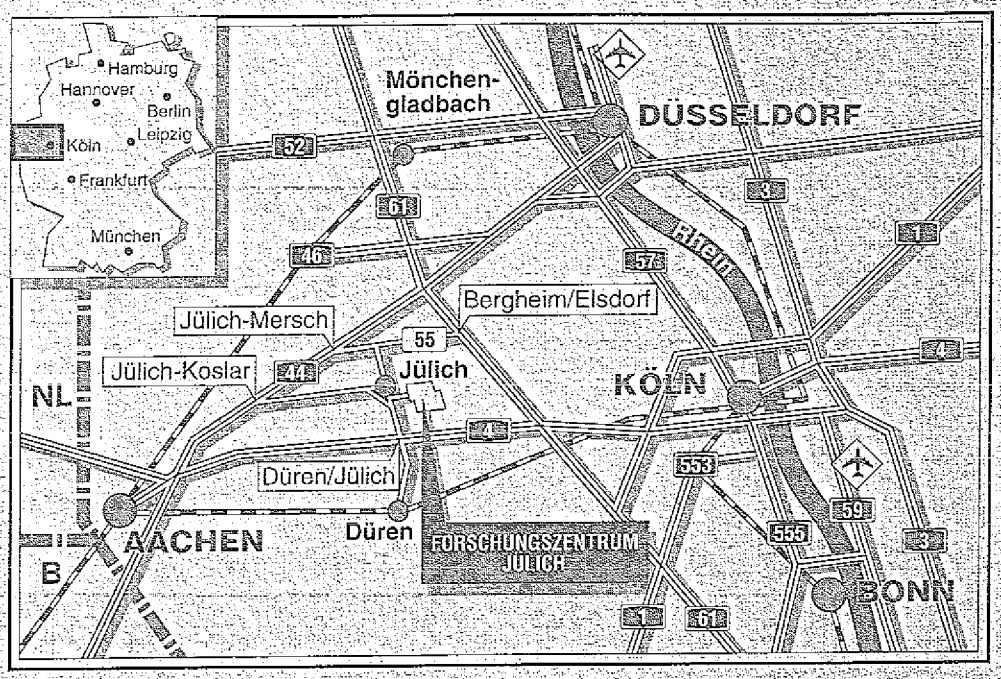
FORSCHUNGSZENTRUM JÜLICH GmbH

Institut für Festkörperforschung

Vereinfachung generierender Partitionen von chaotischen Attraktoren

Michael Eisele



Berichte des Forschungszentrums Jülich ; 3021
ISSN 0944-2952
Institut für Festkörperforschung Jüi-3021
D 38 (Diss. Universität Köln)

Zu beziehen durch: Forschungszentrum Jülich GmbH · Zentralbibliothek
D-52425 Jülich · Bundesrepublik Deutschland
Telefon: 02461/61-6102 · Telefax: 02461/61-6103 · Telex: 833556-70 kfa d

Vereinfachung generierender Partitionen von chaotischen Attraktoren

Michael Eisele

Zusammenfassung

Auch in niedrigdimensionalen Phasenräumen von dissipativen Systemen können Attraktoren mit chaotischer Dynamik existieren. Die wesentlichen koordinatenunabhängigen Eigenschaften dieser Dynamik können mit Hilfe der Symbolischen Dynamik erfaßt werden, bei der Orbits durch Symbolsequenzen gekennzeichnet werden. Dazu benötigt man eine spezielle Unterteilung des Phasenraumes, nämlich eine Partition, die zumindest näherungsweise generierend ist. Solche Partitionen wurden meist konstruiert, indem der Phasenraum in viele kleine Gebiete unterteilt wurde. Eine einfache generierende Partition, die aus wenigen Gebieten besteht, konnte man bisher bloß nach Gespür konstruieren, und zwar nur für Attraktoren, die eine spezielle geometrische Form haben.

Diese Arbeit stellt eine neue, systematische Methode vor, wie man Partitionen vereinfachen kann, ohne daß sie aufhören, generierend zu sein. Dabei wird ein neues Einfachheitskriterium verwendet, welches früheren Einfachheitskriterien ähnelt, aber leichter handhabbar und allgemeiner ist. Um eine Partition zu vereinfachen, muß man nur die wichtigsten Regeln ihrer Symbolischen Dynamik kennen. Zunächst wird die Partition mit Hilfe neuer mathematischer Konzepte (Bündel von expandierenden und kontrahierenden Blättern einer zwiefachen Partition) beschrieben. Diese Beschreibung wird schrittweise abgeändert. Zuletzt wird aus ihr eine Partition konstruiert, die in ebenso guter Näherung generierend und mindestens so einfach ist wie die ursprüngliche Partition.

Die Rechnung selbst beruht auf rein algebraischen Symbolmanipulationen, bei denen die Geometrie des Attraktors nicht betrachtet zu werden braucht. Viele der Rechenregeln sind neu. Sie lassen sich prinzipiell auch auf unübersichtliche, insbesondere höherdimensionale Attraktoren anwenden. Wenn man während der Rechnung auf einige plausible Bedingungen achtet, dann ist auch in diesem allgemeinen Fall ein erfolgreicher Abschluß der Rechnung garantiert. Als Beispiel wird aber nur die zweidimensionale dissipative Hénon Abbildung betrachtet (und zwar bei den Parameterwerten $a = 1.0; b = 0.54$ und den Standardparameterwerten $a = 1.4; b = 0.3$). Die Vereinfachungsmethode ist so gründlich, daß sie sogar bei Standardparameterwerten eine bislang übersehene binäre Partition fand, die ebenso einfach ist wie die bekannte binäre Standardpartition. Als weitere Anwendung der neuen Rechenmethoden wird demonstriert, wie man eine senkrechte Ordnung auf Attraktorpunkten konstruieren kann, sofern deren stabile Mannigfaltigkeiten eindimensional sind.

Die Vereinfachungsmethode ist systematisch genug, um sich prinzipiell als Algorithmus programmieren zu lassen, aber leider ziemlich komplex. Es wird angedeutet, wie sie in der Gestalt eines neuronalen Netzes implementiert werden könnte. Einige neurobiologische Fakten weisen darauf hin, daß die Natur eine solche Implementierung durch die neocortikalen Pyramidenzellen realisiert haben könnte.

Inhaltsverzeichnis

Wandelt sich rasch auch die Welt
wie Wolkengestalten,
alles Vollendete fällt
heim zum Uralten.
R. M. Rilke: „Die Sonnette an Orpheus“ 1.19.

Kapitel 1

Einleitung

1.1 Motivation

Systeme mit komplexer nichtlinearer Dynamik werden auf zwei verschiedene Weisen analysiert: durch numerische Simulation oder durch bloßes Betrachten. Trotz ihrer hohen Spezialisierung und Geschwindigkeit sind Computer nicht immer überlegen. Denn sie können nicht wie das Auge Strukturen erkennen. Sturmfronten auf Satellitenbildern, Konvektionsrollen in Flüssigkeiten, Rauchwirbel in der Luft und all die anderen Erscheinungen, auf die sich unser kognitives Verständnis stützt, können von Computern zwar simuliert werden, so wie jeder dynamische Ablauf simuliert wird, doch sie werden dabei nicht als besondere Strukturen erkannt.

Ob die Computeranalyse oder unser kognitives Verständnis überlegen ist, hängt von der Dynamik ab. Wenn die experimentelle Meßreihe so zufällig erscheint wie das Schwanken der Börsenkurse oder der Lottozahlen, dann kann sie nur auswendig gelernt und mit sehr allgemeinen statistischen Methoden untersucht werden, wie Variationsberechnung, Faktoranalyse usw. . Dazu wird ein großer, unspezialisierter Speicher benötigt. Wenn umgekehrt eine Dynamik in solchem Maße vorhersagbar ist wie das Schwingen eines Pendels oder das Kreisen eines Mondes, dann kann ein Schaltkreis, der so spezialisiert ist wie ein Mikroprozessor, sie genau vorausberechnen.

Zwischen diesen beiden Extremen der zufälligen Dynamik und der Lehrbuchdynamik liegt der Bereich, in dem das Gehirn überlegen ist. Außerdem ähneln große Teile des Gehirns weder einem passiven und homogenen Speicher noch einem präzise verschalteten Mikroprozessor. Das überlegene Beispiel des Gehirns läßt vermuten, daß es noch unbekannte, aber effiziente Methoden gibt, mit denen die Dynamik typischer physikalischer Systeme analysiert, vorausgesehen und gesteuert werden kann.

Doch welche Dynamik ist „typisch“? Diese Frage tauchte schon vor hundert Jahren in der statistischen Physik auf. In typischen Systemen schien sich die Energie gleichmäßig auf die Moden zu verteilen. Doch die Komplexität der Dynamik, die zum Beispiel im Dreikörperproblem gefunden wurde, schreckte Physiker lange Zeit von dieser Frage ab. Nur Mathematiker beschäftigten sich beharrlich damit, vor allem in der Ergodentheorie [20]. Bis heute gibt es jedoch keine allgemein akzeptierte Definition für „typisch“, nur einen gewissen Konsens unter den Wissenschaftlern der Nichtlinearen Dynamik.

Wie der Stammbaum in Abb. 1.1 zeigt, hat die nichtlineare Dynamik neben der statistischen Physik eine zweite wichtige Wurzel: die Computerexperimente. Numerische Simulationen von physikalischen Systemen überraschten mit unerwarteten Phänomenen, die inzwischen als typisch gelten. Dazu gehört das Chaos, das heißt das exponentiell schnelle Anwachsen kleiner Störungen, und die seltsamen Attraktoren, das sind fraktale Attraktoren im Phasenraum dissipativer Systeme. Manche „universellen“ Phänomene blieben auch dann noch erhalten, wenn die Bewegungsgleichungen radikal vereinfacht, jedoch nicht linearisiert wurden. Nur wegen dieser Universalität fand man zum Beispiel in Konvektionsexperimenten dieselben Konstanten, auf die Feigenbaum in einfachen eindimensionalen Abbildungen gestoßen war.

Leider erwies sich in Computersimulationen eine Eigenschaft dynamischer Systeme, die Ma-

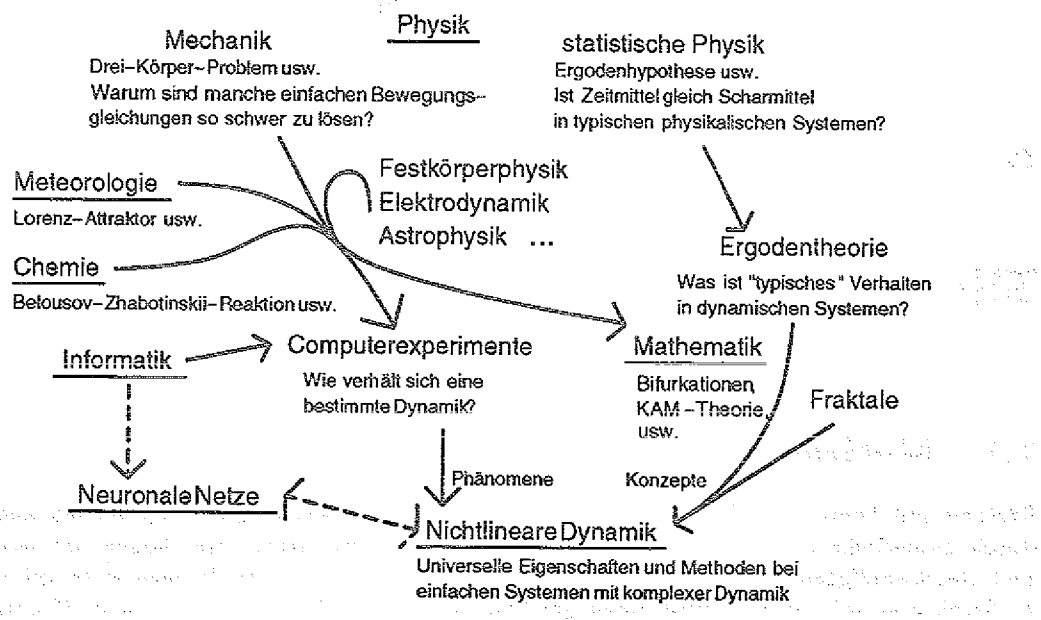


Abbildung 1.1: Vereinfachter Stammbaum der Nichtlinearen Dynamik. Weiterführende Zweige sind gestrichelt.

thematiker oft voraussetzen, als untypisch: die Hyperbolizität. Daher besteht der Beitrag der Mathematik zur Nichtlinearen Dynamik heute weniger aus exakt bewiesenen Sätzen, wie dem KAM-Theorem, als vielmehr aus einem Sammelsurium von Begriffen, mit denen Meßreihen beschrieben werden können. Dazu gehören vor allem diverse Bifurkationstypen, die Lyapunov–Exponenten, normale und verallgemeinerte Dimensionen und dynamische Entropien. Ein empfehlenswertes Lehrbuch stammt von E. A. Jackson [45].

Unser kognitives Verständnis einer komplexen Dynamik vollzieht sich in zwei Schritten. Erst lernen wir, bestimmte Strukturen zu erkennen, wie zum Beispiel eine Welle oder eine Wolke. Dann benutzen wir diese Strukturen, um die Dynamik zu verstehen und zu beschreiben. Auch bei mathematischen Methoden findet man eine Unterteilung in zwei derartige Schritte. Bei der Fourieranalyse, zum Beispiel, wird der Zustand des Systems zunächst in Fourierkomponenten zerlegt, und dann erst wird die Dynamik durch die zeitliche Entwicklung der Fourieramplituden beschrieben. In der nichtlinearen Dynamik besteht der erste Schritt normalerweise darin, den rekonstruierten Phasenraum in kleine, meist quaderförmige Gebiete zu unterteilen. Die Dynamik kann dann durch die Übergänge zwischen diesen kleinen Gebieten beschrieben werden.

Bei diesen mathematischen Methoden ist der erste Schritt unabhängig von der gerade untersuchten Dynamik. Darin unterscheiden sie sich wesentlich von unserem kognitiven Verständnis: Dessen Erfolg hängt davon ab, ob Strukturen erkannt werden, die zur Beschreibung geeignet sind. Zum Beispiel wird man die Bewegung einer Meeresoberfläche nicht verstehen können, solange man nur die einzelnen Lichtreflexe sieht und in diesem flüchtigen Muster keine Wellenstrukturen erkennt. Was wir brauchen, ist also nicht eine beliebige Strukturierung der Dynamik, sondern eine, die an die Dynamik selbst angepaßt ist.

Gibt es auch eine mathematische Methode, die eine an die Dynamik angepaßte Strukturierung der Dynamik verwendet? Ja, und zwar jede „Symbolische Dynamik", die sich auf „generierende Partitionen" stützt. Wie Kap. 2 im Detail erklären wird, beruhen auch diese Methoden auf einer Unterteilung des Phasenraums in kleine Gebiete. Doch diese Unterteilung, die als Partition bezeichnet wird, ist nicht willkürlich, sondern eben „generierend": Sie erfüllt die Bedingung 2.2.1

(aus Kap. 2.2.3), die von der Dynamik abhängt.

Die Eigenschaft „generierend" definiert daher ein präzises mathematisches Kriterium dafür, welche Partitionen an die Dynamik angepaßt sind. Leider kannte man bisher noch keinen systematischen Weg, diese generierenden Partitionen zu finden. In speziellen Fällen, zum Beispiel dem Lorenz Attraktor oder dem Rößler Attraktor, strukturierte man die Dynamik auf einem rein intuitiven Weg, indem man sie durch Begriffe wie „Strecken", „Falten" und „Schneiden" beschrieb und den Phasenraum entlang der „Faltungsflächen" und „Schnittflächen" unterteilte. Diese Methode ist seit langem bekannt [64, 61]. Doch leider stützt sie sich gerade auf die kognitiven Fähigkeiten zur Strukturerkennung, die wir Physiker gerne einem Computer einprogrammieren würden, aber bisher nicht können. Wenn es dieses ungelöste Problem, eine generierende Partition zu finden, nicht gäbe, dann könnte die Symbolische Dynamik in nichtlinearen Systemen eine ähnliche Bedeutung gewinnen wie die Fourieranalyse in linearen Systemen.

1.2 Wichtige Merkmale der neuen Methode

Koordinateninvarianz

Diese Arbeit stellt eine neue Methode zum Finden von generierenden Partitionen vor. Diese Methode benutzt eher algebraische als geometrische Werkzeuge, was einen wichtigen Vorteil hat: Die Methode hängt nicht von der Wahl des Koordinatensystems ab. Diese Unabhängigkeit vom Koordinatensystem ist in der Nichtlinearen Dynamik allgemein übliche Forderung: Selbst wenn die Koordinaten eine tiefere physikalische Bedeutung haben, will man doch Methoden anwenden, die universell sind und nicht von dieser Bedeutung abhängen.

Anwendbarkeit auf experimentelle Daten

Neben der Unabhängigkeit vom Koordinatensystem sind noch weitere Forderungen in der Nichtlinearen Dynamik üblich. Neue Methoden sollten sich im Prinzip auf experimentelle Meßreihen aller typischen, auch höherdimensionalen Systeme anwenden lassen. „Im Prinzip" bedeutet, daß es durchaus üblich ist, die Methoden erst am Beispiel von Systemen mit bekannter und einfacher Dynamik zu testen. In diesem Text wird dasselbe Testobjekt gewählt wie in bisherigen Veröffentlichungen: die Hénon Abbildung bei verschiedenen Parameterwerten. Diese invertierbare und dissipative zweidimensionale Abbildung ist auch das einzige nichthyperbolische dynamische System, in dem die Existenz seltsamer Attraktoren für eine große Menge von Parametern bewiesen ist [7]. Zur numerischen Berechnung von Illustrationen wurde die analytische Form der Hénon Abbildung verwendet, doch die Methode selbst erfordert nur die Kenntnis der zulässigen Symbolsequenzen. Welche Symbolsequenzen zulässig sind, kann man in physikalischen Experimenten tatsächlich messen (siehe z. B. [23]).

Das Ergebnis ist nur näherungsweise generierend

Selbst bei dem Hénon Attraktor ist es noch niemandem gelungen, für eine bestimmte Partition zu beweisen, daß sie generierend ist. Einen exakten mathematischen Beweis werde ich im folgenden gar nicht versuchen. Ich werde mich mit Partitionen begnügen, die gute Näherungen für generierende Partitionen sind. Bei diesen Partitionen dürfen mehrere Orbits dieselbe Symbolsequenz besitzen, aber nur wenn diese Orbits zu allen Zeiten nahe beieinander liegen. Bei verrauschten Daten können solche benachbarten Orbits sowieso nicht unterschieden werden, so daß das Streben nach einer exakten generierenden Partition überflüssig wäre.

Die Methode besteht im Vereinfachen von Partitionen

Es gibt einen trivialen Weg, eine Partition zu konstruieren, die näherungsweise generierend ist. Dazu unterteilt man den Phasenraum einfach in viele, genügend kleine Gebiete. Eine solche feine Partition ist unbefriedigend, da die so gewonnene Strukturierung des Phasenraumes überhaupt

nicht an die Dynamik angepaßt ist. Sie ist aber als Ausgangspunkt der Methode geeignet: Eine solche feine Partition kann dadurch an die Phasenraumdynamik angepaßt werden, daß man sie vereinfacht. Diese Vereinfachung besteht im Vereinigen, Schneiden und Abbilden der einzelnen Gebiete, in die der Phasenraum unterteilt wurde. Als Ergebnis wird eine Partition mit wenigen, großen Gebieten angestrebt, die in genauso guter Näherung generierend ist wie die anfängliche Partition mit ihren vielen kleinen Gebieten.

Mehrere Partitionen können gleich einfach sein

Wenn man eine generierende Partition gefunden hat, dann kann man aus ihr viele andere generierende Partitionen konstruieren. Ich werde in Kap. 2.3.3 argumentieren, daß man nicht immer eine generierende Partition als die einfachste auszeichnen kann, wenn man nur koordinatenunabhängige Kriterien der Einfachheit verwendet. Ich werde mich in Kap. 4.1 aber auch bemühen, unser intuitives Verständnis von „Einfachheit" zu präzisieren. Bei der Vereinfachung generierender Partitionen werde ich dann das koordinatenunabhängige Einfachheitskriterium 4.1.2 (aus Kap. 4.1.2) verwenden. Nach diesem Kriterium kann eine Partition selbst dann noch weiter vereinfacht werden, wenn sie schon die minimale Zahl von Symbolen hat. Dadurch reduziere ich die Zahl der Partitionen, die sich nicht mehr weiter vereinfachen lassen.

1.3 Überblick

Diese Arbeit besteht aus drei Teilen. Die grundlegenden Konzepte werden in Kap. 2 bis 4 dargestellt und ausführlich motiviert. Die zentralen Kapitel 5 bis 7 erbauen dann aus diesen Konzepten eine Methode, durch die eine näherungsweise generierende Partition vereinfacht werden kann. Sowohl der allgemeine Fall als auch ein spezielles Beispiel werden sehr detailliert dargestellt. Die Kap. 8, 9 und große Teile des Anhangs sind wesentlich knapper geschrieben. Sie enthalten eher Ausblicke darauf, was man mit den neuen Konzepten noch alles anfangen kann, als Darstellungen systematischer Methoden.

Kap. 2 erläutert den Hintergrund dieser Arbeit. Zuerst wird an ein paar elementare Begriffe der Nichtlinearen Dynamik erinnert. Danach werden die bisher veröffentlichten Bemühungen zusammengefaßt, die Suche nach generierenden Partitionen zu systematisieren. Dabei erweist es sich als entscheidende Schwierigkeit, Suchkriterien zu finden, die nicht von der Wahl des Koordinatensystems abhängen.

Mit Kapitel 3 beginnt die Darstellung neuer Konzepte. Die normalen Partitionen lassen sich, wenn sie generierend bleiben sollen, nicht flexibel genug abändern. Deshalb wird das Konzept der „Partition" verallgemeinert zu dem Konzept der „zwiefachen Partition". Eine zwiefache Partition besteht aus zwei verschiedenen Partitionen, die jeweils zur Beschreibung von Vergangenheit oder Zukunft verwendet werden. Beide Partitionen hängen nur wenig voneinander ab, und keine muß für sich allein generierend sein. Erst diese Trennung von Vergangenheit und Zukunft schafft die Flexibilität, die nötig ist, um die Partitionen in kleinen Schritten zu verbessern. Dies ist vielleicht die wichtigste Lehre aus dieser Arbeit.

Um mit zwiefachen Partitionen und den eng damit zusammenhängenden Konzepten „Bündel von kontrahierenden Blättern" und „Bündel von expandierenden Blättern" effizient rechnen zu können, müssen selten benutzte Rechensymbole verwendet und neue Rechensymbole definiert werden. Dies geschieht in Kap. 3.2 und noch einmal ausführlicher und präziser in Anhang B. Alle Rechnungen können dann durchgeführt werden, indem man Symbolsequenzen manipuliert. Aber viele Rechenschritte werde ich auch geometrisch veranschaulichen.

Außerdem benötigt man ein Kriterium der Einfachheit, um Partitionen vereinfachen zu können. Kap. 4 argumentiert, daß die Zahl der Symbole kein ausreichender Maßstab ist. Das neue Kriterium 4.1.2 der Einfachheit wird ausführlich motiviert. In Anhang C wird es auf Standardbeispiele angewendet.

Kap. 5 stellt die Vereinfachung einer zwiefachen Partition dar, und zwar sowohl im allgemeinen Fall als auch anhand eines speziellen Beispiels. Und Kap. 6 zeigt, wie man aus der vereinfachten

zwiefachen Partition eine normale, nicht zwiefache Partition rekonstruieren kann, die in genauso guter Näherung generierend ist. Die Kapitel 5.1 und 6.1 geben einen Überblick über diese Methode zur Vereinfachung generierender Partitionen.

Die anschließende Diskussion in Kap. 7 stellt Vorteile und Nachteile dieser Methode gegenüber. Ein Vorteil besteht darin, daß sich die Vereinfachungsschritte ganz systematisch anwenden lassen. Im Falle des Hénon Attraktors bei Standardparameterwerten ergibt sich ein kleines, aber überraschendes Ergebnis: eine neue binäre Partition, die in ebenso guter Näherung generierend und ebenso einfach ist wie die wohlbekannte Standardpartition.

Die Methoden aus Kap. 3 und 4 können mehr als nur generierende Partition vereinfachen. In Kap. 8 werden sie angewendet, um räumliche Ordnungen zu konstruieren. Solche Ordnungen wurden in vielen Veröffentlichungen untersucht, aber fast ausschließlich für eindimensionale oder zweidimensionale dynamische Systeme. Mit den neuen Konzepten dieser Arbeit könnten prinzipiell auch höherdimensionale Systeme geordnet werden, wenn entweder die stabile oder die instabile Mannigfaltigkeit eindimensional ist. Als Beispiel werden ich aber wieder nur die zweidimensionale Hénon Abbildung betrachten.

Zuletzt gibt Kap. 9 einen spekulativen Ausblick, wie sich der hier vorgeschlagene Algorithmus in einem Netzwerk von parallel rechnenden Neuronen implementieren ließe. Dabei geschieht etwas Erstaunliches. Das Gehirn findet Strukturen in dynamischen Systemen, die weit komplexer sind als die primitiven Dynamiken, die in dieser Arbeit als Beispiele verwendet werden. Man sollte erwarten, daß sich diese Komplexität auch in dem Aufbau des Gehirns widerspiegelt und das hier vorgeschlagene neuronale Netzwerk viel zu primitiv ist, um mit dem Gehirn verglichen zu werden. In Wirklichkeit ergeben sich aber viele Analogien zwischen der Anatomie des Neocortex, eines Teils des Säugetiergehirns, und der Architektur des neuronalen Netzwerkes, das nach Prinzipien der Symbolischen Dynamik konstruiert wurde. Diese neurobiologischen Analogien führen aus dem Feld der Physik heraus. Sie wurden aber trotzdem in den Anhang D aufgenommen, weil ich meine, daß man umgekehrt aus der Anatomie des Neocortex einiges über die Analyse nichtlinearer dynamischer Systeme lernen kann.

Andere Anwendungen der neuen Konzepte werden gar nicht diskutiert. Die Suche nach generierenden Partitionen wurde auf die Beispiele beschränkt, die zur Begründung der Methode notwendig sind. Insbesondere wurde kein Versuch unternommen, die Methode auch in mehr als zwei Dimensionen anzuwenden. Erstens hat sich keine dreidimensionale Abbildung als repräsentatives Testobjekt so durchgesetzt wie die zweidimensionale Hénon Abbildung. Und zweitens würden Rechnungen per Hand in höheren Dimensionen äußerst langwierig sein. Im Prinzip könnte man den größten Teil der Rechnung numerisch durchführen. Denn die einzige willkürliche Wahl ist die der anfänglichen (zwiefachen) Partition. Danach läuft die Rechnung ebenso zwangsläufig ab wie andere Optimierungsverfahren (Kap. 7.2.1). Dieser Algorithmus ist allerdings so komplex, daß ich ihn nicht programmierte.

Wenn man generierende Partitionen erst einmal gefunden hat, dann kann man sie anwenden, um zum Beispiel dynamische Entropien, Zeta–Funktionen oder andere über den Attraktor gemittelte Größen zu berechnen. Diese Anwendungen werde ich in dieser Arbeit nicht behandeln, da die Suche nach generierenden Partitionen an sich schon schwierig genug ist.

„Ich habe hier" erzählte er Ulrich, indem er gleichzeitig die dazugehörigen Blätter vorwies, „ein Verzeichnis der Ideenbefehlshaber anlegen lassen, das heißt, es enthält alle Namen, welche in letzter Zeit sozusagen größere Heereskörper von Ideen zum Siege geführt haben; hier dieses andere ist eine Ordre de bataille, das da ein Aufmarschplan; dieses ein Versuch, die Depots oder Waffenplätze festzulegen, aus denen der Nachschub an Gedanken kommt." R. Musil „Der Mann ohne Eigenschaften"; Kap. 85: „General Stumms Bemühung, Ordnung in den Zivilverstand zu bringen."

Kapitel 2

Bisherige Suche nach generierenden Partitionen

2.1 Elementare Begriffe der Nichtlinearen Dynamik

Bevor ich mit der Suche nach generierenden Partitionen beginne, will ich in diesem Abschnitt an ein paar elementare Begriffe der Nichtlinearen Dynamik erinnern. Wem diese Begriffe fremd sind, der kann in Lehrbüchern wie [45] nachschlagen. In [20] und [33] kann er auch spezielle Abschnitte über Symbolische Dynamik finden.

Außerdem will ich in diesem Abschnitt noch eine zweite Motivation für die Suche nach generierenden Partitionen geben: Die von generierenden Partitionen erzeugten Symbolischen Dynamiken stellen eine besonders effiziente Methode dar, um komplexe Dynamiken zu kodieren und zu speichern.

Was hat man sich unter diesem effizienten Speichern vorzustellen? Und wann ist es überhaupt möglich? Eine völlig zufällige Meßreihe läßt sich nicht effizient speichern. Denn Zufälligkeit ist gerade dadurch definiert, daß es keine Möglichkeit gibt, die Meßreihe zu speichern, die wesentlich kürzer als die Meßreihe selbst ist.

Eine rein deterministische Dynamik mit einfachen Bewegungsgleichungen wäre dann am effizientesten gespeichert, wenn man die analytische Form dieser Bewegungsgleichungen erraten hat. Das dynamische System kann entweder ein Fluß $\dot{\underline{x}} = \underline{F}(\underline{x})$ im Phasenraum mit Orbits $\underline{x}(t); t \in \mathbb{R}$ oder eine Abbildung $\underline{x}_{t+1} = \underline{f}(\underline{x}_t)$ mit Orbits $(\underline{x}_t)_{t\in\mathbb{Z}}$ sein. Durch **Poincaré–Schnitte** läßt sich ein Fluß in eine Abbildung umwandeln. Dabei reduziert sich die Dimension des Phasenraumes um eins, was eine große praktische Erleichterung ist. Deshalb werden in dieser Arbeit nur Abbildungen betrachtet.

Doch in der Nichtlinearen Dynamik interessiert der Fall, in dem man Meßreihen kennt aber die Bewegungsgleichungen nicht errät. Die Rekonstruktion einer Nichtlinearen Dynamik durch **Einbettung** garantiert, daß alle Koordinaten des d_p-dimensionalen Phasenraumes meßbar sind. Aus der Meßreihe kann man dann erkennen, von welchem Ort des Phasenraumes aus man unter der Abbildung f zu welchem anderen Ort des Phasenraumes gelangen kann. Wegen des unvermeidlichen Meßfehlers sind kleine Abweichungen ϵ von diesen Orten nicht zu erkennen. Auf diese Weise hat man die Dynamik f näherungsweise rekonstruiert. Allerdings kann die Meßreihe sehr lang sein und viel Speicherplatz beanspruchen.

Eine etwas bessere Weise, die rekonstruierte Dynamik f zu speichern, besteht darin, den Phasenraum im kleine Würfel der Kantenlänge ϵ zu unterteilen. Die Bilder eines solchen Würfels werden in den nächsten n Zeitschritten, falls n klein und die Abbildung f stetig differenzierbar

ist, im allgemeinen nur wenige andere Würfel schneiden. Der Meßreihe kann man entnehmen, welche Übergänge möglich sind. Durch die Kenntnis aller möglichen Übergänge ist die Dynamik hinreichend genau beschrieben. Man braucht etwa ϵ^{-d_p} Würfel und muß daher etwa ebensoviele Übergänge speichern.

Will man die Lage jedes möglichen Orbits nicht nur auf ϵ, sondern beliebig genau abspeichern, so benötigt man auch einen unendlichen Speicher. Dies wird auch bei den nun folgenden, raffinierteren Methoden nicht anders sein. Doch die Rate, mit der der Speicheraufwand wächst, kann reduziert werden, indem man noch etwas geschicktere Weisen wählt, die Dynamik zu beschreiben.

Solange wir uns nicht für Transienten, sondern nur für die Dynamik auf dem Attraktor interessieren, brauchen nur die Würfel gespeichert zu werden, die Attraktorpunkte enthalten. Bei dissipativer Dynamik sind das etwa $\epsilon^{-d_a} \ll \epsilon^{-d_p}$ Würfel, wobei d_a die **fraktale Dimension** des Attraktors ist. Etwa ebensoviele Übergänge müssen dann gespeichert werden.

Diese Konstruktion wird durchaus angewendet, um zum Beispiel **dynamische Entropien** in niedrigen Dimensionen d_p numerisch zu berechnen. Diese Entropien beschreiben, wie viele Sequenzen von n Würfeln in dieser Reihenfolge auch tatsächlich von einem Orbit geschnitten werden. Diese Zahl wächst exponentiell mit n, und zwar mit einer Rate, die um so größer ist, je kleiner man die Würfel wählt. Im Grenzfall $\epsilon \to 0$ ist diese Rate die **topologische dynamische Entropie** der Dynamik. Die Abweichung der bei endlichem ϵ berechneten Wachstumsrate vom Grenzfall der exakten Entropie ist also ein Maß dafür, wie gut man die Dynamik genähert hat. Analog zeigt die Abweichung von der **Kolmogorov–Sinai–Entropie** an, wie gut die Übergangswahrscheinlichkeiten bekannt sind.

Eine effiziente Methode, um die Lage eines Orbits zu einer bestimmten Zeit darzustellen, bietet die übliche Koordinatenschreibweise. Dabei wird die Lage eines Orbits zu jeder Zeit durch die d_p Koordinaten $x^1, \ldots x^{d_p}$ festgelegt. Bei endlicher Genauigkeit ϵ wird der Phasenraum also nicht in kleine Würfel der Kantenlänge ϵ, sondern in lange Streifen der Breite ϵ unterteilt. Jeder dieser Streifen ist durch eine Beziehung der Gestalt $a \leq x^i < a + \epsilon$ definiert. Und man braucht nur d_p sich kreuzende Streifen anzugeben, um die Lage des Orbits auf ϵ genau zu bestimmen.

Um nicht bloß die Lage des Orbits zu einer bestimmten Zeit, sondern auch die möglichen Übergänge zwischen verschiedenen Zeiten auf dieselbe Weise anzugeben, muß man die Übergänge zwischen den Streifen angeben. Leider wird das Bild eines solchen Streifens im allgemeinen nicht mehr parallel zu den Koordinatenachsen, sondern irgendwie quer im Phasenraum liegen. Es wird daher, im Gegensatz zu dem Bild eines kleinen Würfels, viele andere Streifen schneiden, so daß viele Übergänge möglich sind. Die möglichen Übergänge zwischen Streifen sagen also nur wenig darüber aus, welche möglichen Wege ein Orbit nehmen kann.

Durch eine besondere Wahl des Koordinatensystems kann man die Vorteile der Koordinatenschreibweise teilweise retten. Ideal wäre es, wenn die Streifen $a \leq x^i < a + \epsilon$ so liegen würden, daß ihre Bilder unter f wieder parallel zu solchen Streifen liegen und deshalb nur wenige dieser Streifen schneiden. Dieser Idealfall kann praktisch nie erreicht werden. Aber bei besonderen dynamischen Systemen kann man d_s **stabile Richtungen** und d_u **instabile Richtungen** definieren, die stetig vom Ort abhängen und zusammen den Tangentialraum aufspannen $d_s + d_u = d_p$. Dabei bildet die Abbildung f den durch die stabilen bzw. instabilen Richtungen gebildeten Unterraum wieder in solche Unterräume ab. Diese Eigenschaft ist Teil der Definition **hyperbolischer dynamischer Systeme**, die zum Beispiel in [33, Kap. 5.2] diskutiert wird. Dort ist auch im Detail beschrieben, wie man die stabilen und instabilen Richtungen als Richtungen der Achsen spezieller lokaler Koordinatensysteme auffassen kann.

Die möglichen Wege eines Orbits können dann dadurch beschrieben werden, daß man getrennte Abbildungen für den stabilen Unterrraum und den instabilen Unterraum angibt. Jede dieser beiden Abbildungen kann wieder auf die primitive Weise abgespeichert werden, indem man die beiden Unterräume in Würfel der Größe ϵ zerteilt. Statt der Übergänge zwischen ϵ^{-d_p} Würfeln im Phasenraum braucht man dann nur die Übergänge zwischen ϵ^{-d_s} Würfeln des stabilen Unterraumes und ϵ^{-d_u} Würfeln des instabilen Unterraumes abzuspeichern. Wegen $\epsilon^{-d_s} + \epsilon^{-d_u} \ll \epsilon^{-d_s} * \epsilon^{-d_u} = \epsilon^{-d_p}$ muß man also viel weniger Übergänge abspeichern.

Das Standardbeispiel für eine hyperbolische Menge, das Hufeisen von Smale [66], ist kein Attraktor. Tatsächlich gibt es nur relativ komplizierte hyperbolische Attraktoren in zwei Dimensionen

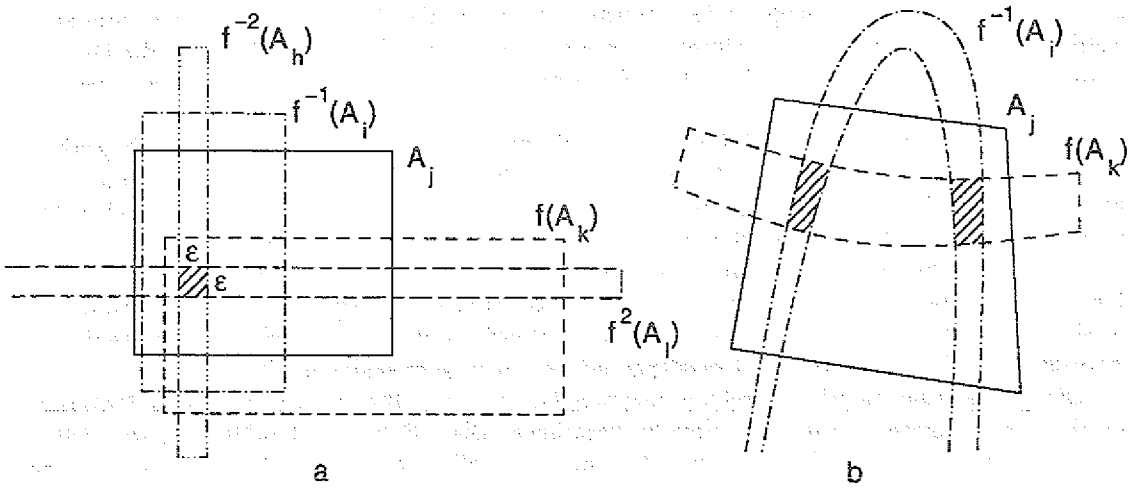


Abbildung 2.1: Abb. a zeigt ein Beispiel für eine günstige Wahl Symbolischer Gebiete. Ein Orbit schneidet bei aufeinanderfolgenden Zeitschritten die Symbolischen Gebiete, A_l, A_k, A_j, A_i und A_h. Seine möglichen Lagen werden durch den Schnitt von Abbildern und Urbildern dieser Symbolischen Gebiete auf ein schraffiertes Gebiet beschränkt, das viel kleiner als die Symbolischen Gebiete selbst ist. Abb. b illustriert ebenso schematisch, warum bei einer ungünstigen Wahl der Symbolischen Gebiete, das heißt bei einer nicht generierenden Partition, die Lage des Orbits nur sehr ungenau bestimmt sein kann.

(wie den von Plykin, siehe [33, Kap. 5.5]). Auf nicht hyperbolischen Attraktoren gibt es diese stabilen und instabilen Richtungen zwar auch fast überall, doch diese hängen dann unstetig vom Ort ab. Als Koordinatenachsen sind solche unstetigen Richtungen unbrauchbar, doch eine weitere ihrer Eigenschaften kann benutzt werden, um die Dynamik effizient zu speichern: Unter der t–fach iterierten Abbildung f^t kontrahieren Entfernungen r entlang der stabilen Richtungen exponentiell schnell $r \sim \exp(\lambda_u t)$, während sie entlang der instabilen Richtungen exponentiell schnell expandieren $r \sim \exp(\lambda_s t)$. Dabei ist $\lambda_u > 0$ ein positiver **Lyapunov Exponent** und $\lambda_s < 0$ negativer Lyapunov Exponent.

Wegen dieser Kontraktion und Expansion ist es möglich, die Lage eines Orbits auch dann auf ϵ genau festzulegen, wenn der Phasenraum in Gebiete unterteilt wird, die sehr viel größer als ϵ sind. Dazu muß jedes dieser Gebiete mit einem **Symbol** gekennzeichet werden. Deshalb werde ich die Gebiete, in die eine Partition den Phasenraum unterteilt, auch **Symbolische Gebiete** nennen. Wenn man die Symbole der letzten m Zeitschritte kennt, so kann man daraus im Prinzip die Lage des Orbits entlang der stabilen Richtungen auf etwa $\exp(\lambda_s \cdot m)$ genau festlegen. Analog folgt aus den nächsten n Zeitschritten die Lage entlang der instabilen Richtung auf etwa $\exp(-\lambda_u \cdot n)$ genau. Um die Lage auf ϵ genau festzulegen muß man deshalb $m \approx \ln(1/\epsilon)/(-\lambda_s)$ vergangene und $n \approx \ln(1/\epsilon)/\lambda_u$ zukünftige Symbole speichern. Die Lage ist dann durch **Symbolsequenzen** der Länge $m + n$ und die möglichen Übergänge durch Symbolsequenzen der Länge $m + n + 1$ hinreichend genau festgelegt.

Was ist der Vorteil dieser Symbolsequenzen? Wenn man sich für jede einzelne Symbolsequenz der Länge $m+n+1$ merken müßte, ob dieser Übergang in der Dynamik vorkommt oder nicht, dann bräuchte man wieder ebensoviel Speicherplatz, wie wenn man den Phasenraum in kleine Würfel zerlegt hätte. Glücklicherweise läßt sich die **Symbolische Dynamik**, das ist die zeitliche Abfolge der Symbole, sehr viel effizienter beschreiben. Wenn man z. B. weiß, daß die Symbolsequenz $AABA$ niemals auftritt, dann kann man daraus folgern, daß die Symbolsequenz $ABAABAABB$ und alle anderen Symbolsequenzen, die $AABA$ enthalten, auch niemals auftreten. Meist genügen schon wenige kurze dieser „verbotenen" Symbolsequenzen, um die Dynamik genau genug zu beschreiben.

Die ganze Schwierigkeit liegt darin, die Symbolischen Gebiete so zu legen, daß der Schnitt ihrer Abbilder und Urbilder die Lage des Orbits auf ϵ genau festlegt, wie in Abb. 2.1.a. Denn auch wenn

dieser Schnitt nur ein kleines Volumen hat, kann er aus mehreren, voneinander entfernten Teilgebieten bestehen, wie in Abb. 2.1.b, und deshalb die Lage des Orbits nur ungenau festlegen. Solche ungüstigen Schnitte zu vermeiden ist das Hauptziel bei der Konstruktion einer generierenden **Partition**.

2.2 Symbolische Dynamik und generierende Partitionen

2.2.1 Historische Referenzen

Hao Bai–Lin [70, 37] und Wiggins [67] beschreiben die Geschichte der Symbolischen Dynamik, die ich hier der Vollständigkeit halber wiedergeben will: Schon früh (1851 [60] und 1901 [34]) verwendeten Mathematiker Sequenzen von Symbolen um Kurven zu charakterisieren. Auf Dynamische Systeme wurden diese Methoden dann von Morse [55] und Birkhoff [10] angewendet. Zum eigenständigen Thema wurde Symbolische Dynamik durch die detaillierte Übersicht von Morse und Hedlund 1938 [56]. Levinsons Arbeit über den getriebenen Van–der–Pol–Oszillator (1950 [48]) und Smales Konstruktion der Hufeisenabbildung (1963 [66]) benutzten Symbolische Dynamik um Sachverhalte zu beweisen, die heute als charakteristisch für deterministisches Chaos gelten. Noch bis 1980 blieb das Feld eine Domäne der Mathematiker. Sinai, Ruelle, Bowen und andere [12, 11, 62] entwickelten die Ergodentheorie hyperbolischer Systeme. Physiker erfuhren davon durch die Zusammenfassung von Alekseev und Yakobson [1], die breitere Übersicht von Eckmann und Ruelle [20] und das Buch von Arnold und Avez [2]. Bei der Analyse experimenteller Zeitreihen steht die Symbolische Dynamik heute im Schatten der Berechnung von Dimensionen, Lyapunov Exponenten etc.. Die detailierteste Anwendung, von der ich weiß, wurde an einem NMR-Laser durchgeführt, bei dem man Millionen von Zeitschritten mit nur etwa einem Prozent Rauschen ausmessen konnte [23]. Nicht nur wurde in der nicht trivialen Geometrie des seltsamen Attraktors eine gute Näherung für eine generierende Partition gefunden [23], auch die plötzliche Änderung („Heterokline Krise") des seltsamen Attraktors bei Änderung eines Parameters konnte mit dem Auftauchen einer zuvor verbotenen Symbolsequenz erklärt werden [22].

2.2.2 Symbole und Symbolsequenzen

Abb. 2.2 zeigt eine stetig differenzierbare Abbildung $f : (0,1) \to (0,1)$ des Einheitsintervalles auf sich selbst mit einem zentralen Maximum bei x_c. Solche Abbildungen wurden von theoretischen Physikern und angewandten Mathematikern (siehe z. B. [13, 18]) in großem Detail untersucht, seitdem Feigenbaum seine universellen Konstanten fand. Doch schon vorher waren qualitative Gesetzmäßigkeiten bekannt, die universell sind, das heißt, die nicht von der speziellen analytischen Form der Abbildung f abhängen. Dazu gehört die Reihenfolge, in der verschiedene periodische Orbits auftauchen, wenn bei Parameteränderung das Maximum der Funktion anwächst [63, 50]. Um diese universelle Reihenfolge zu erklären, wurden die periodischen Orbits durch Symbolsequenzen charakterisiert und dann die Menge der möglichen Symbolsequenzen bestimmt. Beide Schritte wurden im Laufe der Jahre immer weiter vereinfacht und verallgemeinert [51, 19]. Zuletzt entwickelten Zheng Wei-mou und Hao Bai-lin besonders einfache und anschauliche Methoden [38] und stellten sie in lesenswerten Übersichten verschiedener Länge dar [36, 70, 37]. An ihre Schreibweise der Symbolischen Dynamik lehnt sich die Notation an, die im folgenden verwendet wird.

Ein Orbit ist eine unendliche Punktfolge $(x_t)_{t\in\mathbb{Z}}$ mit $f(x_t) = x_{t+1}$. Jede einzelne Lage x_t legt ein einzelnes Symbol in der unendlichen Symbolsequenz des Orbits fest. In dieser Arbeit werden Symbole mit den Zahlen $0, 1$ oder mit Großbuchstaben $A, B, \ldots$ bezeichnet. Üblicherweise verwendet man bei eindimensionalen Abbildungen die beiden Symbole R und L für rechts und links. Um bei der Schreibweise zwischen Symbolsequenzen wie $\ldots RLLR \ldots$ und den entsprechenden Orbits zu unterscheiden, werden letztere durch einen **Punkt** $\circ$ gekennzeichnet. Die Menge aller Orbits, die zur Zeit $t = 0$ rechts von dem Maximum x_c liegen, bezeichnet man als $\circ R$. Analog ist $\circ L$ die Menge aller Orbits $(x_t)_{t\in\mathbb{Z}}$ mit $x_0 \in (0, x_c)$. Eine solche Unterteilung des Phasenraumes $(0,1)$ in zwei Teilintervalle nennt man eine binäre Partition. Eine besondere Rolle spielen die Orbits, die zu

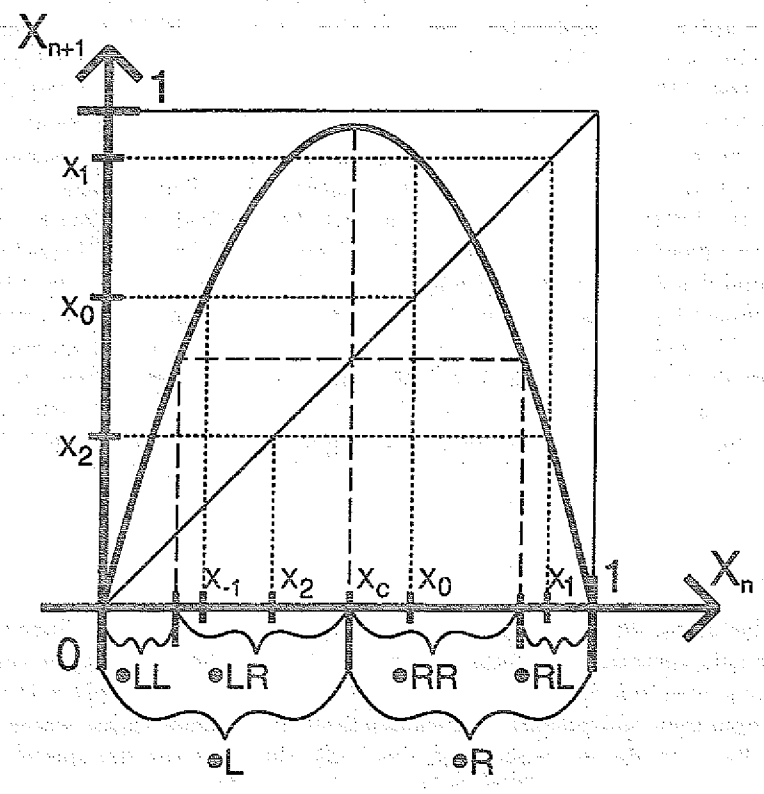


Abbildung 2.2: Beispiel für eine Abbildung $f(x_n) = x_{n+1}$ auf dem Einheitsintervall $[0, 1]$. Das Maximum liegt bei x_c. Durch diesen kritischen Punkt und seine Urbilder wird das Einheitsintervall in kleinere Teilintervalle zerlegt, die mit kurzen Symbolsequenzen gekennzeichnet sind. Ein willkürlich gewählter Orbit $(x_t)_{t\in\mathbb{Z}}$ ist zu den Zeiten $t = -1, 0, 1$ und 2 eingetragen.

irgendeiner Zeit t die Partitionsgrenze schneiden: $x_t = x_c$. Da die Partitionsgrenzen im allgemeinen vom Maß Null sind, spielen diese Orbits bei der Suche nach generierenden Partitionen keine Rolle, obwohl sie für manche Anwendungen wichtig sein können. In dieser Arbeit werde ich nur Orbits behandeln, die keine Partitionsgrenzen schneiden. Die Menge aller dieser Orbits bezeichne ich mit $\circ E$.

Die Lage eines Orbits zu anderen Zeiten $t \neq 0$ wird durch Aneinanderreihung weiterer Symbole gekennzeichnet. Dabei ist $\ldots S_{-2}S_{-1} \circ S_0 S_1 \ldots$ die Menge aller Orbits $(x_t)_{t\in\mathbb{Z}}$ mit

$$S_i = \begin{cases} R & \text{falls } x_i < x_c \\ L & \text{falls } x_i > x_c \end{cases} . \tag{2.1}$$

Wenn man einen Orbit unter f abbildet, so erhält man einen um einen Zeitschritt verschobenen Orbit. Es folgt $f(\ldots S_{-1} \circ S_0 S_1 \ldots) = \ldots S_{-1} S_0 \circ S_1 \ldots$. Die Abbildung f verschiebt also die Symbolsequenz einfach um einen Platz nach links beziehungsweise den Punkt $\circ$ um einen Platz nach rechts.

