

**Güte von Projektionsverfahren in der Dynamik
relevanter kollektiver Koordinaten in ausgewählten
physikalischen Modellen**

Inaugural-Dissertation

zur

**Erlangung des Doktorgrades der
Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf**

vorgelegt von

Sylvie Defrasne

aus Calais

Düsseldorf

2004

Gedruckt mit der Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Referent: Prof. Dr. K.-H. Spatschek

Koreferent: Prof. Dr. A. Pukhov

Tag der mündlichen Prüfung: 22. Januar 2004

Inhaltsverzeichnis

1. Einleitung	1
1.1. Gleichgewichtsnahe Dynamik	2
1.2. Gleichgewichtsferne(re) Dynamik	5
1.3. “Optimal Prediction” Methode 0. Ordnung	9
1.4. Ziel der Arbeit	16
I. Hamiltonsche Systeme	21
2. Schrödingersche Systeme	21
2.1. Gaußsches Maß	30
2.2. Nicht-Gaußsches Maß	32
2.3. Anwendung der “Optimal Prediction” auf den Kollapsfall	34
2.3.1. Bestimmung von m_0 und b	35
2.3.2. Zeitentwicklung der V_α	42
2.4. Einfluß der Anfangsverteilung: Untersuchung der kubisch nichtlinearen Schrödingerschen Gleichung	49
3. Diskretes Modell	59
3.1. 0. Ordnung “Optimal Prediction”	63
3.2. Höhere Zeitentwicklungen der “Optimal Prediction”	66
3.3. “Linearized memory” Approximation	71

II. Nicht-Hamiltonsche dissipative Systeme	77
4. Demonstration der Güte des “finite-time memory”-Terms an einem einfachen mathematischen Modell	79
4.1. Reduktion über die Zentrumsmannigfaltigkeitstheorie	80
4.2. Anwendung der “Optimal Prediction” Methode	81
5. L-H-Übergang	87
5.1. Modell	89
5.2. Modell zur x -Elimination	97
5.3. L-H-Übergang bei anschließender Stabilität des H-Zustandes	99
5.4. L-H-Übergang bei anschließender Instabilität des H-Zustandes	102
5.5. Einfluß und Wichtigkeit der Parameterwerte	106
6. Zusammenfassung	113
A. Bedingte Erwartungswerte einer reellen, normal verteilten Variable	116
B. Struktur der Korrekturterme der “Optimal Prediction” Methode bei höheren Nichtlinearitäten	119
C. Bestimmung von m_0 und b im Fall $p = 2$	120
Literatur	125

1. Einleitung

In dieser Arbeit werden Methoden untersucht, die zur Reduktion und optimalen Voraussage der Dynamik in komplexen Systemen benutzt werden können. Unsere Motivation ist die Tatsache, daß ein komplexes nichtlineares System in der Regel nie exakt beschrieben werden kann. Zunächst ist eine sinnvolle Modellbildung gefordert, aber auch vereinfachte Modelle sind meist noch nicht exakt lösbar. Dann sind angemessene mathematische Näherungen erforderlich. Oft werden als Modell für physikalische Vorgänge und unter gewissen Voraussetzungen Systeme partieller Differentialgleichungen angesetzt. Diese partiellen Differentialgleichungen sind dann die Bewegungsgleichungen der für das Modell ausgewählten Variablen. Unter "kollektiven Koordinaten" versteht man ausgesuchte Freiheitsgrade, auf die das unendlich-dimensionale System "reduziert" wird und von denen man hofft, daß sie für die Dynamik des Systems die relevanten sind und seine zeitliche Entwicklung richtig beschreiben. Je nach dynamischer Situation bieten sich verschiedene Möglichkeiten an, diese Koordinaten und das System der relevanten Differentialgleichungen zu bestimmen. Hier wird auf die detaillierte Darstellung von Plaschko und Brod verwiesen [42]. Entscheidend dabei ist einerseits die Festlegung der Anzahl der relevanten Koordinaten, d. h. die Dimension des daraus folgenden und zu lösenden Systems von Differentialgleichungen. Andererseits ist die Frage nach der eventuellen Berücksichtigung der nicht als relevant eingestuften Freiheitsgrade von großem Interesse: inwiefern kann deren Einfluß auf die kollektiven Koordinaten in Form von gekoppelten Termen in den Differentialgleichungen miteinbezogen werden? Ein vernünftiger Ansatz dabei ist die Mittelung dieses Einflusses über eine große Zahl möglicher Konfigurationen. Dies ermöglicht die Bildung eines in dem Sinne "universellen" Kopplungsterms zwischen un- und berücksichtigten Koordinaten, daß er nicht von spezifischen Anfangsbedingungen abhängt. Die Art der Mittelung - ob arithmetisch oder von einer gegebenen systemspezifischen Statistik bestimmt - hängt vom betrachteten Problem ab.

1.1. Gleichgewichtsnahe Dynamik

In unmittelbarer Nähe eines Gleichgewichtszustands gibt eine Linearisierung erste Aufschlüsse. Die Zeitentwicklung eines autonomen physikalischen Systems sei durch die folgende Differentialgleichung gegeben:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}), \quad \mathbf{x}, \mathbf{F} \in \mathbf{R}^n; \quad (1.1)$$

Betrachtet man die n Komponenten des Vektors \mathbf{x} , hat man ein System von n partiellen Differentialgleichungen. Fixpunkte, auch Gleichgewichtszustände genannt, sind durch die Bedingung $\mathbf{f}(\mathbf{x}) = 0$ gegeben. Deren Stabilität wird mit der Methode der Linearisierung untersucht [42]. Eine betrachtete kleine Störung des Fixpunktes führt dann auf das Eigenwertproblem

$$(\mathbf{J}_\xi - \sigma I)\mathbf{e} = 0,$$

wobei σ der Eigenwert und \mathbf{e} der Eigenvektor ist. \mathbf{J} ist die Jacobi-Matrix des Systems (1.1) und \mathbf{J}_ξ die Jacobi-Matrix an der Stelle $\xi = 0$. Hier wurde o. B. d. A. angenommen, $\xi = 0$ sei Fixpunkt. Die Lösung dieses Eigenwertproblems führt auf eine Klassifikation der verschiedenen Fixpunkte nach deren Eigenwerten. Das Verhalten und die Struktur von Lösungen in der Nähe von Fixpunkten, denen keine rein imaginären Eigenwerte entsprechen, können anhand des Theorems von Hartman-Grobman untersucht werden [42, 30]. Das sogenannte “Stabilitätstheorem” beweist die asymptotische Stabilität bei Senken und Knoten, deren Eigenwerte alle negative Realteile besitzen. Treten aber sogenannte zentrale Eigenwerte ($\Re\sigma = 0$) auf, verliert die Linearisierungsmethode jedoch an Aussagekraft. Dann kann für kleine (lokale Theorie) aber endliche Abweichungen vom Gleichgewicht die Theorie der Zentrumsmannigfaltigkeit herangezogen werden. Diese geht von einer Versklavung der stabilen Moden aus.

Im folgenden wird das Vorgehen exemplifiziert [42]. Nehmen wir an, daß \mathbf{f} ein C^r -Diffeomorphismus ist. Mit Hilfe der Linearisierungsmethode werden der stabile, bzw. instabile und zentrale Eigenraum bestimmt. Es existieren nun entsprechend stabile, bzw. instabile Mannigfaltigkeiten in Form von C^r -Diffeomorphismen und eine zentrale Mannigfaltigkeit in Form eines C^{r-1} -Diffeomorphismus, die durch den Fixpunkt gehen und tangential zu dem jeweiligen Ei-

genraum liegen [30]. Vorausgesetzt wird jetzt, der instabile Eigenraum, der den Eigenwerten mit positiven Realteilen entspricht, sei leer. Das Theorem der zentralen Mannigfaltigkeit besagt nun, daß das Problem (1.1) sich auf die folgende Form bringen läßt

$$\dot{u} = Au + F(u, v) \quad (1.2)$$

$$\dot{v} = Bv + G(u, v) \quad (1.3)$$

Dabei ist $(u, v) = (0, 0)$ Fixpunkt. Die Amplituden u_1, \dots, u_n und v_1, \dots, v_m entsprechen den n linear marginalen und m linear stabilen Moden $[(u_1, \dots, u_n) =: u, (v_1, \dots, v_m) := v]$. Demnach verschwinden die Realteile der Eigenwerte der Matrix A ; die Realteile der Eigenwerte von Matrix B sind negativ. Nichtlineare Beiträge werden durch $N(u, v), M(u, v) \in C^r$ auf den rechten Seiten von (1.2) und (1.3) repräsentiert. Es sei E^c der n -dimensionale (verallgemeinerte) sogenannte "zentrale" Eigenraum von A und E^s der m -dimensionale (verallgemeinerte) sogenannte "stabile" Eigenraum von B . F und G sind C^r -Diffeomorphismen. Die zentrale Mannigfaltigkeit M^c kann lokal durch die Funktion

$$M^c = \{(u, v) \in \mathbf{R}^n \mid v = h(u)\}, \text{ mit } h(0) = 0 \text{ und } J_u(h(u))|_{u=0} = 0 \quad (1.4)$$

beschrieben werden.

Setzt man (1.4) in (1.2) und (1.3) ein, so erhält man

$$\dot{u} = Au + F(u, h(u)) \quad (1.5)$$

und dann

$$J(h(u)) [Au + F(u, h(u))] = Bh(u) + G(u, h(u)) \quad (1.6)$$

Diese letzte Gleichung ist eine nichtlineare partielle Differentialgleichung zur Berechnung der Funktion $h(u)$, für deren Lösung eine Potenzreihe angesetzt werden kann. Das folgende Theorem drückt die Bedeutung der Reduzierung über die Zentrumsmannigfaltigkeitstheorie [11] aus:

Existiert eine Umgebung U^c von $(u, v) = (0, 0)$ auf M^c , so daß jede Trajektorie, die in U^c anfängt, U^c auch nie verläßt, dann existiert eine Umgebung U von $(u, v) =$

$(0, 0)$ in $\mathfrak{R}^n \times \mathfrak{R}^m$, so daß jede Trajektorie, die in U anfängt, gegen eine Trajektorie auf der Zentrumsmannigfaltigkeit konvergiert.

Zur Diskussion der Dynamik reicht daher eine Untersuchung von (1.5) aus, insbesondere um das asymptotische Verhalten von (1.2) und (1.3) herauszufinden. Die Funktion $F(u, h(u))$ muß bis zu einer “relevanten” Ordnung entwickelt werden. Normalformen ergeben sich z.B. in dritter Ordnung [42].

Verallgemeinerungen von (1.2) und (1.3) sind in zwei Richtungen möglich: Einmal kann das endliche System Parameter enthalten; zum anderen kann das System unendlich-dimensional werden. Im Falle von zusätzlichen Parametern erweitern wir zu

$$\dot{u} = A(\Lambda)u + F(u, v, \Lambda), \quad (1.7)$$

$$\dot{v} = B(\Lambda)v + G(u, v, \Lambda), \quad (1.8)$$

$$\dot{\Lambda} = 0, \quad (1.9)$$

wobei $\Lambda = (\lambda_1, \dots, \lambda_p)$ p Parameter repräsentiert. Die Zentrumsmannigfaltigkeit wird dann $(n + p)$ -dimensional. Man spricht anstelle von einem Fixpunkt von einem “Bifurkationspunkt”, der für den Bifurkationsparameter $\Lambda = \Lambda_c$ vom physikalischen System erreicht wird. In diesem Falle sichern bestimmte Theoreme die Konvergenz der Störungsrechnung in der Nähe des Einsatzpunktes der Instabilität [3]; somit ist für ein kleines Gebiet des Parameterraumes die Langzeitdynamik von der Zentrumsmannigfaltigkeitstheorie richtig wiedergegeben. Die Untersuchung des Verhaltens eines idealen zylindrisch eingeschlossenen Plasmas in der Nähe des Einsatzpunktes der resistiven Austausch-Instabilität kann mit diesem Verfahren unternommen werden [5].

Unendlich-dimensionale Systeme kann man beispielsweise nach einer Diskretisierung auf endlich-dimensionale zurückführen.

1.2. Gleichgewichtsferne(re) Dynamik

Entfernt man sich weiter vom Gleichgewicht, stehen bislang (abgesehen von numerischen Verfahren) keine systematischen Methoden zur Verfügung.

Bedenkt man die Tatsache, daß ein nicht-autonomes System der Dimension p auf ein autonomes System der Dimension $p + 1$ zurückgeführt werden kann, geht es also darum, ein System partieller Differentialgleichungen der Form (1.1) zu lösen. Das naheliegendste Verfahren zur Lösung eines solchen Systems beruht auf finiten Differenzen; allerdings erfordert es eine sehr lange Rechenzeit. Sogenannte “ Reduktionsverfahren” bieten gegenüber dieser direkten Integration des Systems gewisse Vorteile.

Prinzipiell werden schwerpunktmäßig

- kollektive Koordinaten

- endliche Basissysteme
 - Galerkin Approximation
 - Karhunen Loeve Moden
 - Wavelet Basis

eingesetzt. Dies geschieht meist *ad hoc*.

Solch eine Reduktion von Systemen partieller Differentialgleichungen auf niedrig-dimensionale Systeme gewöhnlicher Differentialgleichungen besteht meistens in einer Entwicklung der Lösung nach einem bestimmten ausgesuchten Funktionensatz (Moden) und in dem Abschneiden dieser Entwicklung nach endlich vielen Moden [24].

Die gesuchte Lösung wird als unendliche Summe über alle Moden ins Ausgangssystem eingesetzt und nach einer Projektion auf den Satz der adjungierten Moden bekommt man ein äquivalentes System unendlich vieler gewöhnlicher Differentialgleichungen für die zeitliche Entwick-

lung der Amplituden der Moden.

Die Galerkin-Approximation besteht dann darin, aus dieser unendlichen Menge die Galerkinbasis, d. h. endlich viele Moden auszuzeichnen und alle anderen gleich Null zu setzen. Meist bieten sich als unmittelbar zugängliche Moden die Eigenmoden des linearen Operators des Systems an. Die Beschreibung des Systems wird also auf diese linearen Moden und deren Wechselwirkung beschränkt; aus dieser Tatsache folgt, daß die Dynamik von Systemen, bei denen im nichtlinearen Bereich viele Moden relevant sind, von einem Galerkin-Verfahren wahrscheinlich nicht zufriedenstellend wiedergegeben wird [3].

Handelt es sich bei den ausgesuchten Moden um Fourier-Moden, hat man den Spezialfall einer abgeschnittenen Spektraldarstellung.

Der Schlüsselpunkt in der Methode ist also das Aussuchen der für die zeitliche Dynamik des Systems “relevanten” Moden. Dies entspricht der physikalischen Fragestellung, in welchen Moden sich die meiste Energie des Systems konzentriert. Je größer die Galerkinbasis, desto eher konvergiert die Lösung des so gewonnenen endlich-dimensionalen Systems von gewöhnlichen Differentialgleichungen gegen die tatsächliche physikalische Lösung, aber desto höher ist auch der Rechenaufwand. Deswegen eignen sich Systeme, die grundsätzlich durch viele Eigenmoden simuliert werden müssen (cf. Problem der hydrodynamischen Turbulenz, bei der der ganze Inertialbereich über viele Größenordnungen von Wellenlängen, d. h. über viele Fourier-Moden modelliert wird), weniger für dieses Verfahren.

Die Güte einer gegebenen Galerkin-Basis kann durch ein sogenanntes dynamisches “Siebe”-Verfahren bestimmt werden [3].

Die Karhunen-Loeve-Methode, auch “proper orthogonal decomposition” genannt, ist ein Verfahren, mit dem systematisch nach dem optimalen Satz von Ansatzmoden für eine Galerkin-Projektion gesucht wird [24, 46]. Wegen der großen nötigen Rechnerkapazitäten wurde die

Methode, wenn auch früher bekannt, erst seit relativ kurzer Zeit auf konkrete physikalische Probleme angewandt, darunter auf die Modellierung und Simulation des sogenannten “L-H-Übergangs” im Tokamak [6], worauf wir im letzten Teil dieser Arbeit zurückkommen werden. Für eine Darstellung des Verfahrens wird auf die Arbeit von Eickermann [24] verwiesen. Kurz kann man sagen, daß der wesentliche Schritt der Methode in der Bestimmung der Eigenwerte und Eigenvektoren einer reellen symmetrischen Matrix besteht. Mathematisch ist abgesichert, daß die Karhunen-Loeve-Moden die optimale Basis für eine Galerkin-Approximation bilden und zwar für den Parametersatz, für den sie aus der numerischen Simulation gewonnen wurden. In der Praxis ist die Methode also nur sinnvoll anwendbar, wenn man davon ausgehen kann, daß sich die gleichen Funktionen auch bei einem anderen Parametersatz als Entwicklungsfunktionen eignen. Allerdings gibt es hier wie auch im Falle des Galerkin-Verfahrens kein Kriterium, mit dem sich abschätzen läßt, wieviele Moden nötig sind, um die Dynamik des Systems richtig zu approximieren. Dies ist eine große Einschränkung dieser beiden Methoden.

Sogenannte “Wavelets” schließlich entsprechen dem Wunsch nach einer kombinierten Zeit- und Frequenzbeschreibung. Sie können ähnlich wie bei Galerkin- und Karhunen-Loeve-Verfahren als Basisfunktionen zur Diskretisierung von partiellen Differentialgleichungen benutzt und als eine Erweiterung der Fourier-Analyse angesehen werden [39]. Die Untersuchung eines physikalischen Problems mittels Wavelets bietet aber gegenüber einer Entwicklung nach Fourier-Moden den Vorteil, daß sie nicht-stationäre Prozesse besser beschreiben kann, deren physikalische Bedeutung (zeitlich begrenzte Übergänge, Anlaufphasen von Signalen, etc.) oft wichtig ist. Mit anderen Worten geht im Vergleich zu der Fourier-Analyse bei der Wavelets-Beschreibung die Information über die zeitliche Lokalisierung einer bestimmten Frequenzanregung nicht verloren. Zudem besitzen Fourier-Analysen folgende Nachteile: bezüglich Störungen sind sie sehr instabil, die Addition eines zusätzlichen Terms kleiner Amplitude zu einer linearen Superposition von Sinuswellen wird das gesamte Signal kaum verändern, dessen Fourierspektrum aber sehr [1]; bei Unstetigkeiten eines Signals konvergiert seine Fourier-Reihe schlecht und

den Fourier-Koeffizienten ist nicht anzusehen, wo sich die Sprungstelle befindet [9]. Deshalb wurde die Wavelets-Beschreibung als eine unter verschiedenen sogenannten “Zeit-Frequenz”-Darstellungen entwickelt, denen die “gefensterte” Fourier-Transformation auch gehört. Die Wavelet-Moden kann man also als Filter auffassen, die von einer sowohl im Raum (oder in der Zeit) wie auch in der Frequenz lokal begrenzten Funktion dargestellt werden. Diese Lokalisierung kann als “maßgeschneidert” in dem Sinne bezeichnet werden, daß kurzlebige Details der Lösung wie Sprungstellen oder ausgeprägte Spitzen genau lokalisierbar sind, während langzeitliche Eigenschaften naturgemäß in kleinerem Maßstab, sprich weniger genau lokalisiert beschrieben werden [9]. In der Praxis muß zuerst ein sogenanntes “Mutter”-Wavelet in Form einer Funktion $x \mapsto \psi(x)$ gewählt werden. Die zur Analyse von Signalen f verwendeten dilatierten und verschobenen Kopien des Mutter-Wavelets sind dann die Funktionen [9]

$$\Psi_{a,b} : \quad \mathbf{R} \rightarrow \mathbf{C}, \quad t \mapsto \frac{1}{|a|^{1/2}} \psi\left(\frac{t-b}{a}\right)$$

a und b sind dabei respektive der Skalen- und der Verschiebungsparameter. Die Wavelet-Transformierte $Wf(a,b)$ des Signals $f(x)$ ist dann durch

$$f(x) \mapsto Wf(a,b) = \int_{-\infty}^{\infty} \overline{\Psi_{a,b}(x)} f(x) dx$$

gegeben. Analog zu der Fourier-Transformation spricht man von kontinuierlichen und diskreten Wavelet-Transformationen; auch kann das untersuchte Signal wie bei der Fourier-Synthese als Linearkombination der Waveletbasisfunktionen dargestellt, bzw. zurückgewonnen werden. Allerdings sind dabei die Basisfunktionen nicht mehr andauernde Oszillationen, sondern physikalisch einfacher nachvollziehbare zeitlich lokalisierte Funktionen [39]. Als besonderer Vorteil der Wavelet-Methode muß die Tatsache erwähnt werden, daß sie im Gegensatz zur “gefensterten” Fourier-Transformation, die von konstanter Bandbreite ist, mit konstanter *relativer* Bandbreite arbeitet, deshalb bei großen Frequenzen, d. h. bei kleinen Skalen, sehr effizient ist, weshalb sie zur Untersuchung von Singularitäten in physikalischen Signalen besonders geeignet ist [1]. Interessant ist außerdem die gute numerische Stabilität der Methode, die aus dem lokalen Charakter folgt. Diese Methode eignet sich also zur Beschreibung von physikalischen

Problemen, die durch zeitlich lokalisierte Veränderungen charakterisiert sind. Als typische Anwendungsbeispiele bieten sich turbulente Systeme an, deren Eigenschaften auf einer kurzen Zeitskala variieren, bzw. dissipative Systeme, da die Wavelets den realen abklingenden Anregungen eher entsprechen [39]. Dabei bietet eine Wavelet-Analyse lokal eine lineare Zerlegung in relevante Moden, vorausgesetzt die Nichtlinearität des Systems ist nicht zu groß. Weitere Details und insbesondere mögliche Anwendungen auf die Untersuchung des L-H-Übergangs in toroidal eingeschlossenen Plasmen, auf den wir im letzten Teil dieser Arbeit zurückkommen werden, sind in [39] nachzulesen.

1.3. “Optimal Prediction” Methode 0. Ordnung

Mangelnde numerische Auflösung sorgt oft dafür, daß die Zeitentwicklung eines dynamischen Systems in geschlossener Form nicht möglich ist. Dies ist insbesondere der Fall, wenn in komplexen Systemen nur ein paar von vielen Freiheitsgraden berücksichtigt werden können. Dies kann verschiedene Gründe haben: entweder zwingt die eingeschränkte Kapazität der zur Simulation eingesetzten Computer, die vielen Freiheitsgrade auf ein paar, wie man hofft “ausschlaggebende” zu reduzieren; oder es sind nur die Anfangsbedingungen ein paar kollektiver Variablen bekannt, auch wenn im Prinzip alle Freiheitsgrade zugänglich wären [19].

Die “Optimal Prediction” Methode versucht, diesen Mangel an Auflösung in der numerischen Simulation der dynamischen Zeitentwicklung durch die Ausnutzung einer statistischen Information über das System auszugleichen.

Diese Methode wurde in verschiedenen Arbeiten von Chorin, Kupferman, Kast und Levy entwickelt [16, 18, 17].

Zuerst wird das Prinzip der Methode in Anlehnung an eine Arbeit von Chorin *et al.* erläutert, in der diese Methode auf Hamiltonsche partielle Differentialgleichungen angewandt wird [19].

Untersucht wird ein gegebenes Cauchy-Problem:

$$\begin{cases} u_t = R(x, u, u_x, u_{xx}, \dots) \\ u(x, 0) = u_0(x) \end{cases} \quad (1.10)$$

für eine Funktion $u(x)$ in einem p -dimensionalen Raum. Dabei ist $R(x, u, u_x, u_{xx}, \dots)$ eine im allgemeinen nichtlineare Funktion von $u = u(x, t)$ und seinen Raumableitungen.

Wir nehmen an, daß wir nicht in der Lage sind, das Problem in geschlossener Form zu lösen, entweder mangels ausreichender Rechnerkapazitäten, oder weil man sich auf bestimmte Eigenschaften der Lösung beschränken möchte und sie dafür nicht in allen Details kennen braucht.

Vorausgesetzt wird zusätzlich, daß eine bestimmte statistische Information über das System bekannt ist, beispielsweise in der Form eines explizit gegebenen Maßes μ (sogenanntes “*prior measure*” bei Chorin *et al.*). Es ist uns also möglich, einen Satz von vielen Anfangsdaten (typische Größenordnung: $10^2 - 10^3$) zu produzieren, die einer bestimmten Statistik gehorchen. Ihr Mittelwert zur Zeit $t = 0$ ist unmittelbar bekannt. Gesucht wird jetzt deren Mittelwert zu einer weiteren Zeit $t > 0$. Prinzipiell folgt er aus der Mittelung der einzeln integrierten Anfangswerte des ursprünglichen Ensembles. Dennoch ist diese Integration in der Praxis zu aufwendig und daher nicht durchführbar. Wie kann die statistische Verteilung, der der Satz der Anfangsdaten zur Zeit $t = 0$ aber weiterhin zu einer späteren Zeit $t > 0$ gehorcht, genutzt werden, um den Mittelwert $\langle u \rangle$ für jede Zeit $t > 0$ zu finden?

Die Anfangsdaten werden gemäß einer Wahrscheinlichkeitsverteilung bestimmt, die durch das invariante Maß μ gegeben ist. “Invariant” bedeutet dabei, daß wenn man einen bestimmten Satz von Anfangsdaten gemäß den Bewegungsgleichungen integriert, dann erfüllen für $t > 0$ die Daten die gleiche Verteilung. Im Falle eines Hamiltonschen Systems bietet sich das kanonische Maß an; die entsprechende Wahrscheinlichkeitsdichte ist gegeben durch:

$$f(u) = \frac{1}{Z} e^{-\frac{H(u)}{T}}$$

Dabei ist T die Temperatur und bestimmt die Varianz der Dichte. Im folgenden wird hier $T = 1$ gesetzt. Der Einfluß von verschiedenen Werten von T auf die Aussagekraft der “Optimal Prediction” Methode wurde bereits untersucht [19].

Vorausgesetzt sei zur Zeit $t = 0$ ein bestimmter Satz von sogenannten “kollektiven Koordinaten”

V_α bekannt, die gegeben sind durch:

$$V_\alpha = (g_\alpha, u) = \int g_\alpha(x)u(x, 0)dx \quad (1.11)$$

Dabei kann man die Funktionen g_α als “Filter” auffassen. Werden beispielsweise $g_\alpha(x) = \delta(x - x_\alpha)$ ausgewählt, dann sind die V_α die von der Funktion $u(x)$ an den Stellen x_α angenommenen Werte.

Entscheidend dabei ist der folgende Punkt: man setzt voraus, daß im Prinzip das System wohl in geschlossener Form gelöst werden könnte, d. h. daß alle Daten im Prinzip zugänglich sind. Dies bedeutet, daß man freie Wahl der kollektiven Koordinaten hat. In der Praxis besteht einer der Hauptschritte der Methode also darin, die für das zu untersuchende Problem am besten geeigneten g_α zu finden. Insofern wird der Anwendungsbereich der Methode aber auch eingeschränkt: wenn *a priori* nur bestimmte kollektive Koordinaten zugänglich sind, wie es zum Beispiel in geophysikalischen Anwendungen der Fall ist [19], dann muß das Risiko eingegangen werden, daß diese nicht unbedingt allein ausschlaggebend für die weitere dynamische Entwicklung des Systems sind; die “Optimal Prediction” Methode würde dann an Aussagekraft verlieren.

Die Auswahl der “richtigen” Filterfunktionen und damit der kollektiven Koordinaten ist also ein wichtiges Kriterium für den Erfolg der Methode. Wenn außerdem die Anzahl und die Auswahl der kollektiven Koordinaten, die nötig sind, um unsere Lösung ausreichend zu charakterisieren, in der Zeit veränderlich sind, dann hat das für den Erfolg der “Optimal Prediction” Methode erhebliche Konsequenzen, die am Beispiel des Kollapses im Falle der nichtlinearen Schrödinger-Gleichung weiter unten erörtert werden.

Wäre nur der Anfangssatz (1.11) von kollektiven Koordinaten bekannt, und würde man von der statistischen Verteilung keinen Gebrauch machen, dann wäre man auf übliche Regressionsmethoden angewiesen [41]: dabei wird über alle möglichen Lösungen gemittelt, die der invarianten Wahrscheinlichkeitsverteilung gehorchen, und gleichzeitig die Bedingung (1.11) zur Zeit $t = 0$ erfüllen. Diese Mittelung erfolgt für jeden x ; daraus folgt der Mittelwert

$$v(x) = E[u(x)|V_1, \dots, V_N],$$

d. h. der Erwartungswert der Funktion $u(x)$, seien die V_1, \dots, V_N zur Zeit $t = 0$ bekannt.

Das invariante Maß μ kann aber benutzt werden, wenn man bedenkt, daß es auf den Raum der Funktionen u , die die Anfangsbedingungen erfüllen, ein sogenanntes “bedingtes” Maß μ_V induziert, das aber nicht mehr invariant ist. Für $t > 0$ sind die Lösungen $u(x, t)$ von (1.10) durch das Maß $\mu_V(t)$ bedingt, dabei sind im allgemeinen $V_\alpha(t) \neq V_\alpha(0)$ und $g_\alpha(x, t) \neq g_\alpha(x, 0)$.

Die Grundidee der “Optimal Prediction” Methode kann in der folgenden Gleichung ausgedrückt werden:

$$\mu_V(t) = \mu_{V(t)},$$

d. h. das bedingte Maß $\mu_V(t)$ zur Zeit t wird approximiert durch das invariante Maß, bedingt zur Zeit t durch einen Satz von neuen kollektiven Koordinaten

$$V_\alpha(t) = (g_\alpha(t), u(t)) = \int g_\alpha(x, t) u(x, t) dx. \quad (1.12)$$

Dabei werden im allgemeinen Fall die $g_\alpha(x, t)$ von den Filtern zur Zeit $t = 0$ verschieden sein. Deren Anzahl und die affine Form der somit bestimmten kollektiven Koordinaten sollen aber gleich bleiben.

Wie funktioniert also die “Optimal Prediction” Methode? Die Zeitentwicklung des bedingten Maßes μ_V kann mit Hilfe der Bewegungsgleichung (1.10) für die Zeit $t \rightarrow (t + \Delta t)$ bestimmt werden. Daß nur ein infinitesimaler Zeitschritt möglich ist, rührt daher, daß für größere Schritte das bedingte Maß nicht mehr durch das invariante bedingt durch die anfänglichen kollektiven Koordinaten ausgedrückt werden kann. Dies kann man formal schreiben als:

$$t \rightarrow t + \Delta t, \quad \mu_V(t) \rightarrow \mu_V(t + \Delta t)$$

Gleichzeitig wird versucht, für die Zeit $t + \Delta t$ neue passende affine Bedingungen zu finden, die ausgehend vom invarianten μ das gleiche bedingte Maß liefern:

$$t \rightarrow t + \Delta t, \quad \mu \text{ (über neue Konditionen } g_\alpha \text{ und } V_\alpha) \rightarrow \mu_{V(t+\Delta t)}$$

Das Grundprinzip der “Optimal Prediction” Methode besteht darin, diese zwei auf verschiedenen Wegen bestimmten Maße gleichzustellen; somit erhält man Bewegungsgleichungen für die Bedingungen V_α . In der Praxis gelangt man dazu über eine Momentengleichung; Momente des bedingten Maßes zur Zeit $t + \Delta t$, die über Regression von den alten Momenten zur Zeit t gerechnet werden, werden den Momenten, die vom invarianten Maß bedingt durch neue kollektive Koordinaten herrühren, gleichgesetzt.

Wir beschränken uns im folgenden auf die Berechnung von Momenten 1. Ordnung; wie das Verfahren bei Momenten höherer Ordnung aussehen würde, wurde von Chorin *et al.* erläutert [19].

Wir gehen zuerst davon aus, wie in [19], daß N kollektive Koordinaten V_α durch die inneren Produkte (g_α, u) gegeben sind. Wir betrachten nur den 1. Moment, also den Mittelwert $\langle (g_\alpha, u) \rangle_V = (g_\alpha, \langle u \rangle_V)$. Gleichzeitig wird als einziger Parameter pro Bedingung der Wert V_α modifiziert.

Es muß gelten:

$$\frac{d}{dt}V_\alpha = \langle (g_\alpha, u_t) \rangle_V = \langle (g_\alpha, R(u)) \rangle_V = (g_\alpha, \langle R(u) \rangle_V)$$

Die Gleichung also, die der “Optimal Prediction” Methode zugrundeliegt, ist folgende:

$$\frac{d}{dt}V_\alpha = (g_\alpha, \langle R(u) \rangle_V), \quad \alpha = 1, \dots, N \tag{1.13}$$

Dieses kann hier auch wieder als der Erwartungswert des inneren Produktes von g_α und u bei gegebenen kollektiven Variablen V_α aufgefaßt werden:

$$\frac{d}{dt}V_\alpha = E[(g_\alpha, R(u)) | V_1, \dots, V_N], \quad \alpha = 1, \dots, N$$

Man ersetzt also die Funktion $u(t)$ durch ihren bedingten Mittelwert $\langle u \rangle_V^0$, seien die kollektiven Koordinaten $V_\alpha(0)$ gegeben.

Wenn die V_α zur Zeit t bekannt sind, dann kann man mittels einer Regression, nach dem oben schon erwähnten Verfahren, den Mittelwert für jeden x finden. Dabei ist die Wahl von untereinander orthogonalen Kernfunktionen von Vorteil, da sie die Anzahl der nichtverschwindenden

Terme in der Entwicklung reduziert [19].

Falls aber die zeitliche Entwicklung, die von der Bewegungsgleichung (1.10) herrührt, eine ergodische bezüglich des invarianten Maßes ist, geht asymptotisch das bedingte Maß in das invariante über und der Einfluß der Anfangsbedingungen und der kollektiven Koordinaten verschwindet [19]. Es ist also zu erwarten, daß die 0. Ordnung der “Optimal Prediction” nur für kurze Zeiten gute Ergebnisse liefert; dies steht im Einklang mit der Tatsache, daß bei der Berechnung der bedingten Erwartungswerte für diese kurzen Zeiten thermisches Gleichgewicht der unaufgelösten Skalen vorausgesetzt werden kann [12].

Untersucht wurde schon, wann das eben geschilderte Verfahren erfolgsversprechend ist [19]: dabei spielt die Auswahl der Kernfunktionen g_α eine große Rolle. Ist das Anfangswertproblem tatsächlich ein lineares, sprich die Bewegungsgleichung ist durch $u_t = L(u)$ gegeben, wobei L ein linearer Operator ist, dann sind die Gleichungen (1.13) exakt und keine Approximation mehr, sobald die g_α Eigenfunktionen des adjungierten Operators L^\dagger . Dies ist weiterhin der Fall, wenn der von den g_α aufgespannte Raum unter $u_t + L^\dagger = 0$ invariant ist. Die “Optimal Prediction” wird also dann gute Ergebnisse liefern, wenn die g_α einen Raum aufspannen, der fast invariant unter dem von L^\dagger induzierten Fluß ist [31].

Eine ähnliche Aussage kann man bei nichtlinearen Problemen treffen: sei der von den g_α aufgespannte Raum der “aufgelöste” Teil unserer Lösung, dann liefert die “Optimal Prediction” Methode 0. Ordnung eine gute Approximation für die zeitliche Entwicklung dieses Teils der Lösung, die den Einfluß der nicht-aufgelösten Freiheitsgrade berücksichtigt, solange kein wesentlicher Energie- und Informationsaustausch zwischen dem aufgelösten und dem unaufgelösten Teil der Lösung stattfindet [19, 40, 54]. Findet dieser doch statt, dann sind *a priori* zusätzliche sogenannte “memory”-Terme nötig [13, 14, 15].

Auf dem Gebiet der Reduktion von komplexen Systemen partieller Differentialgleichungen sind

schon ähnliche Ansätze verfolgt worden. Immer ging es dabei darum, den Einfluß der nicht-aufgelösten Freiheitsgrade auf die Dynamik der aufgelösten Freiheitsgrade zu berücksichtigen. Bei physikalischen Prozessen, deren dynamische Gleichungen Lösungen hervorbringen, die sich über weite Längen- und Zeitskalen erstrecken, und bei denen empirische Beobachtungen eine Skaleninvarianz erwiesen haben, sobald sich ein bestimmter Parameter einem kritischen Wert nähert, kann diese Fraktalität dazu ausgenutzt werden, über sogenannte “fraktale Interpolation” für kleine Skalen ein skaleninvariantes Feld zu konstruieren, dessen Einfluß auf die großen, aufgelösten Skalen dann analytisch untersucht wird [45]. Dieses Verfahren wurde an der “forced” Burgers-Gleichung untersucht; der Vorteil dieser Methode liegt darin, daß das kleinskalige invariante Feld nicht explizit bekannt sein muß, da die “fraktale Interpolation” es ermöglicht, seinen Einfluß auf die großen Skalen direkt zu untersuchen. Die kleinen Skalen sind dann den aufgelösten untergeordnet und “versklavt”, die so erzeugte Simulation berücksichtigt diese kleinskaligen Fluktuationen.

Eine stochastische Antwort auf das Problem wurde von Vaillant angeführt: damit wurden über eine stochastische Teilchenmethode statistische Lösungen einer McKean-Vlasov-Gleichung mit zufälliger Anfangsbedingung berechnet, wobei die Wichtungskoeffizienten aus “*nonparametric estimators*” einer Regressionsfunktion folgen. Die Konvergenz dieser Methode in Abhängigkeit der Anzahl der simulierten Teilchen und des Zeitdiskretisierungsschritts wurde untersucht [49]. Die “Optimal Prediction” Methode, wenn sie auf ein Hamiltonsches System angewandt wird, ist auch mit den sogenannten “Renormierungsmethoden” verwandt, die die Dimension eines Systems reduzieren, indem über ungewollte Skalen integriert wird [34, 10, 43]. Der vollständige Hamiltonian des Systems steht nämlich im gleichen Verhältnis zum Hamiltonian des reduzierten Systems, wie der Hamiltonian zu seiner renormierten Form in Renormierungsgruppentransformationen. Dieses macht man sich zunutze, indem man eine “kleinzellige” Monte-Carlo Methode der Renormierungsgruppe zur rekursiven Berechnung der bedingten Erwartungswerte einsetzt [12]. Dies wurde an dem Fall einer Anordnung von Ising-Spins geprüft, bei der die strenge Wechselwirkung ohne Skalenseparation “vorsichtige” Methoden der Reduzierung erfordert.

Das typische Anwendungsgebiet solcher Renormierungsmethoden ist die Simulation turbulenter Ströme; das Ausintegrieren der ungewollten Freiheitsgrade ist dabei wegen der nichtlinearen Skaleninterferenz nicht trivial [33]. Dennoch konnten nichtlokale Theorien entwickelt werden, bei denen keine Voraussetzung über die statistische Verteilung der kleinen Skalen *a priori* gemacht werden muß [33].

Schließlich weist die “Optimal Prediction” Methode gewisse Parallelen mit dem Feynman-Kleinert-Näherungsansatz bei der approximativen Berechnung von Pfadintegralen auf [37, 28]. Dabei wird an Stelle des tatsächlichen Potentials ein effektives, klassisches Potential bestimmt, das eine optimale Majorante für das exakte darstellt. Daraus folgt eine effektive Zustandssumme; die physikalischen Erwartungswerte werden dann mit Hilfe dieser Approximationen ausgerechnet. Die Bestimmung des effektiven Potentials ist das Pendant zur weiter unten beschriebenen Partition des Hamiltonians des physikalischen Systems, die bei der “Optimal Prediction” Methode die beste Gaußsche Approximation für den exakten Hamiltonian liefert. Allerdings ist im Unterschied zum Feynman-Kleinert-Ansatz, der *a priori* nur den Hamiltonian des Systems berücksichtigt, ein wichtiger Punkt der “Optimal Prediction” Methode die Einbeziehung des Einflusses der anfangs bekannten kollektiven Koordinaten.

