tau) - X(t)]^n \right\rangle_{X(t)=x} \quad (1.68b)$$

lassen sich bei Prozessen, die einer Langevin-Dynamik gehorchen, aus einer gegebenen Zeitreihe schätzen, wie erstmalig in [49] ausführlich dargestellt wurde. Der darin beschriebene Algorithmus besteht darin, die cpdf im Integral von (1.68a) mit Hilfe von Histogrammen aus dem Datensatz zu schätzen und anschließend das Integral numerisch auszuwerten. Erstmalig angewandt wurde diese Methode bei der Untersuchung statistischer Eigenschaften der turbulenten Kaskade in [10]. Mittlerweile existiert eine Vielzahl an Veröffentlichungen, die Methode selbst und Weiterentwicklungen betreffend [2, 12, 28–30, 34–36, 41, 55], sowie zu Anwendungen [11, 16, 18, 32]. In diesem Abschnitt wird lediglich die grundlegende Methode so dargestellt, wie sie in dieser Arbeit verwendet wird.

### 1.8.1. Die generelle Methode

Anstatt das Integral von (1.68a) auszuwerten, lässt sich auch der bedingte Mittelwert in (1.68b) direkt aus den Daten schätzen, wie im Folgenden dargestellt wird. Wir gehen davon aus, dass wir über die Daten von  $N$  Realisationen  $X_1(t), \dots, X_N(t)$  des stochastischen Prozesses  $X(t)$  verfügen, wobei jede Realisation aus  $M$  Datenpunkten für die Zeiten  $t = i\Delta t$  mit  $i = 0, 1, \dots, M - 1$  besteht. Wir unterteilen nun die Menge der möglichen Realisationen von  $X(t)$  in Bins

$$I_j = [x_0 + (j - 1)\Delta x, x_0 + j\Delta x]$$

mit  $j = 1, \dots, K$ . Dabei sind der Ursprung  $x_0$ , sowie die Breite  $\Delta x$  der Bins zunächst willkürlich festzulegen. Führen wir die *Finite-Time-Koeffizienten*

$$\tilde{D}_\tau^{(n)}(x, t) = \frac{1}{n!} \frac{1}{\tau} \left\langle [X(t + \tau) - X(t)]^n \right\rangle_{X(t)=x} \quad (1.69)$$

ein, so können wir dafür den folgenden Schätzer definieren:

$$\hat{D}_m^{(n)}(x \in I_j, t) = \frac{\sum_{\ell=1}^N [X_\ell(t + m\Delta t) - X_\ell(t)]^n \Theta_j(X_\ell(t))}{n! m \Delta t \sum_{\ell=1}^N \Theta_j(X_\ell(t))}. \quad (1.70)$$

Dabei ist

$$\Theta_j(x) = \begin{cases} 1, & \text{für } x \in I_j, \\ 0 & \text{sonst.} \end{cases} \quad (1.71)$$

Das Zeitinkrement  $m\Delta t$  aus (1.70) entspricht der Größe  $\tau$  aus (1.68). Um den Grenzübergang  $\tau \rightarrow 0$  zu vollziehen, kann (1.70) für verschiedene  $m$  berechnet werden. Den Schätzer für die Kramers-Moyal-Koeffizienten erhält man dann durch Extrapolation für  $m \rightarrow 0$ :

$$\hat{D}_m^{(n)}(x \in I_j, t) \xrightarrow{m \rightarrow 0} \hat{D}^{(n)}(x \in I_j, t). \quad (1.72)$$

Welche Art von Extrapolationsverfahren sinnvollerweise gewählt wird, muss im Einzelfall entschieden werden. Häufig reicht es aus, den Finite-Time-Koeffizienten für  $m = 1$  zu verwenden.

Die numerische Berechnung des Schätzers (1.70,1.72) liefert eine Stufenfunktion, die anschließend bei Bedarf durch einen analytischen Ausdruck approximiert werden kann.

### 1.8.2. Stationäre, nicht stationäre und periodische Prozesse

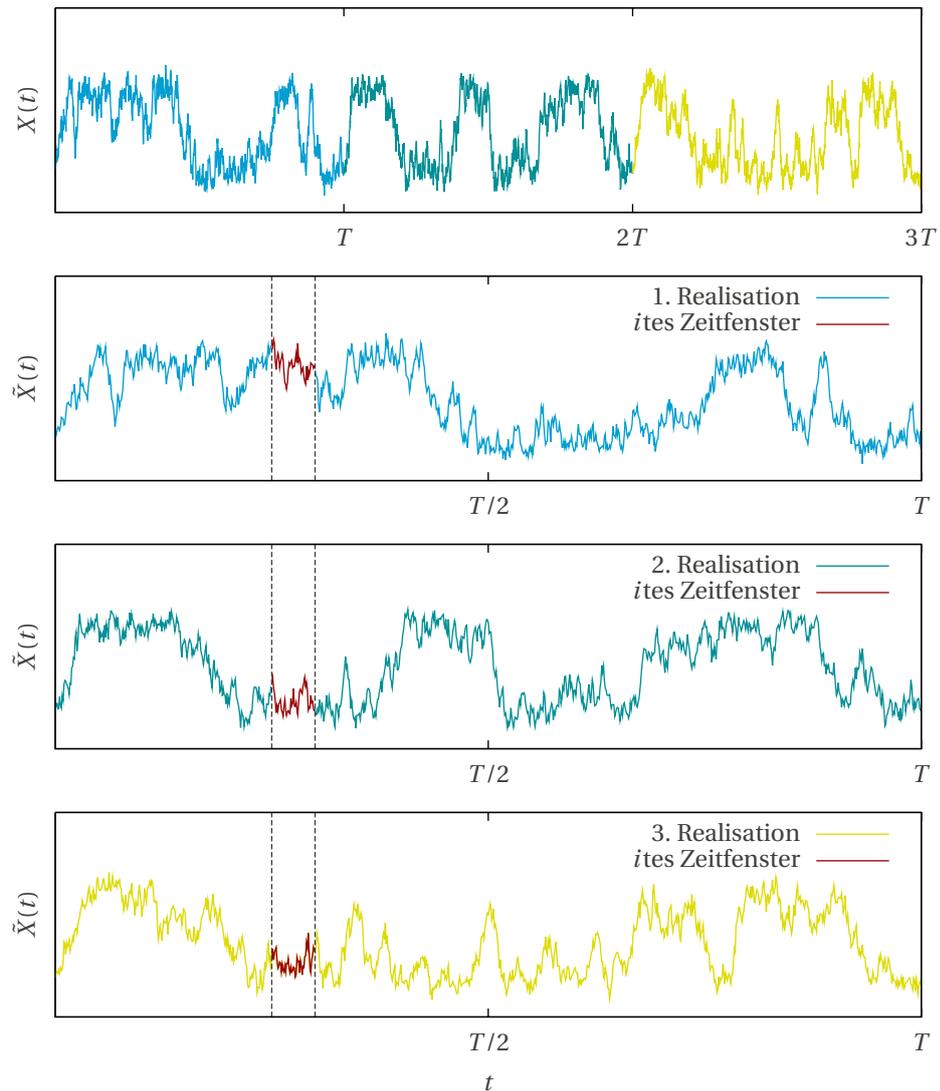
Ist der betrachtete Prozess stationär und ergodisch, so hängen die Kramers-Moyal-Koeffizienten nicht von der Zeit ab und das Ensemblemittel kann durch ein Zeitmittel ersetzt werden. Anstelle von Gl. (1.70) lautet dann der entsprechende Schätzer:

$$\hat{D}_m^{(n)}(x \in I_j) = \frac{\sum_{\ell=1}^N [X(t_\ell + m\Delta t) - X(t_\ell)]^n \Theta_j(X(t_\ell))}{n!m\Delta t \sum_{\ell=1}^N \Theta_j(X(t_\ell))}. \quad (1.73)$$

Ist der Prozess nicht stationär und verfügt man nicht über ein Ensemble von Realisationen (oder über kein hinreichend großes), muss die gemessene Zeitreihe in Zeitfenster unterteilt werden. Man kann dann jedes Zeitfenster als einen von den anderen unabhängigen stationären Prozess betrachten und die Schätzung wie oben beschrieben für jeden dieser Prozesse durchführen. Dabei müssen die Zeitfenster einerseits klein genug gewählt werden, so dass man näherungsweise von Stationarität innerhalb eines Zeitfensters ausgehen kann, andererseits aber hinreichend groß, damit die Statistik auskonvergiert. In diesem Fall hängt die Anwendbarkeit des Verfahrens letztlich von der Samplingfrequenz der verfügbaren Daten ab.

Bei stochastischen Prozessen mit periodischem Treiber (ein solcher wird in Kapitel 3 untersucht) nehmen Drift und Diffusion nach einer bestimmten "Einschwingphase" die Periodizität des Treibers an. In diesem Fall kann die Schätzung so durchgeführt werden, als entspräche jede gemessene Periode des Prozesses  $X(t)$  einer Realisation *eines* (nicht stationären) Prozesses  $\tilde{X}(t)$ . Alle gemessenen Perioden zusammen ergeben dann ein Ensemble dieses Prozesses, so dass, wie oben beschrieben, Ensemblemittel geschätzt werden können (vgl. Abb. 1.2).

Ist die Periodenlänge des Treibers groß und ist dieser nur langsam veränderlich, kann jede Periode in Zeitfenster unterteilt werden. Man betrachtet dann das *ite* Zeitfenster einer Periode (die roten Bereiche in Abb. 1.2) als eine Realisation des *iten* stationären Prozesses  $\tilde{X}^{(i)}(t)$ . Die *iten* Zeitfenster aller gemessenen Perioden bilden zusammen ein Ensemble des *iten* Prozesses. Bei der Bestimmung von Drift und Diffusion für diesen Prozess können dann Ensemble- und Zeitmittel kombiniert werden. Drift- und Diffusionsschätzung von Prozessen mit periodischem Treiber werden ausführlich in [55] diskutiert.



**Abb. 1.2** Drift- und Diffusionsschätzung bei Prozessen mit periodischem Treiber. Jede Periode eines Prozesses  $X(t)$  mit periodischem Treiber  $F(t) = F(t + T)$  (oben) wird als Realisation *eines* Prozesses  $\tilde{X}(t)$  interpretiert. Diese Realisationen (die unteren drei Diagramme) bilden zusammen ein Ensemble des Prozesses  $\tilde{X}(t)$ . Aus diesem Ensemble können die Kramers-Moyal-Koeffizienten dann mit dem Schätzer (1.70) und (1.72) bestimmt werden. Bei großer Periodenlänge und einem langsam veränderlichen Treiber kann jede Realisation des Prozesses  $\tilde{X}(t)$  in Zeitfenster unterteilt werden. Jedes dieser Zeitfenster wird dann als Realisation eines stationären Prozesses interpretiert. So bilden die roten Bereiche zusammen ein Ensemble des näherungsweise stationären Prozesses  $\tilde{X}^{(i)}(t)$ . Bei der Schätzung der Kramers-Moyal-Koeffizienten des Prozesses  $\tilde{X}^{(i)}(t)$  können dann Ensemble- und Zeitmittel kombiniert werden.

## 2. Synergetik

Dieses Kapitel widmet sich einigen wichtigen Grundlagen aus der Synergetik, die in den darauf folgenden Kapiteln von zentraler Bedeutung sind. Nach einer kurzen Einführung werden in Abschnitt 2.2 zunächst einige für das Weitere unverzichtbare Begriffe aus der Theorie der nichtlinearen Dynamik eingeführt. Die folgenden Abschnitte behandeln dann zwei zentrale Prinzipien der Synergetik, nämlich das Versklavungsprinzip und das Prinzip der unvoreingenommenen Schätzung. Der letzte Abschnitt über die Information von Systemen in der Nähe von Instabilitäten bildet die theoretische Grundlage für eine der beiden in Kapitel 4 entwickelten Methoden zur Identifikation von Ordnungsparametern.

### 2.1. Einführung

Der Begriff *Synergie* leitet sich aus dem griechischen *συνεργία* ab, was sich als “Zusammenarbeit” oder “Zusammenwirken” übersetzen lässt. Unter dem Begriff *Synergetik* versteht man somit “die Lehre vom Zusammenwirken”. Während der Begriff gelegentlich auch in eher esoterischen Denkmodellen auftritt, ist hier die in den 1970er Jahren von Hermann Haken [23–25] entwickelte, interdisziplinäre Theorie der Entstehung von Selbstorganisation in komplexen Systemen gemeint.

#### 2.1.1. Komplexe Systeme und Selbstorganisation

Unter *komplexen Systemen* seien hier offene Systeme verstanden, die sich fernab vom thermodynamischen Gleichgewichtszustand befinden und darüber hinaus aus einer großen Anzahl von Untersystemen bestehen, die miteinander auf mehr oder weniger komplexe Art und Weise wechselwirken. *Selbstorganisation* bezeichnet die spontane Herausbildung von räumlichen, zeitlichen, raumzeitlichen oder funktionellen Strukturen, also geordneten Zuständen, ohne dass diese dem System von außen aufgeprägt werden [21].

Beispiele für Selbstorganisation in komplexen Systemen findet man in vielen verschiedenen Wissenschaftsdisziplinen. Aus dem Bereich der Physik seien hier der Laser [23] und die Rayleigh-Bénard-Konvektion [9] genannt, aus der Chemie die Belousov-Zhabotinsky-Reaktion [46] und aus der belebten Natur die Bildung von Fischschwärmen sowie

von Mustern in Bakterienkolonien [4] oder Hirnstromaktivitäten [20]. Aber auch jenseits der Naturwissenschaften gibt es Beispiele, wie die Bildung öffentlicher Meinungen aus der Soziologie [51] oder die Selbstorganisation in Städten aus der Geographie [44].

### 2.1.2. Nichtgleichgewichts-Phasenübergänge

In Systemen, die das Phänomen der Selbstorganisation zeigen, gibt es meist bestimmte kritische Werte von Kontrollparametern, die eine Grenze zwischen einem geordneten und einem ungeordneten Zustand darstellen. Im Beispiel der Rayleigh-Bénard-Konvektion ist ein solcher Kontrollparameter die Temperaturdifferenz  $\Delta T$  zwischen zwei horizontalen Platten, zwischen denen sich eine Flüssigkeiten befindet. Oberhalb eines bestimmten kritischen Wertes  $\Delta T_c$  bilden sich spontan typische Konvektionsrollen heraus, die unterhalb von  $\Delta T_c$  nicht vorhanden sind. Solche qualitativen Änderungen des Systemzustands auf makroskopischer Skala bezeichnet man in offenen Systemen als *Nichtgleichgewichts-Phasenübergänge*.

### 2.1.3. Was leistet die Theorie?

Das Ziel der Synergetik besteht darin, vereinigende Prinzipien komplexer, selbstorganisierender Systeme zu entdecken [21]. Das wichtigste Prinzip, das die Theorie dabei hervorgebracht hat, ist das Ordnungsparameterkonzept. Dieses erklärt die spontane Entstehung geordneter Strukturen dadurch, dass wenige kollektive Größen, sogenannte *Ordnungsparameter*, die Dynamik dominieren und die verbleibenden Freiheitsgrade des Systems unterwerfen und somit dem System eine gewisse Ordnung aufzwingen. Das Ordnungsparameterkonzept ermöglicht so eine Beschreibung komplexer Systeme mit einer drastisch reduzierten Anzahl von Freiheitsgraden. Die beobachteten Nichtgleichgewichts-Phasenübergänge treten in den Ordnungsparametergleichungen in Form von Instabilitäten auf.

Die Theorie beinhaltet zwei mögliche Herangehensweisen, die Evolutionsgleichungen, denen die Ordnungsparameter gehorchen, bei Systemen in der Nähe von Nichtgleichgewichts-Phasenübergängen abzuleiten.

### 2.1.4. Von unten nach oben oder andersherum

Die erste Herangehensweise setzt voraus, dass die Modellgleichungen auf mikroskopischer Ebene bekannt sind. In diesem Fall bietet das Versklavungsprinzip einen mathematischen Formalismus, die Ordnungsparametergleichungen aus den mikroskopischen Modellgleichungen abzuleiten. Dieses Prinzip beruht auf einer Hierarchie der typischen Zeitskalen zwischen der Ordnungsparameterdynamik und der Dynamik der verbleibenden Freiheitsgrade. Diese Hierarchie ist in der Nähe von Nichtgleichgewichts-Phasenübergängen gegeben. Da man durch das Versklavungsprinzip von einer mikroskopischen zu einer makroskopischen Beschreibung des System gelangt, bezeichnet man diese Herangehensweise auch als *Bottom-Up-Ansatz*.

Sind die mikroskopischen Wechselwirkungen nicht bekannt oder derart komplex, so dass ein Bottom-Up-Ansatz aussichtslos erscheint (wie z.B. beim menschlichen Gehirn), eignet sich die zweite Herangehensweise. Das *Prinzip der unvoreingenommenen Schätzung*<sup>1</sup>, das 1957 von E. T. Jaynes in die statistische Mechanik eingeführt wurde [27], erlaubt es, aus Messungen makroskopischer Observablen phänomenologische Ordnungsparametermodelle abzuleiten, aus denen dann möglicherweise Rückschlüsse auf die Wechselwirkungen auf mikroskopischer Ebene gezogen werden können. Dementsprechend bezeichnet man diese Herangehensweise als *Top-Down-Ansatz*.

Um die nähere Darstellung dieser beiden Ansätze etwas zu entlasten, werden zunächst einige Grundbegriffe der nichtlinearen Dynamik eingeführt.

## 2.2. Einige Grundbegriffe der nichtlinearen Dynamik

Die Darstellung dieses Abschnittes orientiert sich an den einführenden Lehrbüchern [52], [1] und [19].

### 2.2.1. Dynamische Systeme

Untersuchungsgegenstand der nichtlinearen Dynamik sind Systeme, die durch einen Satz von Größen  $q_i$ ,  $i = 1, \dots, N$ , beschrieben werden können, die sich mit der Zeit  $t$  ändern, d. h.  $q_i = q_i(t)$ . Handelt es sich um räumlich ausgedehnte Systeme, können die Größen  $q_i$  zusätzlich vom Ort  $\mathbf{r}$  abhängen. Dieser Fall wird hier jedoch nicht behandelt.

Der Zustand des Systems zu einem Zeitpunkt  $t$  sei durch Angabe des *Zustandsvektors*

$$\mathbf{q}(t) = (q_1(t), \dots, q_N(t))^T \quad (2.1)$$

eindeutig bestimmt. Der Zustandsvektor beschreibt für einen festen Zeitpunkt  $t$  einen Punkt im  $N$ -dimensionalen *Zustandsraum*. Die Menge solcher Punkte für alle Zeiten  $t \in [t_0, \infty)$  bezeichnet man als *Trajektorie*.

Die Rate, mit der sich  $\mathbf{q}(t)$  ändert, also die zeitliche Ableitung  $\dot{\mathbf{q}}(t)$ , kann sowohl von den Komponenten  $q_i(t)$  als auch durch äußere Einflüsse explizit von der Zeit abhängen. Ist Letzteres nicht der Fall, bezeichnet man das System als *autonom*. Nichtautonome Systeme lassen sich jedoch mathematisch immer durch Einführen einer neuen Komponente  $q_{N+1} = t$  in ein autonomes System überführen. Wir können die zeitliche Entwicklung des Zustandsvektors also beschreiben durch

$$\dot{\mathbf{q}}(t) = \mathbf{f}(q_1(t), \dots, q_N(t), t). \quad (2.2)$$

Gl. (2.2) bezeichnet man als *dynamisches System*.

<sup>1</sup>In der überwiegend englischsprachigen Literatur werden Begriffe wie *maximum information principle* [25] oder *maximum-entropy principle* [27] verwendet. Hier wird der Begriff *Prinzip der unvoreingenommenen Schätzung* aus [22] übernommen.

### 2.2.2. Lineare, autonome dynamische Systeme

Wir betrachten zunächst lineare, autonome Systeme. Dabei bedeutet *linear*, dass die Funktion  $\mathbf{f}$  linear in allen Komponenten von  $\mathbf{q}$  ist, und *autonom*, dass  $\mathbf{f}$  nicht explizit von der Zeit abhängt. In diesem Fall können wir (2.2) schreiben als

$$\dot{\mathbf{q}}(t) = \mathbf{A}\mathbf{q}(t), \quad (2.3)$$

wobei  $\mathbf{A}$  eine nicht-singuläre  $N \times N$ -Matrix aus konstanten Koeffizienten ist<sup>2</sup>. Die Lösungen des linearen Systems haben die Form

$$\mathbf{q}_i(t) = c_i e^{\lambda_i t} \mathbf{v}_i, \quad (2.4)$$

wobei  $\mathbf{v}_i$  ein Eigenvektor von  $\mathbf{A}$  zum Eigenwert  $\lambda_i$  ist, wie sich leicht durch Einsetzen von (2.4) in (2.3) überprüfen lässt. Existieren  $N$  linear unabhängige Eigenvektoren, so ist die allgemeine Lösung von (2.3) eine Linearkombination der speziellen Lösungen (2.4), deren Koeffizienten  $c_i$  von den Anfangsbedingungen abhängen. Andernfalls müssen *verallgemeinerte Eigenvektoren* hinzugezogen werden (vgl. [19] und darin enthaltene Referenzen).

### 2.2.3. Nichtlineare autonome Systeme

Wir beschäftigen uns weiterhin mit autonomen Systemen, lassen nun jedoch zu, dass die Funktion  $\mathbf{f}$  in Gl. (2.2) nichtlinear in den Komponenten von  $\mathbf{q}$  ist. In den meisten Fällen kann man nun keine expliziten Lösungen mehr angeben. Man ist gezwungen, sich mit qualitativen Aussagen über die möglichen Trajektorien zu begnügen oder sich den Methoden der Numerik zu bedienen.

Als ersten Schritt in Richtung einer qualitativen Beschreibung eines nichtlinearen Systems ist es sinnvoll, dessen Ruhelagen und ihre Stabilität zu bestimmen.

#### *Fixpunkte und Stabilität*

Als *Fixpunkt* oder *Ruhelage* bezeichnet man einen Zustand  $\mathbf{q}_0$  mit der Eigenschaft

$$\left. \frac{d\mathbf{q}(t)}{dt} \right|_{\mathbf{q}_0} = \mathbf{f}(\mathbf{q}_0) = 0. \quad (2.5)$$

Hat ein System also einmal einen Fixpunkt erreicht, wird es diesen Zustand für alle Zeiten beibehalten. In der realen Welt ist dies jedoch nicht zwangsläufig der Fall, da Fluktuationen das System aus seiner Ruhelage heraus treiben können. In diesem Fall ist es entscheidend, wie das System auf kleine Auslenkungen aus der Ruhelage reagiert. Werden kleine Abweichungen von einem Fixpunkt abgedämpft, bezeichnet man diesen

<sup>2</sup>Allgemeiner wäre die Gleichung  $\dot{\mathbf{q}}(t) = \mathbf{A}\mathbf{q}(t) + \mathbf{b}$ . Diese lässt sich jedoch durch die Transformation  $\mathbf{q} \rightarrow \mathbf{q} + \mathbf{A}^{-1}\mathbf{b}$  auf (2.3) zurückführen, sofern  $\mathbf{A}$  invertierbar ist, was wir voraussetzen.

als *stabil*. Werden die Störungen hingegen verstärkt, bezeichnet man den Fixpunkt als *instabil*.

Um die Stabilität einer Ruhelage  $\mathbf{q}_0$  zu bestimmen, betrachtet man daher eine Störung

$$\mathbf{q}_s(t) = \mathbf{q}(t) - \mathbf{q}_0. \quad (2.6)$$

Für deren zeitliche Entwicklung gilt:

$$\begin{aligned} \dot{\mathbf{q}}_s(t) &\stackrel{(2.6)}{=} \dot{\mathbf{q}}(t) - \dot{\mathbf{q}}_0 \\ &\stackrel{(2.5)}{=} \dot{\mathbf{q}}(t) \\ &\stackrel{(2.2)}{=} \mathbf{f}(\mathbf{q}(t)) \\ &\stackrel{(2.6)}{=} \mathbf{f}(\mathbf{q}_0 + \mathbf{q}_s(t)). \end{aligned} \quad (2.7)$$

Ist  $\mathbf{q}_s(t)$  hinreichend klein, lässt sich  $\mathbf{f}(\mathbf{q}_0 + \mathbf{q}_s(t))$  durch eine Taylorentwicklung bis zur linearen Ordnung annähern. Es ist dann

$$\begin{aligned} \dot{\mathbf{q}}_s(t) &\approx \mathbf{f}(\mathbf{q}_0) + \mathbf{L}(\mathbf{q}_0)\mathbf{q}_s(t) \\ &\stackrel{(2.5)}{=} \mathbf{L}(\mathbf{q}_0)\mathbf{q}_s(t), \end{aligned} \quad (2.8)$$

wobei

$$\mathbf{L}(\mathbf{q}) = \{L_{ij}(\mathbf{q})\} = \left\{ \left. \frac{\partial f_i}{\partial q_j} \right|_{\mathbf{q}} \right\} \quad (2.9)$$

die Jacobi-Matrix von  $\mathbf{f}$  an der Stelle  $\mathbf{q}$  bezeichnet. Gl. (2.8) ist äquivalent zu (2.3), deren Lösungen wir bereits diskutiert haben.

Ob kleine Störungen anwachsen oder abgedämpft werden, hängt nach (2.4) von den Realteilen der Eigenwerte der Jacobi-Matrix ab. Sind die Realteile sämtlicher Eigenwerte kleiner als null, werden kleine Störungen exponentiell abgedämpft; der Fixpunkt ist stabil<sup>3</sup>. Existiert mindestens ein Eigenwert mit positivem Realteil, wird eine Störung in Richtung des zugehörigen Eigenvektors anwachsen; der Fixpunkt ist instabil. Falls der Realteil ein oder mehrerer Eigenwerte null ist<sup>4</sup> und alle anderen negativ sind, müssen höhere Ordnungen der Taylor-Entwicklung von  $\mathbf{f}$  betrachtet werden, um eine Aussage über die Stabilität zu machen. Eine detailliertere Klassifikation von Fixpunkten findet sich z. B. in [1] und [52].

<sup>3</sup>Genau genommen ist hier *asymptotisch stabil* gemeint. Es existieren außerdem die Begriffe *Lyapunov stabil* und *exponentiell stabil*. Eine genaue Definition dieser Begriffe findet sich in [19].

<sup>4</sup>Fixpunkte, bei denen die Realteile sämtlicher Eigenwerte von null verschieden sind, heißen *hyperbolische Fixpunkte*. Alle anderen heißen *degenerierte Fixpunkte*. Der Satz von Hartman-Grobman garantiert, dass der Fluss des nichtlinearen Systems in der Nähe hyperbolischer Fixpunkte topologisch äquivalent zu dem Fluss des linearisierten Systems ist (vgl. [1]).

### 2.3. Das Versklavungsprinzip

Das Versklavungsprinzip besteht in einem mathematischen Formalismus, mit dessen Hilfe in der Nähe von Instabilitäten die explizite Zeitabhängigkeit von schnell relaxierenden Größen eliminiert werden kann. Diese schnell relaxierenden Größen werden dann analytische Funktionen der langsamen Größen, wodurch die Anzahl der Freiheitsgrade in vielen Fällen drastisch reduziert wird.

Der Formalismus kann sowohl auf Systeme gewöhnlicher Differentialgleichungen als auch auf Systeme partieller Differentialgleichungen angewendet werden. An dieser Stelle soll nur Ersteres mit dem Spezialfall autonomer Systeme behandelt werden. Außerdem wird die adiabatische Elimination schnell relaxierender Variabler aus der Fokker-Planck-Gleichung dargestellt.

#### 2.3.1. Anwendung auf autonome dynamische Systeme

Ausgangspunkt des Formalismus ist die in Abschnitt 2.2.3 dargestellte lineare Stabilitätsanalyse. Dazu betrachten wir ein  $N$ -komponentiges autonomes dynamisches System  $\dot{\mathbf{q}} = \mathbf{f}(\mathbf{q})$ , mit einem Fixpunkt  $\mathbf{q}_0$ . Linearisiert man um diesen Fixpunkt, erhält man nach Gl. (2.8)

$$\dot{\mathbf{q}}_s(t) = \mathbf{L}(\mathbf{q}_0) \mathbf{q}_s(t), \quad (2.10)$$

wobei  $\mathbf{L}$  die Jacobi-Matrix von  $\mathbf{f}$  ist. Der nächste Schritt besteht darin, die Eigenwerte  $\lambda_i$  und die zugehörigen Eigenvektoren  $\mathbf{v}_i$  von  $\mathbf{L}$  zu bestimmen. Existieren  $N$  linear unabhängige Eigenvektoren, können wir diese als Basisvektoren für ein neues Koordinatensystem verwenden. Damit können wir schreiben

$$\mathbf{q}_s(t) = \sum_{i=1}^N \xi_i(t) \mathbf{v}_i. \quad (2.11)$$

Werden die Indizes derart sortiert, dass die entsprechenden Eigenwerte

$$\operatorname{Re}\{\lambda_1\} \geq \operatorname{Re}\{\lambda_2\} \geq \dots \geq \operatorname{Re}\{\lambda_\ell\} \geq 0 > \operatorname{Re}\{\lambda_{\ell+1}\} \geq \dots \geq \operatorname{Re}\{\lambda_N\} \quad (2.12)$$

erfüllen, können wir die Vektoren

$$\mathbf{u}(t) := (\xi_1(t), \dots, \xi_\ell(t))^T, \quad (2.13a)$$

$$\mathbf{s}(t) := (\xi_{\ell+1}(t), \dots, \xi_N(t))^T \quad (2.13b)$$

definieren. Da für die Amplituden  $\xi_i$  der Eigenvektoren  $\mathbf{v}_i$

$$\xi_i(t) = \xi_i(0) e^{\lambda_i t} \quad (2.14)$$

gilt, werden alle Komponenten des Vektors  $\mathbf{s}$  im linearisierten System mit der Zeit exponentiell abgedämpft (stabile Amplituden), wohingegen die Komponenten des Vektors  $\mathbf{u}$  exponentiell anwachsen (instabile Amplituden, engl. *unstable*).