Obwohl jede mit einem Punkt versehene Symbolsequenz $\circ S$ eine Menge von Orbits ist, werde ich bei invertierbaren Abbildungen f diese Orbitmenge auch manchmal als Punktmenge oder als Gebiet im Phasenraum interpretieren. Diese Sprechweise ist so zu verstehen, daß der ganze Orbit $(x_t)_{t\in\mathbb{Z}}$ mit dem einen Punkt x_0 gleichgesetzt wird. Dies ist möglich, weil bei invertierbaren Abbildung wegen $x_t = f^t(x_0)$ der ganze Orbit aus dem Punkt x_0 folgt.

Bei nichtinvertierbaren Abbildungen folgt nur der halbe Orbit $(x_t)_{t\geq 0}$ aus dem Punkt x_0. Dann wird häufig nur mit halbunendlichen Sequenzen gearbeitet: Die x_0 aller Orbits in der Menge $\circ S_0 S_1 \ldots S_t$ bilden eine durch diese Symbolsequenz definierte Punktmenge im Phasenraum. Und wenn keiner dieser Orbits in der anderen Menge $\circ T_0 T_1 \ldots T_t$ enthalten ist, dann sind nicht nur die Orbitmengen disjunkt, sondern auch die entsprechenden Punktmengen. Für Orbitmengen der Gestalt $S_{-1}\circ$ und $S_{-2}\circ$ gilt dies nicht mehr, weshalb ich solche Orbitmengen auch nur bei invertierbaren Abbildungen als Punktmengen interpretieren werde.

Im allgemeinen Fall kann es n Symbole geben, die mit $A_1 \ldots A_n$ bezeichnet werden sollen. Die entsprechenen offenen Gebiete des Phasenraumes sollen die „Symbolischen Gebiete" $\circ A_1, \ldots \circ A_n$ genannt werden. In der mathematischen Literatur nennt man sie auch die Elemente der Partition $\{\circ A_1, \ldots \circ A_n\}$. Ihre Vereinigung $\bigcup_{i=1}^n \circ A_i$ soll den Attraktor im Phasenraum bis auf eine Menge vom Maß Null überdecken. Die Menge aller betrachteten Orbits ist gleich $\circ E = (\bigcup_{i=1}^n A_i)^{-\infty} \circ (\bigcup_{i=1}^n A_i)^{+\infty}$. Dabei sind $X^{-\infty} = \ldots XXX$ und $X^{+\infty} = XXX \ldots$ halbunendliche Wiederholungen des einen Symbols X. Das Zeichen E (everywhere) kann dann auch als Platzhalter für unbekannte Symbole dienen, wie in $f(A_1\circ) = A_1 E\circ$.

Zum Beispiel kann man aus Abb. 2.2 entnehmen, daß $x_{-1} \in \circ LR$, woraus folgt, daß $x_0 \in f(\circ LR) = L \circ R$. Zusammen mit $x_0 \in \circ RR$ und $x_2 \in \circ LR$ ergibt sich, daß der Orbit in dem Schnitt der Mengen

$$f(\circ LR) \cap \circ RR \cap f^{-2}(\circ LR) = L\circ R \cap \circ RR \cap \circ EELR = L\circ RRLR \tag{2.2}$$

liegt.

2.2.3 Generierende Partitionen in einer Dimension

Was zeichnet eine bestimmte Partition vor anderen aus? Betrachten wir dazu das Beispiel in Abb. 2.2. Wenn das Intervall an der Stelle x_c unterteilt wird, wo f maximal ist, dann ist die Funktion f auf beiden Teilintervallen invertierbar. Mit diesen beiden inversen Funktionen kann man den gesamten Orbit x_t bestimmen, wenn man außer seiner Lage x_0 zum Zeitpunkt $t = 0$ auch noch seine ganze Symbolsequenz kennt. (Durch x_0 selbst ist nur die zukünftige Symbolsequenz festgelegt.) Für einen periodischen Orbit unbekannter Lage und bekannter Symbolsequenz ergibt sich analog eine Gleichung zur Bestimmung von x_0. So kommt man von Symbolsequenzen zurück zu den Koordinaten des Phasenraumes („word–lifting–technique" [70]).

Es stellt sich die Frage, ob Orbits eindeutig durch ihre unendlichen Symbolsequenzen charakterisiert sind. Für die exakte mathematische Behandlung, aber auch für die praktische Kodierung

einer Dynamik (Kap. 2.1) ist dies wichtig. Da aber nicht alle Partitionen diese Bedingung erfüllen, wie wir noch an vielen Beispielen sehen werden, definiert man:

Definition 2.2.1 (Generierende Partition:) *Eine Partition $\{\circ A_1; \ldots \circ A_n\}$ ist dann generierend, wenn bei jeder Wahl der Symbole $S_i \in \{A_1, \ldots A_n\}, i \in \mathbb{Z}$ die Menge $\ldots S_{-2}S_{-1} \circ S_0 S_1 \ldots$ nur höchstens einen Orbit enthält.*

Es können natürlich leere Mengen $\ldots S_{-2}S_{-1} \circ S_0 S_1 \ldots = \emptyset$ auftreten. Derartige Symbolsequenzen heißen **verbotenen**, alle anderen heißen **zulässig**. Ihre Gesamtheit bestimmt die **grammatikalischen Regeln** der Symbolischen Dynamik. Zur vollständigen Beschreibung der **Symbolische Dynamik** fehlen dann nur noch die Wahrscheinlichkeiten aller Symbolsequenzen. Diese hängen von dem ergodischen Maß auf dem Attraktor ab.

Die Definition 2.2.1 stimmt nicht ganz mit dem allgemeinen Sprachgebrauch überein: In allen mir bekannten Veröffentlichungen werden die Symbolischen Gebiete so gewählt, daß ihr Inneres disjunkt ist: $\circ A_i \cap \circ A_j = \emptyset$ für $i \neq j$. In dieser Arbeit dürfen die Symbolischen Gebiete sich schneiden. Einem Orbit können dann viele Symbolsequenzen entsprechen, doch jeder dieser Sequenzen entspricht nur dieser eine Orbit. Zu disjunkten Symbolischen Gebieten kann man jederzeit übergehen, indem man die Schnitte Symbolischer Gebiete mit neuen Symbolen bezeichnet. Ein Vorteil überlappender Symbolischer Gebiete ist, daß dadurch schon in einer Dimension Mehrdeutigkeiten bei der Wahl der generierenden Partition vermieden werden können. Dies wird in Anhang A.1 an einem Beispiel illustriert.

Üblicherweise wird auch noch eine weitere Bedingung in der Definition von generierenden Partitionen gefordert: Die Größe der durch Symbolsequenzen bestimmten Punktmengen soll sogar gleichmäßig gegen Null konvergieren:

$$0 = \lim_{t \to \infty} \max\{ \ \mathrm{diam}(f^t(\circ S_0 \ldots S_{2t})) \ \mid \ S_i \in \{A_1, \ldots A_m\}\} \quad . \tag{2.3}$$

Diese Zusatzbedingung wird im folgenden ignoriert. Denn eine ungleichmäßige Konvergenz wäre in realen Systemen äußerst ungewöhnlich. Und um das Ziel aus Kap. 1.1, eine an die Dynamik angepaßte Strukturierung der Dynamik, zu präzisieren, genügt Def. 2.2.1 in jedem Fall. Mathematiker verwenden diese Zusatzbedingung vor allem deshalb, um zu garantieren, daß die Symbolische Dynamik dieselben dynamischen Entropien besitzt wie die ursprüngliche Dynamik im Phasenraum.

2.2.4 Mathematische Kommentare zur Existenz und Eindeutigkeit generierender Partitionen

An dieser Stelle trennen sich die Wege von reinen Mathematikern und theoretischen Physikern. Der Mathematiker sucht eine exakte Isomorphie zwischen der symbolischen Dynamik und der Dynamik im Phasenraum Ein ernstes Problem sind für ihn die Orbits, denen keine eindeutige Symbolsequenz zugewiesen werden kann, da sie Partitionsgrenzen schneiden (diskutiert in [1]). In der Ergodentheorie setzt man meistens voraus, daß die Partitionsgrenzen das ergodische Maß Null haben, so daß fast allen Orbits eine eindeutige Symbolsequenz zugewiesen werden kann.

Der Physiker dagegen wird argumentieren, daß man in experimentellen Daten die Partitionsgrenzen sowieso nur mit endlicher Genauigkeit festzulegen braucht. Er wird sich nicht um all die komplizierten Phänomene sorgen, die der Attraktor erst bei starker Vergrößerung zeigt. Zum Beispiel könnte in drei oder mehr Dimensionen der Ausnahmefall auftreten, daß eine ganze Linie von Sattelpunkten zum Attraktor gehört. Durch endlich viele Symbole ließen sich diese Fixpunkte nicht mehr unterscheiden. Doch zum Verständnis der globalen Dynamik ist diese Unterscheidung auch nicht wichtig.

Der Mathematiker vereinfacht sich die Aufgabe, indem er beliebige offene Symbolische Gebiete erlaubt. Dann kann er das Krieger Generator Theorem [46] benutzen, das für eine breite Klasse dynamischer Systeme aus Ergodizität die Existenz einer generierenden Partition ableitet. Diese generierende Partition enthält nicht mehr Symbole, als auf Grund der topologischen Entropie h nötig sind ($\exp(h) \leq$ Symbolzahl $\leq \exp(h) + 1$). Die untere Schranke ist schon deswegen nötig, um mit nur einem Symbol pro Zeitschritt genügend viele verschiedene Symbolsequenzen erzeugen

zu können. Schwerer wird die Aufgabe für den Mathematiker dann wieder dadurch, daß er für viele Zwecke eine generierende Partition benötigt, deren Grammatik vollständig ist. „Vollständig" heißt, daß man nicht über eine endliche Länge n hinausgehen muß um festzustellen, ob eine Symbolsequenz erlaubt ist. Eine unendliche Symbolsequenz ist in dieser Grammatik genau dann erlaubt, wenn alle ihre Teilsequenzen der Länge n erlaubt sind. Die Existenz derartiger „Markov–Partitionen" (siehe z. B. [1]) konnte tatsächlich für ganz spezielle dynamische Systeme bewiesen werden, etwa für Automorphismen des Torus auf sich selbst. Aber schon auf dem 3-Torus zeigte Bowen [12], daß die Ränder einer Markov–Partition niemals glatt sind; heute würde er sie wohl fraktal nennen.

Der Physiker widerum kann den Phasenraum einer experimentellen Zeitreihe nur durch relativ einfache, nicht fraktale Symbolische Gebiete strukturieren. In gewissen hyperbolischen dynamischen Systemen, zum Beispiel bei speziellen eindimensionalen Abbildungen oder bei der Bewegung eines Masseteilchens in einer speziellen Geometrie (Billiard), konnten einfache Markov–Partitionen gefunden werden. Dann können elegante und genaue Methoden angewendet werden, die als Entwicklung um instabile periodische Orbits bekannt sind („Cycle Expansion" [3, 4]). Es gilt aber als typisch für dynamische Systeme, daß sie nichthyperbolisch sind. Meistens kann man dann die grammatikalischen Regeln der Symbolischen Dynamik auch nicht durch endlich viele Symbolsequenzen ausdrücken. In diesem Fall kann man zwar immer noch um periodische Orbits entwickeln, doch die numerischen Ergebnisse für Lyapunov-Exponenten oder Dimensionen konvergieren nicht deutlich schneller als bei einfacheren Algorithmen. Am Hénon Attraktor wurden diese Methoden besonders gründlich miteinander verglichen [31, 17]. Instabile periodische Orbits werden daher im folgenden keine wichtige Rolle spielen.

Zur Klassifikation seltsamer Attraktoren und zur Kodierung ihrer Dynamik (Kap. 2.1) genügen generierende Partitionen, die nach irgendwelchen koordinateninvarianten Kriterien die „einfachsten" sind. In invertierbaren Abbildungen sind eine Partition $\{\circ A_1, \ldots \circ A_n\}$ und ihre m-ten Bilder $\{f^m(\circ A_1), \ldots f^m(\circ A_n)\}$ gleich einfach, da sie durch die Koordinatentransformation f^m ineinander übergehen und dieselbe Grammatik haben. Diese Partitionen werden als äquivalent betrachtet.

2.3 Partitionen in zweidimensionalen Phasenräumen

2.3.1 Stabile und Instabile Mannigfaltigkeiten

Im Abschnitt 2.1 wurde schon erwähnt, daß fast jeder Attraktorpunkt stabile und instabile Richtungen besitzt, entlang denen der Tangentialraum im Lauf der Zeit kontrahiert oder expandiert. Die stabilen Richtungen lassen sich im Phasenraum zu Linien, oder allgemeiner zu Mannigfaltigkeiten fortsetzen (siehe z. B. [20]). Diese kontrahieren in der Zukunft exponentiell schnell. In beliebig dimensionalen Phasenräumen definiert man daher:

Definition 2.3.1 (Stabile Mannigfaltigkeit) *Die stabile Mannigfaltigkeit $W^s(x)$ eines Punktes x ist die Menge aller Punkte, deren Orbits in der Zukunft exponentiell schnell gegen den Orbit von x konvergieren:*

$$W^s(x) := \{y | \lim_{t\to\infty} \frac{1}{t} \ln |f^t(x) - f^t(y)| < 0\} \quad . \tag{2.4}$$

Die exponentielle Konvergenzgeschwindigkeit ist keine wichtige Zusatzbedingung. Fast alle Orbits von typischen Abbildungen f haben d_p Lyapunov Exponenten, die von Null verschieden sind, und erfüllen diese Zusatzbedingung automatisch. Wenn die Dynamik f invertierbar ist, dann definiert man analog:

Definition 2.3.2 (Instabile Mannigfaltigkeit) *Die instabile Mannigfaltigkeit $W^u(x)$ eines Punktes x ist die Menge aller Punkte, deren Orbits in der Vergangenheit exponentiell schnell gegen den Orbit von x konvergieren:*

$$W^u(x) := \{y | \lim_{t\to\infty} \frac{1}{t} \ln |f^{-t}(x) - f^{-t}(y)| < 0\} \quad . \tag{2.5}$$

Die Existenz stabiler Mannigfaltigkeiten ist für Fixpunkte x wohlbekannt, wurde aber von Ruelle sogar für fast alle Punkte in Bezug auf das ergodische Maß ρ bewiesen ([62], Theorem 6.3): Wenn f ein Diffeomorphismus auf einem kompakten Raum ist, also stetig differenzierbar mit stetig differenzierbarem Inversen f^{-1}, dann ist für ρ–fast jeden Punkt x die Punktmenge $W^s(x)$ eine differenzierbare Mannigfaltigkeit. Das Beweisschema ist das folgende: Aus dem multiplikativen ergodischen Theorem von Oseledec [57] folgt die Existenz von Lyapunov–Exponenten, das heißt von Richtungen im Tangentialraum, die unter der linearisierten Dynamik exponentiell schnell expandieren beziehungsweise kontrahieren. Im Phasenraum müssen auch die nichtlinearen Terme in der Taylorentwicklung von f um den Orbit von x berücksichtigt werden. Wenn der Abstand zu dem Orbit von x aber mit der Zeit exponentiell klein wird, werden auch diese nichtlinearen Terme exponentiell klein, und ihr summierter Einfluß auf die Lage von W^s konvergiert. Falls f invertiert werden kann, so folgt analog der Existenzsatz für instabile Mannigfaltigkeiten.

Bei bekannter Dynamik f kann man den bei x gelegenen Teil der stabilen Mannigfaltigkeit $W^s(x)$ relativ leicht ausrechnen, indem man erst das Abbild $f^n(x)$ bestimmt, dort in ungefährer stabiler Richtung eine kleine Linie von Punkten verlegt und diese Punktes dann n-mal rückiteriert. Wenn n groß genug ist, dann liegen diese Punkte sehr nahe an der stabilen Mannigfaltigkeit. Der Abstand der Punkte auf der Linie wird so klein gewählt, daß die Lage der Mannigfaltigkeit sich hinreichend genau aus ihnen bestimmen läßt. Da bei der Rückiteration der Abstand der Punkte im allgemeinen exponentiell schnell wächst, müssen bei Überschreiten eines kritischen Abstandes neue Punkte durch lineare Interpolation bestimmt werden. Mit dem hier skizzierten Algorithmus und einem interaktiven Graphiksystem wurden Abbildungen wie 2.3 erstellt. Diese Standardtechnik wird hier nicht weiter erläutert. Sie ist so schnell, daß n viel größer und der kritische Abstand viel kleiner gewählt werden kann, als es für die Auflösung der hier gezeigten Abbildungen nötig ist. Dieselbe Aufgabe wird auch schon durch „public domain"–Programmpaketen („dstool: A Dynamical System Toolkit with an Interactive Graphical Interface" von J. Guckenheimer, M. R. Myers, F. J. Wicklin & P. A. Worfolk am „Center of Applied Mathematics", Cornell Universität, Ithaca NY 14853). und kommerziellen Programmpaketen („Dynamics" von Prof. J. A. Yorke und Mitarbeitern am „Institute for Physical Science and Technology" Maryland Universität, College Park MD 20742; vertrieben vom Springer Verlag) gelöst. Und verbesserte Algorithmen werden in der Literatur diskutiert [41].

In dissipativen Systemen laufen die Orbits des ganzen Attraktionsgebietes auf den Attraktor zu. Wegen der typischerweise exponentiellen Trennung benachbarter Orbits nimmt man an, daß man durch kleine Änderungen des Startpunktes x_0 dem Orbit $(x_t)_{t\in\mathbb{Z}}$ alle möglichen asymptotische Bahnen geben kann. Insbesondere kann man ihn durch diese kleine Änderung gegen jeden anderen Orbit $(y_t)_{t\in\mathbb{Z}}$ konvergieren lassen. Deshalb vermutet man, daß die stabile Mannigfaltigkeit $W^s(y_0)$ eines typischen Attraktorpunktes y_0 im Attraktionsgebiet dicht liegt. Einen exakten Beweis für allgemeine Attraktoren gibt es nicht.

Um in Abbildungen nicht das ganze Attraktionsgebiet schwarz auszufüllen, trage ich nur einzelne Abschnitte der stabilen Mannigfaltigkeiten ein, die ich so wähle, daß sie im Attraktionsgebiet etwa gleichmäßig dicht liegen. Ob es sich dabei um Abschnitte von derselben Mannigfaltigkeit oder von verschiedenen Mannigfaltigkeiten handelt, ist wegen der endlichen Rechengenauigkeit nicht festgelegt. Den Verlauf der nicht eingezeichneten Abschnitte kann man ungefähr erahnen, da stabile Mannigfaltigkeiten sich nicht gegenseitig schneiden dürfen. Normalerweise mißt man experimentelle Daten nur nachdem Transienten abgeklungen sind, das heißt in der unmittelbaren Umgebung des Attraktors. Daraus kann man dann auch nur den Teil der stabilen Mannigfaltigkeit in dieser Attraktorumgebung bestimmen. Zur Illustration werden im folgenden trotzdem große Abschnitte der stabilen Mannigfaltigkeit eingezeichnet. Die Bestimmung der Partitionen selbst kommt ohne diese Illustrationen aus.

2.3.2 Kritische Punkte und Tangentialpunkte

In einer Dimension ist die Lage der generierenden Partition durch die Extrema $f'(x) = 0$ der differenzierbaren Abbildung f festgelegt. Dies gilt auch noch dann, wenn es mehr als einen solchen kritischen Punkt gibt (siehe z. B. [70]). Kann dieser Zusammenhang auf höhere Dimensionen

verallgemeinert werden? Die Standardbeispiele, an denen diese Frage in der Literatur diskutiert wird, sind zweidimensionale, dissipative und invertierbare Abbildungen $f(x_n, y_n) = (x_{n+1}, y_{n+1})$, insbesondere die Hénon Abbildung:

$$x_{n+1} = 1 - a * x_n^2 + y_n \tag{2.6}$$

$$y_{n+1} = b * x_n \quad . \tag{2.7}$$

Für sehr starke Dissipation ähneln sie eindimensionalen nichtinvertierbaren Abbildungen und sind deshalb leichter zu behandeln als Hamiltonsche Systeme.

Die Hénon Abbildung wurde schon 1976 von Hénon [40] eingeführt. Sie dient seitdem als Standardbeispiel für seltsame Attraktoren. Frühe Analysen dieser Abbildung finden sich in [21, 14, 65]. Die Existens eines seltsamen Attraktors wurde kürzlich in [7] bewiesen. Wer sich die Hénon Abbildung unbedingt als physikalisches System vorstellen will, der kann sie (für $b < 0$) als Poincaré Schnitt eines periodisch gekickten nichtlinearen Rotators interpretieren [39].

Welche Eigenschaften der kritischen Punkte lassen sich auf zwei Dimensionen übertragen? Die wichtigste Eigenschaft in Kap. 2.2.3 war die Invertierbarkeit der Abbildung f zwischen kritischen Punkten. Auf den ersten Blick scheint sich dieses Kriterium nicht auf invertierbare Abbildungen übertragen zu lassen. Doch in Kap. 3.2.1 wird gerade das getan werden. In bisherigen Veröffentlichungen wurde versucht, eine andere Eigenschaft kritischer Punkte zu verallgemeinern. Sie betrifft die linearisierte Dynamik im Tangentialraum eines Orbits, der nahe an einem kritischen Punkt x_c vorbeiläuft. Wenn der Orbit einen positen Lyapunov Exponenten hat, dann besteht diese Dynamik aus einer exponentiell schnellen Expansion. Nahe des kritischen Punktes ist aber die Ableitung $f'(x)$ sehr klein, so daß es zu einer abrupten Kontraktion kommt, bevor die Expansion wieder beginnt. Ein Orbit durch den kritischen Punkt x_c gehört zu den Ausnahmen, für die kein Lyapunov Exponent definiert ist: Die vergangene Expansion wird durch eine unendlich starke Kontraktion im Tangentialraum abrupt beendet.

In zwei Dimensionen besitzen typische Orbits einen positiven und einen negativen Lyapunov Exponenten, denen eine expandierende und eine kontrahierende Richtung im Tangentialraum entspricht. Dies sind auch die Richtungen der instabilen und der stabilen Mannigfaltigkeit an der Stelle des Orbits im Phasenraum. Bei nicht periodischen Orbits stehen diese beiden Richtungen in keiner offenkundigen Beziehung zueinander. Denn laut Definition 2.3.1 und 2.3.2 (aus Kap. 2.3.1) ist die stabile Mannigfaltigkeit W^s durch die asymptotische Zukunft und die instabile Mannigfaltigkeit W^u durch die asymptotische Vergangenheit festgelegt. An besonderen Punkten des Phasenraumes können beide Richtungen übereinstimmen. Für solche **Tangentialpunkte** sind die englischen Namen „tangencies“ oder auch „homoclinic tangencies“ gebräuchlich, (was nicht so mißverstanden werden sollte, als seien nur homokline Punkte eines periodischen Orbits gemeint [68]). Wenn stabile und instabile Richtung im Tangentialraum eines Orbits zu einem Zeitpunkt übereinstimmen, dann müssen sie zu allen Zeiten übereinstimmen, und der Orbit besteht aus lauter Tangentialpunkten.

Grassbereger und Kantz [30] schlugen vor, die Partitionsgrenze in zwei Dimensionen so zu legen, daß der Attraktor nur in Tangentialpunkten geschnitten wird. Dabei soll jeder Orbit $(x_t)_{t\in\mathbb{Z}}$ nur zu einem einzigen Zeitpunkt T getroffen werden. Dieser Punkt x_T wurde als **primärer Tangentialpunkt** bezeichnet. Sie konstruierten so eine binäre Partition für den Hénon Attraktor bei den Standardparameterwerten $a = 1.4$ und $b = 0.3$ (Abb. 2.4) sowie bei schwächerer Dissipation $b = 0.54$ und $a = 1.0$ (Abb. 2.6) und testeten numerisch, ob sich die Entropie aus dieser Partition korrekt ergibt. In beiden Fällen kann man sich die Wirkung der Hénon Abbildung so vorstellen, daß der seltsame Attraktor gestreckt und gefaltet wird. Dies wird in Kap. 8 durch Abb. 8.2 und 8.3 illustriert werden. Auch für den Duffing–Oszillator konstruierten Giovannini und Politi [26] eine generierende Partition, indem sie die Partitionsgrenze durch Tangentialpunkte legten.

Schon in [64] wurde ein heuristischer Grund angedeutet, warum die Partitionsgrenzen jeden typischen Orbit von Tangentialpunkten treffen sollte. Man nimmt dazu an, daß stabile und instabile Mannigfaltigkeit verschiedene Krümmung haben, die Tangente also quadratisch ist, und daß benachbarte Abschnitte der stabilen Mannigfaltigkeit in etwa parallel zueinander liegen (Abb. 2.7). Nahe des Tangentialpunktes x_c treten dann doppelte Schnitte (P_i, Q_i) der stabilen mit der

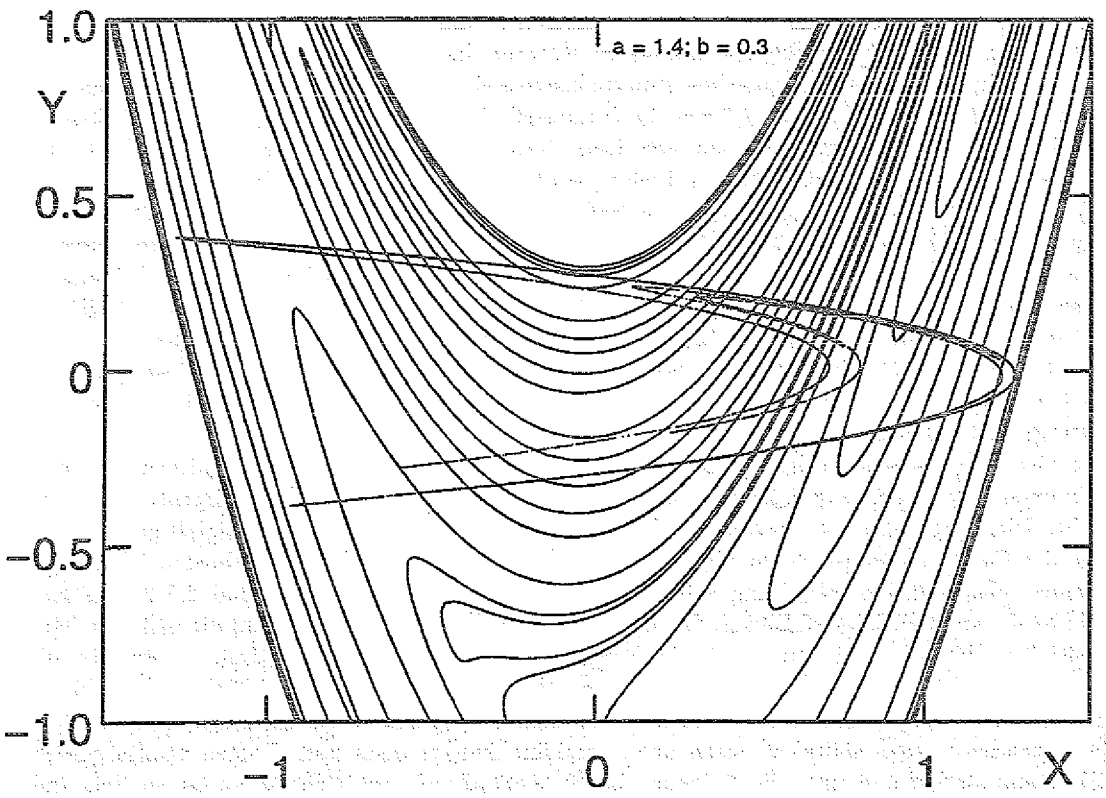


Abbildung 2.3: Der Hénon Attraktor bei Standardparameterwerten $a = 1.4$ und $b = 0.3$. Zehntausend Einzelpunkte des Attraktors zeigen den ungefähr waagrechten Verlauf der instabilen Mannigfaltigkeiten. Willkürlich gewählte Abschnitte der stabilen Mannigfaltigkeiten sind als durchgezogene Linien innerhalb des Attraktionsgebietes eingetragen. Die großen weißen Flächen liegen außerhalb des Attraktionsgebietes.

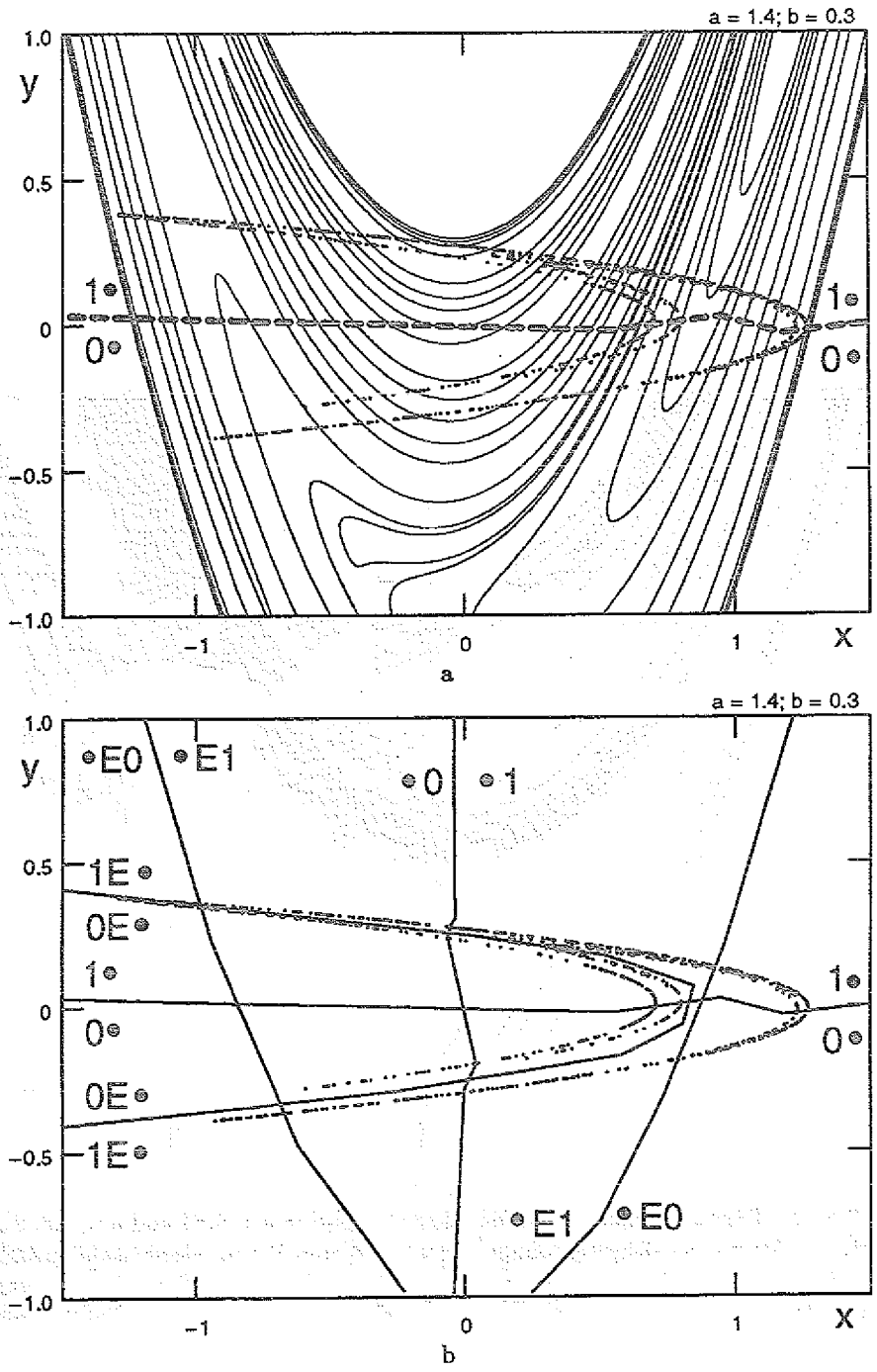


Abbildung 2.4: Die 01–Partition des Hénon Attraktors bei Standardparameterwerten $a = 1.4$ und $b = 0.3$. In Abb. a sind wie in Abb. 2.3 Abschnitte der stabilen Mannigfaltigkeit als dünne Linien eingetragen. Tausend Einzelpunkte kennzeichnen die Lage des Attraktors. Die gestrichelte Linie ist die Partitionsgrenze zwischen den Symbolischen Gebieten $0\circledcirc$ und $1\circledcirc$. Diese Partitionsgrenze wurde in [30] durch alle „primären" Tangentialpunkte gelegt, bei denen stabile und instabile Mannigfaltigkeit parallel und relativ schwach gekrümmt sind. In Abb. b sind außerdem Partitionsgrenzen eingetragen, die $0\bar{E}\circledcirc$ von $1E\circledcirc$ trennen, sowie $\circledcirc 0$ von $\circledcirc 1$ und $\circledcirc E0$ von $\circledcirc E1$. Diese Partitionsgrenzen sind nicht auf ganzer Länge Bilder voneinander. Denn da ihre Lage nur auf dem Attraktor festgelegt ist, müssen diese vier Partitionsgrenzen, wenn sie unter der Dynamik f abgebildet werden, nur auf dem Attraktor ineinander übergehen.

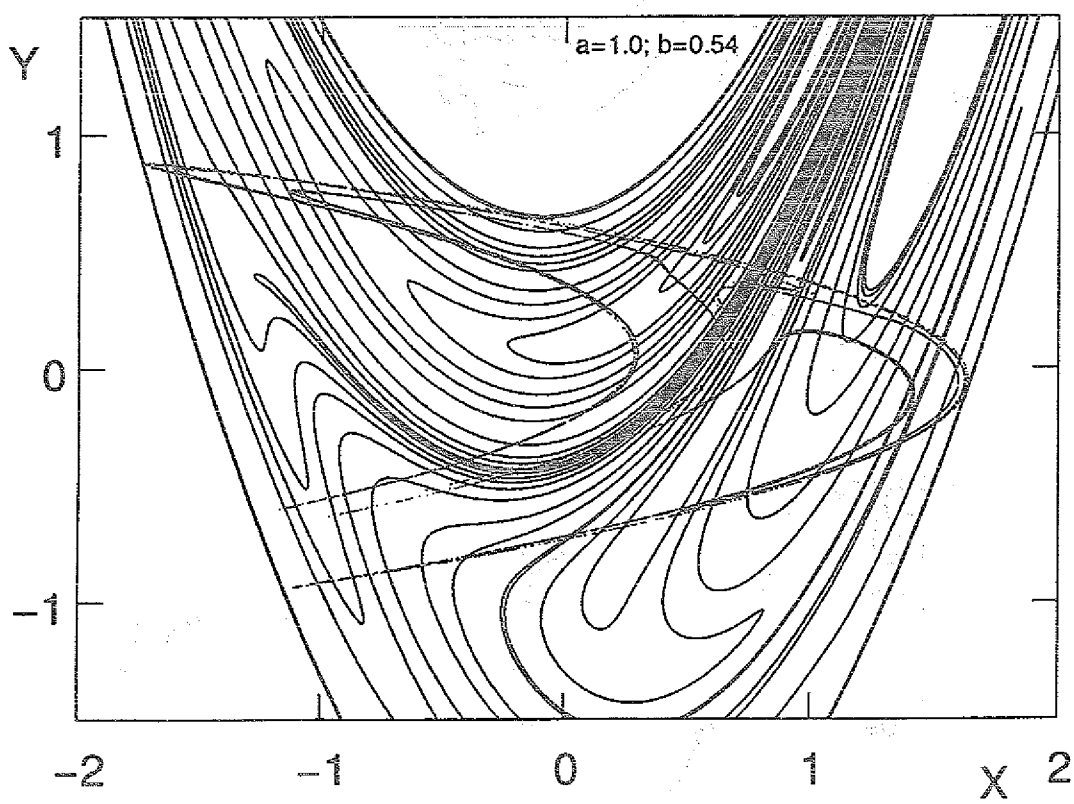


Abbildung 2.5: Der Hénon Attraktor bei schwacher Dissipation $b = 0.54$ und $a = 1.0$. Zehntausend Einzelpunkte des Attraktors zeigen den ungefähr waagrechten Verlauf der instabilen Mannigfaltigkeit. Willkürlich gewählte Abschnitte der stabilen Mannigfaltigkeit sind als durchgezogene Linien innerhalb des Attraktionsgebietes eingetragen. Die großen weißen Flächen liegen außerhalb des Attraktionsgebietes.

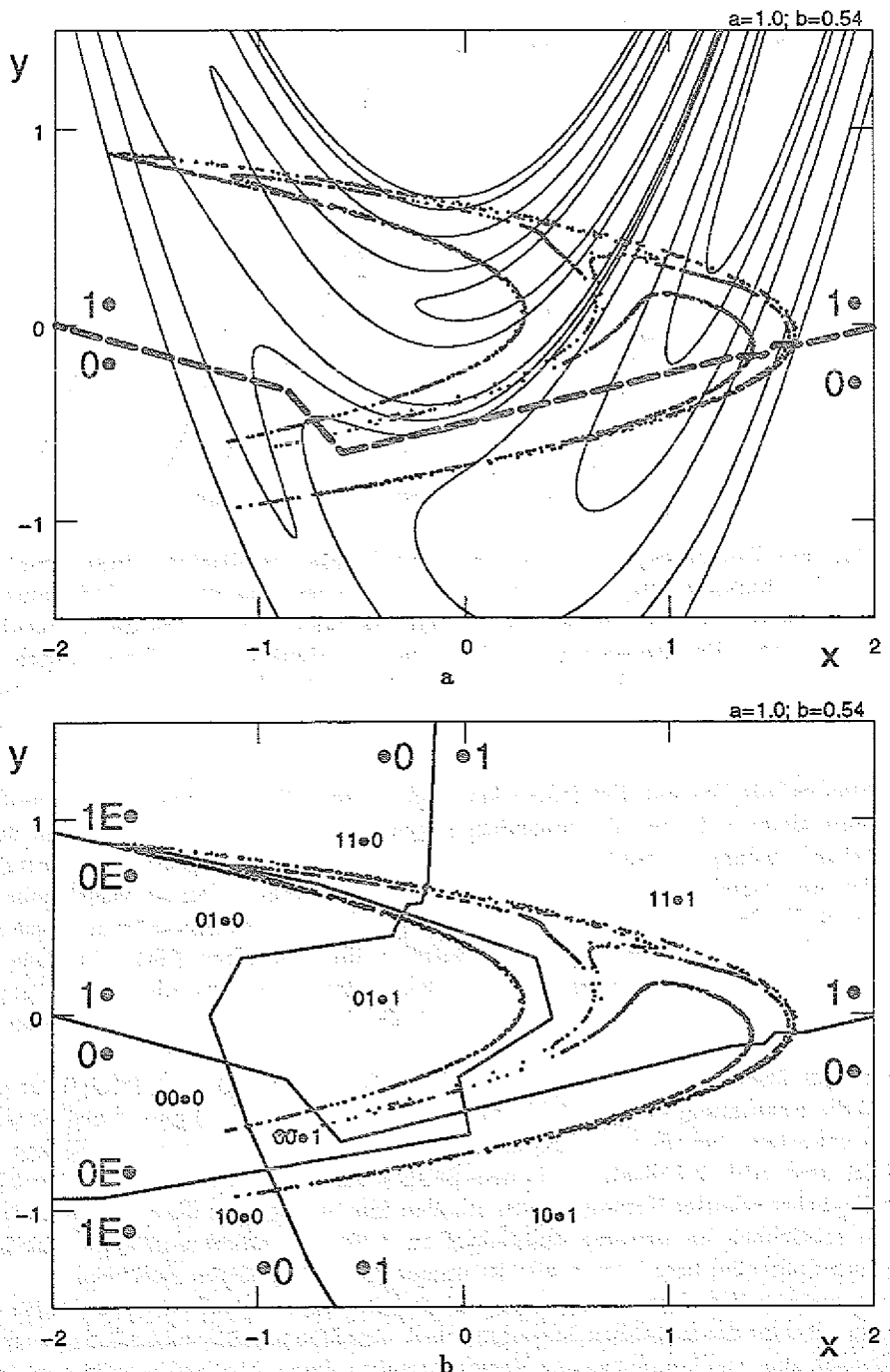


Abbildung 2.6: Die 01–Partition des Hénon Attraktors bei schwacher Dissipation $b = 0.54$ und $a = 1.0$. Tausend Attraktorpunkte und die stabile Mannigfaltigkeit sind wie in Abb. 2.5 eingetragen. Die Partitionsgrenze zwischen $0\circ$ und $1\circ$ wurde in [30] durch alle „primären" Tangentialpunkte gelegt, bei denen stabile und instabile Mannigfaltigkeit parallel und relativ schwach gekrümmt sind. Dadurch ist nur die Lage auf dem Attraktor festgelegt, außerhalb des Attraktors ließe sich die Partitionsgrenze verschieben. Außerdem sind Partitionsgrenzen eingetragen, die $0E\circ$ von $1E\circ$ trennen, sowie $\circ 0$ von $\circ 1$. Diese beiden Partitionsgrenzen schneiden den Attraktor im ersten Abbild oder im ersten Urbild der „primären" Tangentialpunkte.

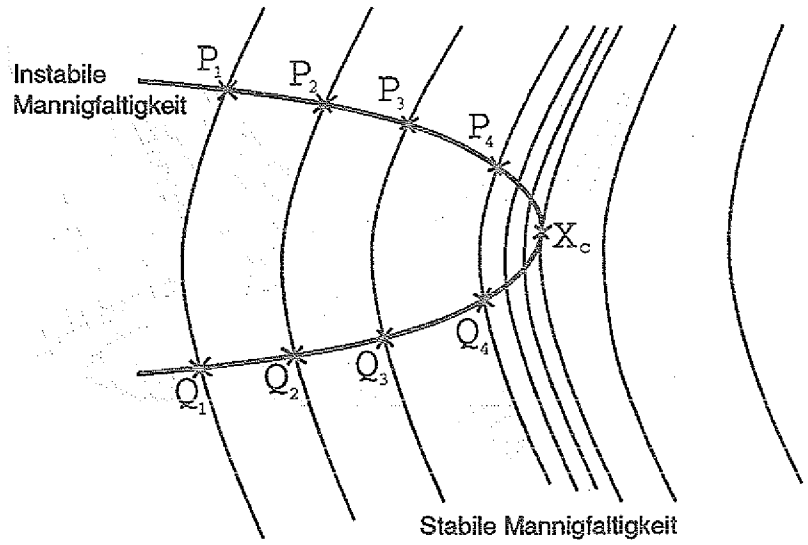


Abbildung 2.7: Ein Tangentialpunkt x_c, dessen instabile Mannigfaltigkeit (dicke Linie) tangential zu seiner stabilen Mannigfaltigkeit (dünne Linie) ist. Außerdem sind die Krümmungen dieser Mannigfaltigkeiten verschieden. Im allgemeinen gibt es dann in der Nähe des Tangentialpunktes eine Folge homokliner Punktepaare (P_i, Q_i), so daß die Punkte jedes Paares durch ein kurzes Stück der stabilen Mannigfaltigkeit und durch ein kurzes Stück der instabilen Mannigfaltigkeit miteinander verbunden sind.

instabilen Mannigfaltigkeit auf. Für jedes i konvergieren die beiden Orbits durch P_i und Q_i sowohl in der Zukunft als auch in der Vergangenheit gegeneinander, sie sind homoklin. In der Zukunft wird ihr Abstand entlang der stabilen Mannigfaltigkeit, in der Vergangenheit ihr Abstand entlang der instabilen Mannigfaltigkeit klein. Zu irgendeinem Zeitpunkt sollen sie verschiedene Symbole haben, sonst ist die Partition nicht generierend. Wegen ihrer Nähe zueinander ist diese Bedingung am einfachsten durch eine Partitionsgrenze zu erfüllen, die zum selben Zeitpunkt t alle $f^t(P_i)$ von allen $f^t(Q_i)$ trennt. Wenn dies der Zeitpunkt $t = 0$ ist, dann ist x_c der „primäre" Tangentialpunkt: Weil sich bei ihm die homoklinen Punktepaare (P_i, Q_i) häufen, muß die Partitionslinie durch ihn verlaufen.

Gleichzeitiges Zusammenlaufen aller homoklinen Punktepaare $(f^t(P_i), f^t(Q_i))$ für großes t erfordert, daß die Krümmung der instabilen Mannigfaltigkeit bei $f^t(x_c)$ immer größer wird. Umgekehrt ist zu erwarten, daß für $t \to -\infty$ die Krümmung der stabilen Mannigfaltigkeit bei $f^t(x_c)$ immer größer wird. Abb. 2.8 illustriert an dem Beispiel eines Orbits, der auf den Attraktor zuläuft, daß bei anfänglicher scharfer Krümmung der stabilen Mannigfaltigkeit die stabile Richtung im Tangentialraum tatsächlich im weiteren Zeitverlauf zunächst expandiert und dann erst kontrahiert. Giovanini und Politi [26] berechneten alle Krümmungen eines anderen Beispiels.

In den numerisch berechneten Abbildungen scheinen die verschiedenen Abschnitte der stabile Mannigfaltigkeit oder der instabilen Mannigfaltikeit ungefähr parallel nebeneinander zu liegen. Im Prinzip könnten aber viel kompliziertere Verschlingungen dieser Abschnitte auftreten. Ihre Häufigkeit wird in Anhang A.2 mit einigen groben Näherungen abgeschätzt. Dort wird argumentiert, daß die numerischen Rechnungen ein zuverlässiges Bild ergeben und kompliziertere Verschlingungen nur einen winzigen Bruchteil des Phasenraumes einnehmen. Denn jeder Bereich des Phasenraumes, in dem die stabile Mannigfaltigkeit stark gekrümmt ist, streckt sich unter der inversen Abbildung f^{-1} langsamer, als seine Breite schrumpft. Analog sind auch Krümmungen der instabilen Mannigfaltigkeit in umso kleineren Bereichen des Phasenraumes konzentriert, je stärker ihre Krümmung ist.

Der plötzliche Wechsel von Expansion auf Kontraktion entlang eines Orbits von Tangentialpunkten hat weitere wichtige Auswirkungen. Erstens hängt die Expansion entlang der instabilen

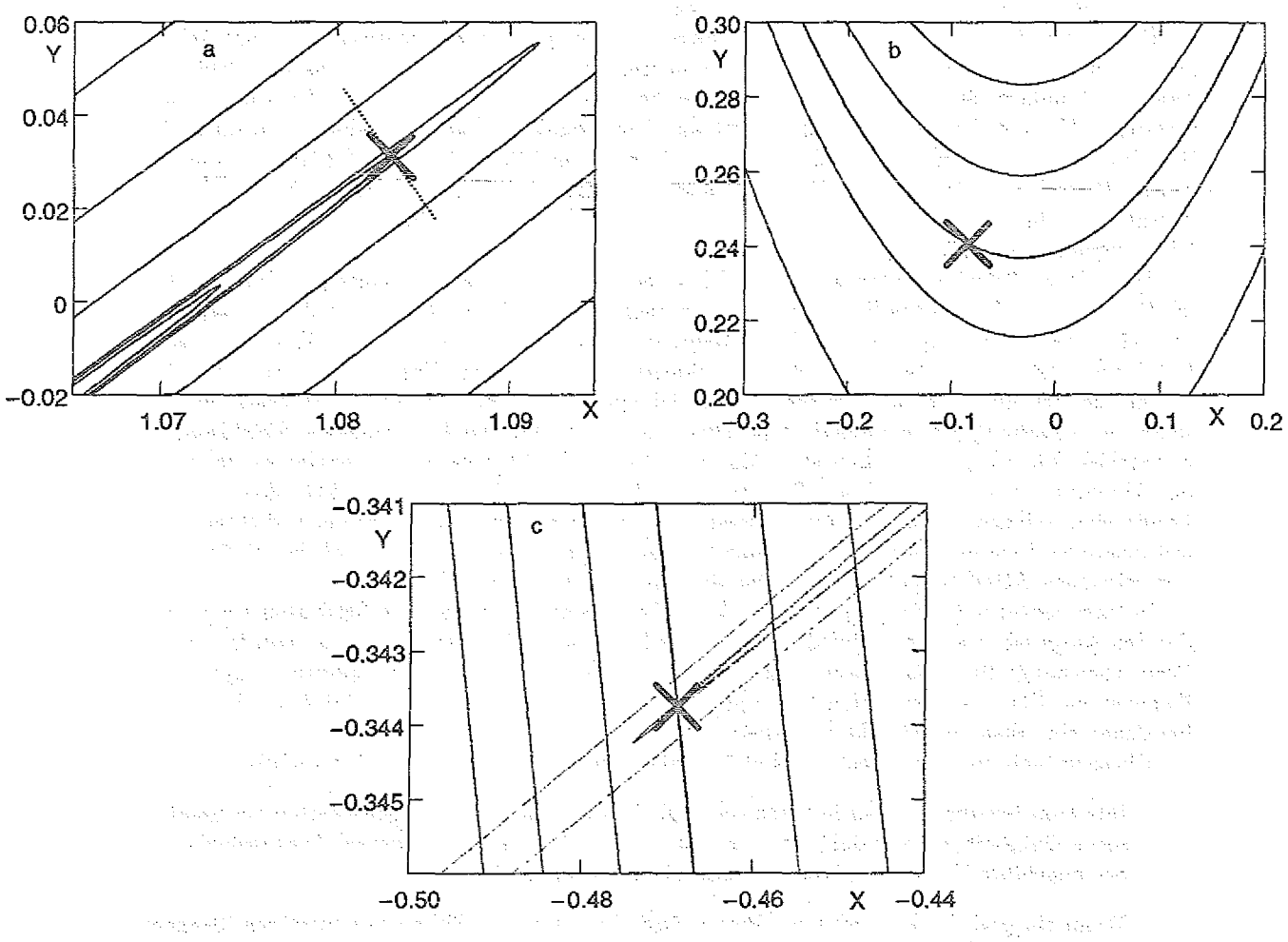


Abbildung 2.8: Der Wechsel von Expansion zu Kontraktion entlang derselben Richtung wird in der Nähe eines Beispielorbits illustriert. Die Lage des Orbits ist zu jedem der drei Zeitschritte durch ein Kreuz gekennzeichet. In Abb. a ist außerdem die stabile Richtung des Orbits als gepunktete Linie eingetragen. Und Abb. c zeigt einzelne Attraktorpunkte. Abschnitte der stabilen Mannigfaltigkeit sind als durchgezogene Linien eingetragen. Zwischen Abb. a und b sowie Abb. b und c liegen jeweils drei Zeitschritte der Hénon Abbildung ($a = 1,4; b = 0,3$). In Abb. a ist die stabile Mannigfaltigkeit des Orbits stark gekrümmt. Diese Krümmung verschwindet während der nächsten drei Zeitschritte, in denen die stabile Richtung des Orbits (gepunktete Linie) um den Faktor 11 expandiert. Erst nach Abb. b, die trotz des großen Maßstabes eine wenig gekrümmte stabile Mannigfaltigkeit zeigt, kann der Tangentialraum entlang der stabilen Richtung kontrahieren, und zwar um den Faktor 1/315 in den nächsten drei Schritten. Dabei wird die instabile Mannigfaltigkeit (Einzelpunkte in Abb. c) gekrümmt.