1.4. Ziel der Arbeit

In dieser Arbeit wird die Anwendbarkeit der eben präsentierten “Optimal Prediction” Methode auf verschiedene wichtige physikalische Phänomene untersucht, die beispielsweise aus Gründen eingeschränkter Rechnerkapazitäten keine geschlossene Beschreibung zulassen. Insbesondere die Güte der Methode soll mit derjenigen anderer gängiger Reduktionsverfahren verglichen werden, wobei im wesentlichen die Galerkin-Approximation, die Benutzung einer Karhunen-Loeve-Basis und die Zentrumsmannigfaltigkeitsreduktion in Betracht gezogen werden.

Das Galerkin- wie auch das Karhunen-Loeve-Verfahren und die “Optimal Prediction” in 0. Ord-

nung haben gemeinsam, daß die gesuchte Lösung auf den von einem Satz von ausgesuchten Moden (den kollektiven Koordinaten) aufgespannten Raum projiziert wird. Sie unterscheiden sich aber in der Art dieser Projektion.

Die “Optimal Prediction” Methode versucht in 0. Ordnung, zusätzlich zu den anfangs bekannten Werten der kollektiven Koordinaten, von der Kenntnis einer bestimmten, vom System erfüllten statistischen Verteilung der Lösungen Gebrauch zu machen. Dabei wird der Einfluß der unaufgelösten Moden in der Projektion insofern berücksichtigt, als die Amplituden dieser Moden in den Bewegungsgleichungen durch ihre bedingten Erwartungswerte bezüglich des bekannten gegebenen Maßes μ ersetzt werden, seien die Anfangswerte der kollektiven Koordinaten bekannt. Wir erhoffen uns von der “Optimal Prediction” Methode Verbesserungen gegenüber den Galerkin- und Karhunen-Loeve-Approximationen, die diese statistische Information über das System nicht beachten.

Im **ersten Teil** der Arbeit wenden wir uns **Hamiltonschen Systemen** zu.

In **Paragraph 2** wird in einem ersten Schritt die nichtlineare und insbesondere die quintisch nichtlineare Schrödinger-Gleichung mit der eben präsentierten “Optimal Prediction” Methode 0. Ordnung untersucht. Der Anreiz dabei ist die Universalität dieser Gleichung, die in weiten Bereichen der Physik (Plasmaphysik, nichtlineare Optik, etc.) auftritt. Über eine Diskretisierung gelangt man von der partiellen auf gewöhnliche Differentialgleichungen. Dabei wird die große Bedeutung der Wahl der kollektiven Koordinaten und deren Anzahl klar. Es stellt sich heraus, daß davon sowohl bei quintischer als auch bei kubischer Nichtlinearität das Gelingen der Methode bei der Approximation der Dynamik des Systems abhängt. Bei ungünstiger Wahl der kollektiven Koordinaten kann die Anwendung der “Optimal Prediction” Methode sogar zu deutlich schlechteren Ergebnissen führen, als denjenigen eines Galerkin-Verfahrens. Somit wird die Bestimmung *a priori* der relevanten Freiheitsgrade für die Dynamik des zu untersuchenden Systems eine grundlegende Frage, die es vor Anwendung der Methode zu lösen gilt.

In **Paragraph 3** wird das Hamiltonsche System der diskreten sogenannten “self-trapping”-Gleichung betrachtet. Dieses Modell wird in verschiedenen Kontexten der Untersuchung der

Physik der Materie und der physikalischen Chemie angeführt. Dabei sieht man, daß die 0. Ordnung der “Optimal Prediction” zur Beschreibung der Dynamik nicht ausreicht und auf weitere sogenannte “memory”-Terme erweitert werden muß. Mathematische in der Literatur veröffentlichte Herleitungen werden aufgegriffen und an unser Modell angepaßt, bzw. modifiziert, um eine qualitativ gute Beschreibung für die zeitliche Entwicklung des physikalischen Systems zu bekommen. Es stellt sich heraus, daß nicht die vollständige Zeitentwicklung der Methode notwendig ist, sondern daß die erste Ordnung in der Zeit, die “linearized memory”-Approximation, schon gute Ergebnisse liefert.

Im **zweiten Teil** dieser Arbeit gehen wir von der Überlegung aus, daß ein unterbestimmtes Hamiltonsches System sich im Mittel wie ein dissipatives System verhält, um die “Optimal Prediction” in Hinblick auf ihre Anwendung auf **dissipative Systeme** zu untersuchen.

Dabei ist entscheidend, daß wir nicht mehr statistische Aussagen machen, sondern einzelne Anfangswertprobleme angehen.

In **Paragraph 4** werden zunächst die Güte der “finite-time memory”-Approximation und ihr Vorteil in der Beschreibung der Dynamik auf längeren Zeitskalen gegenüber der “linearized memory”-Approximation anhand eines einfachen zweidimensionalen Modells mit einer Dissipation gezeigt. Dieses analytisch exakt lösbare Modell eignet sich ganz besonders für diesen Beweis, da die Entwicklung des Systems bekannt ist.

In **Paragraph 5** wird das physikalische Problem toroidal eingeschlossener Plasmen behandelt. Dabei knüpfen wir an bestimmte, schon entwickelte Modelle an, die darauf zielen, den für die Fusionsforschung wichtigen L-H-Übergang (von der “*low confinement*”- zu der “*high confinement*”- Mode) und die eventuelle Instabilität der H-Mode und anschließende Entstehung von sogenannten ELMs zu beschreiben. Diese Modelle bestehen aus einem System von drei gekoppelten Differentialgleichungen für die aus physikalischen Erwägungen als relevant bezeichneten Moden. Wir erhoffen uns von der “Optimal Prediction” Methode Verbesserungen gegenüber einer Reduktion über die Zentrumsmannigfaltigkeitstheorie. Das Anwenden der Methode bringt zwar deutlich genauere qualitative Ergebnisse, dennoch stellt sich heraus, daß

es hiermit auch nicht möglich ist, die Entstehung der ELMs nachzuvollziehen. Zudem zeigt sich, daß die hervorragenden Ergebnisse des mathematischen Modells von Paragraph 4 auf die Anwesenheit einer einzigen Dämpfung zurückzuführen sind. In dem Fall mehrerer miteinander konkurrierender Dämpfungen, wie es der Fall bei unserer Modellierung des L-H-Übergangs ist, kommt man zu der Schlußfolgerung, daß auch hier gesichert werden muß, daß die kollektiven Koordinaten wirklich die “relevanten” des Systems sind. D. h., damit eine Reduktion über die “Optimal Prediction” gelingt, ist es notwendig, die am meisten gedämpfte Mode zu eliminieren, und diejenige mit niedrigerer Dämpfungsrate als kollektive Koordinate mit hinzuzuziehen. Die Arbeit wird durch eine Zusammenfassung in **Paragraph 6** abgeschlossen.

Teil I.

Hamiltonsche Systeme

2. Schrödingersche Systeme

Die “Optimal Prediction” 0. Ordnung wurde schon von Chorin *et al.* auf die komplexe lineare Schrödinger-Gleichung $u_t = -u_{xx} + m^2 u$ angewandt [17, 18, 16]. Die kollektiven Koordinaten wurden durch innere Produkte von u mit Gaußschen Filterfunktionen gebildet. Verglichen wurden die Ergebnisse zur Zeit $t > 0$ der “Optimal Prediction” Methode 0. Ordnung für einen Satz von $A \gg 1$ Anfangsbedingungen mit dem Mittelwert zum gleichen Zeitpunkt der A integrierten Lösungen. Diese Integration kann beispielsweise so durchgeführt werden, daß zuerst eine Regression die mittlere Lösung zur Zeit $t = 0$ liefert, deren Zeitentwicklung dann mit Hilfe der Greenschen Funktion bestimmt wird. Dabei soll angemerkt werden, daß dies nur möglich ist, da im Falle eines linearen Problems die Zeitintegration und die Mittelung kommutieren [17]. Beobachtet wurde eine gute Übereinstimmung zwischen den Ergebnissen der Integration und der “Optimal Prediction” Methode, die zudem deutlich weniger Speicherplatz benötigt [18]. Darüberhinaus wurde der Einfluß der Wahl der gaußschen Filterfunktionen untersucht: die besten Ergebnisse liefern Berechnungen mit Filterfunktionen, deren Standardabweichung gleich dem für die Variable x benutzten Diskretisierungsschritt ist; wird die Standardabweichung kleiner als der Diskretisierungsschritt, dann weichen die Ergebnisse der “Optimal Prediction” schneller von denjenigen der Integration ab [17].

Die stochastische Konvergenz der “Optimal Prediction” Methode gegen die “richtige”, integrierte Lösung wurde von Hald am Beispiel der Klein-Gordon-Gleichung gezeigt [31].

Anhand der Methode können auch reduzierte Modelle zur Beschreibung der Dynamik von Atomen innerhalb eines Moleküls entwickelt werden; dabei spielen die im allgemeinen nichtlinearen Wechselwirkungen der Atome untereinander eine große Rolle [35].

Wir interessieren uns nun für die folgende komplexe eindimensionale elliptische nichtlineare Schrödinger-Gleichung:

$$uu_t = -u_{xx} + A|u|^p u \quad p \geq 2 \quad (2.1)$$

auf dem x -Intervall $[0, 2\pi]$ mit periodischen Randbedingungen, bekanntem Anfangswert $u(x, 0)$ und entweder attraktiver ($A < 0$) oder repulsiver ($A > 0$) Nichtlinearität.

Dabei ist $p = 2\sigma$, $\sigma \geq 1$ beliebig.

Diese Gleichung ist ein Spezialfall der allgemeinen folgenden Schrödinger-Gleichung

$$uu_t = -\nabla^2 u + A|u|^p u \quad p \geq 2, \quad (2.2)$$

wobei ∇^2 der d -dimensionale Laplace-Operator ist.

Der Name “nichtlineare Schrödinger-Gleichung” (NLSG) beruht auf einer formalen Analogie mit der Schrödinger-Gleichung der Quantenmechanik; dort ist die Zeitentwicklung der Wellenfunktion ψ durch $i\hbar\psi_t = H\psi$ gegeben. Setzt man dann für den Hamilton-Operator $H = -\frac{\hbar^2}{2m}\nabla^2 + V$ ein, wo V das Potential ist, bekommt man:

$$i\hbar\psi_t = \left(-\frac{\hbar^2}{2m}\nabla^2 + V \right) \psi, \quad (2.3)$$

d. h.

$$i\hbar\psi_t = -\frac{\hbar^2}{2m}\Delta\psi + V\psi, \quad (2.4)$$

Δ ist dabei der Laplace-Operator, der im eindimensionalen Fall gleich der zweiten Ableitung nach der Raumkoordinate x ist; ist V ein nichtlineares Wechselwirkungspotential, erkennt man eine formale Analogie zu der betrachteten Gleichung (2.1), die die Bezeichnung “nichtlineare Schrödinger-Gleichung” rechtfertigt.

Die Existenz und die Eindeutigkeit einer Lösung eines solchen nichtlinearen Schrödinger-Anfangswertproblems (2.2) sind eine wichtige mathematische Frage; die globale Existenz einer Lösung

kann für $p < 4/d$ gezeigt werden, man spricht von “subkritischen” Fällen, während für $p = 4/d$ (“kritischer” Fall) und $p > 4/d$ (“superkritische Fälle”) die globale Existenz im allgemeinen nicht bewiesen werden kann. Wir werden gleich sehen, daß diese Bezeichnungen mit der möglichen Entstehung von Singularitäten der Lösung einhergehen.

Die Bewegungsgleichung (2.1) ist der folgenden Integraldarstellung äquivalent

$$u(t) = U(t)u(x, 0) + i(-A) \int_0^t U(t-t')|u(t')|^p u(t') dt', \quad (2.5)$$

wobei $U(t)$ der Schrödinger-Operator $e^{it\Delta}$ ist, und gibt die zeitliche Entwicklung eines Hamiltonschen Systems wieder, bei dem die kanonisch konjugierten Koordinaten u und $\pi = iu^*$ sind. Die entsprechende Hamiltonsche Dichte ist gegeben durch

$$\mathcal{H}[u, u^*] = |u_x|^2 + A \frac{2}{p+2} |u|^{p+2} \quad (2.6)$$

und die Bewegungsgleichung (2.1) ist zu den Hamiltonschen Bewegungsgleichungen äquivalent:

$$u_t = \frac{\delta \mathcal{H}[u, \pi]}{\delta \pi}, \quad u_t^* = -\frac{\delta \mathcal{H}[u, \pi]}{\delta u},$$

wobei

$$\begin{aligned} u_t &= \frac{\delta \mathcal{H}[u, u^*]}{\delta u^*} \frac{\delta u^*}{\delta \pi} \\ &= (-i) \frac{\delta \mathcal{H}[u, u^*]}{\delta u^*} \end{aligned}$$

Für eine eingehende Untersuchung der nichtlinearen Schrödinger-Gleichung, unter anderem deren Lagrange’schen und Hamiltonschen Formalismus wird auf das Werk von Sulem und Sulem verwiesen [48].

Im Falle einer attraktiven Nichtlinearität und unter der Bedingung, daß $u \rightarrow 0$ für $|x| \rightarrow \infty$, besitzt die Gleichung (2.1) spezielle Lösungen, die als “stationär”, “gebunden” oder als “Solitonlösungen” bezeichnet werden, der Form $u(x, t) = v(x)e^{i\lambda^2 t}$ mit dem zeitunabhängigen Profil $v(x)$. Die Bezeichnung “stationär” bezieht sich dabei auf die daraus resultierende Zeitinvarianz

der Wellenintensität [48].

Das Interesse an der NLSG wird von deren universellem Charakter hervorgerufen; beispielsweise liefert sie die kanonische Beschreibung von Wasserwellen an der Oberfläche eines idealen Fluids, von Plasmawellen, oder in der Optik von der Dynamik der Einhüllenden einer quasi-monochromatischen ebenen Welle $\varepsilon\psi e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)}$ mit kleiner ($\varepsilon \ll 1$) aber finiter Amplitude, die sich in einem schwach nichtlinear dispersiven Medium bei vernachlässigter Dissipation ausbreitet [48]. Für kurze Zeiten und kleine Ausbreitungsdistanzen ist die Dynamik linear, kumulative nichtlineare Wechselwirkungen haben aber für große Zeit- und Längenskalen eine signifikante Modulation der Wellenamplitude zur Folge. Die NLSG drückt den Effekt auf der linearen Dispersionsrelation der mit der Modulation dicker werdenden Spektrallinien und der nichtlinearen Resonanzen aus [48]. Eine klassische Untersuchungsmethode ist dabei eine Multiskalen-Analyse, bei der die komplexe Amplitude des Wellenträgers nicht mehr als konstant angesehen werden kann. Dabei wird die Gleichung auf die neuen langsamen Variablen $T = \varepsilon t$ und $\mathbf{X} = \varepsilon \mathbf{x}$ umgeschrieben und die Resonanzen durch Lösungsbedingungen eliminiert. Eine wichtige physikalische Anwendung ist die Ausbreitung eines Laserstrahls in einem nichtlinearen Medium, bei dem der Brechungsindex linear mit der Wellenintensität variiert (Kerr-Effekt).

Die kubisch nichtlineare Schrödinger-Gleichung, die man erhält, wenn man in (2.1) $p = 2$ setzt, besitzt exakte Lösungen der Form $\psi = \psi_0 e^{i\gamma|\psi_0|^2 t}$. Die Untersuchung der linearen Stabilität einer solchen Lösung bezüglich infinitesimalen Störungen der Amplitude und der Phase führt zu der sogenannten “Modulations-” oder Benjamin-Feir-Instabilität: bei attraktiver Nichtlinearität sind Störungen mit Wellenzahl κ , so daß $\kappa^2 < 2\gamma|\psi_0|^2$, linear instabil, während Störungen mit großer Wellenzahl über die Dispersion stabilisiert werden. Die nichtlineare Entwicklung der Modulationsinstabilität kann zu einem Wellenkollaps führen, auf den wir in höherer Dimension gleich zurückkommen werden [48].

Die Anwendung der “Optimal Prediction” Methode auf eine Schrödinger-Gleichung mit kubischer Nichtlinearität wurde im Eindimensionalen von Chorin *et al.* ausgehend von der folgenden

etwas abgewandelten Gleichung untersucht:

$$uu_t = -u_{xx} + \frac{1}{4} [3|u|^2u + u^{*3}] , \quad (2.7)$$

für eine komplexe Variable $u(x) = q(x) + ip(x)$ auf dem x -Intervall $[0, 2\pi]$ mit periodischen Randbedingungen [19]. Für diesen Fall ($p = 2$, $d = 1$, also $p < 4/d$) ist die Existenz einer eindeutigen globalen Lösung für jede Anfangsbedingung gegeben.

Die gewählten kollektiven Koordinaten waren dabei ein Satz von vier “fast”-Fourier-Moden V_α (im Sinne einer diskreten Fourier-Transformation) zur Zeit $t = 0$, der auf die vier kleinsten, d. h. die zur Zeit $t = 0$ für die Beschreibung der großen Skalen wichtigsten “fast”-Fourier-Koeffizienten beschränkt wurde.

Diese spezielle Wahl der Filterfunktionen ermöglicht einen direkten Vergleich zwischen der “Optimal Prediction” Methode und einer pseudo-spektralen Methode zur Integration des unterdeterminierten Systems; darüberhinaus erlaubt sie quantitative Aussagen. Chorin *et al.* kamen zu dem Schluß, daß die “Optimal Prediction” Methode einen wesentlichen Fortschritt gegenüber einer Galerkin-Approximation bietet; die Zeitentwicklung der vier ersten “fast”-Fourier-Koeffizienten wurde sehr gut wiedergegeben und die sich daraus ergebende Regression [41] lieferte zusätzlich zum ersten Moment $\langle u_i \rangle_V$ auch eine sehr gute Approximation für die zweiten $\langle (u_i)^2 \rangle_V$ und dritten $\langle |u_i|^2 u_i \rangle_V$ Momente, im Gegensatz zu dem Galerkin-Verfahren [19].

Wir interessieren uns nun für den grundsätzlich anderen Fall eines eindimensionalen ($d = 1$) Problems mit attraktiver Nichtlinearität der Ordnung $p = 4$:

$$uu_t = -u_{xx} + A|u|^4u \quad (2.8)$$

Das untersuchte x -Intervall bleibt $[0, 2\pi]$ und wir gehen von periodischen Randbedingungen aus.

Im Unterschied zum Fall der eindimensionalen kubisch nichtlinearen Schrödinger-Gleichung

ist bei dieser Nichtlinearität nicht mehr garantiert, daß eine beliebige Anfangsverteilung stabil bleibt. Tatsächlich haben wir nun $p = 4/d$. Wir befinden uns in dem sogenannten “kritischen” Fall und, wie oben schon erwähnt, ist der Beweis der globalen Existenz einer Lösung für eine gegebene Anfangsbedingung aus dem Sobolev-Raum nicht möglich. Wenn kein physikalischer Prozeß (z. B. Dissipation) der Nichtlinearität entgegenkommt, dominiert diese letzte gegenüber der Dispersion und je nach Anfangsbedingung kann eine Singularität, ein sogenannter “Kollaps” der Wellenfunktion stattfinden: zusammen mit einer Konzentration der Verteilung im Ortsraum tritt ein Anwachsen der Amplitude ein, das als “blow-up” bezeichnet wird [48]. Dieser Prozeß entspricht einem Energietransfer von den großen zu den kleinen Skalen, wo Dissipation stattfindet, und wurde für einen solchen Fall einer periodischen eindimensionalen Schrödinger-Gleichung mit quintischer Nichtlinearität schon numerisch nachgewiesen (cf. [48] und Zitate hierin).

Wichtige physikalische Anwendungen dieses Phänomens sind beispielsweise in der Plasma-physik der Energietransfer zu Teilchen über den “Langmuir”-Kollaps oder in der nichtlinearen Optik die Selbstfokussierung von Lichtpulsen in nichtlinearen dispersiven Medien.

Es konnte von Zakharov schon in den 1970er Jahren gezeigt werden, daß sich im Eindimensionalen für den kritischen Fall $p = 4$, wie auch für superkritische Regimes $p > 4$ eine fokussierende Singularität in endlicher Zeit entwickeln kann [53].

Mathematisch kann der “Blow-up” wie folgt definiert werden: betrachten wir den Sobolev-Raum der Funktionen f , deren Norm $\|f\|_{W^{1,2}}$ endlich ist ($< \infty$), wobei $\|f\|_{W^{1,2}}$ als

$$\|f\|_{W^{1,2}} = (\|f\|_2^2 + \|\nabla f\|_2^2)^{\frac{1}{2}}$$

gegeben ist. $\|f\|_2^2$ ist die übliche L^2 -Norm: $\|f\|_2^2 = (\int dx |f(x)|^2)^{\frac{1}{2}}$.

Dann existieren Lösungen $u(x,0)$ des Sobolev-Raums und eine von der Anfangsbedingung abhängige Konstante $t_c(u(x,0))$ mit $0 < t_c < \infty$, so daß

$$\lim_{t \rightarrow t_c} \|u\|_{W^{1,2}} = \infty$$

Äquivalent ist dabei die Aussage, daß $\lim_{t \rightarrow t_c} \|\nabla u\|_{W^{1,2}} = \infty$, da $\|u\|_2^2$ ein konstantes Bewegungsintegral der nichtlinearen Schrödinger-Gleichung ist.

Die untersuchte Gleichung (2.8) gehört zu den Modellen, bei denen $p \geq 4/d$ und deshalb die globale Existenz einer Lösung nicht bewiesen werden kann. Interessant ist die Frage, ob es für diese Fälle dennoch möglich ist, ausreichende Bedingungen für einen “Kollaps” zu finden. Und tatsächlich konnte von Zakharov, Vlasov *et al.* [53, 50] anhand des sogenannten “Virial-Theorems” die Existenz eines Kollapses für bestimmte Anfangswertprobleme bewiesen werden. Dieses Theorem, dessen Herleitung in [48, 47] nachzulesen ist, ist auch als “Varianz-Identität” bekannt und gegeben durch

$$\frac{d^2V}{dt^2} \leq 8H + 4 \frac{4-pd}{p+2} \int |u|^{p+2} d\Gamma \quad (2.9)$$

mit den schon benutzten Notationen p und d und der Varianz: $V := \int |u|^2 |x|^2 d\Gamma$. Dabei ist mit der Integration über Γ die Integration über den ganzen d -dimensionalen Raum gemeint.

Schließlich benennt folgendes Theorem die gesuchten ausreichenden Bedingungen für den Kollaps, allerdings keine detaillierte Struktur der dann auftretenden Singularität:

Sei $p \geq 4/d$ vorausgesetzt. $u(x, 0)$ sei eine Anfangsbedingung mit endlicher Varianz

$V(0)$, die eine der folgenden Bedingungen erfüllt:

- $H(u(x, 0)) < 0$,
- $H(u(x, 0)) = 0$ und $\text{Im}(\int x \cdot (u^* \nabla u) d\Gamma) < 0$,
- $H(u(x, 0)) > 0$ und $\text{Im}(\int x \cdot (u^* \nabla u) d\Gamma) < -4\sqrt{H(u)} \|xu\|_2^2$,

dann wird sich in endlicher Verlaufszeit eine Singularität ausbilden [48, 47].

Umgekehrt liefert die Methode der (quasi-)selbstähnlichen Lösungen eine Beschreibung der dynamischen Struktur des “Blow-up”, allerdings ohne Angabe von Kriterien für einen Kollaps.

Kommen wir nun zu der zu untersuchenden Gleichung (2.8) zurück; sei $A = -1$ festgelegt.

Der direkte Integrationsweg einer solchen Gleichung ist der folgende: das Intervall in x wird

in n Punkten diskretisiert, die Differentialoperatoren werden durch Differenzquotienten ersetzt. Die so gewonnenen n gewöhnlichen Differentialgleichungen

$$i \frac{\partial u_k}{\partial t} = -\frac{u_{k-1} - 2u_k + u_{k+1}}{h^2} + (-1)^k |u_k|^4 u_k, \quad (2.10)$$

wobei h der Diskretisierungsschritt ist, werden mittels eines Runge-Kutta-Verfahrens 4. Ordnung mit adaptiver Schrittweitensteuerung numerisch integriert.

Problematisch ist aber die Tatsache, daß die zu einer qualitativ guten Beschreibung oft nötige Diskretisierung meistens numerisch sehr aufwendbar ist. Die Beschränkung auf die Integration weniger ODEs hätte aber einen Informationsverlust zur Folge. Somit ist man auf eine gute Reduktionsmethode angewiesen; wir befinden uns auf dem typischen möglichen Anwendungsgebiet der “Optimal Prediction” und wollen versuchen, den Mangel an Auflösung durch statistische Information zu kompensieren.

Zerteilen wir jetzt die komplexe Variable u nach Real- und Imaginärteil $u = q + ip$, dann ergeben sich aus (2.8) zwei Bewegungsgleichungen für q und p :

$$\begin{cases} q_t = -p_{xx} + (-1)^k \cdot (q^2 + p^2)^2 p \\ p_t = q_{xx} - (-1)^k \cdot (q^2 + p^2)^2 q, \end{cases}$$

Dem diskretisierten System (2.10) entspricht nun folgendes System für die q_i und p_i :

$$\begin{cases} \frac{\partial q_k}{\partial t} = -\frac{p_{k-1} - 2p_k + p_{k+1}}{h^2} + (-1)^k \cdot (q_k^2 + p_k^2)^2 p_k \\ \frac{\partial p_k}{\partial t} = \frac{q_{k-1} - 2q_k + q_{k+1}}{h^2} - (-1)^k \cdot (q_k^2 + p_k^2)^2 q_k, \end{cases}$$

deren Zeitentwicklung nun gesucht wird.

q und p können als zueinander kanonisch konjugierte Koordinaten zum folgenden Hamiltonian aufgefaßt werden:

$$H(q, p) = \frac{1}{2} \int_0^{2\pi} \left((q_x^2 + p_x^2) + (-1)^k \frac{1}{3} (q^2 + p^2)^3 \right) dT \quad (2.11)$$

Nehmen wir mal an, das “*prior measure*” sei das vom Hamiltonian kanonisch induzierte Maß μ , dessen Dichte durch

$$f(u) = \frac{1}{Z} e^{-\mathcal{H}(u)}$$

gegeben ist. Über eine Monte-Carlo-Simulation und speziell einen Metropolis-Algorithmus [8] sind wir in der Lage, einen Satz von vielen Anfangsdaten zu bestimmen, die nach dem invarianten Maß μ verteilt sind, und gleichzeitig die folgenden Bedingungen zur Zeit $t = 0$ erfüllen:

$$\sum_{i=1}^N g_{\alpha}(x_i) u_i(x) = V_{\alpha} \quad \alpha = 1, \dots, N < n,$$

mit anfangs bekannten Filterfunktionen $g_{\alpha}(x)$. Typischerweise muß man mit der Monte-Carlo-Simulation ein Ensemble der Größenordnung von mindestens $10^2 - 10^3$ Anfangsbedingungen würfeln, damit die Aussagen sinnvoll sind, denn der Fehler bei einer Monte-Carlo-Simulation eines Ensembles der Größe N_e ist proportional zu $\frac{1}{\sqrt{N_e}}$ [8, 2].

Nehmen wir an, der Diskretisierungsschritt in x sei: $x_i = i \cdot h = i \cdot \frac{2\pi}{n}$, mit $n = 32$ Stützstellen und es sei uns nur möglich, $N = 4$ Freiheitsgrade statt die zur vollständigen Auflösung nötigen $n = 32$ zu berücksichtigen. Mit diesen anfangs bekannten $N = 4$ kollektiven Koordinaten ist aber das Cauchy-Problem, das aus der Diskretisierung von (2.8) folgt, unterbestimmt. Folgt man der in der Einleitung dargestellten "Optimal Prediction" Methode, so soll die Gleichung (1.13) die Zeitentwicklung der kollektiven Koordinaten V_{α} liefern.

Nun müssen wir die Filterfunktionen $g_{\alpha}(x)$ wählen. Entscheidet man sich für eine diskrete Fourier-Transformation, wird der Satz der $n = 32$ Gleichungen (2.10) über die Funktionen $g_{ki} = \frac{1}{n} e^{-iK_k x_i}$, mit $K_k = -n/2 + k$, auf ein äquivalentes System von 32 gekoppelten Differentialgleichungen für die n so gewonnenen $F_k = \sum_{i=1}^n g_{ki} u_i$, $1 \leq k \leq n$ transformiert. Erforderlich für eine geschlossene Beschreibung wäre hier auch die Lösung der 32 Gleichungen für die F_k . Allerdings beschränkt man sich auf das zur Zeit $t = 0$ bekannte, unterbestimmte System von $N = 4$ kollektiven Koordinaten:

$$V_{\alpha} = \sum_{i=1}^n g_{\alpha i} u_i, \quad \text{mit} \quad g_{\alpha i}(x) = \frac{1}{n} e^{(-iK_{\alpha} x_i)}, \quad \text{wobei} \quad K_{\alpha} = -\frac{N}{2} + 1, \dots, \frac{N}{2} \quad (2.12)$$

Dabei sind

$$V_{\alpha} = F_{n/2 - N/2 + \alpha}, \quad \text{für} \quad \alpha = 1, \dots, N \quad (2.13)$$

und mit u_i ist der Wert der komplexen Funktion u an der Stelle x_i gemeint.

Alle anderen

$$F_\mu \quad \text{mit} \quad \mu \leq \frac{n}{2} - \frac{N}{2} \quad \text{und} \quad \mu \geq \frac{n}{2} + \frac{N}{2} + 1, \quad (2.14)$$

die zur geschlossenen Lösung des (2.10) äquivalenten Systems nötig wären, bleiben unberücksichtigt.

Wir “würfeln” nun mit Hilfe einer Monte-Carlo-Simulation einen Satz der Größenordnung 10^4 von Anfangsdaten, die zur Zeit $t = 0$ diese Fourier-Koeffizienten aufweisen. Wir wollen über die “Optimal Prediction” Methode deren weitere Entwicklung bestimmen, damit anhand der Werte $V_\alpha(t > 0)$ eine Regression Aussagen über die Funktionen $q_i(t), p_i(t)$ zu späteren Zeiten machen kann.

2.1. Gaußsches Maß

Die exakte analytische Berechnung von den Erwartungswerten ($g, \langle R(u) \rangle_V$) auf der rechten Seite von (1.13) kann unter Umständen ein langwieriger Prozeß sein und stellt die Frage nach einem möglichen, etwas vereinfachenden Näherungsverfahren. Der von Chorin *et al.* eingeschlagene Weg besteht darin [19], die gegebene Verteilung über eine Störungstheorie auf eine Gaußsche Verteilung zurückzuführen. Dann sind nämlich die gesuchten bedingten Erwartungswerte bekannt. Wir präsentieren im folgenden die dazu nötigen Ergebnisse für eine komplexe Variable u ; sollte die Variable u reell sein, stehen uns dann ähnliche Theoreme zur Verfügung, die im Anhang angeführt werden.

Die Wahrscheinlichkeitsdichte, daß die normal verteilte Zufallsvariable $\mathbf{u} = (u_1, \dots, u_n) = (q_1 + ip_1, \dots, q_n + ip_n)$ sich in dem Zustand $(s_1, \dots, s_n, r_1, \dots, r_n)$ befindet, ist gegeben durch:

$$\begin{aligned} P(s_1 < q_1 \leq s_1 + ds_1, \dots, s_n < q_n \leq s_n + ds_n, r_1 < p_1 \leq r_1 + dr_1, \dots, r_n < p_n \leq r_n + dr_n) = \\ = \frac{1}{Z} e^{(-\frac{1}{2}(\mathbf{z}, \mathbf{A}\mathbf{z}^*))} dz_1 \dots dz_n, \end{aligned}$$

wobei $z_i = s_i + ir_i$ und $dz_i = ds_i dr_i$ sind.

Die Matrix A ist dabei hermitsch, positiv definit und ihre Inverse A^{-1} ist die Matrix der Kovarianzen, deren Elemente durch:

$$a_{ij}^{-1} = Cov\{u_i, u_j^*\} \equiv \langle u_i u_j^* \rangle - \langle u_i \rangle \langle u_j^* \rangle$$

gegeben sind.

Der Vektor \mathbf{b} hängt mit den Erwartungswerten von \mathbf{u} über die folgende Beziehung zusammen:

$$A^{-1} \mathbf{b} = \langle u \rangle$$

Die Verteilung ist somit durch die Angabe der n Mittelwerte und der $\frac{1}{2}n(n+1)$ unabhängigen Werte der Kovarianzmatrix vollkommen bestimmt. Alle gesuchten Erwartungswerte können also durch diese Parameter ausgedrückt und anhand von den drei folgenden Lemmata berechnet werden, deren Beweise in [17] nachzulesen sind.

- Lemma 1: Der “bedingte” Erwartungswert von u_i ist eine affine Funktion der kollektiven Koordinaten V_α :

$$\langle u_i \rangle_V = q_{i\alpha} V_\alpha \quad (2.15)$$

wobei die Elemente $q_{i\alpha}$ der $n \times N$ Matrix Q durch

$$Q = (CG^\dagger)(GCG^\dagger)^{-1}$$

gegeben sind.

C ist die $n \times n$ Matrix der komplexen Kovarianzen:

$$C_{ij} = \langle u_i u_j^* \rangle = \frac{1}{\pi} \sum_{k=1}^n \frac{e^{ik(x_i - x_j)}}{\frac{4}{h^2} \sin^2 \frac{hk}{2} + m_0^2} \quad (2.16)$$

Die bedingten Erwartungswerte sind also linear in den V_α und unabhängig von Multiplikationsfaktoren in den Kovarianzen.

- Lemma 2: Die “bedingte” Kovarianzmatrix lautet:

$$\begin{aligned} Cov\{u_i, u_j^*\}_V &= \langle u_i u_j^* \rangle_V - \langle u_i \rangle_V \langle u_j^* \rangle_V \\ &= [C - (CG^\dagger)(GCG^\dagger)^{-1}(GC)]_{ij} \end{aligned} \quad (2.17)$$

Die bedingten Kovarianzen sind also nur von den Filterfunktionen g_α abhängig und nicht von den kollektiven Koordinaten V_α .

Schließlich erlaubt das Wick'sche Theorem, den Erwartungswert eines Produktes über folgende Formel zu berechnen:

- Lemma 3:

$$\begin{aligned} & \langle \prod_{p_1=1}^{P_1} (u_{i_{p_1}} - \langle u_{i_{p_1}} \rangle_V) \prod_{p_2=1}^{P_2} (u_{i_{p_2}}^* - \langle u_{i_{p_2}}^* \rangle_V) \rangle_V = \\ & \begin{cases} 0 & P_1 + P_2 \text{ ungerade} \\ \sum_{\text{alle Permutationen}} \text{Cov}\{u_{i_1}, u_{i_2}^*\}_V \cdots \text{Cov}\{u_{i_{P_1+P_2-1}}, u_{i_{P_1+P_2}}^*\}_V & P \text{ gerade} \end{cases} \end{aligned} \quad (2.18)$$

Das Wick'sche Theorem ermöglicht, jeden Mittelwert der Form $\langle u_i^p u_j^{*l} \rangle_V$ als Funktion von ausschließlich $\langle u_i^2 \rangle_V$ und $\langle u_i u_j^* \rangle_V$ auszudrücken [13].

2.2. Nicht-Gaußsches Maß

Ist aber $\mu(u)$ kein Gaußsches Maß, dann bemüht man eine Störungstheorie [27, 43, 36, 19], die die Regression bezüglich des nicht-Gaußschen Maßes auf eine bezüglich eines Gaußschen Maßes zurückführt. Dafür wird der Hamiltonian in einen Gaußschen Anteil \mathcal{H}^0 und einen nicht-Gaußschen Anteil \mathcal{H}^1 partitioniert:

$$\mathcal{H} = \mathcal{H}^0 + \mathcal{H}^1 \quad (2.19)$$

Damit geht eine kanonische Partitionierung des Maßes μ in einen Gaußschen Anteil μ^0 und einen nicht-Gaußschen Anteil einher.

Der bedingte Erwartungswert bezüglich des vollständigen Maßes μ eines Funktionals $\mathcal{F}(u)$ ist definiert als:

$$\langle \mathcal{F} \rangle_V = \frac{1}{Z} \int \mathcal{F}(u) d\mu_V \quad (2.20)$$

Erwartungswerte bezüglich des Gaußschen Maßes μ^0 sind entsprechend gegeben durch:

$$\langle \mathcal{F} \rangle_V^0 = \frac{1}{Z_0} \int \mathcal{F}(u) d\mu_V^0$$

Dabei sind Z und Z_0 die jeweiligen Zustandssummen.

Benutzt man nun die folgende Faktorisierung

$$e^{-\mathcal{H}} = e^{-\mathcal{H}^0} \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} (\mathcal{H}^1)^k,$$

dann kann man die Integration auf der rechten Seite von (2.20) auf eine Integration über $d\mu_V^0$, und somit auf eine ausführbare Integration über ein Gaußsches Maß umschreiben:

$$\int \mathcal{F}(u) d\mu_V = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \int [\mathcal{F}(u) (\mathcal{H}^1)^k] d\mu_V^0$$

Auf der rechten Seite erkennt man, bis auf den Faktor Z_0 , die Definition von $\langle (\mathcal{H}^1)^k \mathcal{F} \rangle_V^0$:

$$\langle \mathcal{F} \rangle_V = \sum_{k=0}^{\infty} \frac{(-1)^k Z_0}{k! Z} \langle (\mathcal{H}^1)^k \mathcal{F} \rangle_V^0$$

Nun kann das Verhältnis beider Zustandssummen explizit berechnet werden:

$$\left(\frac{Z_0}{Z} \right)^{-1} = \frac{\int d\mu_V}{\int d\mu_V^0} = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \langle (\mathcal{H}^1)^k \rangle_V^0$$

Schließlich erhält man für den Erwartungswert $\langle \mathcal{F} \rangle_V$ bezüglich des ursprünglichen nicht-Gaußschen Maßes:

$$\langle \mathcal{F} \rangle_V = \frac{\sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \langle (\mathcal{H}^1)^k \mathcal{F} \rangle_V^0}{\sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \langle (\mathcal{H}^1)^k \rangle_V^0}$$

und somit für die gesuchten Erwartungswerte der Gleichung (1.13):

$$(g_\alpha, \langle R(u) \rangle_V) = \frac{\sum_{k=0}^{\infty} \frac{(-1)^k}{k!} (g_\alpha, \langle R(u) (\mathcal{H}^1)^k \rangle_V^0)}{\sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \langle (\mathcal{H}^1)^k \rangle_V^0} \quad (2.21)$$

Somit ist gezeigt, daß man bei Untersuchung eines Hamiltonschen Problems mit nicht-Gaußschem Hamiltonian die Berechnung mittels einer Störungstheorie auf den Gaußschen Fall zurückführen kann, nachdem man die Partitionierung (2.19) durchgeführt hat.

Die Mittelung auf der rechten Seite von (1.13) bezüglich der normalen Wahrscheinlichkeit \mathcal{H}^0 im Sinne der in (2.19) angeführten Partition liefert außerdem die 0. Ordnung in dieser Entwicklung [19]. Auf sie werden wir uns in dieser Arbeit beschränken.