Wir berücksichtigen nun wieder die nichtlinearen Anteile unseres Ausgangssystems, so dass gilt

$$\dot{\mathbf{q}}_s(t) = \mathbf{L}(\mathbf{q}_0) \mathbf{q}_s(t) + \mathbf{N}(\mathbf{q}_s(t)), \quad (2.15)$$

wobei  $\mathbf{N}(\mathbf{q}_s(t))$  die Nichtlinearitäten bezeichnet. Setzen wir Gl. (2.11) ein, ergibt sich

$$\begin{aligned} \dot{\mathbf{q}}_s &= \sum_{i=1}^N \dot{\xi}_i \mathbf{v}_i = \mathbf{L} \sum_{i=1}^N \xi_i \mathbf{v}_i + \mathbf{N}(\xi_1, \dots, \xi_N) \\ &= \sum_{i=1}^N \lambda_i \xi_i \mathbf{v}_i + \mathbf{N}(\xi_1, \dots, \xi_N). \end{aligned} \quad (2.16)$$

Multiplizieren wir von links mit dem adjungierten Zeilenvektor  $\mathbf{v}_j^\dagger$ , der die Beziehung

$$\mathbf{v}_j^\dagger \cdot \mathbf{v}_i = \delta_{ij} \quad (2.17)$$

erfüllt, so erhalten wir

$$\dot{\xi}_j = \lambda_j \xi_j + \tilde{N}_j(\xi_1, \dots, \xi_N) \quad \text{mit} \quad \tilde{N}_j = \mathbf{v}_j^\dagger \cdot \mathbf{N}. \quad (2.18)$$

Mit den Definitionen (2.13) können wir dies schreiben als

$$\dot{\mathbf{u}} = \mathbf{\Lambda}_u \mathbf{u} + \tilde{\mathbf{N}}_u(\mathbf{u}, \mathbf{s}) \quad (2.19a)$$

$$\dot{\mathbf{s}} = \mathbf{\Lambda}_s \mathbf{s} + \tilde{\mathbf{N}}_s(\mathbf{u}, \mathbf{s}). \quad (2.19b)$$

Dabei sind  $\mathbf{\Lambda}_u$  und  $\mathbf{\Lambda}_s$  Diagonalmatrizen aus Eigenwerten:

$$\mathbf{\Lambda}_u := \text{diag}(\lambda_1, \dots, \lambda_\ell), \quad (2.20a)$$

$$\mathbf{\Lambda}_s := \text{diag}(\lambda_{\ell+1}, \dots, \lambda_N). \quad (2.20b)$$

Die Gln. (2.19) sind immer noch exakt. Bisher wurde nichts weiter als eine Koordinatentransformation durchgeführt. Das Versklavungsprinzip erlaubt es nun, in der Nähe von Instabilitäten das Gleichungssystem (2.19) durch Approximationen zu vereinfachen.

Typische Zeitskalen, die die Dynamik der Amplituden  $\xi_i$ , aus denen die Vektoren  $\mathbf{u}$  und  $\mathbf{s}$  bestehen, charakterisieren, sind durch die entsprechenden Eigenwerte  $\lambda_i$  festgelegt. Deren Realteile können als Raten interpretiert werden, mit denen die Amplituden in der Nähe des betrachteten Fixpunktes anwachsen oder abgedämpft werden. Als typische Zeitskala für die Amplitude  $\xi_i$  können wir daher

$$\tau_i := \frac{1}{|\text{Re}\{\lambda_i\}|} \quad (2.21)$$

definieren.

Eine Instabilität ist dadurch charakterisiert, dass es einen oder mehrere Eigenwerte gibt, deren Realteile einen Nulldurchgang vollführen<sup>5</sup>. Der Gl. (2.21) können wir entnehmen, dass

$$\tau_i \rightarrow \infty \quad \text{falls} \quad \operatorname{Re}\{\lambda_i\} \rightarrow 0. \quad (2.22)$$

Wir betrachten nun den Fall, dass eine Amplitude  $u_j$  gerade instabil geworden ist, d. h.

$$\operatorname{Re}\{\lambda_j\} \gtrsim 0. \quad (2.23)$$

Wegen (2.22) werden sich alle anderen Amplituden im Vergleich zu  $u_j$  sehr schnell verändern. Das hat zur Folge, dass die Trajektorien *stabiler* Amplituden schnell in den entsprechenden Ruhelagen enden, deren Positionen im Zustandsraum nur noch von den aktuellen Werten der instabilen Amplituden abhängen kann. Wir können daher näherungsweise annehmen, dass die stabilen Amplituden einer Änderung der instabilen Amplituden *instantan* folgen. Das heißt, der Wert einer stabilen Amplitude zur Zeit  $t$ ,  $s_k(t)$ , hängt nur noch von den Werten der instabilen Amplituden zur Zeit  $t$  ab:

$$s_k(t) = s_k(u_1(t), \dots, u_\ell(t)) \quad (2.24)$$

In diesem Sinne *versklaven* die instabilen Amplituden die stabilen. Da die Dynamik nun vollständig durch die instabilen Amplituden determiniert ist, werden diese als *Ordnungsparameter* bezeichnet. Die Funktionen (2.24) beschreiben in der Terminologie der nichtlinearen Dynamik *invariante Mannigfaltigkeiten*.

Ist eine Beziehung der Form (2.24) gefunden, ergibt sich durch Einsetzen in Gl. (2.19a) die Ordnungsparameterdynamik

$$\dot{\mathbf{u}} = \Lambda_u \mathbf{u} + \tilde{\mathbf{N}}_u(\mathbf{u}, \mathbf{s}(\mathbf{u})). \quad (2.25)$$

Schließlich bleibt das Problem der Konstruktion der Funktionen  $s_k(u_1(t), \dots, u_\ell(t))$ . Aufgrund der Nichtlinearitäten in (2.19) ist dies im Allgemeinen nicht exakt möglich. Wir können die Funktion jedoch *in der Nähe des betrachteten Fixpunktes*, der sich bei dem transformierten Gleichungssystem (2.19) im Ursprung befindet, annähern, indem wir eine Entwicklung nach Potenzen der  $u_i$

$$s_k(\mathbf{u}) = \sum_i a_{ki} u_i + \sum_{i,j} b_{kij} u_i u_j + \dots \quad (2.26)$$

ansetzen. Beschränken wir uns auf eine Beschreibung in einer kleinen Umgebung des Fixpunktes, können wir annehmen, dass die  $u_i$  klein, und somit höhere Potenzen vernachlässigbar sind. Je mehr Terme man berücksichtigt, desto größer wird die Umgebung

<sup>5</sup>Natürlich sind die Eigenwerte keine dynamischen Variablen, die im Laufe der Zeit verschiedene Werte annehmen, sondern Funktionen der äußeren Kontrollparameter. Wenn gesagt wird, das System befinde sich in der Nähe einer Instabilität, ist damit gemeint, dass die Kontrollparameter so gewählt sind, dass es einen Eigenwert gibt, dessen Realteil fast verschwindet.

des Fixpunktes, in der die berechnete Ordnungsparameterdynamik eine gute Näherung darstellt.

Leitet man (2.26) nach der Zeit ab und vergleicht das Ergebnis mit (2.19), so erhält man ein Gleichungssystem für die Koeffizienten von (2.26). Dieses Verfahren wird nun an einem einfachen Beispiel durchgeführt.

### 2.3.2. Ein Beispiel: Das Haken-Zwanzig-System

Wir betrachten ein Beispielsystem mit einem Ordnungsparameter  $u$  und einer stabilen Amplitude  $s$  [13, 23]:

$$\dot{u} = \varepsilon u - us \quad (2.27a)$$

$$\dot{s} = -\gamma s + u^2 \quad (2.27b)$$

Dabei seien  $\varepsilon$  und  $\gamma$  reelle Konstanten und es sei

$$\gamma > 0. \quad (2.28)$$

Das System (2.27) ist in der Literatur unter dem Begriff *Haken-Zwanzig-System* bekannt. Eine Instabilität tritt auf bei

$$\varepsilon_c = 0. \quad (2.29)$$

Für  $\varepsilon < \varepsilon_c$  gibt es einen Fixpunkt im Ursprung, der stabil ist; für  $\varepsilon > \varepsilon_c$  wird der stabile Fixpunkt im Ursprung zum Sattelpunkt und es entstehen zwei stabile Fixpunkte bei  $(\pm\sqrt{\gamma\varepsilon}, \varepsilon)$ .

“Knapp oberhalb” der Instabilität, d. h.  $\varepsilon \gtrsim \varepsilon_c$ , können wir das Versklavungsprinzip (2.24) anwenden. Dazu machen wir den Ansatz

$$s(u) = au + bu^2 + \mathcal{O}(u^3) \quad (2.30)$$

und vernachlässigen im weiteren Verlauf der Rechnung als erste Näherung alle Terme höherer Ordnung als  $u^2$ . Leitet man (2.30) nach der Zeit ab, ergibt sich

$$\dot{s} = a\dot{u} + 2bu\dot{u}. \quad (2.31)$$

Setzen wir (2.27a) ein, wobei wir den Term proportional zu  $\varepsilon$  vernachlässigen, so erhalten wir

$$\begin{aligned} \dot{s} &= -aus - 2bu^2s \\ &= -au(au + bu^2) - 2bu^2(au + bu^2) \\ &= -a^2u^2 - 3abu^3 - 2b^2u^4 \\ &= -a^2u^2 + \mathcal{O}(u^3). \end{aligned} \quad (2.32)$$

Nach (2.27b) gilt weiterhin

$$\begin{aligned}\dot{s} &= -\gamma(au + bu^2) + u^2 \\ &= -a\gamma u + (1 - b\gamma)u^2.\end{aligned}\tag{2.33}$$

Vergleichen wir die Koeffizienten von (2.32) mit denen von (2.33), so ergibt sich

$$-a\gamma = 0\tag{2.34}$$

$$-a^2 = 1 - b\gamma.\tag{2.35}$$

Dieses Gleichungssystem wird gelöst durch

$$a = 0\tag{2.36}$$

$$b = \gamma^{-1}.\tag{2.37}$$

Damit folgt

$$s(u) = \frac{1}{\gamma}u^2 + \mathcal{O}(u^3).\tag{2.38}$$

In (2.27a) eingesetzt, ergibt sich damit die Ordnungsparameterdynamik

$$\dot{u} = \varepsilon u - \frac{1}{\gamma}u^3 + \mathcal{O}(u^4).\tag{2.39}$$

Berücksichtigt man höhere Potenzen von  $u$ , erhält man Gleichungen, die in einer größeren Umgebung des Fixpunktes im Ursprung gültig sind.

Das Ergebnis (2.38) erhält man auch, indem man in Gl. (2.27b)  $\dot{s} = 0$  setzt und anschließend nach  $s$  auflöst. Diese Art der Näherung bezeichnet man als *adiabatische Approximation*. Dieser Begriff existiert auch in der Quantenmechanik und wird dort als Synonym für die Born-Oppenheimer-Approximation verwendet. Das Versklavungsprinzip ist mit dem Prinzip der Born-Oppenheimer-Approximation durchaus vergleichbar. Letzteres beruht ebenfalls auf der Zeitskalentrennung zwischen den schnellen Elektronen und langsamen Atomkernen oder Ionenrümpfen in einem Molekül oder Festkörper.

### 2.3.3. Anwendung auf die Fokker-Planck-Gleichung

Wir ersetzen nun unser dynamisches System aus Abschnitt 2.3.1 durch ein System von Langevin-Gleichungen, die in Abschnitt 1.6 eingeführt wurden. Der Einfachheit halber beschränken wir uns auf ein zweikomponentiges System, das bereits in der Form (2.19) vorliegt, und fügen additives Rauschen hinzu:

$$\dot{u} = \lambda_u u + N_u(u, s) + \sqrt{2D_u^2} \xi_u(t)\tag{2.40}$$

$$\dot{s} = \lambda_s s + N_s(u, s) + \sqrt{2D_s^2} \xi_s(t).\tag{2.41}$$

Das System befinde sich in der Nähe einer Instabilität, d. h. es sei  $\lambda_s < 0$  und

$$|\lambda_u| \ll |\lambda_s|. \quad (2.42)$$

Die Fokker-Planck-Gleichung für die jpdf  $f_2(u, s)$  lautet

$$\frac{\partial}{\partial t} f_2(u, s) = - \left\{ \frac{\partial}{\partial u} D_u^{(1)} + \frac{\partial}{\partial s} D_s^{(1)} \right\} f_2(u, s) + \left\{ D_u^{(2)} \frac{\partial^2}{\partial u^2} + D_s^{(2)} \frac{\partial^2}{\partial s^2} \right\} f_2(u, s). \quad (2.43)$$

Dabei ist

$$D_u^{(1)}(u, s) = \lambda_u u + N_u(u, s), \quad (2.44)$$

$$D_s^{(1)}(u, s) = \lambda_s s + N_s(u, s). \quad (2.45)$$

Wir machen nun den Ansatz

$$f_2(u, s) = p(s|u) f_1(u) \quad (2.46)$$

mit dem Ziel, das Versklavungsprinzip zu verwenden, um eine FPG für  $f_1(u)$  alleine abzuleiten. Setzen wir dies in (2.43) ein, so ergibt sich

$$\begin{aligned} \frac{\partial f_2}{\partial t} = \dot{p} f_1 + p \dot{f}_1 = & - \frac{\partial}{\partial u} (D_u^{(1)} p f_1) + D_u^{(2)} \frac{\partial^2 f_1}{\partial u^2} p - f_1 \frac{\partial}{\partial s} (D_s^{(1)} p) \\ & + D_s^{(2)} f_1 \frac{\partial^2 p}{\partial s^2} + D_u^{(2)} \left( 2 \frac{\partial f_1}{\partial u} \frac{\partial p}{\partial u} + f_1 \frac{\partial^2 p}{\partial u^2} \right). \end{aligned} \quad (2.47)$$

Eine Integration der gesamten Gleichung über  $s$  führt auf

$$\begin{aligned} f_1 \underbrace{\frac{\partial}{\partial t} \int ds p}_{=1} + \dot{f}_1 \underbrace{\int ds p}_{=1} = & - \frac{\partial}{\partial u} f_1 \int ds D_u^{(1)} p + D_u^{(2)} \frac{\partial^2 f_1}{\partial u^2} \underbrace{\int ds p}_{=1} - f_1 [D_s^{(1)} p]_{s_{\min}}^{s_{\max}} \\ & + D_s^{(2)} f_1 \left[ \frac{\partial p}{\partial s} \right]_{s_{\min}}^{s_{\max}} + D_u^{(2)} \left[ 2 \frac{\partial f_1}{\partial u} \frac{\partial}{\partial u} \underbrace{\int ds p}_{=1} + f_1 \frac{\partial^2}{\partial u^2} \underbrace{\int ds p}_{=1} \right]. \end{aligned} \quad (2.48)$$

Setzen wir voraus, dass  $D_s^{(1)} p$  sowie  $\frac{\partial p}{\partial s}$  an den Rändern des Definitionsbereichs von  $s$  verschwinden, was der generische Fall ist, so erhalten wir

$$\frac{\partial f_1}{\partial t} = - \frac{\partial}{\partial u} f_1 \int ds D_u^{(1)} p + D_u^{(2)} \frac{\partial^2 f_1}{\partial u^2}. \quad (2.49)$$

Dies ist noch keine geschlossene Gleichung, da noch die unbekannte cpdf  $p(u|s)$  enthalten ist.

Vergleichen wir (2.47) und (2.49), so gilt offensichtlich

$$\dot{p}f_1 = -f_1 \frac{\partial}{\partial s} (D_s^{(1)} p) + D_s^{(2)} f_1 \frac{\partial^2 p}{\partial s^2} + D_u^{(2)} \left( 2 \frac{\partial f_1}{\partial u} \frac{\partial p}{\partial u} + f_1 \frac{\partial^2 p}{\partial u^2} \right). \quad (2.50)$$

Da die Variable  $s$  nach Voraussetzung schnell relaxiert, können wir davon ausgehen, dass  $p(s|u)$  sich schneller mit  $s$  ändert als mit  $u$ . Daher können wir in erster Näherung die Terme mit Ableitungen nach  $u$  in (2.50) gegenüber den Termen mit Ableitungen nach  $s$  vernachlässigen. Diese Näherung führt auf

$$\dot{p} = -\frac{\partial}{\partial s} (D_s^{(1)} p) + D_s^{(2)} \frac{\partial^2 p}{\partial s^2}. \quad (2.51)$$

Dies ist eine FPG für  $p(s|u)$ , wenn wir  $u$  als Konstante betrachten. Die stationäre Lösung ist nach Gl. (1.67) gegeben durch

$$p(s|u) = \mathcal{N} \exp \left[ \int^s \frac{D_s^{(1)}(u, s')}{D_s^{(2)}} ds' \right]. \quad (2.52)$$

Setzen wir nun für  $D_s^{(1)}(u, s')$  Gl. (2.45) ein, wobei wir in der nichtlinearen Funktion näherungsweise  $s = 0$  setzen, lässt sich das Integral auswerten. Dies ergibt unter Berücksichtigung der Normierungsbedingung  $\int ds p(s|u) = 1$ :

$$p(s|u) = \sqrt{-\frac{\lambda_s}{2\pi D_s^{(2)}}} \exp \left\{ \frac{\lambda_s}{2D_s^{(2)}} \left( s + \frac{N_s(u)}{\lambda_s} \right)^2 \right\}. \quad (2.53)$$

Damit kann das Integral in Gl. (2.49) für konkrete Funktionen  $N_u$  und  $N_s$  ausgewertet werden. Wir tun dies am Beispiel des im vorherigen Abschnitt behandelten Haken-Zwanzig-Systems. Dafür gilt nach Gln. (2.27):

$$N_u(u, s) = -us, \quad N_s(u) = u^2, \quad \lambda_u = \varepsilon, \quad \lambda_s = -\gamma. \quad (2.54)$$

Damit erhalten wir für das Integral

$$I := \int ds D_u^{(1)} p = \int (\varepsilon u - us) \mathcal{N} \exp \left\{ a \left( s - \frac{u^2}{\gamma} \right)^2 \right\}, \quad (2.55)$$

wobei die Abkürzungen

$$a := -\frac{\gamma}{2D_s^{(2)}} \quad \text{und} \quad \mathcal{N} := \sqrt{\frac{\gamma}{2\pi D_s^{(2)}}} = \sqrt{-\frac{a}{\pi}} \quad (2.56)$$

benutzt wurden. Dies ergibt

$$I = \mathcal{N}u \left[ \underbrace{\varepsilon \int \exp \left\{ a \left( s - \frac{u^2}{\gamma} \right)^2 \right\} ds}_{=\sqrt{-\pi/a}} - \underbrace{\int s \exp \left\{ a \left( s - \frac{u^2}{\gamma} \right)^2 \right\} ds}_{=\frac{u^2}{\gamma} \sqrt{-\pi/a}} \right]. \quad (2.57)$$

Mit (2.56) folgt schließlich

$$I = \varepsilon u - \frac{u^3}{\gamma}. \quad (2.58)$$

Damit erhalten wir als FPG für die pdf des Ordnungsparameters  $f_1(u)$ :

$$\frac{\partial}{\partial t} f_1(u, t) = \left( -\frac{\partial}{\partial u} \left( \varepsilon u - \frac{u^3}{\gamma} \right) + D_u^{(2)} \frac{\partial^2}{\partial u^2} \right) f_1(u, t). \quad (2.59)$$

Die entsprechende Langevin-Gleichung lautet:

$$\dot{u} = \varepsilon u - \frac{u^3}{\gamma} + \sqrt{2D_u^{(2)}} \xi(t). \quad (2.60)$$

Der deterministische Teil dieser Ordnungsparametergleichung entspricht also genau dem Ergebnis (2.39) in adiabatischer Näherung. Auch das Resultat (2.38) für die invariante Mannigfaltigkeit finden wir in unserer Rechnung wieder. Dazu betrachten wir noch einmal Gl. (2.53). Auf das Haken-Zwanzig-System angewendet ergibt sich

$$p(s|u) = \sqrt{\frac{\gamma}{2\pi D_s^{(2)}}} \exp \left\{ -\frac{\gamma}{2D_s^{(2)}} (s - s(u))^2 \right\} \quad (2.61)$$

mit dem  $s(u)$  aus Gl. (2.38). Innerhalb der adiabatischen Approximation ist also die cpdf  $p(s|u)$  eine um die invariante Mannigfaltigkeit  $s(u)$  des deterministischen Falls zentrierte Gauß-Verteilung. Das heißt, dass die schnell relaxierenden Moden um die invariante Mannigfaltigkeit fluktuieren.

## 2.4. Das Prinzip der unvoreingenommenen Schätzung

Wie bereits in der Einführung zu diesem Kapitel erwähnt wurde, erlaubt das Prinzip der unvoreingenommenen Schätzung (im Folgenden mit MEP, aus dem engl. *Maximum-Entropy Principle*, abgekürzt), aus der Kenntnis makroskopischer Observablen Rückschlüsse auf die mikroskopische Dynamik eines Systems zu ziehen. Makroskopische Observablen können beispielsweise Erwartungswerte mikroskopischer Variablen sein. In

diesem Fall lässt sich eine Verteilungsfunktion der mikroskopischen Variablen konstruieren, aus der sich unter bestimmten Annahmen ein mikroskopisches Modell ableiten lässt.

Der Ausgangspunkt der Überlegung ist der Folgende: Durch die Kenntnis bestimmter Erwartungswerte einer Zufallsvariablen verfügt ein Beobachter über eine unvollständige Information der zugrunde liegenden Verteilung. Es existiert nun eine Vielzahl von möglichen Verteilungen, die mit den bekannten Erwartungswerten kompatibel sind. Ein rational handelnder Beobachter würde nun unter all diesen Verteilungen diejenige auswählen, die am besten die vorhandene Information widerspiegelt und in welche am wenigsten Annahmen einfließen, die nicht durch die vorhandene Information zu rechtfertigen sind. Wir nennen diese die *unvoreingenommenste* Verteilung. Sie ist dadurch ausgezeichnet, dass sie gewissermaßen die *Ungewissheit* des Beobachters maximiert.

### 2.4.1. Die Informationsentropie

Ein Maß für die Ungewissheit einer Wahrscheinlichkeitsverteilung liefert die, maßgeblich von C. E. Shannon entwickelte Informationstheorie. In [48] zeigte Shannon, dass das einzige Maß, das bestimmte plausible Bedingungen erfüllt, durch

$$i(p_1, \dots, p_n) = - \sum_{j=1}^n p_j \log p_j \quad (2.62)$$

gegeben ist. Darin ist  $p_j$  die Wahrscheinlichkeit für die Realisation des Zustandes  $j$  bei einem Zufallsexperiment. Die Größe  $i$  wird als *Shannon-Entropie*, *Informationsentropie* oder einfach nur *Information* bezeichnet. Die Basis des Logarithmus in der Definition ist beliebig.

Im Falle kontinuierlicher Zufallsvariablen lässt sich die differentielle Information

$$i(f(x)) = - \int dx f(x) \log f(x), \quad (2.63)$$

definieren, wobei  $f(x)$  eine Wahrscheinlichkeitsdichte ist. Bei mehreren Zufallsvariablen ist  $x$  als Vektor, und  $f$  als jpdf zu betrachten. Die differentielle Information ist jedoch streng genommen nicht als Verallgemeinerung der Shannon-Entropie, die nur für den diskreten Fall definiert ist, zu verstehen. Der Übergang von diskreten zu kontinuierlichen Variablen kann dargestellt werden, indem man die Werte, die eine kontinuierliche Zufallsvariable annehmen kann, in Bins  $I_j$  der Breite  $\varepsilon$  einteilt. Nach dem Mittelwertsatz der Integralrechnung existiert ein  $x_j \in I_j$ , so dass die Wahrscheinlichkeit, dass eine Realisation im Bin  $I_j$  liegt, durch

$$p_j = \varepsilon f(x_j) \quad (2.64)$$

gegeben ist. Setzt man dies in Gl. (2.62) ein, ergibt sich

$$i_\varepsilon = - \sum_j \varepsilon f(x_j) \log(\varepsilon f(x_j)) \quad (2.65)$$

$$= - \sum_j \varepsilon f(x_j) \log(f(x_j)) - \sum_j \varepsilon f(x_j) \log(\varepsilon) . \quad (2.66)$$

Lässt man nun  $\varepsilon$  gegen null gehen, erhält man

$$i = - \int f(x) \log f(x) dx - \log \varepsilon , \quad (2.67)$$

wobei der hintere Term divergiert. Solange es nur um Informationsänderungen geht, oder darum, die Verteilung maximaler Information zu bestimmen, kann der hintere Term als Konstante ignoriert, und die differentielle Information, Gl. (2.63), verwendet werden. Als Maß für die Ungewissheit einer Verteilung kann (2.63) jedoch nicht interpretiert werden, insbesondere da der Ausdruck negativ werden kann.

Diejenige pdf, die der unvoreingenommensten Schätzung entspricht, kann durch Extremalisierung des Funktionals (2.63) berechnet werden. Die Informationen, die wir über die pdf haben, fließen dabei als Nebenbedingungen ein. Mathematisch haben wir es mit einem Problem der Variationsrechnung unter Nebenbedingungen zu tun, wofür sich die Methode der Lagrange-Multiplikatoren eignet. Für eine Darstellung dieser Methode und Beispielrechnungen sei auf [25] verwiesen. Es sei betont, dass diejenige pdf, die die Information unter gegebenen Nebenbedingungen maximiert, nicht unbedingt die tatsächliche pdf im Sinne der Definition (1.10) sein muss. Ob dies der Fall ist oder nicht, hängt von der Wahl der Nebenbedingungen ab.

### 2.4.2. Anwendung auf Systeme in der Nähe von Instabilitäten

Wir beschreiben nun, wie mit Hilfe des MEP Ordnungsparametermodelle von Systemen in der Nähe von Instabilitäten aufgestellt werden können. Für eine generelle Prozedur schlägt Haken Nebenbedingungen der Form [25]

$$c_i = \langle q_i \rangle \quad (2.68a)$$

$$c_{ij} = \langle q_i q_j \rangle \quad (2.68b)$$

$$c_{ijk} = \langle q_i q_j q_k \rangle \quad (2.68c)$$

$$c_{ijkl} = \langle q_i q_j q_k q_l \rangle \quad (2.68d)$$

vor, wobei  $\mathbf{q} = (q_1, \dots, q_N)$  der Zustandsvektor des betrachteten Systems ist. Die Koeffizienten  $c_i, c_{ij}, c_{ijk}, c_{ijkl}$  werden aus Messdaten bestimmt. Als Resultat der Variation des Funktionals (2.63) unter den Nebenbedingungen (2.68) ergibt sich die jpdf

$$f_N(\mathbf{q}) = \exp[-V(\mathbf{q}, \{\lambda\})] , \quad (2.69)$$

mit

$$V(\mathbf{q}, \{\lambda\}) = \lambda + \sum_i \lambda_i q_i + \sum_{ij} \lambda_{ij} q_i q_j + \sum_{ijk} \lambda_{ijk} q_i q_j q_k + \sum_{ijkl} \lambda_{ijkl} q_i q_j q_k q_l . \quad (2.70)$$

Die  $\lambda$ 's sind dabei die noch zu bestimmenden Lagrange-Multiplikatoren, die von den Koeffizienten der Nebenbedingungen (2.68) abhängen. Unter bestimmten Voraussetzungen ist  $V$  als Potential zu interpretieren: Im eindimensionalen Fall zeigt ein Vergleich mit der stationären Lösung der FPG (1.67), dass dies bei additivem Rauschen der Fall ist. Im mehrdimensionalen Fall muss die Voraussetzung des *detaillierten Gleichgewichts* (engl. *detailed balance*) erfüllt sein (s. [25]). Das Potential ergibt sich dann durch ein Wegintegral über den Driftkoeffizienten. Die Bestimmung der Lagrange-Multiplikatoren ist in den meisten Fällen nicht auf analytischem Wege möglich und muss numerisch erfolgen, was an dieser Stelle jedoch nicht diskutiert werden soll.