Richtung eines typischen Orbits empfindlich davon ab, wie nahe an Tangentialpunkten dieser Orbit vorbeiläuft. Zweitens wird das natürliche Maß ρ durch die Kontraktion konzentriert. In einer Dimension führt das zu Wurzelsingularitäten $\rho(x) \sim 1/\sqrt{x}$ bei den Abbildern des kritischen Punktes. In zwei Dimensionen ergibt sich numerisch eine Konzentration des natürlichen Maßes in einem Bereich um die Tangentialpunkte, der etwa so groß ist wie der Krümmungsradius der instabilen Mannigfaltigkeit. Dies wurde sehr detailliert untersucht, als die Erforschung von Multifraktalen und die Berechnung von $f(\alpha)$-Spektren modisch war [29, 59, 58].

Eine dritte Auswirkung der Tangentialpunkte bereitet große mathematische Schwierigkeiten. Da fast alle Orbits dicht auf dem Attraktor liegen, wird wohl auch mindestens ein Orbit von Tangentialpunkten dicht liegen. Deshalb liegen beliebig scharfe Krümmungen der stabilen und instabilen Mannigfaltigkeit vermutlich dicht auf dem Attraktor, und die stabilen und instabilen Richtungen variieren unstetig mit dem Ort. Für derartige nichthyperbolische Attraktoren sind nur wenige Tatsachen exakt bewiesen. Die meisten mathematischen Sätze gelten nur für hyperbolischen Attraktoren. In deren Definition wird unter anderem gefordert, daß die stabilen und instabilen Richtungen stetig variieren.

Trotz dieser Schwierigkeiten gelang es aber Benedicks und Carleson in einer gewaltigen Arbeit die Existenz von chaotischen Hénon Attraktoren nachzuweisen [7]. Auch ihr zentrales Konzept sind kritische Punkte, die eng mit den Tangentialpunkten verwandt sind, aber erst in den zentralen Induktionsschritten des Beweises rekursiv definiert werden können. Die ersten dieser kritischen Punkte können allerdings nur bei sehr starker Dissipation (b sehr nahe bei 0) festgelegt werden, wenn die Abbildung f sich numerisch praktisch nicht von einer eindimensionalen Abbildung unterscheidet. Für ein positives Lebesgue–Maß von Werten des Parameters a beweisen sie dann, daß der Abschluß der instabilen Mannigfaltigkeit eines Fixpunktes ein **seltsamer Attraktor** ist: Alle Orbits einer offenen Umgebung konvergieren gegen ihn, und auf ihm gibt es einen dichten Orbit mit positivem Lyapunov–Exponenten. Man beachte, daß auch andere (aber ähnliche) Definitionen von seltsamen Attraktoren in der mathematischen Literatur zu finden sind.

In einer späteren Arbeit ([8]) wurde auch die Existenz eines sogenannten SRB-Maßes bewiesen, das den Ansprüchen an ein physikalisch sinnvolles Maß entspricht. Unter anderem folgt hieraus die Pesin Gleichung: Die Kolmogorov-Sinai Entropie ist gleich der Summe der positiven Lyapunov–Exponenten. Für eine Dissipation $b = .3$ gibt es zumindest einzelne a–Werte bei $a \approx 1.3924198...$, bei denen ein seltsamer Attraktor existiert [24].

Übrigens sind die „primären" Tangentialpunkte noch aus anderen Gründen wichtig:

- Ihre Lage bestimmt die wichtigsten topologischen und metrischen Eigenschaften der Symbolischen Dynamik, die von der generierenden Partition erzeugt wird. Hieraus folgt insbesondere der ungefähre Verlauf der Korrelationsfunktionen [5, 6].
- Wenn ein periodischer Orbit bei Parameteränderung in die Nähe eines primären Tangentialpunktes wandert, dann kann er stabil werden, was sich deutlich im Bifurkationsdiagramm zeigt [42, 53, 25].
- Nicht nur auf Attraktoren, sondern auch auf chaotischen Satteln können Tangentialpunkte häufig sein [47] und plötzliche Änderungen (Krisen) des Sattels bei Parameteränderung hervorrufen [32].
- Bei der numerischen Berechnung von Lyapunov Exponenten liegt ein Hauptproblem darin, die vorübergehende Kontraktion richtig zu erfassen, die sich nach dem Passieren eines primären Tangentialpunktes einstellt (siehe z. B. [26]). Die Verdichtung des natürlichen Maßes durch diese Kontraktion wurde bereits erwähnt.

2.3.3 Offene Probleme in zwei Dimensionen

Die generierenden Partitionen von Hénon–artigen Attraktoren, die mit Hilfe der Tangentialpunkte konstruiert wurden, sind als Modellsysteme nützlich, an denen Methoden getestet werden können. Doch das Kritierium, den Attraktor in Tangentialpunkten zu schneiden, reicht nicht aus, um in

allgemeinen dynamischen Systemen auf systematische Weise generierende Partitionen zu finden. Genügt es in jedem Fall, den Attraktor mit den Partitionsgrenzen nur in Tangentialpunkten zu schneiden? Und welches sind die „primären" Tangentialpunkte, die geschnitten werden sollen?

Eine Faustregel, die in machen numerischen Simulationen erfolgreich war, wählt die Tangentialpunkte als „primär" aus, bei denen weder die stabile noch die instabile Mannigfaltigkeit besonders stark gekrümmt ist. Genauer: Die Summe der beiden Krümmungen soll kleiner sein als bei allen Abbildern und Urbildern [30]. Diese Definition ist nicht koordinateninvariant: In anderen Koordinatensystemen könnten sich so völlig andere Partitionen ergeben. Und es ist ungewiß, ob eine dieser Partitionen generierend ist. Im Falle des in [26] untersuchten Duffing–Oszillators erhält man nur dann eine generierende Partition, wenn der Attraktor in Tangentialpunkten geschnitten wird, wo die Summe der Krümmungen nicht minimal ist. Und in anderen Fällen muß man vermutlich manche Orbits von Tangentialpunkten mehrmals mit den Partitionsgrenzen schneiden, damit man eine generierende Partition erhält. Beispiele dafür sind zweidimensionale Abbildungen f, die die so stark dissipativ sind, daß sie aussehen wie die eindimensionalen Kreisabbildungen in Anhang A.1, bei denen Schnitte durch alle kritischen Punkte x mit $f'(x) = 0$ nicht ausreichen, um eine generierende Partition zu erzeugen.

Solche mehrfachen Schnitte scheinen auch in höheren Dimensionen unvermeidlich zu sein, falls zum Beispiel die instabile Mannigfaltigkeit zweidimensional ist und sich wie in Abbildung 2.9 faltet. Die Kontraktion des nächsten Zeitschrittes soll senkrecht wirken und so stark sein, daß wir die Feinstruktur der stabilen Mannigfaltigkeit in diesem Beispiel vernachlässigen können. Dann liegen die Tangentialpunkte in etwa auf der Linie, wo die instabile Mannigfaltigkeit eine senkrechte Tangente besitzt, der „cusp"–Linie der Katastrophentheorie. Durch kurze Abschnitte der stabilen Mannigfaltigkeit kann die instabile Mannigfaltigkeit dann dreimal geschnitten werden. Der oberste Punkt P und der unterste Punkt Q liegen auf Orbits, die in Zukunft und Vergangenheit gegeneinander konvergieren. Doch in verschiedenen Symbolischen Gebieten können P und Q nur dann liegen, wenn die Partitionsgrenzen die gezeigte instabile Mannigfaltigkeit auch außerhalb der „cusp"–Linie schneiden. Denn der Weg γ liegt in der instabilen Mannigfaltigkeit und verbindet P und Q, ohne die „cusp"–Linie zu kreuzen.

Wenn man eine generierende Partition gefunden hat, dann lassen sich beliebig viele andere generierende Partitionen finden. Dazu braucht man nur neue Symbolische Gebiete durch Kombination der alten Symbolischen Gebiete konstruieren, so daß sich jedes alte Symbolische Gebiete durch eine Sequenz neuer Symbole ausdrücken läßt. Beispiele dafür werden noch viele kommen. Gibt es eindeutig einfachste generierende Partitionen? Eine einzige generierende Partition, die nach koordinatenunabhängigen Kriterien die eindeutig einfachste ist, kann es nicht immer geben. Denn schon eine Partition $\{\circledast A_1, \ldots \circledast A_n\}$ und ihre Bilder $\{f^t(\circledast A_1), \ldots f^t(\circledast A_n)\}$ müssen nach solchen Kriterien gleich einfach sein. Weniger trivial ist der Fall, wo die Abbildung $f = g^2$ das Quadrat einer anderen invertierbaren Abbildung g mit generierender Partition $\{\circledast B_1, \ldots \circledast B_m\}$ ist. Die Symbolischen Gebiete $\circledast B_i \cap g^{-1}(\circledast B_j)$ bilden dann eine generierende Partition von f, die ebenso einfach ist wie die mit den Symbolischen Gebieten $g(\circledast B_i) \cap \circledast B_j$, sich aber nicht durch f auf sie abbilden läßt.

Ein weitere Typ von Dynamik, der mehrere gleich einfache generierende Partitionen zuläßt, ist definiert durch das direkte Produkt $f = g \otimes h$ zweier invertierbarer Funktionen g und h, die auf Räumen kleinerer Dimension operieren. Wenn $\{\circledast A_1, \ldots \circledast A_n\}$ eine generierende Partition von g und $\{\circledast B_1, \ldots \circledast B_m\}$ eine generierende Partition von h ist, dann bilden für jedes t die Symbolischen Gebiete $\circledast A_i \otimes f^t(\circledast B_j)$ eine generierende Partition von f. Diese Beispiele für gleich einfache generierende Partitionen sind noch relativ harmlos, denn jede Partition ergibt dieselbe Symbolische Dynamik. Sie waren sogar eine wesentliche Motivation für diese Arbeit: Wenn man aus experimentellen Zeitreihen all diese Partitionen ablesen könnte, dann könnte man sehen, ob eine Dynamik in unabhängige Teildynamiken zerfällt. Bisher gibt es noch keine gute Methode, näherungsweise unabhängige Teildynamiken in einer chaotischen Zeitreihe zu erkennen, obwohl so eine Methode gerade für räumlich ausgedehnte Systeme sehr interessant wäre.

Es gibt aber auch generierende Partitionen, die etwa gleich einfach sind aber verschiedene Grammatiken besitzen. Solche fanden Grassberger, Kantz und Mönig [31], als sie auf den Hénon Attraktor bei starker Dissipation (a=1.0; b=0.54) einen Algorithmus von Biham und Wenzel [9]

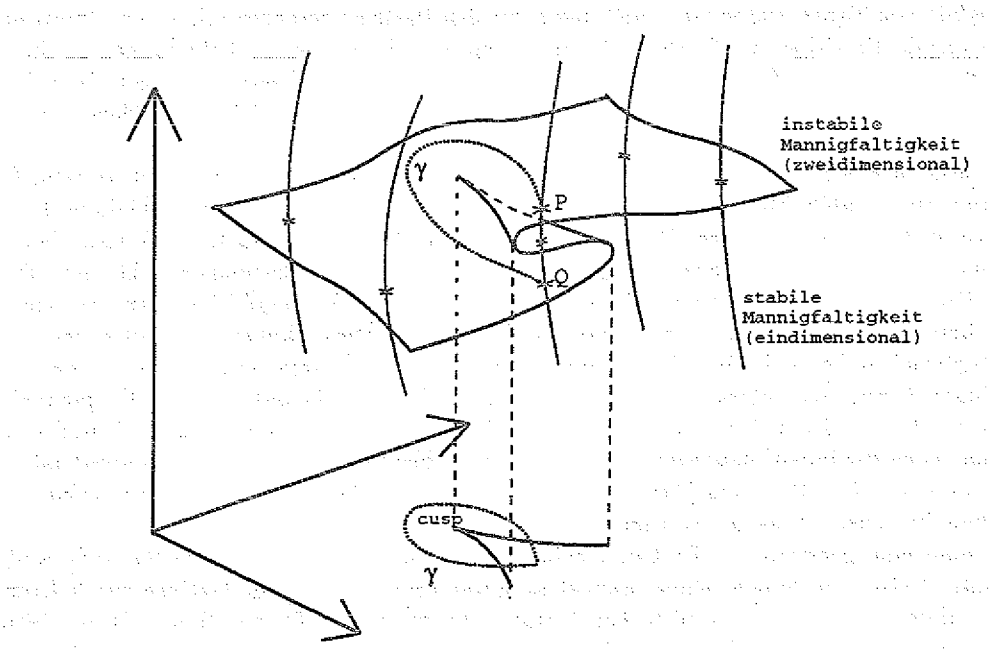


Abbildung 2.9: Ein schematisches Beispiel für die Notwendigkeit, Partitionsgrenzen nicht nur durch Tangentialpunkte zu legen. Die zweidimensionale instabile Mannigfaltigkeit wird im dreidimensionalen Phasenraum von der eindimensionalen stabilen Mannigfaltigkeit geschnitten, die tendenziell von unten nach oben verläuft. Der gezeigte Abschnitt der instabilen Mannigfaltigkeit ist so gefaltet, daß die nach unten projizierten Tangentialpunkte ungefähr auf einer cusp–förmigen Kurve (durchgezogene Linie) liegen. Ein typischer Abschnitt der stabilen Mannigfaltigkeit kann dann die instabile Mannigfaltigkeit in drei nahe benachbarten Punkten schneiden. Der oberste Punkt P und der unterste Punkt Q können entlang der instabilen Mannigfaltigkeit durch eine Kurve γ (gepunktete Linie) verbunden werden, die allen offensichtlichen Tangentialpunkten aus dem Weg geht.

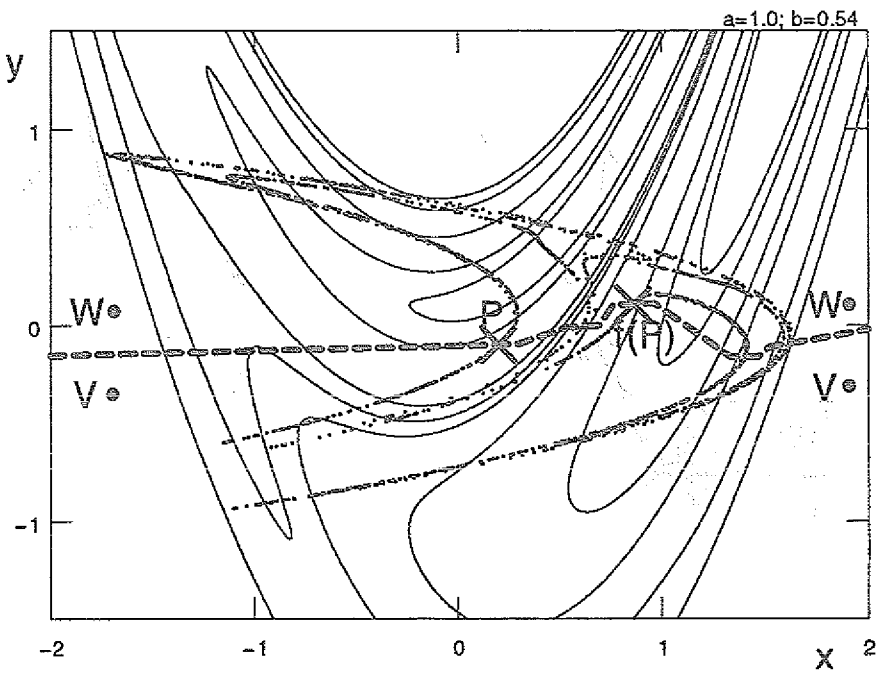


Abbildung 2.10: Die VW–Partition des Hénon Attraktors bei schwacher Dissipation, die in [31] berechnet wurde. Sie ist in ebenso guter Näherung eine generierende Partition wie die 01–Partition in Abb. 2.6. Attraktorpunkte und Abschnitte der stabilen Mannigfaltigkeit sind wie in Abb. 2.6 eingetragen. Die gestrichelten Linie ist die Partitionsgrenze, die $V\circ$ von $W\circ$ trennt.

anwendeten. Abb. 2.10 zeigt die VW-Partitionslinie und Abb. 2.11 eine Vergrößerung davon. Die Schnittpunkte der Partitionslinie mit dem Attraktor bei $f(P)$ scheinen die Abbilder der Schnittpunkte bei P zu sein. Dann läßt sich jede VW-Symbolsequenz durch die Regeln

$$\circ 0 \;=\; W\circ V \;\cup\; V\circ VVV \tag{2.8}$$

$$\circ 1 \;=\; \circ W \;\cup\; V\circ VW \;\cup\; V\circ VVW \tag{2.9}$$

in eine 01–Symbolsequenz (der Abb. 2.6) übersetzen. Die VW-Partition ist daher generierend, falls die 01–Standardpartition in Abb. 2.6 generierend ist. Und die Schnittpunkte der 01–Partitionslinie mit dem Attraktor sind durch Schnittpunkte der VW–Partitionslinie oder deren Urbilder festgelegt.

Regeln für die Rückübersetzung von der 01–Partition in die VW–Partition sind dagegen nicht vollständig bekannt [31]. Insbesondere scheint die Lage der Schnittpunkte bei P nicht genau festgelegt zu sein: Es müssen nicht einmal Tangentialpunkte sein. In den numerischen Daten von Abb. 2.11 sind bei $f(P)$ tatsächlich keine Tangentialpunkte zu erkennen, deren stabile und instabile Mannigfaltigkeiten nur wenig gekrümmt wären. Durch Verschieben der Schnittpunkte bei P und bei $f(P)$ kann man eine ganze Schar von VW–Partitionen erzeugen. Man braucht irgendwelche Kriterien der Einfachheit, um eine bestimmte Lage der Schnittpunkte bei P und somit eine bestimmte VW–Partition auszuzeichnen.

Mehrere übliche Kriterien zeichnen die in Abb. 2.12 gezeigte XY–Partitionslinie aus.

$$\circ X \;=\; \circ 0 \;\cup\; 01\circ 1 \;\cup\; 0\circ 11 \tag{2.10}$$

$$\circ Y \;=\; 0\circ 10 \;\cup\; 11\circ 1 \tag{2.11}$$

In diesem speziellen Fall der VW–Partitionen wird der Schnitt bei $f(P)$ überflüssig, und es gibt nicht mehr Schnittpunkte als bei der 01–Partition. Vermutlich wird jeder Orbit von Tangentialpunkten daher von den Partitionsgrenzen nur zu einer Zeit getroffen. Die Rückübersetzung in

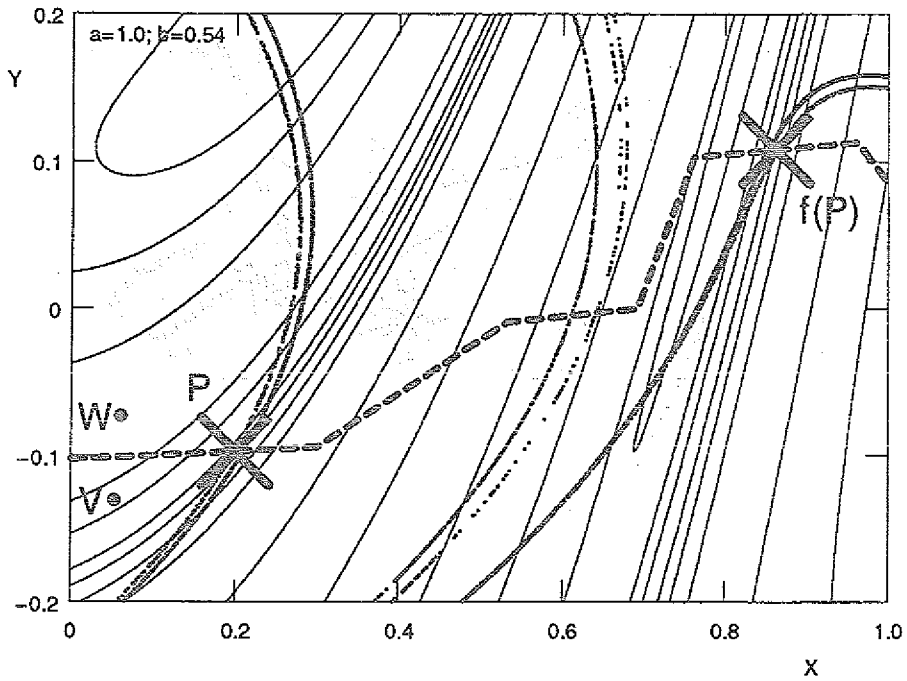


Abbildung 2.11: Ein vergrößerter Ausschnitt aus Abb. 2.10 mit mehr Attraktorpunkten und etwas dichteren Abschnitten der stabilen Mannigfaltigkeit. Die VW–Partitionsgrenze schneidet den Hénon Attraktor bei P und $f(P)$. Dort sind aber keine Tangentialpunkte zu erkennen, deren stabile Mannigfaltigkeit und instabile Mannigfaltigkeit nur schwach gekrümmt sind.

01–Sequenzen ist wie in Gl. 2.8 und 2.9.

$$\circledast 0 \;=\; Y \circledast X \;\cup\; X \circledast XXX \tag{2.12}$$

$$\circledast 1 \;=\; \circledast Y \;\cup\; X \circledast XY \;\cup\; X \circledast XXY \tag{2.13}$$

Wenn man einen Attraktor durch seine generierende Partitionen charakterisieren will, so ist es praktisch, wenn nur eine generierende Partition angegeben werden muß, weil sie eindeutig die einfachste ist. In diesem Fall ist es nicht leicht zu entscheiden, ob die 01–Partition oder die XY–Partition einfacher ist. Im Kapitel 4.1.2 wird das Kriterium 4.1.2 der Einfachheit definiert, dem zufolge die 01–Partition einfacher als die XY–Partition ist. Doch auch dieses Kriterium wird nicht immer eine generierende Partition auswählen können, die eindeutig die einfachste ist. Zum Beispiel wird sich die in Abb. 7.1 gezeigte Partition nach allen mir bekannten koordinatenunabhängigen Kriterien als genauso einfach erweisen wie die Standardpartition (Abb. 2.4) des Hénon Attraktors bei $b = 0.3$ und $a = 1.4$.

Ein naheliegender Versuch, die „primären“ Tangentialpunkte auszuzeichnen, beruht auf stetigen Parameteränderungen. Im eindimensionalen Grenzfall $b = 0$ ist der kritische Punkt ausgezeichnet und die Symbolische Dynamik eindeutig. Entlang eines stetigen Weges im Parameterraum kann diese Symbolische Dynamik auf Werte $b \neq 0$ übertragen werden. Dieser Versuch läuft in unüberwundene Schwierigkeiten. Die Fenster im Parameterraum, in denen kein seltsamer Attraktor sondern nur ein chaotischer Sattel existiert (siehe z. B. [53]), sind noch relativ harmlos, verglichen mit einem Ergebnis von Hansen [35]: Die Symbolische Dynamik, die man erhält, hängt von dem Weg im Parameterraum ab. Und Giovanini und Politi [27] zeigten, daß man die Lage „primärer“ Tangentialpunkte wohl nicht stetig mitverfolgen kann: Sie springen unstetig bei Kollision mit nicht „primären“ Tangentialpunkten.

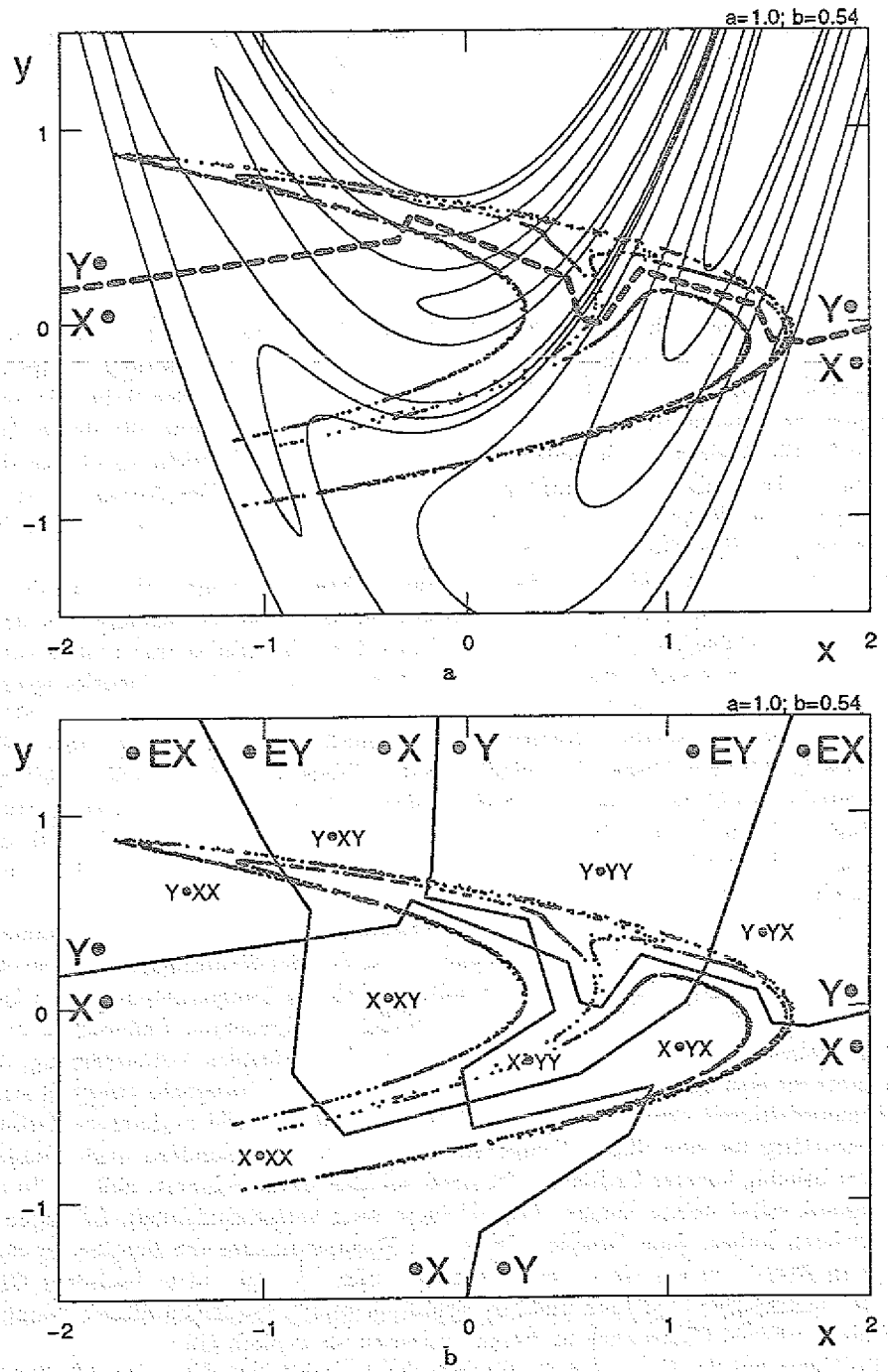


Abbildung 2.12: Die XY–Partition des Hénon Attraktor bei schwacher Dissipation. Sie entsteht aus der VW–Partition in Abb. 2.10 dadurch, daß die Partitionsgrenze an den Stellen P und $f(P)$ so weit verschoben wird, bis bei $f(P)$ der Schnitt mit dem Attraktor entfällt. Die XY–Partition ist in ebenso guter Näherung eine generierende Partition wie die 01–Partition in Abb. 2.6. In Abb. a sind die Attraktorpunkte und die Abschnitte der stabilen Mannigfaltigkeit wie in Abb. 2.6.a eingetragen. Abb. b zeigt die Partitionsgrenzen bei drei aufeinander folgenden Zeiten. Sie trennen $\circ EX$ von $\circ EY$, $\circ X$ von $\circ Y$, bzw. $X\circ$ von $Y\circ$.

2.3.4 Vernachlässigung der Feinstruktur

Die offenen Probleme des letzten Abschnittes lassen sich in zwei Kategorien unterteilen. Probleme mit der Grobstruktur und Probleme mit der Feinstruktur des Attraktors. Abb. 2.13 illustriert den Unterschied anhand der Ikeda Abbildung [44] bei Standardparameterwerten:

$$x_{n+1} = 1.0 + \mu \cdot (x_n \cdot \cos(t_n) - y_n \cdot \sin(t_n)) \quad (2.14)$$

$$y_{n+1} = \mu \cdot (x_n \cdot \sin(t_n) + y_n \cdot \cos(t_n)) \quad (2.15)$$

$$\text{mit} \quad t_n = 0.4 - 6.0/(1 + x_n^2 + y_n^2) \quad (2.16)$$

$$\text{und} \quad \mu = 0.9 \quad (2.17)$$

Diese Abbildung f ist das grob vereinfachte Modell eines lasergepumpten Resonators mit nichtlinearem optischen Medium. Sie wirkt im Phasenraum ungefähr wie eine Faltung, die im Gegensatz zur Hénon Abbildung die Orientierung erhält ($\det(df/dx) > 1$). Die beiden Teile, die aufeinander gefaltet werden, sind durch die Partitionsgrenze in Abb. 2.13.a getrennt, die durch Tangentialpunkte verläuft. Ein Verlauf durch andere Tangentialpunkte würde nicht so gut zu der naiven Interpretation der Dynamik als „Faltung" passen. Die Grobstruktur des Attraktors zu verstehen heißt also die generelle Wirkung der Dynamik f so zu interpretieren, daß man die ungefähre Lage der primären Tangentialpunkte bestimmen kann.

Die Feinstruktur des Attraktors ist dagegen nur wichtig, wenn man die Lage der primären Tangentialpunkte ganz genau kennen will. Dabei können komplexe Verschlingungen der stabilen und instabilen Mannigfaltigkeit wichtig werden. Wegen dieser Verschlingungen ist es zum Beispiel unmöglich zu beweisen und nicht einmal besonders wahrscheinlich, daß die Partitionsgrenzen differenzierbare Linien sind [7]. Und es kann sogar passieren, daß sich bei Betrachtung der Feinstruktur erweist, daß die Partiton zusätzliche Symbole besitzen muß, um generierend zu sein. Als Beispiel dafür zeigt Abb. 2.13.b einen kleinen Ausschnitt aus dem Ikeda Attraktor, wo die stabile und instabile Mannigfaltigkeit so kleine Winkel einschließen, daß zusätzliche Orbits von Tangentialpunkten auftreten könnten. Diese müßten dann durch zusätzliche Partitonsgrenzen (wie z. B. die in sich geschlossene gestrichelte Linie von Abb. 2.13.b) geschnitten werden, die ein kleines zusätzliches Symbolisches Gebiet begrenzen würden.

Durch solche Feinstrukturen wird die exakte mathematische Behandlung von seltsamen Attraktoren erschwert. Zum Beispiel ist durch nichts garantiert, daß der Häufungspunkt x_c in Abbildung 2.7 eine stabile Mannigfaltigkeit besitzt und somit wirklich ein Tangentialpunkt sein kann. Gonchenko, Turarev und Shilnikov [28] zeigten, daß ähnlich unangenehme Phänomene in manchen Parameterintervallen dicht liegen. Ihr Beweis beruht auf der plausiblen Voraussetzung, daß bei einem Parameterwert eine bestimmte quadratische („Newhouse"–) Tangente zwischen stabiler und instabiler Mannigfaltigkeit existiert. Nicht nur unendlich viele stabile periodische Orbits existieren dann gleichzeitig für eine dichte Menge von Parameterwerte, sondern auch Tangenten von kubischer oder beliebig höherer Ordnung. Der Leser sei aber daran erinnert, daß sich in der Nichtlinearen Dynamik selbst dichte Mengen (von kleinem oder verschwindendem Lebesgue Maß) als untypisch erwiesen haben. Zum Beispiel gibt es im Parameterraum die Fenster, in denen statt eines seltsamen Attraktors nur ein chaotischer Sattel existiert, weil ein periodischer Orbit stabil ist. Wenn die dynamischen Systemen nicht hyperbolisch sind, dann liegen diese Parameterfenster vermutlich dicht, obwohl Chaos auch in diesen Systemen als typisch gilt.

Solche Probleme mit der Fein– und der Grobstruktur wecken Zweifel an der Möglichkeit, „einfachste" generierende Partitionen zu finden. Mit den Methoden dieser Arbeit kann man die Grobstruktur auf effiziente Weise vereinfachen. Die Probleme mit der Feinstruktur treten in diesem Zugang zum Teil nicht auf, zum Teil werden sie absichtlich übergangen. Denn wenn solche Orbits, wie die durch P_i und Q_i in Abb. 2.7 festgelegten, zu allen Zeiten sehr nahe benachbart sind, dann ist die mathematische Forderung 2.2.1, sie sollten verschiedene Symbolsequenzen haben, übertrieben streng. Schon wegen des Rauschens kann man im Experiment die Lage eines Orbits nicht genauer bestimmen, als es eine Partition $\{\circledast F_1, \ldots \circledast F_m\}$ kann, die den Phasenraums in viele kleine Gebiete unterteilt. Dabei kann m natürlich sehr groß sein. Wie schon in der Einleitung erwähnt, ist die entscheidende Frage daher nicht „Gibt es generierende Partitionen", sondern „Wie läßt

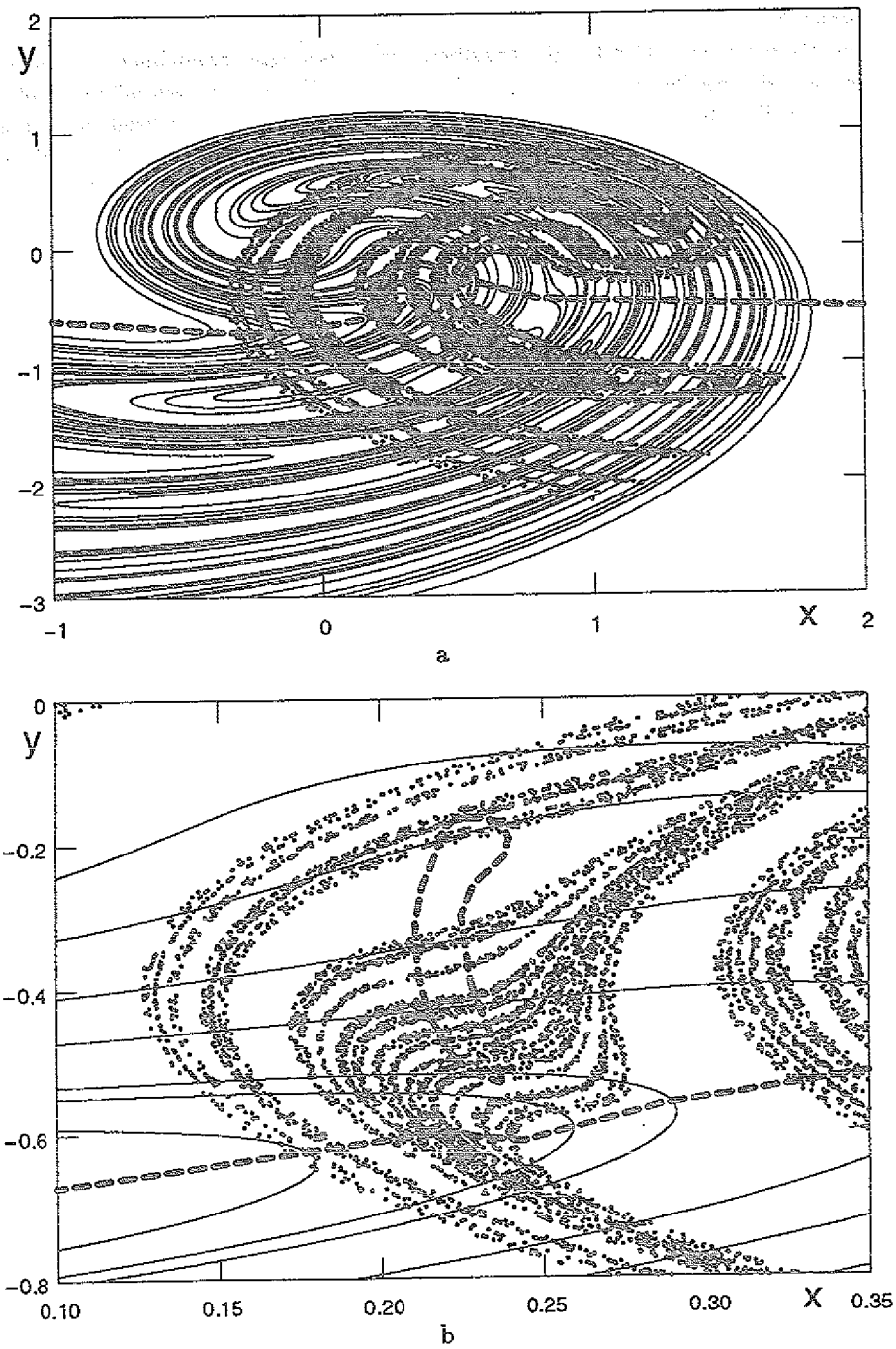


Abbildung 2.13: Die Grobstruktur (Abb. a) und die Feinstruktur (Abb. b) der Ikeda Abbildung. Der Attraktor ist durch zehntausend Einzelpunkte dargestellt; Abschnitte der stabilen Mannigaltigkeit sind durch dünne Linien dargestellt. Die gestrichelte Linie in Abb. a ist eine Partitionsgrenze, die durch Tangentialpunkte verläuft. Sie paßt zu der groben Vorstellung, daß die Dynamik wie eine Faltung wirkt. Abb. b zeigt einen vergrößerten Ausschnitt. Außer bei der Partitionsgrenze von Abb. a (waagrechte gestrichelte Linie) sieht man auch an anderen Stellen nur einen sehr kleinen Winkel zwischen stabiler und instabiler Mannigfaltigkeit (innerhalb des gestrichelt umrahmten Gebietes). Ob dort zusätzliche Tangentialpunkte entstehen, ist bei der verwendeten Rechengenauigkeit nicht festzustellen.

sich die gute Näherung einer generiernden Partition vereinfachen, ohne daß Information über die Orbits verlorengeht?“.

Wenn eine Partition zwei Orbits mit derselben Symbolsequenz bezeichnet, so werde ich sagen, daß die Partition die beiden **Orbits verwechselt**. Das Ziel der Vereinfachung läßt sich dann ausdrücken als ein Partition $\{\circ C_1, \ldots \circ C_m\}$, die nicht mehr Orbits verwechselt, als die anfängliche Partition $\{\circ F_1, \ldots \circ F_m\}$. Die Vereinfachungsmethode kann auch auf nicht generierende Partitionen $\{\circ F_1, \ldots \circ F_m\}$ angewendet werden. Um die Schreibweise aber nicht unnötig zu verkomplizieren, werde ich aber jede Partition $\{\circ C_1, \ldots \circ C_m\}$, die nicht mehr Orbits verwechselt als die anfängliche Partition $\{\circ F_1, \ldots \circ F_m\}$, als generierend bezeichnen.

'Who controls the past,' ran the Party slogan,
'controls the future: who controls the present,
controls the past.'
aus George Orwell: „1984"

Kapitel 3

Bündel von Blättern

In den folgenden Kapiteln wird die neue Methode zum Vereinfachen generierender Partitionen vorgestellt. Sie beruht darauf, daß eine vorgegebene, als generierend unterstellte Partition schrittweise abgeändert wird. Dabei wird sie einfacher, zumindest nach dem Kriterium der Einfachheit, das in Kap. 4 eingeführt wird.

Die Schwierigkeit bei der Wahl der Methode bestand darin, daß es sehr viele ähnliche Möglichkeiten gibt, die Konzepte der Symbolischen Dynamik zu verallgemeinern. Nachdem ich die Möglichkeiten ausgeschlossen hatte, die bei praktischen Rechnungen an Hindernissen scheitern, konnte ich die erfolgreiche Methode zwar einfach definieren. Doch auch im nachhinein konnte ich keinen einfachen Weg sehen, der mich zu dieser Methode geführt hätte. Deshalb ist die Wahl der Methode nur mit einigen allgemeinen Bemerkungen zu motivieren und erst durch erfolgreiche Beispielrechnungen zu rechtfertigen. Ich selber fand die Konzepte „Blätter" und „Bündel von Blättern" erst auf dem Umweg über die Suche nach räumlicher Ordnung, die in Kap. 8 diskutiert wird. Und nur auf diesem Umweg könnte ich jeden einzelnen Schritt motivieren. Die Schreibweise der Formeln habe ich erst im nachhinein an die Rechnungen anpassen können, als sich das neue Kriterium der Einfachheit bereits bewährt hatte. Die Darstellung folgt dem ursprünglichen Erkenntnisweg daher antichronologisch.

3.1 Das grundlegende Problem

3.1.1 Bei welchen Änderungen bleibt eine Partition generierend?

Das Hauptproblem beim Abändern generierender Partitionen liegt darin begründet, daß ein Orbit im Laufe der Zeit beliebig oft in jedes Symbolischen Gebiet zurückkehren kann. Die unendliche Symbolsequenz des Orbits kann sich daher an beliebig vielen Stellen ändern, wenn ein einziges Symbolisches Gebiet abgeändert wird. Deshalb ist es schwer festzustellen, ob zwei verschiedene unendliche Symbolsequenzen auch noch nach der Änderung verschieden sind. Und nur wenn dies für alle möglichen Symbolsequenzen festgestellt werden kann, bleibt die Partition generierend.

Das Beispiel der nicht generierenden KL–Partition

Nehmen wir zum Beispiel an, die 01–Partition des Hénon Attraktors bei schwacher Dissipation in Abb. 2.6 sei generierend. Abb. 3.1 zeigt eine ähnliche Partition, die KL–Partition, die den Attraktor entweder in demselben Punkt schneidet wie die 01–Partition oder in dessen Bildpunkt. Die 01–Symbolsequenz jedes Orbits läßt sich in eine KL–Sequenz übersetzen:

$$\circ K = 1\circ 0 \cup 00\circ \tag{3.1}$$

$$\circ L = 1\circ 1 \cup 10\circ \ . \tag{3.2}$$

(Die Vereinigung $\cup$ ist wie in den vorangegangenen Kapiteln als Vereinigung von Orbitmengen zu verstehen.) Ist diese abgeänderte binäre Partition generierend? Oder gibt es vielmehr KL–Symbolsequenzen, die sich nicht eindeutig in 01–Symbolsequenzen rückübersetzen lassen? Letz-

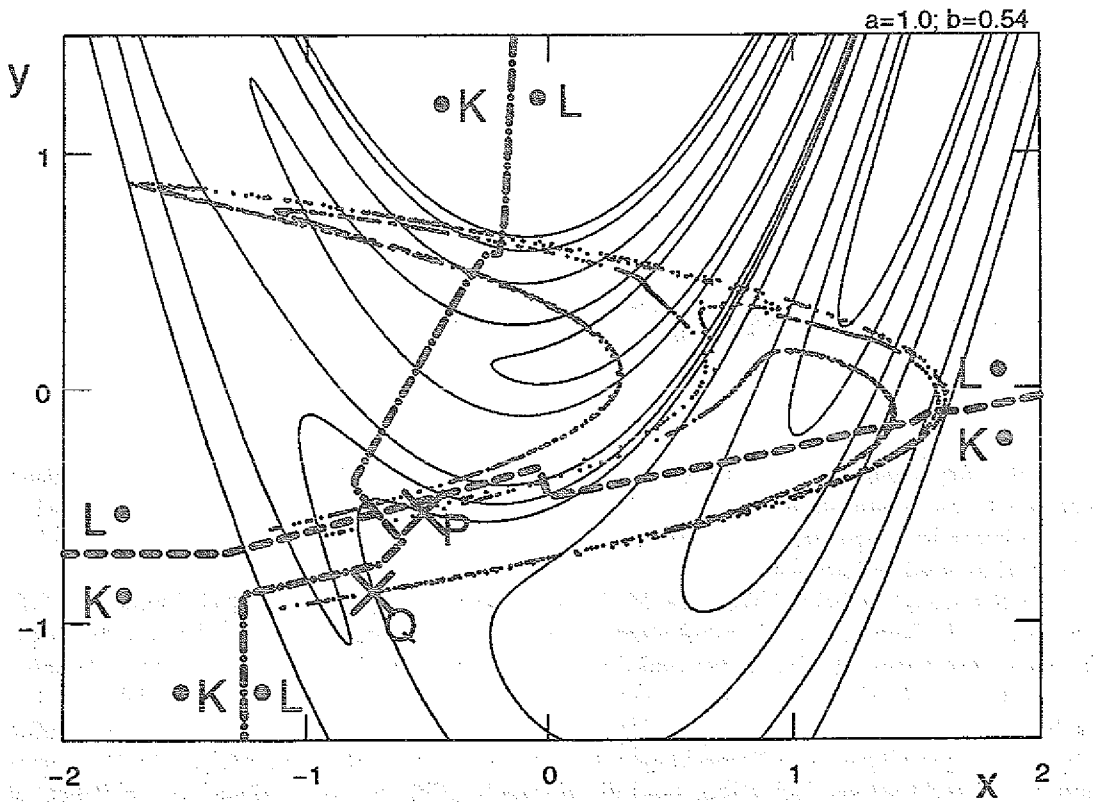


Abbildung 3.1: Beispiel einer nicht generierend Partition des Phasenraumes in die beiden Gebiete $\circ K$ und $\circ L$. Die genaue Lage der Punkte P und Q soll so gewählt sein, daß sie sowohl entlang der stabilen Mannigfaltigkeit als auch entlang der instabilen Mannigfaltigkeit durch möglichst kurze Linien miteinander verbunden werden können. Die Orbits durch die Punkte P und Q konvergieren dann sowohl in der Zunkunft als auch in der Vergangenheit schnell gegeneinander, so daß sie vermutlich zu keiner Zeit in verschiedenen Gebieten $\circ K$ und $\circ L$ liegen.

teres scheint der Fall zu sein: Die Orbitmengen $1010 \circ 111111$ und $1001 \circ 111111$ sind beide in $KLK \circ LLLLLL$ enthalten, wie man durch Einsetzen der Übersetzungsregeln zeigen kann. Wenn sich diese endlichen 01–Sequenzen auf gleiche Weise zu zulässigen unendlichen Sequenzen fortsetzen lassen, dann verwechselt die KL–Partition die beiden entsprechenden Orbits. Ob solche Forsetzungen zulässig sind, kann nicht exakt bewiesen werden, solange nicht alle grammatikalischen Regeln der Symbolischen Dynamik bekannt sind. Doch die Geometrie des Phasenraumes läßt vermuten, daß solche Paare von Symbolsequenzen existieren, zum Beispiel die der Orbits durch die Punkte P und Q in Abb. 3.1. Die KL–Partition ist dann nicht generierend.

Die genaue Rechnung zeigt aber auch umgekehrt, daß wegen der grammatikalischen Regeln $001\circ = 001 \circ 111111$ und $\circ 0E10 = \emptyset$ (siehe Kap. 4.2.1) die oben angegebenen 01–Sequenzen die einzigen sind, die nach der Übersetzung in KL–Symbolsequenzen miteinander verwechselt werden. Da viele Symbolsequenzen dieser Länge zulässig sind, ist es gar nicht leicht, diese Verwechslung zu finden, selbst mit den Rechenregeln für Paare $[,]$ von Orbits aus Kap. B. Dieses Beispiel illustriert, wie schwer es selbst bei einfachen Änderungen der Partition sein kann, nachzuprüfen, ob die Eigenschaft „generierend“ erhalten bleibt.

3.1.2 Der Ansatz zur Lösung des Problems

Um dieses Problem zu lösen werde ich im folgenden die Bedingung 3.2.1 für die Eigenschaft „generierend" so umformulieren, daß sie sich zu einem einzigen Zeitpunkt nachprüfen läßt. Dabei wird die Vergangenheit $\ldots S_{-2}S_{-1}\circ$ durch eine Partition $\{A_1\circ, \ldots A_m\circ\}$ beschrieben ($\forall i < 0 : S_i \in \{A_1, \ldots A_m\}$) und die Zukunft $\circ S_0 S_1 \ldots$ durch eine Partition $\{\circ B_1, \ldots \circ B_n\}$ beschrieben ($\forall i \geq 0 : S_i \in \{B_1, \ldots B_m\}$). Beide Partitionen $\{A_1\circ, \ldots A_m\circ\}$ und $\{\circ B_1, \ldots \circ B_n\}$ zusammen werden als **zwiefache Partition** des Phasenraumes bezeichnet. (Diese soll nicht mit einer binären Partition verwechselt werden, bei der die Zahl der Symbole gleich 2 ist.)

Diese beiden Partitionen sind im allgemeinen verschieden. Von keiner wird gefordert, daß sie generierend ist, so daß sich beide flexibel abändern lassen. Nur zusammen müssen sie die Bedingung 3.2.1 in Abschnitt 3.2.3 erfüllen, die gewährleistet, daß sich nach allen Änderungen aus ihnen beiden zusammen eine generierende Partition rekonstruieren läßt. Nur wegen dieser Bedinung und dem Kriterium der Einfachheit 4.1.2 (aus Kap. 4.1.2) sind diese beiden Partitionen zu einem Zeitpunkt, sozusagen der Gegenwart, voneinander abhängig.

Diese Veralligemeinerung der Symbolischen Dynamik erfordert auch eine Verallgemeinerung der Schreibweise von Symbolsequenzen, so daß sich alle Rechnungen wieder als rein algebraische Symbolmanipulationen ausdrücken lassen. Die geometrischen Beziehungen im Phasenraum sind nur zu Beginn der Rechnung wichtig, beim Bestimmen der grammatikalischen Regeln, und am Ende, beim Interpretieren der Ergebnisse.

3.2 Definition wichtiger neuer Konzepte

3.2.1 Expandierende und kontrahierende Blätter

Was heißt es nun genau, die Vergangenheit durch eine Partition $\{A_1\circ, \ldots A_m\circ\}$ zu beschreiben? Es heißt, den Phasenraum durch Angabe von halbunendlichen Symbolsequenzen in Punktmengen $\ldots S_{-2}S_{-1}\circ$ zu unterteilen, mit $S_i \in \{A_1, \ldots A_m\}$. Jedes dieser Gebiete soll als **expandierendes Blatt** (der Partition $\{A_1\circ, \ldots A_m\circ\}$) bezeichnet werden.

Ich wählte diesen Namen, weil sich diese Punktmengen einfach geometrisch interpretieren lassen, sofern die Partition $\{A_1\circ, \ldots A_m\circ\}$ selbst generierend ist und sogar die verschärfte Bedingung 2.3 in Kap. 2.2.3 erfüllt.

$$0 \quad = \quad \lim_{t\to\infty} \operatorname{diam}(f^{-t}(\ldots S_{-2}S_{-1}\circ)) \tag{3.3}$$

Alle Orbits eines expandierenden Blattes konvergieren daher in der Vergangenheit gegeneinander. In der Regel werden sie exponentiell schnell konvergieren. Wegen Def. 2.3.2 (Kap. 2.3.1) der instabilen Mannigfaltigkeit liegen dann alle Punkte des expandierenden Blattes auf derselben instabilen Mannigfaltigkeit. Abb. 3.2 illustriert solche expandierenden Blätter.