2.3. Anwendung der “Optimal Prediction” auf den Kollapsfall

Der Hamiltonian (2.41) ist nicht Gaußscher Form, deswegen ist eine Partitionierung nötig, wie sie in dem vorigen Abschnitt dargestellt wurde:

$$\mathcal{H} = \mathcal{H}^0 + \mathcal{H}^1, \quad (2.22)$$

mit dem Gaußschen Anteil

$$H^0 = \frac{1}{2} \int_0^{2\pi} (|u_x|^2 + m_0^2 |u|^2 + b) dx \quad (2.23)$$

und dem nicht-Gaußschen Anteil

$$H^1 = \frac{1}{2} \int_0^{2\pi} \left((-1) \frac{1}{3} |u|^6 - m_0^2 |u|^2 - b \right) dx \quad (2.24)$$

Gesucht sind nun Werte von m_0 und b , die garantieren, daß $\mathcal{H}^1(q, p)$ eine Störung zum System mit Hamiltonian $\mathcal{H}^0(q, p)$ darstellt, so daß das Maß μ , die von \mathcal{H} definiert und von den V_α bedingt wird, sich von dem von denselben V_α bedingten Maß μ_0 nur um eine kleine Störung unterscheidet.

Es sind von Chorin *et al.* zwei Möglichkeiten präsentiert worden, die in diesem Sinne optimalen Werte für m_0 und b zu bestimmen [19]: entweder leitet man zuerst m_0 aus der Bedingung her, daß die Varianz $\langle (\mathcal{H}^1)^2 \rangle^0$ als Funktion von m_0 minimiert wird. Dann folgt aus $\langle \mathcal{H}^1 \rangle^0 = 0$ schließlich b . Oder es werden mittels einer Monte-Carlo-Simulation die Korrelationsfunktionen

$\langle u_i u_j^* \rangle$ bezüglich des Maßes μ berechnet, und dann m_0 aus der Forderung bestimmt, daß das so induzierte μ_0 am besten diese Korrelationsfunktionen approximiert, indem der mittlere quadratische Unterschied zwischen den beiden minimiert wird. Dieses letzte Verfahren minimiert den Unterschied zwischen dem vollständigen Maß und seinem quadratischen Anteil, während die vorige Methode nur die ersten Entwicklungsterme minimiert; beide liefern die gleiche optimale Wahl der Werte m_0 und b [19]. Wir wenden zur Bestimmung von m_0 und b die erste Methode an.

2.3.1. Bestimmung von m_0 und b

Der nicht-Gaußsche Anteil des Hamiltonians, ausgedrückt in den Variablen p und q , sieht folgendermaßen aus:

$$H^1(q, p) = \frac{1}{2} \int_0^{2\pi} \left((-1) \frac{1}{3} (q^2 + p^2)^3 - m_0^2 (q^2 + p^2) \right) dT,$$

was nach dem Ausmultiplizieren ergibt:

$$H^1(q, p) = \frac{1}{2} \int_0^{2\pi} \left((-1) \frac{1}{3} (q^6 + 3q^4 p^2 + 3q^2 p^4 + p^6) - m_0^2 (q^2 + p^2) \right) dT$$

Quadriert ergibt diese Gleichung:

$$H^1(q, p)^2 = \frac{1}{4} \int_0^{2\pi} \left[\frac{-1}{3} \left(q^6(x) + 3q^4(x)p^2(x) + 3q^2(x)p^4(x) + p^6(x) \right) - m_0^2 (q^2(x) + p^2(x)) \right] dx \\ * \int_0^{2\pi} \left[\frac{-1}{3} \left(q^6(z) + 3q^4(z)p^2(z) + 3q^2(z)p^4(z) + p^6(z) \right) - m_0^2 (q^2(z) + p^2(z)) \right] dz$$

Die Varianz ist also gegeben durch:

$$\langle H^1(q, p)^2 \rangle = \frac{1}{4} \int_0^{2\pi} dx \int_0^{2\pi} dz \{ \\ \langle \frac{(-1)^2}{9} \left(q^6(x) + 3q^4(x)p^2(x) + 3q^2(x)p^4(x) + p^6(x) \right) * \\ * \left(q^6(z) + 3q^4(z)p^2(z) + 3q^2(z)p^4(z) + p^6(z) \right) \rangle \\ - \frac{2(-1)}{3} m_0^2 \langle \left(q^6(x) + 3q^4(x)p^2(x) + 3q^2(x)p^4(x) + p^6(x) \right) (q^2(z) + p^2(z)) \rangle \\ + m_0^4 \langle (q^2(x) + p^2(x)) (q^2(z) + p^2(z)) \rangle \}$$

Dabei soll angemerkt werden, daß wir der Lesbarkeit wegen in diesem Abschnitt auf den Exponenten 0 in dem Erwartungswert $\langle \cdot \rangle^0$ verzichten.

Die Forderung nach dem Minimieren der Varianz als Funktion von m_0 erfordert die Berechnung der folgenden Terme:

$$\mathcal{A} := \langle \left(q^6(x) + 3q^4(x)p^2(x) + 3q^2(x)p^4(x) + p^6(x) \right) \left((q^6(z) + 3q^4(z)p^2(z) + 3q^2(z)p^4(z) + p^6(z)) \right) \rangle,$$

$$\mathcal{B} := \langle \left(q^6(x) + 3q^4(x)p^2(x) + 3q^2(x)p^4(x) + p^6(x) \right) (q^2(z) + p^2(z)) \rangle,$$

und

$$C := \langle (q^2(x) + p^2(x)) (q^2(z) + p^2(z)) \rangle$$

Bestimmung von C

Wenden wir uns zuerst dem dritten Term zu:

$$C = \langle (q^2(x) + p^2(x)) (q^2(z) + p^2(z)) \rangle$$

$$C = 2 \langle p^2(x)p^2(z) \rangle + 2 \langle q^2(x)p^2(z) \rangle$$

Da q und p bezüglich eines Gaußschen Maßes unkorreliert sind, folgt:

$$C = 2 \langle p^2(x)p^2(z) \rangle + 2 \langle p^2 \rangle^2$$

Aus dem Wick'schen Theorem folgt dann:

$$C = 2(\langle p^2 \rangle^2 + 2 \langle p(x)p(z) \rangle^2) + 2 \langle p^2 \rangle^2,$$

was schließlich für C ergibt:

$$C = 4 \langle p^2 \rangle^2 + 4 \langle p(x)p(z) \rangle^2 \tag{2.25}$$

Bestimmung von \mathcal{B}

Nach dem Ausmultiplizieren und unter Berücksichtigung von:

$$\langle q^6(x)p^2(z) \rangle = \langle p^6(x)q^2(z) \rangle$$

$$\begin{aligned} \langle q^4(x)q^2(z)p^2(z) \rangle &= \langle q^2(x)p^4(z)p^2(z) \rangle \\ &= \langle p^2 \rangle \langle p^4(x)p^2(z) \rangle \end{aligned}$$

$$\begin{aligned} \langle q^4(x)p^2(x)p^2(z) \rangle &= \langle q^2(x)q^2(z)p^4(x) \rangle \\ &= \langle p^4 \rangle \langle p^2(x)p^2(z) \rangle, \end{aligned}$$

da q und p völlig unabhängig und gleichwertig betrachtet werden, erhält man für \mathcal{B}

$$\begin{aligned} \mathcal{B} &= 2 \langle p^6(x)p^2(z) \rangle + 2 \langle p^6 \rangle \langle p^2 \rangle \\ &\quad + 6 \langle p^2 \rangle \langle p^4(x)p^2(z) \rangle + 6 \langle p^4 \rangle \langle p^2(x)p^2(z) \rangle \end{aligned} \quad (2.26)$$

Das Wick'sche Theorem ermöglicht uns jetzt, die Erwartungswerte von Produkten zu vereinfachen, bzw. zu berechnen:

$$\begin{aligned} \langle p^6 \rangle &= 15 \langle p^2 \rangle^3, \\ \langle p^4 \rangle &= 3 \langle p^2 \rangle^2, \\ \langle p^2(x)p^2(z) \rangle &= \langle p^2 \rangle^2 + 2 \langle p(x)p(z) \rangle^2, \\ \langle p^4(x)p^2(z) \rangle &= 3 \langle p^2 \rangle^3 + 6 \langle p^2 \rangle \langle p(x)p(z) \rangle^2, \\ \langle p^6(x)p^2(z) \rangle &= 15 \langle p^2 \rangle^4 + 30 \langle p^2 \rangle^2 \langle p(x)p(z) \rangle^2. \end{aligned}$$

Setzt man nun diese Ergebnisse in (2.26) ein, dann erhält man:

$$\begin{aligned} \mathcal{B} &= 2 (15 \langle p^2 \rangle^4 + 30 \langle p^2 \rangle^2 \langle p(x)p(z) \rangle^2) \\ &\quad + 2 (15 \langle p^2 \rangle^3) \langle p^2 \rangle \\ &\quad + 6 \langle p^2 \rangle (3 \langle p^2 \rangle^3 + 6 \langle p^2 \rangle \langle p(x)p(z) \rangle^2) \\ &\quad + 6 (3 \langle p^2 \rangle^2) (\langle p^2 \rangle^2 + 2 \langle p(x)p(z) \rangle^2) \end{aligned}$$

und schließlich:

$$\mathcal{B} = 96 \langle p^2 \rangle^4 + 132 \langle p^2 \rangle^2 \langle p(x)p(z) \rangle^2 \quad (2.27)$$

Bestimmung von \mathcal{A}

Die Bestimmung von \mathcal{A} erfolgt jetzt gleichermaßen; berücksichtigt werden müssen die folgenden Gleichungen:

$$\begin{aligned} \langle q^6(x)q^4(z)p^2(z) \rangle &= \langle q^2(x)p^4(x)p^6(z) \rangle \\ &= \langle p^6(x)q^2(z)p^4(z) \rangle \\ &= \langle q^4(x)p^2(x)q^6(z) \rangle \\ &= \langle p^2 \rangle \langle p^6(x)p^4(z) \rangle, \end{aligned}$$

$$\begin{aligned} \langle q^6(x)q^2(z)p^4(z) \rangle &= \langle q^4(x)p^2(x)p^6(z) \rangle \\ &= \langle q^2(x)p^4(x)q^6(z) \rangle \\ &= \langle p^6(x)q^4(z)p^2(z) \rangle \\ &= \langle p^4 \rangle \langle p^6(x)p^2(z) \rangle, \end{aligned}$$

$$\begin{aligned} \langle q^4(x)p^2(x)q^4(z)p^2(z) \rangle &= \langle q^2(x)p^4(x)q^2(z)p^4(z) \rangle \\ &= \langle q^2(x)q^2(z) \rangle \langle p^4(x)p^4(z) \rangle \\ &= \langle p^4(x)p^4(z) \rangle \langle p^2(x)p^2(z) \rangle \end{aligned}$$

und

$$\begin{aligned} \langle q^4(x)p^2(x)q^2(z)p^4(z) \rangle &= \langle q^2(x)p^4(x)q^4(z)p^2(z) \rangle \\ &= (\langle p^4(x)p^2(z) \rangle)^2 \end{aligned}$$

Man gelangt zu folgendem Ausdruck:

$$\begin{aligned} \mathcal{A} &= 2 \langle p^6(x)p^6(z) \rangle + 2 \left(\langle p^6 \rangle \right)^2 \\ &\quad + 12 \langle p^2 \rangle \langle p^6(x)p^4(z) \rangle + 12 \langle p^4 \rangle \langle p^6(x)p^2(z) \rangle \\ &\quad + 18 \langle p^4(x)p^4(z) \rangle \langle p^2(x)p^2(z) \rangle \\ &\quad + 18 \left(\langle p^4(x)p^2(z) \rangle \right)^2 \end{aligned} \tag{2.28}$$

Das Wick'sche Theorem liefert Vereinfachungen:

$$\langle p^6 \rangle = 15 \langle p^2 \rangle^3, \quad (2.29)$$

$$\begin{aligned} \left(\langle p^6 \rangle \right)^2 &= 225 \langle p^2 \rangle^6, \\ \langle p^4 \rangle &= 3 \langle p^2 \rangle^2, \end{aligned} \quad (2.30)$$

$$\langle p^6(x)p^2(z) \rangle = 15 \langle p^2 \rangle^4 + 30 \langle p^2 \rangle^2 \langle p(x)p(z) \rangle^2,$$

$$\langle p^6(x)p^4(z) \rangle = 45 \langle p^2 \rangle^5 + 180 \langle p^2 \rangle^3 \langle p(x)p(z) \rangle^2 + 360 \langle p^2 \rangle (\langle p(x)p(z) \rangle)^4,$$

$$\begin{aligned} \langle p^6(x)p^6(z) \rangle &= 225 \langle p^2 \rangle^6 + 1350 \langle p^2 \rangle^4 (\langle p(x)p(z) \rangle)^2 \\ &\quad + 5400 \langle p^2 \rangle^2 (\langle p(x)p(z) \rangle)^4 + 10800 (\langle p(x)p(z) \rangle)^6, \end{aligned}$$

und schließlich

$$\langle p^4(x)p^4(z) \rangle = 9 \langle p^2 \rangle^4 + 36 \langle p^2 \rangle^2 \langle p(x)p(z) \rangle^2 + 72 \langle p(x)p(z) \rangle^4$$

Setzen wir diese Ergebnisse in (2.28) ein, folgt:

$$\begin{aligned} \mathcal{A} &= 2304 \langle p^2 \rangle^6 + 7560 \langle p^2 \rangle^4 \langle p(x)p(z) \rangle^2 \\ &\quad + 18360 \langle p^2 \rangle^2 \langle p(x)p(z) \rangle^4 + 24192 \langle p(x)p(z) \rangle^6 \end{aligned} \quad (2.31)$$

Nun können wir das Minimum der Varianz $\langle (H^1)^2 \rangle$ als Funktion von m_0 bestimmen.

$$\langle (H^1)^2 \rangle = \frac{1}{4} \int_0^{2\pi} dx \int_0^{2\pi} dz \left(\frac{1}{9} \mathcal{A} + \frac{2}{3} m_0^2 \mathcal{B} + m_0^4 \mathcal{C} \right)$$

Setzen wir nun die eben in (2.31), (2.27) und (2.25) berechneten Ausdrücke für \mathcal{A} , \mathcal{B} und \mathcal{C} ein, bekommt man:

$$\begin{aligned} \langle (H^1)^2 \rangle &= [256\pi^2 \langle p^2 \rangle^6 + 210 \langle p^2 \rangle^4 \iint dx dz \langle p(x)p(z) \rangle^2 \\ &\quad + 510 \langle p^2 \rangle^2 \iint dx dz \langle p(x)p(z) \rangle^4 + 672 \iint dx dz \langle p(x)p(z) \rangle^6] \\ &\quad + 64m_0^2\pi^2 \langle p^2 \rangle^4 + 22m_0^2 \langle p^2 \rangle^2 \iint dx dz \langle p(x)p(z) \rangle^2 \\ &\quad + 4m_0^4\pi^2 \langle p^2 \rangle^2 + m_0^4 \iint dx dz \langle p(x)p(z) \rangle^2 \end{aligned}$$

Dabei wurde die Tatsache ausgenutzt, daß $\langle p^2 \rangle$ unabhängig von der Ortsvariable x oder z ist, und sich deshalb einfach ausintegrieren läßt: $\int_0^{2\pi} dx \int_0^{2\pi} dz \langle p^2 \rangle^k = 4\pi^2 \langle p^2 \rangle^k$, mit k beliebig.

Sei nun das folgende Integral definiert:

$$I_m(q) := \int_0^{2\pi} dx \int_0^{2\pi} dz (\langle p(x)p(z) \rangle)^{2q}$$

Mit diesem Integral läßt sich $\langle (H^1)^2 \rangle$ etwas übersichtlicher schreiben:

$$\begin{aligned} \langle (H^1)^2 \rangle &= \left[256\pi^2 \langle p^2 \rangle^6 + 210 \langle p^2 \rangle^4 I_m(1) + 510 \langle p^2 \rangle^2 I_m(2) + 672 I_m(3) \right] \\ &\quad + m_0^2 [64\pi^2 \langle p^2 \rangle^4 + 22 \langle p^2 \rangle^2 I_m(1)] \\ &\quad + m_0^4 [4\pi^2 \langle p^2 \rangle^2 + I_m(1)], \end{aligned}$$

Zur numerischen Berechnung von $\langle (H^1)^2 \rangle$ als Funktion von m_0 brauchen wir nun die Werte der Integrale $I_m(q)$. Eine analytische Berechnung mit Hilfe einer Rekursion liefert:

$$I_m(q) = \frac{1}{(2\pi m_0^2)^{2q}} \sum_{l=0}^{2q} C_{2q}^l (2m_0^2)^l I_l$$

mit

$$I_l = \begin{cases} 4\pi^2 & \text{für } l = 0 \\ 0 & \text{für } l = 1 \\ \frac{2\pi^2}{2^{l-1}} \sum_{k_1=-\infty, k_1 \neq 0}^{+\infty} \frac{1}{k_1^2 + m_0^2} \cdots \sum_{k_{l-1}=-\infty, k_{l-1} \neq 0}^{+\infty} \frac{1}{k_{l-1}^2 + m_0^2} \frac{1}{(k_1 + \dots + k_{l-1})^2 + m_0^2} & \text{für } l \geq 2 \end{cases}$$

Somit hat man also die Varianz $\langle (H^1)^2 \rangle$ auf eine mit m_0 parametrisierte Funktion von den Autokorrelationen:

$$\langle p^2 \rangle = \langle q^2 \rangle = \frac{1}{2\pi} \sum_{k=-\infty}^{+\infty} \frac{1}{k^2 + m_0^2}$$

und den Korrelationen:

$$\langle p(x)p(z) \rangle = \frac{1}{2\pi} \sum_{k=-\infty}^{+\infty} \frac{e^{ik(x-z)}}{k^2 + m_0^2}.$$

umgeschrieben.

Diese Funktion kann bezüglich m_0 minimiert werden; daraus ergibt sich der "optimale" Wert

$$m_0 = 1.703$$

Daraus folgt für b :

$$\langle H^1(q, p) \rangle = 0 \implies \int_0^{2\pi} \left\langle \left(\frac{-1}{3} (q^2 + p^2)^3 - m_0^2 (q^2 + p^2) - b \right) \right\rangle dx = 0,$$

$$b = \frac{1}{2\pi} \int_0^{2\pi} \left\langle \left(\frac{-1}{3} (q^6 + 3q^4 p^2 + 3q^2 p^4 + p^6) - m_0^2 (q^2 + p^2) \right) \right\rangle dx,$$

$$b = \frac{1}{2\pi} \int_0^{2\pi} \left\langle \left(\frac{-1}{3} (2 \langle q^6 \rangle + 6 \langle q^4 \rangle \langle q^2 \rangle) - 2m_0^2 \langle q^2 \rangle \right) \right\rangle dx,$$

$\langle q^4 \rangle$ und $\langle q^6 \rangle$ werden nach den Gleichungen (2.29) und (2.30) umgeschrieben:

$$b = \frac{1}{2\pi} \int_0^{2\pi} \left\langle \left(\frac{-1}{3} (2 \cdot 15 \langle q^2 \rangle^3 + 6 \cdot 3 \langle q^2 \rangle^2 \langle q^2 \rangle) - 2m_0^2 \langle q^2 \rangle \right) \right\rangle dx,$$

$$b = \frac{1}{2\pi} \int_0^{2\pi} \left\langle (-16 \langle q^2 \rangle^3 - 2m_0^2 \langle q^2 \rangle) \right\rangle dx,$$

und schließlich erhält man:

$$b = - (16 \langle q^2 \rangle^3 + 2m_0^2 \langle q^2 \rangle),$$

woraus für den Wert $m_0 = 1.703$,

$$b = -1.957$$

folgt.

Somit haben wir die Partition des ursprünglichen Hamiltonians (2.41) in der Form, die in den Gleichungen (2.22), (2.23) und (2.24) gegeben ist, mit dem "optimalen" Satz von Parametern

$$\begin{cases} m_0 = 1.703, \\ b = -1.957 \end{cases} \quad (2.32)$$

bestimmt. "Optimal" meint hier, wie am Ende des vorigen Abschnittes erklärt, daß für diese Wahl von Parametern der mittlere quadratische Unterschied zwischen dem reellen Hamiltonian und seiner Gaußschen Approximation bis zur dritten Ordnung minimiert ist. Wir können also nun die Erwartungswerte auf der rechten Seite von (1.13) in 0. Ordnung berechnen, indem wir sie den Erwartungswerten bezüglich des Gaußschen Hamiltonians (2.23) gleichstellen und daraus die von der "Optimal Prediction" vorhergesagte Zeitentwicklung unserer kollektiven Koordinaten folgern.

2.3.2. Zeitentwicklung der V_α

Die Zeitentwicklung der kollektiven Koordinaten ist durch

$$i \frac{dV_\alpha}{dt} = \sum_{k=1}^n g_{\alpha k} < -\frac{u_{k-1} - 2u_k + u_{k+1}}{h^2} + (-1)|u_k|^4 u_k >_V^0 \quad (2.33)$$

gegeben.

In die rechte Seite dieser Gleichung setzt man für u_i , $u_i = n \sum_{\beta=1}^N g_{i\beta}^* V_\beta$, ein. Nach dem Ausmultiplizieren und Aufsummieren folgt für den ersten Term:

$$\frac{1}{h^2} g_{\alpha k} < u_{k-1} - 2u_k + u_{k+1} >_V^0 = \frac{4}{h^2} \sin^2 \left(\frac{K_\alpha h}{2} \right) V_\alpha \quad (2.34)$$

Den zweiten Term berechnet man, indem man zuerst mit Hilfe des Wick'schen Theorems den bedingten Erwartungswert des Produktes $< u_k^3 u_k^{*2} >_V^0$ umschreibt:

$$< u_k^3 u_k^{*2} >_V^0 = < u_k >_V^3 < u_k^* >_V^2 + 6 < u_k >_V^2 < u_k^* >_V < u_k^* >_V + 6 < u_k >_V < u_k^* >_V \text{Cov}\{u_k, u_k^*\}_V + 6 < u_k >_V \text{Cov}\{u_k, u_k^*\}_V^2 \quad (2.35)$$

Dabei ist auf der rechten Seite dieser Gleichung der Exponent 0 der Lesbarkeit wegen weggelassen worden. Dennoch sind alle bedingten Erwartungswerte, bzw. Kovarianzen bezüglich des Gaußschen Maßes μ_0 zu verstehen.

Aus (2.17) und der Tatsache, daß $GG^\dagger = \frac{1}{n}I_N$ folgt:

$$\text{Cov}\{u_k, u_k^*\}_V^0 = [C - nG^\dagger GC]_{ii} = c \quad (2.36)$$

Man erhält schließlich

$$\begin{aligned} g_{\alpha k} < u_k^3 u_k^{*2} >_V^0 &= (-i) 6c^2 (-1) \delta_{\alpha, \beta} \cdot V_\beta + 6c (-1) (-i) \sum_{\beta, \gamma, \delta} \delta_{\alpha, \beta - \gamma + \delta} V_\beta V_\gamma^* V_\delta \\ &+ (-1) (-i) \sum_{\beta, \gamma, \delta, \lambda, \mu} \delta_{\alpha, \beta + \gamma + \delta - \lambda - \mu} V_\beta V_\gamma V_\delta V_\lambda^* V_\mu^* \end{aligned} \quad (2.37)$$

wobei hier die Einsteinsche Summenkonvention verwendet wurde.

Setzen wir die Ergebnisse (2.34) und (2.37) in (2.33) ein, folgt für die zeitliche Entwicklung der

Fourier-Koeffizienten:

$$\begin{aligned} \frac{dV_\alpha}{dt} = & -i \left[\frac{4}{h^2} \sin^2 \frac{K_\alpha h}{2} \right] \cdot V_\alpha + (-1)(-i) \sum_{\beta, \gamma, \delta, \lambda, \mu} \delta_{\alpha, \beta + \gamma + \delta - \lambda - \mu} V_\beta V_\gamma V_\delta V_\lambda^* V_\mu^* \\ & + 6c(-1)(-i) \sum_{\beta, \gamma, \delta} \delta_{\alpha, \beta - \gamma + \delta} V_\beta V_\gamma^* V_\delta + (-i) 6c^2(-1) \cdot V_\alpha \end{aligned} \quad (2.38)$$

An dieser Gleichung wird klar, was der Vorteil der "Optimal Prediction" Methode ist: die zwei ersten Termen auf der rechten Seite entsprechen einer Galerkin-Approximation, bei der nur die N Fourier-Koeffizienten berücksichtigt und die Amplituden aller anderen Fourier-Moden V_μ , $N < \mu \leq n$ gewöhnlich gleich Null gesetzt würden; die zwei letzten Terme sind aber Korrekturen, parametrisiert durch c , die über m_0 Kenntnisse über höhere Ordnungen im Hamiltonian und über die $(n - N)$ unberücksichtigten Fourier-Moden miteinbeziehen. Verallgemeinernd ist die Struktur der Korrekturterme interessant, die die "Optimal Prediction" Methode für eine beliebige Potenz der Nichtlinearität in der Gleichung (2.1) liefert. Sie wird im Anhang angeführt. Der Parameter $c(N)$ ist eine abnehmende Funktion von N :

$$c = \frac{1}{\pi} \sum_{k=1}^n \frac{1}{\frac{4}{h^2} \sin^2 \left(\frac{kh}{2} \right) + m_0^2} - \frac{1}{\pi} \sum_{\alpha=1}^N \frac{1}{\frac{4}{h^2} \sin^2 \left(\frac{K_\alpha h}{2} \right) + m_0^2}; \quad (2.39)$$

für $N \rightarrow n$, $c \rightarrow 0$ und die "Optimal Prediction" Methode wird zur geläufigen Galerkin-Approximation [19].

Es ist zu beachten, daß die Parameter m_0 und b , die die Partition des Hamiltonians festlegen, nur vom Intervall abhängig sind, auf dem die Bewegungsgleichung untersucht wird, und nicht von der Anzahl der Stützpunkte bei der Diskretisierung. Der Parameter c des Korrekturterms hängt aber von der Art dieser Diskretisierung ab.

Numerisch wurde nun zuerst mit einer einzigen Anfangsverteilung simuliert. Dies kann als Grenzfall einer Monte-Carlo-Simulation angesehen werden, bei der alle "gewürfelten" Konfigurationen verworfen werden, und eine einzige, und zwar die Anfangskonfiguration, beibehalten wird.

In den Abbildungen (2.2) und (2.1) einerseits, und (2.3) und (2.4) andererseits, wird die Zeitentwicklung des Real- und des Imaginärteils der zwei ersten beibehaltenen kollektiven Koordinaten V_1 und V_2 dargestellt. Die verschiedenen Kurven, die man für jede dieser Entwicklungen aus der Integration des geschlossenen Systems, einem Galerkin-Verfahren, und aus der Vorhersage der “Optimal Prediction” erhält, werden miteinander verglichen.

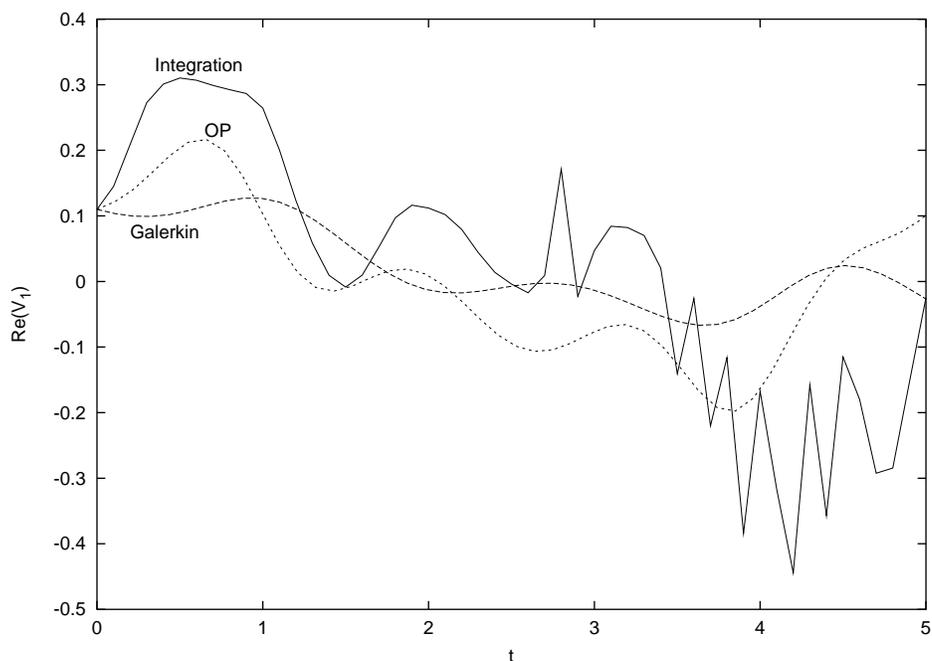


Abbildung 2.1: Zeitentwicklung des Realteils von V_1 : die verschiedenen Kurven sind jeweils über eine numerische Integration, ein Galerkin-Verfahren nach Abschneiden des geschlossenen Systems der Fourier-Koeffizienten auf die vier V_α , und die Anwendung der “Optimal Prediction” Methode 0. Ordnung erzielt worden. Die Galerkin-Approximation ist schlecht, während die “Optimal Prediction” Methode gute qualitative Ergebnisse liefert, allerdings nur bis zu einem gegebenen Zeitpunkt.

An diesen Abbildungen wird klar, daß die “Optimal Prediction” Methode auch bei einer quintischen Nichtlinearität sehr gute Ergebnisse für die Zeitentwicklung der relevanten kollektiven Koordinaten liefert, insbesondere deutlich bessere als ein Galerkin-Verfahren, das keinen Gebrauch von der Tatsache macht, daß die unberücksichtigten Koordinaten einer bestimmten vorgegebenen Statistik gehorchen.

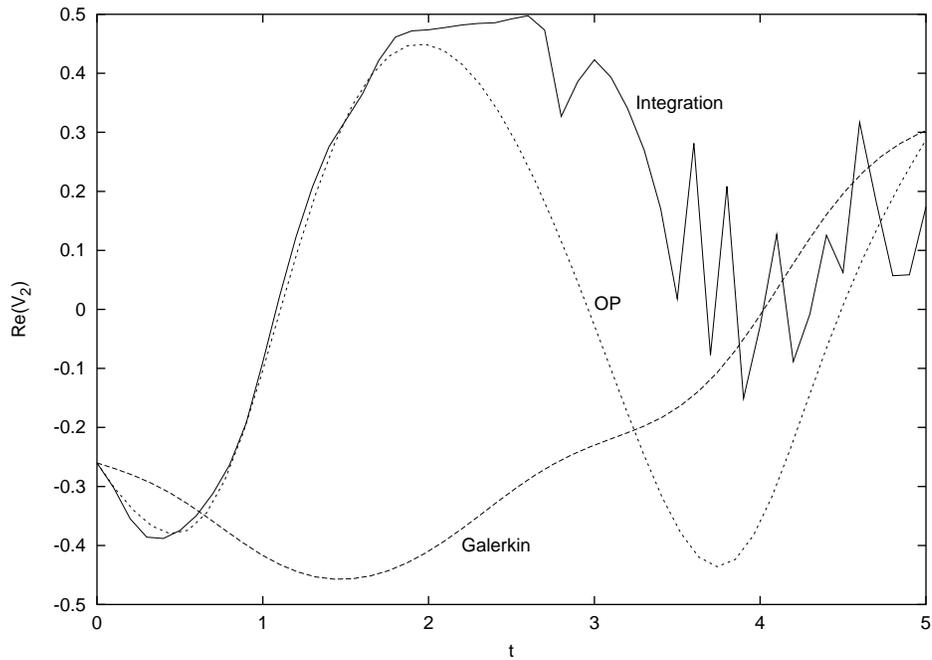


Abbildung 2.2: Zeitentwicklung des Realteils von V_2 : die verschiedenen Kurven sind jeweils über eine numerische Integration, ein Galerkin-Verfahren nach Abschneiden des geschlossenen Systems der Fourier-Koeffizienten auf die vier V_α , und die Anwendung der "Optimal Prediction" Methode 0. Ordnung erzielt worden. Die Galerkin-Approximation ist schlecht, während die "Optimal Prediction" Methode sehr gute Übereinstimmung mit der integrierten Kurve liefert, allerdings auch nur bis zu einem gegebenen Zeitpunkt.

Der qualitative Verlauf von $Re(V_1)$, $Im(V_1)$ und $Im(V_2)$ ist gut wiedergegeben und die Übereinstimmung ist bei $Re(V_2)$ sogar perfekt; allerdings sind diese Ergebnisse nur so gut, wenn oder solange, bis kein Kollaps eintritt.

Wir haben oben schon erwähnt, daß nämlich, wenn doch ein sogenannter "Blow-up" der Amplitude stattfindet, sich die Energie von den großskaligen Fourier-Moden mit $|n/2 - k|$ klein - unseren kollektiven Koordinaten V_α - auf die kleinen Skalen mit $|n/2 - k|$ größer verteilt.

Mit anderen Worten wird der Verlauf der Amplitude im Ortsraum steiler und es sind mehr Fourier-Moden nötig, um das System zu beschreiben. Dies wird an der Abbildung (2.5) deutlich: da wird der Realteil von $V_2 = F_{16}$ mit demjenigen von F_9 verglichen, wobei angemerkt werden soll, daß F_9 hier als Beispiel herausgegriffen worden ist, aber ähnliche Ergebnisse mit den anderen unberücksichtigten Fourier-Moden darstellbar sind, die auf die gleiche Zeitskala

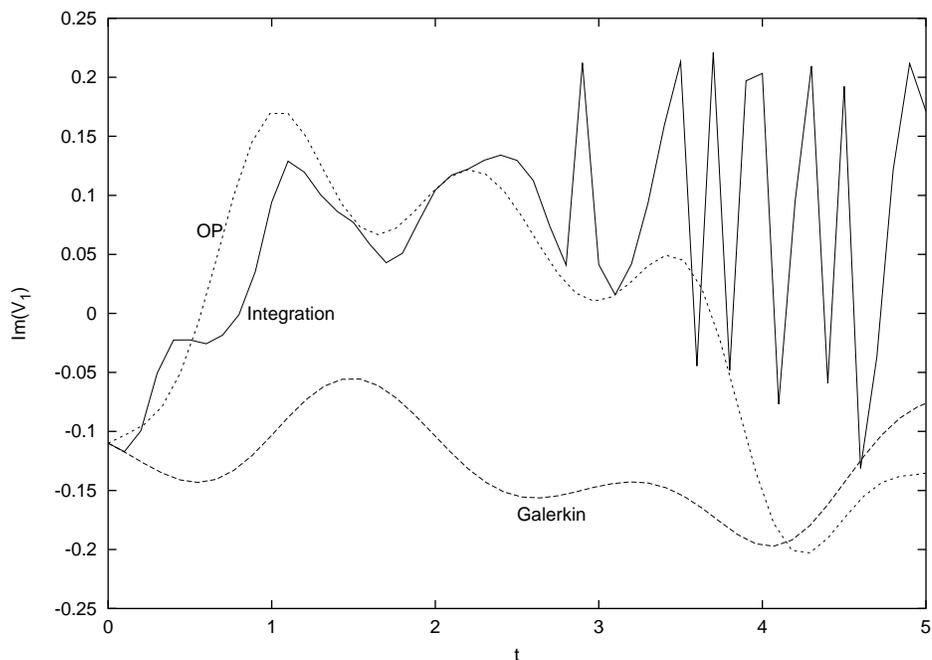


Abbildung 2.3: Zeitentwicklung des Imaginärteils von V_1 : die verschiedenen Kurven sind jeweils über eine numerische Integration, ein Galerkin-Verfahren nach Abschneiden des geschlossenen Systems der Fourier-Koeffizienten auf die vier V_α , und die Anwendung der “Optimal Prediction” Methode 0. Ordnung erzielt worden. Die Galerkin-Approximation ist schlecht, während die “Optimal Prediction” Methode gute qualitative Ergebnisse liefert, allerdings nur bis zu einem gegebenen Zeitpunkt.

hinweisen. Solange F_9 unbedeutend und gegenüber V_2 , d. h. gegenüber den kollektiven Koordinaten vernachlässigbar ist, sind die von der “Optimal Prediction” Methode gelieferten Voraussagen sehr gut; sobald aber die Bedeutung der unberücksichtigten Moden nicht mehr vernachlässigt werden kann, weil die Amplitude im Ortsraum steil anwächst, ist es nicht mehr der Fall. Somit liefert auch die “Optimal Prediction” Methode *a contrario* eine untere Abschätzung für die Zeit des Kollaps: solange sie gute Ergebnisse liefert, kann die Amplitude des Signals noch nicht kollabiert haben.

Bis zu der Zeit $t_c \approx 2.7$ eignen sich also die gewählten kollektiven Koordinaten sehr gut: sie geben tatsächlich das großskalige Verhalten des Systems wieder. Für $t > t_c$ allerdings sind sie nicht mehr unbedingt allein ausschlaggebend für das Verhalten des Systems, bzw. sind mehr Fourier-Koeffizienten nötig.

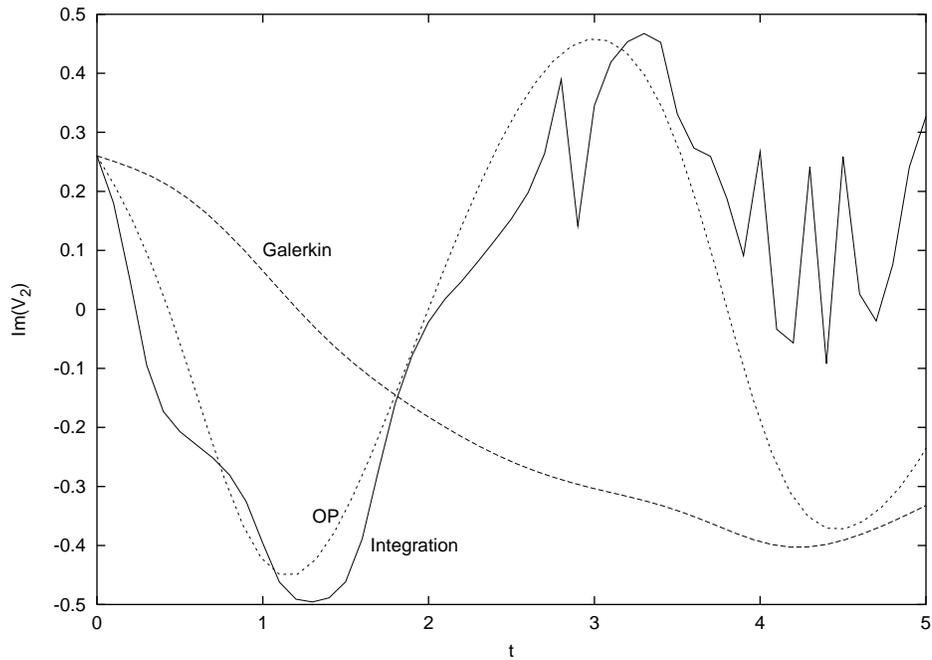


Abbildung 2.4: Zeitentwicklung des Imaginärteils von V_2 : die verschiedenen Kurven sind jeweils über eine numerische Integration, ein Galerkin-Verfahren nach Abschneiden des geschlossenen Systems der Fourier-Koeffizienten auf die vier V_α , und die Anwendung der "Optimal Prediction" Methode 0. Ordnung erzielt worden. Die Galerkin-Approximation ist schlecht, während die "Optimal Prediction" Methode gute qualitative Ergebnisse liefert, allerdings nur bis zu einem gegebenen Zeitpunkt.