Sind die Lagrange-Multiplikatoren bestimmt, lässt sich mit der Interpretation von  $V(\mathbf{q}, \{\lambda\})$  als Potential analog zu der unter 2.3.1 beschriebenen Prozedur eine Koordinatentransformation durchführen, so dass sich die neuen Koordinaten in Ordnungsparameter und stabile Amplituden einteilen lassen. Der erste Schritt bei dieser Koordinatentransformation besteht darin, den Ursprung des neuen Koordinatensystems in ein geeignetes Extremum von  $V$  zu legen. Im zweiten Schritt wird die Matrix der entsprechend transformierten Lagrange-Multiplikatoren ( $\tilde{\lambda}_{ij}$ ) diagonalisiert. Das neue Potential lautet dann

$$\hat{V}(\hat{\mathbf{q}}, \{\hat{\lambda}\}) = \hat{\lambda} + \sum_i \hat{\lambda}_i \hat{q}_i^2 + \sum_{ijk} \hat{\lambda}_{ijk} \hat{q}_i \hat{q}_j \hat{q}_k + \sum_{ijkl} \hat{\lambda}_{ijkl} \hat{q}_i \hat{q}_j \hat{q}_k \hat{q}_l . \quad (2.71)$$

Die transformierten Lagrange-Multiplikatoren  $\hat{\lambda}_i$  entsprechen dabei den Eigenwerten der symmetrischen Matrix ( $\tilde{\lambda}_{ij}$ ) und sind daher reell. Wir identifizieren nun die Ordnungsparameter  $u_i$  mit denjenigen Komponenten  $\hat{q}_i$ , für die

$$\hat{\lambda}_i \leq 0 \quad (2.72)$$

ist, und die stabilen Amplituden  $s_j$  mit den Komponenten  $\hat{q}_j$ , mit der Eigenschaft

$$\hat{\lambda}_j > 0 . \quad (2.73)$$

In Analogie zur Anwendung des Versklavungsprinzips auf die FPG, s. Abschnitt 2.3.3, machen wir nun den Ansatz

$$f(\mathbf{u}, \mathbf{s}) = e^{\hat{V}} = f(\mathbf{u})p(\mathbf{s}|\mathbf{u}) , \quad (2.74)$$

mit dem Ziel, einen Ausdruck für die Verteilung der Ordnungsparameter  $f(\mathbf{u})$  abzuleiten. Dazu zerlegen wir  $\hat{V}$  in einen Anteil, der nur von den Ordnungsparametern abhängt, und einen, der sowohl von den stabilen Amplituden als auch von den Ordnungsparametern abhängt:

$$\hat{V}(\mathbf{u}, \mathbf{s}, \{\hat{\lambda}\}) = \hat{\lambda} + \hat{V}_u(\mathbf{u}, \{\hat{\lambda}^{(u)}\}) + \hat{V}_s(\mathbf{u}, \mathbf{s}, \{\hat{\lambda}^{(s)}\}) . \quad (2.75)$$

Dabei sind  $\hat{V}_u$  und  $V_s$  wie folgt definiert:

$$\hat{V}_u = \sum_i \hat{\lambda}_i^{(u)} u_i^2 + \sum_{ijk} \hat{\lambda}_{ijk}^{(u)} u_i u_j u_k + \sum_{ijkl} \hat{\lambda}_{ijkl}^{(u)} u_i u_j u_k u_\ell \quad (2.76)$$

$$\hat{V}_s = \sum_i \hat{\lambda}_i^{(s)} s_i^2 + 3 \sum_{ijk} \hat{\lambda}_{ijk}^{(s)} s_i u_j u_k + 4 \sum_{ijkl} \hat{\lambda}_{ijkl}^{(s)} s_i u_j u_k u_\ell + \dots \quad (2.77)$$

Es gilt:

$$\begin{aligned} f(\mathbf{u}) &= \int d\mathbf{s} f(\mathbf{u}, \mathbf{s}) \\ &= \int d\mathbf{s} e^{\hat{V}(\mathbf{u}, \mathbf{s}, \{\hat{\lambda}\})} \\ &= e^{\hat{\lambda} + \hat{V}_u(\mathbf{u}, \{\hat{\lambda}^{(u)}\})} \int d\mathbf{s} e^{\hat{V}_s(\mathbf{u}, \mathbf{s}, \{\hat{\lambda}^{(s)}\})}. \end{aligned} \quad (2.78)$$

Für das Integral verwenden wir die Abkürzung

$$\int d\mathbf{s} e^{\hat{V}_s(\mathbf{u}, \mathbf{s}, \{\hat{\lambda}^{(s)}\})} := e^{-h(\mathbf{u})}, \quad (2.79)$$

so dass wir

$$f(\mathbf{u}) = \exp \{ \hat{\lambda} + \hat{V}_u(\mathbf{u}, \{\hat{\lambda}^{(u)}\}) - h(\mathbf{u}) \} := e^{W_u(\mathbf{u})} \quad (2.80)$$

erhalten. Mit Gl. (2.74) folgt daraus:

$$\begin{aligned} p(\mathbf{s}|\mathbf{u}) &= \exp \{ \hat{V}(\mathbf{u}, \mathbf{s}, \{\hat{\lambda}\}) - W_u(\mathbf{u}) \} \\ &= \exp \{ \hat{V}_s(\mathbf{u}, \mathbf{s}, \{\hat{\lambda}^{(s)}\}) + h(\mathbf{u}) \}. \end{aligned} \quad (2.81)$$

Wir können nun das Potential  $\hat{V}_s$  aus Gl. (2.77) annähern, um das Ergebnis der Anwendung des Versklavungsprinzips auf die FPG, Gl. (2.61), zu verallgemeinern. Dazu gehen wir davon aus, dass die Ordnungsparameter  $u_i$  klein sind und, analog zu der Bedingung (2.42), dass die Parameter  $\hat{\lambda}_i^{(s)}$  betragsmäßig groß im Vergleich zu den Parametern  $\hat{\lambda}_i^{(u)}$  sind, so dass wir annehmen können, dass die  $s_i$  von der Größenordnung  $u_i^2$  sind. Alle Terme aus Gl. (2.77), die durch die Punkte abgekürzt sind, sind dann von der Ordnung  $u^5$  oder höher. Wenn wir diese vernachlässigen, entfallen sämtliche Produkte aus verschiedenen Komponenten  $s_i$ , so dass wir die bedingte Wahrscheinlichkeitsdichte  $p(\mathbf{s}|\mathbf{u})$  durch ein Produkt der Form

$$p(\mathbf{s}|\mathbf{u}) = \prod_i p(s_i|\mathbf{u}) \quad (2.82)$$

darstellen können. Die einzelnen Faktoren haben dann genau wie (2.61) Gauß-Gestalt. Nach [25] lassen sie sich explizit darstellen als

$$p(s_i|\mathbf{u}) = \exp \left\{ -\hat{\lambda}_i^{(s)} (s_i - g_i(\mathbf{u}))^2 + \hat{\lambda}_i^{(s)} g_i^2(\mathbf{u}) \right\} \quad \text{mit} \quad (2.83)$$

$$g_i(\mathbf{u}) = \frac{1}{2\hat{\lambda}_i^{(s)}} \left[ 3 \sum_{jk} \hat{\lambda}_{ijk}^{(s)} u_j u_k + 4 \sum_{jkl} \hat{\lambda}_{ijkl}^{(s)} u_j u_k u_\ell \right]. \quad (2.84)$$

## 2.5. Die Information von Systemen in der Nähe von Instabilitäten

Die Informationsentropie, die in Abschnitt 2.4.1 eingeführt wurde, ist nicht nur im Zusammenhang mit dem Prinzip der unvoreingenommenen Schätzung von Bedeutung. Im Zusammenhang mit dem Ordnungsparameterkonzept zeigt sich, dass sich die Information in zwei Summanden aufspalten lässt, von denen der eine nur von der pdf der Ordnungsparameter abhängt. In der Nähe von Instabilitäten nimmt dieser Teil der Information drastisch zu, wohingegen der andere Term eher konstant bleibt [25]. Diese Tatsache wird in Kapitel 4 eine große Rolle spielen.

### 2.5.1. Aufspaltung der Information

Im letzten Abschnitt wurde mittels des MEP gezeigt, dass sich die Verteilungsfunktion von Systemen in der Nähe von Instabilitäten gemäß

$$f(\mathbf{u}, \mathbf{s}) = f(\mathbf{u}) \prod_i p(s_i | \mathbf{u}) \quad (2.85)$$

aufspalten lässt. Setzen wir diesen Ansatz in die Formel für die Information des Systems an, so ergibt sich:

$$i(f(\mathbf{u}, \mathbf{s})) = - \int d\mathbf{u} d\mathbf{s} f(\mathbf{u}, \mathbf{s}) \log f(\mathbf{u}, \mathbf{s}) \quad (2.86)$$

$$= - \int d\mathbf{u} d\mathbf{s} f(\mathbf{u}) \prod_i p(s_i | \mathbf{u}) \left[ \log f(\mathbf{u}) + \sum_j \log p(s_j | \mathbf{u}) \right]. \quad (2.87)$$

Wir sehen, lässt sich das Integral aufspalten in zwei Summanden

$$i = i_u + i_{\text{Rest}}, \quad (2.88)$$

wobei der erste Summand

$$i_u = - \int d\mathbf{u} f(\mathbf{u}) \log f(\mathbf{u}) \int d\mathbf{s} \prod_i p(s_i | \mathbf{u}) \quad (2.89)$$

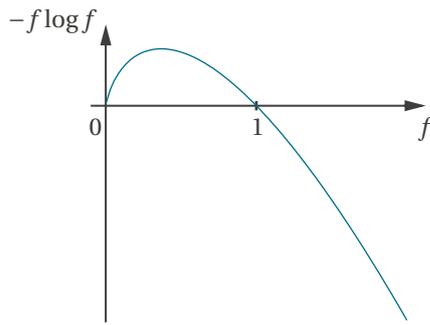
$$= - \int d\mathbf{u} f(\mathbf{u}) \log f(\mathbf{u}) \quad (2.90)$$

nur von der Verteilungsfunktion der Ordnungsparameter abhängt. Wir nennen ihn die Ordnungsparameterinformation. Für den zweiten Summanden ergibt sich:

$$i_{\text{Rest}} = - \int d\mathbf{u} f(\mathbf{u}) \int d\mathbf{s} \prod_i p(s_i | \mathbf{u}) \sum_j \log p(s_j | \mathbf{u}) \quad (2.91)$$

$$= - \int d\mathbf{u} f(\mathbf{u}) \sum_j \int d\mathbf{s} \log p(s_j | \mathbf{u}) \prod_i p(s_i | \mathbf{u}) \quad (2.92)$$

$$= - \int d\mathbf{u} f(\mathbf{u}) \sum_j \int ds_j p(s_j | \mathbf{u}) \log p(s_j | \mathbf{u}). \quad (2.93)$$

Abb. 2.1 Die Funktion  $-f \log f$ .

Dabei entspricht das letzte Integral, über das summiert wird, der Definition der Information einer versklavten Amplitude, wenn man die pdf durch eine auf die Ordnungsparameter bedingte cpdf ersetzt. Dies entspricht in gewisser Weise der Philosophie des Versklavungsprinzips, das besagt, dass die versklavten Amplituden den Ordnungsparametern instantan folgen. Bezeichnen wir die Information der versklavten Amplituden als

$$i_s(\mathbf{u}) = - \sum_j \int ds_j p(s_j|\mathbf{u}) \log p(s_j|\mathbf{u}), \quad (2.94)$$

so entspricht  $i_{\text{Rest}}$  dem Erwartungswert von  $i_s$  bzgl. der Ordnungsparameter:

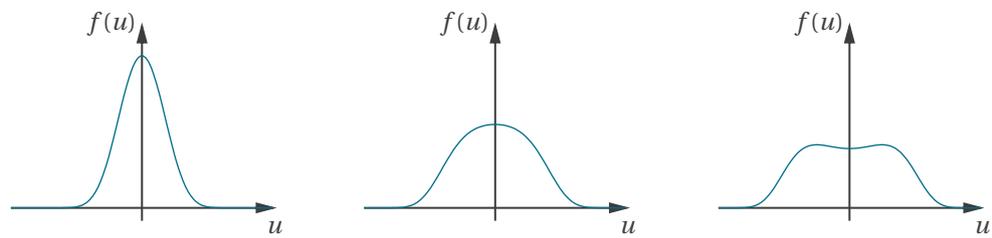
$$i_{\text{Rest}} = \langle i_s(\mathbf{u}) \rangle_{\mathbf{u}} = \int d\mathbf{u} f(\mathbf{u}) i_s(\mathbf{u}). \quad (2.95)$$

### 2.5.2. Die Ordnungsparameterinformation in der Nähe von Instabilitäten

Um ein Gefühl dafür zu bekommen, wie sich die Information für verschiedene Verteilungen verhält, ist es zweckmäßig, einmal den Integranden  $-f \log f$  als Funktion von  $f$  zu betrachten (s. Abb. 2.1). Demnach liefert eine Verteilung  $f$  positive Beiträge zur Information für  $0 < f < 1$  und negative für  $f > 1$ . Das heißt, dass für stark lokalisierte Verteilungen eine geringe oder stark negative Information im Vergleich zu breiten Verteilungen zu erwarten ist.

Abb. 2.2 zeigt das typische Verhalten einer pdf eines Ordnungsparameters weit unterhalb, knapp unterhalb und knapp oberhalb einer Instabilität. Für die Information ist zu erwarten, dass sie von links nach rechts zunimmt. Eine typische cpdf für versklavte Amplituden ist durch Gl. (2.61) gegeben, wobei in der Nähe von Instabilitäten angenommen werden kann, dass der Parameter  $\gamma$  groß ist, so dass die Verteilung stark lokalisiert ist. Entsprechend ist die Information der versklavten Amplituden, Gl. (2.94), für gewöhnlich klein oder stark negativ.

Dieser quantitative Unterschied zwischen der Information der Ordnungsparameter und der Information der versklavten Amplituden wird in Abschnitt 4.4 ausgenutzt, um eine Methode zur Identifikation von Ordnungsparametern und versklavten Amplituden aus Daten abzuleiten.



**Abb. 2.2** Typische pdf eines Ordnungsparameters in der Nähe einer Instabilität. Links: Weit unterhalb der Instabilität. Mitte: Knapp unterhalb der Instabilität. Rechts: Knapp oberhalb der Instabilität.

### 3. Ordnungsparameterdynamik für ein System gekoppelter Oszillatoren

In diesem Kapitel wird das Verhalten einer kollektiven Variablen eines hochdimensionalen Modellsystems analysiert. Diese kann im Kontext einer Darstellung des Systems als Hierarchie von Bewegungsgleichungen fluktuierender Momente (vgl. Abschnitt 3.2) als Ordnungsparameter interpretiert werden kann. Es folgt zunächst eine kurze Einführung in das Modellsystem und die Motivation der durchgeführten Analyse.

#### 3.1. Einführung

Das betrachtete Modellsystem besteht aus  $N$  identischen Oszillatoren in einem bimodalen überdämpften Potential. Dabei ist jeder Oszillator mit allen anderen gekoppelt und wird außerdem stochastischen Fluktuationen unterworfen. Zusätzlich wird das gesamte System durch eine periodische Kraft  $F(t) = F(t + T)$  angetrieben (vgl. Abb. 3.1). Das System stochastischer Evolutionsgleichungen lautet (in dimensionsloser Form):

$$\dot{x}_i = x_i - x_i^3 + \frac{\Theta}{N} \sum_{j=1}^N (x_j - x_i) + \sqrt{2D} \xi_i(t) + F(t) \quad , \quad i = 1, \dots, N. \quad (3.1)$$

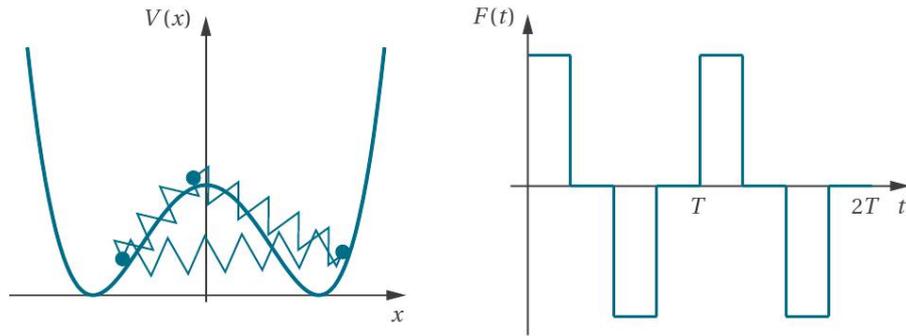
Dabei ist  $x_i(t)$  die Koordinate des  $i$ ten Oszillators,  $\Theta$  eine Kopplungskonstante und  $D$  die Amplitude der gaußverteilten Langevin-Kraft  $\xi_i(t)$ , für die

$$\langle \xi_i(t) \rangle = 0 \quad \text{und} \quad \langle \xi_i(t) \xi_j(s) \rangle = \delta_{ij} \delta(t - s) \quad (3.2)$$

gilt.

Ohne die periodische Kraft  $F(t)$  wurde dieses System erstmals als Beispielmodell für Selbstorganisation in offenen komplexen Systemen durch Rückkopplungsschleifen von Kometani und Shimizu [31] aufgestellt. Unter Rückkopplungsschleifen versteht man dabei die Wechselwirkung zwischen der kollektiven makroskopischen Variablen

$$S(t) = \frac{1}{N} \sum_{i=1}^N x_i(t) , \quad (3.3)$$



**Abb. 3.1** Illustration des betrachteten Modellsystems. Links:  $N$  Oszillatoren ( $N = 3$  in der Abbildung), beschrieben durch die Variable  $x_i$ , bewegen sich in einem bimodalen überdämpften Potential. Die lineare attraktive Kopplung der Oszillatoren untereinander wird durch die Federn veranschaulicht. Rechts: Die Form der periodischen Kraft  $F(t)$  aus Gl. (3.1), die in dieser Arbeit verwendet wurde.

die ein Ensemblemittel über die Subsysteme  $x_i$  darstellt, und den Abweichungen der Subsysteme vom Ensemblemittel

$$a_i(t) = x_i(t) - S(t) . \quad (3.4)$$

Die Form der Wechselwirkung der Subsysteme untereinander

$$\frac{\Theta}{N} \sum_{j=1}^N (x_j - x_i) \quad (3.5)$$

entstammt einem Modell für die Wechselwirkung zwischen Myosin und Aktinfilamenten bei der Muskelkontraktion.

Später untersuchten Desai und Zwanzig die dynamischen Eigenschaften sowie den Gleichgewichtszustand dieses Systems im thermodynamischen Grenzfall  $N \rightarrow \infty$  [7]. Dabei entwickelten sie eine systematische Hierarchie von stochastischen Bewegungsgleichungen für die Kumulanten von  $a_1$ .

Die Antwort des Systems auf ein externes periodisches Signal wurde erstmals von Morillo, Gómez Ordóñez und Casado untersucht [38]. Dabei entdeckten sie, dass das Phänomen der stochastischen Resonanz durch die Wechselwirkung der Subsysteme (3.5) erheblich verstärkt wird. Unter stochastischer Resonanz (SR) versteht man die Verstärkung oder Optimierung der Antwort eines nichtlinearen Systems auf ein schwaches kohärentes externes Signal durch Rauschen [14]. In den letzten Jahren wurde auch der Einfluss der anderen Systemparameter auf die SR im Falle endlicher Systemgröße  $N$  eingehend erforscht [5, 6, 43]. Dabei wurde das Verhalten der kollektiven Größe  $S(t)$ , Gl. (3.3), betrachtet.

In einem kürzlich erschienenen Artikel [40] untersuchen die Autoren die SR-Effekte der Variablen  $S(t)$  bei schwachem Treiber. Schwach bedeutet dabei, dass ein *einzelnes*

System ohne Kopplung, das man dem gleichen Treiber unterwirft, mit den Methoden der linearen Antworttheorie untersucht werden kann. In diesem linearen Regime sind die SR-Effekte, wie die Verstärkung der mittleren Amplitude der Systemantwort auf die treibende Kraft oder das Signal-Rausch-Verhältnis, sehr klein. Das interessante Ergebnis dieser Arbeit ist, dass die SR-Effekte der Variablen  $S(t)$  dennoch groß sind, so dass sich das  $N$ -dimensionale gekoppelte System nicht im linearen Regime befinden kann.

Um diesen und andere Effekte besser verstehen zu können, ist es wünschenswert, eine effektive eindimensionale Langevin-Gleichung der Form

$$\dot{S}(t) = -U'_{\text{eff}}(S) + F(S, t) + \xi(S, t) \quad (3.6)$$

für den stochastischen Prozess  $S(t)$  zu formulieren. Dies wurde bereits früher von Pivovskiy und Zaikin [43] auf analytischem Wege versucht, indem ähnlich wie in [7] eine Hierarchie von stochastischen Bewegungsgleichungen für die fluktuierenden Momente

$$M_k(t) = \frac{1}{N} \sum_{i=1}^N [x_i(t) - S(t)]^k \quad (3.7)$$

aufgestellt wurde. Unter Vernachlässigung aller Momente mit  $k \geq 3$  und der Anwendung des Versklavungsprinzips gelangten die Autoren zu einer Gleichung der Form (3.6), was in Abschnitt 3.2 dargestellt wird. Es stellt sich jedoch heraus, dass diese den ursprünglichen Prozess nicht zufriedenstellend beschreibt und außerdem nicht in allen Parameterbereichen existiert [40].

Im Rahmen der vorliegenden Arbeit wird zum einen überprüft, ob der stochastische Prozess  $S(t)$  Markov-Charakter hat, was die Voraussetzung dafür ist, dass eine Gleichung der Art (3.6) überhaupt formuliert werden kann. Zum anderen wird das in Abschnitt 1.8 beschriebene numerische Verfahren zur Bestimmung von Drift und Diffusion angewendet, um das effektive Potential  $U_{\text{eff}}$  sowie den Charakter des effektiven Rauschens zu bestimmen. Abschließend wird der geschätzte effektive Prozess mit dem tatsächlichen verglichen. Ein Teil der Ergebnisse ist bereits in [17] veröffentlicht.

## 3.2. Die gaußsche Approximation

Mit der Definition von  $S$  in Gl. (3.3) und (3.1) ergibt sich

$$\begin{aligned} \dot{S} &= \frac{1}{N} \sum_{i=1}^N \dot{x}_i \\ &= \frac{1}{N} \sum_{i=1}^N \left\{ x_i - x_i^3 + \frac{\Theta}{N} \sum_{j=1}^N (x_j - x_i) + \sqrt{2D} \xi_i(t) + F(t) \right\}, \end{aligned}$$

was sich zu

$$\dot{S} = S - \frac{1}{N} \sum_{i=1}^N x_i^3 + \sqrt{\frac{2D}{N}} \tilde{\xi}(t) + F(t) \quad (3.8)$$

umformen lässt. Dabei ist  $\tilde{\xi}(t)$  eine Langevin-Kraft mit den gleichen Eigenschaften wie  $\xi_i(t)$ , was aus dem zentralen Grenzwertsatz folgt. Der Summationsterm in (3.8) lässt sich nicht als Funktion von  $S$  darstellen; Gl. (3.8) ist also nicht geschlossen. In [6] wird mit Hilfe der Definition von sogenannten *fluktuierenden Momenten*

$$M_k(t) = \frac{1}{N} \sum_{i=1}^N [x_i(t) - S(t)]^k \quad (3.9)$$

eine unendliche Hierarchie von Langevin-Gleichungen aufgestellt. Mit Hilfe von (3.9) lässt sich der Summationsterm in (3.8) durch

$$\frac{1}{N} \sum_{i=1}^N x_i^3 = M_3 + 3SM_2 + S^3. \quad (3.10)$$

ausdrücken. Damit ergibt sich:

$$\dot{S} = S - S^3 - 3SM_2 - M_3 + \sqrt{\frac{2D}{N}} \tilde{\xi}(t) + F(t). \quad (3.11)$$

Berechnet man die zeitlichen Ableitungen der fluktuierenden Momente, so ergibt sich, dass  $\dot{M}_k(t)$  eine Funktion von  $M_{k-1}$ ,  $M_k$ ,  $M_{k+1}$  und  $M_{k+2}$  ist. Möchte man ein geschlossenes niedrigdimensionales System aus Langevin-Gleichungen formulieren, bleibt lediglich die Hoffnung, dass die Beiträge der fluktuierenden Momente höherer Ordnungen vernachlässigbar klein sind, so dass man diese unendliche Hierarchie von Gleichungen durch eine endliche approximieren kann.

Eine mögliche Näherung besteht darin, alle fluktuierenden Momente  $M_k$  mit  $k \geq 3$  zu vernachlässigen, was in der Literatur als gaußsche Approximation bezeichnet wird. Mit dieser sowie einigen anderen für große  $N$  gültigen Näherungen formulierten Pikovsky und Zaikin [43] das geschlossene zweikomponentige System

$$\dot{S} = S - S^3 - 3MS + \sqrt{\frac{2D}{N}} \xi(t) + F(t) \quad (3.12a)$$

$$\frac{1}{2} \dot{M} = (1 - 3S^2 - \Theta)M - 3M^2 + D. \quad (3.12b)$$

Dabei ist  $M := M_2$  mit der Definition (3.9) und somit stets positiv. Betrachtet man den thermodynamischen Limes  $N \rightarrow \infty$ , so verschwindet der Rauschterm in (3.12a). Wird weiterhin die Kraft  $F(t)$  vorübergehend vernachlässigt, so erhält man ein deterministisches, autonomes, nichtlineares dynamisches System. Dieses besitzt einen Fixpunkt bei

$$S_1^* = 0 \quad (3.13a)$$

$$M_1^* = \left[ 1 - \Theta + \sqrt{(1 - \Theta)^2 + 12D} \right] / 6 \quad (3.13b)$$

sowie zwei weitere bei

$$S_{2,3}^* = \pm \left\{ 0.5 - 0.25 \left[ \Theta - \sqrt{(2+\Theta)^2 - 24D} \right] / 4 \right\}^{1/2} \quad (3.14a)$$

$$M_{2,3}^* = \left[ 2 + \Theta - \sqrt{(2+\Theta)^2 - 24D} \right] / 12. \quad (3.14b)$$

Wir betrachten eine Störung  $\tilde{M} = M - M_1^*$  des ersten Fixpunktes. Dafür ergibt sich das transformierte System

$$\dot{S} = (1 - 3M_1^*)S - 3S^3 - 3S\tilde{M} \quad (3.15a)$$

$$\dot{\tilde{M}} = \alpha\tilde{M} - 6\tilde{M}^2 - 6M_1^*S^2 - 6S^2\tilde{M} \quad (3.15b)$$

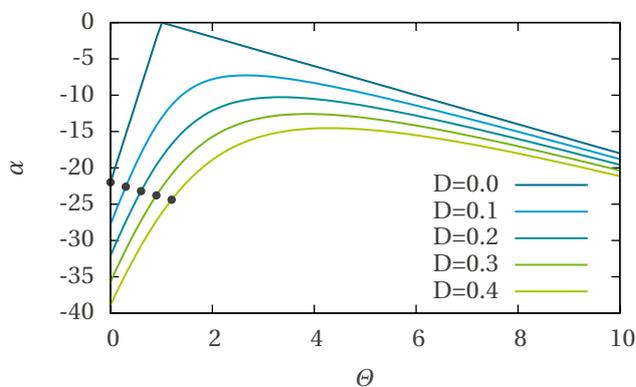
mit dem Parameter

$$\alpha = -2 \left[ 5(1 - \Theta) + 6\sqrt{(1 - \Theta)^2 + 12D} \right]. \quad (3.16)$$

Die Jacobi-Matrix des Systems (3.15) bei  $S = 0$ ,  $\tilde{M} = 0$  hat bereits Diagonalgestalt und lautet:

$$\mathbf{L} = \begin{pmatrix} 0.5 \left( 1 + \Theta - \sqrt{(1 - \Theta)^2 + 12D} \right) & 0 \\ 0 & \alpha \end{pmatrix}. \quad (3.17)$$

Der obere linke Eintrag von  $\mathbf{L}$  hat einen Nulldurchgang bei  $\Theta = 3D$  und ist positiv für  $\Theta > 3D$ . Der Koeffizient  $\alpha$  ist als Funktion von  $\Theta$  für verschiedene  $D$  in Abb. 3.2 dargestellt. Für  $D > 0$  ist  $\alpha$  immer negativ und in der Nähe der Instabilitätspunkte bei  $\Theta = 3D$ , die



**Abb. 3.2** Der Koeffizient  $\alpha$  als Funktion der Kopplungsstärke  $\Theta$  für verschiedene Rauschamplituden  $D$ . Die schwarzen Punkte entsprechen den jeweiligen Instabilitätspunkten  $\Theta = 3D$ .

durch die schwarzen Punkte markiert sind, betragsmäßig hinreichend groß, so dass das in Abschnitt 2.3 behandelte Versklavungsprinzip verwendet werden kann, um die Variable  $\tilde{M}$  aus dem System (3.15) zu eliminieren. Nach [43] ergibt sich dabei das effektive einkomponentige System

$$\dot{S} = aS - bS^3 + \sqrt{\frac{2D}{N}} \xi(t) + F(t) \quad (3.18)$$

mit den Koeffizienten

$$a = 0.5 \left( 1 + \Theta - \sqrt{(1 - \Theta)^2 + 12D} \right) \quad \text{und} \quad (3.19)$$

$$b = \frac{4a}{2 - \Theta + \sqrt{(2 + \Theta)^2 - 24D}}. \quad (3.20)$$

Da der Ausdruck unter der Wurzel in Gl. (3.20) negativ werden kann, existiert das System nicht in allen Parameterbereichen.