Auch bei nicht invertierbaren Abbildungen kann $\ldots S_{-2}S_{-1}\circ$ als expandierendes Blatt definiert werden, allerdings kann diese Orbitmenge nicht als Punktmenge im Phasenraum interpretiert werden (siehe Kap. 2.2.2). Im Hauptteil dieser Arbeit werde ich nur invertierbare Abbildungen betrachten und expandierende Blätter als Punktmengen interpretieren.

Die Ränder eines expandierenden Blattes liegen dort, wo ein t–faches Abbild $f^t(\delta(A_i\circ))$ der Partitionsgrenzen $\delta(A_i\circ)$ die instabile Mannigfaltigkeit schneidet (mit $t \geq 0$). Unter der Abbildung f können sich diese Blätter vergrößern, werden aber durch die Grenzen $\delta(A_i\circ)$ wieder in kleinere expandierende Blätter unterteilt. Die durch f auf der Menge der expandierenden Blätter induzierte Abbildung ist daher mehrdeutig. Sie kann als ein „Strecken und Teilen" interpretiert werden. Die inverse Dynamik f^{-1} induziert dagegen eine eindeutige Abbildung auf den expandierenden Blättern, da sie jedes expandierende Blatt $\ldots S_{-3}S_{-2}S_{-1}\circ$ in eine Teilmenge des expandierenden Blattes $\ldots S_{-3}S_{-2}\circ$ abbildet.

In dem Beispiel von Abb. 3.2 scheint jedes expandierende Blatt einfach zusammenhängend zu sein. In anderen Fällen, vor allem in höheren Dimensionen, ist das wohl eher die Ausnahme als die Regel. Denn schon ein zweidimensionales expandierendes Blatt braucht nicht selbst

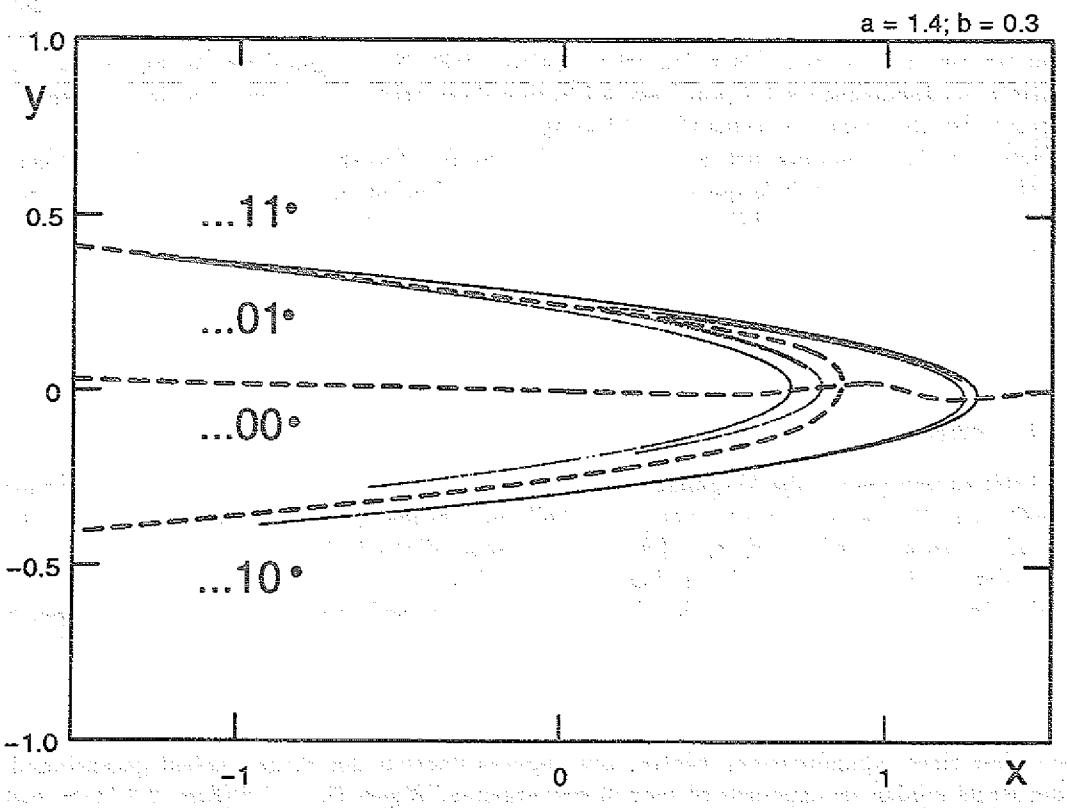


Abbildung 3.2: Expandierende Blätter der 01–Partition des Hénon Attraktors bei Standardparameterwerten (siehe Abb. 2.4). Jedes Blatt ist eine ungefähr waagrechte Linie, die sich aus Einzelpunkten des Attraktors zusammensetzt und die durch eine halbunendliche Symbolsequenz $\ldots S_{-4}S_{-3}S_{-2}S_{-1}\circ$ charakterisiert werden kann. In diesem Beispiel scheint jedes Blatt ein zusammenhängender Abschnitt der instabilen Mannigfaltigkeit zu sein. Die Blätter enden dort, wo die Partitionsgrenze γ, die $0\circ$ von $1\circ$ trennt, oder eines ihrer Abbilder $f^t(\gamma)$ (mit $t > 0$) den Attraktor schneidet. Nur manche dieser Grenzen sind als gestrichelte Linien eingetragen.

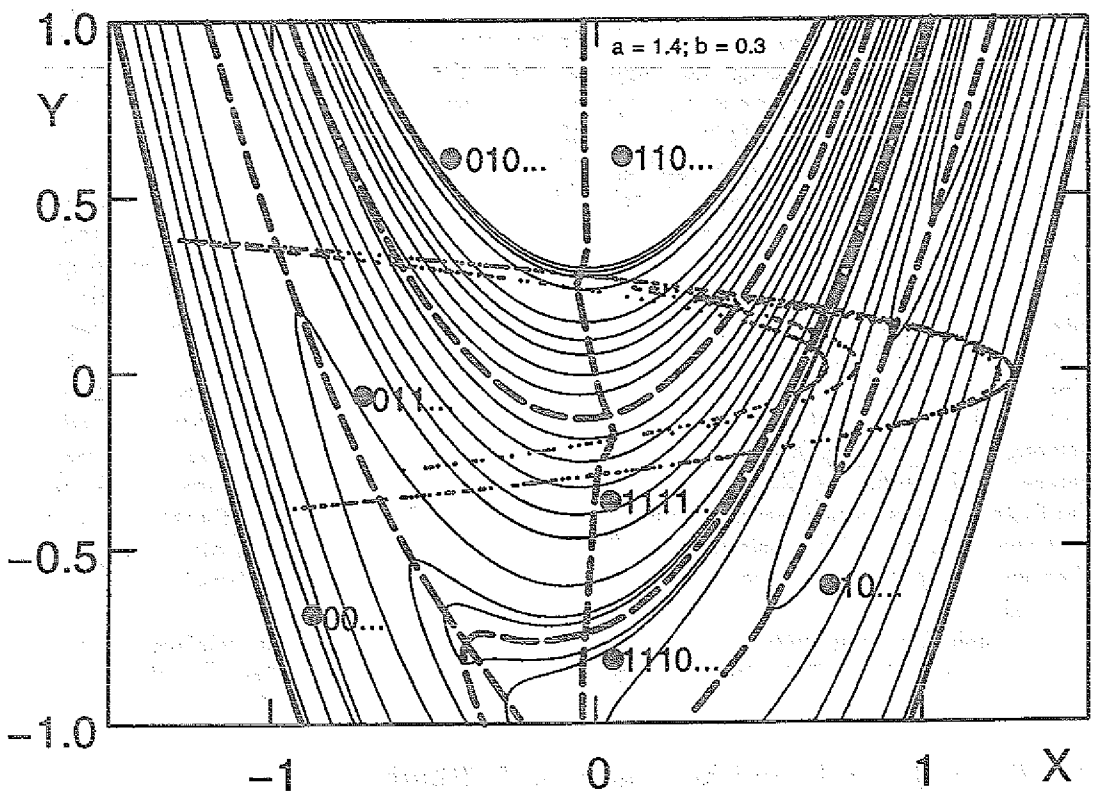


Abbildung 3.3: Geometrische Konstruktion der kontrahierenden Blätter der 01–Partition des Hénon Attraktors bei Standardparameterwerten (siehe Abb. 2.4). Die gestrichelten Linien trennen alle ein Symbolisches Gebiet $\circ E^t 0$ von $\circ E^t 1$ ($t \geq 0$). Abgesehen davon, daß sie nur auf dem Attraktor genau festgelegt sind, sind diese Linien also Urbilder der Partitionsgrenze, die $\circ 0$ und $\circ 1$ trennt. Man erkennt, daß diese Linien die stabile Mannigfaltigkeit nahe starker Krümmungen schneiden. Alle derartige Linien zerlegen den Phasenraum in kleine Punktmengen, die in diesem Beispiel anscheinend zusammenhängende Abschnitte der stabilen Mannigfaltigkeit sind. Die Attraktorpunkte auf einem solchen Abschnitt bilden zusammen ein kontrahierendes Blatt $\circ S_0 S_1 \ldots$.

stark gekrümmt zu sein, sondern nur einen stark gekrümmten Rand zu besitzen, um mit einiger Wahrscheinlichkeit in mehrere Zusammenhangskomponenten zu zerfallen, wenn es durch die Partitionsgrenzen $\delta(A_i\circ)$ unterteilt wird.

Ganz analog werden die Punktmengen $\circ S_0 S_1 \ldots (S_i \in \{B_1, \ldots B_m\})$ einer Partition $\{\circ B_1, \ldots \circ B_n\}$ als **kontrahierende Blätter** bezeichnet. Sie sind in Abb. 3.3 illustriert. Die Abbildung f bildet jedes kontrahierende Blatt $\circ S_0 S_1 S_2 \ldots$ in eine Teilmenge des kontrahierenden Blattes $\circ S_1 S_2 \ldots$ ab. Die inverse Abbildung f^{-1} kann dagegen wieder als „Strecken und Teilen" interpretiert werden. Sie bildet ein kontrahierendes Blatt im allgemeinen auf die Vereinigung mehrerer kontrahierender Blätter ab. Man muß daher nicht nur das kontrahierende Blatt $\circ S_0 S_1 \ldots$, sondern auch das Symbolische Gebiet $B_i\circ$ kennen, in dem ein Punkt z liegt, um folgern zu können, in welchem kontrahierenden Blatt $\circ B_i S_0 S_1 \ldots$ das Urbild $f^{-1}(z)$ liegt. Die Kenntnis des Symbolischen Gebietes $B_i\circ$ ist also Bedingung für die Invertierbarkeit von f, oder genauer gesagt für die Inverteirbarkeit der von f auf der Menge der kontrahierenden Blättern induzierten Abbildung. Diese Bedingung für Invertierbarkeit ist eine Verallgemeinerung der Bedingung, die bereits in Kap. 2.2.3 für eindimensionale, stückweise invertierbare Abbildungen diskutiert wurde.

Jeder Abschnitt der stabilen Mannigfaltigkeit kontrahiert in der Zukunft unter der Dynamik f. Und jeder Abschnitt der instabilen Mannigfaltigkeit expandierte in der Vergangenheit unter der Dynamik f. Aber nicht jeder dieser Abschnitte ist ein expandierendes oder ein kontrahierendes Blatt. Welche Punktmengen expandierende und kontrahierende Blätter sind, hängt ganz entscheident von der zwiefachen Partition ab. Zum Beispiel gehören Abbilder oder Urbilder von expandierenden oder kontrahierenden Blättern im allgemeinen nicht selbst zu den expandierenden oder kontrahierenden Blättern. Sie lassen sich aber trotzdem durch Symbolsequenzen beschreiben. So ist $\ldots S_{-2} S_{-1} E\circ$ das Abbild und $\ldots S_{-2} \circ S_{-1}$ das Urbild des expandierenden Blattes $\ldots S_{-2} S_{-1}\circ$. Um sie von beliebigen anderen Punktmengen zu unterscheiden, werde ich solche Punktmengen, die sich in naheliegender Weise durch halbunendliche Symbolsequenzen beschreiben lassen, als **Blätter** bezeichnen. Dies ist keine präzise Definition. Mathematisch exakt sind nur die Begriffe „expandierendes Blatt" und „kontrahierendes Blatt" definiert. Punktmengen, die keine Blätter sind, werde ich oft als **Gebiete** bezeichnen, auch wenn sie nur aus Attraktorpunkten bestehen und eine fraktale Struktur besitzen. Dazu gehören insbesondere Symbolische Gebiete wie $A_i\circ$ und $\circ B_j$.

Eine normale Partition $\{\circ G_1, \ldots \circ G_q\}$ des Phasenraumes kann auch als zwiefache Partition $\{A_1\circ, \ldots A_m\circ\}$, $\{\circ B_1, \ldots \circ B_n\}$ aufgefaßt werden, indem man $\circ A_i = \circ B_i = \circ G_i$ für alle $i \leq m = n = q$ setzt. Dadurch sind auch für sie expandierende und kontrahierende Blätter definiert.

3.2.2 Expandierende und kontrahierende Bündel

Die Partitionen $\{A_1\circ, \ldots A_m\circ\}$ und $\{\circ B_1, \ldots \circ B_n\}$, die Vergangenheit und Zukunft beschreiben, werden im folgenden nur selten erwähnt. Wichtiger als die zwiefache Partition selbst sind die durch sie definierten kontrahierenden und expandierenden Blätter. Mengen von expandierenden bzw. kontrahierenden Blättern werden so oft auftreten, daß sie als **expandierende Bündel** bzw. **kontrahierende Bündel** bezeichnet werden sollen. Eine Menge von beliebigen Blättern werde ich als **Bündel von Blättern** bezeichnen, wobei die Blätter weder kontrahierend noch expandierend zu sein brauchen.

Um solche Bündel durch Symbolsequenzen zu beschreiben, verwende ich das Rechenzeichen $\sqcup$ (sprich „oder"). Anhang B enthält die genaue Definition und leitet daraus Rechenregeln ab. Da aber auch alle längeren Rechnungen in den Anhang verlegt wurden, genügt eine grobe Erklärung dieses Rechenzeichens, damit der Hauptteil dieser Arbeit verständlich wird.

Betrachten wir einen Ausdruck $\alpha \sqcup \beta$, wobei α und β Blätter oder andere Punktmengen sind. Wenn er für sich allein steht, dann bezeichnet ein solcher Ausdruck ein Element einer abstrakten Algebra, die der Booleschen Mengenalgebra ähnlich ist (siehe Anhang B). Im folgenden wird ein solcher Ausdruck aber nie allein, sondern immer umgeben von Klammern $\{\}'$ auftreten. So ein Gebilde ist viel anschaulicher. Es bezeichnet eine Menge von zwei Punktmengen:

$$\{\alpha \sqcup \beta\}' \quad := \quad \{\alpha\,,\,\beta\} \quad . \tag{3.4}$$

Der unauffällige Strich ' an der geschweiften Klammer }' ist nur deshalb nötig, um diese Menge von Punktmengen von einer Menge zu unterscheiden, die als einziges Element den Ausdruck $\alpha \sqcup \beta$ enthält. Mengen von letzteren Typ treten in dieser Arbeit nicht auf, so daß der Strich ' überlesen werden kann.

Man beachte den Unterschied zwischen dem Operator $\sqcup$ und dem Vereinigungsoperator $\cup$ von Mengen. Im Gegensatz zu $\{\alpha \sqcup \beta\}'$ ist $\{\alpha \cup \beta\}$ eine Menge, die nur eine einzige Punktmenge enthält, nämlich $\alpha \cup \beta$. Eine Identität besteht nur dann, wenn der Vereinigungsoperator $\cup$ außerhalb der geschweiften Klammern $\{\}'$ steht:

$$\{\alpha \sqcup \beta\}' \quad = \quad \{\alpha\} \quad \cup \quad \{\beta\} \quad . \tag{3.5}$$

Wegen dieser Identität wird es in Anhang B möglich sein, Assoziativ– und Distributivgesetze des Vereinigungsoperators $\cup$ auf den neuen Operator $\sqcup$ zu übertragen.

Mit dem neuen Operator $\sqcup$ können auch mehr als nur zwei Punktmengen miteinander verknüpft werden. Zum Beispiel bezeichnet

$$\{(00\circ) \sqcup (01\circ) \sqcup (10\circ) \sqcup (11\circ)\}' \quad = \quad \{00\circ, 01\circ, 10\circ, 11\circ\} \tag{3.6}$$

eine Menge, die vier Punktmengen enthält. Die volle Definition in Anhang B erlaubt es auch, den Operator $\sqcup$ auf die folgende Weise auszuklammern:

$$\{(00\circ) \sqcup (10\circ)\}' \quad = \quad \{(0 \sqcup 1)0\circ\}' \quad . \tag{3.7}$$

In diesem Ausklammern liegt der praktische Nutzen des Operators $\sqcup$. Während der Operator $\sqcup$ auf der linken Seite noch wie in Gl. 3.6 durch ein einfaches Komma ersetzt werden könnte, ist dies auf der rechten Seite nicht mehr möglich.

Die glatten Klammern () um Punktmengen wie $00\circ$ und $10\circ$ werden im folgenden weggelassen. Dies ist möglich, wenn man dem Rechenzeichen $\sqcup$ eine geringere Priorität gibt als der Aneinanderreihung von Symbolen. Die glatten Klammern auf der rechten Seite von Gl. 3.7 sind dagegen unbedingt notwendig. Denn bei dem unsinnigen Ausdruck $0 \sqcup 11\circ$ wüßte man nicht, welche Position das Symbol 0 relativ zu dem Punkt $\circ$ einnimmt.

Der wirkliche Nutzen des Ausklammerns offenbart sich erst dann, wenn man, wie schon in Kap. 2.2.2, die n–fache Wiederholung einer Sequenz α mit $(\alpha)^n$ bezeichnet. Die Menge in Gl. 3.6, die aus vier Punktmengen besteht, kann dann auch geschrieben werden als:

$$\begin{aligned} \{(00\circ) \sqcup (01\circ) \sqcup (10\circ) \sqcup (11\circ)\}' &= \{0(0 \sqcup 1)\circ \sqcup 1(0 \sqcup 1)\circ\}' && (3.8)\\ &= \{(0 \sqcup 1)(0 \sqcup 1)\circ\}' && (3.9)\\ &= \{(0 \sqcup 1)^2\circ\}' \quad . && (3.10) \end{aligned}$$

Die beiden halbunendlichen Wiederholungen der Sequenz α sollen wieder mit $(\alpha)^{-\infty} = \ldots\alpha\alpha$ und $(\alpha)^{+\infty} = \alpha\alpha\ldots$ bezeichnet werden. Das Bündel aller expandierenden Blätter einer zwiefachen Partition $\{A_1\circ, \ldots A_m\circ\}$, $\{\circ B_1, \ldots \circ B_n\}$ wurde als die Menge definiert, die alle Punktmengen $\ldots S_{-2}S_{-1}\circ$ mit $S_i \in \{A_1, \ldots A_m\}$ enthält. Dieses Bündel läßt sich nun schreiben als:

$$\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty}\circ\}' \quad . \tag{3.11}$$

Ebenso kurz läßt sich das Bündel aller kontrahierenden Blätter schreiben:

$$\{\circ(\bigsqcup_{i=1}^{n} B_i)^{+\infty}\}' \quad . \tag{3.12}$$

Es sei abschließend daran erinnert, daß bei nicht invertierbaren Abbildungen f eine Symbolsequenz $\ldots S_{-2}S_{-1}\circ$ nicht als Punktmenge, sondern nur als Orbitmenge interpretiert werden kann. Der Operator $\sqcup$ kann aber auch in diesem Fall verwendet werden, mit dem kleinen Unterschied, daß er auf Orbitmengen statt auf Punktmengen operiert.

3.2.3 Generierende zwiefache Partitionen

In Kap. 2.3.4 wurde die Bedingung „generierend" für eine Partition $\{\circledast G_1, \ldots \circledast G_q\}$ dahin abgeschwächt, daß nur noch verlangt wurde, diese Partition solle nicht mehr Orbits verwechseln als eine Referenzpartition $\{\circledast F_1, \ldots \circledast F_p\}$, die als fein genug gilt. Diese Bedingung kann nun durch Formeln ausgedrückt werden. Wenn die Partition $\{\circledast F_1, \ldots \circledast F_p\}$ generierend ist, dann bezeichnet jede zulässige unendliche Symbolsequenz einen einzigen Orbit. Dasselbe muß dann von der Partition $\{\circledast G_1, \ldots \circledast G_q\}$ gefordert werden.

$$\{(\bigsqcup_{i=1}^{q} G_i)^{-\infty} \circledast (\bigsqcup_{j=1}^{q} G_j)^{+\infty}\}' \quad = \quad \{(\bigsqcup_{i=1}^{p} F_i)^{-\infty} \circledast (\bigsqcup_{j=1}^{p} F_j)^{+\infty}\}' \tag{3.13}$$

ist dann eine Menge, die keine Punktmenge mit mehr als einem Punkt enthält. In anderen Worten: Jedes expandierende Blatt der Partition $\{\circledast G_1, \ldots \circledast G_q\}$ schneidet jedes kontrahierende Blatt in höchstens einem Punkt.

Diese Bedingung läßt sich auf beliebige zwiefache Partitionen verallgemeinern.

Bedingung 3.2.1 (Generierende zwiefache Partition)
Die zwiefache Partition $\{A_1\circledast, \ldots A_m\circledast\}$, $\{\circledast B_1, \ldots \circledast B_n\}$ heißt generierend, wenn sie nicht mehr Orbits verwechselt als die Referenzpartition $\{\circledast F_1, \ldots \circledast F_p\}'$, das heißt, wenn:

$$\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} \circledast (\bigsqcup_{j=1}^{n} B_j)^{+\infty}\}' \quad = \quad \{(\bigsqcup_{i=1}^{p} F_i)^{-\infty} \circledast (\bigsqcup_{j=1}^{p} F_j)^{+\infty}\}' \tag{3.14}$$

Es wird sich für die Rechnungen als nützlich erweisen, den linken Ausdruck in Gl. 3.14 als Schnitt eines expandierenden und eines kontrahierenden Bündels zu schreiben. Ohne weiteres ist dies aber nicht möglich. Denn der Schnitt des expandierenden Bündels in Gl. 3.11 mit dem kontrahierenden Bündel in Gl. 3.12 ist im allgemeinen leer:

$$\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty}\circledast\}' \quad \cap \quad \{\circledast(\bigsqcup_{j=1}^{m} B_j)^{+\infty}\}' \quad = \quad \{\emptyset\} \quad . \tag{3.15}$$

Denn im allgemeinen wird kein expandierendes Blatt, als Punktmenge betrachtet, mit einem kontrahierenden Blatt übereinstimmen.

Um einen nichtleeren Schnitt bilden zu können, müssen etwas andere Bündel definiert werden. Diese Änderung wird dadurch ausgedrückt, daß der schwarzen Punkt $\circledast$ durch den **offenen Punkt** $\circ$ ersetzt wird. Die formale Definition lautet:

$$\{\nu \circ \mu\}' \quad := \quad \{\alpha \subseteq \circledast E \mid \exists \beta \in \{\nu \circledast \mu\}' : \alpha \subseteq \beta\} \quad . \tag{3.16}$$

(In dieser Definition dürfen die Sequenzen ν und μ außer Symbolen auch Klammern und die Operatoren $\cup, \cap$ und $\sqcup$ enthalten.) Die Definition besagt, daß die Menge $\{\nu \circ \mu\}'$ alle Punktmengen enthält, die auch in $\{\nu \circledast \mu\}'$ enthalten sind, aber außerdem noch alle Teilmengen dieser Punktmengen. Zum Beispiel enthält $\{\circ(0 \sqcup 1)\}'$ all die Punktmengen, die entweder ganz in $\circledast 0$ oder ganz in $\circledast 1$ enthalten sind. Und $\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty}\circ\}'$ enthält all die Punktmengen, die ganz in einem expandierenden Blatt enthalten sind. Man beachte, daß die Blattstruktur in den mit den offenen Punkt $\circ$ geschriebenen Bündeln genauso wichtig ist wie in den mit dem schwarzen Punkt $\circledast$ geschriebenen Bündeln. Aus diesem Grund bleiben viele Beziehungen gültig, wenn man überall in ihnen den schwarzen Punkt $\circledast$ durch den offenen Punkt $\circ$ ersetzt, oder umgekehrt den offenen Punkt $\circ$ durch den schwarzen Punkt $\circledast$.

Weitere Folgerungen aus der Definition 3.16 des offenen Punktes $\circ$ finden sich in in Anhang B. Insbesondere kann man jetzt wegen Gl. B.29 einen nichtleeren Schnitt bilden:

$$\{\nu \circ \mu\}' \quad = \quad \{\nu\circ\}' \quad \cap \quad \{\circ\mu\}' \quad . \tag{3.17}$$

Schließlich läßt sich Bedingung 3.2.1, wie gewünscht, als Schnitt zweier Bündel schreiben:

$$\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} \circ\}' \quad \cap \quad \{\circ(\bigsqcup_{j=1}^{n} B_j)^{+\infty}\}' \quad = \quad \{(\bigsqcup_{i=1}^{p} F_i)^{-\infty} \circ (\bigsqcup_{j=1}^{p} F_j)^{+\infty}\}' \quad . \tag{3.18}$$

Diese Beziehung folgt sofort aus 3.14, wenn man Gl. 3.16 und 3.17 verwendet. Umgekehrt sieht man aber auch leicht, daß diese Beziehung 3.18 verletzt wird, falls die zwiefache Partition $\{A_1\circ, \ldots A_m\circ\}$, $\{\circ B_1, \ldots \circ B_n\}$ mehr Orbits verwechselt als die Referenzpartition $\{\circ F_1, \ldots \circ F_p\}'$. Denn dann gibt es ein expandierendes Blatt und ein kontrahierendes Blatt, die sich doppelt schneiden. Wir bilden die Menge der beiden Schnittpunkte z_1, z_2. Diese Punktmenge $\{z_1, z_2\}$ ist in den Mengen auf der linken Seite von Gl. 3.18 enthalten, aber nicht in der Referenzmenge auf der rechten Seite, die nach Voraussetzung nur Punktmengen mit höchstens einem Punkt enthält.

Die Gl. 3.14 und 3.18 sind daher zwei äquivalente Weisen, die Eigenschaft „generierend" auszudrücken. Wenn in späteren Abschnitten zwiefache Partitionen abgeändert werden, dann wird immer darauf geachtet, die Eigenschaft „generierend" zu erhalten.

Der Weise sucht nur eins,
und zwar das höchste Gut.
Ein Narr nach vielerley,
und kleinem streben thut.
aus Angelus Silesius:
„Cherubinischer Wandersmann"

Kapitel 4

Kriterium für die Einfachheit von Partitionen

4.1 Allgemeine Überlegungen

4.1.1 Wozu ein Kriterium der Einfachheit?

Die Vereinfachung generierender Partitionen wird nicht nur dadurch behindert, daß generierende Partitionen sich erst dann flexibel abändern lassen, wenn expandierende und kontrahierende Blätter unterschieden werden (siehe Kap.3). Ein weiteres Hindernis ist die große Zahl der möglichen Änderungen einer Partition. Ein Kriterium der Einfachheit wird benötigt, damit man sich bei den Änderungen nicht in beliebig komplizierte Partitionen verrennt.

Ein sinnvolles Kriterium der Einfachheit sollte mehrere Bedingungen erfüllen. Erstens sollte es invariant unter Koordinatentransformationen sein. Diese Bedingung wurde bereits in Kap. 1.1 begründet. Wenn in Veröffentlichungen koordinatenabhängige Kriterien vorgeschlagen wurden, wie insbesondere die Summe der Krümmungen der stabilen und der instabilen Mannigfaltigkeit [30], so wiesen die Autoren selbst auf diesen Mangel hin.

Eine zweite Bedingung ist, daß die Relation „einfacher" transitiv ist. Wenn eine erste Partition zu einer zweiten vereinfacht wurde, und diese zu einer dritten, dann soll die dritte Partition auch wirklich einfacher sein als die erste. Sonst kann man kaum ein Ende der Vereinfachungsschritte garantieren, ein Ende, das jeder Algorithmus haben muß. Diese Bedingung wird im folgenden dadurch erfüllt, daß das Kriterium 4.1.2 der Einfachheit ein globales Kriterium ist, mit dem sich nicht nur ähnliche, sondern auch ganz unterschiedliche Partitionen vergleichen lassen. Allerdings wird sich bei diesem Vergleich nicht immer eine Partition als die einfachere erweisen, mathematisch gesprochen definiert das Kriterium 4.1.2 nur eine Halbordnung.

Die dritte Bedingung hat praktische Bedeutung: Das Kriterium soll sich möglichst leicht ausrechnen lassen. Die Rechnungen in Kap. 5 werden erst dadurch handhabbar, daß die Kenntnis der Einfachheit einer Partition benutzt werden kann, um die Einfachheit anderer, aber ähnlicher Partitionen leichter zu bestimmen. Außerdem läßt sich die Einfachheit allein durch Rechnen mit Symbolsequenzen bestimmen. Das Kriterium, ob expandierende oder kontrahierende Blätter zusammenhängend sind, oder andere geometrische Kriterien, die sich nur sehr schwer durch Symbolsequenzen ausdrücken lassen, ließen sich in praktischen Rechnungen nicht bewältigen.

K. Hansen versuchte Partitionen dadurch auszuzeichnen, daß sie sich bei Parameteränderung (z. B. $b \to 0$ bei der Hénon Abbildung) auf bestimmte Weise verändern ([35] und private Mitteilung). Doch in praktischen Rechnungen scheint es unmöglich, alle möglichen Änderungen einer Partition entlang aller möglichen Verbindungswege im Parameterraum zu bestimmen. Dazu ist die Struktur des Parameterraumes [53] viel zu komplex. Auch all die vielen Komplexitätsmaße, die auf der Kenntnis aller grammatikalischen Regeln beruhen, lassen sich wohl praktisch kaum anwenden, da im allgemeinen nur die Zulässigkeit von kurzen Symbolsequenzen bekannt ist.

Eine vierte naheliegende Bedingung wäre, daß es für jede Dynamik nur eine eindeutig einfachste Partition geben soll. Diese Bedingung wird hier aus gutem Grund (siehe Kap. 2.3.3) weder gefordert noch erfüllt (Kap. 7.1.2). Seltsame Attraktoren können dann nicht dadurch klassifiziert werden, daß man die einfachste Partition angibt, sondern vielmehr dadurch, daß man all die „einfachsten" Parititionen angibt, die man fand und nicht weiter vereinfachen konnte. In abgeschwächter Form ist diese vierte Bedingung aber sinnvoll: Ein Kriterium der Einfachheit sollte streng genug sein, daß nach dem Vergleich nur wenige „einfachste" Partitionen übrigbleiben. Ein naheliegendes Kriterium wird in Kap. 6 zusätzlich zu dem Kriterium 4.1.2 verwendet: Die Zahl der Symbole soll möglichst klein sein. Allein für sich genommen ist dieses Kriterium nicht besonders streng. Denn selbst beim Hénon Attraktor bei Standardparameterwerten fand ich neben der wohlbekannten binären 01–Partition (Abb. 2.4) noch zwei weitere binäre Partitionen (Abb. 4.2 und 7.1), die auch nicht mehr Orbits verwechseln. Die Zahl der notwendigen Symbole ist außerdem den Bündeln von Blättern nicht leicht anzusehen (Kap. 6.3), so daß dieses Kriterium während der Vereinfachung der zwiefachen Partition in Kap. 5.4 nicht beachtet werden wird.

Damit endet die Liste der Bedingungen. Das Kriterium der Einfachheit ist durch sie natürlich nicht eindeutig festgelegt. Deshalb sind die beiden dynamischen Systeme wichtig, die in der Literatur so lange diskutiert wurden, bis sich ein ungefährer Konsens darüber ergab, welches die einfachste generierende Partition ist. Diese beiden Standardbeispiele sind der Hénon Attraktor bei Standardparameterwerten (Abb. 2.3) bzw. bei schwacher Dissipation (Abb. 2.5). An ihnen wird auch das neue Kriterium 4.1.2 getestet werden. Letztlich genügt es, ein Kriterium zu haben, das sich bewährt, auch wenn es nicht das einzig mögliche ist. Denn was in der Einleitung 1.1 gefordert wurde, ist nur eine der Dynamik angepaßte Strukturierung der Dynamik. Dazu braucht die Anpassung nicht perfektioniert zu werden.

4.1.2 Das neue Kriterium der Einfachheit

Die Wahl des hier verwendeten Kriteriums der Einfachheit erscheint nur dann plausibel, wenn man die Dynamik als eine Dynamik auf Bündeln von Blättern betrachtet: Die Partitionen unterscheiden sich darin, wie groß ihre expandierenden und kontrahierenden Blätter sind. Wenn diese zu groß sind, so daß sich ein expandierendes und ein kontrahierendes Blatt zweimal schneiden, dann ist Bedingung 3.2.1 (aus Kap. 3.2.3) verletzt und Orbits werden verwechselt. Man kann das vermeiden, indem man alle expandierenden und kontrahierenden Blätter ganz klein wählt. Dies entspricht einer Partition des Phasenraumes in viele kleine Symbolische Gebiete. Wie in Abschnitt 2.1 diskutiert, erhält man so zwar eine mögliche Strukturierung der Dynamik, aber sie ist nicht an die Dynamik angepaßt.

Eine einfache Partition muß zwischen diesen beiden Extremen liegen. Ihre expandierenden und kontrahierenden Blätter sind groß, aber nicht so groß, daß sie sich zweimal schneiden. Was sich jedoch zweimal schneiden darf, das ist ein kontrahierendes Blatt mit dem Bild eines expandierenden Blattes. Die Orbits durch zwei solche Schnittpunkte werden im folgenden so wichtig sein, daß sie einen besonderen Namen erhalten:

Definition 4.1.1 (Orbitpaar) *Gegeben sei eine zwiefache Partition* $\{A_1\circ, \ldots A_m\circ\}$, $\{\circ B_1, \ldots \circ B_n\}$. *Wenn zwei verschiedene Orbits* $(x_t)_{t\in\mathbb{Z}}$ *und* $(y_t)_{t\in\mathbb{Z}}$ *zu jedem Zeitpunkt* t *entweder auf demselben expandierenden oder auf demselben kontrahierenden Blatt liegen, dann bilden sie zusammen ein „Orbitpaar" dieser zwiefachen Partition.*

Wenn zwei Orbits zum Zeitpunkt $t_0 + 1$ auf demselben kontrahierenden Blatt liegen, dann auch zu allen Zeiten $t > t_0$. Wenn sie zum Zeitpunkt t_0 auf demselben expandierenden Blatt liegen, dann auch zu allen Zeiten $t \leq t_0$. Zwei Orbits bilden also genau dann ein Orbitpaar, wenn es eine Zeit t_0 gibt, so daß sie zum Zeitpunkt t_0 auf demselben expandierenden Blatt $\ldots S_{-2}S_{-1}\circ$ liegen und zum Zeitpunkt $t_0 + 1$ auf demselben kontrahierenden Blatt $\circ S_1 S_2 \ldots$ liegen. Zum Zeitpunkt t_0 liegen sie daher in

$$\ldots S_{-2}S_{-1}\circ \;\cap\; f^{-1}(\circ S_1 S_2 \ldots) \;=\; \ldots S_{-2}S_{-1}\circ E S_1 S_2 \ldots \qquad (4.1)$$

(mit $S_{-1}, S_{-2}, \ldots \in \{A_1, \ldots A_m\}$ und $S_0, S_1, \ldots \in \{B_1, \ldots B_n\}$). Es sei daran erinnert, daß bei einer zwiefachen Partition $\{A_1\circ, \ldots A_m\circ\}$, $\{\circ B_1, \ldots \circ B_n\}$ alle vergangenen Symbole S_{-t} (mit $t > 0$) eines Orbits aus der A_i–Partition und alle zukünftigen Symbole S_{+t} (mit $t \geq 0$) aus der B_i–Partition stammen. Falls die beiden obigen Orbitpaare zum Zeitpunkt $t_0 = 0$ zum letztenmal auf einem gemeinsamen expandierenden Blatt liegen, dann können sich ihre beiden Symbolsequenzen nur in dem einen Symbol S_0 an der Stelle unterscheiden, wo in Gl. 4.1 das Symbol E steht. Kurz gesagt: Genau dann bilden ein expandierendes Blatt $\alpha = \ldots S_{-2}S_{-1}\circ$ und ein kontrahierendes Blatt $\gamma = \circ S_1 S_2 \ldots$ ein Orbitpaar, wenn α das Urbild $f^{-1}(\gamma) = \circ E S_1 S_2 \ldots$ des kontrahierenden Blattes γ mehr als einmal schneidet.

Ob zwei Blätter α und γ auf diese Weise ein Orbitpaar bilden, ist wichtig für die Änderung der zwiefachen Partition. Denn dabei ändern sich nicht nur die Symbolischen Gebiete, sondern auch die Mengen der expandierenden und kontrahierenden Blätter. Es werden zusätzliche Blätter als expandierend oder kontrahierend deklariert. Und umgekehrt verlieren einige Blätter, die vorher expandierend oder kontrahierend waren, diese Eigenschaft während der Änderung der zwiefachen Partition.

Zum Beispiel ist das Urbild $f^{-1}(\gamma)$ eines kontrahierenden Blattes γ im allgemeinen zu groß, um selbst kontrahierend zu sein. Wenn dieses Urbild $f^{-1}(\gamma)$ als neues kontrahierendes Blatt deklariert wird, dann muß darauf geachtet werden, die Bedingung 3.2.1 (aus Kap. 3.2.3) für eine generierende Partition nicht zu verletzen: Das Urbild $f^{-1}(\gamma)$ darf kein expandierendes Blatt mehr als einmal schneiden. Dies ist genau dann der Fall, wenn das kontrahierende Blatt γ mit keinem expandierenden Blatt Orbitpaare bildet. Um das Urbild $f^{-1}(\gamma)$ als kontrahierend deklarieren zu können, muß man daher zunächst alle Blätter, mit denen γ Orbitpaare bildet, als nicht mehr expandierend deklarieren. Hierin liegt die zentrale Bedeutung der Orbitpaare. Auf dieselbe Weise werden die Orbitpaare in den späteren Rechnungen auch bestimmen, welche Abbilder $f(\alpha)$ von expandierenden Blättern als neue expandierende Blätter deklariert werden dürfen.

Wenn die Größen der expandierenden Blätter nicht aufeinander abgestimmt sind, dann werden die kleinen expandierenden Blätter kaum Orbitpaare bilden können, weil die großen expandierenden Blätter die Existenz großer kontrahierender Blätter verhindern. Umgekehrt wird die Zahl der Orbitpaare dann groß sein, wenn die Größen aller expandierenden Blätter und aller kontrahierenden Blätter aufeinander abgestimmt sind. Die Partition kann also dadurch an die Dynamik angepaßt werden, daß die Menge der Orbitpaare maximiert wird.

Definition 4.1.2 (Kriterium der Einfachheit) *Gegeben seien zwei zwiefache Partitionen. Wenn jedes Orbitpaar der ersten zwiefachen Partition auch ein Orbitpaar der zweiten zwiefachen Partition ist, dann heißt die zweite zwiefache Partition die einfachere.*

Wenn jede Partition Orbitpaare besitzt, die der anderen fehlen, dann ist keine die einfachere. Die Relation „einfacher" ist also das, was Mathematiker eine Halbordnung nennen.

Wie gut erfüllt dieses Kriterium die Bedingungen aus dem vorangegangenen Abschnitt 4.1.1? Erstens ist es unabhängig von der Wahl des Koordinatensystems, da die Größe der expandierenden Blätter an denen der kontrahierenden Blätter gemessen wird und umgekehrt. Zweitens lassen sich beliebige zwiefache Partitionen vergleichen, und ihre Relationen sind transitiv. Wenn allerdings jede Partition Orbitpaare erzeugt, welche die anderen nicht erzeugen, dann ist keine die einfachere.

Drittens ist das Kriterium relativ leicht zu berechnen. Dies gilt insbesondere bei kleinen Änderungen einer bekannten zwiefachen Partition. Um zu wissen, welche Bilder expandierender Blätter dabei selbst expandierende Blätter werden dürfen, muß man sowieso berechnen, mit welchen kontrahierenden Blättern sie Orbitpaare bilden. Dann ist es leicht, die Orbitpaare selbst zu berechnen. Für beide Rechenschritte wird im Anhang B die Schreibweise $[,]$ eingeführt und es werden Rechenregeln abgeleitet. Zum Beispiel ist die Menge aller Orbitpaare einer 01–Partition gleich $\bigcup_{t=-\infty}^{+\infty} f^t\{(0 \sqcup 1)^{-\infty} \circ [0,1](0 \sqcup 1)^{+\infty}\}$'. Diese Schreibweise wird aber nur im Anhang, nicht im Hauptteil dieser Arbeit verwendet.

Das neue Kriterium 4.1.2 der Einfachheit ist verwandt mit dem in Kap. 2.3.2 erwähnten Kriterium, daß nämlich jeder Orbit von Tangentialpunkten nur einmal durch die Partitionsgrenzen geschnitten werden sollte. In Kap. 2.3.3 wurde argumentiert, daß sich dieses Ziel nicht immer erreichen läßt. Falls der Orbit durch den Tangentialpunkt x_c in Abb. 2.7 zu mehreren Zeiten entlang

der gestrichelten Linie geschnitten wird, dann werden benachbarte Orbits wie die durch P_i und Q_i in Abb. 2.7 bei mehreren Zeiten durch die Partitionsgrenzen getrennt und bilden keine Orbitpaare. Umgekehrt sollte daher eine nach Def. 4.1.2 einfache Partition möglichst viele Orbits von Tangentialpunkten nur einmal schneiden. Tatsächlich erhält man in Kap. 6.3.4 und 7.1.2 solche Partitionen als Ergebnis der Vereinfachung.

Das neue Kriterium der Einfachheit läßt sich auf mehrere Weisen verallgemeinern. In Def. 4.1.2 werden nur die Schnitte zwischen einem expandierenden Blatt α und dem ersten Urbild $f^{-1}(\gamma)$ eines kontrahierenden Blattes γ verwendet. Man könnte auch die Zahl der Schnitte zwischen α und dem zweiten Urbild $f^{-2}(\gamma)$ als zusätzliches Maß für die Einfachheit verwenden. Da die zweiten Urbilder größer sind als die ersten Urbilder, müßte man für dieses sekundäre Kriterium mehr Schnitte berechnen, was aufwendig wäre. Eine zweite Weise, das Kriterium der Einfachheit 4.1.2 zu verallgemeinern, bestünde darin, n–fache Schnitte zwichen α und dem ersten Urbild $f^{-1}(\gamma)$ gesondert zu betrachten, falls $n > 2$. In die Def. 4.1.2 gehen solche n–fachen Schnitte nicht anders ein als $\binom{n}{2}$ doppelte Schnitte. Wieviele ternäre, quaternäre etc. Schnitte eine Partition erzeugt, ließe sich mit den Methoden von Anhang B relativ leicht berechnen. Doch wenn zwei Partitionen dieselben Orbitpaare erzeugen, dann wüßte ich nicht zu sagen, welche die einfachere sein soll: diejenige, die mehr trinäre, quaternäre etc. Schnitte erzeugt und daher in gewissem Sinne die größeren expandierenden oder kontrahierenden Blätter hat, oder die andere, die weniger trinäre, quaternäre etc. Schnitte erzeugt und sich deshalb vielleicht mit weniger Symbolen beschreiben läßt. Bei binären Partitionen und auch bei all den Vereinfachungsschritten in Kap. 5 treten keine trinären Schnitte auf, so daß dieses Zusatzkriterium irrelevant für die Beispielrechnungen wäre.

4.2 Direkter Vergleich generierender Partitionen

Das Kriterium der Einfachheit kann auf zwei Weisen angewendet werden. Zum einen kann eine gegebene Partition schrittweise vereinfacht werden. Das ist Inhalt des nächsten Kapitels 5. Zum anderen können mehrere gegebene Partitionen direkt miteinander verglichen werden. Dies geschieht für zwei Standardbeispiele in Anhang C.1 und Anhang C.2. Die Ergebnisse werden hier zusammengefaßt.

4.2.1 Hénon Attraktor bei schwacher Dissipation

Das erste Beispiel betrifft den Hénon Attraktor bei schwacher Dissipation $b = 0.54$; $a = 1.0$.

$$x_{n+1} = 1 - 1.0 \cdot x_n^2 + y_n \tag{4.2}$$

$$y_{n+1} = 0.54 \cdot x_n \tag{4.3}$$

Für ihn wurden in Abschnitt 2.3.3 die beiden Partitionen in Abb. 4.1 vorgeschlagen. Die 01–Partition (Abb. 2.6) stammt aus [30] und die XY–Partition (Abb. 2.12) ist eine naheliegende Vereinfachung der VW–Partition, die in [31] mit dem Algorithmus von [9] ausgerechnet wurde. Beide Partitionen scheinen jeden Orbit von Tangentialpunkten nur einmal zu schneiden. Außer dem Kriterium 4.1.2 ist mir kein koordinatenunabhängiges Kriterium bekannt, das eine der beiden Partitionen als einfacher auszeichnen würde.

Um zu berechnen, welche Partition einfacher ist, muß man die Symbole beider Partitionen in eine gemeinsame Sprache übersetzen können. Außerdem muß man die wichtigsten grammatikalischen Regeln ermitteln, das heißt die kurzen verbotenen Symbolischen Sequenzen. Es gibt mehrere Möglichkeiten, solche Übersetzungsregeln und grammatikalischen Regeln zu finden. Im Falle der 01–Partition und der XY–Partition genügt eine sehr einfache und direkte Vorgehensweise, die hier exemplarisch vorgeführt werden soll.

Übersetzungsregeln

Zunächst trägt man ein paar Abbilder oder Urbilder der Partitionsgrenzen ein. Sie unterteilen den Phasenraum in kleine Gebiete. Um die Übersetzungsregeln zu finden, muß man dann Gebiete

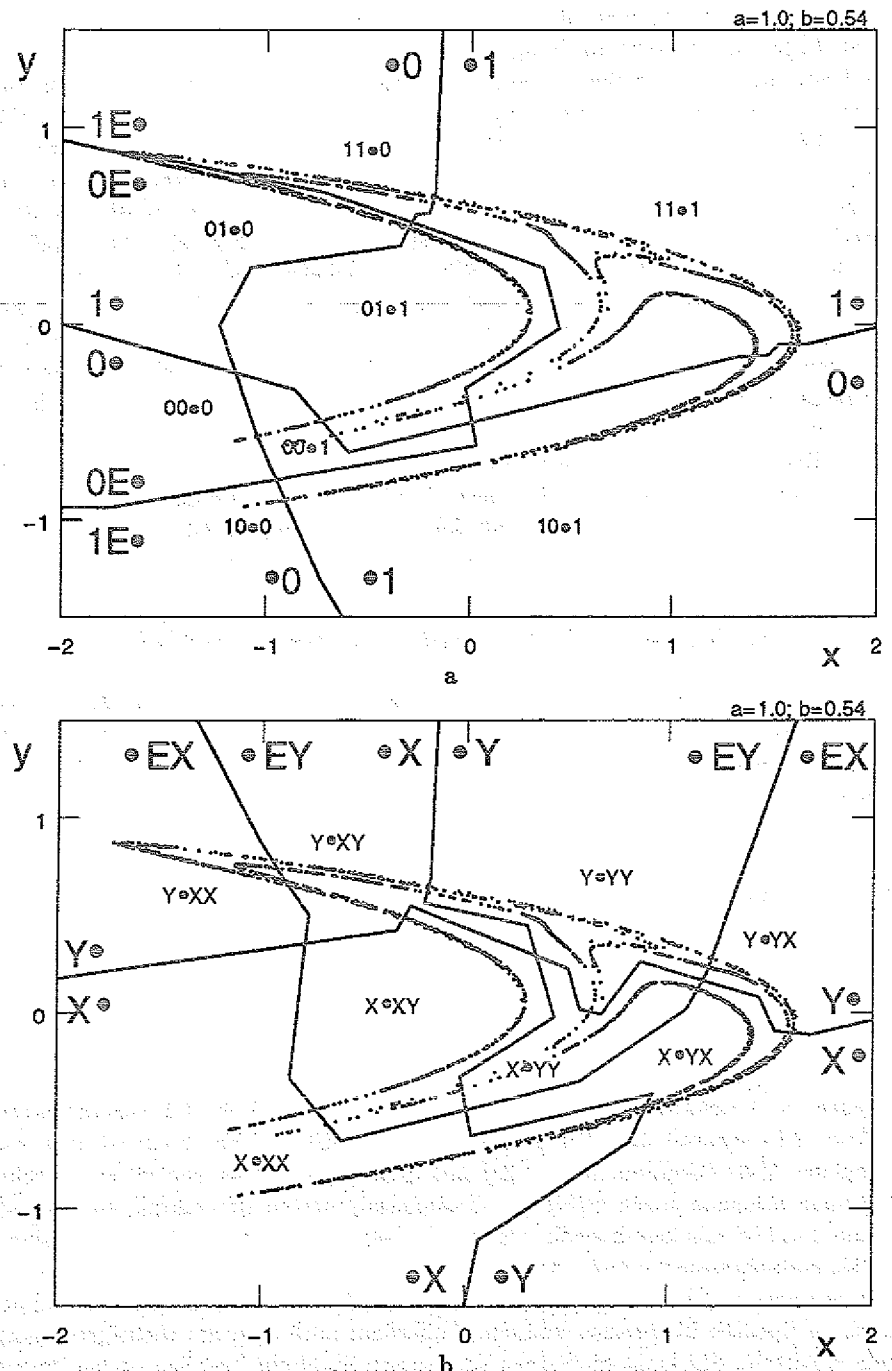


Abbildung 4.1: Vergleich der 01–Partition aus Abb. 2.6 und der XY–Partition aus Abb. 2.12. Die Punkte des Hénon Attraktors bei schwacher Dissipation sind wie in Abb. 2.6 eingetragen. Die 01–Partition erweist sich in Anhang C.1 als einfacher als die XY–Partition, wenn man das Kriterium 4.1.2 verwendet.

identifizieren, die bei beiden Partitionen die gleichen Attraktorpunkte enthalten. In diesem Beispiel sind solche Gebiete besonders leicht zu finden, weil man die Abbildung f nur wenige Male anwenden muß, um Teile der 01–Partitionsgrenze und der XY–Partitionsgrenze aufeinander abzubilden.