Erwähnenswert ist zudem die Tatsache, daß diese Ergebnisse auch bei einer Monte-Carlo-Simulation über hundert Anfangsbedingungen wiederholt und bestätigt worden sind, wenn auch die Übereinstimmung zwischen den Ergebnissen der Integration und denjenigen der "Optimal Prediction" nicht so gut ist, wie im Fall einer einzelnen Trajektorie. Dennoch ist das dynamische Verhalten so, besser als mit einem Galerkin-Verfahren, qualitativ sehr gut approximiert, und die eben gemachten Schlußfolgerungen behalten ihre Gültigkeit.

Dies legt die Vermutung nahe, daß die "Optimal Prediction" Methode nur dann gute Ergebnisse liefert, wenn

- die gewählten kollektiven Koordinaten das großskalige Verhalten des Systems richtig wiedergeben

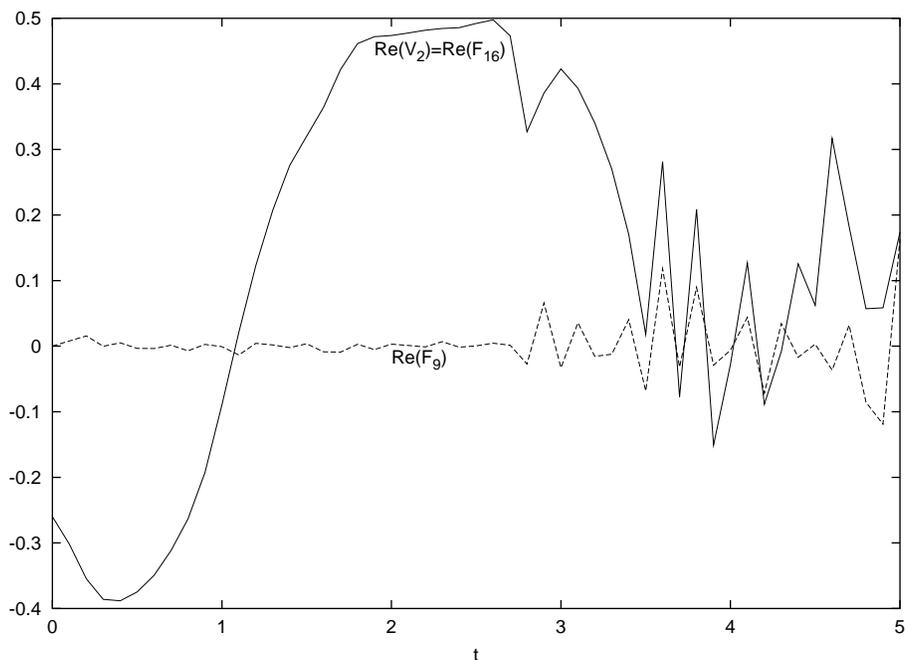


Abbildung 2.5: Vergleich der Zeitentwicklung des Realteils von $V_2 = F_{16}$ und F_9 : es ist deutlich zu erkennen, daß solange $t \leq t_c$, mit $t_c \approx 2.7$, F_9 gegenüber V_2 vernachlässigt werden kann; für $t \geq t_c$ allerdings werden die “kleinen” Skalen um so wichtiger und können nicht mehr unberücksichtigt bleiben.

- und sich die Anzahl der zu dieser großskaligen Beschreibung nötigen Koordinaten nicht mit der Zeit ändert.

Dieses wichtige Ergebnis stellt eine wesentliche Einschränkung des Anwendungsgebiets der “Optimal Prediction” Methode dar.

Gleichzeitig wirft es ein neues Licht auf die von Chorin *et al.* unter anderem in [19] präsentierten Ergebnisse der Anwendung der Methode auf die kubisch nichtlineare Schrödinger-Gleichung. Tatsächlich ist die Frage interessant, ob nun die “Optimal Prediction”, angewandt auf die kubisch nichtlineare Schrödinger-Gleichung, immer gute Ergebnisse liefert, oder ob deren Erfolg etwa von der Festlegung der kollektiven Koordinaten abhängt.

Wir wollen nun im nächsten Abschnitt versuchen, diese Frage zu beantworten.

2.4. Einfluß der Anfangsverteilung: Untersuchung der kubisch nichtlinearen Schrödinger Gleichung

Die Interpretation der Ergebnisse der Anwendung der “Optimal Prediction” auf die quintisch nichtlineare Schrödinger-Gleichung legt die Vermutung nahe, daß das Gelingen dieser Methode, also ihre Verwendbarkeit stark von der mittleren Anfangsverteilung und der Wahl der als “relevant” eingestuften kollektiven Koordinaten abhängt.

Diese Vermutung prüfen wir nun anhand von zwei Beispielen mit deutlich voneinander verschiedenen mittleren Anfangsverteilungen.

Dabei gehen wir von der kubisch nichtlinearen Schrödinger-Gleichung

$$u_t = -u_{xx} + (-2)|u|^2u, \quad (2.40)$$

für eine komplexe Variable $u(x) = q(x) + ip(x)$ auf dem x -Intervall $[0, 2\pi]$ mit periodischen Randbedingungen aus [19].

Diese Gleichung legt über den zugehörigen Hamiltonian

$$H(q, p) = \frac{1}{2} \int_0^{2\pi} \left((q_x^2 + p_x^2) + (-2) \frac{1}{3} (q^2 + p^2)^2 \right) dT \quad (2.41)$$

die zugrundeliegende Statistik fest. Für einen gegebenen Hamiltonian ist auch dessen gewählte Approximation durch einen Gaußschen Hamiltonian nach der Partition (2.19) fest und unabhängig von der spezifischen untersuchten Anfangsverteilung. Je nachdem, wie die kollektiven Koordinaten gewählt und die Monte-Carlo-Simulation durchgeführt werden, können aber die verschiedenen möglichen mittleren Anfangsverteilungen, die zu einem solchen Problem gehören, ziemlich unterschiedlich voneinander ausfallen.

Zuerst wird die Gleichung (2.40) mit den gleichen Stützstellen wie vorhin diskretisiert und das dazugehörige System von 32 Differentialgleichungen über eine diskrete Fourier-Transformation auf ein System von 32 Gleichungen für die F_k gebracht. Dabei übernehmen wir die Notationen der vorigen Abschnitte.

Statt das vollständige System geschlossen zu lösen, beschränken wir uns hier auch auf die 4

kollektiven Koordinaten V_1, \dots, V_4 :

$$\begin{cases} V_1 = F_{n/2-N/2+1}, \\ V_2 = F_{n/2-N/2+2}, \\ V_3 = F_{n/2-N/2+3}, \\ V_4 = F_{n/2-N/2+4}. \end{cases} \quad (2.42)$$

Die Bestimmung der der Gleichung (1.13) entsprechenden Zeitentwicklung der kollektiven Koordinaten erfolgt ähnlich wie bei der quintisch nichtlinearen Schrödinger-Gleichung und wird hier nicht weiter ausgeführt. Die Berechnungen, die zur sogenannten “optimalen” Partition des Hamiltonians führen, werden im Anhang dargestellt. Schließlich erhält man für die Zeitentwicklung der V_α folgende Gleichung:

$$i \frac{dV_\alpha}{dt} = \frac{4}{\hbar^2} \sin^2 \left(\frac{K_\alpha \hbar}{2} \right) V_\alpha + (-2) \sum_{\beta, \gamma, \varepsilon} \delta_{\alpha, \beta - \gamma + \varepsilon} V_\beta V_\gamma^* V_\varepsilon + 2(-2)cV_\alpha \quad (2.43)$$

Hier erkennt man, wie am Beispiel der Schrödinger-Gleichung mit Linearität $p = 4$, die zwei ersten Terme auf der rechten Seite dieser Gleichung als den Galerkin-Beitrag, während der letzte Term die von c parametrisierte Korrektur darstellt. Dabei ist c die in (2.39) definierte Funktion von N .

Unsere zwei Beispiele unterscheiden sich nun durch die Bedeutung der festgehaltenen Fourier-Koeffizienten und die Form der mittleren Anfangsverteilung, die aus der Monte-Carlo-Simulation resultiert.

Untersuchen wir in einem ersten Schritt die aus den folgenden Werten ihrer Fourier-Moden im Fourier-Raum bestimmte Konfiguration: die vier komplexen kollektiven Koordinaten seien

$$\begin{cases} V_1 = F_{15} = (0.01, -0.01) \\ V_2 = F_{16} = (-0.1, 0.1) \\ V_3 = F_{17} = (0.41, -0.41) \\ V_4 = F_{18} = (0.21, -0.21). \end{cases} \quad (2.44)$$

Die Werte der anderen Fourier-Koeffizienten werden zwar nicht festgehalten, dennoch seien in der Monte-Carlo-Simulation die Werte von F_{14} und V_{19} der Größenordnung 10^{-1} und diejenigen aller anderen Koeffizienten der Größenordnung 10^{-6} oder kleiner angesetzt.

Nach einer Monte-Carlo-Simulation über 100 Anfangskonfigurationen zeichnet sich die mittlere Anfangsverteilung durch folgende Eigenschaft aus: die kollektiven Koordinaten haben die eben in (2.44) dargestellten Werte; die Koeffizienten F_{14} und F_{19} nehmen die Werte

$$\begin{cases} F_{14} = (0.1678, 0.321) \\ F_{19} = (-0.147, 0.213) \end{cases} \quad (2.45)$$

an. Alle anderen Fourier-Koeffizienten haben Werte der Größenordnung 10^{-6} oder kleiner und sind damit deutlich kleiner.

Interessant ist dabei, sich den Beitrag der kollektiven Koordinaten zu der gesamten mittleren Anfangsverteilung anzuschauen. Setzt man alle Fourier-Koeffizienten gleich Null außer die vier kollektiven Koordinaten und führt die Rücktransformation $u_i = \sum_{k=1}^n e^{iK_k x_i} V_k$ aus, sieht man, daß die vier niedrigsten Fourier-Koeffizienten zusammen das großskalige Verhalten des Systems gut wiedergeben, und sich somit gut als “kollektive Koordinaten” im Sinne der in der Einleitung dieser Arbeit gegebenen Definition eignen. Deren Anteil an der Gesamtverteilung wird in den Abbildungen (2.6) und (2.7) dargestellt.

Und tatsächlich liefert die 0. Ordnung “Optimal Prediction” in diesem Falle sehr gute Ergebnisse für die Zeitentwicklung der kollektiven Variablen V_2, V_3 und V_4 , ähnlich den Ergebnissen von Chorin *et al.* in [19], wie an den Abbildungen (2.8) und (2.9) ersichtlich ist. Die Symbole stellen die mittleren Werte dar, die aus der Integration der einzelnen hundert Anfangsbedingungen der Monte-Carlo-Simulation folgen; die Voraussagen der “Optimal Prediction” Methode stimmen sehr gut mit denjenigen der Integration überein, während sich die Galerkin-Approximation teilweise schon nach sehr kurzen Zeiten als fehlerhaft erweist.

Was V_1 angeht, ist das Ergebnis allerdings schlecht: sowohl die Galerkin-Approximation, als auch die “Optimal Prediction” Methode liefern keine gute Voraussage für die weitere Zeitentwicklung; dies wird auf die Tatsache zurückgeführt, daß V_1 , bzw. F_{15} , keine gute “kollektive”

Abbildung 2.6: Anteil der vier kollektiven Koordinaten an dem Realteil q der Anfangsverteilung: die dicke durchgezogene Kurve ist das Ergebnis der kompletten diskreten Fourier-Transformation, während die dünne gestrichelte Kurve den Anteil der vier kollektiven Koordinaten darstellt.

Abbildung 2.7: Anteil der vier kollektiven Koordinaten an dem Imaginärteil p der Anfangsverteilung: die dicke durchgezogene Kurve ist das Ergebnis der kompletten diskreten Fourier-Transformation, während die dünne gestrichelte Kurve den Anteil der vier kollektiven Koordinaten darstellt.

Abbildung 2.8: Zeitentwicklung vom Realteil von V_2 , V_3 und V_4 ; die Symbole sind das Ergebnis der Integration über hundert Anfangskonfigurationen; sie stimmen sehr gut mit den Ergebnissen der “Optimal Prediction” Methode (dünne Linien) überein, während die Galerkin-Approximation (dickere Linien) teilweise sehr fehlerhaft ist.

Koordinate darstellt, da die Beiträge von V_2 , V_3 und V_4 , sowie diejenigen von F_{14} und F_{19} , die dazu noch unberücksichtigt bleiben, deutlich größer sind. In einem solchen Fall versagt die “Optimal Prediction” Methode und ist nicht besser, als ein Galerkin-Verfahren. Daß letzteres in dem Falle schlechte Ergebnisse produziert, ist klar, wenn man bedenkt, daß die im Vergleich zu V_1 wichtigen Werte von F_{14} und F_{19} dabei einfach gleich Null gesetzt werden.

Die Vermutung, nach der die “Optimal Prediction” Methode in 0. Ordnung nur dann gute Ergebnisse liefern kann, wenn die kollektiven Koordinaten auch diejenigen sind, die das großskalige Verhalten des dynamischen Systems wiedergeben, oder anders ausgedrückt, wenn der Großteil der Energie des Systems sich in den ausgesuchten Moden konzentriert, wird von unserem zweiten Beispiel bekräftigt.

Wäre allerdings die Energie des Systems größtenteils ausschließlich in den vier kollektiven

Abbildung 2.9: Zeitentwicklung vom Imaginärteil von V_2 , V_3 und V_4 ; die Symbole sind das Ergebnis der Integration über hundert Anfangskonfigurationen; sie stimmen sehr gut mit den Ergebnissen der “Optimal Prediction” Methode (dünne Linien) überein, während die Galerkin-Approximation (dickere Linien) teilweise sehr fehlerhaft ist.

Koordinaten konzentriert, wäre ein Galerkin-Verfahren bei der Voraussage deren weiterer Zeitentwicklung natürlich sehr erfolgreich und der Vorteil der “Optimal Prediction” nicht einsichtig. Somit sollen, damit diese letzte Methode sinnvoll eingesetzt werden kann und tatsächliche Fortschritte gegenüber gängigen Methoden liefert, erstens wenige Moden reichen, um das System zu charakterisieren, zweitens allerdings einige davon, hier beispielhaft F_{14} und F_{19} unberücksichtigt bleiben, so daß deren Einfluß über den vom Parameter c abhängigen Zusatzterm auf der rechten Seite der Gleichung (2.43) erfaßt wird.

In unserem zweiten Beispiel untersuchen wir eine mittlere Anfangsbedingung, zu deren Beschreibung wir wissen, daß viele Moden F_k nötig sind. Dies geschieht, indem wir als Anfangskonfiguration für unsere Monte-Carlo Simulation eine im Ortsraum solitonartige Lösung an-

setzen. Über die oben beschriebene Diskretisierung und anschließende Bildung der Modenkoeffizienten F_k gelangt man wieder zu dem unterbestimmten System, wenn man von den zur geschlossenenen Beschreibung nötigen 32 F_k nur die vier V_α beibehält. Diese werden nun wie vorhin fixiert, während eine Monte-Carlo Simulation 100 Anfangsverteilungen mit denselben kollektiven Koordinaten zur Zeit $t = 0$ liefert. Da aber diese vier Moden völlig unzureichend sind, um das großskalige Verhalten des Systems zu beschreiben, erwarten wir nun schlechte Voraussagen der “Optimal Prediction” Methode. Das ist tatsächlich der Fall, wie nun beispielhaft an dem Realteil von V_2 deutlich wird (cf. dazu Abbildung (2.10)). Bei den anderen kollektiven Koordinaten V_1 , V_3 und V_4 ist das Ergebnis vergleichbar. An der Abbildung (2.10) wird klar, daß

Abbildung 2.10: Zeitentwicklung vom Realteil von V_2 ; sowohl das Galerkin-Verfahren, wie auch die “Optimal Prediction” Methode liefern keine gute Voraussage. Zudem zeichnet sich eine deutliche Verschlechterung der Ergebnisse der “Optimal Prediction” gegenüber einem gängigen Galerkin-Verfahren ab, das rigoros nach den vier kollektiven Koordinaten abschneidet.

die “Optimal Prediction” Methode nicht nur die weitere Zeitentwicklung der kollektiven Koordinaten nicht richtig beschreibt, sondern sogar deutlich schlechtere Ergebnisse als ein Galerkin-

Verfahren liefert. Dies kann man verstehen, indem man sich die Gleichung (2.43) in Erinnerung ruft. Der einzige Unterschied zwischen der Galerkin-Approximation und der “Optimal Prediction” Methode besteht in dem von c abhängigen Zusatzterm, von dem erwartet wird, daß er die Rückwirkung der unberücksichtigten Moden auf die kollektiven Koordinaten ausdrückt. Der Parameter c ist in (2.36) definiert und, wie aus dem Lemma (A.6) einsichtig, unabhängig von den speziellen Anfangswerten $V_\alpha(t=0)$ der kollektiven Koordinaten und alleine von der Matrix G , d. h. von den Filterfunktionen $g_\alpha(x)$ abhängig. Mit diesen Filterfunktionen aber wird festgelegt, welche und vor allen Dingen auch wieviele spezielle Moden (wenn auch ihr spezifischer Wert keine Rolle spielt) erwartungsgemäß für die weitere dynamische Entwicklung des Systems ausschlaggebend sind. Da in unserem Falle die Einschränkung auf die vier beibehaltenen Moden dieser Erwartung nicht gerecht wird, induziert der “ c -Term” der “Optimal Prediction” sogar einen Fehler, der dafür verantwortlich ist, daß das Galerkin-Verfahren in diesem Falle bessere Ergebnisse liefert. Vermutlich würde eine Anwendung der “Optimal Prediction” Methode mit “besseren”, unserer Verteilung eher entsprechenden kollektiven Koordinaten, als den Fourier-Moden, bessere Ergebnisse liefern.

An diesen zwei Beispielen wird die Bedeutung der Anfangsverteilung für das Gelingen der “Optimal Prediction” Methode klar.

In Zusammenfassung der Ergebnisse der Anwendung der Methode im vorigen Abschnitt auf die quintisch und in diesem Abschnitt auf die kubisch nichtlineare Schrödinger-Gleichung kommen wir zu dem Schluß, daß folgende Bedingungen erfüllt werden müssen, damit die “Optimal Prediction” sinnvoll eingesetzt werden kann und gute Ergebnisse liefert:

- die Energie des Systems muß sich hauptsächlich in wenigen Moden konzentrieren,
- die kollektiven Koordinaten müssen unter diesen energiereichen Moden ausgesucht werden,
- die besten Ergebnisse liefert die “Optimal Prediction” Methode, und vor allen Dingen

bessere als ein Galerkin-Verfahren, wenn nur wenige energiereiche Moden unberücksichtigt bleiben, deren Einfluß auf die kollektiven Koordinaten aber über den “ c -Zusatzterm” einbezogen wird,

- die Zahl und die Natur der energiereichen Moden muß in der Zeit konstant bleiben.

Diese letzte Bedingung ist der Aussage äquivalent, daß sobald ein Energietransfer zwischen den berücksichtigten und den unberücksichtigten Moden im Laufe der Zeit stattfindet, wie es der Fall bei kollabierenden Lösungen der nichtlinearen Schrödinger-Gleichung ist, die Methode der “Optimal Prediction” ihrem Namen nicht gerecht wird und ungeeignet ist.

3. Diskretes Modell

In diesem Paragraphen wollen wir nun die Möglichkeit der Anwendung der “Optimal Prediction” Methode auf das Hamiltonsche System der sogenannten diskreten “self-trapping” Gleichung (DST) untersuchen [26], die in der Beschreibung der Physik der kondensierten Materie oft vorkommt.

Der sogenannte “self-trapping” (zu deutsch: “selbst-einfangend”) Charakter basiert auf dem Zusammenspiel von zwei verschiedenen Prozessen; einerseits bewirkt die Lokalisierung eines Integrals der Bewegung (Energie) strukturelle Veränderungen ihrer Umgebung; andererseits reduzieren diese strukturellen Veränderungen die Energie des Integrals der Bewegung und vermeiden somit ihre Dispersion.

Die DST modelliert ein System von gekoppelten klassischen oder quantenmechanischen anharmonischen Oszillatoren [25] und lautet:

$$\left(i \frac{d}{dt} - w_0 \right) \mathbf{A} + \gamma \mathbf{D}(|\mathbf{A}|^2) \mathbf{A} + \varepsilon \mathbf{M} \mathbf{A} = 0 \quad (3.1)$$

Dabei ist \mathbf{A} der n -dimensionale Vektor der komplexen Modenamplituden der n miteinander wechselwirkenden anharmonischen Freiheitsgrade [25].

Die Diagonalmatrix \mathbf{D}

$$\mathbf{D}(|\mathbf{A}|^2) = \begin{pmatrix} |A_1|^2 & \dots & 0 \\ & |A_2|^2 & \dots \\ & & \ddots \\ 0 & \dots & |A_n|^2 \end{pmatrix} \quad (3.2)$$

modelliert den anharmonischen Charakter der Oszillatoren; die Matrix \mathbf{M} wird als reell und symmetrisch angenommen und enthält die linearen dispersiven Wechselwirkungsterme zwischen den n Freiheitsgraden.

Für $\varepsilon = 0$ und $\gamma \neq 0$ stellt die Gleichung (3.1) ein System von nicht miteinander wechselwirkenden anharmonischen Oszillatoren dar, während aus (3.1) für $\gamma = 0$ und $\varepsilon \neq 0$ das Modell eines

einfachen Systems von miteinander gekoppelten linearen Oszillatoren folgt [25].

Es können verschiedene physikalische Motivationen für die Untersuchung der DST angeführt werden:

- erstens vermutete Davydov in den 1970er Jahren, daß die DST den “selbsteingefangenen” Charakter von Vibrationen bestimmter Peptidgruppen entlang der α -Spirale eines Proteins in Wechselwirkung mit niederfrequenten Phononen beschreibt [21, 20]. Bei diesem System handelt es sich um eine Exziton-Phononen-Wechselwirkung, deren Modellierung unter bestimmten Voraussetzungen (adiabatische Approximation für die optischen Phononen und ausschließlich Wechselwirkungen mit den nächsten Nachbarn) auf die DST führt. Die Elemente des Vektors \mathbf{A} sind in diesem Falle die Quantenzahlen der sogenannten Davydov-Moden [25] und die Elemente der Matrix \mathbf{M} Dipol-Dipol-Wechselwirkungsenergien. Experimentelle Bestätigung dafür lieferten später Untersuchungen an Kristallen anhand von Infrarot-Absorption und Raman-Streuung [26]. Eine Untersuchung der durch die quantisierte DST beschriebenen lokalisierten Vibrationsenergie in kleinen polyatomaren Molekülen findet sich in [4].
- zweitens liefert die DST ein klassisches Modell für bestimmte Aspekte der dynamischen Wechselwirkung anharmonischer Moden in kleinen Molekülen wie Wasser, Ammoniak, Methan oder Benzol. Die Elemente von \mathbf{A} sind dann die komplexen Amplituden einer bestimmten molekularen Bindung; ist die komplexe Mode durch die klassischen Koordinaten (q, p) beschrieben, dann sind die A_k durch $A_k = (p_k - iw_0q_k)/\sqrt{2}$ gegeben. Eine Entwicklung des Hamiltonians des Systems in Potenzen von A_k führt bei Vernachlässigung aller nichtlinearer Terme bis auf die Selbstwechselwirkungsterme der Form $|A_k|^4$ dann auf die DST. Insbesondere werden dabei Terme der Form A_k^4 und A_k^{*4} vernachlässigt, was eine weitere Erhaltungsgröße des Systems (die unten definierte Zahl N) ergibt. γ stellt dabei die Aufweichung dieser Bindungen bei Ausdehnungen dar und $\epsilon\mathbf{M}$ die elektromagnetischen und mechanischen Kopplungen zwischen diesen Bindungen [25, 26].

- weitere Beispiele von solchem “self-trapping” sind die Bewegung eines Elektrons in einem Kristallgitter - das lokalisierte Elektron polarisiert den Kristall, während dieser letzte die Energie des Elektrons mindert-, elektromagnetische Energie in einem Plasma oder die Lichtenergie in einer optischen Glasfaser [26].

Aus der unitären Transformation $\mathbf{A} \rightarrow \mathbf{A}e^{-i w_0 t}$ folgt die Standardform der DST:

$$i\dot{\mathbf{A}} + \gamma \mathbf{D}(|\mathbf{A}|^2)\mathbf{A} + \varepsilon \mathbf{M}\mathbf{A} = 0 \quad (3.3)$$

Da jeder konstante Diagonalterm in \mathbf{M} in w_0 miteinbezogen werden kann, ist es vernünftig, zu vermuten, daß, wenn die Moden gleiche Umgebung haben, die Diagonalelemente von \mathbf{M} effektiv gleich Null gesetzt werden können [25]. Dazu kommt, daß das lineare System, das aus (3.1) mit $\gamma = 0$ und $\varepsilon \neq 0$ folgt, schon von Anderson (cf.[26] und Zitate hierin) mit dem Ergebnis untersucht wurde, daß lokalisierte Zustände dann auftreten, wenn die Diagonalelemente von \mathbf{M} größer sind, als der größte der nichtdiagonalen Elemente. Wir wollen uns im folgenden aber auf lokalisierte Zustände konzentrieren, die von der Wirkung der Nichtlinearität herrühren. Daher werden alle Diagonalelemente von \mathbf{M} gleich Null gesetzt [26].

Für eine detaillierte Untersuchung der DST wird auf [26] verwiesen. Wir weisen hier nur auf wichtige Ergebnisse hin. Lösungen von (3.1) besitzen zwei Erhaltungsgrößen: die sogenannte “Zahl” $N = \sum_{i=1}^n |A_i|^2$ und die Energie

$$H = w_0 N - \frac{1}{2} \gamma \sum_{i=1}^n |A_i|^4 - \varepsilon \sum_{i \neq j} m_{ij} A_i^* A_j$$

Üblicherweise wird $N = 1$ durch Skalierung von A und γ erzielt.

Wichtige Lösungen von (3.3) sind die sogenannten stationären Lösungen, die mit

$$\mathbf{A}(t) = \phi e^{i w t},$$

definiert werden können, wobei ϕ ein n -dimensionaler, zeitunabhängiger Vektor ist; für diese Lösungen ist $|A_i(t)|^2$ zeitunabhängig. Von diesen Lösungen erwartet man, daß sie bei Wechselwirkungen des Systems mit elektromagnetischen Wellen wichtig sind [25, 26]. Zudem erfüllen

diese Lösungen die nichtlineare Eigenwertgleichung:

$$-w\phi + \gamma \mathbf{D}(|\phi|^2)\phi + \varepsilon \mathbf{M}\phi = 0$$

Für $\gamma \gg 1$ kann der ε -Term vernachlässigt werden, die Lösungen erfüllen $w\phi_i = \gamma|\phi_i|^2\phi_i$, woraus für reelle γ -Werte $\phi_i = 0, \pm\sqrt{w/\gamma}$ folgt. Diese Lösungen werden in [25, 26] untersucht.

Im folgenden wollen wir nun die DST im Spezialfall eines 10-dimensionalen Vektors $\mathbf{A} \in \mathbf{R}^{10}$ untersuchen.

Inwiefern können spezielle Komponenten dieses Vektors als kollektive Koordinaten für das System ausgesucht werden und deren Zeitentwicklung durch die “Optimal Prediction” Methode wiedergegeben werden?

Die Matrix \mathbf{M} definieren wir nun mit

$$\mathbf{M} = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & & & & & \\ 1 & 0 & 1/2 & 0 & \dots & & & & & \\ 0 & 1/2 & 0 & 1 & 0 & \dots & & & & \\ 0 & 0 & 1 & 0 & 1/2 & 0 & & & & \\ 0 & 0 & 0 & 1/2 & 0 & 1 & & & & \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & \dots & \\ \dots & & & & & & \dots & & & \\ 1 & 0 & 0 & 0 & \dots & & & & 1 & 0 \end{pmatrix} \quad (3.4)$$

Diese Form der Dispersionsmatrix steht für Wechselwirkungen ausschließlich mit den nächstliegenden Strukturen. Außerdem sind die Rückkoppelung der 3. und 4. Strukturen jeweils mit den anderen nur halb so stark wie ihre Wechselwirkung untereinander.

Wären alle Elemente der Nebendiagonalen gleich Eins, erhielte man aus der Gleichung (3.3) und der so definierten Dispersion eine diskrete Form der nichtlinearen Schrödinger-Gleichung mit periodischen Randbedingungen. Unsere Wahl zielt aber auf die Untersuchung eines physikalischen Systems, bei dem die größten Anregungen in den mittleren dritten und vierten Komponenten steckt und alle anderen als “Störungen”, als “Rauschen” betrachtet werden können;

dies erzielt man, indem die numerischen Anfangswerte für diese Wahrscheinlichkeitsamplituden A_i , $i \neq 3, 4$ von kleinerer Größenordnung gewählt werden. Im folgenden bezeichnen wir die Amplituden A_3 und A_4 , die wir als kollektive Koordinaten aussuchen werden, als "Hauptamplituden", um sie von den anderen "rauschenden" Amplituden hervorzuheben.

Die Zerlegung der komplexen Komponenten A_i des Vektors \mathbf{A} in Real- und Imaginärteile q_i, p_i führt auf folgende Differentialgleichungen:

Für $i = 1, 6, \dots, 10$:

$$\begin{cases} \dot{q}_i = \frac{\partial H}{\partial p_i} = -\gamma p_i (q_i^2 + p_i^2) - \varepsilon (p_{i+1} + p_{i-1}) \\ \dot{p}_i = -\frac{\partial H}{\partial q_i} = \gamma q_i (q_i^2 + p_i^2) + \varepsilon (q_{i+1} + q_{i-1}) \end{cases}$$

Für $i = 2$:

$$\begin{cases} \dot{q}_2 = -\gamma p_2 (q_2^2 + p_2^2) - \frac{1}{2}\varepsilon (2p_1 + p_3) \\ \dot{p}_2 = \gamma q_2 (q_2^2 + p_2^2) + \frac{1}{2}\varepsilon (2q_1 + q_3) \end{cases}$$

Für $i = 5$:

$$\begin{cases} \dot{q}_5 = -\gamma p_5 (q_5^2 + p_5^2) - \frac{1}{2}\varepsilon (2p_6 + p_4) \\ \dot{p}_5 = \gamma q_5 (q_5^2 + p_5^2) + \frac{1}{2}\varepsilon (2q_6 + q_4) \end{cases}$$

Und schließlich lauten die Differentialgleichungen für die "Hauptamplituden" (q_3, p_3) und (q_4, p_4) :

$$\begin{cases} \dot{q}_3 = -\gamma p_3 (q_3^2 + p_3^2) - \varepsilon p_4 - \frac{1}{2}\varepsilon p_2 \\ \dot{q}_4 = -\gamma p_4 (q_4^2 + p_4^2) - \varepsilon p_3 - \frac{1}{2}\varepsilon p_5 \\ \dot{p}_3 = +\gamma q_3 (q_3^2 + p_3^2) + \varepsilon q_4 + \frac{1}{2}\varepsilon q_2 \\ \dot{p}_4 = +\gamma q_4 (q_4^2 + p_4^2) + \varepsilon q_3 + \frac{1}{2}\varepsilon q_5 \end{cases} \quad (3.5)$$

3.1. 0. Ordnung "Optimal Prediction"

Die Hamiltonsche Struktur dieses Systems ist bekannt; der Hamiltonian ist durch

$$H = \frac{1}{2} \left[-\frac{1}{2}\gamma \sum_{i=1}^{10} (q_i^2 + p_i^2)^2 \right] - \varepsilon (p_3 p_4 + q_3 q_4) - \frac{1}{2} \varepsilon \sum_{k=1, k \neq 3, k \neq 4}^{10} [p_k (p_{k+1} + p_{k-1}) + q_k (q_{k+1} + q_{k-1})] \quad (3.6)$$

gegeben und wir nehmen von der Verteilung der möglichen Lösungen an, daß sie die vom Hamiltonian kanonisch induzierte Statistik erfüllt. Somit ist das “prior” Maß μ festgelegt, das für die Durchführung der “Optimal Prediction” nötig ist. Wir folgen jetzt deren in der Einleitung dargestellten Prinzip, indem wir als kollektive Koordinaten die “Hauptamplituden” (q_3, p_3) und (q_4, p_4) wählen; die Schlußfolgerungen des letzten Paragraphen legen nahe, daß diese Wahl wahrscheinlich eine gute ist, wenn diese Amplituden tatsächlich das großskalige Verhalten des Systems wiedergeben. Dafür wird folgender Satz V an Anfangsbedingungen für die kollektiven Koordinaten festgelegt

$$\left\{ \begin{array}{l} q_3 = -0.1 \\ q_4 = 0.24 \\ p_3 = -0.87 \\ p_4 = 0.12 \end{array} \right. \quad (3.7)$$

für $t = 0$ und $\varepsilon = \gamma = 1$. Über eine Monte-Carlo-Simulation wird ein Satz von 100 Anfangsbedingungen bestimmt, die nach der Wahrscheinlichkeitsdichte $f(\mu)$ verteilt werden, und zur Zeit $t = 0$ diese Werte für (q_3, p_3) und (q_4, p_4) annehmen. Dabei werden die anderen Amplituden bewußt eine Größenordnung kleiner angesetzt; so wird der Tatsache Rechnung getragen, daß diese als “Rauschen” betrachtet werden sollen.

Ein einfaches Galerkin-Verfahren würde darin bestehen, alle Amplituden gleich Null zu setzen und nur die Hauptamplituden zu behalten; das ursprüngliche System (3.5) von Differentialgleichungen für q_3, p_3, q_4, p_4 wird auf folgendes umgeschrieben:

$$\left\{ \begin{array}{l} \dot{q}_3 = -\gamma p_3(q_3^2 + p_3^2) - \varepsilon p_4 \\ \dot{q}_4 = -\gamma p_4(q_4^2 + p_4^2) - \varepsilon p_3 \\ \dot{p}_3 = +\gamma q_3(q_3^2 + p_3^2) + \varepsilon q_4 \\ \dot{p}_4 = +\gamma q_4(q_4^2 + p_4^2) + \varepsilon q_3 \end{array} \right. \quad (3.8)$$

Wegen der bei diesem Verfahren kompletten Vernachlässigung der Kopplungsterme zwischen den Hauptamplituden und den anderen kann aber diese Methode keine gute Vorhersage liefern. Es geht schließlich darum, diese Kopplungsterme am besten zu approximieren und zu berücksichtigen, auch wenn die Zeitentwicklung der jeweiligen q_2, p_2, q_5, p_5 nicht numerisch bewertet

wird.

Aus (3.5) und der Anwendung der der "Optimal Prediction" zugrundeliegenden Gleichung (1.13) wissen wir, daß die Kopplungsterme q_3, p_3, q_4 und p_4 , die auf der rechten Seite von (3.5) stehen, deren bedingten Erwartungswerten bei dem gegebenen Satz an Anfangsbedingungen (3.7) gleichgestellt werden sollen. Gesucht sind nun in der Terminologie von "Optimal Prediction" die Mittelwerte $\langle q_3 \rangle_V, \langle p_3 \rangle_V, \langle q_4 \rangle_V, \langle p_4 \rangle_V$ für eine spätere Zeit $t > 0$.

Nun ist der Hamiltonian unseres Systems nicht Gaußscher Form; man wäre deshalb auf eine Partitionierung des Hamiltonians nach der im vorigen Paragraphen dargestellten Methode angewiesen. Die gesuchten Erwartungswerte ergeben sich dann aus der Störungsrechnung (2.21). Aus der Tatsache, daß $Cov\{p_i, p_j\}^0 = 0$ und $Cov\{q_i, q_j\}^0 = 0$ sind, und mit Hilfe der Lemmata (2.15) und (2.16) folgt aber, daß die entsprechenden bedingten Erwartungswerte gleich Null sind. D. h. eine Beschränkung auf die 0. Ordnung der "Optimal Prediction" Methode, wie sie bis jetzt gemacht worden ist, würde erstmal keinen zusätzlichen Term, daher auch keine Besserung gegenüber einem Galerkin-Verfahren bringen.

Dennoch versuchen wir in einer etwas veränderten Form der 0. Ordnung die Kopplungsterme zu berücksichtigen, indem wir für sie deren aus der Monte-Carlo Simulation sich ergebenden Mittelwert zur Zeit $t = 0$ einsetzen. D. h. wir versuchen das System (3.5) durch folgendes zu approximieren:

$$\begin{cases} \dot{q}_3 = -\gamma p_3(q_3^2 + p_3^2) - \varepsilon p_4 - \frac{1}{2} \varepsilon \langle p_2 \rangle_{t=0} \\ \dot{q}_4 = -\gamma p_4(q_4^2 + p_4^2) - \varepsilon p_3 - \frac{1}{2} \varepsilon \langle p_5 \rangle_{t=0} \\ \dot{p}_3 = +\gamma q_3(q_3^2 + p_3^2) + \varepsilon q_4 + \frac{1}{2} \varepsilon \langle q_2 \rangle_{t=0} \\ \dot{p}_4 = +\gamma q_4(q_4^2 + p_4^2) + \varepsilon q_3 + \frac{1}{2} \varepsilon \langle q_5 \rangle_{t=0} \end{cases} \quad (3.9)$$

Die Integration dieser Gleichungen bringt bessere Ergebnisse als eine Galerkin-Integration. Die erste Entwicklung wird teilweise sehr gut approximiert, bis der Kopplungsterm sich zu sehr von seinem Wert zur Zeit $t = 0$ entfernt. Stellvertretend hierfür sei die Entwicklung von p_4 in der Abbildung (3.1) dargestellt.

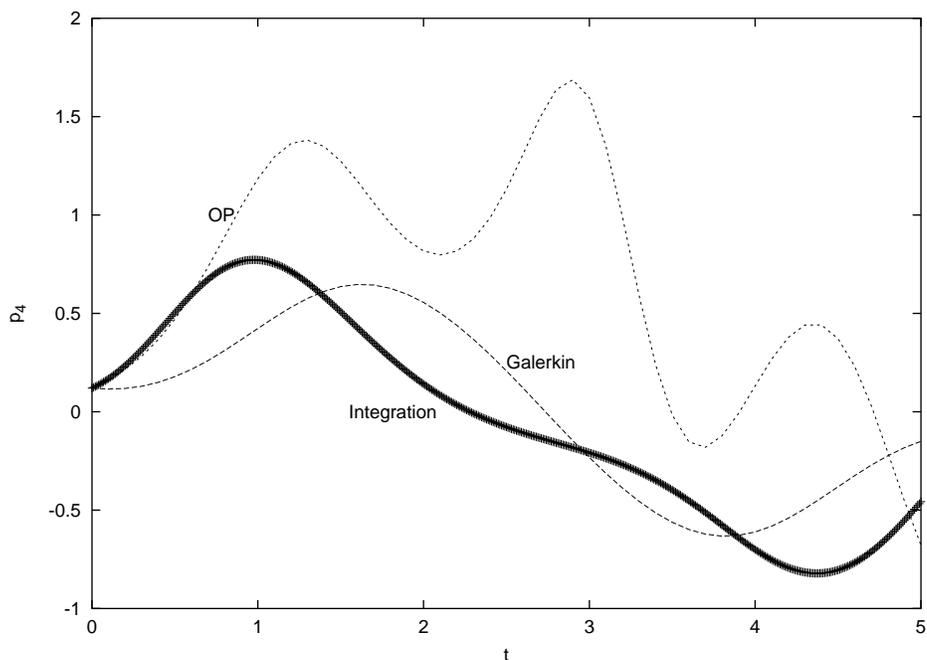


Abbildung 3.1: Zeitentwicklung von p_4 : hier werden die Ergebnisse aus der Integration, aus einem Galerkin-Verfahren und aus der 0. Ordnung der “Optimal Prediction” mit Modellierung des Anfangsrauschens nach (3.9) verglichen; solange der Kopplungsterm sich nicht zu sehr von seinem Wert zur Zeit $t = 0$ entfernt, liefert die “Optimal Prediction” gute Ergebnisse, insbesondere bessere, als ein Abschneiden mit einem Galerkin-Verfahren.