### 3.3. Numerische Simulation

Zur numerischen Simulation des Systems (3.1) wird das einfachste Zeitintegrationsschema, nämlich das explizite Euler-Verfahren, verwendet. Bei einer Schrittweite  $\Delta t$  ergibt sich danach die Koordinate des  $i$ ten Oszillators zum Zeitpunkt  $t + \Delta t$  zu

$$x_i(t + \Delta t) = x_i(t) + \Delta t \left( x_i(t) - x_i^3(t) + \frac{\Theta}{N} \sum_{j=1}^N (x_j(t) - x_i(t)) + F(t) \right) + \sqrt{2D\Delta t} \xi_i(t). \quad (3.21)$$

Nach jedem berechneten Zeitschritt wird die Variable  $S(t)$  gemäß Gl. (3.3) berechnet und gespeichert.

Die Kraft  $F(t) = F(t + T)$  hat die Form

$$F(t) = \begin{cases} A; & 0 \leq t < t_c \\ 0; & t_c \leq t < T/2 \\ -A; & T/2 \leq t < T/2 + t_c \\ 0; & T/2 + t_c \leq t < T \end{cases}. \quad (3.22)$$

Der Übersichtlichkeit halber wird bei den verschiedenen Simulationsläufen anstelle des Parameters  $t_c$  der Parameter

$$r = 2t_c / T \quad (3.23)$$

angegeben. Er dient als Maß für die relative Länge eines Rechteckimpulses im Vergleich zur Periodenlänge.

Die Rauschterme  $\xi_i(t)$  werden mit Hilfe des Box-Muller-Verfahrens [3] erzeugt. Dabei werden aus zwei Standardzufallszahlen  $z_1$  und  $z_2$  zwei stochastisch unabhängige standardnormalverteilte Zufallszahlen  $n_1$  und  $n_2$  erzeugt:

$$\begin{aligned} n_1 &= \sqrt{-2 \ln z_1} \cos(2\pi z_2) \\ n_2 &= \sqrt{-2 \ln z_1} \sin(2\pi z_2). \end{aligned} \quad (3.24)$$

**Tab. 3.1** Simulationsparameter

Datensatz	Parameter							
	$N$	$D$	$\Theta$	$A$	$r$	$\Delta t$	$T$	$T_{\text{sim}}/T$
Sim1	50	0.3	1.0	0.2	0.5	$10^{-4}$	630.0	500
Sim2	10	0.1	0.5	0.1	0.1	$10^{-4}$	630.0	500
Sim3	10	0.2	0.5	0.1	0.1	$10^{-4}$	630.0	500
Sim4	10	0.2645	0.5	0.1	0.1	$10^{-4}$	630.0	500
Sim5	10	0.3	0.5	0.1	0.1	$10^{-4}$	630.0	500
Sim6	10	0.4	0.5	0.1	0.1	$10^{-4}$	630.0	500
Sim7	10	0.3	0.5	0.1	0.5	$10^{-4}$	630.0	3900
Sim8	2	0.3	0.5	0.2	0.5	$10^{-4}$	630.0	500
Sim9	10	0.2	0.5	0.3	0.5	$10^{-3}$	4.0	$5 \cdot 10^7$

Unter *Standardzufallszahlen* versteht man dabei unabhängige Zufallszahlen, die auf dem Intervall  $[0; 1]$  gleichverteilt sind; als *standardnormalverteilt* bezeichnet man Zufallszahlen, die einer Gauß-Verteilung mit einer Varianz von eins unterliegen.

Die Standardzufallszahlen werden mit Hilfe einer frei verfügbaren Fortran-Version des Mersenne Twister-Algorithmus [37] erzeugt. Dieser hat eine Periodizität  $2^{19937} - 1$ , womit sicher gestellt ist, dass die Rauschterme auch bei langen Simulationszeiten  $\delta$ -korreliert sind.

Als Anfangsbedingung wird immer  $x_i(0) = 0 \quad \forall i = 1, \dots, N$  gewählt. Damit die Statistik nicht von den Anfangsbedingungen abhängt, wird die Zeitreihe  $S(t)$  erst nach einer "Einschwingphase" von zehn Periodenlängen der Kraft  $F(t)$  gespeichert.

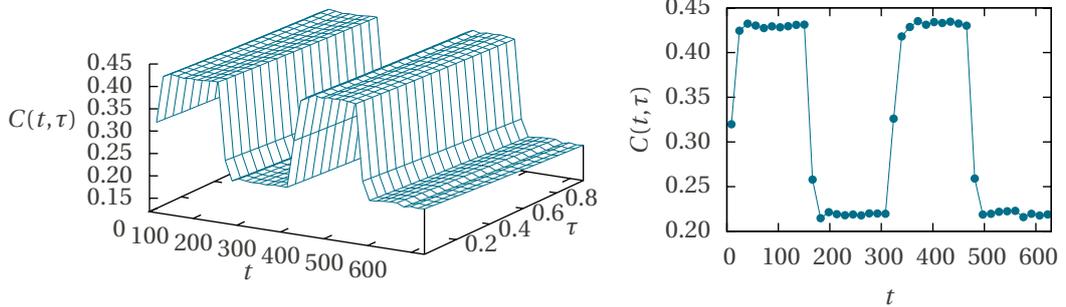
Die Parameter von sämtlichen in dieser Arbeit ausgewerteten Simulationen sind in Tab. 3.1 aufgelistet. Dabei ist  $T_{\text{sim}}$  die physikalische Länge der Simulation und  $T_{\text{sim}}/T$  die Anzahl der simulierten Perioden. Die am häufigsten verwendete Periodenlänge  $T = 630$  entspricht in etwa der in vorherigen Arbeiten [5, 40] verwendeten Kreisfrequenz  $\Omega = 0.01$ .

### 3.4. Stationarität

Bei stochastischen Prozessen mit periodischem Treiber nehmen nach einer bestimmten "Einschwingphase" sämtliche stochastischen Observablen die Periodizität des Treibers an. Daher kann bei der Zeitabhängigkeit dieser Größen, die Transformation

$$t \longrightarrow t \pmod{T} \quad (3.25)$$

angewandt werden. Bei großen Periodenlängen  $T$  der Kraft  $F(t)$ , ist darüber hinaus zu erwarten, dass während jeder der vier Phasen, in denen  $F$  konstant ist, der Prozess nach einer bestimmten Relaxationszeit  $\vartheta_r$  stationär wird. Ist  $\vartheta_r$  sehr viel kleiner als  $t_c$  bzw.  $T/2 - t_c$  in Gl. (3.22), können bei sämtlichen Mittelungsprozessen alle Datenpunkte



**Abb. 3.3** Korrelationsfunktion  $C(t, \tau) = \langle S(t)S(t + \tau) \rangle$  für den Datensatz Sim7. Die rechte Abbildung zeigt einen Schnitt für konstantes  $\tau = 0.1$ .

aus einer Phase verwendet werden. Man kann den Prozess  $S(t)$  dann aufspalten in vier verschiedene stationäre Prozesse, wobei jeder Prozess einer der vier Phasen entspricht. Jede gemessene Phase entspricht dann einer Realisation des entsprechenden stationären Prozesses.

Um die Relaxationszeit abzuschätzen, wird die Korrelation

$$C(t, \tau) = \langle S(t)S(t + \tau) \rangle \quad (3.26)$$

in Abhängigkeit von  $t$  und  $\tau$  untersucht. Dabei wird die neue Zeitachse mit der Transformation (3.25) in 40 Intervalle unterteilt und für jedes Intervall für verschiedene  $\tau$  der Erwartungswert (3.26) geschätzt. Abb. 3.3 zeigt das Ergebnis für den Datensatz Sim7. Danach kann die Relaxationszeit  $\vartheta_r$  abgeschätzt werden auf

$$\vartheta_r \approx 35 \pm 5. \quad (3.27)$$

Bei einer Periodenlänge von  $T = 630$  und einem Parameter  $r = 0.5$  ist die Länge einer Phase konstanter Amplitude des Treibers  $t_c = 157.5$ , so dass es gerechtfertigt ist, näherungsweise von Stationarität während einer solchen Phase auszugehen. Die Ergebnisse aus Abschnitt 3.6.1 bestätigen dies.

### 3.5. Markov-Eigenschaft

Um zu testen, ob der stochastische Prozess  $S(t)$  auf einer bestimmten Zeitskala  $\tau$  Markov-Charakter hat, wird die Gültigkeit der CKG (1.49), die hier die Form

$$p_{2\tau}(S_1, t + 2\tau | S_3, t) = \int dS_2 p_\tau(S_1, t + 2\tau | S_2, t + \tau) p_\tau(S_2, t + \tau | S_3, t) \quad (3.28)$$

annimmt, überprüft. Es sei betont, dass die Gültigkeit der CKG kein Beweis dafür ist, dass  $S(t)$  ein Markov-Prozess ist, sondern lediglich eine *notwendige* Bedingung dafür darstellt. Sie wäre dennoch ein Indiz dafür, dass eine Beschreibung des Prozesses  $S(t)$  als effektiver eindimensionaler Markov-Prozess nicht allzu weit von der Realität entfernt ist.

Zur Überprüfung der Gültigkeit von (3.28) werden separat die linke und rechte Handseite aus einer Zeitreihe geschätzt und miteinander verglichen. Wir bezeichnen die linke Handseite von (3.28) mit  $p_{2\tau}^{\text{LHS}}(S_1|S_3)$  und die rechte Handseite mit  $p_{2\tau}^{\text{RHS}}(S_1|S_3)$ . Eine gute Möglichkeit eines Vergleiches beider Funktionen bieten Konturplots. Dabei werden sowohl von  $p_{2\tau}^{\text{LHS}}(S_1|S_3)$  als auch von  $p_{2\tau}^{\text{RHS}}(S_1|S_3)$  Konturlinien, d. h. Linien, auf denen der entsprechende Funktionswert konstant ist, in die  $(S_1, S_3)$ -Ebene geplottet. Man kann dann mit dem Auge beurteilen, wie gut die Konturlinien übereinanderliegen.

Zum Zwecke eines objektiveren Vergleiches können auch Abstandsmaße, wie z. B. die Kullback-Leibler-Divergenz [33], berechnet werden. Doch auch dann ist es schwierig zu entscheiden, ab welchem Wert des Abstandsmaßes man die CKG für gültig erklärt. Hier wird ein Vergleich mittels Konturplots vorgezogen.

Um die cpdfs aus Gl. (3.28) zu bestimmen, werden zunächst die entsprechenden jpdfs  $f_{2\tau}(S_1, t + 2\tau; S_3, t)$  und  $f_{\tau}(S_2, t + \tau; S_3, t)$  mittels des Schätzers (1.35) unter Verwendung des Epanechnikov-Kerns (1.33) bestimmt. Dabei werden alle Datenpunkte der zweiten Phase des Treibers, in der  $F(t) = 0$  ist, aus allen Perioden verwendet. Die cpdfs werden anschließend mittels Gl. (1.25) bestimmt. Das dabei auftretende Integral im Nenner von (1.25) wird zur numerischen Berechnung mit einer Riemannschen Summe angenähert, ebenso wie das Integral der rechten Handseite der CKG (3.28).

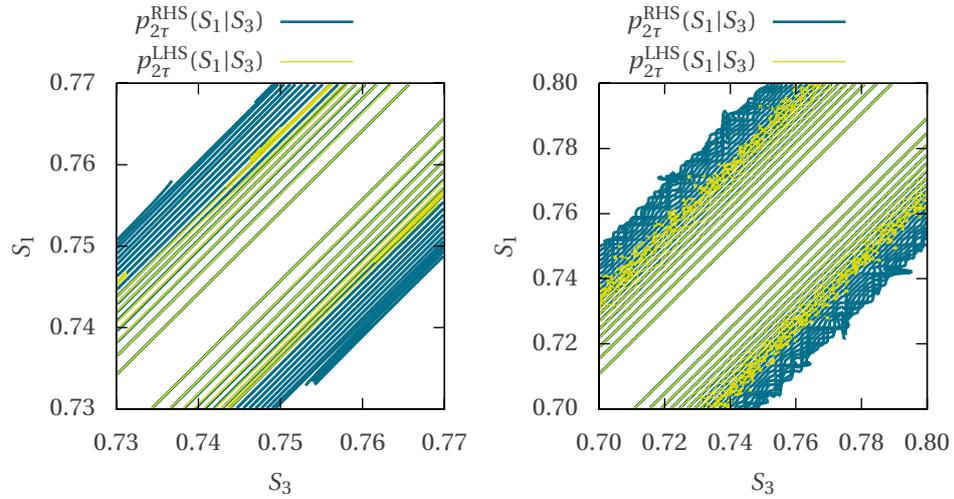
Es stellt sich schlussendlich noch die wichtige Frage, wie das Zeitinkrement  $\tau$  in (3.28) zu wählen ist. Da bei der Schätzung der Kramers-Moyal-Koeffizienten der Grenzwert  $\tau \rightarrow 0$  gebildet wird (vgl. Abschnitt 3.6.2), erscheint es sinnvoll und notwendig die Gültigkeit der CKG für das kleinste verfügbare  $\tau$ , nämlich das Zeitinkrement  $\Delta t$  der numerischen Simulation, zu überprüfen.

Abb. 3.4 zeigt das Ergebnis für die beiden Datensätze Sim7 und Sim8 für  $\tau = \Delta t = 10^{-4}$ . Für die cpdfs ergeben sich sehr schmale gaußförmige Käme entlang der Winkelhalbierenden in der  $(S_1, S_3)$ -Ebene. Um die Konturlinien klar erkennen zu können, wird jeweils ein im Vergleich zur Abb. 3.6 stark vergrößerter Ausschnitt dargestellt. Für beide Datensätze liegen die Konturlinien sehr gut übereinander. Aus Gründen, die im Folgenden genauer dargestellt werden, ist es jedoch fragwürdig, wie aussagekräftig dieses Ergebnis ist.

Für kleine Zeitinkremente  $\tau$  gilt bei einem Langevin-Prozess nach [45] für die Übergangswahrscheinlichkeitsdichte näherungsweise

$$p_{\tau}(S_1, t + \tau | S_2, t) = \frac{1}{2\sqrt{\pi D^{(2)}(S_2, t)\tau}} \exp\left(-\frac{[S_1 - S_2 - D^{(1)}(S_2, t)\tau]^2}{4D^{(2)}(S_2, t)\tau}\right). \quad (3.29)$$

Für beliebige feste Werte von  $S_2$  und  $t$  ergeben sich demnach normierte Gauß-Funktionen, deren Breiten vom Zeitinkrement  $\tau$  abhängen. Im Grenzübergang  $\tau \rightarrow 0$  gehen die



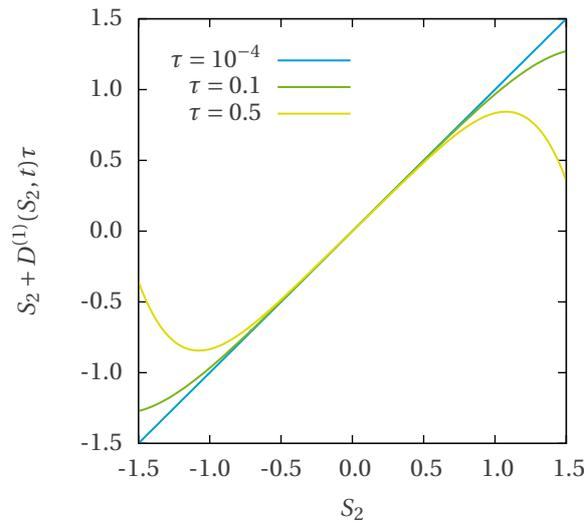
**Abb. 3.4** Vergleich der linken und rechten Handseite der CKG (3.28) für  $\tau = \Delta t = 10^{-4}$ . Die Abbildungen zeigen (im Vergleich zu Abb. 3.6 stark vergrößerte) Ausschnitte von Konturplots. Für die linke Abbildung wurde der Datensatz Sim7 verwendet, für den rechten Sim8 (s. Tab. 3.1). In beiden Fällen liegen die Konturlinien in den Bereichen, in denen die Verteilungen auskonvergiert und daher die Konturlinien glatt sind, gut überein. Die CKG scheint also in beiden Fällen erfüllt zu sein.

Gauß-Funktionen in  $\delta$ -Funktionen über. Die Gauß-Funktionen sind zentriert um

$$S_2 + D^{(1)}(S_2, t)\tau. \quad (3.30)$$

Für  $\tau \approx 0$  sind sie zentriert um  $S_2$ . Für jedes feste  $S_2$  ist die Übergangswahrscheinlichkeitsdichte also eine um  $S_2$  zentrierte Gauß-Funktion, falls das Zeitinkrement  $\tau$  sehr klein ist. Deren einziger “Freiheitsgrad” ist die Breite der Gauß-Funktion. Dieses Grenzverhalten ergibt sich bei allen stochastischen Prozessen, auch bei solchen ohne Markov-Charakter. Daher wird sich in den meisten Fällen für sehr kleine Zeitinkremente auch die Gültigkeit der CKG numerisch bestätigen lassen. Erkennbare Abweichungen sind erst dann zu erwarten, wenn  $\tau$  groß genug ist, so dass die Drift Einfluss auf (3.30) nimmt und somit mehrere “Freiheitsgrade” vorhanden sind. Um abzuschätzen, wann dies der Fall ist, wird (3.30) mit der geschätzten Drift des Datensatzes Sim7 für verschiedene Werte von  $\tau$  geplottet, s. Abb. 3.5. Der Abbildung lässt sich entnehmen, dass möglicherweise erst ab einem Wert der Größenordnung  $10^{-1}$  Abweichungen von der CKG sichtbar werden können.

Aus diesem Grunde wird das Verfahren auch für größere  $\tau$  wiederholt. Abb. 3.6 zeigt exemplarisch die Ergebnisse für  $\tau = 0.25$  für die Datensätze Sim7 und Sim8. Es kann weder hier noch für irgendeinen anderen Wert von  $\tau$  eine signifikante Abweichung von der Gültigkeit der CKG nachgewiesen werden.



**Abb. 3.5** Nach Gl. (3.29) sind die Übergangswahrscheinlichkeiten  $p_\tau(S_1, t + \tau | S_2, t)$  für kleine  $\tau$  für feste Werte von  $S_2$  Gauß-Funktionen, die um (3.30) zentriert sind. Die Abbildung zeigt diese Zentren für den geschätzten Drift des Datensatzes Sim7 für verschiedene  $\tau$ . Bei  $\tau = \Delta t = 10^{-4}$  ist der Einfluss der Drift in (3.30) nicht wahrnehmbar. Als Übergangswahrscheinlichkeit erhält man daher einen Kamm von schmalen Gauß-Kurven entlang der Winkelhalbierenden in der  $(S_1, S_2)$ -Ebene. Da dies auch bei nicht-markovschen Prozessen zu erwarten ist, wird auf einer solch kleinen Zeitskala keine Abweichung zwischen LHS und RHS der CKG (3.28) zu beobachten sein. Bei  $\tau = 0.1$  und  $\tau = 0.5$  wird eine Verformung des Gauß-Kammes sichtbar. Erst auf diesen Zeitskalen scheint daher eine Überprüfung der CKG sinnvoll.

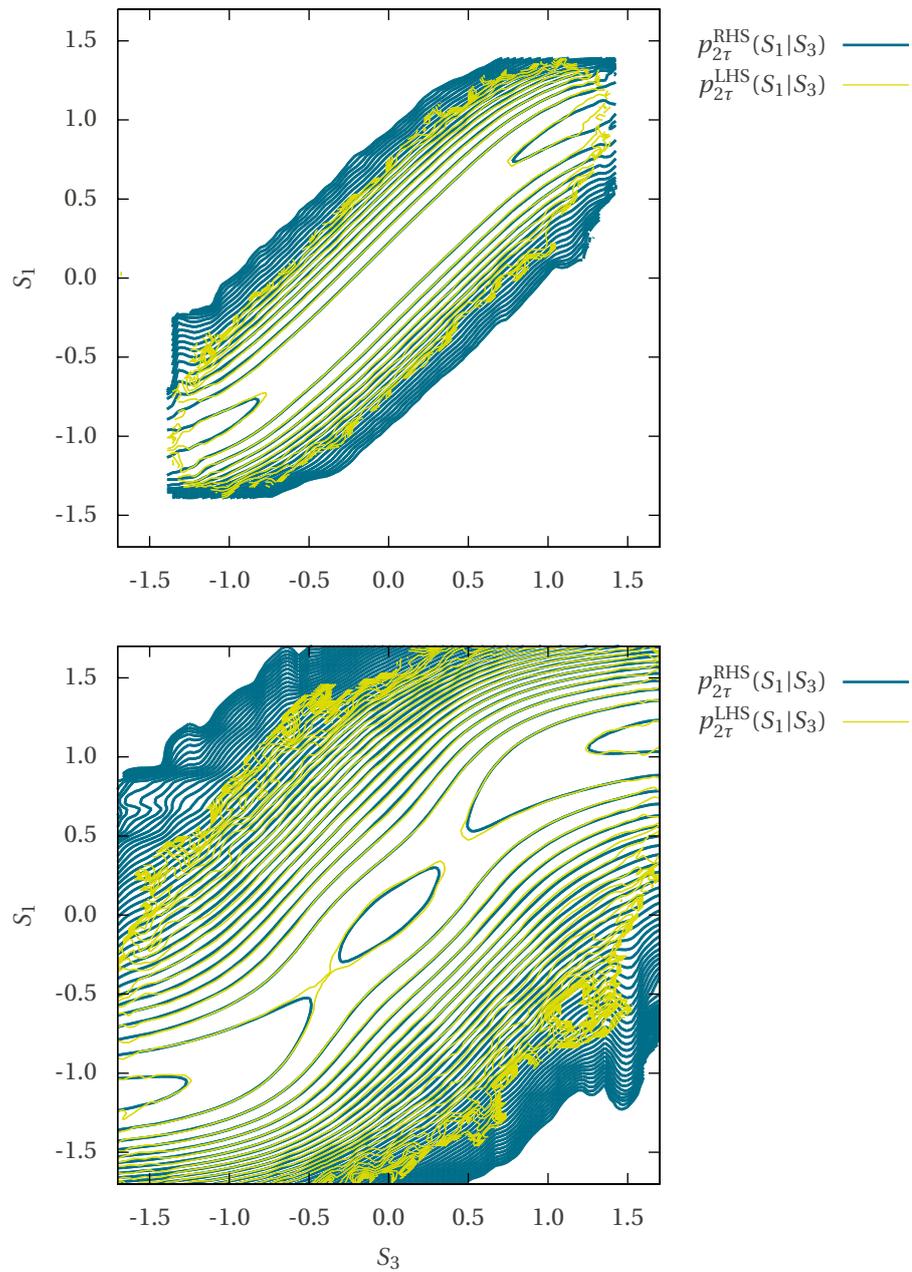
Es sei noch einmal betont, dass diese Ergebnisse keinen *Beweis* der Markov-Eigenschaft des Prozesses  $S(t)$  darstellen. Lediglich der gewissenhaft durchgeführte Versuch, das Gegenteil zu beweisen, ist gescheitert.

### 3.6. Schätzen der Kramers-Moyal-Koeffizienten

Die Resultate des vorangegangenen Abschnittes lassen hoffen, dass es möglich ist, den stochastischen Prozess  $S(t)$  durch eine effektive eindimensionale Langevin-Gleichung zu beschreiben. Der nächste Schritt, die Schätzung der Kramers-Moyal-Koeffizienten, wird in diesem Abschnitt beschrieben.

#### 3.6.1. Zeitabhängigkeit

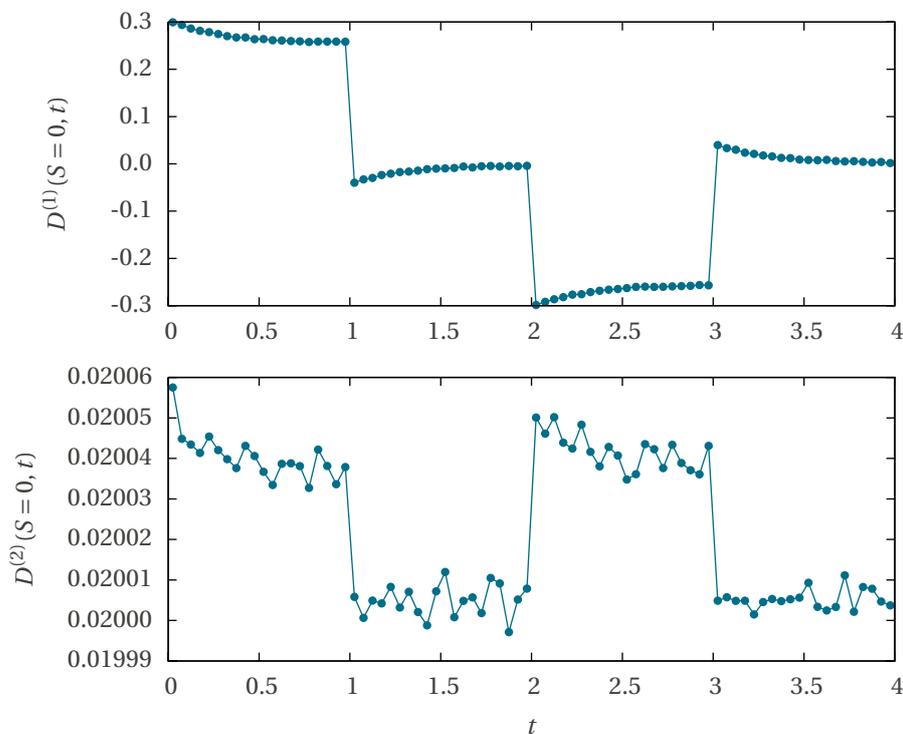
Das Ergebnis aus Abschnitt 3.4 legt die Vermutung nahe, dass Drift und Diffusion während einer Phase konstanter treibender Kraft nach kurzer Zeit stationär werden. Ist dies der Fall, so kann der stochastische Prozess  $S(t)$  in guter Näherung beschrieben werden,



**Abb. 3.6** Vergleich der linken und rechten Handseite der Chapman-Kolmogorov-Gleichung (3.28) für  $\tau = 0.25$  für die Datensätze Sim7 (oben) und Sim8 (unten). Die Konturlinien liegen in den Bereichen, in denen die Verteilungen auskonvergiert sind, gut übereinander. Die Chapman-Kolmogorov-Gleichung scheint also auch hier in beiden Fällen erfüllt zu sein.

indem für jede dieser vier Phasen jeweils zeitlich konstante Drift- und Diffusionsterme geschätzt werden. Um diese Vermutung zu überprüfen, wird die Zeitabhängigkeit beider Koeffizienten für den Datensatz Sim9 überprüft. Die Periodenlänge ist dabei sehr klein gewählt, um mit handhabbarem numerischem Aufwand eine hinreichend gute Zeitauflösung bei der Schätzung der Kramers-Moyal-Koeffizienten zu erzielen. Dadurch werden bei gegebener zeitlicher Auflösung weniger Zeitfenster benötigt.

Die Schätzung der Koeffizienten erfolgt wie in Abschnitt 1.8 beschrieben. Dazu wird jede Periode in 80 Zeitfenster unterteilt, wovon jedes eine Länge von  $t_{ZF} = 0.05$  hat. Abb. 3.7 zeigt die geschätzten Drift- und Diffusionskoeffizienten für konstantes  $S = 0$  in Abhängigkeit von der Zeit. Die Zeit, nach der die Drift nach einem Sprung in der Amplitude des Treibers stationär wird, ist von der Größenordnung 1. Daher kann in sehr guter Näherung davon ausgegangen werden, dass die Drift während einer Phase konstanter Amplitude des Treibers zeitlich konstant ist. Der Einfluss der treibenden Kraft auf die Diffusion ist vernachlässigbar klein. Sie kann in guter Näherung über die ganze Periode hinweg als zeitlich konstant betrachtet werden.



**Abb. 3.7** Drift (oben) und Diffusion (unten) für  $S = 0$  in Abhängigkeit von der Zeit für den Datensatz Sim9. Die Amplitude des Treibers beträgt  $A = 0.3$ . Nach einem Sprung in der Amplitude des Treibers wird die Drift nach einer Zeit der Größenordnung 1 stationär. Bei der Diffusion ist der Einfluss des Treibers vernachlässigbar klein.

Aufgrund dieser Ergebnisse werden bei allen Schätzungen der übrigen Datensätze mit der Periodenlänge  $T = 630$  Zeitfenster der Größe einer Phase konstanter Amplitude der treibenden Kraft verwendet.

### 3.6.2. Der Grenzwert $\tau \rightarrow 0$

Wie in Abschnitt 1.8 bereits dargestellt wurde, wird im Allgemeinen der Grenzwert  $\tau \rightarrow 0$ , der in der Definition der Kramers-Moyal-Koeffizienten (1.68) vorkommt, bestimmt, indem der Schätzer (1.70) für diverse Zeitinkremente  $\tau = m\Delta t$  ausgewertet, und anschließend für  $\tau \rightarrow 0$  extrapoliert wird. Abbildung 3.8 zeigt im oberen Diagramm den *Finite-Time-Drift*  $\tilde{D}_\tau^{(1)}$  für  $S = 0$  in Abhängigkeit von  $\tau$ . Im Bereich zwischen 100 und 1000 Zeitschritten ist  $\tilde{D}_\tau^{(1)}$  annähernd konstant. Im Bereich zwischen 1 und 70 Zeitschritten zeigt sich dann ein annähernd lineares Verhalten. Es wäre möglich, in diesem Bereich linear für  $\tau = 0$  zu extrapolieren. Der Unterschied zwischen dieser Methode und der Methode, einfach den Wert für das kleinst mögliche  $\tau$  zu verwenden, ist jedoch so gering, dass der Aufwand sich nicht lohnen würde.