Zum Beispiel kann man aus Abb. 4.1 entnehmen, daß die beiden Gebiete $01 \circledast 1$ und $X \circledast XY$ die gleichen Attraktorpunkte enthalten, obwohl die Partitionsgrenzen außerhalb des Attraktors etwas anders liegen.

$$01 \circledast 1 \quad = \quad X \circledast XY \tag{4.4}$$

Auch durch Vereinigung von Gebieten kann man Beziehungen erhalten. Zum Beispiel enthält die Vereinigung von $11 \circledast 0$ und $01 \circledast 0$ dieselben Attraktorpunkte wie die Vereinigung von $Y \circledast XY$ und $Y \circledast XX$.

$$1 \circledast 0 \quad = \quad \underline{11} \circledast 0 \quad \cup \quad 01 \circledast 0 \quad = \quad Y \circledast XY \quad \cup \quad Y \circledast XX \quad = \quad Y \circledast X \tag{4.5}$$

Diese Beziehungen genügen noch nicht, um alle 01–Sequenzen in XY–Sequenzen umzuformen. Man kann nun entweder in Abb. 4.1 noch ein Bild einer Partitionsgrenze einzeichnen und nach weiteren Beziehungen suchen. Weniger unübersichtlich ist es aber, eines der kleinen Gebiete selbst abzubilden. Wenn man eine intuitive Vorstellung von der Hénon Abbildung hat, kann man sogar schon in Abb. 4.1 sehen, daß

$$f(X \circledast XX) \quad = \quad 00 \circledast 0 \quad \cup \quad 00 \circledast 1 \quad \cup \quad 01 \circledast 1 \quad . \tag{4.6}$$

Damit wurden genug Beziehungen aus der Phasenraumgeometrie abgelesen. Die Übersetzungsregeln folgen nun durch reine Symbolmanipulation, wenn man triviale Regeln wie $\circledast X \cup \circledast Y = \circledast E = \circledast 0 \cup \circledast 1$ mitverwendet.

$$\begin{aligned}
\circledast X &= Y \circledast X \quad \cup \quad X \circledast XY \quad \cup \quad X \circledast XX && (4.7)\\
&= 1 \circledast 0 \quad \cup \quad 01 \circledast 1 \quad \cup \quad 0 \circledast 00 \quad \cup \quad 0 \circledast 01 \quad \cup \quad 0 \circledast 11 && (4.8)\\
&= \circledast 0 \quad \cup \quad 01 \circledast 1 \quad \cup \quad 0 \circledast 11 && (4.9)
\end{aligned}$$

Das andere Symbolische Gebiet folgt einfach aus

$$\begin{aligned}
\circledast Y &= \circledast E \setminus \circledast X && (4.10)\\
&= \circledast 1 \setminus (01 \circledast 1 \cup 0 \circledast 11) && (4.11)\\
&= 0 \circledast 10 \quad \cup \quad 11 \circledast 1 \quad . && (4.12)
\end{aligned}$$

Diese Regeln stimmen mit den Übersetzungsregeln Gl. 2.10 und 2.11 überein. Längere 01–Symbolsequenzen kann man dann immer durch mehrfaches Anwenden dieser Regeln übersetzen. Die Rückübersetzung folgt ganz analog aus

$$\begin{aligned}
\circledast 0 &= \circledast X \setminus (01 \circledast 1 \cup 0 \circledast 11) && (4.13)\\
&= \circledast X \setminus (X \circledast XY \cup \circledast XXY) && (4.14)\\
&= Y \circledast XY \quad \cup \quad \circledast XXX \quad \text{und} && (4.15)\\
\circledast 1 &= \circledast E \setminus \circledast X && (4.16)\\
&= \circledast Y \quad \cup \quad X \circledast XY \quad \cup \quad \circledast XXY \quad . && (4.17)
\end{aligned}$$

Aus der Existenz der Übersetzungsregeln folgt sofort

$$\{(0 \sqcup 1)^{-\infty} \circ (0 \sqcup 1)^{+\infty}\}' \quad = \quad \{(X \sqcup Y)^{-\infty} \circ (X \sqcup Y)^{+\infty}\}' \quad . \tag{4.18}$$

Wenn eine der beiden Partitionen keine Orbits verwechselt, dann tut es folglich die andere auch nicht.

Grammatikalische Regeln

Die grammatikalischen Regeln wird man normalerweise nicht so vollständig ermitteln können wie die Übersetzungsregeln. Denn normalerweise kann man immer längere Symbolsequenzen finden, die verboten sind, obwohl sie keine verbotenen Teilsequenzen enthalten. Der direkteste Weg, alle verbotenen Sequenzen der Länge n zu finden, besteht darin, die Partitionsgrenzen von n aufeinanderfolgenden Zeitschritten einzutragen und alle Gebiete zu beschriften. In Abb. 4.1.a ergeben sich 8 Gebiete, so daß keine der 2^3 Symbolsequenzen der Länge 3 verboten ist. Hätte man $n = 4$ Partitionsgrenzen eingetragen, so hätte man aber nur $13 = 2^4 - 3$ Gebiete gefunden, die Attraktorpunkte enthalten. Die drei Symbolsequenzen $\circ 0010 = \circ 0110 = \circ 0000 = \emptyset$ der Länge 4 sind verboten.

Meistens sind es aber bestimmte Symbolsequenzen, von denen man wissen will, ob sie verboten oder zulässig sind. Dann ist es bequemer, ein Gebiet abzubilden und nachzuprüfen, welche anderen Gebiete sein Bild schneidet. So kann man z. B. in Abb. 4.1.a erkennen, daß: $f(0E\circ) \cap (1 \circ 0) = 0E1 \circ 0 = \emptyset$.

In die Berechnung aller Orbitpaare der beiden Partitionen gehen noch zwei weitere grammatikalische Regeln ein:

$$\circ XXXXYX \quad = \quad \emptyset \quad = \quad \circ XXXXXXX \tag{4.19}$$

Man kann sie aus Abb. 4.1 ablesen, wenn man das Gebiet $X \circ XX$ dreimal abbildet:

$$\emptyset \quad = \quad f^3(X \circ XX) \quad \cap \quad X \circ XX \quad \supseteq \quad f^3(X \circ XX) \quad \cap \quad X \circ XXX \tag{4.20}$$

$$\emptyset \quad = \quad f^3(X \circ XX) \quad \cap \quad X \circ YX \quad . \tag{4.21}$$

Manche grammatikalische Regeln sind notwendig, damit die Übersetzungsregeln konsistent sind. Denn wenn man eine 01–Sequenz zuerst in eine XY–Sequenz übersetzt und dann in eine 01–Sequenz rückübersetzt, dann muß man als Ergebnis wieder die ursprüngliche Sequenz erhalten. Wenn man insbesondere $\circ 0$ laut Regel 4.15 durch XY–Sequenzen ausdrückt und diese dann mit Hilfe der Rückübersetzungsregeln 4.9 und 4.12 wieder durch 01–Sequenzen ausdrückt, dann erweist sich die grammatikalische Regel

$$\circ 0E10 \quad = \quad \emptyset \tag{4.22}$$

als notwendig, um wieder die Sequenz $\circ 0$ als Ergebnis zu erhalten. Durch zweifache Übersetzung von $\circ Y$ erhält man analog:

$$\circ YXEY \quad = \quad \emptyset \quad . \tag{4.23}$$

Bei genauerer Betrachtung erweisen sich diese grammatikalischen Regeln sogar als äquivalent und als hinreichend für die Konsistenz der Übersetzungsregeln. Sie sind daher die wichtigsten der grammatikalischen Regeln.

Orbitpaare

Die Berechnung der Orbitpaare wird im Anhang C.1 ausgeführt. Eine längere Rechnung (Gleichung C.31 bis Gleichung C.44 in Anhang C.1) ergibt dann, daß alle Orbitpaare der XY–Partition auch Orbitpaare der 01–Partition sind. Das Symbol, in dem sich die beiden Orbits eines Paares unterscheiden, hat dabei in beiden Partitionen nicht unbedingt dieselbe Lage, sondern kann um bis zu zwei Zeitschritte verschoben sein.

Umgekehrt gibt es aber Orbitpaare der 01–Partition, die nicht Orbitpaare der XY–Partition sind (Gleichung C.45, C.47 und C.48). Diese Orbitpaare haben, abgesehen von der Lage des Punktes $\circ$, alle eine der folgenden drei Formen:

$$\overline{\ldots T_{-4}101} \circ \overline{0111T_5 \ldots} \quad = \quad \overline{\ldots S_{-4}YXY} \circ \overline{XXXYS_5 \ldots} \tag{4.24}$$

$$\underline{\ldots T_{-4}101} \circ \underline{1111T_5 \ldots} \quad = \quad \underline{\ldots S_{-4}YXX} \circ \underline{XYYYS_5 \ldots} \tag{4.25}$$

oder

$$\overline{\ldots T_{-5}0111} \circ \overline{0111T_5 \ldots} \quad = \quad \overline{\ldots S_{-5}XXXY} \circ \overline{XXXYS_5 \ldots} \tag{4.26}$$

$$\underline{\ldots T_{-5}0111} \circ \underline{1111T_5 \ldots} \quad = \quad \underline{\ldots S_{-5}XXXY} \circ \underline{YYYYS_5 \ldots} \tag{4.27}$$

oder

$$\overline{\ldots T_{-5}1111} \circledast 0\overline{111T_5\ldots} \;=\; \overline{\ldots S_{-3}YY} \circledast XXX\overline{YS_5\ldots} \tag{4.28}$$

$$\underline{\ldots T_{-5}1111} \circledast 1\underline{111T_5\ldots} \;=\; \underline{\ldots S_{-3}YY} \circledast YYY\underline{YS_5\ldots} \tag{4.29}$$

(mit $T_i \in \{0,1\}$ und $S_i \in \{X,Y\}$). Die waagrechten Striche wurden hier nur als Hilfslinien eingetragen, um den Bereich der Vergangenheit und Zukunft zu kennzeichnen, in dem das Orbitpaar gleiche Symbole hat. Nach dem Kriterium 4.1.2 ist daher die 01–Partition einfacher als die XY–Partition.

4.2.2 Hénon Attraktor bei Standardparameterwerten

Das zweite Beispiel betrifft den Hénon Attraktor bei Standardparameterwerten $b = 0.3$; $a = 1.4$.

$$x_{n+1} \;=\; 1 - 1.4 \cdot x_n^2 + y_n \tag{4.30}$$

$$y_{n+1} \;=\; 0.3 \cdot x_n \tag{4.31}$$

Die einzige binäre Partition, die für diesen Attraktor veröffentlicht und als generierend bezeichnet wurde [30], ist die 01–Partition in Abb. 2.4. Doch mit einigem Herumspielen konnte ich eine weitere Partition finden, die AB–Partition in Abb. 4.2, die auch nicht mehr Orbits verwechselt als die 01–Partition. Sie schneidet einige Orbits von Tangentialpunkten doppelt. Deshalb würde man erwarten, daß sie weniger einfach ist als die 01–Partition.

Grammatikalische Regeln und Übersetzungsregeln

Die Übersetzungsregeln lauten:

$$\circledast 0 \;=\; B \circledast A \;\cup\; A \circledast AA \tag{4.32}$$

$$\circledast 1 \;=\; \circledast B \;\cup\; A \circledast AB \tag{4.33}$$

$$\circledast A \;=\; \circledast 0 \;\cup\; 0 \circledast 11 \tag{4.34}$$

$$\circledast B \;=\; 1 \circledast 1 \;\cup\; 0 \circledast 10 \;. \tag{4.35}$$

Wie Anhang C.2 zeigt, sind diese Regeln konsistent, da die folgenden Sequenzen verboten sind:

$$\circledast 0010 \;=\; \emptyset \;=\; \circledast 0110 \tag{4.36}$$

$$\circledast BABB \;=\; \emptyset \;=\; \circledast AABA \;. \tag{4.37}$$

Daß diese Punktmengen wirklich leer sind, kann man in der Form

$$\emptyset \;=\; 0E \circledast \;\cap\; \circledast 1 \;\cap\; \circledast E0 \;=\; 0E \circledast 10 \tag{4.38}$$

$$\emptyset \;=\; f(B \circledast \cap \circledast A) \;\cap\; \circledast B \;\cap\; \circledast EB \;=\; BA \circledast BB \tag{4.39}$$

$$\emptyset \;=\; f(A \circledast \cap \circledast A) \;\cap\; \circledast B \;\cap\; \circledast EA \;=\; AA \circledast BA \tag{4.40}$$

aus den Phasenraumabbildungen 2.4.b und 4.2.b entnehmen.

Alle 01–Sequenzen bis zur Länge 31 wurden in [17] darauf untersucht, ob sie verboten oder erlaubt sind. Alle Sequenzen bis zur Länge 6 sind erlaubt, falls sie nicht die Sequenz 0010, 0110 oder 0000 enthalten.

Orbitpaare

Alle Orbitpaare beider Partitionen werden im Anhang C.2 verglichen. Die Rechnung von Gl. C.79 bis zu Gl. C.91 ergibt, daß alle Orbitpaare der AB–Partition auch Orbitpaare der 01–Partition sind. Umgekehrt gibt es aber Orbitpaare der 01–Partition, die nicht gleichzeitig Orbitpaare der

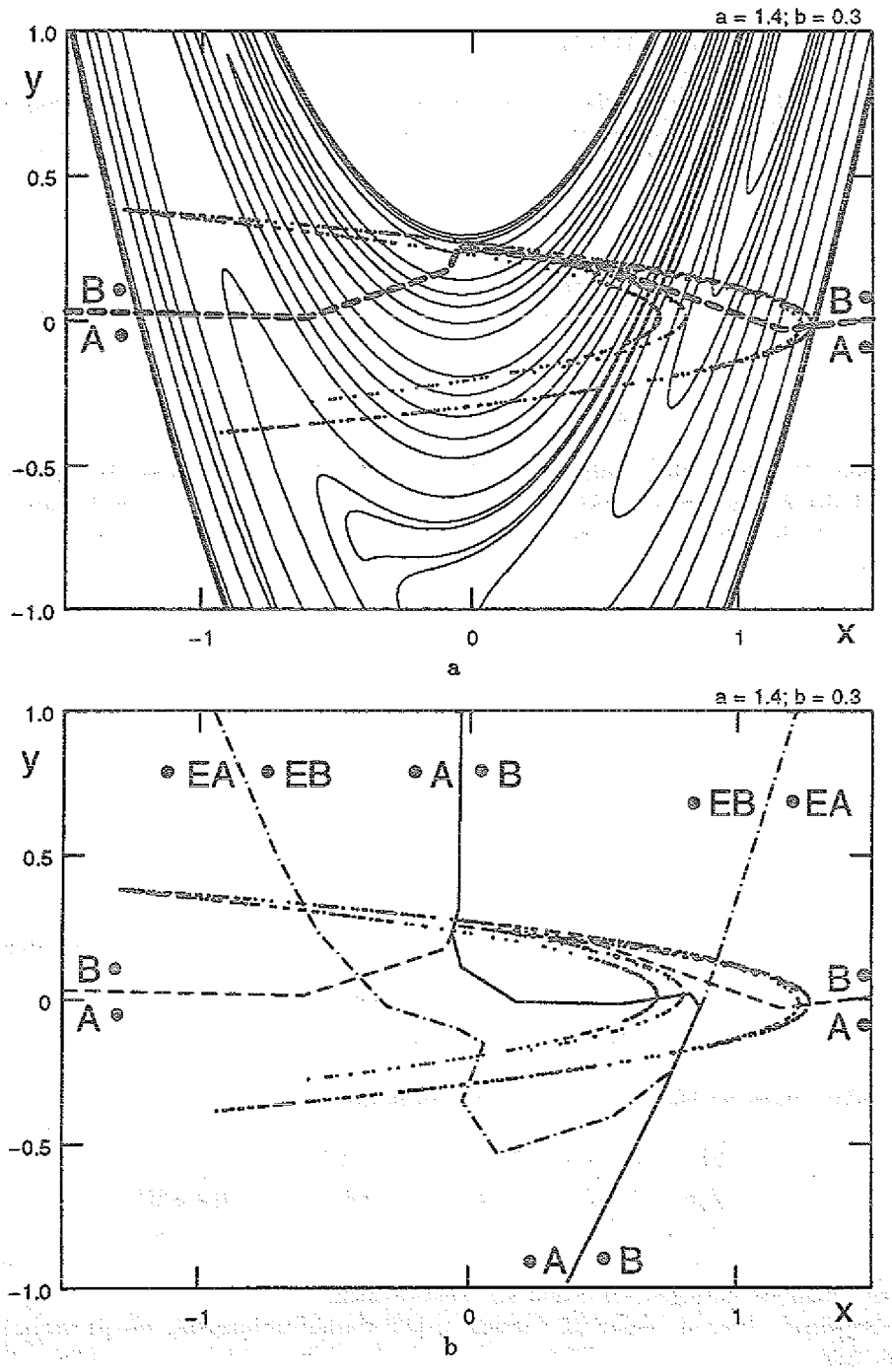


Abbildung 4.2: Die AB–Partition des Hénon Attraktors bei Standardparameterwerten. Sie erweist sich als komplizierter als die 01–Partition in Abb. 2.4, verwechselt aber nicht mehr Orbits als diese. In Abb. a sind die Attraktorpunkte und die Abschnitte der stabilen Mannigfaltigkeit wie in Abb. 2.4 eingetragen. Abb. b zeigt die Partitionsgrenzen bei drei aufeinander folgenden Zeiten. Sie trennen $\circ EA$ von $\circ EB$, $\circ A$ von $\circ B$, bzw. $A\circ$ von $B\circ$.

AB–Partition sind (Gl. C.104 bis C.106). Diese Orbitpaare haben, abgesehen von der Lage des Punktes $\circ$, alle eine der folgenden drei Formen:

$$\overline{\ldots T_{-4}101} \circ \overline{010T_4\ldots} = \overline{\ldots S_{-4}BAB} \circ \overline{AABS_4\ldots} \quad (4.41)$$

$$\underline{\ldots T_{-4}101} \circ \underline{110T_4\ldots} = \underline{\ldots S_{-4}BAA} \circ \underline{BBAS_4\ldots} \quad (4.42)$$

oder

$$\overline{\ldots T_{-4}101} \circ \overline{011T_4\ldots} = \overline{\ldots S_{-4}BAB} \circ \overline{AABS_4\ldots} \quad (4.43)$$

$$\underline{\ldots T_{-4}101} \circ \underline{111T_4\ldots} = \underline{\ldots S_{-4}BAA} \circ \underline{BBBS_4\ldots} \quad (4.44)$$

oder

$$\overline{\ldots T_{-4}111} \circ \overline{011T_4\ldots} = \overline{\ldots S_{-3}BB} \circ \overline{AABS_4\ldots} \quad (4.45)$$

$$\underline{\ldots T_{-4}111} \circ \underline{111T_4\ldots} = \underline{\ldots S_{-3}BB} \circ \underline{BBBS_4\ldots}\,, \quad (4.46)$$

wobei die horizontalen Striche wie in Gleichung 4.29 die Bereiche gleicher Symbole $T_i \in \{0,1\}$ und $S_i \in \{A,B\}$ kennzeichnen.

Das Einfachheitskriterium 4.1.2 stimmt also mit der naiven Erwartung überein: Die 01–Partition ist einfacher als die AB–Partition.

Ein Umweg ist's zum untreuen Freunde,
Wohnt er gleich am Wege;
Aber zum trauten Freunde führt ein Richtsteig,
Weilt er auch weit von hier.
Strophe 34 aus des Hohen Liedes in der älteren Edda

Kapitel 5

Schrittweise Vereinfachung einer zwiefachen Partition

5.1 Überblick

Das Ziel dieses Kapitels ist es, zu demonstrieren, wie eine zwiefache Partition vereinfacht werden kann.

Zunächst muß man die anfängliche Partition wählen, die vereinfacht werden soll. Von der Größe ihrer expandierenden und kontrahierenden Blätter hängt es ab, ob sie sich vereinfachen läßt. Doch auch eine anfängliche Größe der expandierenden und kontrahierenden Bündel muß gewählt werden (Kap. 5.2.2). Von ihr hängt es ab, ob man mögliche Vereinfachungen finden kann. Zu kleine Bündel schaden nicht, doch was klein genug ist, bleibt unbestimmt.

Nach dieser anfänglichen Wahl kann die zwiefache Partition $\{A_1\circ, \ldots A_m\circ\}$, $\{\circ B_1, \ldots \circ B_n\}$ systematisch vereinfacht werden. Dabei sind zwei Eigenschaften der Bündel wichtig:

- In gerichteten Graphen (wie Abb. 5.4) wird dargestellt, wie die Bündel durch die Dynamik f verknüpft sind (siehe Kap. 5.3.2). Um dies auszurechenen, muß man die wichtigsten grammatikalischen Regeln der Symbolischen Dynamik kennen.
- In Tabellen (wie Abb. 5.3) wird dargestellt, welche expandierenden Bündel mit welchen kontrahierenden Bündeln Orbitpaare bilden (siehe Kap. 5.3.1). Um dies auszurechenen, muß man neben den grammatikalischen Regeln auch Übersetzungsregeln kennen, mit deren Hilfe man die beiden Partitionen $\{A_1\circ, \ldots A_m\circ\}$ und $\{\circ B_1, \ldots \circ B_n\}$ in einer gemeinsamen Sprache ausdrücken kann.

Die Orbitpaare lassen sich leichter aus den Bündeln selbst ausrechnen; die zeitlichen Übergänge leichter aus der Partition. Der Abschnitt 5.2 zeigt, wie man die Partition aus den Bündeln berechnet und umgekehrt.

Bei der Vereinfachung der zwiefachen Partition werden die Mengen der expandierenden und kontrahierenden Blätter abgeändert, indem gewisse Blätter aus diesen Mengen entferent werden oder indem zusätzliche Blätter als expandierend oder kontrahierend deklariert werden. Dabei müssen zwei Bedingungen beachtet werden. Laut dem Kriterium 4.1.2 der Einfachheit (aus Kap. 4.1.2) sollen die neuen expandierenden und kontrahierenden Blätter mehr Orbitpaare bilden. Dieses Kriterium befürwortet daher möglichst große expandierende und kontrahierende Blätter. Gleichzeitig muß aber die Bedingung 3.2.1 (aus Kap. 3.2.3) für die Eigenschaft „generierend" erhalten bleiben, das heißt, nach der Änderung sollen nicht mehr Orbits verwechselt werden als davor.

Außer durch diese beiden Bedingungen werde ich die Methode auch noch dadurch beschränken, daß nur ganz bestimmte Punktmengen als neue expandierende oder kontrahierende Blätter deklariert werden dürfen. Im wesentlichen handelt es sich dabei nur um Abbilder von expandierenden oder Urbilder von kontrahierenden Blättern. Im einfachsten Fall (Kap. 5.4.4) ist es ohne weiteres möglich, neue expandierende und kontrahierende Blätter zu deklarieren. In zweiten Fall (Kap.

5.4.6) müssen zuvor andere expandierende oder kontrahierende Blätter entfernt werden. Im dritten und letzten Fall (Kap. 5.4.8) muß zuvor ein bestimmter Teil aus dem expandierenden oder kontrahierenden Blatt herausgeschnitten werden. In jedem Fall werde ich nicht auf einzelnen Blättern, sondern immer nur auf ganzen Bündeln operieren.

Vorteile und Nachteile dieser Vereinfachungsmethode werden in späteren Abschnitten besprochen. Abschnitt 7.2.1 wird erwähnen, wie man garantieren kann, daß die Vereinfachungsschritte zu einem Ende führen. Als Endergebnis erhält man eine zwiefache Partition, die nicht mehr Orbits verwechselt und nicht weniger Orbitpaare bildet als die anfängliche zwiefache Partition. Die Zahl der Symbole kann sich dabei allerdings erhöht haben.

Das Zerlegen eines Bündels in kleinere Teilbündel, kann unter Umständen nötig sein, um letztlich eine normale (das heißt nicht zwiefache) Partition rekonstruieren zu können, die dieselben expandierenden und kontrahierenden Blätter erzeugt wie die zwiefachen Partition. Dies wird in Abschnitt 6.5 diskutiert. Diese Rekonstruktion hängt von den Zusatzbedingungen 6.74, 6.75, 6.79 und 6.80 ab. Da diese Zusatzbedingungen in der folgenden Beispielrechnung immer erfüllt sind, werden sie in diesem Kapitel 5 nicht weiter beachtet.

Unter Zeitinversion wird die Dynamik f zu f^{-1}. Außerdem dreht sich die Reihenfolge jeder Symbolsequenz um. Aus expandierenden Blättern werden kontrahierende Blätter und umgekehrt. Aus der Partition, die die Vergangenheit beschreibt, wird die Partition, die die Zukunft beschreibt, und umgekehrt. Die Methode, die hier vorgestellt wird, erhält diese Symmetrie. Das heißt, man könnte jederzeit die Zeit invertieren, die Methode anwenden, anschließend die Zeit wieder invertieren, und erhielte dasselbe Ergebnis wie ohne jegliche Zeitinversion. Jede Operation auf den expandierenden Blättern kann daher zeitinvertiert auch auf kontrahierende Blätter angewendet werden, und umgekehrt. Es genügt deshalb, wenn ich nur die Operationen, die auf expandierende Blätter wirken, im Detail darstelle.

Als Beispiel werde ich die AB–Partition des Hénon Attraktors bei Standardparameterwerten vereinfachen. Fast alle Rechenschritte werden im Detail vorgeführt, ausgenommen die Berechnung der Orbitpaare. Denn um auch diese einigermaßen effizient vorzuführen, müßte ich die Rechenregeln von Anhang B voraussetzen. Nach jedem Vereinfachungsschritt ergibt sich eine neue zwiefache Partition, die ich immer mit $\{A_1\circ, \ldots A_m\circ\}, \{\circ B_1, \ldots \circ B_n\}$ bezeichnen und durch die unveränderten Symbolischen Gebiete $\circ A$ und $\circ B$ ausdrücken werde. Das Endergebnis wird sich im nächsten Kapitel als die bekannte 01–Partition erweisen.

5.2 Vorbereitende Schritte

Bevor man eine zwiefache Partition systematisch vereinfachen kann, muß man eine zwiefache Partition wählen, von der man ausgeht. Diese Wahl ist wichtig, aber leider ziemlich willkürlich. Dies gilt auch für die Wahl, wie groß die anfänglichen expandierenden und kontrahierenden Bündel sein sollen. Ich begründe zunächst, warum die Größe der Bündel wichtig ist, drücke sie dann mit Hilfe einer Partition aus und diskutiere, wie man sinnvolle Größen finden kann (Kap. 5.2.2). Umgekehrt kann man auch eine Partition durch ihre expandierenden Bündel ausdrücken (Kap. 5.2.3). Dieser eigentlich triviale Rechenschritt wird etwas erweitert, um den Überlapp von Symbolischen Gebieten zu verkleinern (Kap. 5.2.4) und zu garantieren, daß die Symbolischen Gebiete den Phasenraum lückenlos überdecken (Kap. 5.2.5). Als Beispiel wird die AB–Partition des Hénon Attraktors bei Standardparameterwerten eingeführt, deren Vereinfachung uns bis in das nächste Kapitel hinein beschäftigen wird. Die Wahl dieser Partition (Kap. 5.2.6) und die Wahl ihrer Bündelgrößen (Kap. 5.2.7) werden begründet.

5.2.1 Die anfänglichen expandierenden Blätter

Dieser Abschnitt wiederholt einige einfache, aber wichtige Überlegungen aus vorherigen Kapiteln.

Ausgangspunkt der Vereinfachung ist eine zwiefache Partition $\{A_1\circ, \ldots A_m\circ\}, \{\circ B_1, \ldots \circ B_n\}$, die möglichst wenige oder keine Orbits verwechselt. Man kann diese zum Beispiel gewinnen, indem man den Phasenraum in viele kleine Symbolische Gebiete unterteilt. Vielleicht kennt man aber

auch schon eine Partition, die ziemlich einfach erscheint und von der man gerne wüßte, ob sie sich weiter vereinfachen läßt.

Durch diese Partition sind laut Gl. 3.11 die expandierenden Blätter

$$\ldots S_{-2}S_{-1}\circledcirc \in \{(\bigsqcup_{i=1}^{m} A_i)^{-\infty}\circledcirc\}' \tag{5.1}$$

und laut Gl. 3.12 die kontrahierenden Blätter

$$\circledcirc S_0 S_1 \ldots \in \{\circledcirc(\bigsqcup_{i=1}^{n} B_i)^{+\infty}\}' \tag{5.2}$$

festgelegt. Von der Größe dieser Blätter hängt es alleine ab, ob Orbitpaare gebildet oder Orbits verwechselt werden. Es sei noch einmal betont, daß diese Größe nichts mit geometrischen Längen zu tun hat. Ein Blatt ist genau dann größer als ein anderes, wenn es dieses als echte Teilmenge enthält.

Diese unendlich vielen Blätter können nicht alle einzeln betrachtet werden, sondern nur in Bündeln. Jeder Rechenschritt wird auf allen Blättern eines Bündels zugleich operieren. Wenn er auf manchen Blättern operieren soll, und auf anderen nicht, dann müssen diese Blätter in verschiedenen Bündeln enthalten sein. Dies ist nur dann möglich, wenn die Bündel nicht zu groß gewählt wurden. Zu kleine Bündel können nicht schaden. Allenfalls muß man statt auf einem einzelnen Bündel immer auf mehreren Bündeln zugleich operieren. Dies ist bei allen folgenden Rechenmethoden erlaubt.

Im folgenden treten drei verschiedene Größen auf, die nicht miteinander verwechselt werden sollten: die Größe eines Bündels, die Größe eines Blattes und die Größe des Gebietes, das aus der Vereinigung der Blätter eines Bündels entsteht. In allen drei Fällen ist die Größe einer Menge, nicht eine geometrische Größe gemeint. Zum Beispiel ist ein Bündel Γ genau dann so groß wie oder größer als ein Bündel Λ, wenn jedes Blatt von Λ auch ein Blatt von Γ ist: $\Gamma \supseteq \Lambda$. Es genügt dazu nicht, daß die Vereinigung der Blätter in Γ ein mindestens ebenso großes Gebiet abdeckt wie die Vereinigung der Blätter in Λ. Zum Beispiel decken die Blätter des Bündels

$$\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty}(A_1\circledcirc \cup A_2\circledcirc)\}' \tag{5.3}$$

ein ebenso großes Gebiet $A_1\circledcirc \cup A_2\circledcirc$ ab wie die Blätter des Bündels

$$\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty}(A_1\circledcirc \sqcup A_2\circledcirc)\}' \quad . \tag{5.4}$$

Die Blätter aber unterscheiden sich. Die letzteren Blätter sind kleiner, weil sie entweder ganz in $A_1\circledcirc$ oder ganz in $A_2\circledcirc$ liegen müssen. Dagegen darf sich jedes der ersteren Blätter über das gesamte Gebiet $A_1\circledcirc \cup A_2\circledcirc$ erstrecken. Aus demselben Grund soll man nicht den Operator $\cup$ in $\bigcup_{i=1}^{m} A_i\circledcirc = E\circledcirc$ mit dem Operator $\sqcup$ in $\bigsqcup_{i=1}^{m} A_i\circledcirc \neq E\circledcirc$ verwechseln.

5.2.2 Wahl der Bündelgrößen

Dieser Abschnitt gibt Hinweise zu der wichtigen, aber leider relativ willkürlichen Wahl der anfänglichen Bündelgrößen.

Die Größe der expandierenden Blätter hängt von der Partition $\{A_1\circledcirc, \ldots A_m\circledcirc\}$ ab, welche die Vergangenheit beschreibt. Es gibt auch andere Partitionen, die dieselben expandierenden Blätter erzeugen. Diese Freiheit werde ich benutzen, um mit dieser Partition auch gleichzeitig anzugeben, wie groß die expandierenden Bündel Γ_k sind. Sie sollen die Größe

$$\Gamma_k \quad = \quad \{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} A_k\circledcirc\}' \tag{5.5}$$

haben. Ganz analog sind die kontrahierenden Bündel Λ_k durch

$$\Lambda_k \quad = \quad \{\circledast B_k(\bigsqcup_{i=1}^{m} B_i)^{-\infty}\}' \tag{5.6}$$

definiert. Hierdurch wird die Wahl der anfänglichen Bündelgrößen ein wenig eingeschränkt, da sich nicht jedes expandierende Bündel in dieser Form schreiben läßt, das Bündel aller expandierenden Blätter zum Beispiel nicht.

Wenn man derartige expandierende Bündel gegeben hat, dann kann man sie leicht verkleinern, ohne die expandierenden Blätter zu verändern. Man setzt dazu zum Beispiel $\tilde{A}_i\circledast := A_iA_m\circledast$. Dann erzeugt die Partition $\{A_1\circledast, \ldots A_{m-1}\circledast, \tilde{A}_1\circledast, \ldots \tilde{A}_m\circledast\}$ dieselben expandierenden Blätter wie die Partition $\{A_1\circledast, \ldots A_m\circledast\}$, doch kleinere expandierende Bündel.

Wie groß sollen die Bündel sein? Ein eindeutiges, allgemein gültiges Kriterium kenne ich nicht. Je kleinere Bündel man wählt, mit desto mehr Bündeln muß man arbeiten und desto aufwendiger wird die Rechnung. Je größere Bündel man wählt, desto größer ist das Risiko, eine mögliche Vereinfachung zu übersehen. Erst Abschnitt 6.6 wird zeigen, wie sich dieser Aufwand reduzieren läßt. Dort werden Bündel, die bestimmte Kriterien erfüllen, miteinander vereinigt. Diese Kriterien schließen nicht völlig aus, daß nach der Vereinigung der Bündel bestimmte Vereinfachungsschritte unmöglich geworden sind, die vorher möglich waren. Aber sie reduzieren zumindest das Risiko. Als Faustregel würde ich vorschlagen, die Bündel solange zu verkleinern, bis sich die so erhaltenen Bündel nach den Kriterien von 6.6 wieder vereinigen lassen.

Nicht nur am Anfang der Rechnung besteht die Chance, durch Verkleinern der Bündel mehr Vereinfachungen zu ermöglichen, sondern auch bei jedem späteren Stand der Rechnung. Um nicht immer wieder eine willkürliche Wahl treffen zu müssen, werde ich jeweils mit den Bündelgrößen weiterrechnen, die sich aus der vorausgegangenen Rechnung ergaben.

5.2.3 Übersetzung von Bündeln in eine Partition

Durch Gl. 5.5 bzw. Gl. 5.6 sind die expandierenden bzw. kontrahierenden Bündel einer zwiefachen Partition bestimmt. In der späteren Rechnung werden sich diese Bündel ändern. Dabei können recht kompliziert zu schreibende Bündel entstehen, wie zum Beispiel das expandierende Bündel $\Sigma = \{(A \sqcup B)^{-\infty}((A \sqcup B)BB \circledast E \sqcup AAB \circledast E \sqcup (BAA \circledast B \cup BAB \circledast A))\}'$ in Kap. 6.6.2.

Wie findet man die geänderte zwiefache Partition, welche die geänderten Bündel erzeugt? Falls so eine Partition überhaupt existiert, dann erhält man sie, indem man Blätter eines Bündels A miteinander vereinigt.

$$[\mathrm{A}] \quad := \quad \bigcup_{\alpha \in \mathrm{A}} \alpha \tag{5.7}$$

Man beachte, daß A wie jedes Bündel eine Menge von Punktmengen ist, während [A] eine Punktmenge, das heißt ein Gebiet im Phasenraum ist. Dieses Rechenzeichen [A] wird fast ausschließlich in Anhang B verwendet. Wenn man es auf Gl. 5.5 bzw. 5.6 anwendet, dann ergeben sich die gesuchten Symbolischen Gebiete:

$$A_k\circledast \quad = \quad [\Lambda_k] \qquad \text{bzw.} \tag{5.8}$$

$$\circledast B_k \quad = \quad [\Gamma_k] \quad . \tag{5.9}$$

5.2.4 Entfernen zu kleiner Blätter

Die Übersetzung von Bündeln in Symbolische Gebiete bietet auch eine Gelegenheit, bei der man den Überlapp verringern kann, den verschiedene Bündel oder verschiedene Symbolische Gebiete miteinander haben. Dies ist ein relativ trivialer Rechenschritt, den ich wo immer möglich anwende.

Entfernt wird jedes expandierende Blatt α, das ganz in einem anderen expandierenden Blatt $\beta \supset \alpha$ enthalten ist. So ein Blatt α ist nämlich irrelevant für die Bed. 3.2.1 einer generierenden zwiefachen Partition (aus Kap. 3.2.3), für das Kriterium 4.1.2 der Einfachheit (aus Kap. 4.1.2) und alle anderen Beziehungen.

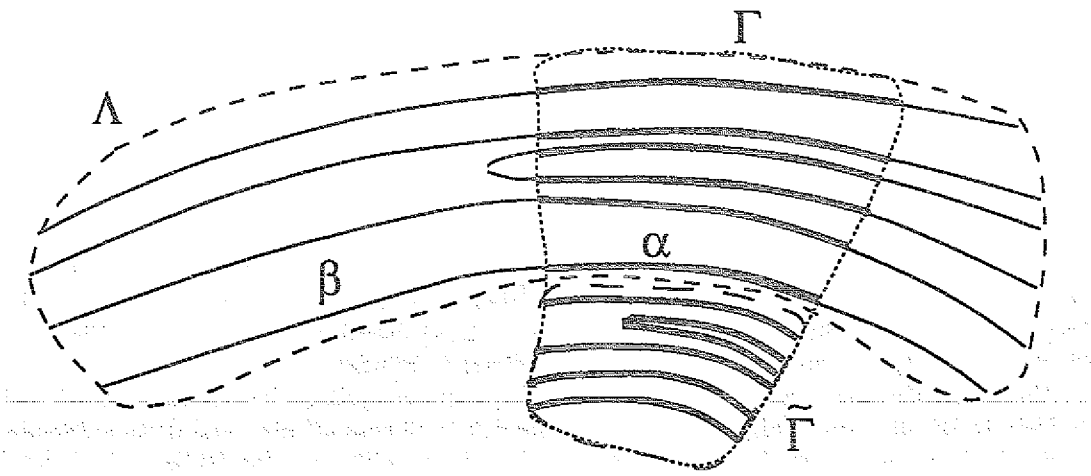


Abbildung 5.1: Entfernen von zu kleinen expandierenden Blättern. Blätter des expandierenden Bündels Γ sind als dicke durchgezogene Linien eingetragen. Manch eines dieser Blätter ist in einem Blatt des expandierenden Bündels Λ enthalten, zum Beispiel α in β. Sie können entfernt werden, wodurch sich das Bündel Γ zu $\tilde{\Gamma}$ verkleinert. Die gestrichelten Linien illustrieren die Größe dieser drei Bündel.

Es kann der Fall vorkommen, daß die beiden expandierenden Blätter α und β gleich sind, obwohl sie aus verschiedenen expandierenden Bündeln stammen. In diesem Fall kann eines der beiden Blätter entfernt werden, aber nicht beide zugleich. Welches Blatt man wählt und entfernt, beeinflußt die Menge der expandierenden Blätter zwar nicht, wohl aber die Größe der Bündel. Und wie bereits gesagt, gibt es kein generelles Kriterium, die Größe der Bündel zu wählen.

In der folgenden Rechnung werde ich diese Wahl nach einer Faustregel treffen, die dafür sorgt, daß ich möglichst wenig Bündel erhalte. Wenn in den beiden expandierenden Bündeln Γ und Λ gleich große expandierende Blätter sind, dann entferne ich sie alle aus demselben Bündel, und zwar möglichst aus dem mit den kleineren Blättern. In anderen Worten: Wenn Γ Blätter enthält, die in größeren Blättern von Λ enthalten sind, aber umgekehrt kein Blatt aus Λ in einem größeren Blatt von Γ enthalten ist, dann entferne ich alle möglichen Blätter, auch die gleich großen, aus Γ.

Durch dieses Entfernen der Blätter verkleinert sich das Bündel Γ zu dem Bündel $\tilde{\Gamma}$, das sich formal schreiben läßt als:

$$\tilde{\Gamma} \quad = \quad \{\alpha \in \Gamma \quad | \quad \forall \beta \in \Lambda : \quad \alpha \not\subseteq \beta\} \tag{5.10}$$

Auf dieselbe Weise lassen sich Blätter auch dann entfernen, wenn die Bündel Γ, $\tilde{\Gamma}$ und Λ kontrahierend sind.

In Abb. 5.1 erkennt man, daß die Blattstruktur der Bündel wirklich wichtig ist. Die von den Bündeln bedeckten Gebiete $[\Gamma]$ und $[\Lambda]$ allein verraten nicht, ob sich das Bündel Γ oder das Bündel Λ verkleinern läßt. In der Beispielrechnung wird es sogar immer möglich sein, all die Blätter von Γ zu entfernen, die in $[\Lambda]$ liegen. Dadurch verschwindet der Überlapp zwischen Symbolischen Gebieten. Im allgemeinen Fall kann allerdings ein Überlapp bestehen bleiben, wenn nämlich verschieden expandierende Blätter teilweise überlappen, ohne daß eines im anderen enthalten ist.

Expandierende Bündel lassen sich auf diese Weise natürlich nur dann verkleinern, wenn zwei Symbolische Gebiete $A_k\circ$ und $A_j\circ$ überlappen: $A_k \circ \cap A_j\circ \neq \emptyset$ und $j \neq k$. Der für die folgende Rechnung wichtigste Fall tritt dann ein, wenn ein Index l existiert, mit $A_l A_j\circ \supseteq A_l A_k\circ \neq \emptyset$. Dann kann das expandierende Bündel

$$\Gamma_k \quad = \quad \{(\bigsqcup_{i=1}^{m} A_i)^{-\infty}(\bigsqcup_{h=1}^{m} A_h A_k\circ)\}' \tag{5.11}$$

deshalb verkleinert werden, weil sein l-tes Teilbündel

$$\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} A_l A_k \circledcirc\}' \tag{5.12}$$

in dem anderen expandierenden Bündel

$$\Gamma_j \quad = \quad \{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} A_j \circledcirc\}' \tag{5.13}$$

enthalten ist. Das verkleinerte Bündel in Gl. 5.10 ergiebt sich als:

$$\tilde{\Gamma}_k \quad = \quad \{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} (\bigsqcup_{\substack{j=1 \\ j \neq l}}^{m} A_j A_k \circledcirc)\}' \quad . \tag{5.14}$$

Ein konkretes Beispiel für das Entfernen zu kleiner expandierender Blätter wird in Kap. 5.4.5 gegeben werden.

5.2.5 Schließen von Lücken im Phasenraum

Bei manchen Änderungen der expandierenden und kontrahierenden Bündel kann es geschehen, daß die neuen Bündel nicht dieselben Gebiete im Phasenraum bedecken wie die alten und daß dadurch Lücken zu entstehen scheinen. Solche Lücken wären fatal, weil dann die Symbolischen Gebiete nicht mehr den ganzen Attraktor überdecken würden. Tatsächlich werde ich die expandierenden und kontrahierenden Bündel nie so massiv abändern, daß Lücken entstehen. Die Lücken, die beim zeitlichen Verschieben von Bündeln zu entstehen scheinen, werden in Wirklichkeit durch Urbilder von expandierenden Bündeln oder durch Abbilder von kontrahierenden Bündeln aufgefüllt.

Es sei daran erinnert, daß das Urbild $f^{-1}(\alpha)$ eines expandierenden Blattes α immer selbst ein expandierendes Blatt oder Teil eines solchen ist. Formal läßt sich das mit den Rechensymbolen aus Kap. 3 schreiben als:

$$f^{-1}\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} \circ\}' \quad = \quad \bigcup_{k=1}^{m} \{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} \circ A_k\}' \quad \subseteq \quad \{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} \circ\}' \quad . \tag{5.15}$$

Wenn die neuen expandierenden Bündel Γ_k eine Lücke offenlassen, dann wird diese geschlossen, indem auch die Urbilder $f^{-t}(\Gamma_k)$ (mit $t \geq 1$) als expandierende Bündel betrachtet werden. Es schadet nichts, wenn zu viele dieser Urbilder $f^{-t}(\Gamma_k)$ als expandierende Bündel deklariert werden. Denn fast alle diese Bündel haben so kleine Blätter, daß sie durch die Methode des Abschnitts 5.2.4 nicht nur verkleinert, sondern vollständig entfernt werden können.

5.2.6 Das Beispiel der AB-Partition

Weil ich nur ein Beispiel im Detail vorrechnen werde, will ich es so wählen, daß es eine der wichtigsten Gefahren illustriert, die bei der Vereinfachung von generierenden Partitionen auftritt. Da die Partitionen Schritt für Schritt vereinfacht werden, liegt der Fall vor, der in der Numerik als Gradientenverfahren bezeichnet wird. Die Hauptgefahr bei Optimierungsmethoden, die Gradientenverfahren sind, besteht darin, in einem lokalen Maximum der „Einfachheit" steckenzubleiben, das heißt bei Partitionen, die sich nur noch durch große Änderungen vereinfachen ließen. Wie groß diese Gefahr ist, hängt von der Art der graduellen Änderungen und dem Kriterium der Einfachheit ab.

Um die neue Methode, die auf Bündel von Blättern operiert, und das neue Kriterium 4.1.2 der Einfachheit zu testen, könnte man mit einer willkürlichen Partition des Attraktors in kleine Symbolische Gebiete beginnen. Diese aufwendige Rechnung wäre aber nicht sehr instruktiv, da die meisten Vereinfachungsschritte auch mit anderen Methoden durchgeführt werden könnten, zum

Beispiel durch das direkte Vereinigen von Symbolischen Gebieten. Es ist deshalb interessanter, mit einer Partition zu beginnen, die sich mit diesen primitiven Methoden nicht weiter vereinfachen läßt, weil sie schon ziemlich einfach ist. Eine solche Partition ist die AB–Partition des Hénon Attraktors bei Standardparameterwerten, die in Kapitel 4.2.2 definiert wurde und die Abb. 4.2 zeigt. Sie kann wie üblich als zwiefache Partition $\{A\circ, B\circ\}$', $\{\circ A, \circ B\}$' aufgefaßt werden. In Gl. 4.37 wurden auch schon die beiden wichtigsten grammatikalischen Regeln angegeben: $\circ AABA = \circ BABB = \emptyset$. Diese AB–Partition unterscheidet sich auf nichttriviale Weise von der 01–Standardpartition in Abb. 2.4.

Beide Partitionen zeigt Abb. 5.2. Laut Kap. 4.2.2 ist diese AB–Partition komplizierter als die 01–Partition, und zwar sowohl nach dem bekannten Kriterium, nach dem die Partitionsgrenze den Attraktor in möglichst wenig Punkten schneiden soll, als auch nach dem neuen Kriterium 4.1.2 der Einfachheit. Mit alten Methoden ist die Vereinfachung zur 01–Partition aber sehr schwer, da in den Übersetzungsregeln 4.32 und 4.33 Symbolsequenzen der Länge 3 auftreten. Man müßte daher die AB-Partition zweimal abbilden, dann sie verfeinern, indem man Symbolische Gebiete zu allen drei Zeiten schneidet, und dann aus den $2^3 = 8$ Schnitten wieder binäre Partitionen bilden. Dies ist auf $2^{(8-1)} - 1 = 127$ Weisen möglich, und bei jeder dieser 127 Partitionen müßte man nachprüfen, ob sie sich entweder in die AB–Partition zurückübersetzen läßt oder mehr Orbits verwechselt als diese. Ersteres würde für die 01–Partition gelten, die sich dann anschließend auch als einfacher erweisen würde als die AB–Partition.

Der Leser wird einsehen, daß eine solche Rechnung praktisch unmöglich ist, es sei denn, man schriebe ein Computerprogramm für all diese Rechenschritte, einschließlich des Tests auf Einfachheit. Möglich, aber unzuverlässig ist die intuitive Suche nach einer einfacheren Partition anhand von Abbildung 4.2. Möglich und zuverlässig ist die folgende Rechnung mit Bündeln von Blättern. Sie ist von Anfang bis Ende mit Symbolen A und B geschrieben. Jeder Schritt führt von einer alten zwiefachen Partition zu einer neuen, die mindestens genauso einfach ist. Die abschließende Rekonstruktion einer nicht mehr zwiefachen generierenden Partition ergibt dann als Endergebnis die Definitionen 6.23 und 6.24 der Symbole 1 und 0, die mit 4.33 und 4.32 übereinstimmen.

5.2.7 Die anfänglichen Bündelgrößen der AB–Partition

Die AB-Partition bestimmt die anfänglichen Mengen der expandierenden Blätter $\{(A \sqcup B)^{-\infty}\circ\}$' und der kontrahierenden Blätter $\{\circ(A \sqcup B)^{+\infty}\}$'.

Wie groß sollen die Bündel sein? Die großen Bündel $\{(A \sqcup B)^{-\infty} A\circ\}$' und $\{(A \sqcup B)^{-\infty} B\circ\}$' ergeben sich, wenn man in Gl. 5.5 die Symbolischen Gebiete $A_{1\circ} = A\circ$ und $A_{2\circ} = B\circ$ wählt. Durch $\circ B_1 = \circ A$ und $\circ B_2 = \circ B$ erhält man analog die kontrahierenden Bündel $\{\circ A(A \sqcup B)^{+\infty}\}$' und $\{\circ B(A \sqcup B)^{+\infty}\}$'. Wenn man mit diesen Bündeln beginnen würde, dann würde es sich nach Berechnung der Orbitpaare herausstellen, daß alle Bündel miteinander Orbitpaare bilden. In diesem Fall wären keine Vereinfachungen möglich.

Man muß die anfänglichen Bündel daher kleiner wählen Es gibt keinen Grund, ein bestimmtes Bündel zu verkleinern. Deshalb sollen alle Bündel gleichermaßen verkleinert werden. Dadurch ergeben sich die vier expandierenden Bündel $\{(A \sqcup B)^{-\infty} S_{-2} S_{-1}\circ\}$' und die vier kontrahierenden Bündel $\{\circ S_0 S_1 (A \sqcup B)^{+\infty}\}$', mit $S_i \in \{A, B\}$. Man beachte, daß sich die Größe der expandierenden und kontrahierenden Blätter nicht ändert. Die entsprechende zwiefache Partition ist $\{AA\circ, AB\circ, BA\circ, BB\circ\}$, $\{\circ AA, \circ AB, \circ BA, \circ BB\}$. Von dieser Partition geht die folgende Rechnung aus.

Da es nicht eindeutig vorgegeben ist, wie groß die Bündel zu wählen sind, könnte man auch auf den Gedanken kommen, noch kleinere Bündel zu wählen, zum Beispiel $\{(A \sqcup B)^{-\infty} S_{-3} S_{-2} S_{-1}\circ\}$' und $\{\circ S_0 S_1 S_2 (A \sqcup B)^{+\infty}\}$'. Dies würde die folgende Rechnung erschweren, das Ergebnis aber nicht ändern.