3.2. Höhere Zeitentwicklungen der “Optimal Prediction”

Die 0. Ordnung der “Optimal Prediction” Methode, die wir bisher präsentiert haben, erzeugt ein System gewöhnlicher Differentialgleichungen: es sind die Bewegungsgleichungen des Satzes gewählter kollektiver Koordinaten; deren zeitliche Entwicklung ist allerdings nur für eine kurze Zeit richtig wiedergegeben, die von der Anzahl und der Wahl der kollektiven Koordinaten und der Varianz der Anfangsdaten, i. e. der Temperatur abhängig ist [13, 19, 15, 14]. Es ist zwar die “optimale” Voraussage in dieser Ordnung für Markov’sche Approximationen [32], dennoch wird dabei das Abklingen mit der Zeit des Einflusses der Anfangsdaten auf die gesamte Verteilung der Lösungen nicht berücksichtigt. Anders ausgedrückt: bei ergodischen Systemen spielen nach gewisser Zeit die Anfangsdaten keine Rolle mehr, und dieser Tatsache trägt die 0. Ordnung nicht Rechnung [19]. In Anlehnung an den Mori-Zwanzig-Formalismus, der seine An-

wendung bei nichtgleichgewichtigen statistischen Systemen findet, wurde von Chorin *et al.* eine verallgemeinerte Langevin-Gleichung entwickelt, die einen nicht-Markov’schen, sogenannten “memory”-Term beinhaltet und damit die nächsthöhere Ordnung der “Optimal Prediction” Methode darstellt. Das Mori-Zwanzig-Verfahren verbindet die Entwicklung von makroskopischen Variablen mit der mikroskopischen Dynamik; das Pendant zu den makroskopischen Variablen sind bei Chorin *et al.* die kollektiven Koordinaten des unterbestimmten Systems [13]. Diese Herleitung wird nun für eine Bewegungsgleichung

$$\frac{d\varphi}{dt} = R(\varphi(t)), \quad \varphi(0) = x \quad (3.10)$$

im Phasenraum Γ dargestellt und erläutert. Dabei sind R und φ n -dimensionale Vektoren. Unser Ziel bleibt die Berechnung von Mittelwerten für $\varphi_i(t)$ für $1 \leq i \leq m$, mit $m < n$. Diese Mittelwerte sollen ja Funktionen $\hat{\varphi}(t)$ ausschließlich der aufgelösten Variablen \hat{x} sein, d. h. man sucht deren Projektionen auf den Raum der Funktionen von x_1, \dots, x_m . Es ist bewiesen, daß die beste Projektion dafür diejenige ist, die in den vorigen Abschnitten benutzt wurde und sich auf die bedingten Erwartungswerte stützt: sie ist die beste Approximation von einer Funktion der gesamten n Variablen x mit einer Funktion nur von den m ersten Variablen \hat{x} ; sie minimiert den mittleren quadratischen Abstand zwischen beiden [14, 13]. Allerdings sind auch andere Projektionen wie die lineare Projektion und die “finite-rank” Projektion geläufig.

Jeder Anfangsbedingung x entspricht eine Trajektorie $\varphi(t) = \varphi(x, t)$. Die Entwicklung einer beliebigen Phasenvariable u ist durch den sogenannten “Entwicklungsoperator” S^t und die Gleichung $(S^t u)(x) = u(\varphi(x, t))$ gegeben [15]. Das Differenzieren dieser Gleichung liefert dann:

$$\frac{\partial}{\partial t}(S^t u)(x) = \sum_i R_i(\varphi(x, t)) \frac{\partial u}{\partial \varphi_i}(\varphi(x, t)) = LS^t u(x), \quad (3.11)$$

wobei der Liouville-Operator durch $L = \sum_i R_i(\varphi) \frac{\partial}{\partial \varphi_i}$ gegeben ist. Gewöhnlich schreibt man für S^t : $S^t = e^{tL}$ [14].

Betrachten wir jetzt den Spezialfall, in dem die Phasenvariable $u(\varphi(t))$ eine der kollektiven Koordinaten $\varphi_j(x, t) = e^{tL} x_j$ ist. Das Mori-Zwanzig-Verfahren besteht nun darin [40, 54], die

Zeitableitung $R_j(\varphi(x, t)) = e^{tL}Lx_j$ in einen auf den Raum der Funktionen $\hat{\varphi}(t)$ projizierten Anteil und den dazu orthogonalen Anteil aufzuspalten:

$$\frac{\partial}{\partial t} e^{tL}x_j = e^{tL}Lx_j = e^{tL}PLx_j + e^{tL}QLx_j, \quad (3.12)$$

wobei P eine der oben genannten Projektionen ist und $Q = I - P$. Der erste Term ist eine Funktion von u und kann geschrieben werden als $e^{tL}PLx_j = \mathcal{R}_j(\hat{\varphi}(x, t))$ mit $\mathcal{R}_j(\hat{x}) := (PR_j)(\hat{x})$ [15]. Der zweite Term rührt von der sogenannten ‘‘orthogonalen’’ Dynamik her [15, 14, 13] und kann mit Hilfe der Dyson-Formel umformuliert werden in:

$$e^{tQL}QLx_j + \int_0^t e^{(t-s)L}PL e^{sQL}QLx_j ds, \quad (3.13)$$

so da man aus (3.11) und mit den Definitionen $F_j(x, t) = e^{tQL}QLx_j$ und $K_j(\hat{x}, t) = PLF_j(x, t)$ schlielich die verallgemeinerte Langevin-Gleichung

$$\frac{\partial}{\partial t} e^{tL}x_j = e^{tL}\mathcal{R}_j(\hat{x}, t) + \int_0^t e^{(t-s)L}K_j(\hat{x}, s) ds + F_j(x, t) \quad (3.14)$$

erhlt [15], oder in einer etwas anderen Form:

$$\frac{\partial}{\partial t} \varphi_j(x, t) = \mathcal{R}_j(\hat{\varphi}(x, t)) + \int_0^t K_j(\hat{\varphi}(x, t-s), s) ds + F_j(x, t). \quad (3.15)$$

Den ersten Term dieser Gleichung erkennen wir als die 0. Ordnung der ‘‘Optimal Prediction’’ Methode. Er entspricht der rechten Seite der Gleichung (1.13), whrend die zwei letzten Terme den sogenannten ‘‘memory’’-Term ausmachen [15, 13, 14]: der zweite Term hngt von den Werten von u zwischen den Zeiten 0 und t ab; der letzte Term hat keine Komponenten auf dem Raum der Funktionen von \hat{x} , er kann als Rauschen, bzw. Zufallsterm angesehen werden, dessen Statistik durch die Anfangsdaten bestimmt wird. Da der ‘‘memory’’-Term von dem Rauschen abhngt, entspricht dem sogenannten ‘‘Fluktuations-Dissipationstheorem’’ der statistischen Physik [15], dessen bliche Form aus der speziellen Wahl der linearen Projektion fr P folgt. Eine Untersuchung des Zusammenhangs der ‘‘Optimal Prediction’’ Methode mit dem Fluktuations-Dissipationstheorem und die daraus folgende Approximation eines unterbestimmten Systems

durch eine stochastische Differentialgleichung kann in [2] nachgelesen werden, wo dies auf die gemittelten Euler-Gleichungen im Zweidimensionalen angewandt wurde.

Durch Projektion erhält man dann die gesuchte Funktion:

$$\frac{\partial}{\partial t} P\varphi_j(x, t) = P\mathcal{R}_j(\hat{\varphi}(x, t)) + \int_0^t PK_j(\hat{\varphi}(x, t-s), s) ds \quad (3.16)$$

Hier wurde also statt der Bewegungsgleichung für die Komponente $\varphi_j(t)$

$$\frac{d\varphi_j}{dt} = R_j(\varphi(t)), \quad (3.17)$$

die direkt aus (3.10) folgt, ihre Projektion $P\varphi_j(x, t)$ auf den Raum der Funktion der aufgelösten Koordinaten \hat{x} betrachtet, und die rechte Seite $P(R_j(\varphi(t)))$ dann durch eine Funktion von \hat{x} approximiert.

Falls für P die auf bedingten Erwartungswerten basierte Projektion angesetzt wird, dann bekommt man genau den Mittelwert $\hat{\varphi}(x, t)$ bedingt durch die anfänglichen Werte der kollektiven Koordinaten, d. h. die optimale Approximation.

Es ist interessant anzumerken, daß die Gleichungen (3.15) und (3.16) exakt und dem ursprünglichen Problem (3.10) völlig äquivalent sind; in der Praxis können sie aber selten exakt gelöst werden, man ist auf Approximationen angewiesen [15].

Die erste mögliche Approximation kann darin bestehen, den bedingten Erwartungswert $E[\varphi_j(x, t)|\hat{x}]$, die optimale Projektion, durch eine “finite-rank” Projektion der Form

$$E[\varphi_j(x, t)|\hat{x}] = \sum_{\nu} (\varphi_j(t), h^{\nu}) h^{\nu}(\hat{x}) \approx \sum_{\nu} c_j^{\nu}(t) h^{\nu}(\hat{x}) \quad (3.18)$$

mit der bekannten, orthonormalen Funktionenbasis $(h^{\nu}(\hat{x}))_{\nu \in I}$ zu ersetzen. Dabei ist auf die Probleme der möglicherweise schwachen Konvergenz einer solchen Summe zu achten (cf. [15] und Zitate darin).

Wenn durch eine Monte-Carlo-Simulation und anschließende Integration eines großen Satzes von Anfangsdaten die Koeffizienten $c_j^{\nu}(t)$ für jede Zeit t berechnet worden sind, dann kann man für ein spezifisches Problem den Erwartungswert $E[\varphi_j(x, t)|\hat{x}]$ durch die rechte Seite der obigen

Gleichung approximieren, in der die $c_j^v(t)$ nun bekannt sind, und für die Argumente \hat{x} der Basisfunktionen die spezifischen Werte des betrachteten Problems eingesetzt werden.

Die Berechnung des “memory”-Terms auf der rechten Seite von (3.16) kann Schwierigkeiten bereiten; in den wenigsten Fällen ist sie exakt möglich, d. h. wenn der bedingte Erwartungswert exakt berechnet werden kann; dann ist tatsächlich aber eine Verbesserung der zeitlichen Voraussage der Dynamik des Systems im Vergleich zur 0. Ordnung der “Optimal Prediction” Methode zu beobachten [13]. Meistens muß aber bei der Berechnung von $K_j(\hat{x}, t)$ eine Approximation gemacht werden; für eine detaillierte Darstellung wird auf [15] verwiesen. Zusammenfassend kann man sagen, daß jede Komponente $F_j(x, t)$ des Rauschens über eine Volterra-Gleichung bestimmt wird und nach der oben beschriebenen “finite-rank”-Projektion projiziert wird. Daraus folgt für $K_j(\hat{x}, s)$:

$$K_j(\hat{x}, s) = PLF_j(x, s) = \sum_v a_j^v(s) h^v(\hat{x}), \quad (3.19)$$

wobei die $a_j^v(s)$ die Korrelationen $a_j^v(s) = (LF_j(s), h^v)$ sind. Durch das Einsetzen dieses letzten Ergebnisses in die Volterra-Gleichung für $F_j(x, t)$ folgt letztendlich wieder eine Volterra-Gleichung für $a_j^\mu(s)$ der Form

$$a_j^\mu(s) = f_j^\mu(t) - \int_0^t \sum_v \in I(s) g^{v\mu}(t-s) ds \quad (3.20)$$

mit $f_j^\mu(t) = (Le^{tL}F_j(0), h^\mu)$ und $g^{v\mu}(t) = (Le^{tL}h^v, h^\mu)$. Hier können die Funktionen $f_j^\mu(t)$ und $g^{v\mu}(t)$ über eine Monte-Carlo-Simulation bestimmt werden; anschliessend ist dann der “memory”-Term durch (3.19) gegeben. Über die zwei Ergebnisse (3.18) und (3.20) haben Chorin *et al.* ein “allgemeines Gerüst” entwickelt, das es ermöglicht, sobald eine Monte-Carlo-Simulation durchgeführt worden ist, dann für jedes spezifisches Problem eine optimale Approximation zu liefern, deren Koeffizienten unabhängig von den spezifischen Anfangsbedingungen sind [15, 14].

Der hier hergeleitete “memory”-Term wird im folgenden mit “finite-time memory” Term bezeichnet, um ihn von seiner linearisierten Form (“linearized memory”-Term), die bei Chorin *et al.* “short-memory”-Approximation genannt wird, zu unterscheiden. Nun werden wir sehen,

daß die Miteinbeziehung dieses “linearized memory”-Terms eine deutliche Verbesserung in der Untersuchung unseres Modells für eine zehndimensionale diskrete DST bringt.

3.3. “Linearized memory” Approximation

Wir interessieren uns also für den zusätzlichen “finite-time memory”-Term, den eine weitere Ordnung der “Optimal Prediction” Methode mit sich bringt. Dabei beschränken wir uns auf die in [15] dargestellte “linearized memory”-Approximation. Der Integrand des “finite-time memory”-Terms auf der rechten Seite der verallgemeinerten Langevin-Gleichung (3.16)

$$\int_0^t K_j(\hat{\phi}(x, t-s), s) ds$$

wird dafür in eine Taylor-Reihe um den Wert $s = 0$ entwickelt. Dabei wird nur der erste Term behalten:

$$\int_0^t K_j(\hat{\phi}(x, t), 0) ds = t S_j(\hat{\phi}(x, t)), \quad (3.21)$$

mit

$$S_j(x) = PLQLx_j.$$

Schreiben wir

$$\phi(t) \approx P\hat{\phi}(x, t),$$

bekommen wir dann als Approximation für die Gleichung (3.16) die gedämpfte Gleichung

$$\begin{cases} \frac{d\phi(t)}{dt} = \mathcal{R}(\phi(t)) + t S(\phi(t)) \\ \phi(0) = \hat{x}, \end{cases} \quad (3.22)$$

die linear in t ist: die “linearized memory”-Approximation.

Wenden wir nun dieses Schema auf unser diskretes Modell an, ergibt sich für die erste Ordnung des Integranden $K_j(\hat{\phi}(x, t-s), s)$ für die verschiedenen kollektiven Koordinaten q_3, q_4, p_3, p_4 :

$$\begin{cases} S_{q_3} = -\frac{\varepsilon^2}{4}q_3 \\ S_{q_4} = -\frac{\varepsilon^2}{4}q_4 \\ S_{p_3} = -\frac{\varepsilon^2}{4}p_3 \\ S_{p_4} = -\frac{\varepsilon^2}{4}p_4 \end{cases}$$

Daraus folgen für den Fall, daß nur die 0. Ordnung “Optimal Prediction” berücksichtigt wird, die gedämpften Gleichungen:

$$\begin{cases} \dot{q}_3 = -\gamma p_3(q_3^2 + p_3^2) - \varepsilon p_4 + t \left(\frac{-\varepsilon^2}{4}q_3\right) \\ \dot{q}_4 = -\gamma p_4(q_4^2 + p_4^2) - \varepsilon p_3 + t \left(\frac{-\varepsilon^2}{4}q_4\right) \\ \dot{p}_3 = +\gamma q_3(q_3^2 + p_3^2) + \varepsilon q_4 + t \left(\frac{-\varepsilon^2}{4}p_3\right) \\ \dot{p}_4 = +\gamma q_4(q_4^2 + p_4^2) + \varepsilon q_3 + t \left(\frac{-\varepsilon^2}{4}p_4\right) \end{cases} \quad (3.23)$$

Da die Beiträge $\langle q_2 \rangle, \langle p_2 \rangle, \langle q_5 \rangle$ und $\langle p_5 \rangle$ gemittelt bezüglich eines Gaußschen Maßes Null ergeben, und nur diese erste Ordnung berücksichtigt wurde, entspricht der erste Teil dieser Gleichungen, der Markov’sche Term $\mathcal{R}(\phi(t))$ von (3.22), allerdings exakt dem Galerkin-Term der Gleichungen (3.8); ihm wurde nun ein in t linearer Dämpfungsterm hinzuaddiert.

Zuerst wurde der Fall eines verschwindenden Anfangsrauschens betrachtet; die Integration des Gleichungssystems (3.23) wurde durchgeführt, wobei der “linearized memory”-Term nur bis zu einer heuristisch bestimmten Zeit t_1 berücksichtigt wurde, was dem Abklingen seines Beitrags in der Zeit Rechnung trägt; der Integrand

$$\int_0^t K_j(\hat{\phi}(x, t-s), s) ds$$

wurde auf

$$\int_0^{t=1.55} K_j(\hat{\phi}(x, t-s), s) ds$$

trunkiert. Am Beispiel der so gewonnenen zeitlichen Entwicklung von p_3 (Abbildung 3.2) wird klar, daß das Trunkieren des Integranden eine Verbesserung bezüglich einer einfachen “linearized memory” Approximation liefert. Das qualitative Verhalten der Lösung wird so über eine

längere Zeit approximiert. Die 0. Ordnung der “Optimal Prediction” fällt im Falle verschwindenden Rauschens mit dem Galerkin-Verfahren zusammen und taucht deswegen nicht gesondert auf.

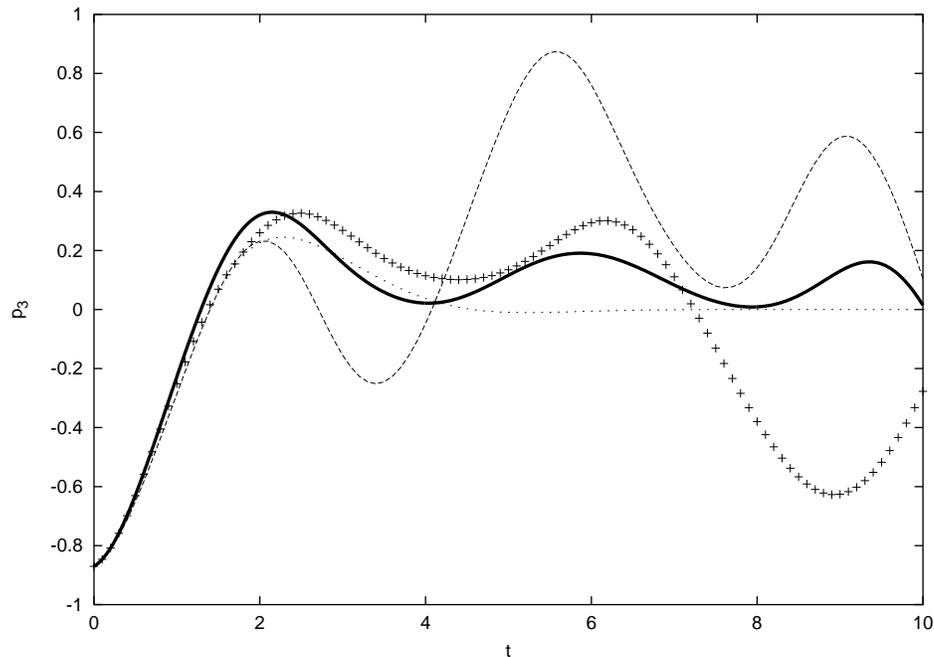


Abbildung 3.2: Zeitentwicklung von p_3 : Vergleich zwischen der Integration (dicke durchgezogene Kurve), dem Galerkin-Verfahren (feine gestrichelte Kurve), der “linearized memory” Approximation der “Optimal Prediction” 0. Ordnung (gepunktete Kurve) und schließlich der “linearized memory”-Approximation mit trunziertem Integranden (Symbole). Diese letzte Kurve bringt die besseren qualitativen Ergebnisse und approximiert recht gut die aus der Integration erhaltene Kurve.

Eine “Mischform” der “Optimal Prediction” Methode wurde nun im Falle nichtverschwindenden Anfangsrauschens implementiert, indem der Einfluß der Anfangsbedingungen insofern berücksichtigt wird, daß man für eine Zeit $t \leq t_1$ zusätzlich zum “linearized memory”-Term die obige Approximation $\langle q_i \rangle \approx \langle q_i \rangle_{t=0}, \langle p_i \rangle \approx \langle p_i \rangle_{t=0}, i = 2, 5$ für die Kopplungsterme einsetzt, und dann für $t \geq t_1$ diese Kopplungsterme einfach gleich Null setzt, entsprechend der 0. Ordnung der störungstheoretischen Entwicklung des Markov’schen Terms. Dies ist eine vertretbare Approximation, insofern der Einfluß der Anfangsdaten mit der Zeit abklingt [15, 14, 19]. Der “linearized memory”-Term wurde nur für eine Zeit $t \leq t_2, t_2 > t_1$ berücksichtigt; dies spie-

gelt die Tatsache wider, daß die erste Ordnung in der Entwicklung des “finite-time memory”-Term nur für die erste Zeit ausschlaggebend ist. Ein solches heuristisches Abschneiden des Integranden, um der Relaxation des Systems in sein Gleichgewichtszustand zu entsprechen, hat schon bei der Untersuchung der kubisch nichtlinearen Schrödinger-Gleichung gute Ergebnisse geliefert [13]. Für spätere Zeiten ist ein bedeutender Einfluß der höheren Terme zu erwarten. Insofern erwartet man, daß die Güte der nun so implementierten Version der “Optimal Prediction” Methode in der Zeit begrenzt ist. Die Zeiten t_1, t_2 werden hier auch heuristisch durch die Wiedergabe des qualitativen reellen Verhaltens begründet.

$$\text{Für } t \leq 0.5 \quad \begin{cases} \dot{q}_3 = -\gamma p_3(q_3^2 + p_3^2) - \varepsilon p_4 - \frac{1}{2} \varepsilon \langle p_2 \rangle_{t=0} + t \left(\frac{-\varepsilon^2}{4} q_3 \right) \\ \dot{q}_4 = -\gamma p_4(q_4^2 + p_4^2) - \varepsilon p_3 - \frac{1}{2} \varepsilon \langle p_5 \rangle_{t=0} + t \left(\frac{-\varepsilon^2}{4} q_4 \right) \\ \dot{p}_3 = +\gamma q_3(q_3^2 + p_3^2) + \varepsilon q_4 + \frac{1}{2} \varepsilon \langle q_2 \rangle_{t=0} + t \left(\frac{-\varepsilon^2}{4} p_3 \right) \\ \dot{p}_4 = +\gamma q_4(q_4^2 + p_4^2) + \varepsilon q_3 + \frac{1}{2} \varepsilon \langle q_5 \rangle_{t=0} + t \left(\frac{-\varepsilon^2}{4} p_4 \right) \end{cases}$$

$$\text{Für } t \leq 2.0 \quad \begin{cases} \dot{q}_3 = -\gamma p_3(q_3^2 + p_3^2) - \varepsilon p_4 + t \left(\frac{-\varepsilon^2}{4} q_3 \right) \\ \dot{q}_4 = -\gamma p_4(q_4^2 + p_4^2) - \varepsilon p_3 + t \left(\frac{-\varepsilon^2}{4} q_4 \right) \\ \dot{p}_3 = +\gamma q_3(q_3^2 + p_3^2) + \varepsilon q_4 + t \left(\frac{-\varepsilon^2}{4} p_3 \right) \\ \dot{p}_4 = +\gamma q_4(q_4^2 + p_4^2) + \varepsilon q_3 + t \left(\frac{-\varepsilon^2}{4} p_4 \right) \end{cases}$$

$$\text{Für } t > 2.0 \quad \begin{cases} \dot{q}_3 = -\gamma p_3(q_3^2 + p_3^2) - \varepsilon p_4 \\ \dot{q}_4 = -\gamma p_4(q_4^2 + p_4^2) - \varepsilon p_3 \\ \dot{p}_3 = +\gamma q_3(q_3^2 + p_3^2) + \varepsilon q_4 \\ \dot{p}_4 = +\gamma q_4(q_4^2 + p_4^2) + \varepsilon q_3 \end{cases}$$

Im Falle nichtverschwindenden Rauschens liefert diese Implementierung für kurze Zeiten sehr gute Ergebnisse; im Gegenteil zu der 0. Ordnung der “Optimal Prediction”, die den “finite-time memory”-Term gar nicht berücksichtigt, liefert sie eine sehr gute qualitative Übereinstimmung mit der gemittelten integrierten Trajektorie, wie an den Abbildungen (3.3) und (3.4) zu beobachten ist.

Das betrachtete Modell der DST, dessen Hamiltonian durch (3.6) gegeben ist, ist ein konservatives System, in dem eine “effektive” Dämpfung der Hauptamplituden - unserer kollektiven Koordinaten-, verursacht durch die Nebenamplituden, auftritt. Wir haben nun gesehen, daß für ein solches physikalisches System die “Optimal Prediction” Methode, wenn sie bis zum “linearized memory”-Term berücksichtigt wird, eine sehr gute Approximation für das qualitative Verhalten des Systems liefert. Dabei sind die Rechnungen und Approximationen, die zu diesem Ergebnis führen, einfach und benötigen nicht die volle Berechnung des “finite-time memory”-Terms.

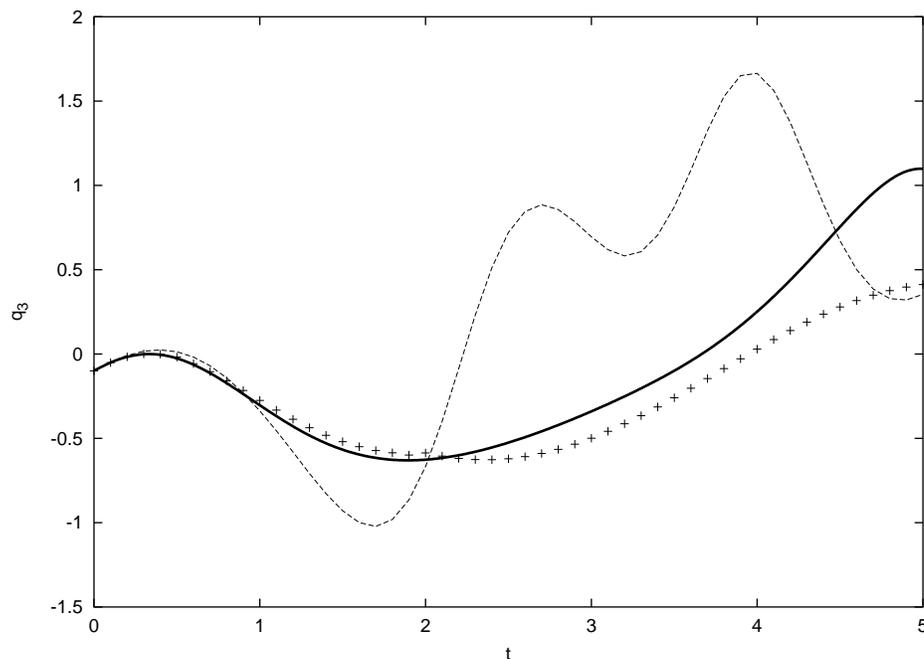


Abbildung 3.3: Zeitentwicklung von q_3 : Vergleich zwischen der Integration (dicke durchgezogene Kurve), der “Optimal Prediction” 0. Ordnung mit Modellierung des Anfangsrauschens (feine gestrichelte Kurve) und der “linearized memory” Approximation mit trunziertem Integranden (Symbole). Eindeutig gibt dieses letzte Verfahren das qualitative Verhalten des Systems am besten wieder.

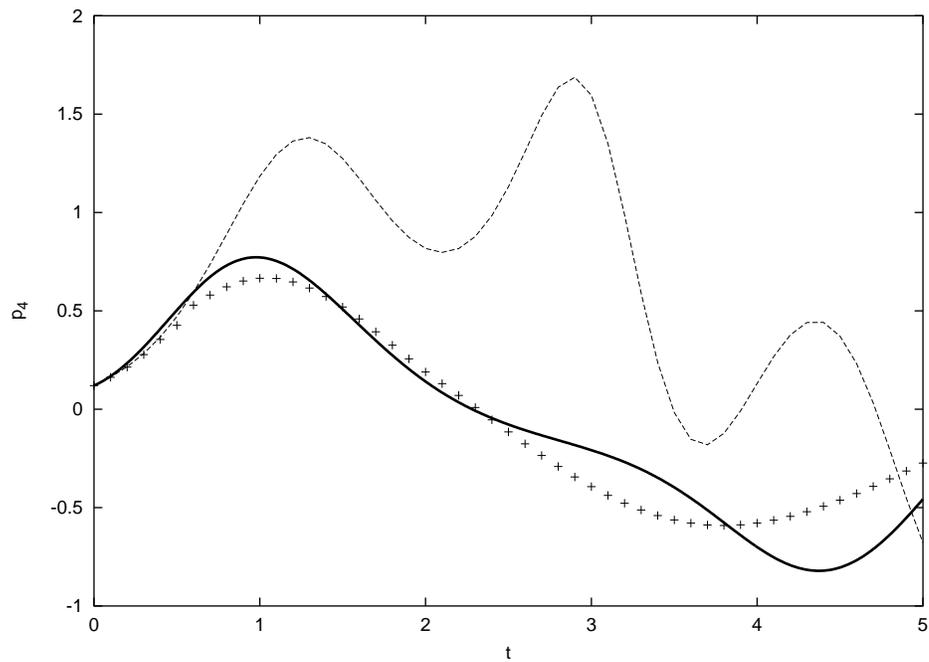


Abbildung 3.4: Zeitentwicklung von p_4 : Vergleich zwischen der Integration (dicke durchgezogene Kurve), der "Optimal Prediction" 0. Ordnung mit Modellierung des Anfangsrauschens (feine gestrichelte Kurve) und der "linearized memory" Approximation mit trunziertem Integranden (Symbole). Eindeutig gibt dieses letzte Verfahren das qualitative Verhalten des Systems am besten wieder.

Teil II.

Nicht-Hamiltonsche dissipative Systeme

Im ersten Teil dieser Arbeit haben wir Hamiltonsche Systeme der Dimension n untersucht (falls $n = \infty$ ist, dann führt eine Diskretisierung des Problems auf ein endlich-dimensionales System zurück), deren Dynamik durch ein System von n gewöhnlichen Differentialgleichungen gegeben ist. Diese sind die Bewegungsgleichungen der n Freiheitsgrade, die zur geschlossenen Beschreibung des Systems nötig sind. Allerdings wurde angenommen, daß uns unmöglich war, die n Gleichungen zu implementieren, und daß zur Zeit $t = 0$ nur $N < n$ Anfangsdaten zur Verfügung stehen, die als “kollektive Koordinaten” bezeichnet werden können. Insofern ist das System unterbestimmt.

Vorausgesetzt wurde die Kenntnis einer zusätzlichen statistischen Information über dieses System in Form eines invarianten Maßes, für das man bei Hamiltonschen Systemen das kanonisch induzierte Maß ansetzen kann.

Es wurde untersucht, inwiefern die “Optimal Prediction” Methode, die von Chorin *et al.* in verschiedenen Schritten entwickelt wurde [16, 18, 17, 19, 13, 14, 15, 2], die gemittelte zeitliche Entwicklung der kollektiven Koordinaten wiedergibt. Dabei wurde als Referenz die gemittelte Trajektorie genommen, die man aus der Integration eines Ensembles von größenordnungsmäßig $A \approx 10^3$ möglichen Anfangsdaten bekommt, die eine Monte-Carlo-Simulation liefert.

Der entscheidende Ausgangspunkt dabei war die Überlegung, daß ein unterbestimmtes konservatives System sich im Mittel wie ein dissipatives System verhält [2]. So wurden Parallelen zwischen der “Optimal Prediction” Methode und den Methoden, insbesondere dem Mori-Zwanzig-Formalismus, der statistischen Physik hervorgehoben [13, 15, 14].

Ausgehend von den vielversprechenden Ergebnissen der Methode, wenn sie auf konservative

Systeme im Sinne der eben beschriebenen Mittelung angewandt wird, werden wir nun untersuchen, ob man die Methode auch bei dissipativen Systemen anwenden kann.

Hier wird ausdrücklich darauf hingewiesen, daß nicht mehr über einen Ensemble gemittelt wird, sondern daß man sich nun für einzelne Trajektorien innerhalb eines dissipativen Systems interessieren wird.

In einem ersten Schritt werden wir ein einfaches, zweidimensionales dissipatives, mathematisches Modell mit einer gedämpften Variable nach der Anwendbarkeit der “Optimal Prediction” Methode untersuchen. Insbesondere wird dabei die Güte der im letzten Paragraphen definierten “finite-time memory”-Approximation gezeigt. Ein solches dissipatives System bietet den Vorteil, daß das Abklingverhalten der gedämpften Variable bekannt ist und das System somit exakt lösbar ist.

Dennoch wird die Situation bei mehreren miteinander gekoppelten Dämpfungen komplizierter, wie wir in einem zweiten Schritt bei dem physikalischen Problem des L-H-Übergangs in toroidal eingeschlossenen Fusionsplasmen sehen werden.

4. Demonstration der Güte des “finite-time memory”-Terms an einem einfachen mathematischen Modell

Betrachten wir erstmal das folgende, zweidimensionale System:

$$\begin{cases} \dot{x} = -xy \\ \dot{y} = -y + x^2 \end{cases} \quad (4.1)$$

Die zwei Freiheitsgrade x und y werden beide gedämpft und die Gesamtenergie des Systems $H = x^2 + y^2$ nimmt mit der Zeit ab:

$$\begin{aligned} \frac{d}{dt}H &= \frac{d}{dt}(x^2 + y^2) \\ &= (2\dot{x}x + 2\dot{y}y) \\ &= (2x(-xy) + 2y(-y + x^2)) \\ &= -2y^2 < 0 \end{aligned}$$

Es handelt sich also um ein dissipatives System. Solche Systeme zeichnen sich im Gegensatz zu Hamiltonschen Systemen dadurch aus, daß das Phasenraumvolumen keine Konstante der Bewegung mehr ist; im Gegenteil kontrahiert es mit fortschreitender Zeit. Dies führt zu einer Bewegung auf einer Fläche von niedrigerer Dimension als der ursprüngliche Phasenraum. Die Bewegung wird also nicht mehr von kanonischen Gleichungen beschrieben, sondern allgemein von einem System von Differentialgleichungen erster Ordnung [38]. Die Anzahl dieser Differentialgleichungen gibt die Dimension des Phasenraums wieder; in unserem Falle also ist er zweidimensional.

Betrachtet man nun die zweite Gleichung; es handelt sich um eine inhomogene Differentialgleichung 1. Ordnung für die Variable y . Ihre Lösung setzt sich aus der Lösung der homogenen und einer speziellen Lösung der inhomogenen Gleichung zusammen.

Die homogene Gleichung $\dot{y} = -y$ hat die Lösung $y = y_0 e^{-t}$, die wir unten in der Anwendung der “Optimal Prediction” Methode wiedererkennen werden. Addiert man hierzu eine spezielle

Lösung der inhomogenen Gleichung, bekommt man:

$$y = y_0 e^{-t} + e^{-t} \int_0^t ds e^s x(s)^2$$

Daraus folgt, daß das Gleichungssystem auch geschrieben werden kann als:

$$\dot{x} = -x \left(y_0 e^{-t} + e^{-t} \int_0^t ds e^s x(s)^2 \right)$$

4.1. Reduktion über die Zentrumsmannigfaltigkeitstheorie

Das System besitzt den Fixpunkt $(x = 0, y = 0)$.

Untersucht man dessen Stabilität mit der Linearisierungsmethode [42], zeigt sich, daß ein solches Gleichungssystem über die Zentrumsmannigfaltigkeitstheorie reduziert werden kann. Die Stabilität des Fixpunktes wird nämlich durch die folgende Eigenwertgleichung bestimmt:

$$(J_\xi - \sigma I) \mathbf{e} = 0,$$

wobei σ der Eigenwert und \mathbf{e} der Eigenvektor sind. J ist die Jacobi-Matrix des Systems, d. h.

$$\mathbf{J} = \begin{pmatrix} -y & -x \\ 2x & -1 \end{pmatrix}$$

und \mathbf{J}_ξ die Jacobi-Matrix an der Stelle $(x = 0, y = 0)$.

Die Eigenwerte sind $\xi = 0$ und $\xi = -1$. Der instabile Eigenraum des Fixpunktes ist leer und die Berechnung der Eigenvektoren zeigt, daß der zentrale Eigenraum die x -Achse und der stabile Eigenraum die y -Achse sind. Die zentrale Mannigfaltigkeit geht durch den Fixpunkt $(0, 0)$ und liegt tangential zum zentralen Eigenraum, d. h. zur x -Achse [42].

Es handelt sich um einen Diffeomorphismus:

$$M^c : y = y(x) \quad \text{mit} \quad y(0) = y_x(0) = 0,$$

für den man als Approximation ein Polynom der Form $\sum_n A_n x^n$ ansetzen kann und insbesondere in niedrigster Ordnung:

$$y = A_0 x^0 + A_1 x^1 + A_2 x^2 + O(x^3),$$

wobei mit $O(x^3)$ bedeutet, daß Polynome des Grades 3 und höheren vernachlässigt werden.

Nach Einsetzen in das Gleichungssystem und unter Berücksichtigung der Bedingungen

$$\begin{cases} y(0) = 0 \implies A_0 = 0 \\ y_x(0) = 0 \implies A_1 = 0 \end{cases}$$

folgt $y = A_2 x^2$; man kann annehmen: $A_2 = 1$. Man bekommt für die erste Gleichung $\dot{x} = -xy$ des Systems (4.1) die Approximation $y \approx x^2$. Durch die Zentrumsmannigfaltigkeitstheorie wird also das ursprüngliche Zweigleichungssystem auf folgende Gleichung für x reduziert:

$$\dot{x} \approx -x^3 \tag{4.2}$$

Gegenüber (4.1) wurde die Ordnung der Differentialgleichung um eins reduziert. Nun liegt ein Fixpunkt bei $x = 0$ vor: der Attraktor des ursprünglichen Systems [38].

4.2. Anwendung der "Optimal Prediction" Methode

Wenden wir uns jetzt der "Optimal Prediction" Methode zu. Die Untersuchung erfolgt ähnlich den Arbeiten von Chorin *et al.* [14] und besonders [15]; unser Gleichungssystem wird an deren Terminologie angepaßt. Sei φ der zweidimensionale Vektor

$$\varphi = \begin{pmatrix} x \\ y \end{pmatrix},$$

und sei die Entwicklung des dynamischen Systems gegeben durch:

$$\begin{cases} \frac{d}{dt}\varphi(t) = R(\varphi(t)) \\ \varphi(0) = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \end{cases}$$

Dabei sind x_0 und y_0 die Anfangswerte und $R(\varphi(t))$ der zweidimensionale Vektor $\begin{pmatrix} -xy \\ -y+x^2 \end{pmatrix}$.

Ziel ist die Reduzierung des obigen Gleichungssystems auf eine einzige Gleichung für x ; x spielt also in der Terminologie von "Optimal Prediction" die Rolle der einzig zugänglichen kollektiven Variable. Somit ist das System unterbestimmt. Wir betrachten also die Bewegungsgleichung

für die aufgelöste Variable x in Anlehnung an die Gleichung (3.16). In 0. Ordnung liefert die "Optimal Prediction" Methode den Markov'schen Term:

$$\mathcal{R} = e^{tL}PLx, \quad (4.3)$$

wobei L der Liouville'sche lineare Differentialoperator

$$L = -xy\frac{\partial}{\partial x} + (-y + x^2)\frac{\partial}{\partial y}$$

und P der Projektionsoperator auf den Raum der Funktionen ausschließlich von x sind.