Bei der Diffusion zeigt sich ein anderes Verhalten, wie in Abb. 3.8 unten zu sehen ist. Die Abhängigkeit der *Finite-Time-Diffusion*  $\tilde{D}_\tau^{(2)}$  von  $\tau$  ist wesentlich stärker. Doch auch hier würde ein Extrapolationsverfahren keine signifikante Verbesserung liefern.

Da bei der aktuellen Analyse die Zeitreihen künstlich erzeugt werden und somit die "Samplingfrequenz" beliebig gewählt werden kann, befinden wir uns in einer komfortablen Situation. Bei der Analyse "echter" Daten ist es hingegen sehr wichtig, die Finite-Time-Koeffizienten in Abhängigkeit von  $\tau$  zu betrachten und gegebenenfalls zu extrapolieren, da ein zu großes  $\tau$  zu drastischen Abweichungen führen kann, wie Abbildung 3.8 verdeutlicht.

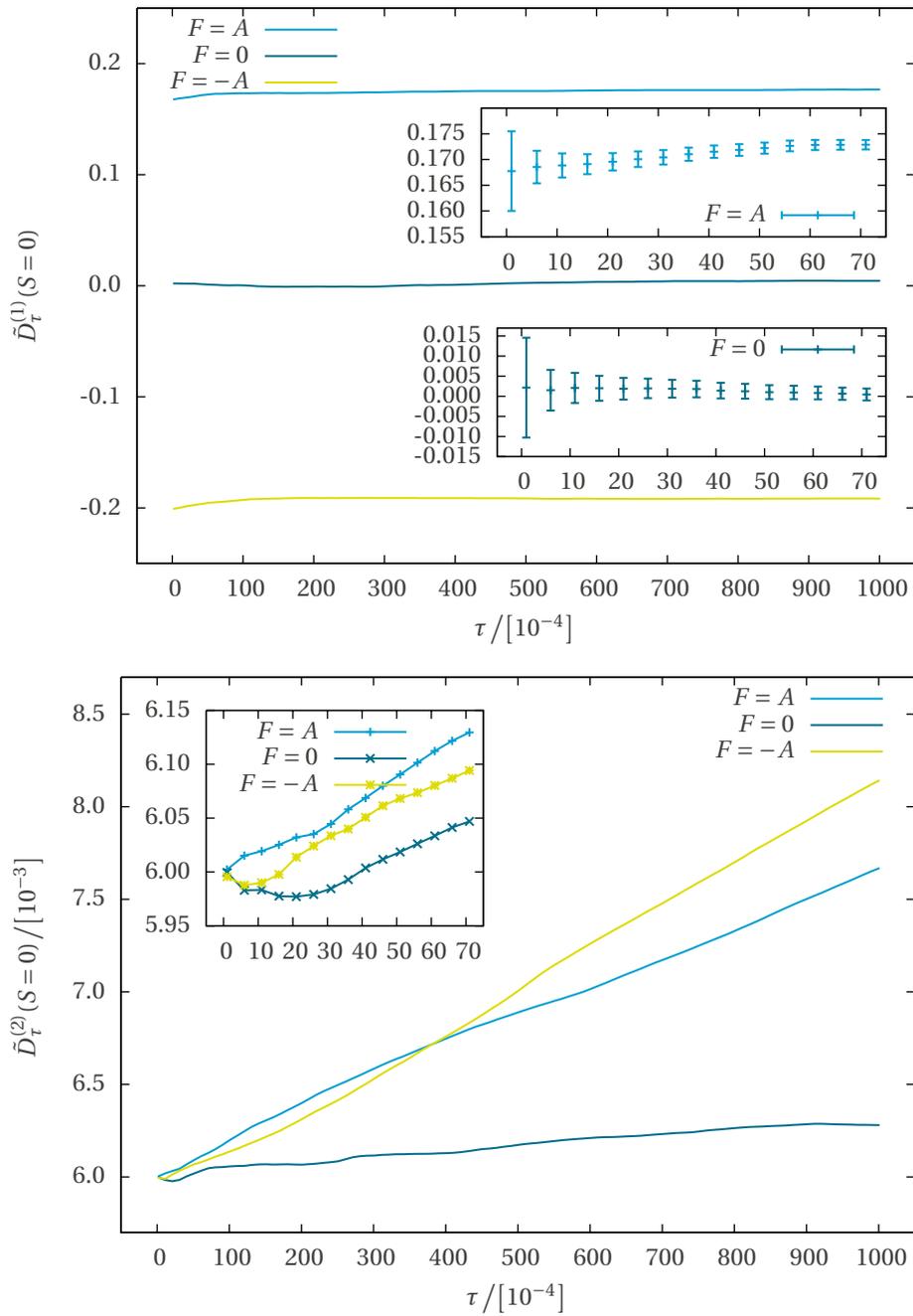
### 3.6.3. Diffusion und höhere Momente

Bei allen untersuchten Datensätzen ergibt sich ein in guter Näherung konstanter Diffusionskoeffizient von

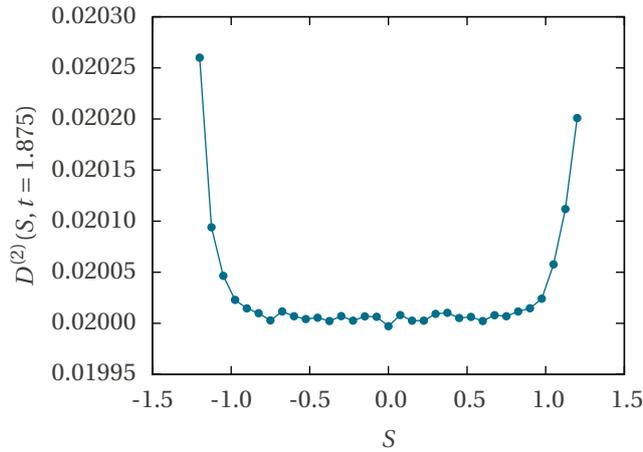
$$D^{(2)} = \frac{D}{N}. \quad (3.31)$$

Ein Beispiel zeigt Abb. 3.9 für den Datensatz Sim9 mit  $D = 0.2$  und  $N = 10$ . Dass die geschätzte Diffusion für größere Beträge von  $S$  ansteigt, liegt daran, dass in diesen Bereichen weniger Datenpunkte vorhanden sind. Die Erfahrung zeigt, dass bei geringerer Datendichte die geschätzte Diffusion meist nach oben abweicht. Die Ursache dafür ist jedoch noch nicht hinreichend geklärt<sup>1</sup>.

<sup>1</sup>Möglicherweise liegt die Ursache darin, dass die Größe, über die gemittelt wird, immer positiv ist, d. h., der Wertebereich ist nach unten beschränkt. Daher kann es in der Statistik größere Ausreißer nach oben als nach unten geben, was bei geringer Datendichte zu asymmetrischen Abweichungen führt.



**Abb. 3.8** Finite-Time-Drift und -Diffusion für konstantes  $S = 0$  in Abhängigkeit des Zeitinkrements  $\tau$ .



**Abb. 3.9** Diffusion für konstantes  $t = 1.875$  in Abhängigkeit von  $S$  für den Datensatz Sim9.

Im Allgemeinen führen Projektionen eines hochdimensionalen stochastischen Prozesses mit additivem Rauschen auf einen Unterraum kleinerer Dimension zu multiplikativem Rauschen. Dass dies hier nicht der Fall ist, ist ein wichtiges Resultat. Es stimmt mit der effektiven Langevin-Gleichung (3.18), die in [43] abgeleitet wurde, überein.

Um zu überprüfen, ob die Kramers-Moyal-Entwicklung tatsächlich abbricht, wird außerdem der vierte Koeffizient bestimmt. Dabei ergibt sich

$$\frac{D^{(4)}}{D^{(2)2}} \sim 10^{-4}, \quad (3.32)$$

d. h.,  $D^{(4)}$  ist praktisch null. Damit garantiert das Pawula-Theorem (s. [45]), dass auch der dritte Koeffizient sowie alle höheren verschwinden.

### 3.6.4. Die qualitative Form der Drift

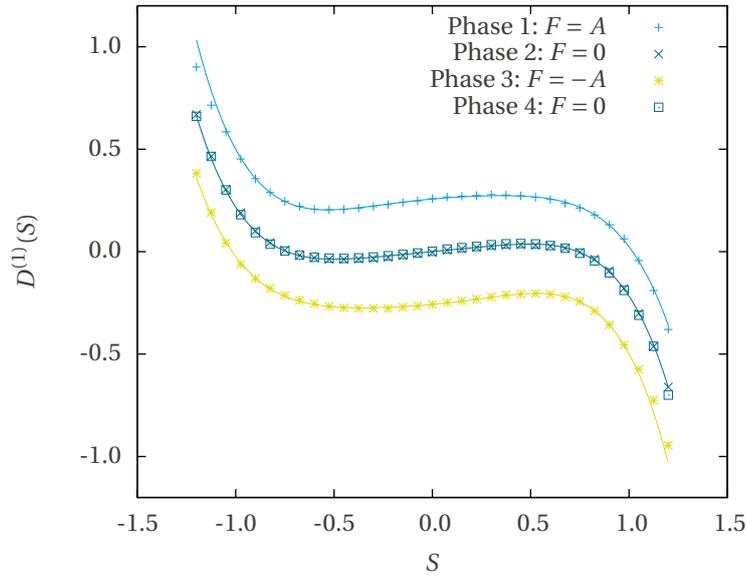
Abb. 3.10 zeigt die geschätzten Driftkoeffizienten für jede der vier Phasen für den Datensatz Sim9. Die Ergebnisse der zweiten und der vierten Phase, bei denen die Kraft  $F = 0$  ist, stimmen bis auf statistische Schwankungen überein. Man kann also weiterhin die zweite und vierte Phase (zumindest für große Periodenlängen) als verschiedene Realisationen des gleichen stochastischen Prozesses betrachten.

Die Drift der zweiten und vierten Phase  $D_{F=0}^{(1)}(S)$  lässt sich durch ein ungerades Polynom fünften Grades darstellen:

$$D_{F=0}^{(1)}(S) = aS + bS^3 + cS^5. \quad (3.33)$$

Die dunkelblaue Kurve in Abb. 3.10 entspricht einem Fit dieser Funktion<sup>2</sup>. Sollte die gaußsche Approximation, mit der Pikovsky und Zaikin das zweikomponentige System (3.12)

<sup>2</sup>Die Koeffizienten werden – wie auch bei allen folgenden Fits – mit Hilfe der Methode der kleinsten Quadrate bestimmt. Dazu wird eine Routine der Open Source-Software GnuPlot verwendet.



**Abb. 3.10** Die Drift für die vier Phasen der Periode des Treibers für den Datensatz Sim9. Jede Kurve entspricht dem Drift des letzten Zeitfensters der jeweiligen Periode. Die Symbole entsprechen den numerischen Ergebnissen, die Kurven stellen Fits von Polynomen fünften Grades dar.

aufstellen, eine gute Näherung darstellen, müssten bei der Anwendung des Versklavungsprinzips auf dieses System gemäß (3.33) Terme höherer Ordnung berücksichtigt werden.

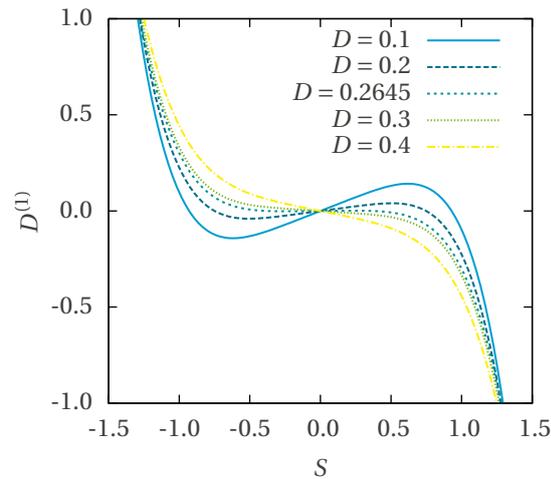
Die Driftkoeffizienten der ersten und dritten Phase lassen sich ebenfalls durch Polynome fünften Grades darstellen; in diesem Fall sind jedoch auch gerade Terme notwendig:

$$D_{F=\pm A}^{(1)}(S) = aS + bS^2 + cS^3 + dS^4 + eS^5 + f. \quad (3.34)$$

Die hellblaue und die grüne Kurve in Abb. 3.10 zeigen Fits dieser Funktion für die erste, bzw. die dritte Phase. An diesem Ergebnis wird deutlich, dass der Treiber  $F(t)$  auf den effektiven Drift der kollektiven Variable  $S(t)$  keinesfalls rein additiv wirkt, was in dem effektiven Modell (3.18) angenommen wird.

### 3.6.5. Der Gleichgewichtszustand

In der Vergangenheit wurde dem Gleichgewichtszustand des nicht getriebenen Systems ( $F = 0$ ) viel Aufmerksamkeit entgegengebracht. Dieser wurde von Desai und Zwanzig im thermodynamischen Grenzfall  $N \rightarrow \infty$  auf analytischem Wege diskutiert [7]. Dabei entdeckten sie verschiedene Regionen im Parameterraum der Kopplungsstärke  $\Theta$  und der Rauschamplitude  $D$ . In zweien dieser Regionen ist die pdf der kollektiven Variablen  $S(t)$  eindeutig und entweder unimodal oder bimodal, jeweils symmetrisch um  $S = 0$ . Im



**Abb. 3.11** Abhängigkeit der Drift von der Rauschamplitude. Die Abbildung zeigt Fits ungerader Polynome fünfter Ordnung der geschätzten Driftkoeffizienten der Phasen, in denen die Kraft  $F = 0$  ist. Es wurden die Datensätze `Sim2` bis `Sim6` verwendet (s. Tab. 3.1).

dritten Bereich gibt es zwei verschiedene pdfs, die beide unimodal und zentriert um einen Wert  $\pm S_0 \neq 0$  sind. Welche der beiden pdfs realisiert wird, hängt vom Anfangszustand des Systems ab.

Im Falle endlicher Systemgröße  $N$  kann diese Uneindeutigkeit der pdf nicht vorkommen. Der Gleichgewichtszustand bei endlichem  $N$  wird auf numerischem Wege ausführlich in [17] diskutiert. Es treten dabei erhebliche Abweichungen zum thermodynamischen Grenzfall auf. Bei fester Systemgröße  $N$  gibt es eine kritische Linie in der  $(D, \Theta)$ -Ebene. Für große  $D$ , oberhalb dieser Linie, ist die stationäre Verteilung unimodal, für kleine  $D$ , unterhalb der Linie, ist sie bimodal. Dieser Übergang spiegelt sich auch im Driftkoeffizienten der effektiven Langevin-Dynamik in Form einer Bifurkation wider.

Dies wird gezeigt, indem für festes  $N = 10$  und  $\Theta = 0.5$  die Drift für  $F = 0$  für variierende  $D$  bestimmt wird. Um ein besseres Sampling des Phasenraums zu erhalten, wird das System trotzdem mit treibender Kraft simuliert. Die Drift wird jedoch nur für die Phasen berechnet, in denen die Amplitude gerade null ist. Um diese Phasen zu verlängern, wird der Parameter  $r$  auf 0.1 gesetzt. Abb. 3.11 zeigt das Ergebnis. Die Bifurkation erfolgt in etwa bei einer Rauschamplitude von  $D = 0.2645$ . Erhöht man die Kopplungsstärke  $\Theta$ , findet die Bifurkation bei einer höheren Rauschamplitude statt. In dem zweikomponentigen System (3.12) liegt der Bifurkationspunkt bei  $\Theta = 3D$  (vgl. Abschnitt 3.2), was in offensichtlichem Widerspruch zu den hier erzielten Ergebnissen steht.

Das Bifurkationsverhalten lässt sich anschaulich folgendermaßen interpretieren: Ist die Amplitude des Rauschterms groß, finden viele Übergänge zwischen den beiden Potentialmulden statt. Sind in einer Mulde mehr Teilchen als in der anderen, wird die Übergangsrate von der stark bevölkerten Mulde in die schwach bevölkerte Mulde größer

sein als andersherum. Dieser diffusive Effekt treibt die kollektive Variable  $S$  in Richtung Null, so dass hier ein stabiler Fixpunkt in der effektiven Dynamik entsteht.

Eine starke Kopplung hat den Effekt, dass die verschiedenen Oszillatoren nahe beieinander gehalten werden. Ist eine Potentialmulde stärker bevölkert als die andere, werden die Oszillatoren aus der schwächer besetzten Mulde in die stärker besetzte gezogen. Dadurch wird der Fixpunkt in der Dynamik der kollektiven Variablen  $S$  bei Null instabil und in der Nähe der beiden Potentialmulden entstehen jeweils stabile Fixpunkte.

### 3.7. Rekonstruktion und Vergleich

Es stellt sich nun die Frage, wie gut das geschätzte eindimensionale Langevin-Modell den tatsächlichen Prozess reproduziert. Betrachtet man den Gleichgewichtszustand des ungetriebenen Systems, so lässt sich dessen numerisch geschätzte Wahrscheinlichkeitsdichte mit der stationären Lösung der FPG, in die die geschätzten Drift- und Diffusionskoeffizienten eingehen, vergleichen, s. Abschnitt 3.7.1. Für den getriebenen Fall werden als Vergleichsgrößen der kohärente und inkohärente Teil der Einpunkt-Korrelationsfunktion untersucht, s. Abschnitt 3.7.2. In beiden Fällen wird nur jeweils ein Satz von Parametern untersucht, nämlich der von Sim7.

#### 3.7.1. Stationäre pdf

Um die stationäre pdf zu schätzen, wird das  $N$ -dimensionale System wie in Abschnitt 3.3 beschrieben simuliert, wobei die Kraft  $F(t)$  zu null gesetzt wird. Sämtliche weiteren Parameter gleichen denen des Datensatzes Sim7, (s. Tab. 3.1). Die Variable  $S(t)$  wird erst nach  $6.3 \cdot 10^6$  Zeitschritten gespeichert, so dass davon ausgegangen wird, dass der Gleichgewichtszustand erreicht ist. Die pdf wird mit dem Schätzer (1.31) und der Kernfunktion (1.33) bestimmt.

In der stationären Lösung der geschätzten FPG ist der Diffusionskoeffizient gemäß (3.31) gegeben durch

$$D^{(2)} = 0.03. \quad (3.35)$$

Als Drift wird das Ergebnis der Schätzung für den Datensatz Sim7 für die Phasen mit  $F = 0$  verwendet. Die Parameter der Fitfunktion (3.33) sind Tab. 3.2 zu entnehmen. Damit

**Tab. 3.2** Koeffizienten der geschätzten Driftfunktionen für den Datensatz Sim7

Phase	Parameter					
	$a$	$b$	$c$	$d$	$e$	$f$
$F = 0$	-0.0292	-0.0673	-0.2354	-	-	-
$F = A$	-0.0287	-0.0269	-0.0402	0.0293	-0.2683	0.0993
$F = -A$	-0.0264	0.0460	-0.0267	-0.0654	-0.2997	-0.1000

lautet die stationäre Lösung der FPG:

$$f_{\text{stat}}(S) = \frac{\mathcal{N}}{D^{(2)}} \exp \left\{ \frac{1}{D^{(2)}} \left( \frac{a}{2} S^2 + \frac{b}{4} S^4 + \frac{c}{6} S^6 \right) \right\}. \quad (3.36)$$

Die Normierungskonstante  $\mathcal{N}$  wird mit Hilfe des Softwarepakets `Mathematica` numerisch bestimmt. Sie ergibt sich zu

$$\mathcal{N} \approx 0.02065 \quad (3.37)$$

Abb. 3.12 zeigt die stationäre pdf (3.36) aus der FPG zusammen mit der numerisch geschätzten pdf. Beide Verteilungen stimmen sehr gut überein.

### 3.7.2. Einpunkt-Korrelationsfunktion

Eine wichtige Größe, die bei der Analyse der stochastischen Resonanz in [5] untersucht wurde, ist die Einpunkt-Korrelationsfunktion

$$L(\tau) = \frac{1}{T} \int_0^T dt \langle S(t) S(t+\tau) \rangle, \quad (3.38)$$

wobei  $T$  die Periodenlänge des Treibers ist. Sie lässt sich aufspalten in einen kohärenten Teil

$$L_c(\tau) = \frac{1}{T} \int_0^T dt \langle S(t) \rangle \langle S(t+\tau) \rangle \quad (3.39)$$

mit der Eigenschaft  $L_c(\tau) = L_c(\tau + T)$  und einen inkohärenten Teil

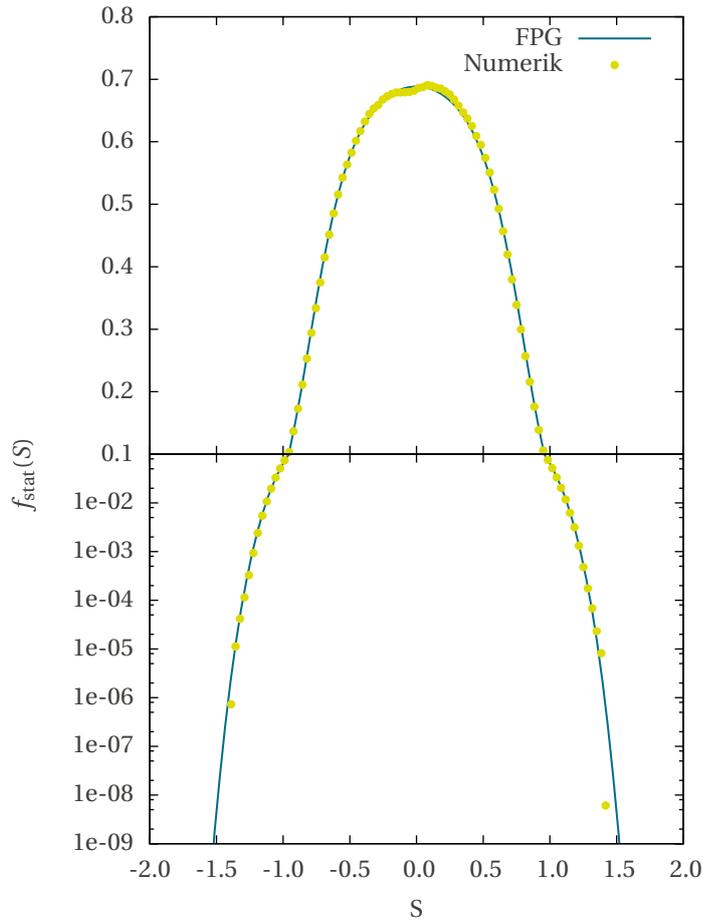
$$L_i(\tau) = \frac{1}{T} \int_0^T dt \langle S(t) S(t+\tau) \rangle - L_c(\tau), \quad (3.40)$$

der für große  $\tau$  asymptotisch gegen null geht.

Die beiden Funktionen  $L_c(\tau)$  und  $L_i(\tau)$  stellen eine gute Möglichkeit zum Vergleich zwischen dem ursprünglichen  $N$ -dimensionalen und dem effektiven eindimensionalen System dar.

Zum Zwecke dieses Vergleiches wird der effektive Prozess mit den geschätzten Kramers-Moyal-Koeffizienten simuliert. Dabei werden die Ergebnisse für den Datensatz `Sim7` verwendet. Wie bereits in Abschnitt 3.7.1 erwähnt, ergibt sich ein effektiver Diffusionskoeffizient von

$$D^{(2)} = 0.03. \quad (3.41)$$

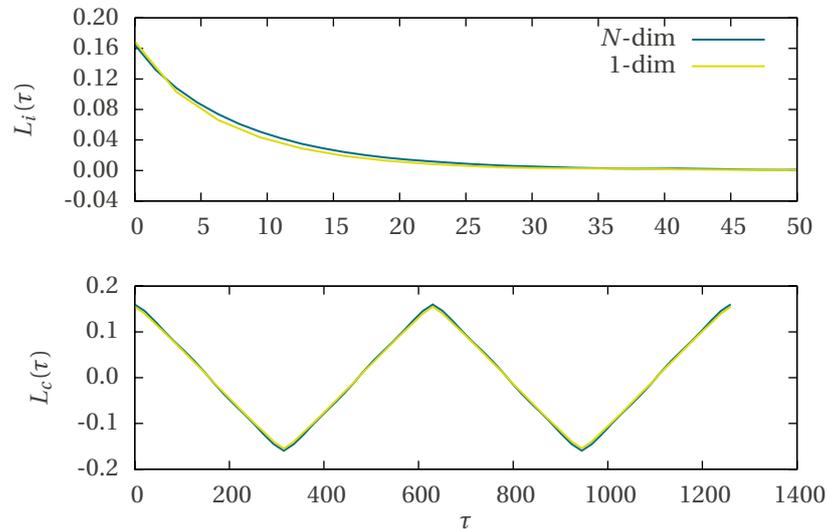


**Abb. 3.12** Die stationäre pdf des ungetriebenen Systems mit den Parametern von Sim7. Die grünen Punkten entsprechen der numerischen Schätzung; die blaue Kurve stellt die stationäre Lösung der FPG mit den geschätzten Drift- und Diffusionskoeffizienten dar. Im oberen Teil des Graphen ist die Ordinate linear skaliert, im unteren Teil logarithmisch.

Die Fitkoeffizienten der drei Driftfunktionen nach den Gln. (3.33) und (3.34) sind in Tab. 3.2 aufgelistet. Das Integrationsschema lautet dann

$$\begin{aligned}
 S(t + \Delta t) = & S(t) + \sqrt{2D^{(2)}} \Delta t \xi(t) \\
 & + \Delta t \cdot \begin{cases} D_{F=A}^{(1)}(S(t)); & t \bmod T \in [0, t_c) \\ D_{F=0}^{(1)}(S(t)); & t \bmod T \in [t_c, T/2) \\ D_{F=-A}^{(1)}(S(t)); & t \bmod T \in [T/2, T/2 + t_c) \\ D_{F=0}^{(1)}(S(t)); & t \bmod T \in [T/2 + t_c, T) \end{cases} \quad (3.42)
 \end{aligned}$$

Die Funktionen (3.39) und (3.40) werden sowohl für die rekonstruierte Zeitreihe der effektiven eindimensionalen Dynamik als auch für den Datensatz Sim7 numerisch berechnet. Abb. 3.13 zeigt das Ergebnis. Die Übereinstimmung zwischen dem ursprünglichen  $N$ -dimensionalen und dem geschätzten effektiven eindimensionalen Modell ist nahezu perfekt.



**Abb. 3.13** Der kohärente (unten) und inkohärente (oben) Teil der Einpunkt-Korrelationsfunktion  $L(\tau)$ . Die blaue Kurve entspricht jeweils dem ursprünglichen  $N$ -dimensionalen System und die grüne dem rekonstruierten effektiven eindimensionalen System. Die verwendeten Systemparameter entsprechen denen des Datensatzes Sim7.

### 3.8. Zusammenfassung und Ausblick

Fassen wir die wichtigsten Resultate zusammen. Zunächst wurde gezeigt, dass die Übergangswahrscheinlichkeitsdichten der kollektiven Variablen  $S(t)$ , Gl. (3.3), des  $N$ -dimensionalen Prozesses (3.1) die Chapman-Kolmogorov-Gleichung erfüllen. Dies gibt Anlass zu der Hoffnung, dass der Prozess  $S(t)$  Markov-Charakter hat und daher im Prinzip durch eine eindimensionale effektive Langevin-Gleichung bzw. eine entsprechende Fokker-Planck-Gleichung beschrieben werden kann.

Die Schätzung der Kramers-Moyal-Koeffizienten des Prozesses  $S(t)$  zeigt zunächst, dass diese bei hinreichend großen Periodenlängen des betrachteten Treibers  $F(t)$  während einer Phase konstanter Amplitude schnell genug stationär werden, so dass  $S(t)$  in guter Näherung durch vier hintereinandergeschaltete stationäre Prozesse, die sich periodisch wiederholen, modelliert werden kann. Da für die zweite und die vierte Phase des Treibers, in denen jeweils die Amplitude null ist, die Kramers-Moyal-Koeffizienten

gleich sind, reicht die Angabe von drei Sätzen zeitunabhängiger Drift- und Diffusionskoeffizienten aus, um den Prozess  $S(t)$  zu beschreiben.

Weiterhin ergibt sich, dass der Diffusionskoeffizient in guter Näherung in allen Phasen des Treibers gleich, und darüber hinaus unabhängig von  $S$  ist. Das Rauschen ist also additiv. Der vierte Kramers-Moyal-Koeffizient verschwindet praktisch, was nach dem Pawula-Theorem zur Folge hat, dass die Kramers-Moyal-Entwicklung auch tatsächlich abbricht und eine Fokker-Planck-Gleichung ergibt.