5.3 Hilfsmittel

Eine Besonderheit der folgenden Beispielrechnung ist, daß meist nur auf wenigen Bündeln zugleich operiert wird. Die meisten Operationen betreffen nur ein oder zwei Bündel. Trotzdem braucht man

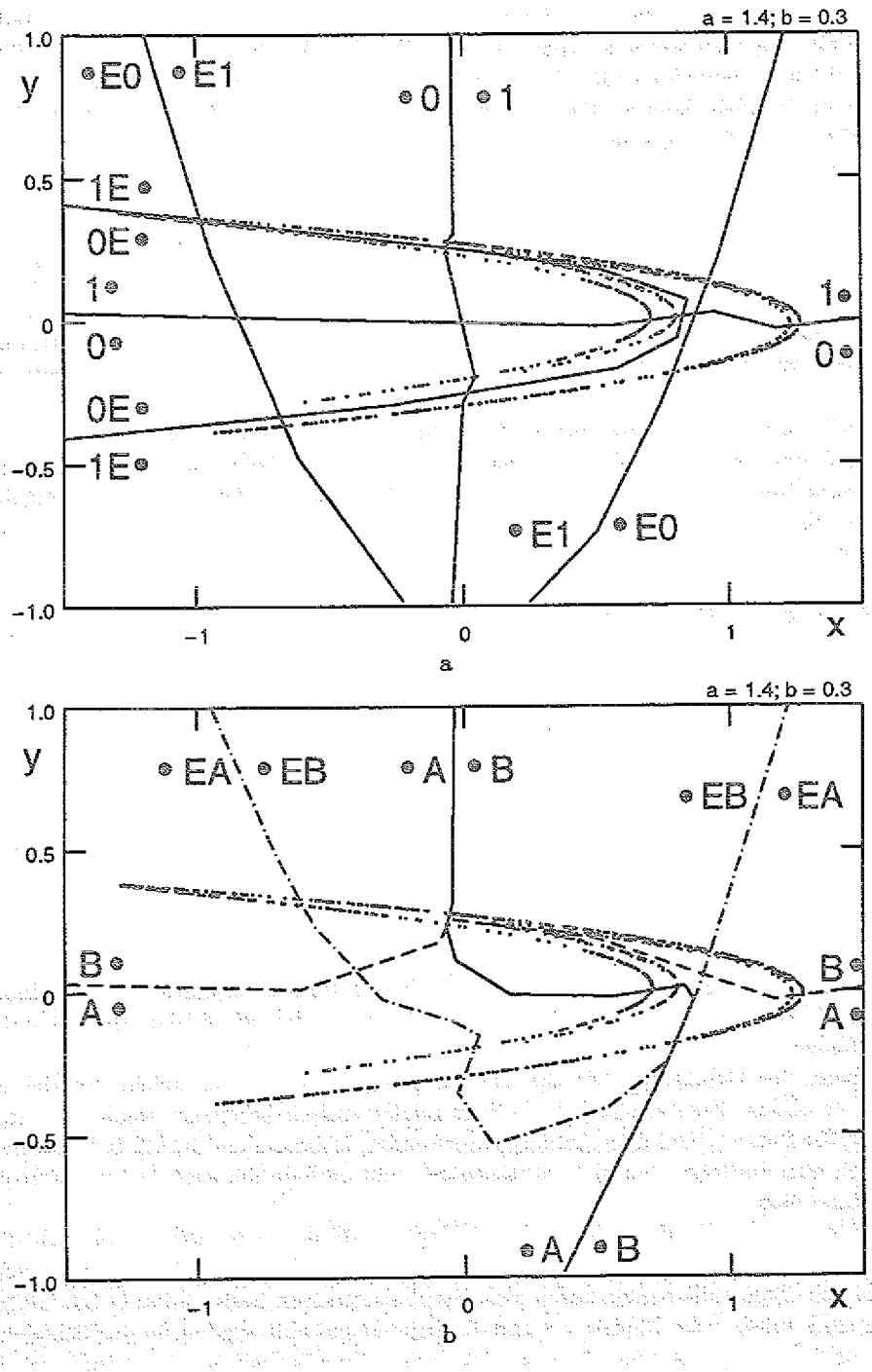


Abbildung 5.2: Vergleich der 01–Partition aus Abb. 2.4 und der *AB*–Partition aus Abb. 4.2. Ausgehend von der *AB*–Partition des Hénon Attraktors bei Standardparameterwerten wird in diesem Kapitel die einfachere 01–Partition berechnet.

nach jedem Vereinfachungsschritt einen Überblick über alle Bündel, um diejenigen zu finden, auf denen der nächste Vereinfachungsschritt operieren kann. Dazu dienen zwei Hilfsmittel: In Tabellen wird zusammengefaßt, welche expandierenden Bündel mit welchen kontrahierenden Bündeln Orbitpaare bilden. Und in Graphen wird illustriert, von welchen expandierenden Bündeln man unter der Abbildung f zu welchen expandierenden Bündeln kommt. Andere Graphen illustrieren analog, von welchen kontrahierenden Bündeln man unter der inversen Abbildung f^{-1} zu welchen kontrahierenden Bündeln kommt. Damit sind alle Beziehungen zwischen Bündeln erfaßt, die von Bedeutung für die Rechnung sind. Wie diese Beziehungen aussehen, hängt von den grammatikalischen Regeln der Symbolischen Dynamik ab.

5.3.1 Die Tabelle der Orbitpaare

Berechnung der Orbitpaare

Die expandierenden und kontrahierenden Bündel sind dadurch miteinander verknüpft, daß sie Orbitpaare bilden. Welche Bündel Orbitpaare bilden, muß vor jedem Vereinfachungschritt berechnet werden.

Dazu muß man Symbolsequenzen betrachten, die Symbole aus beiden Partitionen $\{A_1\circ,\ldots A_m\circ\}$ und $\{\circ B_1,\ldots\circ B_n\}$ miteinander vermischen. Um zu unterscheiden, welche dieser Symbolsequenzen zulässig und welche verboten sind, müssen nicht nur grammatikalische Regeln bekannt sein, sondern auch Regeln, die Symbole aus beiden Partitionen in eine gemeinsame Darstellung übersetzen.

Im Falle der Beispielrechnung ist diese gemeinsame Schreibweise einfach durch die anfängliche AB–Partition gegeben. Die Übersetzungsregeln drücken die zwiefache Partition $\{A_1\circ,\ldots A_m\circ\}$, $\{\circ B_1,\ldots\circ B_n\}$, welche nach jedem Vereinfachungsschritt eine andere ist, durch die immer gleichen Symbolischen Gebiete $\circ A$ und $\circ B$ aus.

Für die zeitliche Verschiebung eines expandierenden Bündels

$$\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty}A_j\circ\}' \tag{5.16}$$

wird es sich später als unwichtig erweisen, wie viele Blätter dieses Bündel Orbitpaare mit kontrahierenden Blättern eines Bündel

$$\{\circ B_k(\bigsqcup_{i=1}^{n} B_i)^{+\infty}\}' \tag{5.17}$$

bildet. Wichtig ist nur, ob die beiden Bündel überhaupt Blätter enthalten, die miteinander Orbitpaare bilden. Wenn dies zutrifft, dann werde ich sagen, daß die beiden Bündel miteinander Orbitpaare bilden.

Eine Tabelle der Orbitpaare wie die Tabelle 5.3 faßt zusammen, welche Bündel miteinander Orbitpaare bilden. Eine solche Tabelle kann relativ einfach berechnet werden, wenn man die Rechenregeln für Paare $[,]$ aus dem Anhang B verwendet, insbesondere die Gl. B.49. Diese Rechenregeln will ich aber im folgenden nicht voraussetzen und deshalb die entsprechenden Rechnungen auch nicht darstellen.

Im Grunde besteht die Berechnung der Orbitpaare darin, nachzuprüfen, ob sich die beiden Symbolsequenzen $\ldots A_j\circ SB_k\ldots$ und $\ldots A_j\circ TB_k\ldots$ bei den Punkten $\ldots$ auf gleiche Weise fortsetzen lassen. Damit die beiden durch diese Symbolsequenzen bestimmten Orbits auch wirklich verschieden sind, müssen die Gebiete $\circ S$ und $\circ T$ disjunkt gewählt werden. In der Beispielrechnung genügt es, einfach $\circ S = \circ A$ und $\circ T = \circ B$ zu setzen. Im allgemeinen Fall können Symbolische Gebiete überlappen. Aber auch dann kann man, zum Beispiel durch $\circ S = \circ A_1 \setminus \circ A_2$ und $\circ T = \circ A_2$, disjunkte Gebiete konstruieren.

Anmerkungen

Wenn keine Sequenzen aus Symbolen A_i und B_j verboten sind, dann bilden alle Bündel miteinander Orbitpaare. In diesem Fall kann man die zwiefache Partition nicht vereinfachen. Auch wenn

	$\circ AA$	$\circ AB$	$\circ BA$	$\circ BB$
$AA\circ$	$\emptyset$	$\emptyset$	$\emptyset$	$\star$
$BA\circ$	$\star$	$\emptyset$	$\emptyset$	$\emptyset$
$AB\circ$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$
$BB\circ$	$\star$	$\emptyset$	$\star$	$\emptyset$

Abbildung 5.3: Die anfänglichen expandierenden und kontrahierenden Bündel und die durch sie gebildeten Orbitpaare. Eine Sequenz $S_{-2}S_{-1}\circ$ kennzeichnet das expandierende Bündel $\{(A \sqcup B)^{-\infty}S_{-2}S_{-1}\circ\}$' und eine Sequenz $\circ S_0S_1$ kennzeichnet das kontrahierende Bündel $\{\circ S_0S_1(A \sqcup B)^{+\infty}\}$'. Die leere Menge $\emptyset$ als Tabelleneintrag bedeutet, daß die beiden Bündel aufgrund der grammatikalischen Regeln 5.18 keine Orbitpaare bilden. Der Stern $\star$ bedeutet, daß Orbitpaare höchstwahrscheinlich gebildet werden.

man sehr viele kleine expandierende und kontrahierende Bündel wählen muß, damit wenigstens ein paar Bündel keine Orbitpaare miteinander bilden, dann ist es unwahrscheinlich, daß Vereinfachungen möglich sind. Dies hängt damit zusammen, daß man eine Mindestzahl von Symbolen braucht, damit die Symbolische Dynamik dieselben dynamischen Entropien hat wie die Phasenraumdynamik (siehe Kap. 2.2.4). Wenn keine oder nur sehr lange Symbolsequenzen verboten sind, dann hat man diese Mindestzahl vermutlich schon erreicht.

Da man normalerweise nicht alle grammatikalischen Regeln kennt, kann man zwar in manchen Fällen die Existenz von Orbitpaaren definitiv ausschließen, aber in den restlichen Fällen die Existenz von Orbitpaaren nicht exakt beweisen. Das Risiko, hierbei zu irren, scheint aber gering. Dies kommt daher, daß ein Symbol S_0 durch grammatikalische Regeln normalerweise einen viel größeren Einfluß auf nahe als auf ferne Symbole derselben Sequenz ausübt. Wenn man zu beiden Seiten von S oder T schon lange zulässige Fortsetzungen gefunden hat, dann spielt es deshalb beim Anfügen weiterer Symbole kaum eine Rolle, ob das Symbol S oder das Symbol T in der Mitte der Sequenz steht. Dann sollte es möglich sein, beide Sequenzen auf dieselbe Weise fortzusetzen. Und vermutlich gibt es nicht nur eine zulässige Fortsetzung, sondern eine große Zahl, die bei positiver dynamischer Entropie exponentiell schnell mit der Entfernung anwächst.

Orbitpaare der AB–Partition

Wir hatten die vier expandierenden Bündel $\{(A \sqcup B)^{-\infty}S_{-2}S_{-1}\circ\}$' und die vier kontrahierenden Bündel $\{\circ S_0S_1(A \sqcup B)^{+\infty}\}$' gewählt (mit $S_i \in \{A, B\}$). Es muß also in 16 Fällen nach Orbitpaaren gesucht werden. Dabei werden die grammatikalischen Regeln aus Gl. 4.37

$$\circ AABA \quad = \quad \circ BABB \quad = \quad \emptyset \tag{5.18}$$

verwendet. Das Ergebnis ist in den 16 Einträgen der Tab. 5.3 dargestellt.

Als Beispiel will ich den Eintrag $\emptyset$ oben links erläutern. Er besagt, daß das expandierende Bündel $\{(A \sqcup B)^{-\infty}AA\circ\}$' keine Orbitpaare mit dem kontrahierenden Bündel $\{\circ AA(A \sqcup B)^{+\infty}\}$' bildet. Denn solche Orbitpaare müßten aus einem Orbit mit der Symbolsequenz $\ldots AA \circ AAA\ldots$ und einem Orbit mit der Symbolsequenz $\ldots AA \circ BAA\ldots$ bestehen. Die zweite Sequenz, bei der das Symbol B in der Mitte steht, ist aber wegen der grammatikalischen Regel $AA \circ BA = \emptyset$ verboten.

5.3.2 Der Graph der zeitlichen Übergänge

Konstruktion des Graphen

In Kap. 3.2.1 wurde die Dynamik f auf den expandierenden Blättern als ein Strecken und Teilen bezeichnet. Ein expandierendes Blatt $\alpha = \ldots S_{-2}S_{-1}A_i\circ$ (mit $S_t \in \{A_1, \ldots A_m\}$) wird zu einem Blatt $f(\alpha) = \ldots S_{-2}S_{-1}A_iE\circ$ gestreckt. Dieses ist im allgemeinen zu groß, um selber zu

den Blättern zu gehören, die als expandierend deklariert wurden. Dagegen gehören seine Teile $\ldots S_{-2}S_{-1}A_iA_j\circledcirc$ zu den expandierenden Blättern. Es ist praktisch, diese Teilung schon vor dem Strecken zu vollziehen. Dann wird das expandierende Blatt $\alpha = \ldots S_{-2}S_{-1}A_i\circledcirc$ in höchstens m Blätter $\alpha_j = \ldots S_{-2}S_{-1}A_i \circledcirc A_j$ geteilt, deren Bilder zu den expandierende Blätter gehören.

Den besten Überblick über dieses Strecken und Teilen erhält man, wenn man die zeitlichen Übergänge graphisch darstellt. Ein Beispiel für so einen Graphen zeigt Abb. 5.4. Die Ecken des Graphen sind mit den Symbolischen Gebieten der Partition

$$\{A_1\circledcirc, \ldots A_m\circledcirc\} \quad = \quad \{AA\circledcirc, AB\circledcirc, BA\circledcirc, BB\circledcirc\} \tag{5.19}$$

beschriftet ($m = 4$). Die Übergänge zwischen diesen Ecken werden durch die Dynamik f induziert: Der Übergang $A_i\circledcirc \to A_j\circledcirc$ existiert, falls

$$f(A_i\circledcirc) \quad \cap \quad A_j\circledcirc \quad \neq \quad \emptyset \quad . \tag{5.20}$$

Dadurch wird garantiert, daß jeder Punkt in $A_i\circledcirc$ entlang mindestens einem dieser Übergänge abgebildet wird, das heißt:

$$f(A_i\circledcirc) \quad \subseteq \quad \bigcup_{j\in J(i)} A_j\circledcirc \quad . \tag{5.21}$$

Hierbei ist $J(i)$ die Menge aller Indices j der Ecken $A_j\circledcirc$, zu denen Übergänge von $A_i\circledcirc$ führen.

Anmerkungen

Wenn von einem Symbolischen Gebiet $A_i\circledcirc$ nur q Übergänge ausgehen, dann wird ein expandierendes Blatt $\alpha = \ldots S_{-2}S_{-1}A_i\circledcirc$ vor dem nächsten Zeitschritt auch nur in höchstens q Teile $\alpha_j = \ldots S_{-2}S_{-1}A_i \circledcirc A_j$ (mit $j \in J(i)$) geteilt.

$$\alpha = \bigcup_{j\in J(i)} \alpha_j \quad . \tag{5.22}$$

Wenn die Symbolischen Gebiete $A_i\circledcirc$ überlappen, kann es prinzipiell passieren, daß ein Teil ganz in einem anderen Teil enthalten ist:

$$\ldots S_{-2}S_{-1}A_i \circledcirc A_j \quad \subseteq \quad \ldots S_{-2}S_{-1}A_i \circledcirc A_k \quad . \tag{5.23}$$

Der kleinere Teil kann dann ignoriert werden (siehe Kap. 5.2.4). Wenn sogar $A_i \circledcirc A_j \subseteq A_i \circledcirc A_k$ gilt, dann können alle expandierenden Blätter dieser Gestalt $\ldots S_{-2}S_{-1}A_iA_j\circledcirc$ entfernt werden. Dies bedeutet, daß der Übergang $A_i\circledcirc \to A_j\circledcirc$ entfernt werden kann, obwohl $A_i \circledcirc A_j \neq \emptyset$. In der Beispielrechnung wird kein Übergang auftreten, der auf diese Weise entfernt werden kann.

Bündel, deren Blätter nicht geteilt werden

Nicht jedes expandierende Blatt α muß beim nächsten Zeitschritt geteilt werden. Wenn es nicht geteilt wird, dann gehört sein Bild $f(\alpha)$ zu den expandierenden Blättern. Da dieses Bild keine doppelten Schnitte mit kontrahierenden Blättern haben kann, ohne die Bedingung 3.2.1 für generierende Partitionen (aus Kap. 3.2.3) zu verletzen, kann das expandierende Blatt α auch keine Orbitpaare mit kontrahierenden Blättern bilden.

Es kann sogar ganze expandierende Bündel geben, die kein Blatt enthalten, das im nächsten Zeitschritt geteilt wird. Diese Bündel bestehen dann nur aus Urbildern expandierender Blätter. Und sie bilden keine Orbitpaare mit kontrahierenden Bündeln. Manche dieser Bündel $\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} A_j\circledcirc\}^{\epsilon}$ kann man daran erkennen, daß von dem entsprechenden Symbolischen Gebiet $A_j\circledcirc$ nur ein einziger Übergang ausgeht. Solche Bündel werden in den Tabellen der Orbitpaare (wie z. B. Tab. 5.13) gar nicht erst aufgeführt.

Aber auch zwei oder mehr Übergänge garantieren nicht, daß irgendwelche Blätter geteilt werden. Dies liegt an den grammatikalischen Regeln. Wenn es für jede Symbolsequenz $\ldots S_{-2}S_{-1}$ einen Index k und eine Regel

$$\ldots S_{-2}S_{-1}A_i\circledcirc \quad = \quad \ldots S_{-2}S_{-1}A_i \circledcirc A_k \tag{5.24}$$

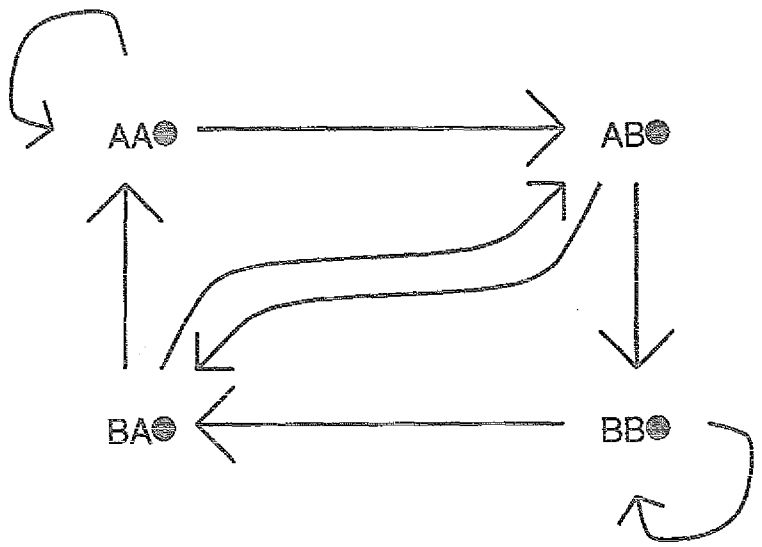


Abbildung 5.4: Die Übergänge zwischen den Symbolischen Gebieten der Partition $\{AA\circ, AB\circ, BA\circ, BB\circ\}$ unter der Abbildung f. Am Beginn der Beispielrechnung werden die expandierenden Bündel $\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} A_j\circ\}$' durch diese Partition erzeugt. Ein Übergang von $A_i\circ$ nach $A_j\circ$ existiert dann, wenn $f(A_i\circ) \cap A_j\circ = A_iA_j\circ \neq \emptyset$.

gibt, dann wird kein Blatt des expandierenden Bündels $\{(\bigsqcup_{l=1}^{m} A_l)^{-\infty} A_i\circ\}$' geteilt. Doch von dem Gebiet $A_i\circ$ geht mehr als ein Übergang aus, es sei denn, der Index k ist für alle Symbolsequenzen $\ldots S_{-2}S_{-1}$ derselbe.

Umgekehrt wirkt die inverse Dynamik f^{-1} auf den kontrahierenden Blättern wie ein Teilen und Strecken. Die graphische Darstellung, zum Beispiel in Abb. 5.6, geschieht ganz analog. Ein Übergang $\circ B_i \to \circ B_j$ existiert, falls

$$f^{-1}(\circ B_i) \quad \cap \quad \circ B_j \quad \neq \quad \emptyset \quad . \tag{5.25}$$

Wenn von einer Ecke $\circ B_i$ nur ein Übergang ausgeht, dann werden die Blätter des kontrahierenden Bündels $\{\circ B_i(\bigsqcup_{l=1}^{m} B_l)^{-\infty}\}$' im nächsten Zeitschritt nicht geteilt.

Blätter der AB-Partition, die nicht geteilt werden

Als Beispiel sind in Abb. 5.4 die Übergänge zwischen den Symbolischen Gebieten der Partition $\{AA\circ, AB\circ, BA\circ, BB\circ\}$ aufgetragen. Von jeder Ecke des Graphen gehen zwei Übergänge aus. Man könnte daher glauben, daß alle vier expandierenden Bündel Blätter enthalten, die im nächsten Zeitschritt geteilt werden. Dies gilt tatsächlich für drei dieser Bündel, jedoch nicht für das Bündel $\{(A \sqcup B)^{-\infty} AB\circ\}$'. Dieses kann wegen $BAB \circ B = AAB \circ A = \emptyset$ umgeschrieben werden zu

$$\{(A \sqcup B)^{-\infty} AB\circ\}' \quad = \quad \{(A \sqcup B)^{-\infty}(BAB \circ A \quad \sqcup \quad AAB \circ B)\}' \quad . \tag{5.26}$$

Es enthält daher nur Blätter, deren Bilder auch wieder expandierende Blätter sind.

$$f(\{(A \sqcup B)^{-\infty} BAB \circ A\}') \quad \subseteq \quad \{(A \sqcup B)^{-\infty} BA\circ\}' \tag{5.27}$$

$$f(\{(A \sqcup B)^{-\infty} AAB \circ B\}') \quad \subseteq \quad \{(A \sqcup B)^{-\infty} BB\circ\}' \tag{5.28}$$

Folglich bildet dieses Bündel keine Orbitpaare (siehe Tab. 5.3).

Man kann auch graphisch ausdrücken, daß diese Blätter im nächsten Zeitschritt nicht geteilt werden. Dazu muß man das Symbolische Gebiet $AB\circ$ in die beiden kleineren Gebiete $AAB\circ$ und $BAB\circ$ aufspalten. Die neue Partition, die in Abb. 5.5 dargestellt ist, erzeugt dieselben expandierenden Blätter wie die alte Partition von Abb. 5.4. Doch von $AAB\circ$ und $BAB\circ$ geht jetzt nur noch

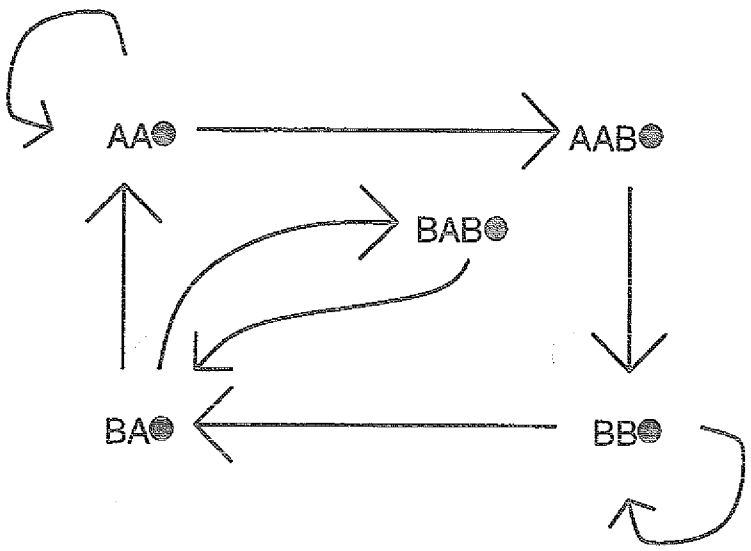


Abbildung 5.5: Die Partition, die beschreibt, welche expandierenden Blätter unter der Abbildung f geteilt werden. Diese Partition erzeugt dieselben expandierenden Blätter wie die Partition in Abb. 5.4. Das expandierende Bündel $\{(A \sqcup B)^{-\infty} AB\circ\}$' wurde in seine Teilmengen $\{(A \sqcup B)^{-\infty} AAB\circ\}$' und $\{(A \sqcup B)^{-\infty} BAB\circ\}$' zerlegt. Keines der darin enthaltenen expandierenden Blätter wird im nächsten Zeitschritt geteilt. Denn von $AAB\circ$ und $BAB\circ$ geht wegen $AAB \circ A = BAB \circ B = \emptyset$ nur jeweils ein Übergang aus.

jeweils ein Übergang aus, so daß man sofort sieht, daß keines der entsprechenden expandierenden Blätter im nächsten Zeitschritt geteilt wird.

Analoges gilt für kontrahierende Bündel. Die Partition $\{\circ AA, \circ AB, \circ BA, \circ BB\}$ definiert ein kontrahierendes Bündel, nämlich $\{\circ AB(A \sqcup B)^{+\infty}\}$', das mit keinem expandierenden Bündel Orbitpaare bildet (Tab. 5.3). Das Fehlen dieser Orbitpaare kann man aus Abb. 5.6 nicht ablesen, wohl aber aus Abb. 5.7, wo das kontrahierende Bündel $\{\circ AB(A \sqcup B)^{+\infty}\}$' in $\{\circ ABB(A \sqcup B)^{+\infty}\}$' und $\{\circ ABA(A \sqcup B)^{+\infty}\}$' zerlegt wurde.

Man braucht diese nützlichen Zerlegungen von Symbolischen Gebieten nicht zu erraten. Denn man stößt automatisch auf sie, wenn man Bündel zeitlich verschiebt (Kap. 5.4.5).

5.4 Vereinfachungsschritte

5.4.1 Überblick

Ich werde im folgenden die einzelnen Vereinfachungsschritte im Detail erklären, und zwar sowohl für den allgemeinen Fall als auch für das Beispiel der AB-Partition. Da man dabei aber leicht den Überblick verliert, werde ich zunächst das Schema zusammenfassen, nach dem zwiefache Partitionen vereinfacht werden. Manche Formulierungen in diesem Überblick werden dem Leser vage erscheinen, doch sie werden alle in späteren Abschnitten präzisiert.

Zu Beginn sind eine zwiefache Partition, ihre expandierenden Bündel und ihre kontrahierenden Bündel gegeben. Vor der ersten Vereinfachung müssen Beziehungen zwischen den expandierenden und den kontrahierenden Bündeln hergestellt werden. Dazu berechnet man mit Hilfe der grammatikalischen Regeln und etwaiger Übersetzungsregeln, welche expandierenden Bündel mit welchen kontrahierenden Bündeln Orbitpaare bilden.

Es sei daran erinnert, daß es nicht nur von der Dynamik, sondern auch ganz entscheidend von der zwiefachen Partition abhängt, ob ein Blatt expandierend ist oder ob es kontrahierend ist. Wenn die zwiefache Partition geändert wird, dann ändern sich manche der expandierenden und kontrahierenden Bündel. Umgekehrt kann man Änderungen der zwiefachen Partition beschreiben,

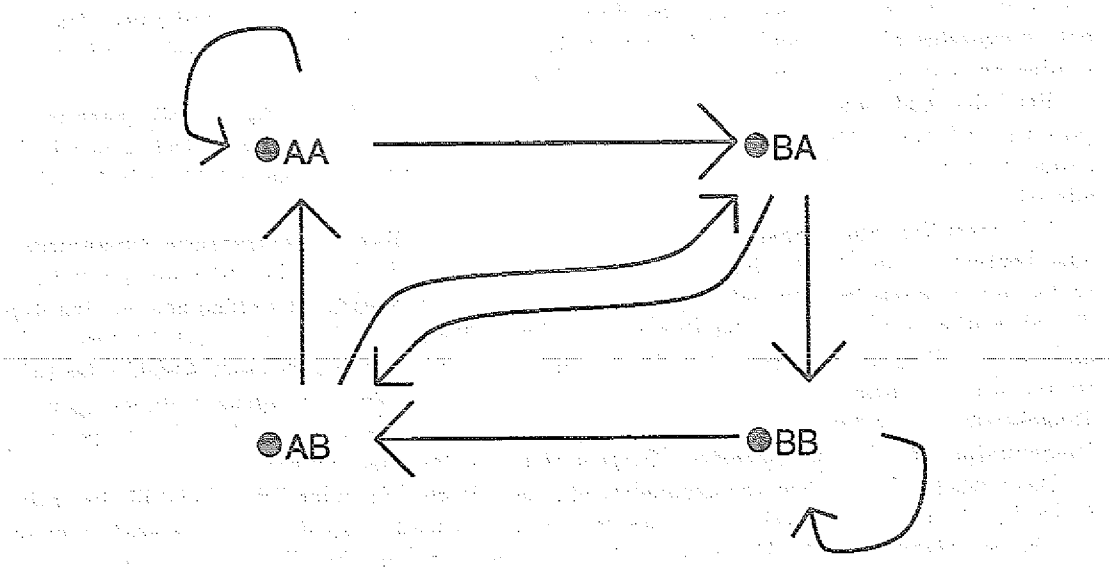


Abbildung 5.6: Die Übergänge zwischen den Symbolischen Gebieten der Partition $\{\circ AA, \circ AB, \circ BA, \circ BB\}$ unter der inversen Abbildung f^{-1}. Am Beginn der Beispielrechnung werden die kontrahierenden Bündel $\{\circ B_j(\bigsqcup_{i=1}^{m} B_i)^{+\infty}\}'$ durch diese Partition erzeugt. Ein Übergang von $\circ B_i$ nach $\circ B_j$ existiert dann, wenn $f^{-1}(\circ B_i) \cap \circ B_j = \circ B_j B_i \neq \emptyset$.

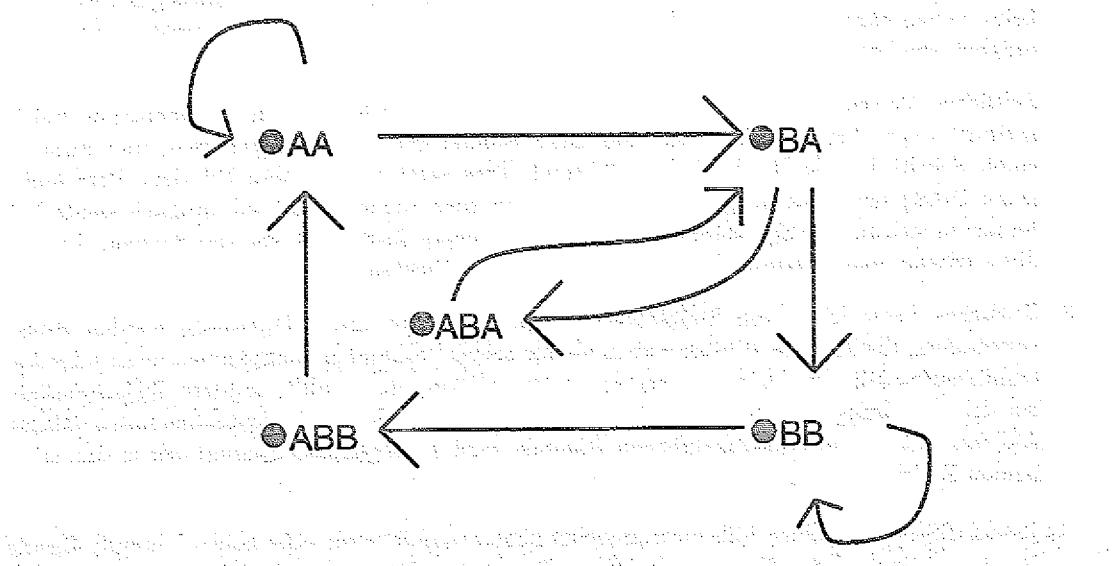


Abbildung 5.7: Die Partition, die beschreibt, welche kontrahierenden Blätter unter der inversen Abbildung f^{-1} geteilt werden. Diese Partition erzeugt dieselben kontrahierenden Blätter wie die Partition in Abb. 5.6. Das kontrahierende Bündel $\{\circ AB(A \sqcup B)^{+\infty}\}'$ wurde in seine Teilmengen $\{\circ ABA(A \sqcup B)^{+\infty}\}'$ und $\{\circ ABB(A \sqcup B)^{+\infty}\}'$ zerlegt. Keines der darin enthaltenen kontrahierenden Blätter unter der inversen Abbildung f^{-1} geteilt. Denn von $\circ ABA$ und $\circ ABB$ geht wegen der grammatikalischen Regel $A \circ ABA = B \circ ABB = \emptyset$ nur jeweils ein Übergang aus.

indem man die Änderungen der expandierenden und kontrahierenden Bündel angibt. Bei diesen Änderungen werden entweder expandierende Bündel als nicht mehr expandierend deklariert oder umgekehrt nicht expandierende Bündel als expandierend deklariert. Kontrahierende Bündel werden auf analoge Weise entfernt oder neu deklariert.

Bei jeder Änderung der Partition muß die Bedingung 3.2.1 für die Eigenschaft „generierend" (aus Kap. 3.2.3) beachtet werden. Außerdem sollen die zusätzlichen Bündel neue Orbitpaare bilden, damit die neue zwiefache Partition nach dem Kriterium 4.1.2 (aus Kap. 4.1.2) einfacher als die alte ist.

Bei jedem Vereinfachungsschritt werden nicht willkürliche Bündel als zusätzliche expandierende oder kontrahierende Bündel deklariert, sondern nur solche, die durch die Dynamik f aus bereits vorhandenen expandierenden oder kontrahierenden Bündeln konstruiert werden können. Ich werde dies als zeitliches Verschieben der Bündel bezeichnen. Man kann sich dieses zeitliche Verschieben grob so vorstellen, daß ein Stück einer Partitionsgrenze entfernt und durch sein Abbild oder Urbild ersetzt wird. In Kap. 2.3 wurde der Fall erwähnt, daß die Partitionsgrenzen durch „primäre" Tangentialpunkte gehen. In diesem Fall würden also die Abbilder oder Urbilder dieser „primären" Tangentialpunkte als neue „primäre" Tangentialpunkte deklariert werden.

Das zeitliche Verschieben von expandierenden oder kontrahierenden Bündeln ist die Operation, durch die eine zwiefache Partition schrittweise vereinfacht wird. Vor jedem dieser Schritte muß man sich fragen: Welche Bündel können zeitlich verschoben werden, ohne Bed. 3.2.1 für generierende zwiefache Partitionen zu verletzen? Wie findet man diese Bündel? Welche neuen expandierenden oder kontrahierenden Bündel ergeben sich? Ist der Vereinfachungsschritt erfolgreich, das heißt, werden zusätzliche Orbitpaare gebildet? Diese Fragen lassen sich besser beantworten, wenn drei Varianten des zeitlichen Verschiebens unterschieden werden.

1. Zeitliches Verschieben eines einzelnen Bündels (Kap. 5.4.4). Diese Operation setzt voraus, daß das expandierende (bzw. kontrahierende) Bündel mit keinem kontrahierenden (bzw. expandierenden) Bündel Orbitpaare bildet. Dann wird das Bild des expandierenden Bündels als expandierend deklariert. Es kann passieren, daß diese Operation überflüssig ist, weil dabei keine neuen expandierenden Blätter entstehen. Anderenfalls wird die Operation so oft wie möglich wiederholt.

2. Zeitliches Verschieben von mehreren Bündeln (Kap. 5.4.6). Bei dieser Operation werden mehrere expandierende und kontrahierende Bündel gemeinsam verschoben, und zwar um einen Schritt in dieselbe zeitliche Richtung. Dies setzt voraus, daß bei dem Verschieben keine Orbitpaare verloren gehen. Die Operation wird wieder so oft wie möglich wiederholt. Es lassen sich immer alle Bündel gemeinsam um einen Zeitschritt vor verschieben, doch auf diese triviale und nutzlose Operation sollte man verzichten.

3. Zeitliches Verschieben von Teilblättern (Kap. 5.4.8). Bei dieser Operation werden Bündel verschoben, die kleinere Blätter haben als die schon vorhandenen expandierenden oder kontrahierenden Bündel. Diese Blätter sind nicht willkürlich gewählt, sondern Zwischenstufen bei der Zerteilung eines schon vorhandenen expandierenden oder kontrahierenden Blattes. Auf den so konstruierten zusätzlichen Bündeln wird dann genauso operiert wie in den ersten beiden Fällen.

In jedem dieser drei Fälle erhält man zunächst neue expandierende oder kontrahierende Bündel. Aus diesen muß man dann die neue zwiefache Partition $\{A_1\circ, \ldots A_m\circ\}$, $\{\circ B_1, \ldots \circ B_n\}$ rekonstruieren, bevor man den nächsten Vereinfachungsschritt einleiten kann. Bei der Berechnung der Orbitpaare dieser neuen Partition nützt es sehr, die Orbitpaare der alten Partition zu kennen.

Ansonsten verläuft die Methode so wie andere Gradientensuchverfahren: Durch jeden Einzelschritt kommt man zu einer einfacheren zwiefachen Partition. Man ist am Ende angelangt, wenn sich keine der drei vereinfachenden Operationen mehr anwenden läßt. Kap. 7.2.1 wird erwähnen, wie garantiert werden kann, daß dieses Verfahren endet.

Das Endergebnis kann von der Reihenfolge abhängen, mit der die Operationen angewendet werden. Um alle möglichen Endergebnisse, das heißt alle einfachsten Partitionen, zu erhalten,

müßte man im Prinzip alle möglichen Reihenfolgen ausprobieren. Der Rechenaufwand könnte dabei kombinatorisch explodieren. In der Beispielrechnung wird die AB–Partition nur auf eine mögliche Weise vereinfacht. Eine andere Reihenfolge der Operationen kann zu einem anderen Ergebnis führen. Diese andere einfachste Partition wird erst in Kap. 7.1.2 dargestellt werden.

5.4.2 Zeitliches Verschieben von Blättern

Man kann jedes expandierende bzw. kontrahierende Blatt α in zwei Zeitrichtungen verschieben: vor oder zurück. Für alle vier Fälle soll nun genau definiert werden, welche expandierenden (bzw. kontrahierenden) Blätter aufhören, expandierend zu sein, und welche neuen expandierenden Blätter deklariert werden. Dabei sollte man beachten, daß Urbilder von expandierenden Blättern selbst wieder expandierende Blätter oder Teilmengen derselben sind.

1. Vorverschieben eines expandierenden Blattes α: Das Blatt $f(\alpha)$ wird als neues expandierendes Blatt deklariert. Das Blatt α bleibt expandierend.

2. Rückverschieben eines expandierenden Blattes α: Das Blatt α wird als nicht expandierend deklariert. Das Blatt $f^{-1}(\alpha)$ bleibt ein expandierendes Blatt bzw. ein Teil eines solchen.

3. Vorverschieben eines kontrahierenden Blattes β: Das Blatt β wird als nicht kontrahierend deklariert. Das Blatt $f(\beta)$ bleibt ein kontrahierendes Blatt bzw. ein Teil eines solchen.

4. Rückverschieben eines kontrahierenden Blattes β: Das Blatt $f^{-1}(\beta)$ wird als neues kontrahierendes Blatt deklariert. Das Blatt β bleibt kontrahierend.

Im ersten und im vierten Fall wird ein neues Blatt als expandierend oder kontrahierend deklariert, doch kein Blatt hört auf, expandierend oder kontrahierend zu sein. Im zweiten und dritten Fall hört ein Blatt auf, expandierend oder kontrahierend zu sein. Vor diesen Änderungen bedecken die expandierenden oder kontrahierenden Blätter (fast) alle Attraktorpunkte. Diese vollständige Bedeckung soll bei den Änderungen erhalten bleiben. Daher müssen im zweiten und dritten Fall neue expandierende oder kontrahierende Blätter deklariert werden.

Bei dem Rückverschieben des expandierenden Blattes α müssen neue expandierende Blätter deklariert werden, welche die Lücke bedecken, die bei α entsteht. Solche expandierenden Blätter lassen sich problemlos finden, da laut Kap. 5.2.5 jedes Urbild eines expandierenden Blattes als expandierendes Blatt betrachtet werden kann. Und die Urbilder aller expandierenden Blätter, einschließlich des Urbildes $f^{-1}(\alpha)$, bedecken zusammen alle Attraktorpunkte, also auch die Lücke bei α. Die dort liegenden Urbilder werden bei dem Rückverschieben eines expandierenden Blattes α als neue expandierende Blätter deklariert, um die Lücke zu schließen. Analog erfordert das Vorverschieben eines kontrahierenden Blattes β die explizite Deklaration von neuen kontrahierenden Blättern, die zuvor Urbilder von expandierenden Blättern waren. Diese neuen expandierenden und kontrahierenden Blätter sind kleiner als die Blätter α und β, deren Platz sie einnehmen, und verletzen daher die Bed. 3.2.1 für eine generierende Partition nicht. Allerdings kann sich in diesen Fällen 2 und 3 die Zahl der Orbitpaare verringern.

Im Fall 1 bleibt das Blatt α expandierend, wenn $f(\alpha)$ als neues expandierendes Blatt deklariert wird. Die einzigen expandierenden Blätter, die bei diesem Vorverschieben des expandierenden Blattes α wegfallen, sind diejenigen, die als Teilmengen in dem neuen expandierenden Blatt $f(\alpha)$ enthalten sind und nach der Methode von Kap. 5.2.4 entfernt werden. Analoges gilt für kontrahierende Blätter, die um einen Zeitschritt zurück verschoben werden. Da bei solchen Änderungen keine expandierenden oder kontrahierenden Blätter verschwinden, die nicht als Teilmengen in anderen expandierenden oder kontrahierenden Blättern enthalten sind, kann sich die Zahl der Orbitpaare nicht verringern. Allerdings muß darauf geachtet werden, die Bedingung 3.2.1 für die Eigenschaft „generierend“ nicht zu verletzen.

5.4.3 Zeitliches Verschieben von Bündeln

Da nur auf ganzen Bündeln operiert wird, können auch nur ganze expandierende oder kontrahierende Bündel, und nicht einzelne Blätter, in der Zeit verschoben werden. Das zeitliche Verschieben eines Bündels besteht aus dem zeitlichen Verschieben all seiner Blätter. Der letzte Abschnitt definierte für zeitliches Vorverschieben und Rückverschieben, welche neuen Blätter als expandierend oder kontrahierend deklariert oder welche alten expandierenden oder kontrahierenden Blätter entfernt werden. Hieraus folgt auch, welche neuen Bündel als expandierend oder kontrahierend deklariert oder welche alten expandierenden oder kontrahierenden Bündel entfernt werden.

Vorverschieben eines expandierenden Bündels

Beim zeitlichen Vorverschieben des expandierenden Bündels

$$\Gamma_k \quad = \quad \{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} A_k \circledcirc\}' \tag{5.29}$$

bleibt dieses Bündel expandierend. Außerdem wird das Bündel

$$f(\Gamma_k) \quad = \quad \{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} A_k E \circledcirc\}' \tag{5.30}$$

als neues expandierendes Bündel deklariert. Man beachte, daß die neuen expandierenden Blätter nicht auf ein Symbolisches Gebiet $A_i\circledcirc$ beschränkt sind, sondern sich durch alle diese Symbolischen Gebiete hindurch erstrecken dürfen. Zusammen bedecken diese neuen expandierenden Blätter das Gebiet $[f(\Gamma_k)] = A_k E\circledcirc$. Falls $A_k A_l\circledcirc \neq \emptyset$, dann werden manche der neuen expandierenden Blätter mit Blättern des alten expandierenden Bündels

$$\Gamma_l \quad = \quad \{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} A_l \circledcirc\}' \tag{5.31}$$

überlappen. Die neuen expandierenden Blätter sind zumindest nicht kleiner als die alten. Ob sie gleich groß sein und deshalb mit der Methode von Kap. 5.2.4 entfernt werden können, wird später untersucht werden.

Rückverschieben eines kontrahierenden Bündels

Auf analoge Weise kann ein kontrahierendes Bündel

$$\Lambda_k \quad = \quad \{\circledcirc B_k(\bigsqcup_{i=1}^{n} B_i)^{+\infty}\}' \tag{5.32}$$

in der Zeit zurück verschoben werden. Dabei bleibt dieses Bündel kontrahierend und das Bündel

$$f^{-1}(\Lambda_k) \quad = \quad \{\circledcirc E B_k(\bigsqcup_{i=1}^{n} B_i)^{+\infty}\}' \quad , \tag{5.33}$$

wird als neues kontrahierendes Bündel deklariert. Diese neuen kontrahierenden Blätter sind zumindest nicht kleiner als die bereits vorhandenen kontrahierenden Blätter.

Rückverschieben eines expandierenden Bündels

Wenn das expandierende Bündel

$$\Gamma_k \quad = \quad \{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} A_k \circledcirc\}' \tag{5.34}$$

in der Zeit zurück verschoben wird, dann bleibt Γ_k nicht expandierend. Eine Lücke im Phasenraum, die nicht durch expandierende Blätter abgedeckt würde, entsteht dadurch nicht. Denn als Urbilder von alten expandierenden Blättern bleiben die Blätter in

$$\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} A_k \circ A_l\}' \quad = \quad f^{-1}(\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} A_k A_l \circ\}') \tag{5.35}$$

sogar für $l = k$ expandierend. Und alle diese expandierenden Blätter zusammen decken das Gebiet

$$\bigcup_{l=1}^{m} A_k \circ A_l \quad = \quad A_k \circ \quad = \quad [\Gamma_k] \tag{5.36}$$

ab. Ein Bündel aus Gl. 5.35 muß allerdings nur dann explizit betrachtet werden, wenn seine Blätter nicht alle in größeren Blättern anderer expandierender Bündel enthalten sind. Ansonsten können alle seine Blätter nach der Methode von Kap. 5.2.4 entfernt werden.

Vorverschieben eines kontrahierenden Bündels

Auf analoge Weise kann ein kontrahierendes Bündel

$$\Lambda_k \quad = \quad \{\circ B_k (\bigsqcup_{i=1}^{n} B_i)^{+\infty}\}' \tag{5.37}$$

in der Zeit vor verschoben werden. Das Bündel Λ_k hört dabei auf, kontrahierend zu sein. Eine Lücke im Phasenraum ensteht nicht, weil sein Abbild

$$f(\Lambda_k) \quad = \quad \{B_k \circ (\bigsqcup_{i=1}^{n} B_i)^{+\infty}\}' \quad , \tag{5.38}$$

kontrahierend bleibt. Durch dieses Abbild, aber auch durch die Abbilder anderer kontrahierender Bündel, kann die Lücke $\circ B_k$ in jedem Fall aufgefüllt werden.

5.4.4 1. Variante: Zeitliches Verschieben von einem einzelnen Bündel

Die einfachste Variante, eine zwiefache Partition zu vereinfachen, besteht darin, ein einzelnes expandierendes oder kontrahierendes Bündel zeitlich zu verschieben, während alle anderen expandierenden oder kontrahierenden Bündel zunächst unverändert bleiben. Diese alten Bündel ändern sich nur dadurch, daß möglicherweise einige ihrer Blätter in neuen und größeren expandierenden und kontrahierenden Blättern enthalten sind und entfernt werden können.

Im letzten Abschnitt wurden vier Weisen unterschieden, auf die Bündeln zeitlich verschoben werden können. Neue Orbitpaare können nur dann gebildet werden, wenn ein expandierendes Bündel in der Zeit vor verschoben wird oder wenn ein kontrahierendes Bündel in der Zeit zurück verschoben wird.

Auswählen und Verschieben des Bündels

Dabei muß man aber darauf achten, die Bedingung 3.2.1 für die Eigenschaft „generierend" zu erhalten. Glücklicherweise kann man aus den Orbitpaaren ablesen, bei welchen zeitlichen Verschiebungen diese Bedingung erhalten bleibt. Denn genau dann, wenn ein expandierendes Blatt α mit einem kontrahierenden Blatt β ein Orbitpaar bildet, gibt es auch einen doppelten Schnitt zwischen dem kontrahierenden Blatt β und dem Bild $f(\alpha)$ des expandierenden Blattes. Ein expandierendes Bündel Γ kann daher genau dann in der Zeit vor verschoben werden, wenn es mit keinem der kontrahierenden Bündel Orbitpaare bildet (siehe Abb. 5.8).

Wenn man bereits alle Orbitpaare berechnet und in der Form von Tab. 5.3 aufgetragen hat, dann erkennt man expandierende Bündel, die in der Zeit vor verschoben werden können, sofort

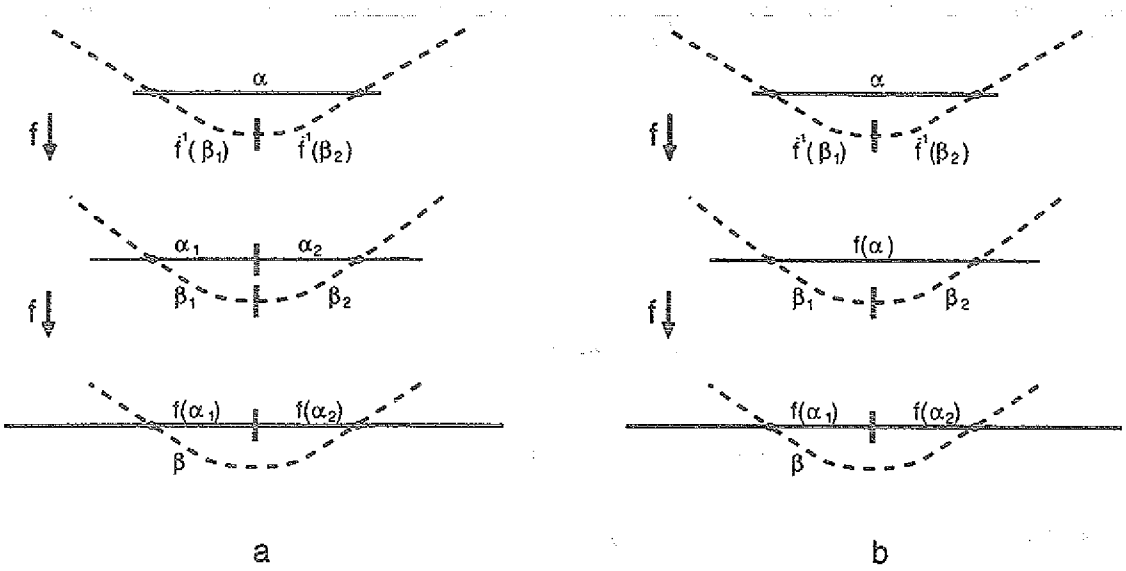


Abbildung 5.8: Bei dem zeitlichen Vorverschieben eines einzelnen expandierenden Blattes α ändert sich die Menge der als expandierend deklarierten Blätter. Abb. a zeigt expandierende Blätter vor der Änderung und Abb. b expandierende Blätter nach der Änderung. Bei drei aufeinanderfolgenden Zeitschritten sind das expandierende Blatt α, seine Bilder $f(\alpha) = \alpha_1 \cup \alpha_2$ und $f^2(\alpha) = f(\alpha_1) \cup f(\alpha_2)$, das kontrahierende Blatt β und dessen Urbilder $f^{-1}(\beta) = \beta_1 \cup \beta_2$ und $f^{-2}(\beta) = f^{-1}(\beta_1) \cup f^{-1}(\beta_2)$ dargestellt. Diese Bilder befinden sich im allgemeinen an verschiedenen Stellen des Phasenraums. Der fette Strich kennzeichnet die Stelle, an der ein expandierendes bzw. kontrahierendes Blatt endet. Das Blatt $f(\alpha)$ ist nur in Abb. b expandierend. In Abb. a sind nur seine Teile α_1 und α_2 expandierende Blätter. Nur in Abb. b bildet daher das expandierende Blatt $f(\alpha)$ ein Orbitpaar mit dem kontrahierenden Blatt β.

daran, daß in ihrer Zeile nur leere Mengen $\emptyset$ stehen. Analog erkennt man kontrahierende Bündel, die in der Zeit zurück verschoben werden können, daran, daß in ihrer Spalte nur leere Mengen stehen.