Speziell nimmt in unserem Fall der Markov'sche Term folgende Gestalt an:

$$PLx = P(-xy) = -x \langle y \rangle$$

Dieser Term entspricht der 0. Ordnung der Lösung $y = y_0 e^{-t}$ der homogenen Gleichung $\dot{y} = -y$.

Unter $\langle y \rangle$ versteht man nun nicht mehr den Erwartungswert von y bezüglich einer gewissen Wahrscheinlichkeitsdichte, sei x gegeben, sondern einfach den Anfangswert der Variable y , also $\langle y \rangle = y_0$.

Der "finite-time memory"-Term wird durch die allgemeine Formel bestimmt:

$$\int_0^t e^{(t-s)L}K(s,x)ds$$

wobei $K_j(s,x)$ wie im Abschnitt (3.2) definiert ist:

$$K(s,x) = PLF(s,x) = PL e^{sQL}QLx$$

In einem ersten Schritt bestimmen wir die sogenannte "linearized memory"-Approximation, die im Paragraphen (3.3) im allgemeinen Fall abgeleitet wurde.

Die Anwendung der ersten zwei Operatoren auf der rechten Seite der vorangegangenen Gleichung liefert

$$QLx = (1 - P)Lx = -(1 - P)xy = -xy + xy_0,$$

anschließend wird der Operator PL angewandt:

$$\begin{aligned}
 PL[-xy + xy_0] &= Pxy^2 - P(-y + x^2)x + P(-xyy_0) \\
 &= x(y^2)_0 + xy_0 - x^3 - xy_0^2 \\
 &= x((y^2)_0 - y_0^2) + xy_0 - x^3.
 \end{aligned}$$

Daraus folgt die "Optimal Prediction" für die kollektive Variable x (im Sinne von relevanter Variable) in der "linearized memory"-Approximation und bei Berücksichtigung einer einzelnen Trajektorie ($(y^2)_0 = y_0^2$):

$$\dot{x} = -x[y_0 - y_0 t] - x^3 t$$

Wir erkennen die t -Linearität der Gleichung (3.22). Man kann diese Gleichung auch umschreiben in

$$\dot{x} = -xy_0(1 - t) - x^3 t \quad (4.4)$$

Die Abbildung (4.1) zeigt den Vergleich zwischen der direkten Integration von x ausgehend vom ursprünglichen System und der "linearized memory"-Approximation. Daran ist gut zu erkennen, daß die "linearized memory"-Approximation für kurze Zeiten gute Ergebnisse liefert; allerdings muß sie in der Zeit begrenzt sein, was an der Abbildung bestätigt wird, da sie einem approximierten Integranden entspricht, wie er aus dem Abschneiden der Entwicklung in der Form der Gleichung (3.21) folgt. Für spätere Zeiten erwarten wir einen Integranden der Form:

$$e^{-Ls} K_j(s, x),$$

bei dem die s -Abhängigkeit nicht vernachlässigt wird.

Der lineare Term in t ist deswegen zu ersetzen durch eine allgemeinere Funktion von t :

$$f(t) = 1 - e^{-t},$$

von der nur der erste Term bei der "linearized memory"-Approximation behalten wird; i. e. der "finite-time memory"-Term wird dadurch

$$\int_0^t ds e^{(t-s)L} K_j(s, x) \approx \{xy_0 - x^3\} f(t)$$

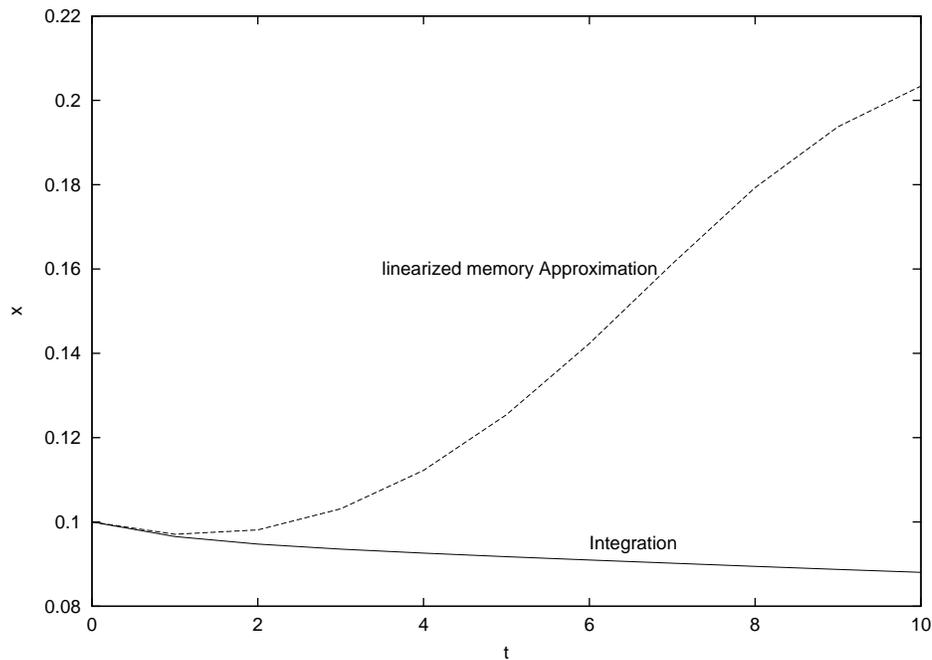


Abbildung 4.1: Vergleich für die Zeitentwicklung für x zwischen der direkten Integration und der “linearized memory”-Approximation

Und schließlich sieht die reduzierte Gleichung für x folgendermaßen aus:

$$\dot{x} \approx -xy_0 \{1 - f(t)\} - x^3 f(t),$$

oder, wenn man für $f(t)$ den Ausdruck $(1 - e^{-t})$ einsetzt:

$$\dot{x} \approx -xy_0 e^{-t} - x^3 (1 - e^{-t}) \quad (4.5)$$

An dieser letzten Gleichung wird klar, was der Unterschied zwischen der “Optimal Prediction” Methode und der Reduktion über die Zentrumsmannigfaltigkeit ist. Für $t \rightarrow \infty$ liefern sie zwar asymptotisch die gleichen Ausdrücke, dann ist nämlich in der Gleichung (4.5)

$$e^{-t} \rightarrow 0 \quad \text{und} \quad \dot{x} \approx -x^3 \quad (4.6)$$

Man erkennt das Ergebnis aus der Zentrumsmannigfaltigkeitsreduktion (ZMR) der Gleichung (4.2), die in der Abbildung (4.2) wiedergegeben ist. Allerdings bietet die “Optimal Prediction”

Methode im Unterschied zu dieser Reduktion den großen Vorteil, daß sie die Anfangsbedingungen in der Form des Terms $y_0 e^{-t}$ mitberücksichtigt; wie im ersten Teil dieser Arbeit schon erwähnt, findet man hier wieder das Abklingen des Einflusses der Anfangsbedingung in Form einer mit der Zeit abnehmenden Exponentialfunktion. Und tatsächlich wird an der Abbildung (4.3) deutlich, daß die “Optimal Prediction” Methode über die ganze Zeitskala perfekt mit der integrierten Kurve übereinstimmt. Die beiden Kurven sind praktisch nicht auseinanderzuhalten, während die Zentrumsmannigfaltigkeitsreduktion anfangs nicht den Effekt der spezifischen Anfangsdaten miteinbezieht und daher erst für große Zeitskalen gute Übereinstimmung mit der Integration liefert.

Bei einem Problem der Form (4.1) eignet sich also die “Optimal Prediction” Methode hervorragend, um die Zeitentwicklung der dominierenden Koordinate vorherzusagen. Allerdings werden wir an dem Gegenbeispiel des L-H-Übergangs im folgenden Abschnitt sehen, daß die Methode hier deswegen so gut funktioniert, da man eine einzige gedämpfte Mode hat, und diese eliminiert wird; daraus folgt eine gute Vorhersage für die dominierende Mode, was aber nicht mehr zwangsläufig der Fall ist, wenn beispielsweise zwei Moden gleichzeitig gedämpft werden.

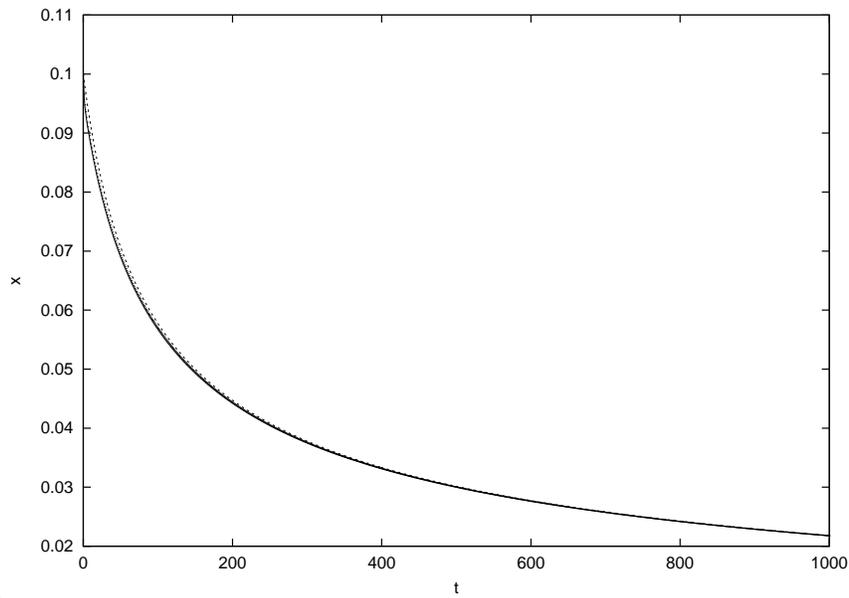


Abbildung 4.2: Globaler Vergleich für die Zeitentwicklung von x zwischen der direkten Integration, der ZMR und der "finite-time memory" Approximation der "Optimal Prediction" Methode: die drei Kurven sind asymptotisch nicht auseinanderzuhalten.

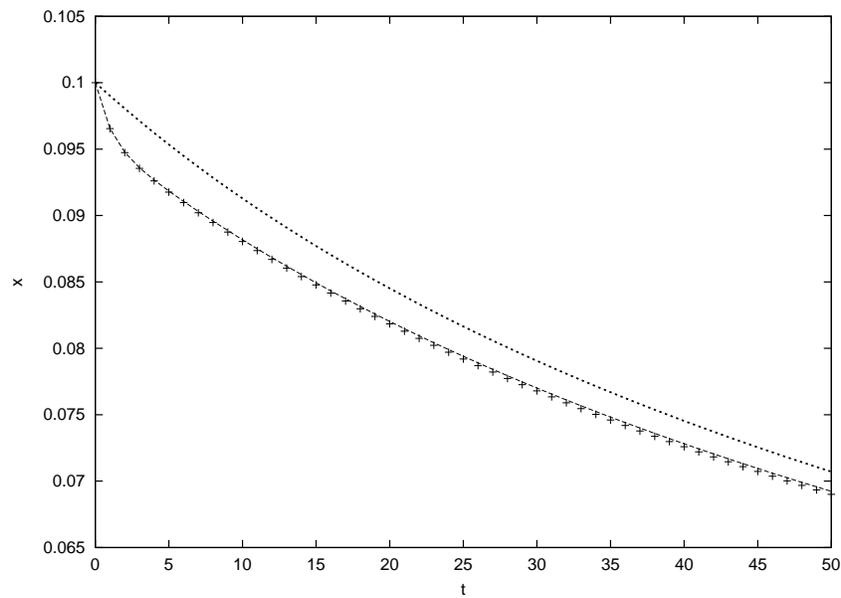


Abbildung 4.3: Vergleich für die Zeitentwicklung von x für kleine Zeitskalen zwischen der direkten Integration (Symbole), der ZMR (gepunktete Kurve) und der "Optimal Prediction" Methode mit "finite time memory" Approximation (feine gestrichelte Kurve): die Ergebnisse der "Optimal Prediction" sind praktisch mit denjenigen der Integration nicht auseinanderzuhalten, während die ZMT zwar das qualitative Verhalten von x wiedergibt, dennoch numerisch schlechtere Ergebnisse liefert.

5. L-H-Übergang

Gegenstand vieler Untersuchungen der Fusionsforschung ist der Einschluß von Hochtemperatur-Plasmen in magnetischen Feldern, wie er zum Beispiel in Tokamaks oder Stellaratoren geschieht. Er wird in der L-Mode (“low confinement”) durch hohe Teilchen- und Energieverluste eingeschränkt, die man auf Fluktuationen aufgrund von Instabilitäten zurückführt [5] und die zu Belastungsproblemen der Wände führen [23]. Da einige aus dem Experiment bestimmte Transportkoeffizienten um Größenordnungen über den Vorhersagen der (neo)klassischen Theorie liegen, spricht man von “anomalem Transport”. Die Energie-Einschlußzeit ist dabei relativ niedrig.

Im Jahre 1982 wurde im Tokamak-Experiment ASDEX (“axisymmetric divertor experiment”) durch starke Neutralteilchen-Zusatzheizung der spontane Übergang in ein neues Einschluß - Regime beobachtet [51], in die sogenannte H-Mode, die eine ungefähr doppelte Energie-Einschlußzeit bei deutlich reduziertem Teilchen- und Wärmetransport ermöglicht [5, 3]. Seitdem konnte die H-Mode in weiteren Tokamaks und Stellaratoren erreicht werden und der entscheidende Übergang von der L- in die H-Mode bei eingeschlossenen Hochtemperatur-Plasmen dadurch unabhängig von der experimentellen Struktur beschrieben werden [3, 29]: er gehört zu den sogenannten Bifurkationsproblemen. Dabei werden, wenn ein bestimmter Bifurkationsparameter einen kritischen Wert überschreitet, einzelne lineare Moden des Systems instabil und wachsen an. Es entstehen neue Strukturen. Wichtig wird dabei die Untersuchung der linearen Instabilität, deren Wechselspiel mit nichtlinearer Sättigung und die nichtlineare Wechselwirkung instabiler Moden untereinander. Im Falle des L-H-Übergangs hat die Überschreitung eines bestimmten Schwellwertes für die Heizleistung eine zeitlich verzögerte Anregung einer radialen Komponente des elektrischen Feldes zur Konsequenz; sie wird im Bereich des Plasmarandes negativer [5]. Über die die Plasmaströmung dominierende $\vec{E} \times \vec{B}$ -Drift bewirkt dieses radial abfallende Feld eine verscherte poloidale Rotationsbewegung. Die Fluktuationen, die für den anomalen Transport verantwortlich sind, verschwinden gleichzeitig und es entsteht am Rande des Plasmas eine etwa 1cm breite Transportbarriere, die durch starke Dichte- und Temperatur-

gradienten gekennzeichnet ist. Auf einer viel größeren Zeitskala verbessert sich der Einschluß im Inneren des Plasmas.

Die entscheidende Rolle des radial abfallenden elektrischen Feldes wurde seitdem gezielt ausgenutzt, um über *edge biasing*, d. h. über Erzeugung radialer Felder in der Randschicht mittels Plasmasonden den H-Zustand zu erreichen [3, 52].

Von großer Bedeutung für den L-H-Übergang ist also die Dynamik des Plasmas in Randnähe; die verschiedenen theoretischen Modelle zielen darauf hin, die experimentellen Ergebnisse zu erklären; es wurden verschiedene Situationen beobachtet: knapp oberhalb des Schwellwertes können zum Beispiel sogenannte *dithering cycles* auftreten, und zwar eine Folge von periodischen L-H-L Übergängen aufgrund von Hysterese-Effekte.

Weiterhin werden Oszillationen während des Betriebs der H-Mode, die *edge-localized modes* (ELMs), beobachtet, eine weitere Bifurkation also, die durch verschiedene Instabilitäten am Plasmarand bewirkt wird. Grundsätzlich unterscheidet man zwischen den ELMs vom Typ I, bei denen die Frequenz mit steigender Heizleistung wächst und den ELMs vom Typ III, bei denen das Gegenteil der Fall ist. Während der Oszillationen vom Typ I befindet sich das Plasma länger im H-Zustand. Der zeitlich gemittelte Transport ist dabei geringer als während der Oszillationen vom Typ III, die bis auf vorangegangene Oszillationen der Magnetfeldstörung sich nicht von den oben erwähnten *dithering cycles* unterscheiden [5].

Zur Beschreibung des L-H-Übergangs wurden schon verschiedene Modelle entwickelt; für eine detaillierte Aufzählung wird hier auf [5] verwiesen. Bei einer ganzen Reihe von Modellen wurde eine qualitative Beschreibung des Übergangs und dessen Zusammenhangs mit der Anregung des radialen elektrischen Feldes möglich, dennoch mußten dabei *ad hoc* Ausdrücke zu Diffusion, Wärmeleitung und Teilchenverlusten eingeführt werden. Weitere Modelle gehen von einer magnetohydrodynamischen Beschreibung des Plasmas aus. Der Druckgradient am Plasmarand führt in Kombination mit der Krümmung des Magnetfeldes zu Instabilitäten; daraus entwickeln sich turbulente Fluktuationen, die unter gewissen Bedingungen die verscherte poloidale Bewegung anregen.

Im folgenden werden wir an ein bereits entwickeltes Modell anschließen [6], das die dreidimensionalen resistiven Ballooning-Fluktuationen im Tokamak beschreibt.

5.1. Modell

Die Methode, die in [6] dargestellt wird, ist eine allgemeine Methode zur Untersuchung eines nichtlinearen dynamischen Systems in der Nähe des Einsatzpunktes einer Instabilität; sie kombiniert eine Zerlegung in eine Karhunen-Loeve-Modenbasis und eine Galerkin-Projektion, um daraus ein reduziertes Modell von wenigen gewöhnlichen Differentialgleichungen zu bekommen.

Wir gehen von einem reduzierten magnetohydrodynamischen (MHD) Modell des Plasmas aus; dabei werden die resistiven MHD Gleichungen unter der Annahme berücksichtigt, daß die Fluktuationen approximativ entlang der magnetischen Feldlinien erfolgen. Das Modell ermöglicht die Beschreibung der resistiven Ballooning-Fluktuationen, bzw. der ‘‘Ballooning’’-Instabilität, die aus der Kombination des Druckgradienten mit der toroidalen Krümmung der magnetischen Feldlinien resultiert und durch ihre poloidale Lokalisierung gekennzeichnet ist. Als lokalisierte Moden wachsen die ‘‘Ballooning-Moden’’ auf der stabilitätsmäßig ungünstigen Außenseite des Plasmas [44]. Unter bestimmten Bedingungen (Plasmadruck $p \ll \frac{B^2}{2\mu_0}$ und inverses Aspektverhältnis $\frac{a}{R} \ll 1$) beschränkt sich das Modell auf zwei Feldgleichungen für die Vortizität- und Druckfluktuationen, deren normalisierte Form folgende ist:

$$\begin{cases} \frac{d}{dt} \nabla_{\perp}^2 \phi = -\nabla_{\parallel}^2 \phi - Gp + \nu \nabla_{\perp}^4 \phi \\ \frac{dp}{dt} = -\frac{\partial \phi}{\partial y} + \chi_{\parallel} \nabla_{\parallel}^2 p + \chi_{\perp} \nabla_{\perp}^2 p, \end{cases} \quad (5.1)$$

wobei ϕ das normalisierte elektrostatische Potential ist [6].

In diesem Modell bleiben diamagnetische Effekte und magnetische Fluktuationen unberücksichtigt. Dennoch ermöglicht es einen Einblick in die dynamischen Grundprozesse am Plasmarand, und zwar eine qualitative Beschreibung der beobachteten physikalischen Phänomene und die Erforschung der Bifurkationen, die das System im nichtlinearen Regime durchläuft: der L-H-Übergang und der Übergang zu den ELMs können charakterisiert werden.

Diese Bifurkationen wurden auch in [7] anhand eines zweidimensionalen MHD Modells untersucht, bei dem eine Zentrumsmannigfaltigkeits-Approximation in der Nähe des Einsatzpunktes der Instabilität und eine sogenannte “inertial manifold”-Approximation [30] jenseits dieses Bereichs angewandt wurden. Dabei wurde gezeigt, daß eine Zentrumsmannigfaltigkeit-Approximation für eine einzige marginale Mode nicht in der Lage ist, den L-H-Übergang zu beschreiben; aus numerischen Simulationen weiß man, daß dafür mehrere Moden instabil werden sollen. Deshalb wurde das ursprüngliche Modell einer einzigen marginalen Mode erweitert: der Einfluß mehrerer, aber weniger, instabiler Moden wurde in Betracht gezogen. Durch die Einbeziehung eines effektiven Heizungsquellterms konnte zudem die experimentelle Beobachtung von *dithering cycles* und EMLs analytisch nachvollzogen werden [7].

Die Variablen ϕ und p werden in den poloidalen und toroidalen Richtungen Fourier-entwickelt; die toroidale Krümmung der Feldlinien bewirkt eine lineare Kopplung der Harmonischen (m, n) mit ihren nächsten Nachbarn $(m \pm 1, n)$, die die Bildung sogenannter “globaler Moden” zur Folge hat. Eine “globale” Eigenmode ist also eine Kombination über einem breiten radialen Gebiet von mehreren Fourier-Moden mit gleicher toroidaler Wellenzahl n und verschiedenen poloidalen Wellenzahlen m . Numerische Simulationen ermöglichen dann die Suche nach den zur Beschreibung der Entstehung eines verscherten Flußes geeigneten Karhunen-Loeve-Eigenwerten und -Eigenmoden. Die Fourier-Moden werden dann in dieser Eigenmodenbasis zerlegt; aus der Numerik läßt sich bestimmen, daß die relevanten Moden hierzu trotz der nichtlinearen Wechselwirkungen die linearen globalen Moden sind [6]. Die fünf wichtigsten werden zur Bildung des Modells beibehalten, und zwar die drei instabilsten globalen Moden ϕ_0, ϕ_1, ϕ_2 und die zwei nächsten instabilen Moden $\phi_{\bar{1}}, \phi_{\bar{2}}$. Über eine Galerkin-Projektion läßt sich ein reduziertes System bilden: das Potential ϕ wird als Summe über die fünf behaltene Moden angesetzt:

$$\phi = a_0\phi_0 + a_1\phi_1 + a_{\bar{1}}\phi_{\bar{1}} + a_2\phi_2 + a_{\bar{2}}\phi_{\bar{2}} \quad (5.2)$$

Durch Multiplikation mit jeder Mode ϕ_i entstehen dann Gleichungen für die Amplituden a_i ; aus der Erkenntnis aus numerischen Simulationen, daß das Gewicht der Moden mit zunehmender Modenzahl stark abnimmt, folgt die Skalierung $a_0, a_1, a_{\bar{1}} \sim \varepsilon$ und $a_2, a_{\bar{2}} \sim \varepsilon^2$; eine weitere Re-

duzierung dieser fünf Gleichungen auf drei erfolgt dadurch, daß das bekannte Argument aus der Zentrumsmannigfaltigkeitsreduktion benutzt wird, nach dem die Dynamik der stabilen Moden durch diejenige der dominierenden Moden bestimmt wird. Mit anderen Worten wird angenommen, daß die Moden $a_2, a_{\bar{2}}$ “versklavt” sind. Ihre Amplituden können also durch diejenigen der dominanten Moden $a_0, a_1, a_{\bar{1}}$ ausgedrückt werden. Es folgt bei Vernachlässigung von Termen der Ordnung ε^4 und unter der Voraussetzung, daß die versklavten Moden schnell die von den dominierenden Moden bedingten Werte annehmen, ein vereinfachtes System von drei gekoppelten, gewöhnlichen Differentialgleichungen für die Amplituden, das nach einfachen Umschreibungen [6] auf die folgende Form gebracht werden kann:

$$\begin{cases} \dot{x} = \gamma_0 x - \gamma_0 yz \\ \dot{y} = \gamma_1 y - xz - \nu_1 z^2 y - \nu_2 y^3 \\ \dot{z} = \delta z + xy - \nu_1 y^2 z \end{cases} \quad (5.3)$$

Dabei ist $x \sim \phi_0$, $y \sim \phi_1$ und $z \sim \phi_{\bar{1}}$. Ein möglicher typischer Parametersatz, den wir (P1) nennen, für die Koeffizienten des Systems ist folgender:

$$(P_1) \quad \gamma_0 = -0.1; \gamma_1 = 0.03; \delta = -0.005; \nu_1 = 0.4; \nu_2 = 0.1; \quad (5.4)$$

Wir befinden uns also in dem dissipativen komplizierteren Fall, bei dem zwei Variablen, x und z , gleichzeitig gedämpft werden. Dieses System soll in der Lage sein, die Wechselwirkungsmechanismen der Moden untereinander in Hinblick auf die Fluktuationsreduktion und die Bildung eines verscherten Flußes zu modellieren.

Die Suche nach einer weiteren Reduzierungsmöglichkeit für das Systems (5.3) wirft die Frage auf, ob ein Galerkin-Verfahren anwendbar ist. Unser Ziel ist die Elimination der Variable x ; d. h. wir wollen zu einem reduzierten System für die Variablen y und z gelangen, von dem wir hoffen, daß es in der Lage ist, das physikalische System und insbesondere den L-H-Übergang, bzw. den Übergang in die ELMs zu beschreiben.

Nun würde die Anwendung eines Galerkin-Verfahrens das Abschneiden in der Entwicklung der gesuchten Lösung bei den Moden y und z bedeuten; x wird identisch gleich Null gesetzt und

nicht weiter berücksichtigt. Es folgt das reduzierte System in Galerkin-Approximation:

$$\text{Galerkin-Approximation} \quad \begin{cases} \dot{x} \equiv 0 \\ \dot{y} = \gamma_1 y - \nu_1 z^2 y - \nu_2 y^3 \\ \dot{z} = \delta z - \nu_1 y^2 z \end{cases} \quad (5.5)$$

Interessant ist der Vergleich dieses letzten Systems mit dem reduzierten System, das man bekommt, wenn man annimmt, daß die Variable x nicht mehr dominierend sondern “versklavt” und der Größenordnung ε^2 ist. Dann ist nämlich die Anwendung der Zentrumsmannigfaltigkeitstheorie möglich: aus der Bedingung $\dot{x} = 0$, die der Tatsache entspricht, daß der von den dominierenden Moden y und z bestimmte Wert für x schnell erreicht wird, folgt der Zusammenhang $x \approx yz$. Das ursprüngliche System (5.3) reduziert sich mit Hilfe der Zentrumsmannigfaltigkeitstheorie also auf

$$\text{Zentrumsmannigfaltigkeitstheorie} \quad \begin{cases} \dot{x} = yz \\ \dot{y} = \gamma_1 y - (1 + \nu_1) z^2 y - \nu_2 y^3 \\ \dot{z} = \delta z + (1 - \nu_1) y^2 z \end{cases} \quad (5.6)$$

An dem Vergleich zwischen den Systemen (5.5) und (5.6) wird klar, wieso das Galerkin-Verfahren nicht geeignet ist, um den L-H-Übergang zu beschreiben: schaut man sich beispielsweise die Gleichung für die Variable z an, sieht man, daß mit dem oben genannten Parametersatz (5.4) das Vorzeichen von dem Term $y^2 z$ auf der rechten Seite in beiden Systemen unterschiedlich ist, und sein Absolutbetrag von verschiedenen Größenordnungen sein kann. Während beim Galerkin-Verfahren $(-\nu_1)$ ein negatives Vorzeichen hat, ist bei der Zentrumsmannigfaltigkeitstheorie $(1 - \nu_1) > 0$. Benutzt man einen anderen Parametersatz, bei dem ν_1 noch kleiner ist, beispielsweise $\nu_1 = 0.1$, dann ist der Unterschied zwischen beiden Termen noch größer: man kann sich vergewissern, daß die Terme $-0.1y^2 z$ und $+0.9y^2 z$ einer völlig anderen Dynamik entsprechen. Von der ZMT weiß man aber [22], daß sie den L-H-Übergang gut beschreibt. Es ist also zu erwarten, daß ein einfaches Abschneiden der gesuchten Lösung nach endlich vielen Moden, wie es beim Galerkin-Verfahren üblich ist, für dieses Problem keine guten Ergebnisse liefert.

Das System (5.6), das aus der Reduktion über die Zentrumsmannigfaltigkeitstheorie folgt, ist dem zweidimensionalen Ansatz von Diamond *et al.* äquivalent, der im Falle starker Turbulenz aus einer statistischen Analysis folgt [22]. Dabei wurde angenommen, daß die Bifurkation vom L- in den H-Zustand stattfindet, wenn die Turbulenz stark genug ist, damit die Reynolds-Spannung die Dämpfung des $\vec{E} \times \vec{B}$ -Flußes überwindet. Ein Modell für den Übergang unter gewissen Voraussetzungen wurde in Form von zwei gekoppelten nichtlinearen Envelope-Gleichungen für das Dichtefluktuationsniveau E und den verscherten Fluß U hergeleitet:

$$\begin{cases} \frac{1}{2} \frac{dE}{dt} = \gamma_0 E - \alpha_1 E^2 - \alpha_2 U E \\ \frac{1}{2} \frac{dU}{dt} = -\mu U + \alpha_3 U E \end{cases} \quad (5.7)$$

Dabei sind die drei Koeffizienten α_1 , α_2 und α_3 nicht unabhängig voneinander: die Energieerhaltung fordert $\alpha_2 \propto \alpha_3$. Die Gleichungssysteme (5.6) und (5.7) sind äquivalent, was an folgender Variablentransformation klar wird: $E \rightarrow \sqrt{y}$, $U \rightarrow \sqrt{z}$.

Bei Skalierung der Zeit und der Variablen E und U kann man zeigen, daß das Modell (5.7) nur von zwei dimensionslosen Parametern abhängt, unter denen ein sogenannter Kontrollparameter, γ_0 , der die Anwachsrate der Instabilität wiedergibt, zur Untersuchung der Bifurkation gewählt werden kann [22].

Abgesehen von der trivialen Lösung besitzt das System zwei Fixpunkte, die dem L-, bzw. H-Zustand entsprechen. Die L-Mode ist für $\gamma_0 < \alpha_1 \mu / \alpha_3$ stabil, während die H-Mode für $\gamma_0 > \alpha_1 \mu / \alpha_3$ stabil ist. Der Übergang von der L- in die H-Mode erfolgt für $\gamma_0 = \alpha_1 \mu / \alpha_3$, was bedeutet, daß die Energie, die in das System hineingepumpt wird, größer als die Dämpfungsrate des Flußes ist, so daß Energie in dem verscherten Fluß gespeichert werden kann. Dies führt zu einem Schwellkriterium für die Heizleistung, das den Übergang möglich macht. Die Gleichungen des Systems (5.7) erfüllen das Kolmogorov-Theorem; deren Lösungen sind entweder ein stabiler Fixpunkt oder ein nichtlinear stabiler Grenzzyklus [22]. Eine lokale Stabilitätsanalyse bekommt man über Linearisierung um die zwei Fixpunkte L und H. Die Stabilitätsgleichung besitzt zwei Lösungen: die eine entspricht der Fluktuation, die andere dem Fluß. Bei dem L-Zustand entsprechen beide Wurzel gedämpften Moden mit Abklingraten $\gamma = -(\mu - \alpha_3 \gamma_0 / \alpha_1)$

und $\gamma = -\gamma_0$. Bei dem H-Zustand und weit vom kritischen Punkt sind beide Moden oszillatorisch und gedämpft. In der Nähe des kritischen Punktes dagegen ist die Frequenz dieser Oszillationen gleich Null, die Dämpfungsrate ist klein für den Fluß und groß für die Fluktuationen, der Übergang ist glatt, was der glatten superkritischen, experimentell beobachteten, für die “*dithering Mode*” charakteristischen Bifurkation entspricht [22].

Daß in der Nähe des kritischen Punktes die Dämpfungsrate für den Fluß klein und für die Fluktuationen groß ist, legt die Vermutung nahe, daß die Fluktuationen gegenüber dem Fluß als versklavt betrachtet werden können; so kann das dynamische Modell (5.7) über Mittelung über schnelle Fluktuationsmoden auf eine einzige Gleichung reduziert werden, die dem Landau-Modell eines Phasenübergangs zweiter Ordnung äquivalent ist.

Der L-H-Übergang kann also als ein Phasenübergang zweiter Ordnung betrachtet werden. Der poloidale Flußgradient spielt die Rolle des Ordnungsparameters. Die L-Mode, die durch zufällig verteilte turbulente Wirbel charakterisiert ist, ist der ungeordnete Zustand, während die H-Mode, bei der der globale verscherte Fluß die zufällige Konvektion überwindet, der geordnete Zustand ist. Der Kontrollparameter für die Bifurkation spielt die Rolle der Temperatur im Phasenübergang, das Bifurkationskriterium entspricht der kritischen Temperatur [22].

Mit dem Modell (5.7), bzw. (5.6) kann zwar der L-H-Übergang gut beschrieben werden, dennoch können die Instabilität des H-Zustandes und dessen Entwicklung in oszillierende Zustände nicht reproduziert werden.

Im Vergleich zu dieser Herangehensweise, der die Zentrumsmannigfaltigkeitstheorie zugrundeliegt, wollen wir nun im folgenden untersuchen, inwiefern das System (5.3) weiterhin reduziert werden kann, indem die “Optimal Prediction” Methode angewandt wird. Wir wollen versuchen, anhand von zwei kollektiven Koordinaten, und zwar y und z , und deren von der “Optimal Prediction” vorhergesagten weiteren Zeitentwicklung, das physikalische System zu charakterisieren, in der Hoffnung, daß so eine Beschreibung des L-H-Übergangs, inklusive der Instabilität des H-Zustands und des Übergangs in die ELMs möglich wird. Die Wahl von y und z als kollektive Koordinaten begründen wir *a priori* damit, daß der Parametersatz (P_1) uns nahelegt, daß x die

am weitesten gedämpfte Variable ist und wahrscheinlich für die dynamische Entwicklung des Systems weniger relevant ist; diese Wahl wird *a posteriori* durch richtige Approximation der Dynamik gerechtfertigt.

Wir gehen nun wieder von unserem ursprünglichen System (5.3) aus, und werden zunächst die Fixpunkte des Systems und die dazugehörigen Stabilitätskriterien des L-, bzw. H-Zustandes angeben; (5.3) besitzt zwei nichttriviale Fixpunkte, die erfüllen:

$$\begin{cases} x = yz \\ z(\delta + (1 - v_1)y^2) = 0 \\ y(\gamma_1 - (1 + v_1)z^2 - v_2y^2) = 0 \end{cases}$$

d. h.

$$\begin{cases} x = yz \\ z = 0 \quad \text{oder} \quad y^2 = \frac{\delta}{v_1 - 1} \\ y = 0 \quad \text{oder} \quad z^2(1 + v_1) = -v_2y^2 + \gamma_1 \end{cases}$$

Der erste ist der sogenannte L-Zustand, bei dem kein verscherter Fluß stattfindet:

$$\text{L-Zustand} \quad \begin{cases} x = 0 \\ y = \sqrt{\frac{\gamma_1}{v_2}} \\ z = 0 \end{cases} \quad (5.8)$$

Wenn allerdings die Dissipation unter einen kritischen Wert fällt, so daß

$$\frac{\gamma_1}{v_2} > \frac{-\delta}{1 - v_1}, \quad (5.9)$$

dann ist dieser Zustand instabil; das System geht über in den zweiten stationären Zustand, in den sogenannten H-Zustand, in dem ein nichtverschwindender verscherter Fluß und ein reduziertes Fluktuationsniveau zu beobachten sind. Der H-Zustand ist gegeben durch:

$$\text{H-Zustand} \quad \begin{cases} x = yz \\ y = \sqrt{\frac{\delta}{v_1 - 1}} \\ z = \sqrt{\frac{v_2}{1 + v_1} \left(\frac{\gamma_1}{v_2} - \frac{\delta}{v_1 - 1} \right)} \end{cases} \quad (5.10)$$

Wiederum kann für ein Unterschreiten eines weiteren kritischen Wertes der Dissipation dieser Zustand instabil werden, und zwar wenn:

$$\frac{2v_1^2}{1+v_1} \left(\frac{\gamma_1}{v_2} + \frac{\delta}{1-v_1} \right) > \frac{-\delta}{1-v_1} \quad (5.11)$$

Dann treten ELMs auf, die den Energieeinschluß erschweren [3]. Die Fusionsreaktionen können dadurch gestoppt werden. Das könnte man in der Praxis vermeiden, indem die Dissipation möglichst genau auf das Stabilitätsregime der H-Mode eingestellt wird. Allerdings bewirkt der verbesserte Teilcheneinschluß eine Akkumulation verschiedener Verunreinigungen und Abfallprodukte, die an die Randschicht abgegeben werden müssen. Dies geschieht aber im ELMs-Regime. Daher versucht die sogenannte “dynamische Kontrolle” die Vorteile beider Regimes auszunutzen: der Tokamak funktioniert in einem Parameterbereich, in dem die H-Mode instabil gegenüber der ELMs ist; allerdings sorgt ein geeigneter Kontrollmechanismus für ihre Stabilisierung. Sollte die Verunreinigungskonzentration einen gewissen Schwellwert erreichen, wird dieser Kontrollmechanismus kurzzeitig ausgeschaltet. Die Verunreinigungen werden dann über die ELMs nach außen hin getrieben; der Kontrollmechanismus wird dann wieder eingesetzt und die H-Mode stellt sich erneut für einen solchen Zyklus wieder ein. Dadurch wird die Aufrechterhaltung des H-Mode-Regimes in der Praxis gewährleistet [3].

Im folgenden wird das Prinzip der “Optimal Prediction” Methode angewandt, wobei der “finite-time memory”, wie er im ersten Teil dieser Arbeit dargestellt wurde, miteinbezogen wird. Die so erzeugten Aussagen, von denen man hofft, daß sie den in der Praxis gewünschten L-H-Übergang und die eventuelle Instabilität des H-Zustandes wiedergeben, werden mit den numerischen Simulationen verglichen, die aus der Integration von (5.3) folgen. Es wurde schon bestätigt [6], daß diese numerische Integration die experimentellen Befunde richtig approximiert, und insofern ein gutes Modell für die physikalischen Vorgänge darstellt.