Was die qualitative Form der Drift betrifft, gibt es zwei wesentliche Unterschiede zu dem Modell von Pikovsky und Zaikin. Der Erste besteht darin, dass ein  $S^5$ -Term in allen Phasen des Treibers unverzichtbar ist. Der zweite Unterschied besteht in dem Einfluss des Treibers auf die Driftfunktion  $D^{(1)}(S)$ . Im Modell von Pikovsky und Zaikin ist dieser Einfluss rein additiv, d. h., es gilt:

$$D^{(1)}(S) = D_{F=0}^{(1)}(S) + F(t) . \quad (3.43)$$

Nach dem hier erzielten Ergebnis hingegen beeinflusst der Treiber auch den  $S$ -abhängigen Teil der Drift (vgl. Gl. (3.34)) und könnte daher als *multiplikativ* bezeichnet werden. Diese beiden Punkte unterscheiden die Drift der Variablen  $S(t)$  auch qualitativ von der eines einzelnen Oszillators ohne Kopplung, welcher durch die Langevin-Gleichung

$$\dot{x} = x - x^3 + F(t) + \sqrt{2D}\xi(t) \quad (3.44)$$

beschrieben wird. Dies liefert eine Erklärung dafür, dass die SR-Effekte von  $S(t)$  bei schwacher Amplitude der treibenden Kraft wesentlicher stärker sind, als die SR-Effekte eines einzelnen Oszillators bei gleichem Treiber, was in [40] beobachtet wurde (vgl. Abschnitt 3.1).

Es wurde gezeigt, dass sich die beobachteten Phasenübergänge zwischen uni- und bimodaler pdf von  $S$  im Gleichgewichtszustand, die bei Variation der Parameter  $D$  und  $\Theta$  auftreten, im Drift als Bifurkation wiederfinden.

Die stationäre pdf, die sich aus der Fokker-Planck-Gleichung mit den geschätzten Drift- und Diffusionskoeffizienten des ungetriebenen Systems ergibt, stimmt mit der numerisch gemessenen pdf von  $S$  des tatsächlichen  $N$ -dimensionalen Systems überein. Im getriebenen Fall reproduziert das geschätzte effektive eindimensionale System korrekt die Einpunkt-Korrelationsfunktion des Originalsystems.

Dass die Beschreibung des stochastischen Prozesses  $S(t)$  durch eine effektive eindimensionale Langevin-Gleichung gut funktioniert, ist im Kontext des in Abschnitt 3.2 erwähnten hierarchischen Gleichungssystems der fluktuierenden Momente so zu verstehen, dass die Letzteren von der Variablen  $S(t)$  versklavt werden. Dabei ist keineswegs garantiert, dass diese Versklavung in allen Bereichen des Parameterraumes von  $N$ ,  $D$  und  $\Theta$  funktioniert. Die in Abschnitt 3.7 dargestellten Tests wurden aufgrund des hohen Aufwandes nur für *einen* Satz von Parametern durchgeführt. Ebenso wurde die CKG nur für zwei Sätze von Parametern überprüft. Es wäre also noch interessant zu untersuchen, ob das Verfahren in anderen Parameterbereichen scheitert. In diesem Zusammenhang wäre

weiterhin interessant, die im nächsten Kapitel entwickelte Methode zur Rekonstruktion von invarianten Mannigfaltigkeiten anzuwenden, um zu sehen, wie die fluktuierenden Momente sich in Abhängigkeit der Systemparameter verhalten.

Unter der Annahme, dass die in Abschnitt 3.2 dargestellte gaußsche Approximation eine gute Näherung darstellt, was nach [40] zumindest in bestimmten Parameterbereichen der Fall zu sein scheint, könnte man versuchen, das Versklavungsprinzip in höherer Ordnung auf das System (3.12) anzuwenden, um den wichtigen  $S^5$ -Term zu erhalten. Die Ergebnisse bei der Untersuchung des Gleichgewichtszustandes in Abschnitt 3.6.5 zeigen jedoch, dass schon der Bifurkationspunkt  $\theta = 3D$  des Systems (3.12) nicht demjenigen des echten Systems entspricht. Man müsste also zusätzlich fluktuierende Momente höherer Ordnung berücksichtigen. An dieser Stelle ist es fraglich, ob der dadurch erzielte Erkenntnisgewinn noch den rechentechnische Aufwand rechtfertigen würde.

## 4. Identifikation von Ordnungsparametern und invarianten Mannigfaltigkeiten aus Daten

Im vorherigen Kapitel wurde eine effektive eindimensionale Ordnungsparametergleichung für ein hochdimensionales Modellsystem aus (synthetisch erzeugten) Daten geschätzt. In diesem Fall waren sowohl die Bewegungsgleichungen des mikroskopischen Modellsystems als auch der Ordnungsparameter bekannt. In diesem Kapitel wird der Fall betrachtet, dass Messdaten eines Systems vorliegen, bei dem weder die zugrunde liegende mikroskopische Dynamik noch die Ordnungsparameter bekannt sind. Es wird aber vorausgesetzt, dass sich das System in der Nähe einer Instabilität befindet und wenige Ordnungsparameter die Dynamik dominieren. Das Ziel bei einer Analyse solcher Messdaten besteht darin, die Ordnungsparameter, die versklavten Modenamplituden und die invariante Mannigfaltigkeit, um die die versklavten Amplituden fluktuieren, zu identifizieren und die Ordnungsparameterdynamik durch Langevin-Gleichungen zu beschreiben. Die Methodik des letzten Schrittes ist bereits bekannt und wurde im vorangegangenen Kapitel angewendet. Für die ersten Schritte werden neue Ansätze, die sich aus der Theorie der Synergetik (s. Kapitel 2) begründen, entwickelt und an einem einfachen Modellsystem getestet.

### 4.1. Motivation

Eine mögliche Anwendung der hier entwickelten Methoden besteht darin, ausgedehnte Systeme zu beschreiben, die räumliche oder raumzeitliche Muster ausbilden. Solche Systeme werden üblicherweise durch partielle Differentialgleichungen der Form

$$\dot{\mathbf{q}}(\mathbf{x}, t) = N(\mathbf{q}(\mathbf{x}, t), \nabla, \boldsymbol{\alpha}, \mathbf{x}, t) \quad (4.1)$$

beschrieben. Dabei ist  $\mathbf{q}(\mathbf{x}, t)$  ein Vektorfeld, dessen Komponenten messbare Größen darstellen, die das betrachtete System charakterisieren. Eine solche messbare Größe kann z. B. eine Geschwindigkeitskomponente einer Flüssigkeitsströmung, die Konzentration einer chemischen Substanz oder die Anzahl einer Bakterienart pro Flächenelement sein. Die zeitliche Entwicklung von  $\mathbf{q}(\mathbf{x}, t)$  wird durch die Funktion  $N$  bestimmt, die von  $\mathbf{q}(\mathbf{x}, t)$  selbst, räumlichen Ableitungen von  $\mathbf{q}(\mathbf{x}, t)$ , die durch den  $\nabla$ -Operator symbolisiert werden, einem Satz von externen Kontrollparametern  $\boldsymbol{\alpha}$  und schließlich explizit vom Ort  $\mathbf{x}$  und der Zeit  $t$  abhängen kann.

In der Praxis bestehen raumzeitliche Daten aus einer Menge von Observablen, die an verschiedenen diskreten Punkten  $\mathbf{x}_i$  im Raum zu verschiedenen diskreten Zeitpunkten  $t_j$  gemessen wurden. Von den unendlich vielen Freiheitsgraden des Vektorfeldes sind also experimentell nur endlich viele Freiheitsgrade zugänglich. Das gleiche gilt für numerische Simulationen. Daher ersetzen wir das Vektorfeld  $\mathbf{q}(\mathbf{x}, t)$  durch einen Zustandsvektor  $\mathbf{q}(t_j)$ , dessen  $i$ te Komponente  $q_i(t_j)$  der zur Zeit  $t_j$  am Ort  $\mathbf{x}_i$  gemessenen Observable<sup>1</sup> entspricht. Dabei gehen wir davon aus, dass die Messungen an verschiedenen Orten jeweils zu gleichen Zeiten erfolgen, und nehmen an, dass der zeitliche Abstand zwischen zwei Messungen konstant ist.

Es wird angenommen, dass die zeitliche Entwicklung dieses Zustandsvektors durch ein System stochastischer Bewegungsgleichungen beschrieben werden kann:

$$\dot{\mathbf{q}}(t) = \mathbf{N}(\mathbf{q}(t), \{\varepsilon\}) + \mathbf{H}(\mathbf{q})\boldsymbol{\xi}(t). \quad (4.2)$$

Dabei ist  $\mathbf{N}$  eine nichtlineare Funktion der Komponenten von  $\mathbf{q}$ , die von einem Satz von Kontrollparametern  $\{\varepsilon\}$  abhängt. Bei dem Term  $\boldsymbol{\xi}(t)$  handelt es sich um einen gaußverteilten Rauschterm mit den Eigenschaften (1.51). Die Diffusionsmatrix  $\mathbf{H}(\mathbf{q})$  hat üblicherweise Diagonalgestalt.

In der Nähe von Instabilitäten können wir den Zustandsvektor durch eine Überlagerung von stabilen und instabilen räumlichen Moden darstellen:

$$\mathbf{q}(t) = \sum_k s_k(t) \mathbf{q}_{s,k} + \sum_\ell u_\ell(t) \mathbf{q}_{u,\ell}. \quad (4.3)$$

Nach dem in Abschnitt 2.3 beschriebenen Versklavungsprinzip können die Amplituden  $s_k(t)$  der stabilen Moden  $\mathbf{q}_{s,k}$  durch die Amplituden  $u_\ell(t)$  der instabilen Moden  $\mathbf{q}_{u,\ell}$  ausgedrückt werden, so dass sich die Bewegungsgleichungen auf

$$\dot{\mathbf{u}}(t) = \mathbf{f}(\mathbf{u}(t), \mathbf{s}(\mathbf{u}(t))) + \mathbf{G}(\mathbf{u}(t), \mathbf{s}(\mathbf{u}(t)))\boldsymbol{\xi}(t), \quad \mathbf{u}(t) = \{u_\ell(t)\} \quad (4.4)$$

reduzieren. Wir bezeichnen dann die Amplituden  $u_\ell(t)$  als Ordnungsparameter. Da die räumlichen Moden im Allgemeinen nicht orthogonal sind, führen wir die adjungierten Moden  $\mathbf{q}_{u,i}^\dagger$  und  $\mathbf{q}_{s,i}^\dagger$  ein, so dass

$$\begin{aligned} \mathbf{q}_{u,i}^\dagger \cdot \mathbf{q}_{u,j} &= \delta_{ij}, & \mathbf{q}_{s,i}^\dagger \cdot \mathbf{q}_{u,j} &= 0, \\ \mathbf{q}_{u,i}^\dagger \cdot \mathbf{q}_{s,j} &= 0, & \mathbf{q}_{s,i}^\dagger \cdot \mathbf{q}_{s,j} &= \delta_{ij} \end{aligned} \quad (4.5)$$

gilt, wobei  $\delta_{ij}$  das Kronecker-Delta bezeichnet. Damit ergibt sich für die Ordnungsparameter und die Amplituden der stabilen Moden

$$u_i(t) = \mathbf{q}_{u,i}^\dagger \cdot \mathbf{q}(t), \quad (4.6)$$

$$s_j(t) = \mathbf{q}_{s,j}^\dagger \cdot \mathbf{q}(t). \quad (4.7)$$

<sup>1</sup>Bei mehreren verschiedenen Observablen müssen weitere Indizes eingeführt werden, was jedoch keinen großen Unterschied bedeutet.

Um das betrachtete System mit Hilfe von Ordnungsparametern beschreiben zu können, benötigen wir also die adjungierten Moden  $\mathbf{q}_{u,i}^\dagger$  und  $\mathbf{q}_{s,i}^\dagger$ , die durch  $\mathbf{s}(\mathbf{u})$  beschriebene invariante Mannigfaltigkeit sowie den Driftvektor  $\mathbf{f}(\mathbf{u})$  und die Diffusionsmatrix  $\mathbf{G}(\mathbf{u})$  der Ordnungsparametergleichung (4.4).

Ohne die Berücksichtigung von Rauschen wurde bereits in den 1990er Jahren von C. Uhl, R. Friedrich und H. Haken eine Methode zur Lösung dieses Problems entwickelt [53, 54]. Dabei wurden für die Funktionen  $s_i(\mathbf{u})$  und  $\mathbf{f}(\mathbf{u})$  parametrische Ansätze gemacht, deren Parameter sowie die Komponenten der adjungierten Moden  $\mathbf{q}_{u,i}^\dagger$  und  $\mathbf{q}_{s,i}^\dagger$  durch ein Extremalprinzip bestimmt wurden.

Die in dieser Arbeit entwickelten Methoden unterscheiden sich davon zum einen durch die explizite Einbeziehung von Rauschen und zum anderen dadurch, dass keine parametrischen Ansätze für die zu bestimmenden Funktionen gemacht werden.

## 4.2. Testsystem

Als Testsystem für die entwickelten Ansätze wird das folgende zweikomponentige Modellsystem verwendet:

$$\dot{q}_1(t) = \varepsilon q_1(t) - q_1(t)q_2(t) + \sqrt{2D_1}\xi_1(t) \quad (4.8a)$$

$$\dot{q}_2(t) = -\gamma q_2(t) + q_1^2(t) + \sqrt{2D_2}\xi_2(t). \quad (4.8b)$$

Dabei sind  $\varepsilon, \gamma > 0$  und  $\varepsilon \ll \gamma$ . Dies entspricht dem bereits in den Abschnitten 2.3.2 und 2.3.3 behandelten Haken-Zwanzig-System mit Rauschen. Dementsprechend handelt es sich bei  $q_1 = u$  um einen Ordnungsparameter und bei  $q_2 = s$  um eine schnell relaxierende Modenamplitude. Die gesuchten adjungierten Moden sind daher

$$\mathbf{q}_u^\dagger = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{q}_s^\dagger = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (4.9)$$

Die Amplitude  $s(t)$  fluktuiert nach Gln. (2.61) und (2.38) um die invariante Mannigfaltigkeit, die durch

$$s(u) = \frac{1}{\gamma}u^2 + \mathcal{O}(u^3) \quad (4.10)$$

beschrieben wird. Die stochastische effektive Ordnungsparametergleichung lautet nach (2.60) in adiabatischer Näherung

$$\dot{u}(t) = \varepsilon u(t) - \frac{u^3(t)}{\gamma} + \sqrt{2D_1}\xi(t). \quad (4.11)$$

Dementsprechend ist die gesuchte Funktion

$$f(u) = \varepsilon u - \frac{u^3}{\gamma}. \quad (4.12)$$

Die Matrix  $\mathbf{G}$  hat nur einen Eintrag:

$$G = \sqrt{2D_1}. \quad (4.13)$$

Zunächst wird das Problem der Identifikation der adjungierten Moden  $\mathbf{q}_u^\dagger$  und  $\mathbf{q}_s^\dagger$  angegangen. Dazu werden in den Abschnitten 4.3 und 4.4 zwei verschiedene Ansätze diskutiert. Beiden Ansätzen gemein ist, dass jeweils eine bestimmte Größe in Abhängigkeit einer Testmode  $\mathbf{q}^\dagger$  aus einer Zeitreihe  $\mathbf{q}(t)$  geschätzt wird. Die Zeitreihe wird numerisch erzeugt, was im folgenden Abschnitt 4.2.1 erläutert wird. Durch Extremalisierung dieser Größe bezüglich  $\mathbf{q}^\dagger$  werden  $\mathbf{q}_u^\dagger$  und  $\mathbf{q}_s^\dagger$  identifiziert. Die Extremalisierung erfolgt dabei immer unter der Nebenbedingung, dass die Moden  $\mathbf{q}^\dagger$  normiert sind:

$$\mathbf{q}^\dagger \cdot \mathbf{q}^\dagger = 1. \quad (4.14)$$

#### 4.2.1. Erzeugung der Zeitreihe

Die numerische Erzeugung der Zeitreihe erfolgt durch das explizite Euler-Verfahren:

$$q_1(t + \Delta t) = (\varepsilon q_1(t) - q_1(t)q_2(t))\Delta t + \sqrt{2D_1\Delta t}\xi_1(t) \quad (4.15a)$$

$$q_2(t + \Delta t) = (-\gamma q_2(t) + q_1^2(t))\Delta t + \sqrt{2D_2\Delta t}\xi_2(t). \quad (4.15b)$$

Die standardnormalverteilten Zufallszahlen  $\xi_1(t)$  und  $\xi_2(t)$  werden dabei auf die bereits in Abschnitt 3.3 beschriebene Art erzeugt. Die verwendeten Parameter sind in Tabelle 4.1 aufgelistet. Die Parameter  $\varepsilon$  und  $\gamma$  werden so gewählt, dass die Voraussetzung des Versklavungsprinzips erfüllt ist: Es gilt  $|\varepsilon/\gamma| = 0.1 \ll 1$ . Die Rauschamplituden sind derart festgelegt, dass einerseits der Determinismus noch dominant ist, andererseits aber das System noch hinreichend stark aus seinen Ruhelagen herausgetrieben wird. Für weitergehende Untersuchungen wäre es interessant, die Funktionalität der im Folgenden diskutierten Ansätze in Abhängigkeit dieser Parameter zu untersuchen.  $N$  ist die Anzahl der Datenpunkte der erzeugten Zeitreihe.

**Tab. 4.1** Simulationsparameter

Parameter	Wert
$\Delta t$	$10^{-4}$
$\varepsilon$	0.1
$\gamma$	1
$D_1$	$10^{-3}$
$D_2$	$10^{-3}$
$N$	$10^8$

### 4.3. Der Driftansatz

#### 4.3.1. Motivation

Der Driftansatz basiert direkt auf der Trennung der Zeitskalen zwischen Ordnungsparametern und versklavten Amplituden. Dass sich die Ersteren wesentlich langsamer verändern als die Letzteren, wird sichtbar, wenn man das Driftvektorfeld schätzt. Die Komponenten des Driftvektors sind hier definiert als

$$D_i^{(1)}(\mathbf{q}) = \lim_{\tau \rightarrow 0} \frac{1}{\tau} \langle q_i(t+\tau) - q_i(t) \rangle_t \Big|_{\mathbf{q}(t)=\mathbf{q}}, \quad i = 1, 2. \quad (4.16)$$

Dabei stellt  $\langle \dots \rangle_t$  ein Zeitmittel dar. Somit lautet der entsprechende Schätzer:

$$\hat{D}_i^{(1)}(\mathbf{q} \in I_{jk}) = \frac{1}{\Delta t} \frac{\sum_{\ell=1}^{N-1} [q_i(t_{\ell+1}) - q_i(t_\ell)] \Theta_{jk}(\mathbf{q}(t_\ell))}{\sum_{\ell=1}^{N-1} \Theta_{jk}(\mathbf{q}(t_\ell))}, \quad i = 1, 2. \quad (4.17)$$

Dabei sind

$$I_{jk} = [q_{1,0} + (j-1)\Delta q, q_{1,0} + j\Delta q_1] \times [q_{2,0} + (k-1)\Delta q, q_{2,0} + k\Delta q_2] \quad (4.18)$$

zweidimensionale Bins und

$$\Theta_{jk}(\mathbf{q}) = \begin{cases} 1, & \text{für } \mathbf{q} \in I_{jk}, \\ 0 & \text{sonst.} \end{cases} \quad (4.19)$$

ist die entsprechende Verallgemeinerung von (1.71).

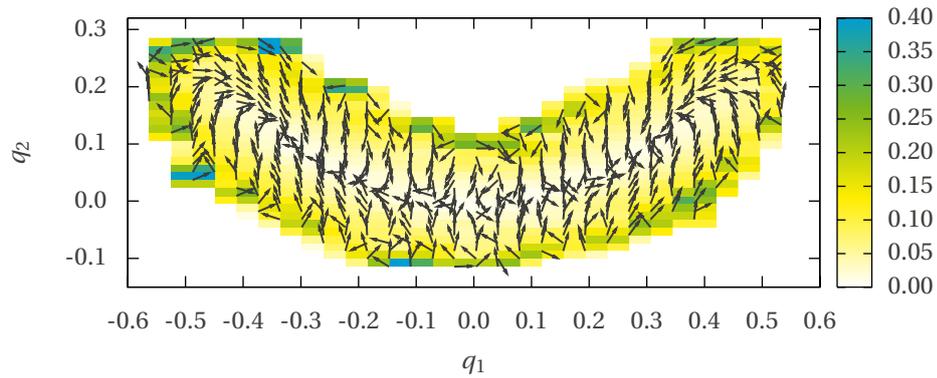
Abb. 4.1 zeigt das geschätzte Driftvektorfeld unseres Testsystems, das auch in der Abbildung auf der Titelseite zusammen mit der farbkodierten jpdf  $f_2(u, s)$  sowie der Approximation der invarianten Mannigfaltigkeit in vierter Ordnung,  $s(u) = \frac{1}{\gamma} u^2 + \frac{2}{\gamma} u^4$ , zu sehen ist. Die Beträge der Komponenten der Vektorpfeile in  $q_2$ -Richtung sind im Mittel größer als die in  $q_1$ -Richtung, da  $q_1$  sich langsamer verändert als  $q_2$ .

Um diese Beobachtung zur Identifikation von  $\mathbf{q}_u^\dagger$  und  $\mathbf{q}_s^\dagger$  zu verwenden, wird die Größe

$$\zeta(\mathbf{q}^\dagger) := \left\langle \left| \mathbf{q}^\dagger \cdot \mathbf{D}^{(1)}(\mathbf{q}) \right| \right\rangle_{\mathbf{q}} \quad (4.20)$$

definiert. Dabei bezeichnet  $\langle \dots \rangle_{\mathbf{q}}$  das Mittel über den Phasenraum, also praktisch das Mittel über alle Bins, für die der Driftvektor geschätzt wurde.  $\zeta(\mathbf{q}^\dagger)$  dient als Maß für das Inverse der typischen Zeitskala der Dynamik projiziert auf  $\mathbf{q}^\dagger$ . Es ist wichtig, den Betrag des Skalarproduktes zu verwenden, damit Driftvektoren, die in entgegengesetzte Richtungen zeigen, sich nicht gegenseitig herausmitteln. Es wird folgende Hypothese formuliert:

$$\zeta(\mathbf{q}^\dagger) = \begin{cases} \text{minimal,} & \text{für } \mathbf{q}^\dagger = \mathbf{q}_u^\dagger \\ \text{maximal,} & \text{für } \mathbf{q}^\dagger = \mathbf{q}_s^\dagger \end{cases}. \quad (4.21)$$



**Abb. 4.1** Das mittels (4.17) geschätzte Driftvektorfeld für das System (4.8). Die Pfeile sind auf Einheitslänge normiert. Der Betrag der Vektoren ist farblich kodiert dargestellt. An den Stellen, an denen keine Pfeile zu sehen sind, gibt es nicht genügend Datenpunkte in der simulierten Zeitreihe. Da die Richtungen der Vektoren von Bedeutung sind, wurde die Skalierung der Achsen so gewählt, dass gleiche Intervalle auf beiden Achsen jeweils die gleiche Länge haben.

#### 4.3.2. Numerische Umsetzung

Die Hypothese (4.21) wird für das in Abschnitt 4.2 beschriebene System getestet, indem  $\zeta(\mathbf{q}^\dagger)$  bezüglich  $\mathbf{q}^\dagger$  extremalisiert wird. Der Nebenbedingung (4.14) wird dabei durch den Ansatz

$$\mathbf{q}^\dagger = \mathbf{q}^\dagger(\phi) = \begin{pmatrix} \cos \phi \\ \sin \phi \end{pmatrix} \quad (4.22)$$

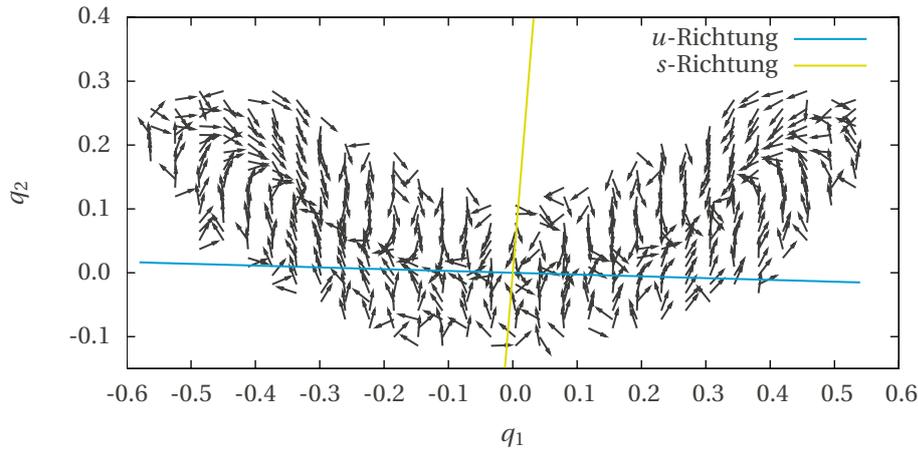
Rechnung getragen. Das nun zu lösende Problem besteht darin, numerisch das Minimum und das Maximum einer reellen Größe bezüglich einer reellen Variable aus einem beschränkten Intervall zu bestimmen. Bei dessen Lösung wird der Weg des geringsten Aufwandes unter Verzicht auf numerische Präzision und Eleganz eingeschlagen. Das Intervall  $[-\pi/2, \pi/2]$  wird mit einer Schrittweite von  $\Delta\phi = 5 \cdot 10^{-4}$  abgetastet und dabei für jedes  $\phi$  die Größe  $\zeta(\mathbf{q}^\dagger(\phi))$  berechnet. Anschließend werden Minimum und Maximum von  $\zeta(\mathbf{q}^\dagger(\phi))$  bestimmt.

Für die Anwendung auf Probleme mit mehr als zwei Dimensionen müsste ein etwas fortgeschritteneres Optimierungsverfahren verwendet werden, wie beim Informationsansatz geschehen (s. Abschnitt 4.4.2).

#### 4.3.3. Ergebnis

Aus dem Minimum respektive Maximum ergibt sich:

$$\mathbf{q}_u^\dagger = \begin{pmatrix} -0.99959 \\ 0.028 \end{pmatrix}, \quad \mathbf{q}_s^\dagger = \begin{pmatrix} 0.081 \\ 0.9967 \end{pmatrix}. \quad (4.23)$$



**Abb. 4.2** Resultierende Projektionsrichtungen beim Driftansatz. Die blaue Gerade entspricht der Projektionsrichtung des Ordnungsparameters  $u$ . Die zugehörige Geradengleichung lautet  $q_2(q_1) = (q_{u,2}^\dagger / q_{u,1}^\dagger) q_1$ . Die grüne Gerade mit der Gleichung  $q_2(q_1) = (q_{s,2}^\dagger / q_{s,1}^\dagger) q_1$  entspricht der Projektionsrichtung der versklavten Amplitude  $s$ . Das Driftvektorfeld ist das gleiche wie in Abb. 4.1 und dient nur der Orientierung.

Um das Ergebnis graphisch zu veranschaulichen, werden die damit verbundenen Projektionsrichtungen zusammen mit dem Driftvektorfeld in Abb. 4.2 dargestellt. Die Projektionsrichtungen lassen sich durch Ursprungsgeraden der Steigung

$$q_{i,2}^\dagger / q_{i,1}^\dagger \quad \text{mit } i = u, s \quad (4.24)$$

darstellen.

Die geringe Abweichung vom gesuchten Ergebnis (4.9) lässt sich vermutlich auf die durch die begrenzte Anzahl von Datenpunkten bedingte Ungenauigkeit der Schätzung des Driftvektorfeldes zurückführen.

## 4.4. Der Informationsansatz

### 4.4.1. Motivation

Mit den Argumenten aus Abschnitt 2.5.2 lässt sich annehmen, dass in der Nähe von Instabilitäten die Information der Ordnungsparameter groß, und die der versklavten Amplituden klein ist. Letztere wurde in Gl. (2.94) definiert als

$$i_s(\mathbf{u}) = - \sum_j \int ds_j p(s_j | \mathbf{u}) \log p(s_j | \mathbf{u}) . \quad (4.25)$$

Für die jpdf aller versklavten Amplituden gilt

$$\begin{aligned}
 f(\mathbf{s}) &= \int d\mathbf{u} f(\mathbf{u}, \mathbf{s}) \\
 &= \int d\mathbf{u} f(\mathbf{u}) p(\mathbf{s}|\mathbf{u}) \\
 &= \langle p(\mathbf{s}|\mathbf{u}) \rangle_{\mathbf{u}} .
 \end{aligned} \tag{4.26}$$

Das heißt,  $f(\mathbf{s})$  kann als Erwartungswert der bedingten Verteilung  $p(\mathbf{s}|\mathbf{u})$  bzgl.  $\mathbf{u}$  aufgefasst werden. Analog gilt für die pdf einer einzelnen, oBdA der ersten versklavten Amplitude:

$$\begin{aligned}
 f(s_1) &= \int d\mathbf{u} f(\mathbf{u}) \int ds_2 \dots ds_n p(\mathbf{s}|\mathbf{u}) \\
 &= \int d\mathbf{u} f(\mathbf{u}) \int ds_2 \dots ds_n \prod_{i=1}^n p(s_i|\mathbf{u}) \\
 &= \int d\mathbf{u} f(\mathbf{u}) p(s_1|\mathbf{u}) \\
 &= \langle p(s_1|\mathbf{u}) \rangle_{\mathbf{u}} .
 \end{aligned} \tag{4.27}$$

Wir betrachten noch einmal die Gleichungen (2.83) und (2.84). Diese lauten:

$$p(s_i|\mathbf{u}) = \exp \left\{ -\hat{\lambda}_i^{(s)} (s_i - g_i(\mathbf{u}))^2 + \hat{\lambda}_i^{(s)} g_i^2(\mathbf{u}) \right\} , \tag{4.28}$$

$$g_i(\mathbf{u}) = \frac{1}{2\hat{\lambda}_i^{(s)}} \left[ 3 \sum_{jk} \hat{\lambda}_{ijk}^{(s)} u_j u_k + 4 \sum_{jk\ell} \hat{\lambda}_{ijk\ell}^{(s)} u_j u_k u_\ell \right] . \tag{4.29}$$

In der Nähe von Instabilitäten können wir annehmen, dass die Parameter  $\hat{\lambda}_i^{(s)}$  groß sind. Daraus folgt zum einen, dass die cpdfs  $p(s_i|\mathbf{u})$  scharf um die Funktionen  $g_i(\mathbf{u})$  lokalisiert sind. Aufgrund der Vorfaktoren  $1/2\hat{\lambda}_i^{(s)}$  der Funktionen  $g_i(\mathbf{u})$  folgt daraus zum anderen, dass sich die Letzteren für kleine  $\mathbf{u}$  nur langsam ändern, und damit auch die  $\mathbf{u}$ -Abhängigkeit der cpdfs  $p(s_i|\mathbf{u})$  nur sehr schwach ist. Nimmt man beide Folgerungen zusammen, lässt sich mit Gl. (4.27) ableiten, dass auch die pdfs  $f(s_i)$  scharf lokalisiert sind, was zur Folge hat, dass die entsprechende Information

$$i = - \int ds_i f(s_i) \log f(s_i) \tag{4.30}$$

klein, oder stark negativ ist.