Wie ein expandierendes Bündel

$$\Gamma_k \quad = \quad \{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} A_k \circledast\}' \tag{5.39}$$

in der Zeit vor verschoben wird, wurde bereits in Kap. 5.4.3 erklärt: Das Bündel

$$f(\Gamma_k) \quad = \quad \{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} A_k E \circledast\}' \tag{5.40}$$

wird als neues expandierendes Bündel deklariert.

Die neuen Orbitpaare

Orbitpaare können bei dem Vorverschieben eines expandierenden Blattes nicht verloren gehen, da keine Blätter ihre Eigenschaft verlieren, expandierend zu sein. Doch entstehen neue Orbitpaare? Zwei Fälle sind möglich:

1. Das neue expandierende Bündel enthält mindestens ein Blatt, das zuvor nicht expandierend waren. Es gibt nun wieder drei Möglichkeiten:
   (a) Wenn dieses neue expandierende Bündel neue Orbitpaare bildet, dann ist die neue zwiefache Partition nach Kriterium 4.1.2 einfacher geworden.
   (b) Wenn dieses neue expandierende Bündel keine Orbitpaare bildet, dann kann man es noch um einen weiteren Zeitschritt vor verschieben.
   (c) Wenn dieses neue expandierende Bündel nur Orbitpaare bildet, die auch von anderen expandierenden Bündeln schon gebildet wurden, dann ist die zwiefache Partition nach dem Kriterium 4.1.2 eigentlich nicht einfacher geworden. Wir haben aber größere expandierende Blätter erhalten, was laut Kap. 4.1 zumindest ein Teilerfolg ist und die zeitliche Verschiebung rechtfertigt.
2. Das neue expandierende Bündel $f(\Gamma_k)$ enthält nur Blätter, die bereits zuvor expandierend waren. Formal heißt das:

$$f(\Gamma_k) \quad = \quad \{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} A_k E \circledast\}' \quad \subseteq \quad \{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} \circledast\}' \quad . \tag{5.41}$$

   Anschaulich bedeutet das, daß die links stehenden Symbole A_i und die grammatikalischen Regeln genügen, um das rechts stehende Symbol E durch ein Symbol A_l zu ersetzen.

   Das zeitliche Verschieben des Bündels Γ_k war daher überflüssig. Wir können vermeiden, diesen überflüssigen Schritt später in der Rechnung zu wiederholen, wenn wir uns merken, daß das Bündel Γ_k nur aus Urbildern von expandierenden Blättern besteht. Am besten merkt man sich dies mit Hilfe der graphischen Darstellung, die in Kap. 5.3.2 eingeführt wurde. Dazu muß man das Symbolische Gebiet $A_k \circledast$ in kleinere Teilgebiete zerlegen, von denen jeweils nur ein Übergang ausgeht. Diese Teilgebiete sind $A_k \circ A_j$ (für alle j mit $A_k \circ A_j \neq \emptyset$). Aus der Gl. 5.41 folgt, daß die neue Partition

$$\{A_1\circledast, \ldots A_{k-1}\circledast, A_{k+1}\circledast, \ldots A_m\circledast, A_k \circ A_1, \ldots A_k \circ A_m\} \tag{5.42}$$

   dieselben expandierenden Blätter erzeugt wie die alte Partition $\{A_1\circledast, \ldots A_m\circledast\}$. Wie das Beispiel im nächsten Abschnitt zeigt, ist diese Zerlegung keine eigentlich neue Methode, sondern eine Kombination der Methoden aus Kap. 5.2.4 und 5.2.5.

Durch ein analoges Verfahren kann auch ein kontrahierendes Bündel um einen Zeitschritt zurück verschoben werden.

	$\circ AA$	$\circ AB$	$\circ BA$	$\circ BB$
$AA\circ$	$\emptyset$	$\emptyset$	$\emptyset$	$\star$
$BA\circ$	$\star$	$\emptyset$	$\emptyset$	$\emptyset$
$AB\circ$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$
$BB\circ$	$\star$	$\emptyset$	$\star$	$\emptyset$

Abbildung 5.9: Wiederholung der Tab. 5.3: Die anfänglichen expandierenden und kontrahierenden Bündel und die durch sie gebildeten Orbitpaare.

5.4.5 Zeitliches Verschieben einzelner Bündel der AB–Partition

Keine neuen Orbitpaare entstehen

In Tab. 5.3, die in Abb. 5.9 wiederholt wird, gibt es nur eine Zeile, die aus lauter leeren Mengen $\emptyset$ besteht. Es gibt daher auch nur ein expandierendes Bündel, das man zeitlich vor verschieben kann, nämlich das Bündel $\{(A \sqcup B)^{-\infty} AB\circ\}'$. Doch wenn man dieses Bündel um einen Zeitschritt verschiebt, erhält man keine neuen expandierenden Blätter. Denn in Gl. 5.26 bis 5.28 wurde gezeigt, daß man bei all den Blättern in $\{(A \sqcup B)^{-\infty} ABE\circ\}'$ das Symbol E entweder durch A oder durch B ersetzen kann, ohne das Blatt zu verkleinern. Deshalb sind alle diese Blätter in $\{(A \sqcup B)^{-\infty}\circ\}'$ enthalten und von vorneherein expandierend. Das zeitliche Vorverschieben dieses expandierenden Bündels ändert daher die Menge der expandierenden Blätter nicht und ist nutzlos. Nützlich wäre nur das Vorverschieben eines expandierenden Bündels, daß keine Orbitpaare bildet und außerdem nicht nur aus Urbildern expandierender Blätter besteht. Ein solches Bündel tritt in Tab. 5.9 nicht auf.

In Tab. 5.9 gibt es außerdem eine Spalte, die aus lauter leeren Mengen $\emptyset$ besteht, und daher ein kontrahierendes Bündel Λ, das mit keinem der expandierenden Bündel Orbitpaare bildet. Man kann daher, ohne die Bed. 3.2.1 für die Eigenschaft „generierend" zu verletzen, dieses Bündel $\Lambda = \{\circ AB(A \sqcup B)^{+\infty}\}'$ um einen Zeitschritt zurück verschieben, das heißt $f^{-1}(\Lambda) = \{\circ EAB(A \sqcup B)^{+\infty}\}'$ als neues expandierendes Bündel deklarieren. Doch wieder sind diese Blätter bereits in anderen kontrahierenden Bündeln enthalten (Kap. 5.3.2), so daß diese zeitliche Verschiebung die Menge der kontrahierenden Blätter nicht ändert.

Da alle anderen Bündel in Tab. 5.9 Orbitpaare bilden, kommen wir mit dieser Methode nicht weiter, was nicht gegen diese Methode spricht. Schließlich wurde die AB–Partition in Kap. 5.2.6 so gewählt, daß sie schon recht einfach ist. Wenn man mit einer feinen Partition begonnen hätte, die viele Symbole und dementsprechend kleine expandierende und kontrahierende Blätter besitzen würde, dann hätte man die Partition durch diese Methode vermutlich stark vereinfachen können.

Aufspaltung des Symbolischen Gebietes $AB\circ$

Das zeitliche Verschieben des expandierenden Bündels $\{(A \sqcup B)^{-\infty} AB\circ\}'$ war nicht völlig nutzlos. Es hat uns gezeigt, daß dieses Bündel nur Urbilder von expandierenden Blättern enthält. Durch die Methode von Kap. 5.3.2 kann dieser Sachverhalt auch graphisch dargestellt werden. Man spaltet das Gebiet $AB\circ$ in kleinere Gebiete auf, von denen jeweils nur noch ein Übergang ausgeht.

Diese kleineren Gebiete wurden im allgemeinen Fall als $A_k \circ A_j$ bezeichnet. In diesem Beispiel sind es demnach die Gebiete

$$AB\circ \cap B\circ A = AB\circ A = BAB\circ A = BAB\circ \quad \text{und} \tag{5.43}$$

$$AB\circ \cap B\circ B = AB\circ B = AAB\circ B = AAB\circ \tag{5.44}$$

(da $AAB \circ A = BAB \circ B = \emptyset$). Abb. 5.5 zeigt die neue Partition, welche die Vergangenheit beschreibt.

Analoges gilt für kontrahierende Bündel. Zeitliches Zurückverschieben des Bündels $\Lambda = \{\circ AB(A \sqcup B)^{+\infty}\}'$ ergab, daß dieses Bündel nur aus Bildern kontrahierender Blätter besteht. Wenn man

das Symbolische Gebiet $\circ AB$ von Abb. 5.6 in die Gebiete $\circ ABB$ und $\circ ABA$ von Abb. 5.7 zerlegt, dann geht von diesen Gebieten nur noch jeweils ein Übergang aus.

Beispiel für das Entfernen zu kleiner expandierender Blätter

Man kann diese kleineren Gebiete aber auch mit den Methoden von Kap. 5.2.4 und 5.2.5 erzeugen. Dies will ich hier darstellen, um ein Beispiel für diese Methoden zu geben. Laut Abschnitt 5.2.5 können die Urbilder expandierender Bündel

$$\{(A \sqcup B)^{-\infty} B \circ A\}' = f^{-1}(\{(A \sqcup B)^{-\infty} BA\circ\}') \quad \text{und} \tag{5.45}$$

$$\{(A \sqcup B)^{-\infty} B \circ B\}' = f^{-1}(\{(A \sqcup B)^{-\infty} BB\circ\}') \tag{5.46}$$

selbst wieder als expandierende Bündel betrachtet werden. Wegen Gl. 5.26 haben sie einige ihrer Blätter mit dem Bündel $\{(A \sqcup B)^{-\infty} AB\circ\}'$ gemeinsam. Diese Blätter können laut Abschnitt 5.2.4 aus einem der Bündel entfernt werden. Wenn diese Blätter aus den Bündeln $\{(A \sqcup B)^{-\infty} B \circ A\}'$ bzw. $\{(A \sqcup B)^{-\infty} B \circ B\}'$ entfernt werden, dann kommt man letztlich wieder zurück zu der Partition in Abb. 5.4. Wenn man sich aber entscheidet, diese Blätter aus dem Bündel $\{(A \sqcup B)^{-\infty} AB\circ\}'$ zu entfernen, so wird dieses Bündel leer. In beiden Fällen werden gleich viele Blätter entfernt, so daß man sich frei für einen Fall entscheiden kann. Im letzteren Fall kann die Methode von Abschnitt 5.2.4 noch ein weiteres Mal angewendet werden: Aus den beiden Bündeln $\{(A \sqcup B)^{-\infty} B \circ A\}'$ und $\{(A \sqcup B)^{-\infty} B \circ B\}'$ werden alle Blätter entfernt, die in Blättern des Bündels $\{(A \sqcup B)^{-\infty} BB\circ\}'$ enthalten sind. Als Endergebnis erhält man die beiden neuen expandierenden Bündel $\{(A \sqcup B)^{-\infty} AB \circ A\}'$ und $\{(A \sqcup B)^{-\infty} AB \circ B\}'$. Diese kleineren expandierenden Blätter überdecken die Gebiete $AB \circ A$ und $AB \circ B$. Wegen der grammatikalischen Regeln $BAB \circ B = AAB \circ A = \emptyset$ sind dies dieselben Gebiete $BAB\circ$ und $AAB\circ$, die weiter oben auf dem direkten Wege konstruiert wurden.

5.4.6 2. Variante: Zeitliches Verschieben von mehreren Bündeln

Der zweite Vereinfachungsschritt ist auch dann möglich, wenn alle expandierenden und kontrahierenden Bündel Orbitpaare bilden. Denn auch expandierende Bündel, die Orbitpaare bilden, können um einen Zeitschritt vor verschoben werden, aber nur zusammen mit all den kontrahierenden Bündeln, mit denen sie Orbitpaare bilden.

Auswählen und Verschieben der Bündel

Zunächst muß man die Menge der m expandierenden Bündel

$$\Gamma_k = \{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} A_k\circ\}' \tag{5.47}$$

unterteilen in die p expandierenden Bündel Γ_i^V, die um einen Zeitschritt vor verschoben werden sollen, und die $q = m-p$ expandierenden Bündel Γ_i^W, die unverändert bleiben. Wie man geeignete Bündel auswählt, wird später diskutiert.

Dann unterteilt man auch die n kontrahierenden Bündel

$$\Lambda_k = \{\circ B_k(\bigsqcup_{i=1}^{n} B_i)^{+\infty}\}' \tag{5.48}$$

in zwei Mengen. Mit Λ_i^V $(1 \le i \le r)$ bezeichnet man die Bündel, die mindestens ein Orbitpaar mit einem der expandierenden Bündel Γ_j^V bilden. Mit Λ_i^W $(1 \le i \le s)$ bezeichnet man die Bündel, die keine Orbitpaare mit den Bündeln Γ_j^V bilden $(r+s=n)$.

Wenn man die expandierenden Bündel Γ_i^V einen Zeitschritt vor verschieben würde, ohne die kontrahierenden Bündel Λ_i^V, mit denen sie Orbitpaare bilden, zu entfernen, dann würde man

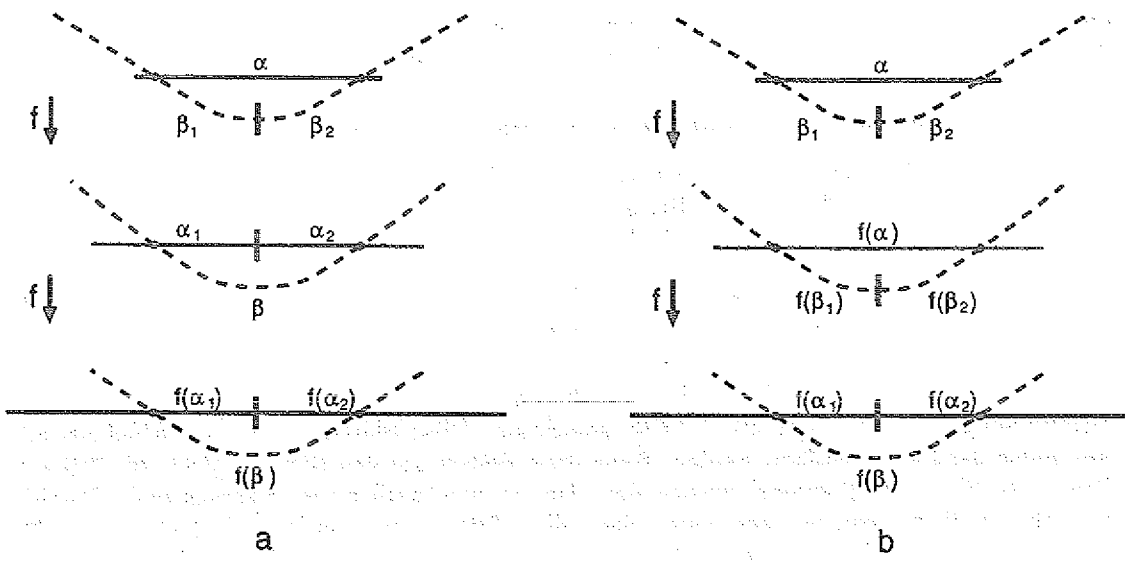


Abbildung 5.10: Zeitliches Vorverschieben eines expandierenden Blattes α und eines kontrahierenden Blattes β. Das Schema ist genauso aufgebaut wie in Abb. 5.8. Abb. a zeigt die Blätter bevor und Abb. b die Blätter nachdem α und β um einen Zeitschritt vor verschoben wurden. Nur der mittlere der drei Zeitschritte ändert sich: Das Blatt $f(\alpha)$ wird expandierend, das Blatt β zerfällt in die kleineren kontrahierenden Blätter $f(\beta_1)$ und $f(\beta_2)$.

doppelte Schnitte erzeugen und die Bed. 3.2.1 für die Eigenschaft „generierend" verletzen. Diese kontrahierenden Bündel Λ_i^V müssen daher auch einen Zeitschritt vor verschoben werden. Abb. 5.10 zeigt die gemeinsame Verschiebung von einem expandierenden und einem kontrahierenden Blatt.

Die neuen Orbitpaare

Was geschieht bei diesem Vereinfachungsschritt mit den Orbitpaaren? Vor dem Vereinfachungsschritt werden alle Orbitpaare von den expandierenden Bündeln Γ_i^V und Γ_i^W mit den kontrahierenden Bündeln Λ_i^V und Λ_i^W gebildet. Diese jeweils zwei Typen von Bündeln können in vier Kombinationen gepaart werden. Da die Bündel Γ_i^V nach Voraussetzung keine Orbitpaare mit den Bündeln Λ_i^W bilden, brauchen nur drei Kombinationen betrachtet zu werden.

Nach dem Vereinfachungsschritt existieren die expandierenden Bündel $f(\Gamma_i^V)$ und Γ_i^W, die kontrahierenden Bündel $f(\Lambda_i^V)$ und Λ_i^W und möglicherweise noch andere Bündel, die nötig sind, um Lücken im Phasenraum zu schließen. Letztere Bündel enthalten nur Urbilder von expandierenden Blättern und Abbilder von kontrahierenden Blättern. Wie bereits erwähnt, können solche Blätter keine Orbitpaare bilden. Nach dem Vereinfachungsschritt werden daher alle Orbitpaare von den expandierenden Bündeln $f(\Gamma_i^V)$ und Γ_i^W mit den kontrahierenden Bündeln $f(\Lambda_i^V)$ und Λ_i^W gebildet. Von den vier möglichen Kombinationen brauchen wieder nur drei betrachtet zu werden, da Γ_i^W nur dann Orbitpaare mit $f(\Lambda_i^V)$ bilden könnte, wenn die Partition vor dem Vereinfachungsschritt nicht generierend gewesen wäre.

Die Orbitpaare wurden in Def. 4.1.1 (von Kap. 4.1.2) so definiert, daß Urbilder und Abbilder eines Orbitpaares auch wieder Orbitpaare sind. Die Orbitpaare, die von den verschobenen Bündeln $f(\Gamma_i^V)$ und $f(\Lambda_i^V)$ gebildet werden, sind gerade die Bilder der Orbitpaare, die von den ursprünglichen Bündeln Γ_i^V und Λ_i^V gebildet wurden. Von diesen Orbitpaaren gehen also keine verloren, es werden aber auch keine neuen gebildet. Ebenso bleiben die Orbitpaare unverändert, die von den expandierenden Bündeln Γ_i^W mit den kontrahierenden Bündeln Λ_i^W gebildet werden.

Somit bleiben vor und nach dem Vereinfachungsschritt nur noch je eine Kombination von Bündeln übrig, die auf ihre Orbitpaare hin untersucht werden muß:

- Die zeitlich verschobenen expandierenden Bündel $f(\Gamma_i^V)$ können Orbitpaare mit den unveränderten kontrahierenden Bündeln Λ_i^W bilden. Von diesen Orbitpaaren hängt das weitere Vorgehen ab, das genauso abläuft wie bei dem Verschieben eines einzelnen expandierenden Bündels in Kap. 5.4.4. Insbesondere kann man, wenn keine derartigen Orbitpaare gebildet werden, dieselben Bündel $f(\Lambda_i^V)$ und $f(\Gamma_i^V)$ noch einen weiteren Zeitschritt vor verschieben.

- Umgekehrt können auch Orbitpaare verlorengehen, nämlich diejenigen, die von den unveränderten expandierenden Bündeln Γ_i^W mit den zu verschiebenden kontrahierenden Bündeln Λ_i^V gebildet werden. Dieser Fall muß vermieden werden, da die zwiefache Partition sonst nicht einfacher werden würde.

Auswahlkriterium

Die Auswahl der zu verschiebenden expandierenden Bündel Γ_i^V muß daher so getroffen werden, daß alle Orbitpaare, die von den anderen expandierenden Bündeln Γ_i^W mit den kontrahierenden Bündeln Λ_i^V gebildet werden, auch von Γ_i^W mit Λ_i^W oder von Γ_i^V mit Λ_i^V gebildet werden. In dem besonderen Fall, daß alle Symbolischen Gebiete disjunkt sind, können verschiedene Bündel nicht dieselben Orbitpaare bilden. In diesem Fall verlangt die Auswahlbedingung einfach, daß die expandierenden Bündel Γ_i^W keine Orbitpaare mit den kontrahierenden Bündeln Λ_i^V bilden dürfen. Dies ist das Gegenstück zu der obigen Voraussetzung, daß die expandierenden Bündel Γ_i^V keine Orbitpaare mit den kontrahierenden Bündeln Λ_i^W bilden dürfen.

Wie findet man Bündel, die dieser Auswahlbedingung genügen? Am besten mit Hilfe der Tabelle, in der die Orbitpaare eingetragen sind. Man geht von einem expandierenden Bündel aus, das in der Zeit verschoben werden soll, prüft nach, welche kontrahierenden Bündel dann wegen der Bed. 3.2.1 für generierende Partitionen verschoben werden müssen, folgert, welche anderen expandierenden Bündel dann wegen der Auswahlbedingung verschoben werden müssen, und so weiter. Die Suche bricht erfolglos ab, falls sich am Ende herausstellt, daß alle Bündel verschoben werden müßten. Denn die ganze Partition unter f abzubilden ist eine triviale und überflüssige Operation. Man kann dann eine neue Suche mit einem anderen expandierenden Bündel beginnen, welches in der Zeit verschoben werden soll. Alle expandierenden Bündel auf diese Weise durchzuprobieren klingt aufwendig, ist es aber nicht, da wir von allen schon ausprobierten Bündeln wissen, ob sie sich in der Zeit verschieben lassen, wodurch die Suche vorzeitig beendet werden kann. Wenn alle Symbolischen Gebiete disjunkt sind, dann ist das Verfahren analog zu dem, welches eine quadratische Matrix durch Permutation ihrer Zeilen oder Spalten auf Blockstruktur bringt.

Auch um einen Zeitschritt zurück können expandierende Bündel Γ_i und kontrahierende Bündel Λ_j verschoben werden. Das Resultat ist dasselbe, als würde man erst alle anderen Bündel um einen Zeitschritt vor verschieben, und dann die ganze zwiefache Partition durch ihr Urbild f^{-1} ersetzen.

5.4.7 Zeitliches Verschieben mehrerer Bündel der AB–Partition

Verschieben der Bündel

In diesem Beispiel sind die Symbolischen Gebiete disjunkt. Deshalb müssen verschiedene Bündel auch verschiedene Orbitpaare bilden. Um Bündel zu finden, die sich gemeinsam in der Zeit verschieben lassen, braucht man daher nur zu wissen, welche Bündel überhaupt Orbitpaare bilden. Aus der Suche in Tab. 5.9 ergibt sich eine Möglichkeit, Bündel gemeinsam zu verschieben, und zwar die folgende:

In der Spalte des kontrahierenden Bündels $\{\circ BB(A \sqcup B)^{+\infty}\}$' gibt es nur einen Eintrag, der nicht die leere Menge $\emptyset$ ist. Dieses Bündel bildet also nur mit dem expandierenden Bündel $\{(A \sqcup B)^{-\infty} AA\circ\}$' Orbitpaare. Beide Bündel können daher gemeinsam um einen Zeitschritt zurück verschoben werden, ohne daß die Bedingung 3.2.1 für die Eigenschaft „generierend" verletzt wird. Umgekehrt bildet auch das expandierende Bündel $\{(A \sqcup B)^{-\infty} AA\circ\}$' nur mit dem kontrahierenden Bündel $\{\circ BB(A \sqcup B)^{+\infty}\}$' Orbitpaare. Deshalb gehen bei der zeitlichen Verschiebung keine Orbitpaare verloren.

Zurückverschieben um einen Zeitschritt ergibt das neue kontrahierende Bündel $\{\circ EBB(A \sqcup B)^{+\infty}\}$' und entfernt das alte expandierende Bündel $\{(A \sqcup B)^{-\infty} AA\circ\}$'. Eine kurze Rechnung bestätigt, daß das neue kontrahierende Bündel tatsächlich neue Orbitpaare mit den unveränderten expandierenden Bündeln bildet (Tab. 5.13). Daher wird die Partition gemäß dem Kriterium 4.1.2 einfacher, obwohl die Zahl der Bündel nicht kleiner wird. Die Verschiebung derselben beiden Bündel in die andere Zeitrichtung ist auch möglich und führt im Endergebnis zu einer anderen generierenden Partition. Dies wird in Kap. 7.1.2 diskutiert werden.

Die neuen expandierenden Bündel

Diese veränderten expandierenden und kontrahierenden Bündel können mit den Methoden von Kap. 5.2 auch durch Partitionen ausgedrückt werden. Betrachten wir zunächst die expandierenden Bündel: Das Bündel $\{(A \sqcup B)^{-\infty} AA\circ\}$' ist nicht mehr expandierend. Die größten expandierenden Blätter, welche die Lücke $AA\circ$ ausfüllen, sind die der Bündel

$$f^{-1}(\{(A \sqcup B)^{-\infty} AA\circ\}') = \{(A \sqcup B)^{-\infty} AA \circ A\}' \cup \{(A \sqcup B)^{-\infty} BA \circ A\}' \quad \text{und} \quad (5.49)$$

$$f^{-1}(\{(A \sqcup B)^{-\infty} AB\circ\}') = \{(A \sqcup B)^{-\infty} AA \circ B\}' \cup \{(A \sqcup B)^{-\infty} BA \circ B\}' \; . \quad (5.50)$$

Die anderen expandierenden Bündel $\{(A \sqcup B)^{-\infty} AAB\circ\}'$, $\{(A \sqcup B)^{-\infty} BAB\circ\}'$, $\{(A \sqcup B)^{-\infty} BA\circ\}'$ und $\{(A \sqcup B)^{-\infty} BB\circ\}'$ aus Abb. 5.5 bleiben unverändert. Die expandierenden Blätter in $\{(A \sqcup B)^{-\infty} BA\circ\}'$ sind mindestens so groß wie die in $\{(A \sqcup B)^{-\infty} BA \circ A\}$ und $\{(A \sqcup B)^{-\infty} BA \circ B\}$. Deshalb können die letzteren Blätter nach der Methode von Kap. 5.2.4 entfernt werden. Übrig bleiben die kleineren expandierenden Bündel $\{(A \sqcup B)^{-\infty} AA \circ A\}'$ und $\{(A \sqcup B)^{-\infty} AA \circ B\}'$. Demnach ist $\{AA \circ A, AA \circ B, AAB\circ, BAB\circ, BA\circ, BB\circ\}$. die neue Partition, die die expandierenden Bündel erzeugt. Abb. 5.11 zeigt die Übergänge zwischen diesen Symbolischen Gebieten unter der Abbildung f.

Die neuen kontrahierenden Bündel

Was geschieht mit den kontrahierenden Bündeln? Hier ist ein neues kontrahierendes Bündel $\{\circ EBB(A \sqcup B)^{+\infty}\}$' entstanden. Dessen Blätter nehmen den Platz der kleineren Blätter in $\{\circ BBB(A \sqcup B)^{+\infty}\}$' und $\{\circ ABB(A \sqcup B)^{+\infty}\}$' ein. Die Methode von Kap. 5.2.4 verkleinert deshalb das kontrahierende Bündel $\{\circ BB(A \sqcup B)^{+\infty}\}$' zu dem Bündel $\{\circ BBA(A \sqcup B)^{+\infty}\}$' und entfernt außerdem alle Blätter des kontrahierenden Bündels $\{\circ ABB(A \sqcup B)^{+\infty}\}$'. Die anderen kontrahierenden Bündel $\{\circ BA(A \sqcup B)^{+\infty}\}$' und $\{\circ AA(A \sqcup B)^{+\infty}\}$' von Abb. 5.7 bleiben unverändert. Daher ist $\{\circ AA, \circ BA, \circ ABA, \circ BBA, \circ EBB\}$ die neue Partition, die die Zukunft beschreibt.

Abb. 5.12 zeigt die Übergänge zwischen diesen Symbolischen Gebieten unter der inversen Abbildung f^{-1}. Und Tab. 5.13 zeigt, welche expandierenden und kontrahierenden Bündel Orbitpaare bilden. In der Tat bildet das neue kontrahierende Bündel $\{\circ EBB(A \sqcup B)^{+\infty}\}$' neue Orbitpaare mit den unveränderten expandierenden Bündeln $\{(A \sqcup B)^{-\infty} BA\circ\}'$ und $\{(A \sqcup B)^{-\infty} BB\circ\}'$. Die zwiefache Partition ist also nach Kriterium 4.1.2 tatsächlich einfacher geworden.

5.4.8 3. Variante: Zeitliches Verschieben von Teilblättern

Warum müssen Teilblätter betrachtet werden?

Bisher beruhte jede Änderung einer zwiefachen Partition auf dem Verschieben von ganzen expandierenden oder kontrahierenden Blättern. Wie schon mehrfach erwähnt, kann aber auch jede Teilmenge eines expandierenden Blattes selbst als expandierendes Blatt aufgefaßt werden. Im folgenden werden auch solche Teilblätter verschoben werden, aber nur ganz bestimmte. Denn alle möglichen Teilblätter der vorhandenen expandierenden Blätter zu betrachten wäre praktisch unmöglich.

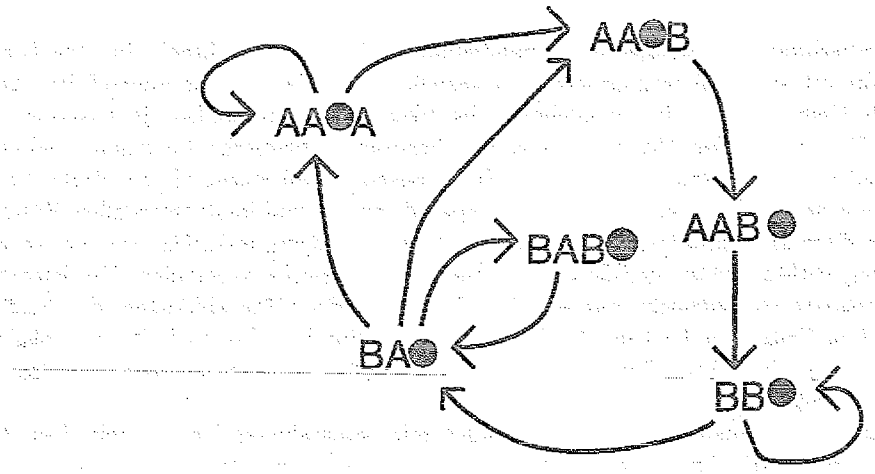


Abbildung 5.11: Die erste Vereinfachung der Partition, die die expandierenden Bündel beschreibt. Die expandierenden Blätter des Bündels $\{(A \sqcup B)^{-\infty} AA\circ\}$' wurden in die kleineren expandierenden Blätter der Bündel $\{(A \sqcup B)^{-\infty} AA \circ A\}$' und $\{(A \sqcup B)^{-\infty} AA \circ B\}$' zerlegt. Die anderen Bündel aus Abb. 5.5 blieben unverändert. Die Bezeichnungen wurden in Abb. 5.4 erklärt.

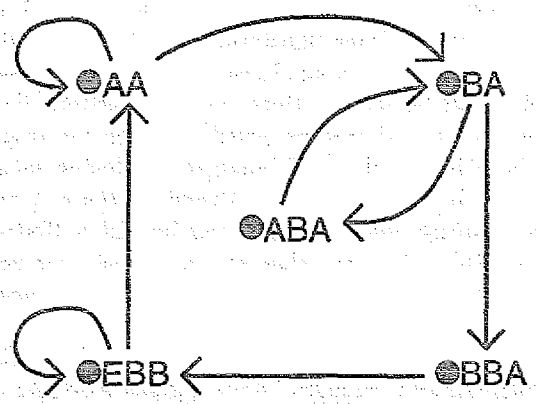


Abbildung 5.12: Die erste Vereinfachung der Partition, die die kontrahierenden Bündel beschreibt. Die Blätter des neuen kontrahierenden Bündels $\{\circ EBB(A \sqcup B)^{+\infty}\}$' nahmen den Platz der kleineren kontrahierenden Blätter in $\{\circ ABB(A \sqcup B)^{+\infty}\}$' und $\{\circ BBB(A \sqcup B)^{+\infty}\}$' ein. Die anderen kontrahierenden Blätter aus Abb. 5.7 blieben unverändert. Die Bezeichnungen wurden in Abb. 5.6 erklärt.

	$\circ AA$	$\circ BA$	$\circ EBB$
$AA \circ A$	$\emptyset$	$\emptyset$	$\star$
$BA\circ$	$\star$	$\emptyset$	$\star$
$BB\circ$	$\star$	$\star$	$\star$

Abbildung 5.13: Die expandierenden und kontrahierenden Bündel nach dem ersten Vereinfachungsschritt und die durch sie gebildeten Orbitpaare. Die zwiefache Partition ist diejenige von Abb. 5.11 und Abb. 5.12. Nur die Bündel sind eingetragen, deren Blätter beim nächsten Zeitschritt geteilt werden, da die anderen Bündel keine Orbitpaare bilden. Die Bezeichnungen sind wie in Tab. 5.3.

Motivation

Warum muß man überhaupt Teilmengen der expandierenden Blätter betrachten? Um eine Forderung zu erfüllen, die oft an Optimierungsverfahren gestellt wird. Die Optimierungsschritte kann man sich meist als kleine Schritte in irgendeinem abstrakten Raum vorstellen. In unserem Fall entspräche jedem Punkt des abstrakten Raumes eine bestimmte zwiefache Partition. Zwischen diesen Punkten sind nur bestimmte Schritte möglich. In unserem Fall wurde die zwiefache Partition bisher nur durch das zeitliche Verschieben von expandierenden und kontrahierenden Bündeln abgeändert. Durch diese Einschränkung kann eine bestimmte Richtung festgelegt werden, in welche die Optimierung verläuft. Aber gerade das will man normalerweise vermeiden. Die Richtung der Optimierungsschritte soll vielmehr nur von dem Kriterium der Güte abhängen, das heißt in unserem Fall von dem Kriterium 4.1.2 der Einfachheit. Wenn man von diesem Kriterium absieht, dann soll überall in dem abstrakten Raum, wo ein Schritt möglich ist, auch der inverse Schritt in die umgekehrte Richtung möglich sein.

Erfüllt das zeitliche Verschieben von expandierenden oder kontrahierenden Bündeln diese Forderung? Auf den ersten Blick scheint ein solcher Schritt reversibel zu sein. Denn wenn man bestimmte Bündel erst in der Zeit vor und anschließend in der Zeit zurück verschiebt, dann erhält man die ursprünglichen expandierenden und kontrahierenden Bündel zurück. Dies gilt aber nur, sofern keine expandierenden und kontrahierenden Blätter als Teile von größeren expandierenden und kontrahierenden Blättern entfernt werden. Betrachten wir als Beispiel den Schritt, mit dem in Kap. 5.4.7 die AB–Partition vereinfacht wurde. Abb. 5.14.a zeigt einige expandierende Blätter vor dem Vereinfachungsschritt und Abb. 5.14.b. dieselben Blätter nach dem Vereinfachungsschritt. Das Blatt $\ldots BAA\circledast$, das um einen Zeitschritt zurück verschoben wird, gehört nur vor diesem Vereinfachungsschritt zu den expandierenden Blättern. Aber auch sein Urbild $\ldots BA \circledast A$ erscheint in Abb. 5.14.b nicht mehr explizit als expandierendes Blatt, da es als Teil des größeren expandierenden Blattes $\ldots BA\circledast$ entfernt wurde. Doch nur die Blätter, die explizit als expandierend deklariert wurden, und nicht ihre Teilmengen wurden in den vorausgegangenen Abschnitten in der Zeit verschoben. Deshalb ist es mit den bisherigen Methoden nicht möglich, ausgehend von den expandierenden Blättern in Abb. 5.14.b das Urbild $\ldots BA \circledast A$ um einen Zeitschritt vor zu verschieben. Dies wäre aber nötig, um von den expandierenden Blättern in Abb. 5.14.b zu den expandierenden Blättern in Abb. 5.14.a zu gelangen. Der Schritt von der zwiefachen Partition von Abb. 5.14.a zu der zwiefachen Partition von Abb. 5.14.b ist daher nach den bisherigen Methoden irreversibel.

Bei genauerer Betrachtung zeigt sich sogar, daß die bisherigen Methoden nur solche expandierenden bzw. kontrahierenden Blätter erzeugen können, deren Abbilder oder Urbilder schon zu den expandierenden bzw. kontrahierenden Blättern der anfänglichen zwiefachen Partition gehörten. Diese starke Einschränkung ist der tiefere Grund dafür, daß manche zeitlichen Verschiebungen von Bündeln irreversibel sind. Diese Einschränkung soll nun dadurch überwunden werden, daß auch bestimmte Teilmengen von expandierenden und kontrahierenden Blättern in der Zeit verschoben werden.

Konstruktion der zu verschiebenden Bündel

Man könnte das zeitliche Verschieben von expandierenden und kontrahierenden Bündeln reversibel machen, indem man jede Teilmenge eines expandierenden Blattes auch als expandierend deklariert und prüft, ob sie zeitlich verschoben werden kann. Alle solchen Teilblätter zu betrachten ist aber praktisch unmöglich. Zum Glück genügt es, nur ganz bestimmte Teilblätter zu betrachten.

Wie man in Abb. 5.14.b sieht, verschwindet das expandierende Blatt $\ldots BAA\circledast$ beim Übergang von Abb. 5.14.a zu Abb. 5.14.b nicht spurlos. Seine Grenzen werden zu Grenzen der kleineren expandierenden Blätter $\ldots BAA \circledast A$ und $\ldots BAA \circledast B$. Auch wenn man die anfänglichen expandierenden Blätter in Abb. 5.14.a nicht mehr kennt, kann man diese Grenzen als Anhaltspunkt nehmen, um das benötigte Teilblatt $\ldots BA \circledast A$ zu rekonstruieren. Um von der zwiefachen Partition von Abb. 5.14.b zurück zu der zwiefachen Partition von Abb. 5.14.a zu gelangen, muß man ein ganzes Bündel $\{(A \sqcup B)^{-\infty} BA \circledast A\}$‘ solcher Teilblätter um einen Zeitschritt vor verschieben.

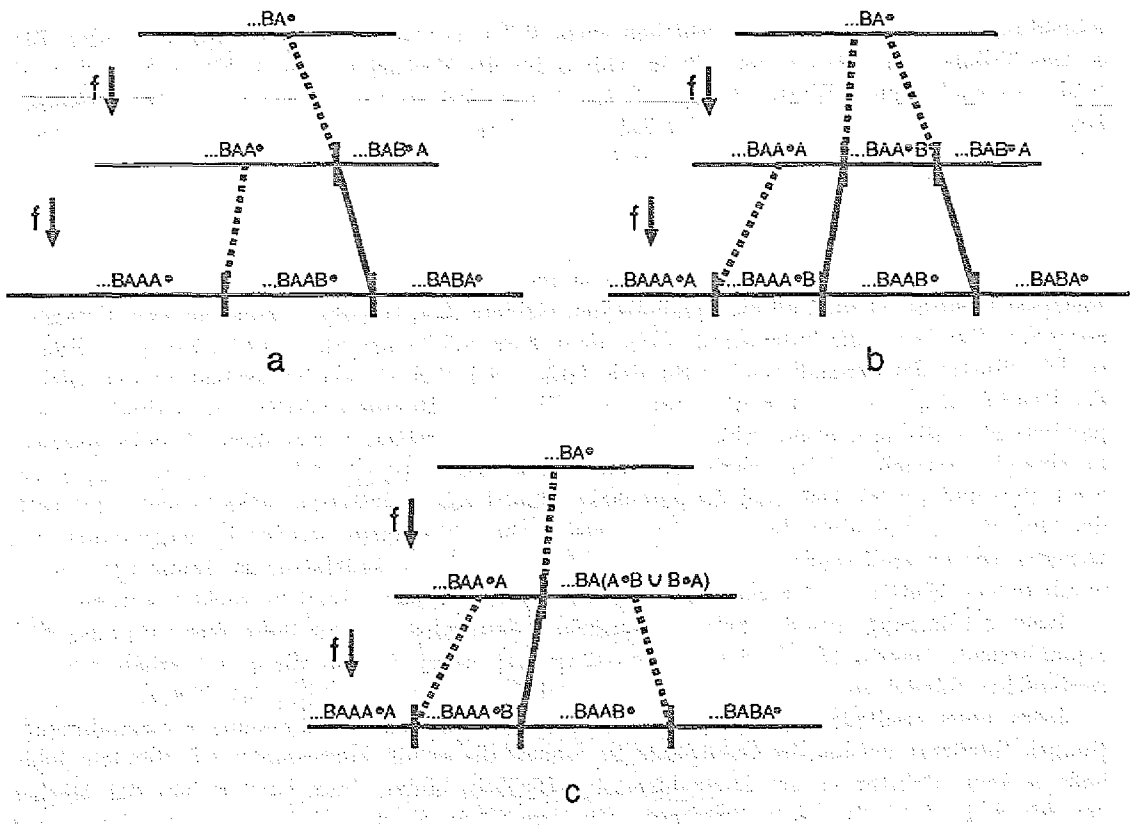


Abbildung 5.14: Übersicht über die Rechnung zur Vereinfachung der AB–Partition. Jedes der drei Schemata zeigt ein paar Blätter, die bei einem bestimmten Stand der Beispielrechnung expandierend sind. Das Schema in Abb. a zeigt expandierende Blätter der ursprünglichen Partition (aus Abb. 5.4 oder 5.5). Die beiden anderen Schemata zeigen die entsprechenden expandierenden Bätter nach dem ersten Vereinfachungsschritt (Abb b, siehe auch Abb. 5.11) und nach dem zweiten und letzten Vereinfachungsschritt (Abb c, siehe auch Abb. 5.16). Das Blatt $\ldots BA\circ$ gehört in allen drei Fällen zu den expandierenden Blättern. Seine beiden nächsten Bilder sind als waagrechte Linien eingetragen. Dicke Linien unterteilen diese Bilder in expandierende Blätter. Gestrichelte Linien zeigen Zeit und Ort einer neuen Teilung.

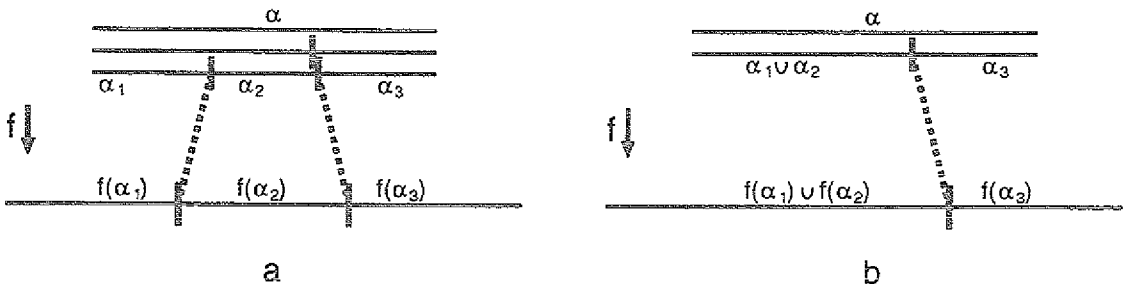


Abbildung 5.15: Zeitliches Vorverschieben eines Teiles $\alpha_1 \cup \alpha_2$ von einem expandierenden Blatt α. Das Teilblatt $\alpha_1 \cup \alpha_2 \subset \alpha$ entsteht in Abb. a aus der Vereinigung der Urbilder α_1 und α_2 der beiden expandierenden Blätter $f(\alpha_1)$ und $f(\alpha_2)$. Der Text erläutert, warum derartige Teilblätter konstruiert und, wenn möglich, in der Zeit verschoben werden sollen. Wenn man das expandierende Teilblatt $\alpha_1 \cup \alpha_2$ um einen Zeitschritt vor verschiebt, dann wird das Blatt $f(\alpha_1) \cup f(\alpha_2)$ expandierend (Abb. b).

In allgemeinen Fall soll man die Bündel von Teilblättern nach folgender Methode konstruieren: Zunächst identifiziert man all die Symbolischen Gebiete $A_i\circledcirc$, von denen mehr als zwei Übergänge ausgehen. Das heißt, die Indexmenge $J(i)$, die in Kap. 5.3.2 eingeführt wurde, hat $q > 2$ Elemente. Die Blätter des expandierenden Bündels $\{(\bigsqcup_{l=1}^{m} A_l)^{-\infty} A_i\circledcirc\}$' werden zerteilt zu den Blättern der Bündel $\{(\bigsqcup_{l=1}^{m} A_l)^{-\infty} A_i \circledcirc A_j\}$' (mit $j \in J(i)$). Diese Bündel enthalten die Urbilder von expandierenden Blättern (siehe Abb. 5.15). Wenn man die Blätter zweier dieser Bündel paarweise miteinander vereinigt, dann erhält man Bündel der Form $\{(\bigsqcup_{l=1}^{m} A_l)^{-\infty} A_i \circledcirc (A_j \cup A_k)\}$' (mit $j, k \in J(i)$ und $j \neq k$). Dies sind die gesuchten Bündel von Teilblättern. Man beachte, daß rechts der Operator $\cup$ und nicht der Operator $\sqcup$ steht. Diese Teilblätter werden im allgemeinen nicht Urbilder von expandierenden Blättern sein. Aber jedes dieser Teilblätter ist Teilmenge eines expandierenden Blattes $\ldots A_i\circledcirc$ und kann deshalb selbst als expandierend betrachtet werden.

Wenn q Übergänge von der Ecke $A_i\circledcirc$ ausgehen, dann wird man auf diese Weise $\binom{q}{2}$ zusätzliche expandierende Bündel $\{(\bigsqcup_{l=1}^{m} A_l)^{-\infty} A_i \circledcirc (A_j \cup A_k)\}$' erhalten. Nur für $q = 2$ erhält man kein zusätzliches Bündel, sondern bloß das bereits vorhandene Bündel $\{(\bigsqcup_{l=1}^{m} A_l)^{-\infty} A_i\circledcirc\}$'.

Jedes dieser zusätzlichen Bündel wird dann ebenso behandelt wie die anderen expandierenden Bündel. Zunächst werden die Orbitpaare berechnet, die es mit kontrahierenden Bündeln bildet. Falls es keine Orbitpaare mit kontrahierenden Bündeln bildet, dann kann es mit der Methode von Abschnitt 5.4.4 um einen Zeitschritt vor verschoben werden. Falls es nur mit bestimmten kontrahierenden Bündeln Orbitpaare bildet, dann können sie wie in Abschnitt 5.4.6 gemeinsam um einen Zeitschritt vor verschoben werden, sofern dabei keine Orbitpaare verlorengehen. Und falls sich beide Methoden nicht anwenden lassen, dann kann man dieses zusätzliche Bündel mit der Methode von Kap. 5.2.4 einfach wieder entfernen.

Kontrahierende Bündel von Teilblättern werden auf eine analoge Weise konstruiert. Wenn von einem Symbolischen Gebiet $\circledcirc B_i$ unter der inversen Dynamik f^{-1} Übergänge zu den Symbolischen Gebieten $\circledcirc B_j$, $\circledcirc B_k$ etc. führen, dann entsteht das Bündel $\{(B_j \cup B_k) \circledcirc B_i(\bigsqcup_{l=1}^{n} B_l)^{+\infty}\}$' durch die paarweise Vereinigung von Teilblättern. Auch für dieses zusätzliche kontrahierende Bündel sollte man testen, ob es nach den Methoden von Abschnitt 5.4.4 und 5.4.6 um einen Zeitschritt rückverschoben werden kann.

Anmerkung

Es wurde noch nicht gezeigt, daß durch die so konstruierten Bündel von Teilblättern das zeitliche Verschieben von Bündeln tatsächlich umgekehrt werden kann. Eine genaue Betrachtung würde zeigen, daß man mit dieser zusätzlichen Methode tatsächlich immer zurück zu den ursprünglichen expandierenden und kontrahierenden Blättern kommen kann. Die so erhaltenen expandierenden und kontrahierenden Bündel wären nicht unbedingt dieselben wie die Bündel der anfänglichen Partition, aber auf jeden Fall nicht größer als diese. In diesem Sinne ist jetzt jeder Vereinfachungs-

schritt reversibel.

Von der genauen Betrachtung will ich nur einen zentralen Argumentationschritt darstellen. Er ist wichtig, wenn $q \geq 4$ Übergänge von einer Ecke ausgehen. Nach der obigen Methode werden nur zusätzliche Bündel der Form $\{(\bigsqcup_{h=1}^{m} A_h)^{-\infty} A_i \circledast (A_j \cup A_k)\}$' konstruiert. Nun mag sich der Leser gefragt haben, warum man nicht auch ein Bündel $\{(\bigsqcup_{h=1}^{m} A_h)^{-\infty} A_i \circledast (A_j \cup A_k \cup A_h)\}$' konstruiert, dessen Blätter aus der Vereinigung von drei (oder mehr) Urbildern expandierender Blätter entstehen.

Es stellt sich als überflüssig heraus, ein solches Bündel $\{(\bigsqcup_{h=1}^{m} A_h)^{-\infty} A_i \circledast (A_j \cup A_k \cup A_h)\}$' explizit zu betrachten. Denn es läßt sich nur dann um einen Zeitschritt vor verschieben, wenn sich auch das Bündel $\{(\bigsqcup_{h=1}^{m} A_h)^{-\infty} A_i \circledast (A_j \cup A_k)\}$' vor verschieben läßt. Letzteres kann nach der obigen Methode konstruiert werden. Wenn man es zeitlich vor verschiebt, dann entsteht das neue Symbolische Gebiet $f(A_i \circledast (A_j \cup A_k))$. Wenn man nun dieselbe Methode erneut auf das Bündel $\{(\bigsqcup_{h=1}^{m} A_h)^{-\infty} A_i \circledast\}$' anwendet, dann erhält man (unter anderem) die Teilblätter in $\{(\bigsqcup_{h=1}^{m} A_h)^{-\infty} A_i \circledast (A_j \cup A_k)\}$' und $\{(\bigsqcup_{h=1}^{m} A_h)^{-\infty} A_i \circledast A_l\}$'. Durch paarweise Vereinigung dieser Teilblätter entsteht das gesuchte Bündel $\{(\bigsqcup_{h=1}^{m} A_h)^{-\infty} A_i \circledast (A_j \cup A_k \cup A_i)\}$'. Paarweise Vereinigung von Urbildern expandierender Blätter genügt daher, um auch alle wichtigen Bündel der Form $\{(\bigsqcup_{h=1}^{m} A_h)^{-\infty} A_i \circledast (A_j \cup A_k \cup A_h)\}$' zu konstruieren, sofern die Methode dieses Abschnittes 5.4.8 nur oft genug angewendet wird.