5.2. Modell zur x -Elimination

Wir gehen vom System (5.3) aus

$$\begin{cases} \dot{x} = \gamma_0 x - \gamma_0 yz \\ \dot{y} = \gamma_1 y - xz - v_1 z^2 y - v_2 y^3 \\ \dot{z} = \delta z + xy - v_1 y^2 z \end{cases}$$

und versuchen nun, x anhand der ‘‘Optimal Prediction’’ Methode zu eliminieren, und somit unser System auf zwei Differentialgleichungen f#r y und z zu reduzieren. Wir gehen nun wie im vorigen Abschnitt vor; in 0. Ordnung ‘‘Optimal Prediction’’ dr#ckt sich die Zeitentwicklung durch einen Markov’schen Term #hnlich dem in der Gleichung (4.3) aus:

$$\begin{cases} \dot{y} = e^{tL} P y \\ \dot{z} = e^{tL} P z \end{cases} \quad (5.12)$$

Der Liouville-Operator L ist dabei

$$L = (\gamma_0 x - \gamma_0 yz) \frac{\partial}{\partial x} + (\gamma_1 y - xz - v_1 z^2 y - v_2 y^3) \frac{\partial}{\partial y} + (\delta z + xy - v_1 y^2 z) \frac{\partial}{\partial z}. \quad (5.13)$$

Es folgt in 0. Ordnung ‘‘Optimal Prediction’’:

$$\begin{cases} \dot{y} = \gamma_1 y - x_0 z - v_1 z^2 y - v_2 y^3 \\ \dot{z} = \delta z + x_0 y - v_1 y^2 z \end{cases}$$

Hier ist wie im vorigen Abschnitt auf der rechten Seite der Gleichung keine Mittelung mehr #ber die unaufgel#ste Variable x , sondern deren Anfangswert x_0 . Nun ist der ‘‘finite-time memory’’-Term durch die folgende allgemeine Integralform gegeben:

$$\int_0^t e^{(t-s)L} K_j(s, \hat{x}) ds$$

wobei

$$K_j(s, \hat{x}) = PLF_j(s, x) = PL e^{sQL} QLx_j$$

Als erste Approximation hierzu dient hier wieder die “linearized memory”-Approximation, bei der $s = 0$ ist:

$$\begin{cases} PLQL \ y = \gamma_0 yz^2 - \gamma_0 x_0 z \\ PLQL \ z = -\gamma_0 y^2 z + \gamma_0 x_0 y \end{cases} \quad (5.14)$$

Daher folgt für die “linearized memory”-Approximation:

$$\begin{array}{l} \text{“linearized memory”-} \\ \text{-Approximation} \end{array} \begin{cases} \dot{y} = \gamma_1 y - x_0 z - v_1 z^2 y - v_2 y^3 + t (\gamma_0 yz^2 - \gamma_0 x_0 z) \\ \dot{z} = \delta z + x_0 y - v_1 y^2 z + t (-\gamma_0 y^2 z + \gamma_0 x_0 y), \end{cases} \quad (5.15)$$

d. h. ein System von zwei Gleichungen für \dot{y}, \dot{z} , die linear in t sind.

Bei der Berechnung des vollen “finite-time memory”-Terms wird die Klammer auf der rechten Seite von (5.15) nicht mehr mit t multipliziert, sondern mit

$$\frac{1}{|\gamma_0|} (1 - e^{-|\gamma_0|t}),$$

um das asymptotische Verhalten richtig wiederzugeben. Der Vorfaktor $1/|\gamma_0|$ vor der Klammer und in der Exponentialfunktion findet seine Rechtfertigung in der Tatsache, daß x wie $\dot{x} \approx \gamma_0 x$ gedämpft wird.

Dies ist äquivalent zu folgender Aussage: man ersetzt in den Gleichungen für \dot{y} und \dot{z} x folgendermaßen

$$x \approx x_0 e^{-|\gamma_0|t} + (1 - e^{-|\gamma_0|t}) yz \quad (5.16)$$

Man erhält dann für die Gleichung der “Optimal Prediction” mit “finite-time memory”

$$\begin{cases} \dot{y} = \gamma_1 y - x_0 z - v_1 z^2 y - v_2 y^3 + \frac{1}{|\gamma_0|} (1 - e^{-|\gamma_0|t}) (\gamma_0 yz^2 - \gamma_0 x_0 z) \\ \dot{z} = \delta z + x_0 y - v_1 y^2 z + \frac{1}{|\gamma_0|} (1 - e^{-|\gamma_0|t}) (-\gamma_0 y^2 z + \gamma_0 x_0 y) \end{cases}$$

Da $|\gamma_0| = -\gamma_0$ folgt die vereinfachte Form:

$$\begin{cases} \dot{y} = \gamma_1 y - x_0 z - v_1 z^2 y - v_2 y^3 + (1 - e^{-|\gamma_0|t}) (-yz^2 + x_0 z) \\ \dot{z} = \delta z + x_0 y - v_1 y^2 z + (1 - e^{-|\gamma_0|t}) (y^2 z - x_0 y) \end{cases} \quad (5.17)$$

oder umgeschrieben:

$$\begin{array}{l} \text{“Optimal Prediction”} \\ \text{mit “finite-time memory”-Term} \end{array} \left\{ \begin{array}{l} \dot{y} = \gamma_1 y - x_0 z e^{-|\gamma_0|t} - v_1 z^2 y - v_2 y^3 - yz^2(1 - e^{-|\gamma_0|t}) \\ \dot{z} = \delta z + x_0 y e^{-|\gamma_0|t} - v_1 y^2 z + y^2 z(1 - e^{-|\gamma_0|t}) \end{array} \right. \quad (5.18)$$

An dem Vergleich zwischen den Systemen (5.3) und (5.18) kann man sich das Prinzip der “Optimal Prediction” Methode noch einmal klar machen: für kleine t wird x mit seinem Anfangswert x_0 gleichgesetzt; dieser Beitrag nimmt aber mit fortschreitender Zeit ab und x tendiert gegen den asymptotischen Wert $x \approx yz$. Im Unterschied zu einer Zentrumsmannigfaltigkeitsreduktion, die für alle Zeiten die Approximation $x \approx yz$ durchführt, berücksichtigt die “Optimal Prediction” Methode also auch den Einfluß der Anfangsbedingungen und dessen Abklingen in der Zeit. Daß die “Optimal Prediction” asymptotisch für $t \rightarrow \infty$ in die Zentrumsmannigfaltigkeitsmethode übergeht, zeigt gleichzeitig auch ihre Einschränkungen: es ist zu erwarten, daß dynamische Bereiche, die von einer Zentrumsmannigfaltigkeitsreduktion angegangen werden können, auch asymptotisch mit Hilfe der “Optimal Prediction” Methode beschrieben werden können; Bereiche, die aber keine Reduktion über die Zentrumsmannigfaltigkeitstheorie zulassen, werden wahrscheinlich auch nicht gut von der “Optimal Prediction” Methode vorausgesagt. Diese Vermutung wird bei der Untersuchung des L-H-Übergangs und des Übergangs vom H-Zustand in Oszillationen bestätigt.

5.3. L-H-Übergang bei anschließender Stabilität des H-Zustandes

Betrachten wir das Ausgangssystem (5.3) mit dem Parametersatz (P_2):

$$\gamma_0 = -0.1; \gamma_1 = 0.03; \delta = -0.08; v_1 = 0.4; v_2 = 0.1; \quad (5.19)$$

Mit diesem Parametersatz ist der L-Zustand instabil, denn das Instabilitätskriterium (5.9) ist erfüllt:

$$\frac{\gamma_1}{v_2} \approx 0.3 > \frac{-\delta}{1 - v_1} \approx 0.1333 \quad (5.20)$$

Das System geht in den H-Zustand über, dessen Stabilität durch folgendes Kriterium (5.11) gezeigt wird:

$$\frac{2v_1^2}{1+v_1} \left(\frac{\gamma_1}{v_2} + \frac{\delta}{1-v_1} \right) \approx 0.038 < \frac{-\delta}{1-v_1} \approx 0.1333 \quad (5.21)$$

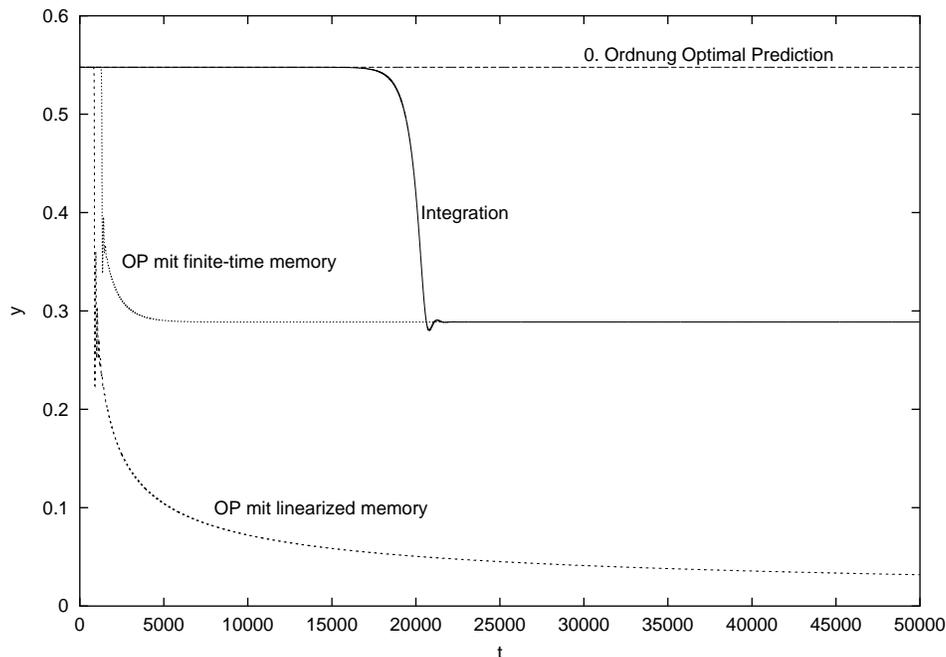


Abbildung 5.1: Zeitentwicklung von y : die Ergebnisse der 0. Ordnung “Optimal Prediction”, der “linearized memory”-Approximation und der “finite-time memory”-Vorhersage werden mit denjenigen der Integration verglichen; die 0. Ordnung “Optimal Prediction” Methode ohne Berücksichtigung des “memory”-Terms gibt die Instabilität des L-Zustandes nicht wieder; die “linearized memory”-Approximation reproduziert zwar diese Instabilität, dennoch ist nach kurzer Zeit die Vorhersage falsch; die “finite-time memory” Approximation reproduziert den Anfangs- und Endzustand gut, allerdings entspricht der Übergang selber nicht der Realität.

Die Abbildungen (5.1) und (5.2) stellen die zeitliche Entwicklung respektive von y und z dar; tatsächlich wird das Langzeitverhalten, d. h. der Übergang vom L- in den H-Zustand und dessen Stabilität, von der vollständigen “Optimal Prediction” Methode, also mit dem “finite-time memory”-Term wiedergegeben, was nicht der Fall bei der Nutzung der 0. Ordnung der Methode oder der “linearized memory”-Approximation ist.

Die 0. Ordnung “Optimal Prediction” bezieht sich auf die Anfangswerte: der L-Zustand wird fälschlicherweise als stabil beschrieben.

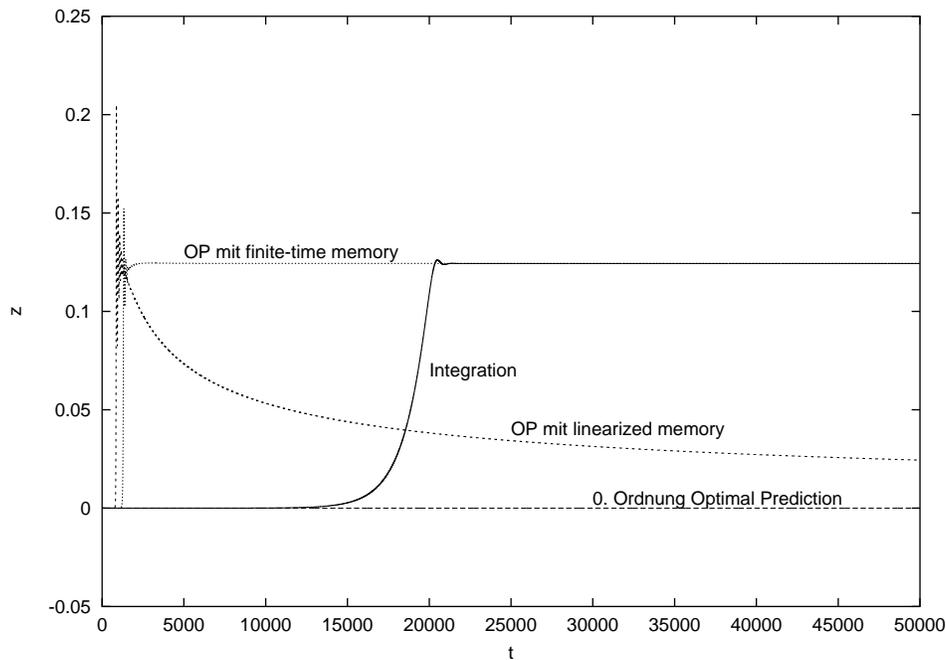


Abbildung 5.2: Zeitentwicklung von z : die Ergebnisse der 0. Ordnung "Optimal Prediction", der "linearized memory"-Approximation und der "finite-time memory"-Vorhersage werden mit denjenigen der Integration verglichen; die 0. Ordnung "Optimal Prediction" Methode ohne Berücksichtigung des "memory"-Terms gibt die Instabilität des L-Zustandes nicht wieder; die "linearized memory"-Approximation reproduziert zwar diese Instabilität, dennoch ist nach kurzer Zeit die Vorhersage falsch; die "finite-time memory" Approximation reproduziert den Anfangs- und Endzustand gut, allerdings entspricht der Übergang selber nicht der Realität.

Die "linearized memory"-Approximation berücksichtigt zwar die weitere Zeitentwicklung des Systems, dennoch nur in erster Ordnung; mit ihr ist die Aussage möglich, nach der der L-Zustand instabil ist und vom System verlassen wird, aber die spätere Zeitentwicklung ist falsch angegeben.

Allein der volle "finite-time memory"-Term ermöglicht die Aussage eines Übergangs in den H-Zustand. Dennoch sind dessen so berechneter Zeitpunkt und Verlauf falsch, was darauf zurückgeführt wird, daß für den Übergang selber die "Optimal Prediction" Methode, wie auch die Zentrumsmannigfaltigkeitsreduktion, nicht benutzt werden können.

Ist man also nur an dem Langzeitverhalten des Systems in diesem Parameterbereich interessiert, ist die "Optimal Prediction" Methode sinnvoll, für eine Beschreibung des Übergangs selber ist

sie hier ungeeignet.

5.4. L-H-Übergang bei anschließender Instabilität des H-Zustandes

Betrachten wir nun das ursprüngliche System (5.3) mit dem Parametersatz (P_1)

$$(P_2) \quad \gamma_0 = -0.1; \gamma_1 = 0.03; \delta = -0.005; \nu_1 = 0.4; \nu_2 = 0.1 \quad (5.22)$$

Der Anfangspunkt unseres Systems entspricht dem L-Zustand. Mit diesem Parametersatz ist er instabil, da

$$\frac{\gamma_1}{\nu_2} = 0.3 > \frac{-\delta}{1 - \nu_1} \approx 0.00833; \quad (5.23)$$

Ausgehend vom L-Zustand geht also das System in den H-Zustand über, der auch instabil ist, wie man an der Stabilitätsbedingung (5.11) nachprüfen kann:

$$\frac{2\nu_1^2}{1 + \nu_1} \left(\frac{\gamma_1}{\nu_2} + \frac{\delta}{1 - \nu_1} \right) \approx 0.0667 > \frac{-\delta}{1 - \nu_1} \approx 0.00833; \quad (5.24)$$

d. h. das System oszilliert nun. Die Bewegungsgleichungen (5.3) wurden mit Anfangsbedingungen integriert, die eine infinitesimal kleine Störung des L-Zustandes darstellen.

Da die 0. Ordnung der “Optimal Prediction” “nur” die Anfangsbedingungen berücksichtigt, kann sie hier auch nur das falsche Ergebnis liefern, der L-Zustand wäre stabil. Deswegen wird sie hier nicht weiter betrachtet, genauso wenig, wie die “linearized memory”-Approximation, von der im vorigen Abschnitt gezeigt wurde, daß sie nur das Anfangsverhalten und die Instabilität des L-Zustandes, nicht aber das asymptotische Verhalten reproduziert.

Wir konzentrieren uns nun auf die “Optimal Prediction” mit “finite-time memory”. Im vorigen Abschnitt hat man gesehen, daß sie im Falle des L-H-Übergangs und bei Stabilität des H-Zustandes das asymptotische Verhalten richtig wiedergibt. Kann diese Methode nun auch die Instabilität des H-Zustandes und den Übergang in oszillierende Zustände reproduzieren?

Die Zeitentwicklung für die Variablen y und z ist in den Abbildungen (5.3) und (5.4) zu erkennen; die dicke durchgezogene Linie ist die Integrationskurve; sie zeigt den Übergang vom L-Zustand über den H-Zustand in Oszillationen, deren Frequenz und Amplitude über die beobachtete Zeitskala als konstant betrachtet werden können. Die dünne gestrichelte Kurve ist die “Optimal Prediction” in “finite-time memory”. Wir sehen, daß die Instabilität des L-Zustandes und der Übergang in oszillierende Zustände reproduziert werden. Die Frequenz und Amplitude dieser Oszillationen stimmen anfangs mit denjenigen der Integration überein. Über eine große Zeitskala, von der anzumerken ist, daß sie viel größer ist, als die entsprechenden Ergebnisse, die für Hamiltonsche Systeme im ersten Teil dieser Arbeit erzielt worden sind, bis $t \approx 1300$, kann man sagen, daß die “Optimal Prediction” das qualitative Verhalten des Systems sehr gut approximiert. Dennoch sind die erzeugten Oszillationen eigentlich gedämpft und die Amplitude im Gegenteil zur Integration nicht konstant: asymptotisch pendelt sich die Lösung der “Optimal Prediction” auf den H-Zustand ein, dessen Instabilität also nicht wiedergegeben wird.

Dieser letzte Punkt entspricht der Tatsache, daß die “Optimal Prediction” mit dem “finite-time memory”-Term asymptotisch für große Zeiten t der Zentrumsmannigfaltigkeitsreduktion entspricht; d. h. asymptotisch kann man mit ihr nur in der Nähe eines Gleichgewichtszustandes eine gute Approximation erwarten, der auch eine Reduktion über die Zentrumsmannigfaltigkeit erlaubt. Daher ist es nicht verwunderlich, daß die Oszillationen mit der “Optimal Prediction” Methode nicht reproduzierbar sind. Mathematisch findet diese Tatsache in folgender Überlegung ihre Rechtfertigung: Betrachten wir die erste Gleichung des ursprünglichen Systems (5.3)

$$\dot{x} = \gamma_0 x - \gamma_0 y z;$$

es handelt sich dabei um eine inhomogene Differentialgleichung 1. Ordnung. Wie im vorigen Abschnitt können wir nun die allgemeine Form der Lösung als Summe von der allgemeinen Lösung der homogenen Gleichung $\dot{x} = \gamma_0 x$, i. e.

$$x = x_0 e^{\gamma_0 t} = x_0 e^{-|\gamma_0| t}$$

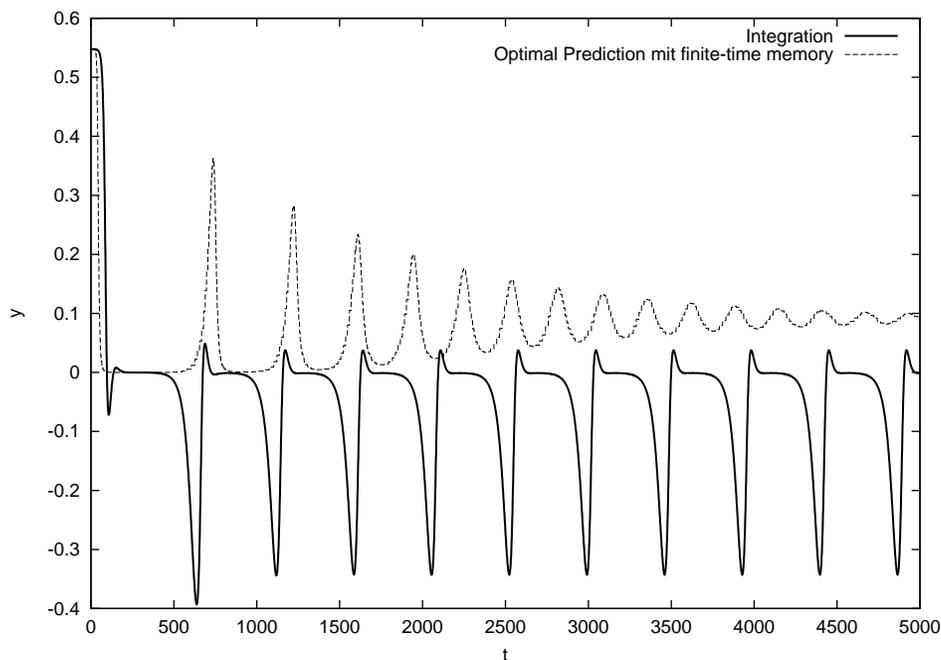


Abbildung 5.3: Zeitentwicklung von y : die Ergebnisse der "Optimal Prediction" mit "finite-time memory" (dünne gestrichelte Kurve) werden mit denjenigen der Integration (dicke durchgezogene Kurve) verglichen; man sieht, daß die anfänglichen Oszillationen qualitativ sehr gut wiedergegeben werden; deren Frequenz ist bei beiden Kurven gleich. Dennoch geht die von der "Optimal Prediction" Methode vorhergesagte Kurve mit weiter fortschreitender Zeit in den H-Zustand über, dessen Instabilität von der Methode nicht reproduziert wird.

und einer speziellen Lösung der inhomogenen Gleichung:

$$\int_0^t ds (-\gamma_0 yz)(s) e^{-|\gamma_0|t + |\gamma_0|s} \quad (5.25)$$

angeben.

In der Nähe eines Fixpunktes kann man folgende Approximation durchführen:

$$yz \approx \text{Konstante},$$

d. h. das Produkt yz im Integranden (5.25) kann vorgezogen werden; das Integral kann nun einfach ausgeführt werden:

$$(-\gamma_0 yz) e^{-|\gamma_0|t} \int_0^t ds e^{|\gamma_0|s},$$

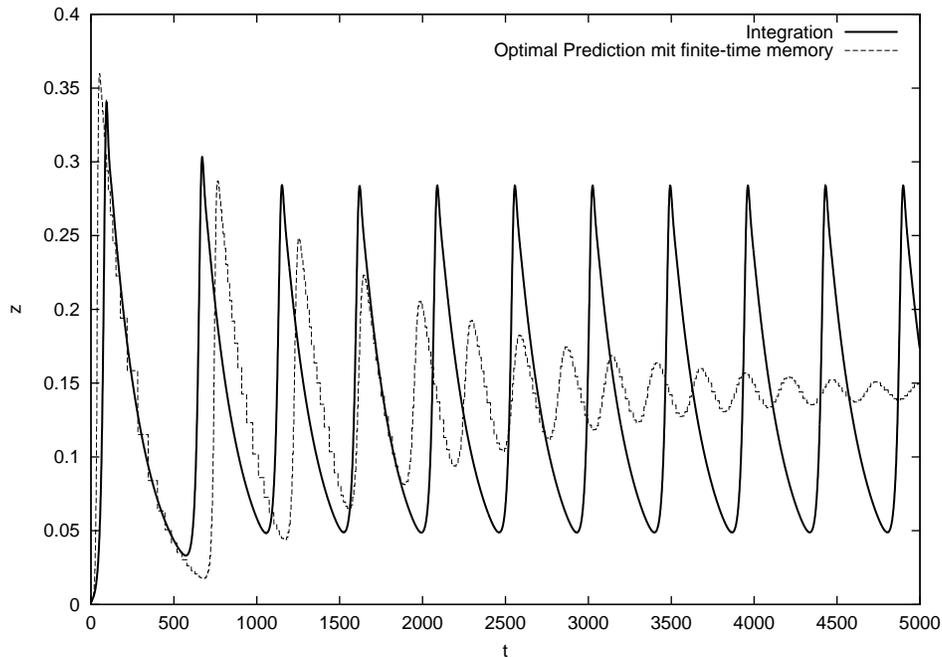


Abbildung 5.4: Zeitentwicklung von z : die Ergebnisse der “Optimal Prediction” mit “finite-time memory” (dünne gestrichelte Kurve) werden mit denjenigen der Integration (dicke durchgezogene Kurve) verglichen; man sieht, daß die anfänglichen Oszillationen qualitativ sehr gut wiedergegeben werden; deren Frequenz ist bei beiden Kurven gleich. Dennoch geht die von der “Optimal Prediction” Methode vorhergesagte Kurve mit weiter fortschreitender Zeit in den H-Zustand über, dessen Instabilität von der Methode nicht reproduziert wird.

i. e.

$$(-\gamma_0 y z) e^{-|\gamma_0| t} \frac{1}{|\gamma_0|} (1 - e^{-|\gamma_0| t})$$

Daraus folgt als Lösung für x :

$$x = x_0 e^{-|\gamma_0| t} - \frac{1}{|\gamma_0|} \gamma_0 y z (1 - e^{-|\gamma_0| t})$$

und, da $\gamma_0 = -|\gamma_0|$ schließlich:

$$x = x_0 e^{-|\gamma_0| t} + y z (1 - e^{-|\gamma_0| t})$$

Dies ist genau der Ausdruck (5.16), der der “Optimal Prediction” Methode mit “finite-time memory”-Term zugrundeliegt; man sieht also, daß er nur zulässig ist, wenn die Approximation

$yz \approx \text{Konstante}$ erlaubt ist, die wiederum einem Gleichgewichtszustand oder Fixpunkt des ursprünglichen Systems entspricht.

5.5. Einfluß und Wichtigkeit der Parameterwerte

Der Parameter γ_0 hat keinen Einfluß auf die Lage und Stabilitätskriterien der L- und H-Zustände. Davon überzeugt man sich, wenn man sich die erste Gleichung anschaut:

$$\dot{x} = \gamma_0 x - \gamma_0 yz$$

Die Fixpunkte L und H erfüllen nämlich die Bedingung $x = yz$, der Parameter γ_0 spielt dabei keine Rolle mehr, weiterhin nicht in den Gleichungen für \dot{y} und \dot{z} , und in den Stabilitätskriterien, die man nach Linearisierung um die Fixpunkte, bzw. mit der Methode der kleinen Störungen erhält.

Dennoch liegt der “Optimal Prediction” Methode folgendes Prinzip zugrunde: die ausgewählten kollektiven Koordinaten sollen tatsächlich auch diejenigen sein, die maßgeblich für die zeitliche Entwicklung der Dynamik sind. Das Verletzen dieses Prinzips kann zu einem Scheitern der Methode führen, wie im ersten Teil dieser Arbeit an Hamiltonschen Systemen im Falle der quintisch und kubisch nichtlinearen Gleichung gezeigt worden ist. Wenn man aber in der Untersuchung des Systems (5.3) (y, z) als kollektive Koordinaten aussucht, und x eliminiert, z aber die am meisten gedämpfte Mode ist, ist zu erwarten, dass die Methode in den Bereichen versagt, in denen z tatsächlich stark gedämpft wird. Da ist nämlich die Approximation $x \approx yz$ schlecht, die der Methode asymptotisch zugrundeliegt. Deshalb ist eine Verschlechterung der gerade erzielten Ergebnisse zu erwarten, wenn γ_0 deutlich kleiner im Betrag ist als δ , und tatsächlich die z -Mode die am weitesten gedämpfte ist.

Dieses wird nun untersucht, indem man vom schon benutzten Parametersatz (5.4) die vier Pa-

parameter

$$\left\{ \begin{array}{l} \gamma_1 = 0.03 \\ \delta = -0.005 \\ \nu_1 = 0.4 \\ \nu_2 = 0.1, \end{array} \right.$$

unverändert läßt und die Brauchbarkeit der “Optimal Prediction” Methode in Hinsicht auf den L-H-Übergang für verschiedene Werte von γ_0 untersucht. Aus der vorangegangenen Analyse wissen wir, daß der L-Zustand mit diesem Parametersatz instabil ist und, daß das System in den H-Zustand und anschließend wegen dessen Instabilität in Oszillationen übergeht.

- Der Fall $\gamma_0 = -0.1$ wurde im vorigen Abschnitt und anhand der Abbildungen (5.3) und (5.4) schon behandelt. In diesem Falle lieferte die “Optimal Prediction” anfangs eine sehr gute qualitative und sogar quantitative Beschreibung der Dynamik des Systems.

- Betrachten wir jetzt γ_0 eine Größenordnung kleiner: $\gamma_0 = -0.01$

Die Ergebnisse der Abbildungen (5.5) und (5.6) für die Zeitentwicklung von y und z zeigen eine etwas schlechtere Wiedergabe der Oszillationen im Vergleich zum Fall $\gamma_0 = -0.1$; insbesondere stimmt die Oszillationsfrequenz nicht mehr mit der Frequenz aus der Integration überein. Hier sind γ_0 und δ fast der gleichen Größenordnung: $\gamma_0 = 2\delta$, d. h. die Dämpfung der Variable z ist nur unwesentlich geringer als diejenige der Variable x . Daraus folgt, daß die Annahme einer einzigen gedämpften Variable x , die unserer Anwendung der “Optimal Prediction” Methode zugrundeliegt, unzulässig ist. Natürlich stabilisiert auch hier der H-Zustand für größere t , und die Oszillationen werden nicht reproduziert.

- Wählen wir nun $\gamma_0 = -0.001$.

Der Vergleich zwischen der Integration und der “Optimal Prediction” Methode für die Zeitentwicklung von y und z wird in den Abbildungen (5.7) und (5.8) dargestellt. Tatsächlich wird die Vermutung bestätigt, daß die “Optimal Prediction” Methode in der Vorher-

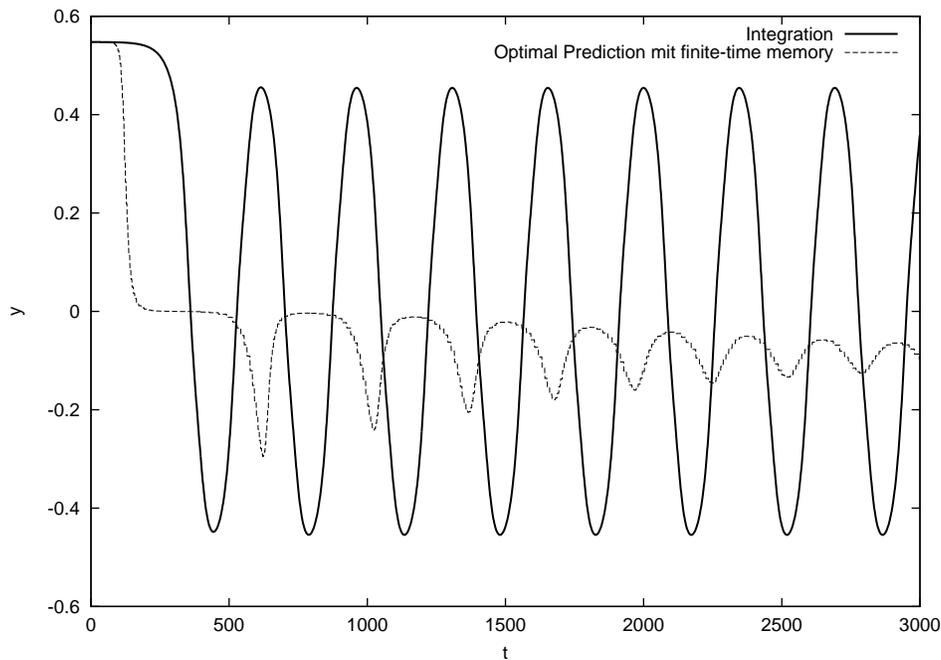


Abbildung 5.5: Zeitentwicklung von y für $\gamma_0 = -0.01$: die dicke durchgezogene Kurve folgt aus der Integration, während die dünne gestrichelte Kurve die Vorhersage der "Optimal Prediction" darstellt.

sage der dynamischen Entwicklung des Systems versagt, wenn im Falle mehrerer Dämpfungen sich unter den kollektiven Koordinaten die am meisten gedämpfte Mode befindet, und eine Mode "eliminiert" wurde, die eher als dominant bezeichnet werden könnte. Hier werden die Art des Übergangs und die Oszillationen nicht mehr gut wiedergegeben. Die Werte für y und z pendeln sich sehr schnell auf die Werte des H-Zustandes ein; da seine Instabilität grundsätzlich von der Methode nicht erfaßt werden kann, wie wir vorhin gesehen haben, kann man sagen, daß bei dieser ungünstigsten Wahl der kollektiven Koordinaten und den gegebenen Dämpfungen die einzige Aussage der "Optimal Prediction" Methode in der Instabilität des L-Zustandes besteht, der von dem physikalischen System verlassen wird; anschließend finden Oszillationen statt, deren Frequenz und Amplitude aber nicht richtig erfaßt werden.

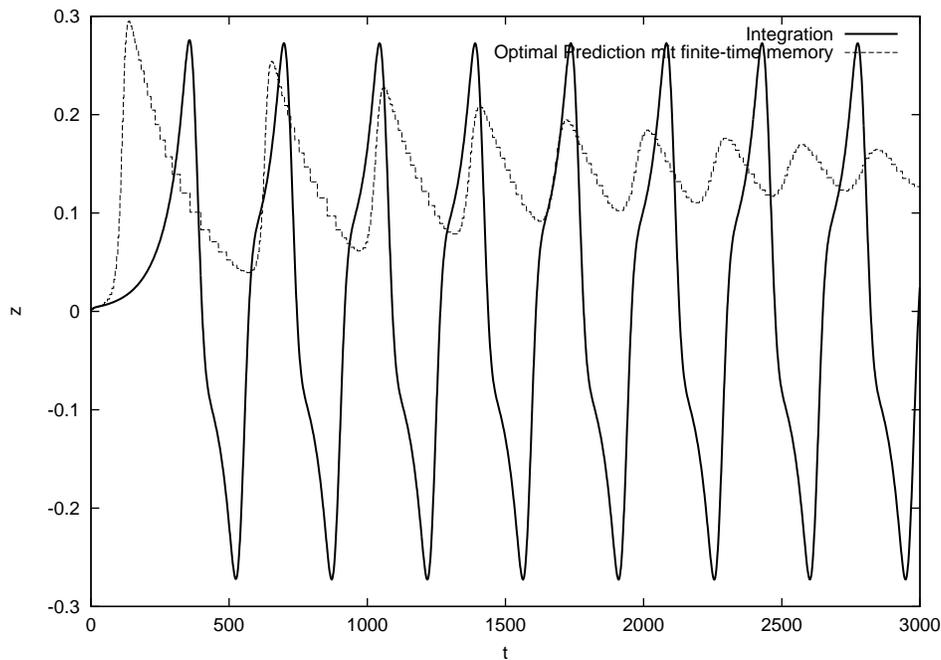


Abbildung 5.6: Zeitentwicklung von z für $\gamma_0 = -0.01$: die dicke durchgezogene Kurve folgt aus der Integration, während die dünne gestrichelte Kurve die Vorhersage der “Optimal Prediction” darstellt.

Abschließend kann man sagen, daß γ_0 zwar keinen Einfluß auf die Stabilität der jeweiligen Zustände L und H hat, dennoch ausschlaggebend für die Art des L-H-Übergangs ist, und deshalb für die Brauchbarkeit der “Optimal Prediction” Methode.

Das Gelingen der Methode hängt davon ab, wie günstig die kollektiven Koordinaten gewählt worden sind. Beim zweidimensionalen mathematischen Modell, das im vorigen Paragraphen untersucht worden ist, waren die Ergebnisse hervorragend, denn man war mit nur einer gedämpften Mode konfrontiert, die auch eliminiert und gegenüber der anderen, einer “kollektiven Koordinate” entsprechend, vernachlässigt werden konnte.

Dennoch haben wir gerade gesehen, daß im Falle mehrerer Dämpfungen, wie beim physikalischen L-H-Übergang in toroidal eingeschlossenen Plasmen, der Auswahl der kollektiven Koordinaten größere Bedeutung zukommt. Damit die “Optimal Prediction” Methode die dynamische Entwicklung des Systems qualitativ und für kurze Zeitskalen sogar quantitativ gut beschreibt, muß unbedingt die am meisten gedämpfte Mode eliminiert werden. Im Falle zwei gleich großer,

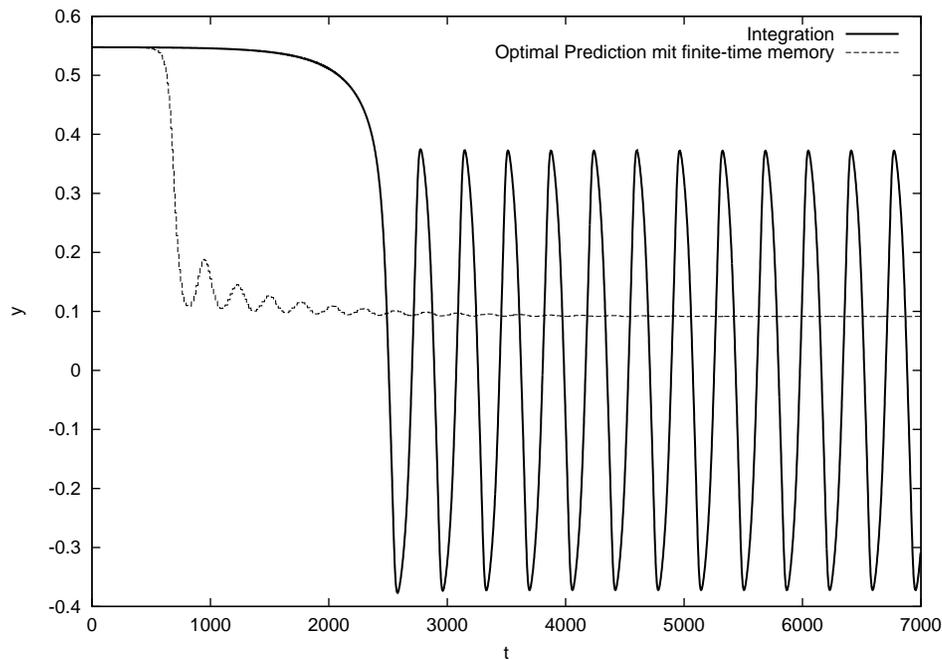


Abbildung 5.7: Zeitentwicklung von y für $\gamma_0 = -0.001$: die dicke durchgezogene Kurve folgt aus der Integration, während die dünne gestrichelte Kurve die Vorhersage der “Optimal Prediction” darstellt.

deshalb konkurrierender Dämpfungen eignet sich die Methode nicht. Je kleiner das Verhältnis $|\gamma_0|/|\delta|$ bei der Untersuchung des Systems (5.3) ist, desto besser wird die Zeitskala des L-H-Übergangs, die Frequenz und Amplitude der anschließenden Oszillationen von der “Optimal Prediction” Methode approximiert.

Mit anderen Worten muß die Frage nach den für die Dynamik des untersuchten Systems relevanten Moden und die dazugehörige Unterteilung in “dominante” und “versklavte” Moden vor Anwendung der Methode geklärt werden.

In diesem Zusammenhang ist das gleiche Anwendungsgebiet der “Optimal Prediction” Methode festzustellen, wie das der Zentrumsmannigfaltigkeitsreduktion. Ihr gegenüber hat aber die “Optimal Prediction” Methode den eindeutigen Vorteil der Berücksichtigung des für kleine Zeiten wichtigen Einflusses der Anfangsbedingungen. Die Zentrumsmannigfaltigkeitsreduktion liefert bei Untersuchung des L-H-Übergangs nur einen “abrupten” Übergang, der der Tatsache entspricht, daß diese Reduktionsmethode nur in der Umgebung beider Fixpunkte, d. h. in der

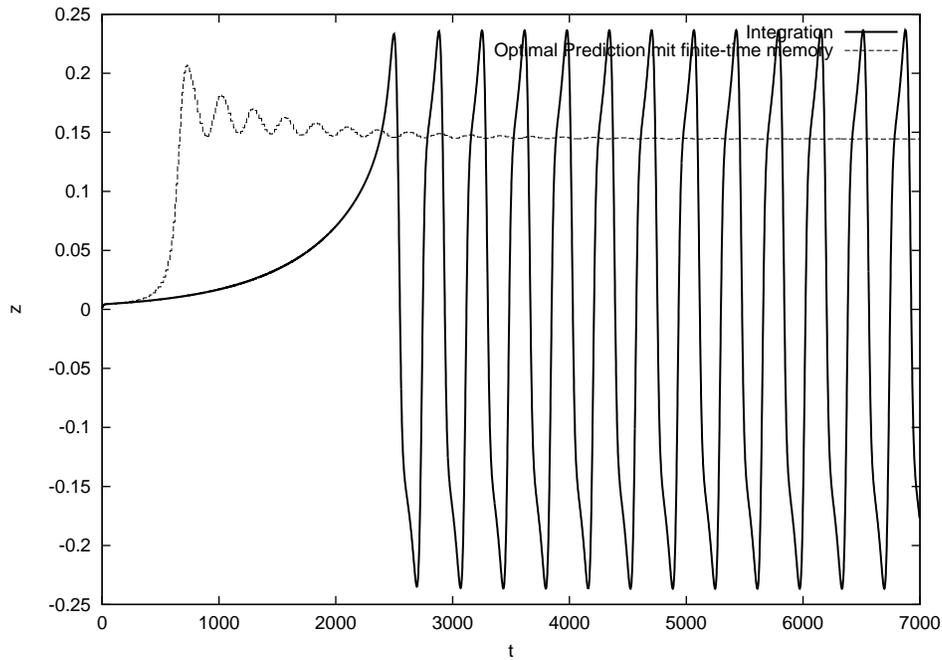


Abbildung 5.8: Zeitentwicklung von z für $\gamma_0 = -0.001$: die dicke durchgezogene Kurve folgt aus der Integration, während die dünne gestrichelte Kurve die Vorhersage der “Optimal Prediction” darstellt.