Aus dieser Überlegung kann ein weiteres Kriterium zur Identifikation der adjungierten Moden abgeleitet werden. Dazu wird die Information in Abhängigkeit von  $\mathbf{q}^\dagger$  definiert:

$$i(\mathbf{q}^\dagger) = - \int f_{X_{\mathbf{q}^\dagger}}(x) \log f_{X_{\mathbf{q}^\dagger}}(x) dx \tag{4.31}$$

Der stochastische Prozess  $X_{\mathbf{q}^\dagger}$  ist dabei definiert als

$$X_{\mathbf{q}^\dagger}(t) = \mathbf{q}^\dagger \cdot \mathbf{q}(t). \quad (4.32)$$

Es wird nun die folgende Annahme getroffen:

$$i(\mathbf{q}^\dagger) = \begin{cases} \text{maximal,} & \text{für } \mathbf{q}^\dagger = \mathbf{q}_u^\dagger \\ \text{minimal,} & \text{für } \mathbf{q}^\dagger = \mathbf{q}_s^\dagger \end{cases}. \quad (4.33)$$

Um die adjungierten Moden  $\mathbf{q}_u^\dagger$  und  $\mathbf{q}_s^\dagger$  zu identifizieren, wird daher numerisch die Information (4.31) unter Berücksichtigung der Nebenbedingung (4.14) extremalisiert. Letztere ist wichtig, da eine breitere Verteilung zu einer größeren Information führt. Dazu wird ein modifiziertes Gradientenverfahren verwendet, das im Gegensatz zu der im letzten Abschnitt verwendeten Methode auch problemlos auf mehr als zwei Dimensionen erweitert werden kann.

#### 4.4.2. Numerische Umsetzung

Die Zeitreihe  $\mathbf{q}(t)$  wird wie in Abschnitt 4.2.1 beschrieben erzeugt. Der Algorithmus zur Extremalisierung der Information (4.31) bezüglich  $\mathbf{q}^\dagger$  unter der Nebenbedingung (4.14) lässt sich in folgende Teilschritte gliedern:

1. Setze  $\mathbf{q}^\dagger$  auf einen normierten Anfangswert.
2. Berechne die Zeitreihe des Prozesses  $X_{\mathbf{q}^\dagger}(t) = \mathbf{q}^\dagger \cdot \mathbf{q}(t)$ .
3. Schätze die pdf  $f_{X_{\mathbf{q}^\dagger}}(x)$  aus der Zeitreihe  $X_{\mathbf{q}^\dagger}(t)$ .
4. Berechne die Information  $i(\mathbf{q}^\dagger)$  aus der pdf  $f_{X_{\mathbf{q}^\dagger}}(x)$ .
5. Berechne den Gradienten  $\nabla_{\mathbf{q}^\dagger} i(\mathbf{q}^\dagger)$ .
6. Berechne den Hilfsvektor  $\mathbf{g} = \nabla_{\mathbf{q}^\dagger} i(\mathbf{q}^\dagger) - \left( \nabla_{\mathbf{q}^\dagger} i(\mathbf{q}^\dagger) \cdot \mathbf{q}^\dagger \right) \mathbf{q}^\dagger$ .
7. Setze neues  $\mathbf{q}^\dagger$  auf  $\mathbf{q}^\dagger + \eta \mathbf{g}$ .
8. Normiere neues  $\mathbf{q}^\dagger$ .
9. Wiederhole die Schritte 2 – 8 bis das Extremum erreicht ist.

Im Folgenden werden diese Schritte im Detail erläutert.

*Schritt 1*

Der Anfangswert von  $\mathbf{q}^\dagger$  wird willkürlich auf  $q_1^\dagger = 0.5$ ,  $q_2^\dagger = \sqrt{1 - 0.5^2}$  festgesetzt.

*Schritt 2*

Da Schätzungen von univariaten Wahrscheinlichkeitsdichten in der Regel schon bei weniger Datenpunkten konvergieren als Drift- und Diffusionsschätzungen, wird an dieser Stelle zur Minderung des Rechenaufwandes von der ursprünglichen Zeitreihe mit  $10^8$  Datenpunkten nur jeder 1000. Datenpunkt verwendet. Damit besteht die verwendete Zeitreihe  $\mathbf{q}(t)$  aus  $10^5$  Datenpunkten. Aus diesen Datenpunkten wird die Zeitreihe  $X_{\mathbf{q}^\dagger}(t) = \mathbf{q}^\dagger \cdot \mathbf{q}(t)$  berechnet.

*Schritt 3*

Im dritten Schritt werden zunächst das Minimum  $X_{\min}$  und Maximum  $X_{\max}$  der Zeitreihe  $X_{\mathbf{q}^\dagger}(t)$  bestimmt. Anschließend wird  $X_{\min}$  um  $0.1(X_{\max} - X_{\min})$  nach unten verschoben und  $X_{\max}$  um denselben Betrag nach oben, da aufgrund der endlichen Bandbreite der Kernfunktion auch außerhalb des Datenbereichs die geschätzte pdf noch von null verschiedene Werte annehmen kann. Über das Intervall  $[X_{\min}, X_{\max}]$  werden  $n_s = 100$  Stützstellen  $x_i$  im Abstand

$$\Delta x = \frac{X_{\max} - X_{\min}}{n_s - 1} \quad (4.34)$$

eingerrichtet. Bevor die pdf geschätzt werden kann, muss eine Bandbreite  $h$  gewählt werden. Dazu wird zunächst der Mittelwert  $\bar{X}$  und anschließend die Standardabweichung  $\sigma$  nach Gl. (1.28) bestimmt. Anschließend wird die Daumenregel (1.34) zur Festlegung der Bandbreite verwendet. An jeder Stelle  $x_i$  wird dann die pdf  $f_{X_{\mathbf{q}^\dagger}}(x_i)$  mittels des Schätzers (1.31) und des Epanechnikov-Kerns (1.33) geschätzt.

*Schritt 4*

Im vierten Schritt wird die Information über

$$i(\mathbf{q}^\dagger) = - \sum_{i=1}^{n_s} f_{X_{\mathbf{q}^\dagger}}(x_i) \log f_{X_{\mathbf{q}^\dagger}}(x_i) \Delta x \quad (4.35)$$

berechnet. Es wird der natürliche Logarithmus verwendet.

*Schritt 5*

Die Ableitung der Information bezüglich einer Komponente von  $\mathbf{q}^\dagger$  ist formal gegeben durch

$$\frac{\partial i}{\partial q_j^\dagger} = - \int \left[ \log f_{X_{\mathbf{q}^\dagger}} \frac{\partial}{\partial q_j^\dagger} f_{X_{\mathbf{q}^\dagger}} + \frac{\partial}{\partial q_j^\dagger} f_{X_{\mathbf{q}^\dagger}} \right] dx. \quad (4.36)$$

Dabei kann die Ableitung  $\partial f_{X_{q^\dagger}} / \partial q_j^\dagger$ , wie in Anhang A dargestellt wird, direkt aus den Daten geschätzt werden, ohne sie durch einen endlichen Differenzenquotienten anzunähern. Dies ist jedoch sowohl programmiertechnisch als auch numerisch äußerst aufwendig. Stattdessen wird direkt die Ableitung der Information durch einen Differenzenquotienten angenähert:

$$\frac{\partial i(q^\dagger)}{\partial q_j^\dagger} \approx \frac{i(q^\dagger + \varepsilon \hat{e}_j) - i(q^\dagger - \varepsilon \hat{e}_j)}{2\varepsilon}. \quad (4.37)$$

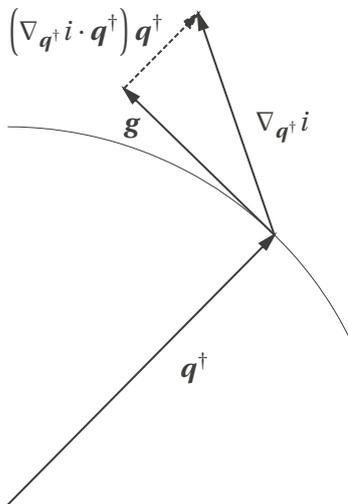
Dabei bezeichnet  $\hat{e}_j$  den  $j$ ten kanonischen Basisvektor des Phasenraums. Der Parameter  $\varepsilon$  wird auf 0.01 festgelegt.

#### Schritt 6

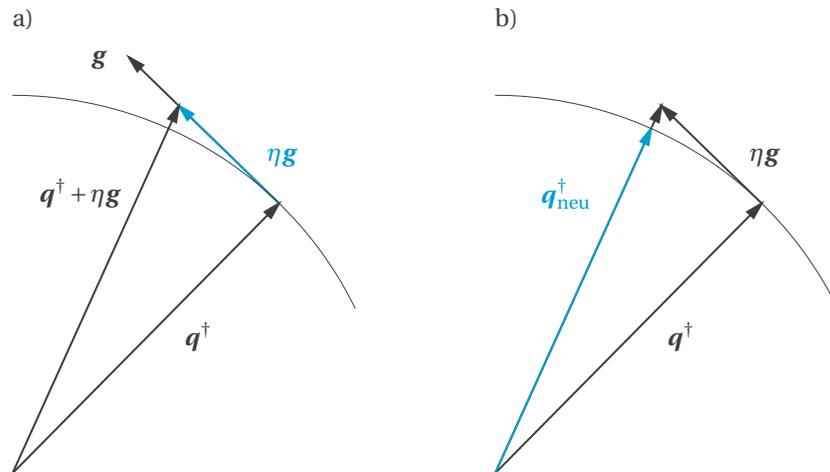
Die Nebenbedingung (4.14) beschreibt eine  $(n-1)$ -dimensionale Einheitshypersphäre in dem  $n$ -dimensionalen Raum, der durch die adjungierten Moden  $q^\dagger$  aufgespannt wird. Um bei den sukzessiven Iterationen des Gradientenverfahrens diese Hypersphäre nicht zu verlassen, wird von dem unter Schritt 5 berechneten Gradienten diejenige Komponente, die senkrecht auf der Hypersphäre steht, subtrahiert, vgl. Abb. 4.3. Das heißt, der Gradient wird durch den Hilfsvektor

$$\mathbf{g} = \nabla_{q^\dagger} i - (\nabla_{q^\dagger} i \cdot q^\dagger) q^\dagger \quad (4.38)$$

ersetzt.



**Abb. 4.3** Der Hilfsvektor  $\mathbf{g} = \nabla_{q^\dagger} i - (\nabla_{q^\dagger} i \cdot q^\dagger) q^\dagger$ , der in Schritt 6 berechnet wird.



**Abb. 4.4** Veranschaulichung der Schritte 7 und 8 des Algorithmus. Wird ein *Maximum* der Information gesucht, ist der Parameter  $\eta$  positiv wie in den Abbildungen. Wird ein *Minimum* gesucht, wählt man einen negativen Wert für  $\eta$ .

#### Schritte 7 und 8

In Schritt 7 wird der Vektor  $\eta \mathbf{g}$  zu  $\mathbf{q}^\dagger$  addiert, s. Abb. 4.4 a). Der Parameter  $\eta$  ist bei der Suche nach einem Maximum der Information positiv und bei der Suche nach einem Minimum negativ. Ein geeigneter Betrag von  $\eta$  wird durch “ausprobieren” bestimmt. Ist  $\eta$  zu groß, kann es passieren, dass der Algorithmus nicht konvergiert; ist  $\eta$  zu klein, konvergiert der Algorithmus sehr langsam.

Schließlich besteht der achte Schritt darin, den Vektor  $\mathbf{q}^\dagger + \eta \mathbf{g}$  zu normieren und als neues  $\mathbf{q}^\dagger$  zu definieren, s. Abb. 4.4 b).

#### Schritt 9

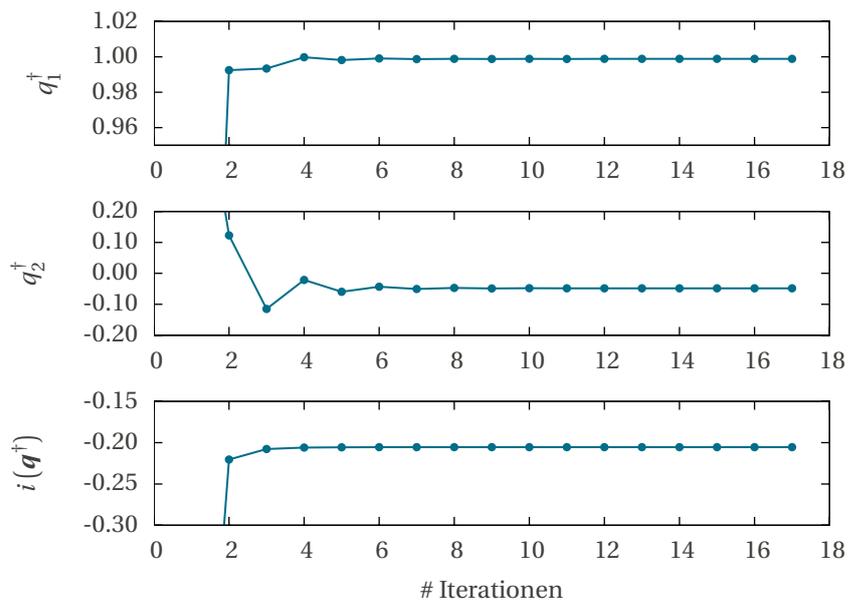
Die Schritte 2 – 8 werden in einer Schleife wiederholt, bis der Algorithmus konvergiert, was bedeutet, dass sich am Ende einer Iteration  $\mathbf{q}^\dagger_{\text{neu}}$  hinreichend gering von dem alten  $\mathbf{q}^\dagger$  unterscheidet. Dies wird durch die Abbruchbedingung

$$\left| \mathbf{q}^\dagger_{\text{neu}} - \mathbf{q}^\dagger \right| < \delta \quad (4.39)$$

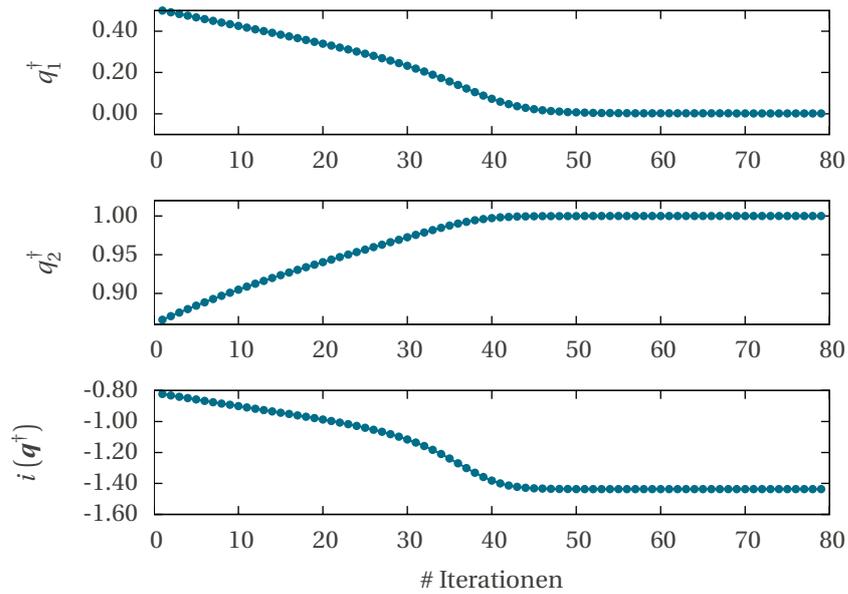
realisiert. Es wird  $\delta = 10^{-6}$  gewählt.

#### 4.4.3. Ergebnis

Abbn. 4.5 und 4.6 stellen dar, wie sich  $\mathbf{q}^\dagger$  und  $i(\mathbf{q}^\dagger)$  mit den Iterationen der Maximierung respektive Minimierung entwickeln. Dass der Algorithmus bei der Maximierung schneller



**Abb. 4.5** Maximierung der Information. Die Diagramme zeigen die Entwicklung der Komponenten der adjungierten Mode  $q_1^+$  (oben) und  $q_2^+$  (mitte) sowie der Information (unten) mit den Iterationen des Maximierungsalgorithmus.

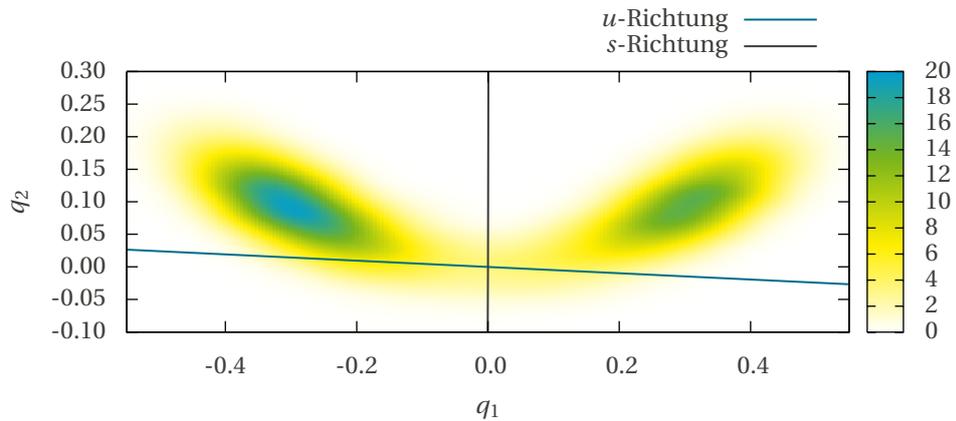


**Abb. 4.6** Minimierung der Information. Die Diagramme zeigen die Entwicklung der Komponenten der adjungierten Mode  $q_1^+$  (oben) und  $q_2^+$  (mitte) sowie der Information (unten) mit den Iterationen des Minimierungsalgorithmus.

konvergiert als bei der Minimierung, ist durch die Breite der Extrema und der jeweiligen Wahl des Parameters  $\eta$  bedingt. Als Endergebnis ergibt sich:

$$\mathbf{q}_u^\dagger = \begin{pmatrix} 0.998837 \\ -0.04806 \end{pmatrix}, \quad \mathbf{q}_s^\dagger = \begin{pmatrix} 0.0021017 \\ 0.9999978 \end{pmatrix}. \quad (4.40)$$

Dies wird analog zur Abb. 4.2 in Abb. 4.7 graphisch veranschaulicht. Dazu werden die mit den adjungierten Moden verbundenen Projektionsrichtungen durch Ursprungsgeraden der Steigung (4.24) dargestellt. Zur Orientierung wird zusätzlich die geschätzte jpdf  $f_2(q_1, q_2)$  farbkodiert geplottet, da der Informationsansatz auf Wahrscheinlichkeitsdichten basiert.



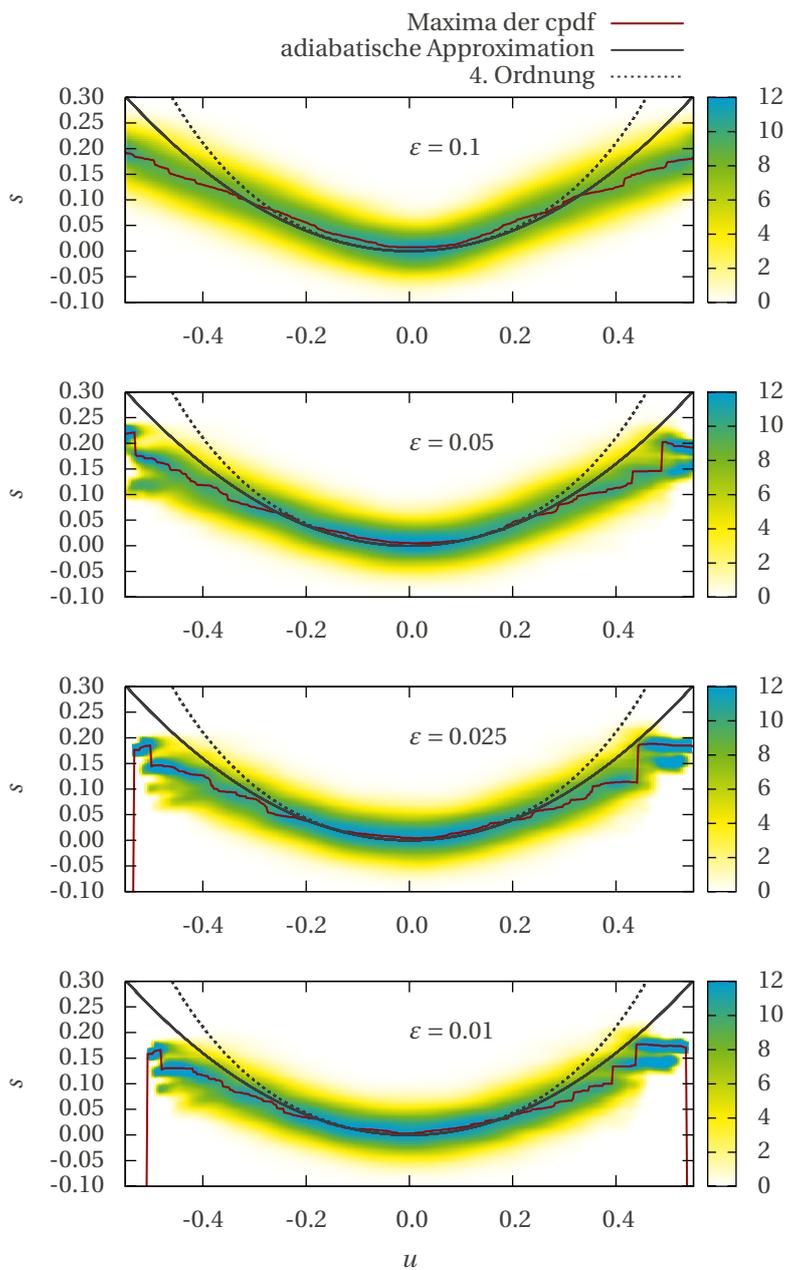
**Abb. 4.7** Resultierende Projektionsrichtungen beim Informationsansatz. Die Darstellung der Geraden gleicht der von Abb. 4.2. Anstelle des Vektorfeldes ist hier farbkodiert die jpdf  $f_2(q_1, q_2)$  zur Orientierung dargestellt (vgl. Abschnitt 4.5). Die Achsen sind derart skaliert, dass gleiche Intervalle auf beiden Achsen jeweils die gleiche Länge haben.

## 4.5. Invariante Mannigfaltigkeit

Abschnitt 2.3.3 behandelt die Anwendung des Versklavungsprinzips auf die FPG. Für das Haken-Zwanzig-System ergibt sich dabei die cpdf (Gl. (2.61)):

$$p(s|u) = \sqrt{\frac{\gamma}{2\pi D_s^{(2)}}} \exp \left\{ -\frac{\gamma}{2D_s^{(2)}} (s - s(u))^2 \right\}. \quad (4.41)$$

Dabei ist  $s(u)$  durch Gl. (4.10) gegeben. Für festes  $u$  liegt demnach das Maximum von  $p(s|u)$  bzgl.  $s$  auf der Mannigfaltigkeit  $s(u)$ . Sind die adjungierten Moden  $\mathbf{q}_u^\dagger$  und  $\mathbf{q}_s^\dagger$  bekannt, kann demnach die invariante Mannigfaltigkeit  $s(u)$  identifiziert werden, indem  $p(s|u)$  aus den Daten geschätzt, und anschließend für jedes  $u$  das Maximum bzgl.  $s$  bestimmt wird.



**Abb. 4.8** Rekonstruktion der invarianten Mannigfaltigkeit für verschiedene Parameter  $\varepsilon$ . Farbkodiert ist jeweils die geschätzte cpdf  $p(s|u)$  dargestellt. Die rote Kurve geht jeweils durch die Maxima der cpdf bzgl.  $s$  für alle festen  $u$  und stellt somit die rekonstruierte Mannigfaltigkeit dar. Die schwarze durchgezogene Kurve beschreibt die invariante Mannigfaltigkeit, die nach der adiabatischen Approximation durch Gl. (4.10) gegeben ist. Die schwarze gepunktete Kurve beschreibt die Approximation von  $s(u)$  in 4. Ordnung in  $u$ .

Ersteres geschieht, indem zunächst die jpdf  $f_2(u, s)$  mittels (1.35) und (1.33) geschätzt wird. Das Ergebnis ist farbkodiert in Abb. 4.7 zu sehen. Daraus wird  $p(s|u)$  gemäß Gl. (1.25) über

$$p(s|u) = \frac{f_2(u, s)}{\int f_2(u, s) ds} \quad (4.42)$$

berechnet.

Abb. 4.8 zeigt das Ergebnis zusammen mit der sich daraus ergebenden invarianten Mannigfaltigkeit für verschiedene Werte des Parameters  $\varepsilon$ . Zusätzlich ist die nach der adiabatischen Approximation zu erwartende Mannigfaltigkeit (4.10) geplottet. Für den Fall  $\varepsilon = 0.1$ , der bisher betrachtet wurde, gibt es auch in der direkten Umgebung des Sattelpunktes bei  $s = u = 0$  kleine aber wahrnehmbare Abweichungen zwischen der rekonstruierten Mannigfaltigkeit und (4.10). Diese verschwinden für kleinere  $\varepsilon$ , also wenn sich das System noch näher am kritischen Punkt  $\varepsilon_c = 0$  befindet. Für größere  $|u|$  weicht die rekonstruierte Mannigfaltigkeit in allen Fällen stärker von (4.10) ab. Diese Abweichung wird größer, wenn anstelle der adiabatischen Approximation eine Näherung 4. Ordnung in  $u$  der Mannigfaltigkeit  $s(u)$  betrachtet wird. Diese ergibt sich nach [24] zu

$$s(u) = \frac{1}{\gamma} u^2 + \frac{2}{\gamma} u^4 + \mathcal{O}(u^6) \quad (4.43)$$

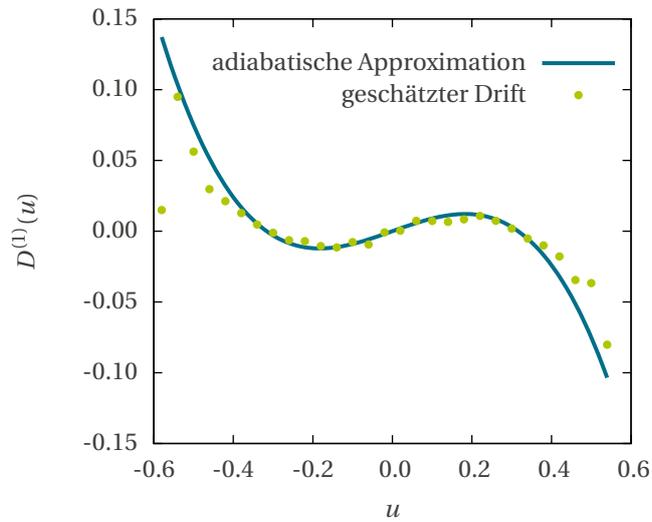
und ist ebenfalls in Abb. 4.8 dargestellt. Gl. (4.41) scheint daher nur in der Nähe des Sattelpunktes eine gute Approximation zu sein. Dies lässt sich verstehen, wenn man sich daran erinnert, welche Näherungen bei der Ableitung von Gl. (4.41) gemacht wurden. In Abschnitt 2.3.3 wurden bei der adiabatischen Elimination von  $s$  aus der FPG Terme mit Ableitungen von  $p(s|u)$  nach  $u$  gegenüber solchen mit Ableitungen nach  $s$  in Gl. (2.50) vernachlässigt. Wie man in Abb. 4.8 erkennen kann, ist diese Näherung in der Umgebung des Sattelpunktes auch legitim. Je weiter man sich jedoch vom Sattelpunkt wegbewegt, desto mehr ändert sich  $p(s|u)$  in  $u$ -Richtung und die Näherung wird schlechter. Auch das Resultat (2.83) aus dem MEP gilt nur für kleine  $|u|$  und  $|s|$ , da es unter Vernachlässigung hoher Potenzen der Komponenten von  $u$  und  $s$  abgeleitet wurde.

## 4.6. Schätzen der Ordnungsparametergleichung

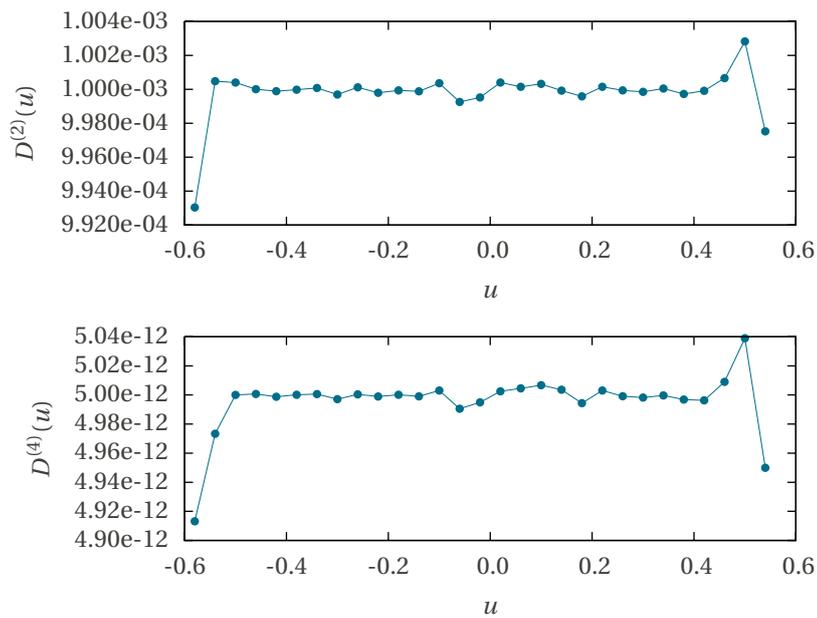
Schlussendlich wird das in Abschnitt 1.8 beschriebene Verfahren zur Schätzung von Drift und Diffusion des Ordnungsparameters  $u(t) = \mathbf{q}_u^\dagger \cdot \mathbf{q}(t) = q_1(t)$  angewendet. Das Ergebnis der Driftschätzung ist in Abb. 4.9 zusammen mit der Driftfunktion (4.12), die sich aus der adiabatischen Elimination der versklavten Amplitude  $s = q_2$  aus der FPG ergibt, dargestellt. Beides passt offensichtlich gut zusammen.

Abb. 4.10 zeigt den zweiten und vierten Kramers-Moyal-Koeffizienten. Der zweite Koeffizient, die Diffusion, ist bis auf statistische Fluktuationen konstant und ergibt sich zu

$$D^{(2)}(u) \approx 10^{-3} = D_1, \quad (4.44)$$



**Abb. 4.9** Drift der Ordnungsparametergleichung. Die grünen Punkte entsprechen dem Ergebnis der Driftschätzung. Die blaue Kurve entspricht dem nach Gl. (4.11) zu erwartenden Drift.



**Abb. 4.10** Der zweite (oben) und vierte (unten) Kramers-Moyal-Koeffizient des Ordnungsparameters. Der zweite Koeffizient ist praktisch konstant und stimmt mit der Diffusion aus Gl. (4.11) überein. Der vierte Koeffizient ist von der Größenordnung  $10^{-12}$ , also praktisch null.

was genau wie die Driftschätzung mit Gl. (4.11) übereinstimmt. Der vierte Koeffizient ist von der Größenordnung  $10^{-12}$ , also praktisch null. Damit garantiert das Pawula-Theorem (s. [45]), dass auch alle weiteren Koeffizienten  $D^{(n)}(u)$  mit  $n \geq 3$  verschwinden.

#### 4.7. Zusammenfassung und Ausblick

Mit dem Driftansatz und dem Informationsansatz wurden auf der Grundlage der Theorie der Synergetik zwei Methoden entwickelt, die das Potential besitzen, Ordnungsparameter und versklavte Modenamplituden aus multivariaten Datensätzen von Systemen in der Nähe von Instabilitäten zu identifizieren. Beide Methoden wurden erfolgreich an einem einfachen zweikomponentigen Modellsystem getestet.

Bisher wurde noch nicht untersucht, inwieweit die Methoden funktionieren, wenn wichtige Systemparameter der Gln. (4.8) wie  $\varepsilon$  und  $\gamma$  oder die Diffusionskoeffizienten  $D_1$  und  $D_2$  variiert werden. Während der Driftansatz vom Prinzip her unabhängig von der Art des Rauschens ist, bleibt beim Informationsansatz offen, ob er beispielsweise auch bei multiplikativem Rauschen oder im Falle von asymmetrischen Diffusionskoeffizienten, d. h.  $D_1 \neq D_2$ , funktioniert.

Mit den Argumenten aus Abschnitt (4.4.1) ist zu erwarten, dass eine Hauptkomponentenanalyse (engl. *Principal Component Analysis*, PCA) das Problem ebenfalls lösen würde. Eine PCA liefert jedoch ausschließlich *orthogonale* Moden, die den Eigenvektoren der symmetrischen Kovarianzmatrix entsprechen. Sie versagt also spätestens dann, wenn die gesuchten Moden nicht mehr orthogonal sind. Bei dem hier verwendeten Modellsystem (4.8) sind die gesuchten adjungierten Moden  $\mathbf{q}_u^\dagger$  und  $\mathbf{q}_s^\dagger$  orthogonal zueinander. Der nächste wichtige Schritt, der getan werden muss, besteht deshalb darin, zu testen, ob die beiden Methoden auch dann funktionieren, wenn die gesuchten Moden nicht mehr orthogonal sind.

Sollten die Methoden diesen Test bestehen, müsste als nächstes untersucht werden, inwieweit sie auch auf höherdimensionale Probleme anwendbar sind. Bei der Driftmethode ist die Durchführbarkeit in mehreren Dimensionen stark durch die Anzahl der verfügbaren Datenpunkte limitiert, da die Anzahl der Bins, und damit die Anzahl der Datenpunkte, die für die Schätzung des Driftvektorfeldes benötigt werden, exponentiell mit der Dimension des Datensatzes wachsen. Beim Informationsansatz hingegen müssen unabhängig von der Dimension immer nur eindimensionale pdfs der projizierten Zeitreihe  $X_{q^i}(t)$  geschätzt werden. Daher ist die Anzahl der verfügbaren Datenpunkte weniger entscheidend. Lediglich die Extremalisierung wird in mehreren Dimensionen aufwendiger. Man steht dann vor dem Problem, mehrere Minima und Maxima finden zu müssen. Möglicherweise sind Monte-Carlo-Methoden geeignet, dieses Problem zu lösen. Zunächst muss jedoch untersucht werden, ob die Hypothese, dass Minima und Maxima der Information Projektionsrichtungen von versklavten Amplituden und Ordnungsparametern entsprechen, sich in mehr als zwei Dimensionen überhaupt noch bewahrt.

Es wurde außerdem versucht, die invariante Mannigfaltigkeit mit Hilfe der cpdf  $p(s|u)$  zu rekonstruieren. Dabei stellte sich heraus, dass in der Umgebung des Sattelpunktes, also für kleine  $|u|$  und  $|s|$ , die Maxima von  $p(s|u)$  bzgl.  $s$  auf der Mannigfaltigkeit liegen, was als Kriterium zur deren Identifikation genutzt werden kann. In größerer Entfernung zum Sattelpunkt ist dies jedoch nicht mehr der Fall. Bei der Erweiterung dieser Methode auf mehr als zwei Dimensionen ergibt sich das gleiche Problem wie bei der Driftmethode, nämlich, dass die Anzahl der Datenpunkte, die man braucht, um eine hinreichend gut aufgelöste cpdf zu schätzen, exponentiell mit der Dimension wächst. Was die Rekonstruktion invarianter Mannigfaltigkeiten betrifft, sind daher noch andere Ansätze vonnöten.

Schlussendlich wurde noch eine Langevin-Gleichung für den Ordnungsparameter des Modellsystems geschätzt. Das Ergebnis stimmt gut mit der theoretischen Vorhersage überein. Eine Überprüfung der Markov-Eigenschaft, wie sie im vorangegangenen Kapitel durchgeführt wurde, ist bisher nicht erfolgt. Es wäre an dieser Stelle wichtig, einmal quantitativ den Zusammenhang zwischen Versklavung, also Zeitskalentrennung, und der Markov-Eigenschaft des Ordnungsparameters zu untersuchen. Man würde erwarten, dass die Markov-Länge größer wird, wenn der Unterschied der charakteristischen Zeitskalen von Ordnungsparametern und stabilen Amplituden verringert wird.

Ein interessantes Problem, zu dessen Lösung die hier entwickelten Methoden möglicherweise beitragen könnten, liegt in der Vorhersage von Instabilitäten. In vielen Systemen aus verschiedenen Bereichen der Wissenschaft hat die qualitative Änderung des makroskopischen Systemzustandes, mit der eine Instabilität verbunden ist, dramatische Auswirkungen. Man denke dabei beispielsweise an epileptische Anfälle. Bei solchen Systemen könnte man versuchen, diejenigen Moden zu identifizieren, die, um bei dem Beispiel der Epilepsie zu bleiben, bei einem Anfall instabil werden und gezielt deren Dynamik beobachten. Es besteht dann die Hoffnung, dass die herannahende Instabilität sich beispielsweise in der Drift bemerkbar macht. Auch die Informationsentropie, die bei einer Instabilität stark anwächst, könnte sich als geeigneter Precursor herausstellen.



## A. Gradient der Information

Im Sinne einer übersichtlicheren Darstellung betrachten wir nur den Fall, dass der Zustandsvektor  $\mathbf{q}$  zwei Komponenten besitzt, was bei dem behandelten Beispielsystem auch der Fall ist. Die folgende Rechnung kann aber ohne größere Schwierigkeiten auch auf beliebige Dimensionen verallgemeinert werden.

Der stochastische Prozess  $X_{\mathbf{q}^\dagger}(t)$  ist im zweidimensionalen Fall definiert als

$$X_{\mathbf{q}^\dagger}(t) = q_1^\dagger q_1(t) + q_2^\dagger q_2(t). \quad (\text{A.1})$$

Dessen pdf  $f_{X_{\mathbf{q}^\dagger}}$  ist nach (1.10) definiert als

$$f_{X_{\mathbf{q}^\dagger}}(x) = \langle \delta(x - X_{\mathbf{q}^\dagger}) \rangle. \quad (\text{A.2})$$

Um Ableitungen nach den Komponenten von  $\mathbf{q}^\dagger$  berechnen zu können, schreiben wir dies als

$$\begin{aligned} f_{X_{\mathbf{q}^\dagger}}(x) &= \int \delta(x - q_1^\dagger q_1 - q_2^\dagger q_2) f_2(q_1, q_2) dq_1 dq_2 \\ &= \int \frac{1}{|q_1^\dagger|} \delta\left(\frac{x}{q_1^\dagger} - \frac{q_2^\dagger}{q_1^\dagger} q_2 - q_1\right) f_2(q_1, q_2) dq_1 dq_2 \\ &= \frac{1}{|q_1^\dagger|} \int f_2\left(\frac{x}{q_1^\dagger} - \frac{q_2^\dagger}{q_1^\dagger} q_2, q_2\right) dq_2. \end{aligned} \quad (\text{A.3})$$

Alternativ lässt sich anstelle der  $q_1$ -Integration auch die  $q_2$ -Integration ausführen, was auf

$$f_{X_{\mathbf{q}^\dagger}}(x) = \frac{1}{|q_2^\dagger|} \int f_2\left(q_1, \frac{x}{q_2^\dagger} - \frac{q_1^\dagger}{q_2^\dagger} q_1\right) dq_1 \quad (\text{A.4})$$

führt. Um nicht Gefahr zu laufen durch null zu teilen, sollte bei der numerischen Berechnung für  $|q_1^\dagger| > |q_2^\dagger|$  die erste Variante, Gl. (A.3), verwendet werden, andernfalls die zweite Variante, Gl. (A.4).

Leitet man (A.3) nach  $q_2^\dagger$  ab, so ergibt sich

$$\frac{\partial}{\partial q_2^\dagger} f_{X_{q_2^\dagger}}(x) = \frac{1}{|q_1^\dagger|} \int \left( -\frac{q_2}{q_1^\dagger} \right) \frac{\partial}{\partial \chi_1} f_2(\chi_1, q_2) dq_2 \quad \text{mit} \quad (\text{A.5})$$

$$\chi_1 = \frac{x}{q_1^\dagger} - \frac{q_2^\dagger}{q_1^\dagger} q_2. \quad (\text{A.6})$$

Die Ableitung der zweiten Version ergibt:

$$\frac{\partial}{\partial q_2^\dagger} f_{X_{q_2^\dagger}}(x) = \frac{1}{(q_2^\dagger)^2} \int \left[ \frac{1}{|q_2^\dagger|} \frac{\partial}{\partial \chi_2} f_2(q_1, \chi_2) (q_1 q_1^\dagger - x) \mp f_2(q_1, \chi_2) \right] dq_1 \quad \text{mit} \quad (\text{A.7})$$

$$\chi_2 = \frac{x}{q_2^\dagger} - \frac{q_1^\dagger}{q_2^\dagger} q_1. \quad (\text{A.8})$$

Dabei gilt das Minuszeichen für  $q_2 > 0$  und das Pluszeichen für  $q_2 < 0$ . Die Ableitungen nach  $q_1^\dagger$  ergeben sich analog.

Die Ableitungen der jpdfs können wie die jpdfs selbst direkt aus den Daten geschätzt werden. Der Schätzer (1.35) lautet beispielsweise für  $f_2(\chi_1, q_2)$

$$\hat{f}_2(\chi_1, q_2) = \frac{1}{N h_1 h_2} \sum_{i=1}^N K \left( \frac{\chi_1 - q_1^i}{h_1} \right) K \left( \frac{q_2 - q_2^i}{h_2} \right). \quad (\text{A.9})$$

Damit ergibt sich für die Ableitung

$$\frac{\partial}{\partial \chi_1} \hat{f}_2(\chi_1, q_2) = \frac{1}{N h_1 h_2} \sum_{i=1}^N K' \left( \frac{\chi_1 - q_1^i}{h_1} \right) \frac{1}{h_1} K \left( \frac{q_2 - q_2^i}{h_2} \right). \quad (\text{A.10})$$

Dabei bezeichnet  $K'$  die Ableitung der Kernfunktion nach ihrem Argument. Für den Epanechnikov-Kern (1.33) gilt:

$$K'(y) = \begin{cases} -\frac{3\sqrt{5}}{50} y & \text{für } y^2 < 5 \\ 0 & \text{für } y^2 > 5 \end{cases}. \quad (\text{A.11})$$

An den Stellen  $y = \pm\sqrt{5}$  ist der Epanechnikov-Kern nicht differenzierbar. Die Ableitung kann an diesen Stellen gleich null gesetzt werden.

# Abbildungsverzeichnis

1.1. Beispiel einer Realisation eines stochastischen Prozesses . . . . .	13
1.2. Drift- und Diffusionsschätzung bei Prozessen mit periodischem Treiber . .	20
2.1. Die Funktion $-f \log f$ . . . . .	39
2.2. Typische pdf eines Ordnungsparameters in der Nähe einer Instabilität . .	40
3.1. Illustration des betrachteten Modellsystems . . . . .	42
3.2. Der Koeffizient $\alpha$ . . . . .	45
3.3. Korrelationsfunktion $C(t, \tau) = \langle S(t)S(t + \tau) \rangle$ . . . . .	48
3.4. Vergleich der linken und rechten Handseite der CKG, $\tau = 10^{-4}$ . . . . .	50
3.5. Maxima der Gauß-Funktionen . . . . .	51
3.6. Vergleich der linken und rechten Handseite der CKG, $\tau = 0.25$ . . . . .	52
3.7. Drift und Diffusion für $S = 0$ in Abhängigkeit von der Zeit . . . . .	53
3.8. Finite-Time-Drift und -Diffusion . . . . .	55
3.9. Diffusion für konstantes $t = 1.875$ in Abhängigkeit von $S$ . . . . .	56
3.10. Die Drift für die vier Phasen der Periode des Treibers . . . . .	57
3.11. Abhängigkeit der Drift von der Rauschamplitude . . . . .	58
3.12. Die stationäre pdf . . . . .	61
3.13. Kohärenter und inkohärenter Teil der Einpunkt-Korrelationsfunktion $L(\tau)$	62
4.1. Das Driftvektorfeld für das System (4.8) . . . . .	70
4.2. Resultierende Projektionsrichtungen beim Driftansatz . . . . .	71
4.3. Der Hilfsvektor $\mathbf{g}$ . . . . .	75
4.4. Veranschaulichung der Schritte 7 und 8 des Algorithmus . . . . .	76
4.5. Maximierung der Information . . . . .	77
4.6. Minimierung der Information . . . . .	77
4.7. Resultierende Projektionsrichtungen beim Informationsansatz . . . . .	78
4.8. Rekonstruktion der invarianten Mannigfaltigkeit . . . . .	79
4.9. Drift der Ordnungsparametergleichung . . . . .	81
4.10. Der zweite und vierte Kramers-Moyal-Koeffizient . . . . .	81



# Literaturverzeichnis

- [1] Argyris, J. ; Faust, G. ; Haase, M. ; Friedrich, R. : *Die Erforschung des Chaos*. 2. Auflage. Springer-Verlag, 2009
- [2] Böttcher, F. ; Peinke, J. ; Kleinhans, D. ; Friedrich, R. ; Lind, P. G. ; Haase, M. : Reconstruction of Complex Dynamical Systems Affected by Strong Measurement Noise. In: *Phys. Rev. Lett.* 97 (2006), Nr. 9, S. 090603
- [3] Box, G. E. P. ; Muller, M. E.: A Note on the Generation of Random Normal Deviates. In: *Ann. Math. Statist.* 27 (1958), Nr. 2, S. 610–611
- [4] Camazine, S. ; Deneubourg, J.-L. ; Franks, N. R. ; Sneyd, J. ; Theraula, G. ; Bonabeau, E. : *Self-Organization in Biological Systems*. Princeton University Press, 2003
- [5] Casado, J. M. ; Gómez Ordóñez, J. ; Morillo, M. : Stochastic resonance of collective variables in finite sets of interacting identical subsystems. In: *Phys. Rev. E* 73 (2006), Nr. 011109
- [6] Cubero, D. : Finite-size fluctuations and stochastic resonance in globally coupled bistable systems. In: *Phys. Rev. E* 77 (2008), Nr. 021112
- [7] Desai, R. C. ; Zwanzig, R. : Statistical Mechanics of a Nonlinear Stochastic Model. In: *J. of Stat. Phys.* 19 (1978), Nr. 1
- [8] Einstein, A. : Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. In: *Ann. d. Phys.* 322 (1905), Nr. 8, S. 549 – 560
- [9] Friedrich, R. ; Bestehorn, M. ; Haken, H. : Pattern Formation in Convective Instabilities. In: *Int. J. Mod. Phys. B* 4 (1990), S. 365 – 400
- [10] Friedrich, R. ; Peinke, J. : Statistical properties of a turbulent cascade. In: *Physica D: Nonlinear Phenomena* 102 (1997), Nr. 1-2, S. 147 – 155
- [11] Friedrich, R. ; Peinke, J. ; Renner, C. : How to Quantify Deterministic and Random Influences on the Statistics of the Foreign Exchange Market. In: *Phys. Rev. Lett.* 84 (2000), Nr. 22, S. 5224–5227

- [12] Friedrich, R. ; Siebert, S. ; Peinke, J. ; Lück, S. ; Siefert, M. ; Lindemann, M. ; Raethjen, J. ; Deuschl, G. ; Pfister, G. : Extracting model equations from experimental data. In: *Phys. Lett. A* 271 (2000), S. 217–222
- [13] Friedrich, R. ; Haken, H. : Exact stationary probability distribution for a simple model of a nonequilibrium phase transition. In: *Zeitschrift für Physik B Condensed Matter* 71 (1988), Nr. 4, S. 515 – 517
- [14] Gammaitoni, L. ; Hänggi, P. ; Jung, P. ; Marchesoni, F. : Stochastic resonance. In: *Rev. Mod. Phys.* 70 (1998), Nr. 1, S. 223–287
- [15] Ghasemi, F. ; Peinke, J. ; Sahimi, M. ; Tabar, M. R. R.: Regeneration of stochastic processes: an inverse method. In: *Eur. Phys. J. B* 47 (2005), S. 411 – 415
- [16] Ghasemi, F. ; Sahimi, M. ; Peinke, J. ; Tabar, M. R. R.: Analysis of Non-stationary Data for Heart-rate Fluctuations in Terms of Drift and Diffusion Coefficients. In: *J. Biol. Phy.* 32 (2006), Nr. 2, S. 117 – 128
- [17] Gómez Ordóñez, J. ; Casado, J. M. ; Morillo, M. ; Honisch, C. ; Friedrich, R. : Statistical Mechanics of finite arrays of coupled bistable elements. In: *Europhys. Lett.* 88 (2009), Nr. 40006
- [18] Gradisek, J. ; Govekar, E. ; Grabec, I. : Qualitative and quantitative analysis of stochastic processes based on measured data, II: Applications to experimental data. In: *Journal of Sound and Vibration* 252 (2002), Nr. 3, S. 563 – 572
- [19] Guckenheimer, J. ; Holmes, P. ; John, F. (Hrsg.) ; Marsden, J. E. (Hrsg.) ; Sirovich, L. (Hrsg.): *Applied Mathematical Sciences*. Bd. 42: *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. 2. Auflage. Springer-Verlag, 1986
- [20] Haken, H. : *Brain Dynamics*. Springer, 2002
- [21] Haken, H. : Synergetics: Basic Concepts. In: *Encyclopedia of Complexity and Systems Science*. Springer New York, 2009
- [22] Haken, H. ; Wunderlin, A. : *Die Selbststrukturierung der Materie*. Vieweg, 1991
- [23] Haken, H. : *Synergetics*. Springer-Verlag, 1978 (Springer Series in Synergetics)
- [24] Haken, H. : *Advanced Synergetics*. Springer-Verlag, 1983 (Springer Series in Synergetics)
- [25] Haken, H. : *Information and Self-Organization*. 2nd Enlarged Edition. Springer-Verlag, 2000 (Springer Series in Synergetics)
- [26] Härdle, W. ; Müller, M. ; Sperlich, S. ; Werwatz, A. : *Nonparametric and Semiparametric Models*. Springer, 2004 (Springer Series in Statistics)

- [27] Jaynes, E. T.: Information Theory and Statistical Mechanics. In: *Phys. Rev.* 106 (1957), Nr. 4, S. 620–630
- [28] Kirchner, C. : *Nicht-stationäre Langevin-Gleichungen als Modell kardiologischer Zeitreihen*, Friedrich-Alexander-Universität Erlangen-Nürnberg, Diss., 2008
- [29] Kleinhans, D. ; Friedrich, R. ; Nawroth, A. ; Peinke, J. : An iterative procedure for the estimation of drift and diffusion coefficients of Langevin processes. In: *Physics Letters A* 346 (2005), Nr. 1-3, S. 42 – 46
- [30] Kleinhans, D. : *Analyse stochastischer Prozesse – Eine iterative Methode zur Bestimmung von Drift und Diffusion*, Westfälische Wilhelms-Universität Münster, Diplomarbeit, 2005
- [31] Kometani, K. ; Shimizu, H. : A Study of Self-Organizing Processes of Nonlinear Stochastic Variables. In: *J. of Stat. Phys.* 13 (1975), Nr. 6
- [32] Kriso, S. ; Peinke, J. ; Friedrich, R. ; Wagner, P. : Reconstruction of dynamical equations for traffic flow. In: *Physics Letters A* 299 (2002), Nr. 2-3, S. 287 – 291
- [33] Kullback, S. ; Leibler, R. A.: On Information and Sufficiency. In: *The Annals of Mathematical Statistics* 22 (1951), Nr. 1, S. 79–86
- [34] Lade, S. J.: Finite sampling interval effects in Kramers-Moyal analysis. In: *Phys. Lett. A* 373 (2009), Nr. 41, S. 3705 – 3709
- [35] Lade, S. J.: Geometric and projection effects in Kramers-Moyal analysis. In: *Phys. Rev. E* 80 (2009), Nr. 3, S. 031137
- [36] Lamouroux, D. ; Lehnertz, K. : Kernel-based regression of drift and diffusion coefficients of stochastic processes. In: *Phys. Lett. A* 373 (2009), S. 3507 – 3512
- [37] Matsumoto, M. ; Nishimura, T. : Mersenne Twister: A 623-Dimensionally Equidistributed Uniform Pseudo-Random Number Generator. In: *ACM Transactions on Modeling and Computer Simulation* 8 (1998), Nr. 1, S. 3–30
- [38] Morillo, M. ; Gómez-Ordoñez, J. ; Casado, J. M.: Stochastic resonance in a mean-field model of cooperative behavior. In: *Phys. Rev. E* 52 (1995), Nr. 1, S. 316–320
- [39] Morillo, M. ; Gómez Ordóñez, J. ; Casado, J. M.: System size stochastic resonance in driven finite arrays of coupled bistable elements. In: *The European Physical Journal B* 74 (2010), Nr. 2, S. 211 – 215
- [40] Morillo, M. ; Gómez Ordóñez, J. ; Casado, J. M.: Role of fluctuations in the response of coupled bistable units to weak time-periodic driving forces. In: *Phys. Rev. E* 78 (2008), Nr. 021109

- [41] Nawroth, A. P. ; Peinke, J. ; Kleinhans, D. ; Friedrich, R. : Improved estimation of Fokker-Planck equations through optimization. In: *Phys. Rev. E* 76 (2007), Nr. 5, S. 056102
- [42] Paul, W. ; Baschnagel, J. : *Stochastic Processes*. Springer-Verlag, 1999
- [43] Pikovsky, A. ; Zaikin, A. ; Casa, M. A. I.: System Size Resonance in Coupled Noisy Systems and in the Ising Model. In: *Phys. Rev. Lett.* 88 (2002), Nr. 5, S. 050601
- [44] Portugali, J. : Self-Organization and the City. In: Meyers, R. A. (Hrsg.): *Encyclopedia of Complexity and Systems Science*. Springer, 2009, S. 7953 – 7991
- [45] Risken, H. : *The Fokker-Planck Equation*. Springer-Verlag, 1984
- [46] Runge, F. F. ; Liesegang, R. E. ; Belousov, B. : *Selbstorganisation Chemischer Strukturen*. 2. Auflage. Harri Deutsch, 2002
- [47] Scott, D. W.: *Multivariate density estimation: theory, practice, and visualization*. John Wiley & Sons, 1992
- [48] Shannon, C. E.: A mathematical theory of communication. In: *Bell Syst. Tech. J.* 27 (1948), S. 379 – 423, 623 – 656
- [49] Siegert, S. ; Friedrich, R. ; Peinke, J. : Analysis of data sets of stochastic systems. In: *Phys. Lett. A* 243 (1998), S. 275 – 280
- [50] Silverman, B. W.: *Monographs on Statistics and Applied Probability*. Bd. 26: *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, 1986
- [51] Stauffer, D. : Opinion Dynamics and Sociophysics. In: Meyers, R. A. (Hrsg.): *Encyclopedia of Complexity and Systems Science*. Springer, 2009, S. 6380 – 6388
- [52] Strogatz, S. H.: *Nonlinear Dynamics and Chaos*. Westview Press, 2000
- [53] Uhl, C. ; Friedrich, R. ; Haken, H. : Reconstruction of spatio-temporal signals of complex systems. In: *Z. Phys. B* 92 (1993), Nr. 2, S. 211 – 219
- [54] Uhl, C. ; Friedrich, R. ; Haken, H. : Analysis of spatiotemporal signals of complex systems. In: *Phys. Rev. E* 51 (1995), Nr. 5, S. 3890 – 3900
- [55] Wilmer, A. : *Zeitreihenanalyse stochastischer Systeme mit periodischem Treiber und Delay*, Westfälische Wilhelms-Universität Münster, Diplomarbeit, 2006

## Danksagung

Zunächst bedanke ich mich bei Herrn Prof. Dr. Rudolf Friedrich. Er hat mich überaus freundlich in seine Arbeitsgruppe aufgenommen und mich bei der Anfertigung dieser Arbeit stets gut betreut. Sein Interesse und seine Anerkennung wirkten auf mich sehr motivierend.

Desweiteren danke ich der gesamten Arbeitsgruppe für eine sehr angenehme Arbeitsatmosphäre, allgegenwärtige Hilfsbereitschaft und anregende Kaffeepausen. Besonders hervorzuheben sind dabei die Herren Oliver Kamps und Michael Wilczek. Vielen Dank für die konstruktive Kritik an dieser Arbeit, für die Beantwortung unzähliger Fragen und für viele fruchtbare Diskussionen, bei denen ich viel gelernt habe.

Ferner danke ich Dr. Olav Zimmermann und Dr. Jan Meinke vom Forschungszentrum Jülich. Ihre wertvollen Tipps zu allen Aspekten des wissenschaftlichen Arbeitens möchte ich nicht missen.

Ich möchte außerdem allen Menschen danken, die mich während meines Studiums begleitet haben und dazu beitrugen, dass ich mich an diese Zeit sicherlich immer gerne zurückerinnern werde.

Der allergrößte Dank gilt meinen Eltern. Dafür, dass sie mir ein finanziell sorgenfreies Studium ermöglichten, mich praktisch und emotional unterstützten, wo sie nur konnten, und mir darüber hinaus große Wertschätzung entgegenbrachten.

### Erklärung zur Diplomarbeit

Hiermit versichere ich, diese Arbeit selbständig angefertigt, und keine Hilfsmittel außer den angegebenen verwendet zu haben.

Münster, 23. August 2010