Umgebung des Anfangs- und Endzustandes gültig ist. Somit sind mit ihr nur asymptotische Aussagen möglich, dennoch keine über den Übergang selbst.

Allerdings ist man mit der “Optimal Prediction” Methode wie auch mit der Zentrumsmannigfaltigkeitsreduktion [22] nicht in der Lage, asymptotisch die Oszillationen (ELMs) zu reproduzieren, die nach Instabilität des H-Zustandes auftreten.

6. Zusammenfassung

In dieser Arbeit wurde die Anwendbarkeit der “Optimal Prediction” Methode auf verschiedene wichtige physikalische Modelle untersucht, die beispielsweise aufgrund eingeschränkter Rechnerkapazitäten keine geschlossene Beschreibung zulassen. Dabei wurde besonderer Wert auf die Güte der Methode im Vergleich zu gängigen Reduktionsverfahren wie dem Galerkin- oder dem Karhunen-Loeve-Verfahren oder einer Zentrumsmannigfaltigkeitsreduktion gelegt.

Typisch für das Galerkin-Verfahren und seine verbesserte Form der Karhunen-Loeve-Zerlegung ist die Projektion der gesuchten Lösung auf einen Satz von sogenannten kollektiven Koordinaten (Moden). Dabei besteht die Schwierigkeit darin, sich auf eine bestimmte Anzahl von Moden festzulegen und diese auszusuchen. Leider gibt es bei einem gegebenen Problem kein anderes Kriterium dafür, daß die ausgesuchten kollektiven Koordinaten tatsächlich die relevanten sind, als den nachträglichen Vergleich der so erzeugten Dynamik mit numerischen Ergebnissen der Integration.

Die “Optimal Prediction” Methode versucht in 0. Ordnung, im Gegenteil zu den eben genannten Verfahren, vorab bekannte statistische Eigenschaften des physikalischen Systems besser auszunutzen. Aus der Kenntnis eines statistischen invarianten Maßes folgt eine entsprechende Verteilung der gesuchten Lösungen. Für diese setzt man auch eine Zerteilung in eine Modenbasis an; da nur wenige dieser Moden aber aufgelöst werden sollen, bzw. können, ersetzt man die unaufgelösten Moden in den Bewegungsgleichungen durch ihre bedingten Erwartungswerte bezüglich des bekannten Maßes μ , seien die Anfangswerte der kollektiven Koordinaten bekannt. Die Methode besteht also in den zwei wichtigen Schritten der Projektion der Lösung auf den Raum der kollektiven Koordinaten und der Mittelung des Einflusses der unaufgelösten Freiheitsgrade auf die beibehaltenen Moden.

Der erste Teil dieser Arbeit war der Untersuchung der Dynamik in ausgewählten Hamiltonschen Systemen gewidmet.

In **Paragraph 2** wurde gezeigt, daß die Anwendung der “Optimal Prediction” Methode in 0. Ordnung auf die quintisch nichtlineare Schrödinger-Gleichung gute Ergebnisse liefert, aller-

dings nur bis zum Eintritt des “Blow-up” der Signalamplitude, für den sie eine untere zeitliche Einschätzung liefert. Diese Einschränkung wurde darauf zurückgeführt, daß die anfangs gewählten kollektiven Koordinaten bei einem Kollaps dann nicht mehr oder nicht mehr allein für die Dynamik relevant sind. Bei dieser Analyse wurde klar, daß die Festlegung *a priori* der für die zeitliche Entwicklung relevanten Moden ausschlaggebend für das Gelingen der Methode ist. Dies wurde bei einer Untersuchung der kubisch nichtlinearen Schrödinger-Gleichung in zwei verschiedenen Anfangskonfigurationen bestätigt.

Bei der Untersuchung der diskreten “self-trapping”-Gleichung in **Paragraph 3** mußte die 0. Ordnung der “Optimal Prediction” Methode auf einen weiteren sogenannten “memory”-Term erweitert werden, um die qualitative Entwicklung des Systems wiederzugeben. Es wurden in der Literatur veröffentlichte Herleitungen eingeführt und an unser Problem angepaßt. Es konnte festgestellt werden, daß die erste Ordnung in der Zeit, d. h. die “linearized memory”-Approximation, für die Beschreibung der Dynamik ausreichend ist.

Der zweite Teil dieser Arbeit beschäftigte sich mit dissipativen Systemen. Ausschlaggebend dabei war, daß nicht mehr gemittelte Aussagen gemacht, sondern einzelne Anfangswertprobleme gelöst wurden.

Zuerst wurde in **Paragraph 4** die Güte der “finite-time memory”-Approximation der “Optimal Prediction”, einschließlich ihres Vorteils gegenüber der “linearized memory”-Näherung an einem einfachen zweidimensionalen Beispiel mit einer einzigen Dämpfung bewiesen.

Dann wurde in **Paragraph 5** an bekannte drei-, bzw. zweidimensionale Modelle zur Beschreibung des für die Fusionsforschung wichtigen L-H-Übergangs in toroidal eingeschlossenen Plasmen angeknüpft: es konnte gezeigt werden, daß die “Optimal Prediction” Methode deutlich bessere Ergebnisse erzielt, als eine Reduktion über die Zentrumsmannigfaltigkeitstheorie. Dennoch konnten mit der Methode die experimentell beobachteten Oszillationen der ELMs nicht nachvollzogen werden. Zudem wurde gezeigt, daß im Falle mehrerer miteinander konkurrierender Dämpfungen der Wahl der kollektiven Koordinaten besondere Bedeutung zukommt: es muß die Variable bei dem Verfahren eliminiert werden, deren Dämpfungsrate am größten ist.

Insgesamt kann aus diesen Ergebnissen geschlossen werden, daß sich die “Optimal Prediction” gut eignet, um die dynamische Zeitentwicklung der genannten Modelle zu beschreiben. Insbesondere bietet sie gegenüber einer Zentrumsmannigfaltigkeitsreduktion, gegen die sie asymptotisch für große Zeiten tendiert, den Vorteil, daß sie den mit der Zeit abklingenden Einfluß der Anfangskordinaten berücksichtigt.

Dennoch besteht der schwierigste Schritt und gleichzeitig die größte Einschränkung bei der Anwendung der Methode darin, die richtigen kollektiven Koordinaten, d. h. die für die Dynamik des untersuchten physikalischen Systems relevanten Freiheitsgrade *a priori* zu bestimmen.

A. Bedingte Erwartungswerte einer reellen, normal verteilten Variable

Wir präsentieren hier die Theoreme, die zur Berechnung der Erwartungswerte ($\langle g, \langle R(u) \rangle_V$) auf der rechten Seite von (1.13) führen, wenn \mathbf{u} eine reelle, normal verteilte Variable $\mathbf{u} = (u_1, \dots, u_n)$ ist, deren Wahrscheinlichkeitsdichte durch

$$P(s_1 < u_1 \leq s_1 + ds_1, \dots, s_n < u_n \leq s_n + ds_n) = \frac{1}{Z} e^{(-\frac{1}{2}(\mathbf{s}, A\mathbf{s}) + \mathbf{b} \cdot \mathbf{s})} ds_1 \cdots ds_n$$

gegeben ist.

Die Matrix A ist dabei symmetrisch, positiv definit und ihre Inverse A^{-1} ist die Matrix der Kovarianzen, deren Elemente durch:

$$a_{ij}^{-1} = Cov\{u_i, u_j\} \equiv \langle u_i u_j \rangle - \langle u_i \rangle \langle u_j \rangle$$

gegeben sind.

Der Vektor \mathbf{b} hängt mit den Erwartungswerten von \mathbf{u} über die folgende Beziehung zusammen:

$$A^{-1} \mathbf{b} = \langle \mathbf{u} \rangle$$

Die Verteilung ist somit wie im Falle einer komplexen Variable durch die Angabe der n Mittelwerte und der $\frac{1}{2}n(n+1)$ unabhängigen Werte der Kovarianzmatrix vollkommen bestimmt. Die gesuchten Erwartungswerte werden hier auch durch diese Parameter ausgedrückt und anhand von den 3 folgenden Lemmata berechnet, deren Beweise in [17] nachzulesen sind.

- **Lemma 1:** Der “bedingte” Erwartungswert von u_i ist eine affine Funktion der kollektiven Koordinaten V_α :

$$\langle u_i \rangle_V = q_{i\alpha} V_\alpha + c_i, \quad (\text{A.1})$$

wobei die Elemente $q_{i\alpha}$ der $n \times N$ Matrix Q und des n -Vektors \mathbf{c} gegeben sind durch:

$$Q = (A^{-1} G^+) (G A^{-1} G^+)^{-1} \quad (\text{A.2})$$

$$\mathbf{c} = A^{-1} \vec{b} - (A^{-1} G^+) (G A^{-1} G^+)^{-1} (G A^{-1} \mathbf{b}) \quad (\text{A.3})$$

Die bedingten Erwartungswerte sind also linear in den V_α und unabhängig von Multiplikationsfaktoren in den Kovarianzen.

- Lemma 2: Die “bedingte” Kovarianzmatrix lautet:

$$Cov\{u_i, u_j\}_V = \langle u_i u_j \rangle_V - \langle u_i \rangle_V \langle u_j \rangle_V \quad (\text{A.4})$$

$$= [A^{-1} - ((A^{-1}G^+)(GA^{-1}G^+)^{-1}(GA^{-1}))]_{ij} \quad (\text{A.5})$$

Die bedingten Kovarianzen sind nur von den Filterfunktionen g_α abhängig und nicht von den kollektiven Koordinaten V_α .

Schließlich erlaubt eine direkt vom komplexen auf den reellen Fall übertragbare Form des Wick’schen Theorems, den Erwartungswert eines Produktes über folgende Formel zu berechnen:

- Lemma 3:

$$\langle \prod_{p=1}^P (u_{i_p} - \langle u_{i_p} \rangle_V) \rangle_V = \begin{cases} 0 & P \text{ ungerade} \\ \sum_{\text{alle Permutationen}} Cov\{u_{i_1}, u_{i_2}\}_V \cdots Cov\{u_{i_{p-1}}, u_{i_p}\}_V & P \text{ gerade} \end{cases} \quad (\text{A.6})$$

Somit wird jeder Mittelwert der Form $\langle u_i^p u_j^l \rangle$ als Funktion von ausschließlich $\langle u_i^2 \rangle$ und $\langle u_i u_j \rangle$ ausgedrückt [13].

Sind die Gaußschen Zufallsvariablen \mathbf{u} aber durch eine kontinuierliche Variable x parametrisiert, dann werden die Erwartungswerte $\langle \mathbf{u}(x) \rangle$ und die Kovarianzen $a^{-1}(x, y) = \langle \mathbf{u}(x) \mathbf{u}(y) \rangle - \langle \mathbf{u}(x) \rangle \langle \mathbf{u}(y) \rangle$ in einem geeigneten Bereich für (x, y) definiert. Die Matrix A^{-1} wird zum Integraloperator, deren “Kern” $a^{-1} = a(x, y)$ gegeben ist durch:

$$\int a^{-1}(x, y) a(y, z) dy = \delta(x, z)$$

A. Bedingte Erwartungswerte einer reellen, normal verteilten Variable

Die Filterfunktionen $g_\alpha(x)$ sind jetzt auch kontinuierliche Funktionen von x , und (A.1) wird jetzt zu

$$\langle \mathbf{u}(x) \rangle_V = \langle \mathbf{u}(x) \rangle + c_\beta(x) [V_\beta - \langle \int g_\beta(y) \mathbf{u}(y) dy \rangle], \quad (\text{A.7})$$

mit

$$c_\beta(x) = \left\{ \int a^{-1}(x, y) g_\alpha(y) dy \right\} m_{\alpha\beta}^{-1},$$

und

$$m_{\alpha\beta} = \int \int g_\alpha(x) a^{-1}(x, y) g_\beta(y) dx dy$$

Die Formeln (A.4) und (A.6) werden entsprechend auf kontinuierliche Variablen umgeschrieben.

B. Struktur der Korrekturterme der “Optimal Prediction” Methode bei höheren Nichtlinearitäten

Betrachtet man eine nichtlineare Schrödinger-Gleichung mit beliebig großer Potenz der Nichtlinearität

$$iu_t = -u_{xx} + A|u|^t u \quad t \geq 4, \quad (\text{B.1})$$

kann man allgemein zeigen, daß die Form des zweiten Terms der rechten Seite der (2.33) entsprechenden Gleichung folgende ist:

$$\begin{aligned} g_{\alpha k} \langle u_k^{t+1} u_k^{*t} \rangle_V^0 &= A g_{\alpha k} \langle u_k \rangle^{t+1} \langle u_k^* \rangle^t \\ &+ B \langle u_k \rangle (\text{Cov}\{u_k, u_k^*\})^t \\ &+ \sum_{q,p,s; q+p+2s=(t+1)+t} K_{q,p,s} \langle u_k \rangle^q \langle u_k^* \rangle^p (\text{Cov}\{u_k, u_k^*\})^s, \end{aligned} \quad (\text{B.2})$$

wobei A , B und $K_{q,p,s}$ numerische Koeffizienten sind. Der erste Term liefert die 0. Ordnung in c , d. h. die Galerkin-Approximation in der Form

$$\delta_{\alpha, \underbrace{\beta + \gamma + \dots}_{t+1} + \underbrace{\delta - \varepsilon - \dots - \eta}_t} V_\beta V_\gamma \dots V_\delta V_\varepsilon^* \dots V_\eta^*.$$

Der zweite Term liefert die höchste Ordnung in c in der Form

$$c^t V_\alpha.$$

Der dritte Term schließlich liefert eine Summe von Termen der Form

$$c^s \delta_{\alpha, \underbrace{\beta + \dots + \gamma}_q - \underbrace{\delta - \dots - \varepsilon}_p} V_\beta \dots V_\gamma V_\delta^* \dots V_\varepsilon^*$$

Die allgemeine Form der über die “Optimal Prediction” Methode 0. Ordnung bestimmten zeitlichen Entwicklung der kollektiven Koordinaten, die von (1.13) gegeben ist, enthält also um so mehr Korrekturterme, als die Nichtlinearität der Schrödinger-Gleichung (2.1) größerer Ordnung ist.

C. Bestimmung von m_0 und b im Fall $p = 2$

Für die nichtlineare Schrödinger-Gleichung (2.1) mit $p = 2$ und $A = -2$ erhält man nach der Partitionierung des Hamiltonians für $H^1(q, p)$:

$$H^1(q, p) = \frac{1}{2} \int_0^{2\pi} ((q_x^2 + p_x^2)^2 + m_0^2(q^2 + p^2) + b) dx,$$

und

$$H^1(q, p) = \frac{-1}{2} \int_0^{2\pi} ((q^2 + p^2)^2 + m_0^2(q^2 + p^2) - b) dx$$

Da b aber keine Rolle bei der Mittelung spielt, setzen wir erstmal $b = 0$ für die Berechnung von m_0 .

Quadriert ergibt diese Gleichung:

$$\begin{aligned} H^1(q, p)^2 &= \frac{1}{4} \int_0^{2\pi} \left[(q^2(x) + p^2(x))^2 + m_0^2 (q^2(x) + p^2(x)) \right] dx \\ &\quad * \int_0^{2\pi} \left[(q^2(z) + p^2(z))^2 + m_0^2 (q^2(z) + p^2(z)) \right] dz \end{aligned}$$

Die Varianz ist also gegeben durch:

$$\begin{aligned} \langle H^1(q, p)^2 \rangle &= \frac{1}{4} \int_0^{2\pi} dx \int_0^{2\pi} dz \{ \langle (q^2(x) + p^2(x))^2 \cdot (q^2(z) + p^2(z))^2 \rangle \\ &\quad + 2m_0^2 \langle (q^2(x) + p^2(x))^2 \cdot (q^2(z) + p^2(z)) \rangle \\ &\quad + m_0^4 \langle (q^2(x) + p^2(x)) \cdot (q^2(z) + p^2(z)) \rangle \} \end{aligned}$$

Seien

$$\mathcal{A} = \langle (q^2(x) + p^2(x))^2 \cdot (q^2(z) + p^2(z))^2 \rangle, \quad (\text{C.1})$$

und

$$\mathcal{B} = \langle (q^2(x) + p^2(x))^2 \cdot (q^2(z) + p^2(z)) \rangle \quad (\text{C.2})$$

und

$$\mathcal{C} = \langle (q^2(x) + p^2(x)) \cdot (q^2(z) + p^2(z)) \rangle \quad (\text{C.3})$$

Die Forderung nach dem Minimieren der Varianz als Funktion von m_0 erfordert die Berechnung von \mathcal{A} , \mathcal{B} und \mathcal{C} , die mit dem Wick'schen Theorem erfolgt.

Bestimmung von C

Wenden wir uns zuerst dem dritten Term zu:

$$C = \langle (q^2(x) + p^2(x)) (q^2(z) + p^2(z)) \rangle$$

$$C = 2 \langle p^2(x)p^2(z) \rangle + 2 \langle q^2(x)p^2(z) \rangle$$

Da q und p bezüglich eines Gaußschen Maßes unkorreliert sind, folgt:

$$C = 2 \langle p^2(x)p^2(z) \rangle + 2 \langle p^2 \rangle^2$$

Aus dem Wick'schen Theorem folgt dann:

$$C = 2 (\langle p^2 \rangle^2 + 2 \langle p(x)p(z) \rangle^2) + 2 \langle p^2 \rangle^2,$$

was schließlich für C ergibt:

$$C = 4 \langle p^2 \rangle^2 + 4 \langle p(x)p(z) \rangle^2$$

Bestimmung von \mathcal{B}

Nach dem Ausmultiplizieren und unter Berücksichtigung von:

$$\langle q^4(x)q^2(z) \rangle = \langle p^4(x)p^2(z) \rangle,$$

$$\langle p^4(x)q^2(z) \rangle = \langle q^4(x)p^2(z) \rangle,$$

und

$$\begin{aligned} \langle q^2(x)p^2(x)q^2(z) \rangle &= \langle q^2(x)p^2(x)p^2(z) \rangle \\ &= \langle p^2 \rangle \langle p^2(x)p^2(z) \rangle, \end{aligned}$$

da q und p völlig unabhängig und gleichwertig betrachtet werden, erhält man für \mathcal{B}

$$\mathcal{B} = 2 \langle q^4(x)q^2(z) \rangle + 2 \langle q^4 \rangle \langle q^2 \rangle + 4 \langle q^2 \rangle \langle q^2(x)q^2(z) \rangle \quad (\text{C.4})$$

C. Bestimmung von m_0 und b im Fall $p = 2$

Das Wick'sche Theorem ermöglicht uns jetzt, die Erwartungswerte von Produkten zu vereinfachen, bzw. zu berechnen:

$$\begin{aligned}\langle p^4 \rangle &= 3 \langle p^2 \rangle^2, \\ \langle p^2(x)p^2(z) \rangle &= \langle p^2 \rangle^2 + 2 \langle p(x)p(z) \rangle^2, \\ \langle p^4(x)p^2(z) \rangle &= 3 \langle p^2 \rangle^3 + 6 \langle p^2 \rangle \langle p(x)p(z) \rangle^2,\end{aligned}$$

Setzt man nun diese Ergebnisse in (C.4) ein, dann erhält man:

$$\mathcal{B} = 16 \langle p^2 \rangle^3 + 20 \langle p^2 \rangle \langle p(x)p(z) \rangle^2 \quad (\text{C.5})$$

Bestimmung von \mathcal{A}

Die Bestimmung von \mathcal{A} erfolgt jetzt gleichermaßen; berücksichtigt werden müssen die folgenden Gleichungen:

$$\begin{aligned}\langle q^4(x)q^2(z)p^2(z) \rangle &= \langle q^2(x)p^2(x)p^4(z) \rangle \\ &= \langle q^2(x)p^2(x)q^4(z) \rangle \\ &= \langle p^4(x)q^2(z)p^2(z) \rangle \\ &= \langle p^2 \rangle \langle p^4(x)p^2(z) \rangle,\end{aligned}$$

$$\langle q^4(x)q^4(z) \rangle = \langle p^4(x)p^4(z) \rangle, \quad (\text{C.6})$$

und

$$\langle q^4(x)p^4(z) \rangle = \langle p^4(x)q^4(z) \rangle = \langle p^4 \rangle^2. \quad (\text{C.7})$$

Man gelangt zu folgendem Ausdruck:

$$\mathcal{A} = 2 \langle q^4(x)q^4(z) \rangle + 2 \langle q^4 \rangle^2 + 8 \langle q^2 \rangle \langle q^2(x)q^4(z) \rangle + 4 \langle q^2(x)q^2(z) \rangle^2 \quad (\text{C.8})$$

Zusätzlich zu den oben schon erwähnten liefert das Wick'sche Theorem folgende Vereinfachung:

$$\langle p^4(x)p^4(z) \rangle = 9 \langle p^2 \rangle^4 + 36 \langle p^2 \rangle^2 \langle p(x)p(z) \rangle^2 + 72 \langle p(x)p(z) \rangle^4$$

Setzen wir diese Ergebnisse in (C.8) ein, folgt:

$$\mathcal{A} = 64 \langle q^2 \rangle^4 + 136 \langle q^2 \rangle^2 \langle q(x)q(z) \rangle^2 + 160 \langle q(x)q(z) \rangle^4 \quad (\text{C.9})$$

Bestimmung des Minimums der Varianz $\langle (H^1)^2 \rangle$ als Funktion von m_0

$$\langle H^1(q, p)^2 \rangle = \frac{1}{4} \int_0^{2\pi} dx \int_0^{2\pi} dz \{ \mathcal{A} + 2m_0^2 \mathcal{B} + m_0^4 \mathcal{C} \} \quad (\text{C.10})$$

Setzen wir nun die eben berechneten Ausdrücke für \mathcal{A} , \mathcal{B} und \mathcal{C} ein, bekommt man:

$$\begin{aligned} \langle (H^1)^2 \rangle &= \frac{1}{4} \int_0^{2\pi} dx \int_0^{2\pi} dz \{ [16 \langle q^2 \rangle^4 + 34 \langle q^2 \rangle^2 \langle q(x)q(z) \rangle^2 + 40 \langle q(x)q(z) \rangle^4] \\ &\quad + m_0^2 [8 \langle q^2 \rangle^3 + 10 \langle q^2 \rangle \langle q(x)q(z) \rangle^2] \\ &\quad + m_0^4 [\langle q^2 \rangle^2 + \langle q(x)q(z) \rangle^2] \} \end{aligned}$$

Mit dem schon im Hauptteil dieser Arbeit definierten Integral

$$I_m(q) = \int_0^{2\pi} dx \int_0^{2\pi} dz (\langle p(x)p(z) \rangle)^{2q}$$

schreibt sich $\langle (H^1)^2 \rangle$ etwas übersichtlicher:

$$\begin{aligned} \langle (H^1)^2 \rangle &= [64\pi^2 \langle p^2 \rangle^4 + 34 \langle p^2 \rangle^2 I_m(1) + 40 I_m(2)] \\ &\quad + m_0^2 [32\pi^2 \langle p^2 \rangle^3 + 10 \langle p^2 \rangle I_m(1)] \\ &\quad + m_0^4 [4\pi^2 \langle p^2 \rangle^2 + I_m(1)], \end{aligned}$$

da $\langle p^2 \rangle$ unabhängig von der Ortsvariable x oder z ist.

Die analytische Berechnung von $I_m(q)$ mit Hilfe einer Rekursion liefert:

$$I_m(q) = \frac{1}{(2\pi m_0^2)^{2q}} \sum_{l=0}^{2q} C_{2q}^l (2m_0^2)^l I_l$$

mit

$$I_l = \begin{cases} 4\pi^2 & \text{für } l = 0 \\ 0 & \text{für } l = 1 \\ \frac{2\pi^2}{2^{l-1}} \sum_{k_1=-\infty, k_1 \neq 0}^{+\infty} \frac{1}{k_1^2 + m_0^2} \cdots \sum_{k_{l-1}=-\infty, k_{l-1} \neq 0}^{+\infty} \frac{1}{k_{l-1}^2 + m_0^2} \frac{1}{(k_1 + \dots + k_{l-1})^2 + m_0^2} & \text{für } l \geq 2 \end{cases}$$

C. Bestimmung von m_0 und b im Fall $p = 2$

Somit hat man also die Varianz $\langle (H^1)^2 \rangle$ auf eine mit m_0 parametrisierte Funktion von den Autokorrelationen:

$$\langle p^2 \rangle = \langle q^2 \rangle = \frac{1}{2\pi} \sum_{k=-\infty}^{+\infty} \frac{1}{k^2 + m_0^2}$$

und den Korrelationen:

$$\langle p(x)p(z) \rangle = \frac{1}{2\pi} \sum_{k=-\infty}^{+\infty} \frac{e^{ik(x-z)}}{k^2 + m_0^2}.$$

Diese Funktion kann bezüglich m_0 minimiert werden; daraus ergibt sich der "optimale Wert" $m_0 = 1.016$.

Daraus folgt für b :

$$\langle H^1(q, p) \rangle = 0 \implies \int_0^{2\pi} \langle ((q^2 + p^2)^2 + m_0^2(q^2 + p^2) - b) \rangle dx = 0,$$

$$b = \frac{1}{2\pi} \int_0^{2\pi} (2 \langle q^4 \rangle + 2 \langle q^2 \rangle^2 + m_0^2(2 \langle q^2 \rangle)),$$

$\langle q^4 \rangle$ wird mit dem Wick'schen Theorem umgeschrieben, es folgt:

$$b = \frac{1}{2\pi} \int_0^{2\pi} (8 \langle q^2 \rangle^2 + 2m_0^2 \langle q^2 \rangle),$$

und schließlich:

$$b = (8 \langle q^2 \rangle^2 + 2m_0^2 \langle q^2 \rangle),$$

Mit $m_0 = 1.016$ folgt also $b \approx 2.907$.

Literatur

- [1] J. P. Antoine, *Wavelet Analysis: a New Tool in Physics*, in *Wavelets in Physics*, edited by J. C. V. den Berg, pages 9–21, Cambridge, University Press, 1999.
- [2] J. Bell, A. J. Chorin, and W. Crutchfield, *Stochastic Optimal Prediction with Application to Averaged Euler Equations*, in *Proc. 7th Nat. Conf. Comput. Fluid. Mech.*, edited by C. A. Lin, pages 1–13, Pingtung, Taiwan, 2000.
- [3] M. Berning, *Entstehung kohärenter Strukturen als Folge von Austausch-Instabilitäten*, PhD thesis, Heinrich-Heine-Universität Düsseldorf, 1999.
- [4] L. Bernstein, *Local Modes and Degenerate Perturbation Theory*, in *Davydov's Soliton Revisited, Self-Trapping of Vibrational Energy in Protein*, edited by P. L. Christiansen and A. C. Scott, pages 495–503, NATO ASI Series, Series B: Physics Vol. 243, Plenum Press, NY and London, 1990.
- [5] P. Beyer, *Modellrechnungen zum L-H-Übergang in Tokamak-Plasmen*, PhD thesis, Heinrich-Heine-Universität Düsseldorf, 1995.
- [6] P. Beyer, S. Benkadda, and X. Garbet, *Proper Orthogonal Decomposition and Galerkin Projection for a Three-Dimensional Plasma Dynamical system*, *Physical Review E* **61**(1) (2000) 813–823.
- [7] P. Beyer and K.-H. Spatschek, *Center Manifold Theory for the Dynamics of the L-H-Transition*, *Phys. Plasmas* **3**(3) (1996) 995–1004.
- [8] K. Binder and D. P. Landau, *Monte-Carlo Simulations in Statistical Physics*, Cambridge University Press, 2000.
- [9] C. Blatter, *Wavelets- Eine Einführung*, Vieweg Verlag, Wiesbaden, 1998.
- [10] T. Burkhardt and J. van Leeuwen, *Real-Space Renormalization*, Springer, Berlin, 1982.

- [11] J. Carr, *Applications of Center Manifold Theory*, Springer-Verlag New York Inc., 1981.
- [12] A. J. Chorin, *Conditional Expectations and Renormalization*, *Multiscale Modeling and Simulation* **1** (2003) 105–118.
- [13] A. J. Chorin, O. H. Hald, and R. Kupferman, *Optimal Prediction and the Mori-Zwanzig Representation of Irreversible Processes*, *Proc. Nat. Acad. Sci. USA* **97**(7) (2000) 2968–2973.
- [14] A. J. Chorin, O. H. Hald, and R. Kupferman, *Non-Markovian Optimal Prediction*, *Monte Carlo Meth. Appl.* **99**(7) (2001) 99–109.
- [15] A. J. Chorin, O. H. Hald, and R. Kupferman, *Optimal Prediction with Memory*, *Physica D* **166**(3-4) (2002).
- [16] A. J. Chorin, A. P. Kast, and R. Kupferman, *Optimal Prediction of Underresolved Dynamics*, *Proc. Nat. Acad. Sci. USA* **95**(8) (1998) 4094–4098.
- [17] A. J. Chorin, A. P. Kast, and R. Kupferman, *Prediction of Large-Scale Dynamics Using Unresolved Computations*, *Contemp. Math. AMS* **238** (1999) 53–75.
- [18] A. J. Chorin, A. P. Kast, and R. Kupferman, *Unresolved Computation and Optimal Predictions*, *Comm. Pure Appl. Math.* **52**(10) (1999) 1231–1254.
- [19] A. J. Chorin, R. Kupferman, and D. Levy, *Optimal Prediction for Hamiltonian Partial Differential Equations*, *J. Comput. Phys.* **162**(1) (2000) 267–297.
- [20] A. S. Davydov, *The Theory of Contraction Proteins under their Excitation*, *J. Theor. Biol.* **38** (1973) 559–569.
- [21] A. S. Davydov and N. I. Kislukha, *Solitary Excitations in One-Dimensional Molecular Chains*, *Phys. Status Solidi (B)* **59** (1973) 465–470.

-
- [22] P. H. Diamond, Y. M. Liang, B. A. Carreras, and P. W. Terry, *Self-Regulating Shear Flow Turbulence: a Paradigm for the L to H Transition*, Physical Review Letters **72**(16) (1994) 2565 ff.
- [23] M. Eberhard, *Vergleich von Transportmodellen für Feldlinien und Teilchen in stochastischen Magnetfeldern mittels symplektischer Abbildungen*, PhD thesis, Heinrich-Heine Universität Düsseldorf, 2000.
- [24] T. Eickermann, *Nichtlineare Dynamik driftartiger Moden in schwach ionisierten Plasmen*, PhD thesis, Heinrich-Heine-Universität Düsseldorf, 1994.
- [25] J. C. Eilbeck, *Introduction to the Discrete Self-Trapping Equation*, in *Davydov's Soliton Revisited, Self-Trapping of Vibrational Energy in Protein*, edited by P. L. Christiansen and A. C. Scott, pages 473–483, NATO ASI Series, Series B: Physics Vol. 243, Plenum Press, NY and London, 1990.
- [26] J. C. Eilbeck, P. S. Lomdahl, and A. C. Scott, *The Discrete Self-Trapping Equation*, Physica D **16** (1985) 318–338.
- [27] J. D. Fetter, A. L. and Walecka, *Quantum Theory of many Particle Systems*, McGraw-Hill, 1971.
- [28] R. P. Feynman and H. Kleinert, *Effective Classical Partition Functions*, Phys. Rev. A **34**(6) (1986) 5080–5084.
- [29] R. J. Groebner, *An Emerging Understanding of H-mode Discharges in Tokamaks*, Phys. Fluids **B 5** (1993) 2343–3354.
- [30] J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields*, Springer-Verlag New York Inc., 1983.
- [31] O. Hald, *Optimal Prediction and the Klein-Gordon Equation*, Proc. Nat. Acad. Sci. USA **96**(9) (1999) 4774–4779.

- [32] O. Hald and R. Kupferman, *Convergence of Optimal Prediction for Nonlinear Hamiltonian Systems*, SIAM J. Numer. Anal. **39**(3) (2001) 983–1000.
- [33] Q. Hou, N. Goldenfeld, and A. McKane, *Renormalization Groups and Perfect Operators for Stochastic Differential Equations*, Physical Review E **6303**(3) (2001) 6125 ff.
- [34] L. Kadanoff, *Statistical Physics: Statics, Dynamics and Renormalization*, World Scientific, Singapore, 2000.
- [35] A. P. Kast, *Optimal Prediction of Stiff Oscillatory Mechanics*, Proc. Nat. Acad. Sci. USA **97**(12) (2000) 6253–6257.
- [36] H. Kleinert, *Gauge Fields in Condensed Matter*, volume 1, World-Scientific Singapore, 1989.
- [37] H. Kleinert, *Pfadintegrale in Quantenmechanik, Statistik und Polymerphysik*, B. I. Wissenschaftsverlag, Mannheim, 1993.
- [38] A. J. Lichtenberg and M. A. Lieberman, *Regular and Stochastic Motion*, Springer-Verlag New York, Heidelberg, Berlin, 1983.
- [39] B. P. V. Milligen, *Wavelets, Non-Linearity and Turbulence in Fusion Plasmas*, in *Wavelets in Physics*, edited by J. C. V. den Berg, pages 9–21, Cambridge, University Press, 1999.
- [40] H. Mori, *Transport, Collective Motion, and Brownian Motion*, Prog. Th. Phys. **33** (1965) 423–455.
- [41] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, Inc., 1991.
- [42] P. Plaschko and K. Brod, *Nichtlineare Dynamik, Bifurkation und chaotische Systeme*, Vieweg Verlag, Wiesbaden, 1995.
- [43] M. S.-K., *Modern Theory of Critical Phenomena*, W. A. Benjamin Massachusetts, 1976.

-
- [44] U. Schumacher, *Fusionsforschung, Eine Einführung*, Wissenschaftliche Buchgesellschaft Darmstadt, 1993.
- [45] A. Scotti and C. Meneveau, *Fractal Model for Coarse-Grained Partial Differential Equations*, Phys. Rev. Lett. **78** (1997) 867–870.
- [46] L. Sirovic, *Empirical Eigenfunctions and Low Dimensional Systems*, in *New Perspectives in Turbulence*, edited by L. Sirovic, pages 139–163, Springer-Verlag, New York, 1978.
- [47] K. H. Spatschek, *Theoretische Plasmaphysik*, B. G. Teubner, Stuttgart, 1990.
- [48] C. Sulem and P. L. Sulem, *The Nonlinear Schrödinger Equation, Self-Focusing and Wave Collapse*, Springer, New York, 1999.
- [49] O. Vaillant, *Une méthode particulière stochastique à poids aléatoires pour l'approximation de solutions statistiques d'équations de McKean-Vlasov-Fokker-Planck*, PhD thesis, Université de Provence Aix-Marseille I, 2000.
- [50] S. N. Vlasov, V. A. Petrishchev, and V. I. Talanov, *Averaged Description of Wave Beams in Linear and Nonlinear Media*, Radiophys. Quantum Electron. **14** (1974) 1062–1070.
- [51] F. Wagner, G. Becker, K. Behringer, D. Campbell, A. Eberhagen, W. Engelhardt, G. Fussmann, O. Gehre, J. Gernhardt, G. von Gierke, G. Haas, M. Huang, F. Karger, M. Keilhacker, O. Klüber, M. Kornherr, K. Lackner, G. Lisitano, G. G. Lister, H. M. Mayer, D. Meisel, E. R. Müller, H. Murmann, H. Niedermeyer, W. Poschenrieder, H. Rapp, H. Röhr, F. Schneider, G. Siller, E. Speth, A. Stäbler, K. H. Steuer, G. Venus, O. Vollmer, and Z. Yü, *Regime of Improved Confinement and High Beta in Neutral-Beam-Heated Divertor Discharges of the ASDEX Tokamak*, Phys. Rev. Lett. **49** (1982) 1408–1412.
- [52] R. R. Weynants and G. V. Oost, *Edge Biasing in Tokamaks*, Plasma Phys. Control. Fusion **35** (1993) B177–B189.
- [53] V. E. Zakharov, *Collapse of Langmuir Waves*, Sov. Phys. JETP **35** (1972) 908–914.

- [54] R. Zwanzig, *Problems in Nonlinear Transport Theory*, in *Systems far from Equilibrium*, edited by L. Garrido, pages 198–221, Springer, Berlin, 1980.

Herrn Prof. Dr. K.-H. Spatschek danke ich dafür, mir die Möglichkeit gegeben zu haben, diese Arbeit an seinem Institut anzufertigen, und für seine vielfältige Unterstützung während der letzten Jahre.

Herrn Prof. Dr. A. Pukhov spreche ich meinen Dank für die Übernahme des Koreferats aus.

Herrn Prof. Dr. R. Grauer bin ich für wertvolle Anregungen und Hinweise bei der Einarbeitung in das Thema dankbar.

Bei *Herrn Dr. H. Wenk* bedanke ich mich für die die Betreuung des Computernetzes und bei *Herrn E. Zügge* für die Hilfe bei technischen Problemen.

Für vielfältige und immer freundliche Hilfe bei organisatorischen Fragen bedanke ich mich bei *Frau E. Gröters* und *Frau R. Gerardi*.

Bei allen Kollegen des Instituts für Theoretische Physik, insbesondere bei Dr. W. Laedke, Fabio Mancin, Alexander Posth, Christoph Karle, Dmitri Lesnik, Zelemir Marojevic, Marcus Neuer, Andreas Wingen, Götz Lehmann und Kolja Kauder bedanke ich mich für die sehr gute und freundliche Arbeitsstimmung und die ständige Bereitschaft zu helfen.

Dem Graduiertenkolleg *Hochtemperatur-Plasmaphysik* und dem DAAD danke ich für die finanzielle Unterstützung.

Meiner Familie danke ich für ihre Unterstützung und meinen Bekannten und Freunden dafür, daß ich bei ihnen immer die notwendige Entspannung und Distanz zur Arbeit finden konnte, sei es in Düsseldorf, Singapur, Stuttgart, Grenoble oder Saumur. Und schließlich bin ich *Fabian* für seine großartige und liebevolle Unterstützung in den letzten vier Jahren und insbesondere für seine Geduld und Hilfe in den letzten Wochen sehr dankbar.

An Eidesstatt erkläre ich, daß

- ich die vorgelegte Dissertation selbst und ohne unerlaubte Hilfe angefertigt habe und daß ich diese in der jetzigen oder einer ähnlichen Form noch keiner anderen Fakultät eingereicht habe,
- die Dissertation noch nicht veröffentlicht wurde.