5.4.9 Zeitliches Verschieben von Teilblättern der AB–Partition

Durch die Konstruktion von Teilblättern werden nicht nur die vorherigen Änderungen einer zwiefachen Partition umkehrbar. Es werden auch ganz neue Änderungen möglich. Auch bei der Vereinfachung der AB–Partition wird ein weiterer Vereinfachungsschritt möglich.

Teilblätter können, wie bereits erwähnt, nur dann eine Rolle spielen, wenn von einem Symbolischen Gebiet mehr als zwei Übergänge ausgehen. In unserem Beispiel trifft dies nur auf ein Symbolisches Gebiet zu, nämlich auf $BA\circledast$ in Abb. 5.11. Von ihm gehen drei Übergänge nach $AA\circledast A$, $AA\circledast B$ und $BAB\circledast = BAB\circledast A$ aus. Die entsprechenden Urbilder expandierender Blätter sind in den Bündeln $\{(A \sqcup B)^{-\infty} BA\circledast AA\}$', $\{(A \sqcup B)^{-\infty} BA\circledast AB\}$' und $\{(A \sqcup B)^{-\infty} BA\circledast BA\}$' enthalten. Es gibt $\binom{3}{2} = 3$ verschiedene Weisen, diese drei Bündel der Form $\{(\bigsqcup_{h=1}^{m} A_h)^{-\infty} A_i \circledast A_j\}$' zu den gewünschten Bündeln der Form $\{(\bigsqcup_{h=1}^{m} A_h)^{-\infty} A_i \circledast (A_j \cup A_k)\}$' zu kombinieren. So ergeben sich die drei Bündel von Teilblättern $\{(A \sqcup B)^{-\infty} BA \circledast A\}$', $\{(A \sqcup B)^{-\infty} BA \circledast EA\}$' und $\{(A \sqcup B)^{-\infty} BA \circledast (AB \cup BA)\}$'.

Wie man sieht, ergibt sich durch diese Konstruktion tatsächlich das Bündel $\{(A \sqcup B)^{-\infty} BA \circledast A\}$', das, wie bereits erwähnt, nötig wäre, um den ersten Vereinfachungsschritt umzukehren. Jetzt sind wir aber an weiteren Vereinfachungsschritten interessiert. Dabei können uns dieses Bündel und das zweite Bündel $\{(A \sqcup B)^{-\infty} BA \circledast EA\}$' nicht helfen. Denn die genauere Betrachtung zeigt, daß beide Bündel Orbitpaare bilden, und zwar mit den kontrahierenden Bündeln $\{\circledast EBB(A \sqcup B)^{+\infty}\}$' bzw. $\{\circledast AA(A \sqcup B)^{+\infty}\}$'. Sie können deshalb nicht ohne diese kontrahierenden Bündel in der Zeit verschoben werden. Und wenn man versucht, sie mit diesen kontrahierenden Bündeln in der Zeit zu verschieben, dann stellt man fest, daß Orbitpaare verlorengehen, die zwiefache Partition also nicht einfacher wird.

Dagegen bildet das dritte Bündel $\{(A \sqcup B)^{-\infty} BA \circledast (AB \cup BA)\}$' keine Orbitpaare. Es kann ohne weiteres mit der Methode von Kap. 5.4.4 um einen Zeitschritt vor verschoben werden. Dabei wird das Bündel $\{(A \sqcup B)^{-\infty} BA(A \circledast B \cup B \circledast A)\}$' als neues expandierendes Bündel deklariert.

Das neue expandierende Bündel bildet auch neue Orbitpaare mit kontrahierenden Bündeln (Tab. 5.17). Daher ist die zwiefache Partition nach Kriterium 4.1.2 einfacher geworden, obwohl die Zahl der expandierenden Bündel nicht kleiner geworden ist. Mit den Methoden von Kap. 5.2 kann die Partition gefunden werden, welche die neuen expandierenden Bündel erzeugt. Abb. 5.16 zeigt diese Partition und ihre Übergänge. Das neue expandierende Bündel $\{(A \sqcup B)^{-\infty} BA(A \circledast B \cup B \circledast A)\}$' wird durch das neue Symbolische Gebiet $BAA \circledast B \cup BAB \circledast A$ erzeugt. Und weil die Blätter in dem neuen expandierenden Bündel größer sind, konnte das alte expandierende Bündel $\{(A \sqcup B)^{-\infty} BAB \circledast A\}$' entfernt und das alte expandierende Bündel $\{(A \sqcup B)^{-\infty} AA \circledast B\}$' zu $\{(A \sqcup B)^{-\infty} AAA \circledast B\}$' verkleinert werden.

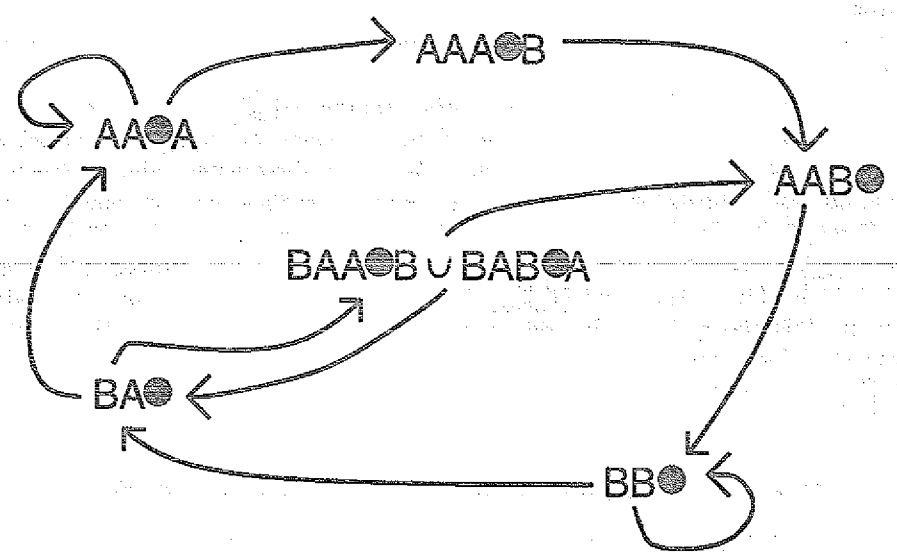


Abbildung 5.16: Die zweite Vereinfachung der Partition, die die expandierenden Bündel beschreibt. Die expandierenden Blätter der Bündel $\{(A \sqcup B)^{-\infty} BAB \circ A\}$' und $\{(A \sqcup B)^{-\infty} BAA \circ B\}$' wurden paarweise zu den Blättern des neuen expandierenden Bündels $\{(A \sqcup B)^{-\infty}(BAB \circ A \cup BAA \circ B)\}$' vereinigt. Die anderen expandierenden Blätter aus Abb. 5.11 blieben unverändert. Die Bezeichnungen wurden in Abb. 5.4 erklärt.

	$\circ AA$	$\circ BA$	$\circ EBB$
$AA \circ A$	$\emptyset$	$\emptyset$	$\star$
$BA\circ$	$\star$	$\emptyset$	$\star$
$BAA \circ B \cup BAB \circ A$	$\emptyset$	$\star$	$\star$
$BB\circ$	$\star$	$\star$	$\star$

Abbildung 5.17: Die expandierenden und kontrahierenden Bündel nach dem zweiten und letzten Vereinfachungsschritt und die durch sie gebildeten Orbitpaare. Nur die Bündel aus Abb. 5.16 und Abb. 5.12 sind eingetragen, die beim nächsten Zeitschritt geteilt werden, da die anderen Bündel keine Orbitpaare bilden. Die Bezeichungen sind wie in Tab. 5.3.

Zusammenfassung

Die Abb. 5.14 faßt die Vereinfachung der AB–Partition am Beispiel eines einzelnen Blattes zusammen. Man erkennt, daß sich durch die Vereinfachungsschritte der Zeitpunkt ändert, zu dem eine Grenze zwischen expandierenden Blättern zum ersten Mal auftaucht. Die Teilung eines expandierenden Blattes wird also nur in der Zeit verschoben und nicht ganz entfernt.

Im allgemeinen Fall können aber auch Teilungen auftreten, die zu allen Zeiten überflüssig sind. Dies trifft zum Beispiel auf den Ausschnitt der VW–Partitionsgrenze zu, der in Abb. 2.10 bei P und $f(P)$ liegt. Dieser Ausschnitt der Grenze wird beim Übergang zur einfacheren XY–Partition in Abb. 2.12 nicht durch ein Abbild oder Urbild ersetzt, sondern völlig entfernt. Mit der hier vorgeschlagenen Methode könnte man eine solche überflüssige Teilung entfernen, indem man sie um unendlich viele Zeitschritte verschiebt.

Die Vereinfachung der AB–Partition hat damit ein Ende erreicht. Wir erhalten die zwiefache Partition, deren Symbolische Gebiete in Abb. 5.12 und Abb. 5.16 stehen. Die Vergangenheit wird durch die Partition

$$\{A_1\circ, \ldots A_m\circ\} = \{AA\circ A, \quad AAA\circ B, \quad AAB\circ, \quad BB\circ, \quad BA\circ, \quad BAA\circ B \cup BAB\circ A\} \tag{5.51}$$

beschrieben und die Zukunft durch die Partition

$$\{\circ B_1, \ldots \circ B_n\} \quad = \quad \{\circ AA, \quad \circ ABA, \quad \circ BA, \quad \circ BBA, \quad \circ EBB\} \ . \tag{5.52}$$

Man kann sich mit einigen einfachen Rechnungen davon überzeugen, daß diese zwiefache Partition mit den Methoden dieser Arbeit nicht weiter vereinfacht werden kann. Denn jedes weitere zeitliche Verschieben der expandierenden oder kontrahierenden Bündel würde entweder Orbitpaare vernichten oder die Eigenschaft „generierend“ verletzen. Und das zeitliche Verschieben von Teilblättern kommt schon deshalb nicht in Frage, weil von jedem Symbolischen Gebiet jetzt nur noch zwei zeitliche Übergänge ausgehen.

Zurückkehren ist die Bewegung des Weges.
Nachgeben ist die Art des Weges.
Die zehntausend Dinge sind aus dem Etwas geboren.
Etwas ist aus dem Nichts geboren.
Lao Tse: „Tao Te King"; Kap. 40;
übersetzt von Gia-Fu Feng und Gabriela Fontél.

Kapitel 6

Rekonstruktion einer nicht zwiefachen Partition

6.1 Überblick

Ziel dieses Kapitels ist es, zu demonstrieren, durch welche Schritte eine normale (d. h. nicht zwiefache) Partition $\{\circledast\tilde{C}_1, \ldots \circledast\tilde{C}_{\tilde{p}}\}$ vereinfacht werden kann, ohne daß sie ihre Eigenschaft verliert, generierend zu sein.

Der erste Schritt ist trivial: Die Partition wird als zwiefache Partition $\{\tilde{C}_1\circledast, \ldots \tilde{C}_{\tilde{p}}\circledast\}$, $\{\circledast\tilde{C}_1, \ldots \circledast\tilde{C}_{\tilde{p}}\}$ aufgefaßt. Diese zwiefache Partition hat offensichtlich dieselben expandierenden und kontrahierenden Blätter wie die nicht zwiefache Partition $\{\circledast\tilde{C}_1, \ldots \circledast\tilde{C}_{\tilde{p}}\}$.

Der zweite Schritt wurde im vorausgegangenen Kapitel demonstriert. Die zwiefache Partition wird vereinfacht, indem ihre expandierenden und kontrahierenden Bündel zeitlich verschoben werden. Als Ergebnis erhält man eine zwiefache Partition $\{A_1\circledast, \ldots A_m\circledast\}$, $\{\circledast B_1, \ldots \circledast B_n\}$, die nach dem Einfachheitskriterium 4.1.2 (aus Kap. 4.1.2) mindestens so einfach ist wie die ursprüngliche Partition und die nach wie vor die Bedingung 3.2.1 für die Eigenschaft „generierend" (aus Kap. 3.2.3) erfüllt.

Der dritte Schritt ist Inhalt dieses Kapitels. Aus der vereinfachten zwiefachen Partition $\{A_1\circledast, \ldots A_m\circledast\}$, $\{\circledast B_1, \ldots \circledast B_n\}$ soll wieder eine nicht zwiefache Partition $\{\circledast C_1, \ldots \circledast C_p\}$ konstruiert werden, die dieselben expandierenden und kontrahierenden Blätter besitzt. Das Kriterium 4.1.2 der Einfachheit und die Bedingung 3.2.1 für die Eigenschaft „generierend" hängen nur davon ab, welche expandierenden oder kontrahierenden Blätter eine Partition erzeugt. Daher ist die rekonstruierte Partition $\{\circledast C_1, \ldots \circledast C_p\}$ mindestens so einfach wie die ursprüngliche Partition $\{\circledast\tilde{C}_1, \ldots \circledast\tilde{C}_{\tilde{p}}\}$ und nach wie vor generierend.

Damit die nicht zwiefache Partition $\{\circledast C_1, \ldots \circledast C_p\}$ dieselben expandierenden und kontrahierenden Blätter besitzt wie die zwiefache Partition $\{A_1\circledast, \ldots A_m\circledast\}$, $\{\circledast B_1, \ldots \circledast B_n\}$, muß sie mehrere Bedingungen erfüllen. Die beiden wichtigsten Bedingungen stehen in Gl. 6.3 und 6.4. Leider weiß man von vorneherein nicht, welche Zahl p der Symbole genügt, um die beiden Bedingungen zu erfüllen. Die Rekonstruktion beruht daher auf Versuch und Irrtum: Man setzt eine möglichst geringe Zahl p der Symbole an und versucht, damit eine nicht zwiefache Partition zu rekonstruieren. Wenn dies mißlingt, erhöht man die Zahl p der Symbole um eins und startet einen neuen Versuch.

Als Beispiel wird die zwiefache Partition betrachtet, die sich am Ende des letzten Kapitels aus der Vereinfachung der AB-Partition ergab. Es gelingt in Abschnitt 6.3, aus ihr eine nicht zwiefache Partition $\{\circledast C_1, \ldots \circledast C_p\}$ zu rekonstruieren, die nur $p = 2$ Symbole hat. Diese Partition erweist sich als die wohlbekannte 01-Partition, von der wir schon in Kap. 4.2.2 sahen, daß sie einfacher als die AB-Partition ist. Die Vereinfachung der AB-Partition liefert also das gewünschte Ergebnis.

Abschnitt 6.4 enthält die weiteren Bedingungen 6.32, 6.35, 6.37 und 6.38, die bei der Rekonstruktion der nicht zwiefache Partition $\{\circledast C_1, \ldots \circledast C_p\}$ erfüllt sein müssen. Die Abschnitte 6.4.1 bis 6.4.3 folgern, daß diese nicht zwiefache Partition dieselben expandierenden und kontrahierenden

Blätter besitzt wie die zwiefache Partition $\{A_1\circ,\ldots A_m\circ\}$, $\{\circ B_1,\ldots\circ B_n\}$, aus der sie konstruiert wurde. Es wird gezeigt, daß diese vier Zusatzbedingungen bei der Beispielrechnung automatisch erfüllt sind.

Diese vier Zusatzbedingungen lassen sich unter Umständen nur dadurch erfüllen, daß expandierende oder kontrahierende Bündel in kleinere Teilbündel zerlegt werden. Die Größe der expandierenden und kontrahierenden Blätter wird dagegen nicht verändert. Nach solchen Zerlegungen ist eine nicht zwiefache Partition immer rekonstruierbar, wenn die Zahl p der Symbole groß genug gewählt wird (Kap. 6.5).

Umgekehrt kann man auch Bündel miteinander vereinigen. Dies ist weder für die Vereinfachung einer zwiefachen Partition noch für die Rekonstruktion einer nicht zwiefachen Partition notwendig. Aber es kann diese Rechnungen verkürzen. Allerdings kann man diese Rechnungen auch blockieren, wenn man zu viele Bündel miteinander vereinigt. Denn — wie bereits erwähnt — können zu große expandierende oder kontrahierende Bündel dazu führen, daß man prinzipiell mögliche Vereinfachungsschritte übersieht. Um dieses Risiko gering zu halten, werde ich in Kap. 6.6 eine heuristische Regel vorschlagen, welche expandierenden oder kontrahierenden Bündel miteinander vereinigt werden sollten. Ganz ausschließen kann ich dieses Risiko zwar nicht. Doch zumindest auf die Beispielrechnung von Kap. 5 läßt sich die heuristische Regel mit Erfolg anwenden. Wie ich in Abschnitt 6.6.2 zeigen werde, verkürzt sich die Rechnung zur Vereinfachung der AB–Partition, wenn man Bündel miteinander vereinigt. Die Rekonstruktion der nicht zwiefachen Partition ist nach diesen Vereinigungen sogar trivial. Das Ergebnis, die 01–Partition, bleibt unverändert.

6.2 Die wichtigsten Bedingungen für die Rekonstruktion einer nicht zwiefachen Partition

6.2.1 Definition der beiden Bedingungen

Gegeben sei eine zwiefache Partition $\{A_1\circ,\ldots A_m\circ\}$, $\{\circ B_1,\ldots\circ B_n\}$. Aus ihr soll eine nicht zwiefache Partition $\{\circ C_1,\ldots\circ C_p\}$ rekonstruiert werden, die dieselben expandierenden und kontrahierenden Blätter erzeugt:

$$\{(\bigsqcup_{i=1}^{p} C_i)^{-\infty}\circ\}' = \{(\bigsqcup_{i=1}^{m} A_i)^{-\infty}\circ\}' \tag{6.1}$$

$$\{\circ(\bigsqcup_{i=1}^{p} C_i)^{+\infty}\}' = \{\circ(\bigsqcup_{i=1}^{n} B_i)^{+\infty}\}' \quad . \tag{6.2}$$

(Der Punkt $\circ$ wird an Stelle des Punktes $\circ$ verwendet, weil sich die beiden Partitionen in expandierenden oder kontrahierenden Blättern unterscheiden dürfen, die so klein sind, daß sie in größeren expandierenden oder kontrahierenden Blättern enthalten sind.) Die Zahl p der Symbole C_i soll möglichst klein sein.

Die Grundidee besteht darin, die Symbole A_i bzw. B_i nicht alle gleichzeitig, sondern nacheinander durch C_i–Symbole zu ersetzen. Damit dies möglich ist, muß die Partition $\{\circ C_1,\ldots\circ C_p\}$ so gewählt werden, daß sie mehrere Bedingungen erfüllt. Die beiden wichtigsten Bedingungen besagen, daß sich in allen A_i bzw. B_i–Sequenzen der Länge 2 das rechte bzw. das linke Symbol durch ein C_i–Symbol ersetzen läßt:

$$\forall i,j \quad \exists k: \qquad A_iA_j\circ = A_iC_k\circ \tag{6.3}$$

$$\forall i,j \quad \exists k: \qquad \circ B_jB_i = \circ C_kB_i \quad . \tag{6.4}$$

Diese Bedingungen werden nur dann gefordert, wenn $A_iA_j\circ \neq \emptyset$ bzw. $\circ B_jB_i \neq \emptyset$. Leere Gebiete brauchen nicht betrachtet zu werden. Unter anderem garantiert die Bed. 6.4, daß die Vereinigung aller Symbolischen Gebiete der nicht zwiefachen Partition den Phasenraum überdeckt.

$$\bigcup_{k=1}^{p}(\circ C_k) \supseteq \bigcup_{j,i=1}^{n}(\circ B_jB_i) = \circ E \tag{6.5}$$

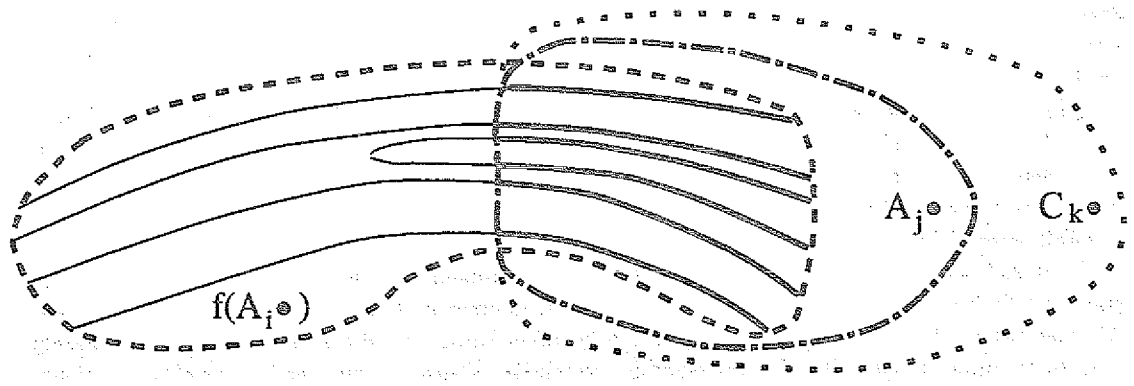


Abbildung 6.1: Die Erzeugung von expandierenden Blättern durch die A_i–Partition, die die Vergangenheit beschreibt, und durch die nicht zwiefache C_i–Partition. Die Blätter des Bündels $\{(\bigsqcup_{l=1}^{m} A_l)^{-\infty} A_i E \circledast\}$' (durchgezogene Linien) sind Bilder von expandierenden Blättern, aber im allgemeinen nicht selbst expandierend. Das Symbolische Gebiet $A_j\circledast$ schneidet aus ihnen Teile heraus, die expandierende Blätter sind (dicke durchgezogene Linien). Das Symbolische Gebiet $C_k\circledast$ muß so gewählt werden, daß es genau dieselben Teile herausschneidet wie das Gebiet $A_j\circledast$ (gestrichelte Linien).

Wozu diese beiden Bedingungen noch dienen und wie sie erfüllt werden können, wird im folgenden diskutiert.

6.2.2 Geometrische Interpretation

Abb. 6.1 illustriert, was die Bedingung 6.3 anschaulich bedeutet. Das Bild $\ldots S_{-3}S_{-2}A_iE\circledast$ eines expandierenden Blattes $\ldots S_{-3}S_{-2}A_i\circledast$ gehört im allgemeinen nicht zu den Blättern, die als expandierend deklariert wurden. Es kann durch den Rand eines Symbolischen Gebietes $A_j\circledast$ geschnitten werden. Durch einen solchen Schnitt entsteht das kleinere Blatt $\ldots S_{-3}S_{-2}A_iA_j\circledast$, das expandierend ist. Aus der Bedingung 6.3 folgt, daß auch das Symbolische Gebiet $C_k\circledast$ dasselbe expandierende Blatt

$$\ldots S_{-3}S_{-2}A_iC_k\circledast \quad = \quad \ldots S_{-3}S_{-2}A_iA_j\circledast \tag{6.6}$$

aus dem großen Blatt $\ldots S_{-3}S_{-2}A_iE\circledast$ herausschneidet. Dabei hängt der Index k von i und j ab, aber nicht von den Symbolen $S_l \in \{A_1, \ldots A_m\}$. Weil der Index k von i abhängt, kann das Symbolische Gebiet $C_k\circledast$ von $A_j\circledast$ verschieden sein. Da die Zahl p der C_k–Symbole klein sein soll, wird das Symbolische Gebiet $C_k\circledast$ normalerweise größer sein als $A_j\circledast$.

6.3 Rekonstruktion der 01–Partition

6.3.1 Aufgabenstellung

In der Beispielrechnung des letzten Kapitels wurde die AB–Partition vereinfacht. Es ergab sich die zwiefache Partition, deren Symbolische Gebiete in Abb. 5.12 und Abb. 5.16 stehen. Die Vergangenheit wird durch die Partition

$$\{A_1\circledast, \ldots A_m\circledast\} = \{AA\circledast A, \quad AAA\circledast B, \quad AAB\circledast, \quad BB\circledast, \quad BA\circledast, \quad BAA\circledast B \cup BAB\circledast A\} \tag{6.7}$$

beschrieben, die Zukunft durch die Partition

$$\{\circledast B_1, \ldots \circledast B_n\} \quad = \quad \{\circledast AA, \quad \circledast ABA, \quad \circledast BA, \quad \circledast BBA, \quad \circledast EBB\} \quad . \tag{6.8}$$

Wie bereits erwähnt, läßt sich diese zwiefache Partition nicht weiter vereinfachen.

Nun soll eine nicht zwiefache Partition $\{\circ C_1, \ldots \circ C_p\}$ rekonstruiert werden, die dieselben expandierenden und kontrahierenden Blätter erzeugt wie die zwiefache Partition in Gl. 6.7 und 6.8. Wie im Überblick 6.1 erwähnt, weiß man von vorneherein nicht, wie groß die Zahl p der Symbole sein muß, um alle Bedingungen zu erfüllen. Der erste Versuch wird mit nur zwei Symbolen gestartet, die hier $C_1 := 0$ und $C_2 := 1$ genannt werden.

In diesem Abschnitt sollen zunächst die beiden Bedingungen 6.3 und 6.4 erfüllt werden. Die anderen Bedingungen, die bei der Rekonstruktion einer zwiefachen Partition beachten werden müssen, werden erst in Kap. 6.4 diskutiert.

6.3.2 Graphische Hilfsmittel

Beschriftungen der Übergänge

In Kap. 5.3.2 wurden Graphen eingeführt, die zeitliche Übergänge zwischen den Symbolischen Gebieten illustrieren. Diese Graphen sind auch jetzt nützliche Hilfsmittel für die Rechnung. Betrachten wir zunächst die Partition, die die Vergangenheit beschreibt. Ihre zeitlichen Übergänge wurden schon in Abb. 5.16 gezeigt. Abb. 6.2 zeigt sie nochmals, aber mit zusätzlichen Beschriftungen. Jeder Übergang von einem Symbolischen Gebiet $A_i\circ$ zu einem Symbolischen Gebiet $A_j\circ$ wird mit dem Gebiet $A_i \circ A_j$ beschriftet. Es sei daran erinnert, daß ein Übergang $A_i\circ \to A_j\circ$ existiert, wenn

$$f(A_i\circ) \cap A_j\circ = A_iA_j\circ \neq \emptyset. \tag{6.9}$$

Deshalb tritt jede Gebiet $A_i \circ A_j$, das nicht leer ist, irgendwo in dem Graphen als Beschriftung eines Überganges $A_i\circ \to A_j\circ$ auf.

Für jeden dieser Übergänge soll die Bedingung 6.3:

$$A_i \circ A_j = A_i \circ C_k \tag{6.10}$$

erfüllt werden. Für das Symbol C_k kann das Symbol 0 oder das Symbol 1 gewählt werden. Diese Wahl wird während der folgenden Rechnung getroffen werden. Sie wird in den Graphen eingetragen, indem der jeweilige Übergang mit dem Symbol 0 oder dem Symbol 1 beschriftet wird. Die Beschriftung in Abb. 6.2 zeigt bereits das Endergebnis der folgenden Rechnung.

Eine Anmerkung zur Schreibweise: Der Ausdruck $A_i\circ C_k$ kann nicht immer als einfache Sequenz der Symbole A, B, 0 und 1 geschrieben werden. Wenn zum Beispiel $A_i\circ = AAA\circ B$ und $\circ C_k = \circ 1$, dann kann man einen solchen Ausdruck nur als Schnitt $AAA \circ B \cup \circ 1$ schreiben, da die Symbole B und 1 dieselbe Stelle einnehmen.

Auf analoge Weise wird die Partition $\{\circ B_1, \ldots \circ B_n\}$, die die Vergangenheit beschreibt, durch einen Graphen dargestellt. Abb. 5.12 zeigte die zeitlichen Übergänge unter der inversen Dynamik f^{-1}. Abb. 6.3 zeigt diese Übergänge mit zusätzlichen Beschriftungen. Jeder Übergang $\circ B_i \to \circ B_j$ ist mit dem Gebiet $\circ B_jB_i$ beschriftet. Außerdem ist das Symbol $C_k \in \{0, 1\}$ eingetragen, das am Ende der Rechnung die Bedingung 6.4

$$\circ C_kB_i = \circ B_jB_i \tag{6.11}$$

erfüllen wird.

Die beiden Beschriftungen erfüllen daher in beiden Graphen eine einfache Beziehung. Das Gebiet $A_i \circ A_j$ oder $\circ B_jB_i$, mit dem ein Übergang beschriftet ist, muß in $\circ C_k$ enthalten sein, falls der Übergang mit C_k beschriftet wird.

Anmerkung: Analogien zur Graphentheorie

Diese Weise, eine nicht zwiefache Partition zu rekonstruieren, ist nicht aus der Luft gegriffen. Sie beruht auf elementaren Konzepten der Graphentheorie und der Theorie formaler Sprachen (siehe z. B. [43]). Betrachten wir dazu ein expandierendes Blatt. Es ist durch eine Symbolfolge $\ldots S_{-2}S_{-1}\circ$ definiert (mit $S_i \in \{A_1, \ldots A_m\}$). Diese Symbolfolge definiert aber auch einen halbunendlichen Weg in dem Graphen, der die zeitlichen Übergänge zwischen den Symbolischen

Gebieten $A_i\circ$ darstellt. Dieser Weg verläuft entlang der Pfeile und endet bei der Ecke $S_{-1}\circ$. Eine Startecke besitzt er nicht, da er halbunendlich ist. Umgekehrt definiert auch jeder halbunendliche Weg ein (möglicherweise leeres) expandierendes Blatt.

Die Konstruktion einer nicht zwiefachen Partition hat das Ziel, daß diese Partition dieselben expandierenden (und kontrahierenden) Blätter erzeugt wie die zwiefache Partition. Es soll also Symbole $T_i \in \{C_1, \ldots C_p\}$ geben, mit

$$\ldots S_{-2}S_{-1}\circ \quad = \quad \ldots T_{-2}T_{-1}\circ \quad . \tag{6.12}$$

Die hier vorgeschlagene Rekonstruktionsmethode beruht darauf, daß die Symbolsequenz $\ldots T_{-2}T_{-1}\circ$ denselben Weg durch den Graphen festlegt wie die Symbolsequenz $\ldots S_{-2}S_{-1}\circ$. Der Weg wird im ersteren Fall durch die Beschriftungen der Übergänge und im letzteren Fall durch die Beschriftungen der Ecken festgelegt.

Wege in einem gerichteten Graphen durch Symbole eindeutig festzulegen ist eine Aufgabe, die aus der Theorie formaler Sprachen her vertraut ist. Auf einfache Weise wird sie in sogenannten „deterministischen Automaten" gelöst. Diese entsprechen gerichteten Graphen mit einigen zusätzlichen Eigenschaften: Ein bestimmter Zustand (d. h. eine Ecke des Graphen) soll als Startecke ausgezeichnet sein. Alle Übergänge sollen mit Symbolen C_i bezeichnet sein. Verschiedene Übergänge sollen auch verschiedene Symbole tragen, falls sie an derselben Ecke starten. Dadurch legt jede Sequenz von C_i–Symbolen einen Weg im Graphen fest. (Man könnte den Weg auch dadurch festlegen, daß man mit den C_i–Symbolen nicht die Übergänge sondern die Ecken beschriftet. Diese andere Art der Beschriftung würde eine alternative Rekonstruktionsmethode ergeben, die der hier vorgestellten ähnlich wäre.)

Im gegenwärtigen Problem entsprechen Übergängen von $A_i\circ$ zu $A_j\circ$ die Gebiete $A_i \circ A_j$. Sie „deterministisch" mit den Symbolen C_i zu kennzeichnen bedeutet, die Symbolischen Gebiete $\circ C_1, \ldots \circ C_p$ so zu wählen, daß $A_i \circ \cap \circ C_k$ genau die Punkte aus $A_i\circ$ umfaßt, deren Bilder in $A_j\circ$ liegen. Dies ist der Grund für die Bedingung 6.3. Eine Startecke auszuzeichnen ist bei halbunendlichen Wegen unmöglich. Man muß aber darauf achten, daß die Startecke $A_i\circ$ jedes endlichen Teilweges durch die vorherigen Symbole festgelegt ist. Dies ist letztlich der Grund, warum später die Bedingung 6.37 gefordert werden muß. Und durch die Bedingung 6.32 wird garantiert, daß für jede erlaubte Sequenz von C_i–Symbolen auch ein Weg im Graphen gefunden werden kann. Analoges gilt für die Bedingungen 6.4, 6.35 und 6.38, die an die kontrahierenden Blätter gestellt wurden. Sie werden garantieren, daß auch der andere Graph, der die Übergänge zwischen den Symbolischen Gebieten $\circ B_i$ unter der inversen Dynamik f^{-1} beschreibt, einem deterministischen Automaten ähnlich wird. Daß beide deterministische Automaten auf der gleichen Dynamik f beruhen, äußert sich letztlich darin, daß die Symbolischen Gebiete $\circ C_i$ in beiden Fällen übereinstimmen müssen.

6.3.3 Die Rekonstruktion

Die eigentliche Rekonstruktion der nicht zwiefachen Partition führt in unserem Beispiel schnell zu einem eindeutigen Ergebnis. Das Rechenschema ist einfach: Die Beschriftungen mit Symbolen 0 und 1 in dem Graphen 6.2 haben wegen Gl. 6.10 Auswirkungen auf die Wahl der Symbolischen Gebiete $\circ 0$ und $\circ 1$. Diese Wahl wiederum beeinflußt durch Gl. 6.11 die Beschriftungen in dem anderen Graphen 6.3. Und auf dem umgekehrten Weg wirken die Beschriftungen des Graphen 6.3 auf die Beschriftungen des Graphen 6.2 zurück.

Da die Symbole 0 und 1 zu Beginn der Rechnung frei vertauschbar sind, können sie an einer Stelle der Graphen frei gewählt werden. Sei dies die Ecke $BB\circ$ in Abb. 6.2. Zwei Übergänge gehen von ihr aus. Die Bedingung 6.10 wird an diesen beiden Übergängen erfüllt, indem man

$$BB \circ A \;=\; BB \circ 0 \tag{6.13}$$

$$BB \circ B \;=\; BB \circ 1 \tag{6.14}$$

setzt. Es folgt wegen $BB \circ A \cap BB \circ B = \emptyset$:

$$BB \circ A \;\subseteq\; \circ 0 \quad ; \quad BB \circ A \;\cap\; \circ 1 \;=\; \emptyset \tag{6.15}$$

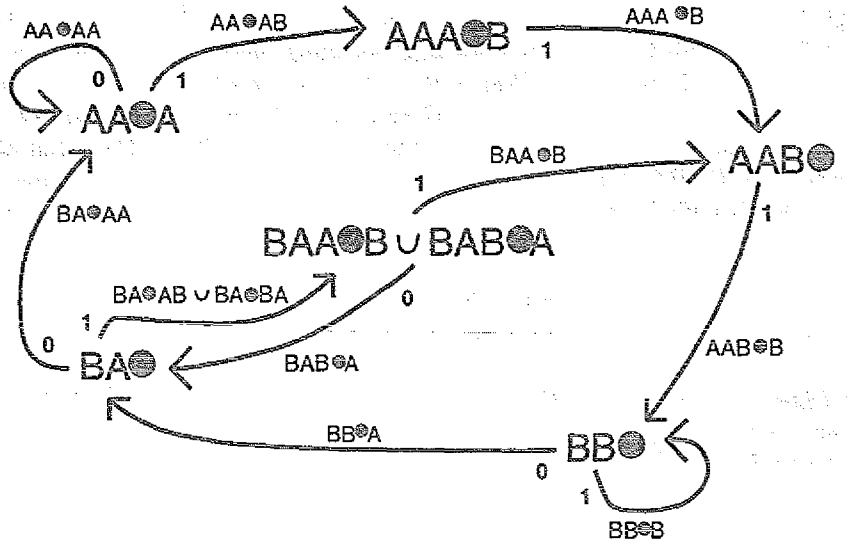


Abbildung 6.2: Die Rekonstruktion von 01–Symbolen in der Partition, die die expandierenden Bündel beschreibt. Der gerichtete Graph ist derselbe wie in Abb. 5.16. Doch jeder Übergang $\mu \to \nu$ trägt zwei zusätzliche Beschriftungen. Die Beschriftung $\mu \cap f^{-1}(\nu)$ bezeichnet das Gebiet im Phasenraum, das diesem Übergang entspricht. Und die Beschriftung mit dem Symbol 0 oder 1 kennzeichnet, ob dieses Gebiet nach der Rekonstruktion der 01–Partition in dem Symbolischen Gebiet $\circledast 0$ oder in $\circledast 1$ enthalten ist.

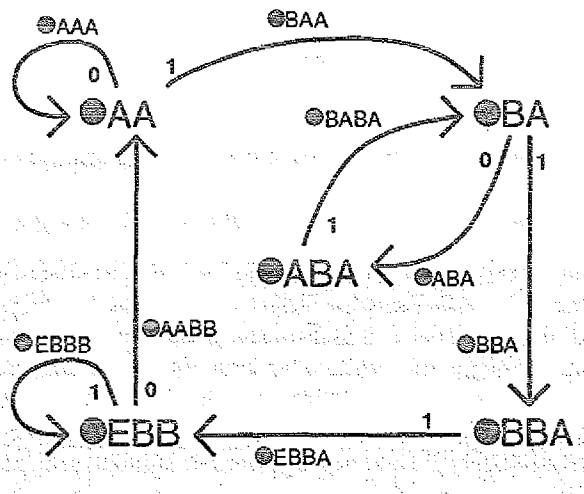


Abbildung 6.3: Die Rekonstruktion von 01–Symbolen in der Partition, die die kontrahierenden Bündel beschreibt. Der gerichtete Graph ist derselbe wie in Abb. 5.12. Doch jeder Übergang $\circledast\alpha \to \circledast\beta$ trägt zwei zusätzliche Beschriftungen. Die Beschriftung $\circledast E\alpha \cap \circledast\beta$ bezeichnet das Gebiet im Phasenraum, das diesem Übergang entspricht. Und die Beschriftung mit dem Symbol 0 oder 1 kennzeichnet, ob dieses Gebiet nach der Rekonstruktion der 01–Partition in dem Symbolischen Gebiet $\circledast 0$ oder in $\circledast 1$ enthalten ist.

$$BB \circledast B \subseteq \circledast 1 \; ; \quad BB \circledast B \cap \circledast 0 = \emptyset \; . \tag{6.16}$$

Dies legt Beschriftungen im anderen Graphen 6.3 fest. Zum Beispiel kann der Übergang $\circledast EBB \to \circledast EBB$ nicht mit 0 beschriftet werden. Denn würde man ihn so beschriften, dann müßte wegen Gl. 6.11 gelten: $\circledast EBBB \subseteq \circledast 0$. Wegen $BB \circledast B \cap \circledast 0 = \emptyset$ wäre dies nur dann möglich, falls $BB \circledast B \cap \circledast EBBB = \emptyset$. Es gibt aber keinen Grund anzunehmen, die Sequenz $BB \circledast BBBB$ sei verboten. Für den numerisch berechneten Attraktor in Abb. 4.2 gilt im Gegenteil die grammatikalische Regel $BB \circledast BBBB \neq \emptyset$. Der Übergang $\circledast EBB \to \circledast EBB$ muß deshalb mit 1 beschriftet werden. Und folglich muß der andere von $\circledast EBB$ ausgehende Übergang $\circledast EBB \to \circledast AA$ mit dem Symbol 0 beschriftet werden.

$$\circledast AABB = \circledast 0EBB \tag{6.17}$$
$$\circledast EBBB = \circledast 1EBB \tag{6.18}$$

Hieraus folgen dann umgekehrt wieder Beschriftungen im Graphen 6.2. Zum Beispiel folgt aus $\circledast EBBB \cap \circledast 0 = \emptyset$ und der grammatikalischen Regel $AA \circledast AB \cap \circledast EBBB \neq \emptyset$, daß der Übergang $AA \circledast A \to AAA \circledast B$ nicht mit 0 beschriftet werden kann. Die weitere Rechnung ist so einfach, daß sie hier nicht ausgeführt werden soll.

6.3.4 Das Endergebnis

Letztlich ergibt sich als einzig mögliche Beschriftung die in Abb. 6.2 und 6.3 gezeigte. Das Symbolische Gebiet $\circledast 1$ ergibt sich aus dieser Beschriftung, indem alle mit 1 beschrifteten Übergänge $A_i \circledast \xrightarrow{1} A_j \circledast$ in Abb. 6.3 zusammengefaßt werden:

$$\begin{aligned}
\circledast 1 &= \bigcup_{A_i} A_i \circledast 1 && (6.19)\\
&= \bigcup_{A_i \circledast \xrightarrow{1} A_j \circledast} A_i \circledast A_j && (6.20)\\
&= \circledast EBBB \cup \circledast BAA \cup \circledast BBA \cup \circledast BABA \cup \circledast EBBA && (6.21)\\
&= \circledast ABB \cup \circledast B \setminus \circledast BABB && (6.22)\\
&= \circledast B \cup A \circledast AB && (6.23)
\end{aligned}$$

(da laut Gl. 4.37 gilt $\circledast BABB = B \circledast ABB = A \circledast ABA = \emptyset$). Bei disjunkten Symbolen $\circledast 0 \cap \circledast 1 = \emptyset$ folgt daraus:

$$\circledast 0 = \circledast E \setminus \circledast 1 = B \circledast A \cup A \circledast AA \; . \tag{6.24}$$

Dieses Endergebnis stimmt tatsächlich mit der Definition der 01–Partition in Gl. 4.32 und 4.33 überein. Die Vereinfachung der AB–Partition liefert das gewünschte Ergebnis.

Wenn die Zusatzbedingung $\circledast 0 \cap \circledast 1 = \emptyset$ nicht wäre, dann könnten weitere Gebiete zu diesem $\circledast 0$ hinzugefügt werden, ohne daß expandierende oder kontrahierende Blätter vergrößert oder andere Bedingungen verletzt würden. Dies ist möglich, weil in Abb. 6.3 und 6.2 von manchen Ecken nur ein einzelner Übergang ausgeht, der mit 1 beschriftet wurde. Da die Gebiete $AAA \circledast BBBA$, $AAB \circledast BABA$ und $AAB \circledast BBBA$, die hinzugefügt werden können, sehr klein sind, habe ich darauf verzichtet, sie geometrisch zu interpretieren. Es kann also verschiedene nicht zwiefache Partitionen geben, die dieselben expandierenden und kontrahierenden Blätter erzeugen.

6.4 Zusatzbedingungen für die Rekonstruktion

Dieses eine Beispiel beweist noch nicht, daß sich immer eine nicht zwiefache Partition $\{\circledast C_1, \ldots \circledast C_p\}$ rekonstruieren läßt, die dieselben kontrahierenden und expandierenden Blätter besitzt wie die zwiefache Partition $\{A_1\circledast, \ldots A_m\circledast\}$, $\{\circledast B_1, \ldots \circledast B_n\}$. Diese Rekonstruktion gelingt nur, wenn

Zusatzbedingungen erfüllt werden, die ich in diesem Abschnitt vorstellen will. Es sind dies die vier Bedingungen 6.32, 6.35, 6.37 und 6.38. Aus ihnen folgt dann insbesondere, daß die rekonstruierte nicht zwiefache Partition generierend ist. Kap. 6.4.4 zeigt, daß diese Zusatzbedingungen bei der Rekonstruktion der 01-Partition erfüllt sind. Wie sich diese Zusatzbedingungen im allgemeinen Fall erfüllen lassen, wird in Kap. 6.5 diskutiert. Dabei kann es sich unter Umständen als nötig erweisen, expandierende oder kontrahierende Bündel in kleinere Teilbündel zu zerlegen.

6.4.1 Die Notwendigkeit von Zusatzbedingungen

Im folgenden bezeichnet T_i ein Symbol aus der nicht zwiefachen Partition $\{\circledast C_1, \ldots \circledast C_p\} \ni \circledast T_i$. Die mit S_i bezeichneten Symbole stammen für $i < 0$ aus der Partition $\{A_1\circledast, \ldots A_m\circledast\} \ni S_i\circledast$, welche die Vergangenheit beschreibt, und für $i \geq 0$ aus der Partition $\{\circledast B_1, \ldots \circledast B_n\} \ni \circledast S_i$, welche die Zukunft beschreibt.

Warum muß man überhaupt noch zusätzliche Bedingungen an die Partition $\{\circledast C_1, \ldots \circledast C_p\}$ stellen? Auf den ersten Blick könnte man meinen, schon allein mit Hilfe der Bedingung 6.3 die expandierenden Blätter $\ldots S_{-2}S_{-1}\circledast$ der zwiefachen Partition in expandierende Blätter $\ldots T_{-2}T_{-1}\circledast$ der nicht zwiefachen Partition übersetzen zu können. Diese Bedingung 6.3 besagte, daß jede Sequenz $S_{-2}S_{-1}\circledast$ mit dem passenden Symbol $T_{-1} \in \{C_1, \ldots C_p\}$ als eine Sequenz $S_{-2}T_{-1}\circledast$ geschrieben werden kann. Durch Induktion folgt, daß sich auch für jede d-stellige Sequenz $S_{-d}\ldots S_{-2}S_{-1}\circledast$ Symbole $T_{1-d}, \ldots T_{-1} \in \{C_1, \ldots C_p\}$ finden lassen, so daß:

$$S_{-d}S_{1-d}\ldots S_{-2}S_{-1}\circledast \quad = \quad S_{-d}T_{1-d}\ldots T_{-2}T_{-1}\circledast \quad . \tag{6.25}$$

Bevor diese Beziehung auf halbunendliche Sequenzen verallgemeinert werden kann, muß das einzeln stehende Symbol S_{-d} entfernt werden.

$$S_{-d}S_{1-d}\ldots S_{-2}S_{-1}\circledast \quad \subseteq \quad T_{1-d}\ldots T_{-2}T_{-1}\circledast \quad , \tag{6.26}$$

Da d beliebig groß sein darf, kann man sogar für jede halbunendliche Symbolsequenz $\ldots S_{-3}S_{-2}S_{-1}$ eine halbunendliche Symbolsequenz $\ldots T_{-3}T_{-2}T_{-1}$ finden, mit

$$\ldots S_{-2}S_{-1}\circledast \quad \subseteq \quad \ldots T_{-2}T_{-1}\circledast \quad . \tag{6.27}$$

Dieselbe Argumentation läßt sich wesentlich knapper schreiben, wenn man sich an die algebraische Schreibweise der Bündel gewöhnt hat. Aus der Bed. 6.3 folgt

$$\{(\bigsqcup_{i=1}^{m} A_i)^2\circ\}' \quad \subseteq \quad \{(\bigsqcup_{i=1}^{m} A_i)(\bigsqcup_{i=1}^{p} C_i)\circ\}' \quad . \tag{6.28}$$

Wenn man diese Beziehung mit f^t abbildet und die Beziehungen für alle $t \geq 0$ miteinander schneidet, erhält man

$$\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty}\circ\}' \quad \subseteq \quad \{(\bigsqcup_{i=1}^{p} C_i)^{-\infty}\circ\}' \quad , \tag{6.29}$$

was zu Gl. 6.27 äquivalent ist.

Auf diese Weisen folgt aus der Bed. 6.3, daß jedes expandierende Blatt der zwiefachen Partition in einem mindestens ebenso großen expandierenden Blatt der nicht zwiefachen Partition enthalten ist. Die Zusatzbedingungen 6.32 und 6.37 werden notwendig sein, um auch das Gegenteil beweisen zu können: Jedes expandierende Blatt der nicht zwiefachen Partition ist in einem mindestens ebenso großen expandierenden Blatt der zwiefachen Partition enthalten. Mit Bündeln läßt sich dies schreiben als

$$\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty}\circ\}' \quad \supseteq \quad \{(\bigsqcup_{i=1}^{p} C_i)^{-\infty}\circ\}' \quad . \tag{6.30}$$

Zusammen mit Gl. 6.29 folgt das gewünschte Ergebnis, daß nämlich die zwiefache und die nicht zwiefache Partition dieselben expandierenden Blätter erzeugen, mit einer kleinen Einschränkung:

Expandierende Blätter α, die so klein sind, daß sie in größeren expandierenden Blättern $\beta \supset \alpha$ enthalten sind, müssen bei den beiden Partitionen nicht übereinstimmen. Diese Einschränkung stört nicht. Denn schon in Kap. 5.2.4 wurde begründet, warum diese expandierenen Blätter α irrelevant sind und entfernt werden können.

6.4.2 Definition der Zusatzbedingungen

Es soll nun garantiert werden, daß jedes expandierende Blatt der nicht zwiefachen Partition $\{\circledast C_1, \ldots \circledast C_p\}$ in einem mindestens ebenso großen expandierenden Blatt der zwiefachen Partition $\{A_1\circledast, \ldots A_m\circledast\}$, $\{\circledast B_1, \ldots \circledast B_n\}$ enthalten ist. Dieses Ziel wird durch Beziehung 6.30 ausgedrückt. Diese Beziehung enthält unendlich lange Symbolsequenzen und ist auf direktem Wege schwer nachzuprüfen. Deswegen werden weiter unten die Bedingungen 6.32 und 6.37 aufgestellt, die leichter nachzuprüfen sind, weil sie nur endliche Symbolsequenzen enthalten. Diese beiden Zusatzbedingungen sind hinreichend, aber nicht unbedingt notwendig, damit Beziehung 6.30 erfüllt ist.

Um die erste Zusatzbedingung geometrisch zu verstehen, muß man sich an die Abb. 6.1 erinnern. Diese Abbildung illustrierte, wie das Symbolische Gebiet $A_j\circledast$ und das Symbolische Gebiet $C_k\circledast$ dieselben expandierenden Blätter aus den Blättern des Bündels $f(\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} A_i\circledast\}')$ herausschneiden. Bisher wurde nur eine Bedingung gestellt, nämlich die in Gl. 6.3:

$$\forall i,j \quad \exists k: \qquad A_iA_j\circledast \quad = \quad A_iC_k\circledast \quad . \tag{6.31}$$

Es soll also in jedem Fall ein Symbolisches Gebiet $C_k\circledast$ geben, das expandierende Blätter der richtigen Größe herausschneidet. Wenn außerdem noch ein Symbolisches Gebiet $C_l\circledast$ existiert, das kleinere expandierende Blätter herausschneidet, so stört es — wie gesagt — nicht.

Noch gefordert werden muß, daß kein Symbolisches Gebiet $C_k\circledast$ zu große Blätter herausschneidet:

$$\forall i,k \quad \exists j: \qquad A_iA_j\circledast \quad \supseteq \quad A_iC_k\circledast \quad . \tag{6.32}$$

Man beachte den Unterschied zu der vorherigen Bedingung 6.3: Der Index k ist nun vorgegeben und ein passender Index j wird gesucht. Mit Bündeln kann man diese Zusatzbedingung schreiben als:

$$\{(\bigsqcup_{i=1}^{m} A_i)(\bigsqcup_{j=1}^{m} A_j)\circ\}' \quad \supseteq \quad \{(\bigsqcup_{i=1}^{m} A_i)(\bigsqcup_{k=1}^{p} C_k)\circ\}' \quad . \tag{6.33}$$

Zusammen mit der alten Bedingung 6.31 ergibt sich eine Übereinstimmung zwischen den Gebieten, die durch Symbolsequenzen $A_iA_j\circledast$ bzw. $A_iC_k\circledast$ definiert sind:

$$\{(\bigsqcup_{i=1}^{m} A_i)(\bigsqcup_{j=1}^{m} A_j)\circ\}' \quad = \quad \{(\bigsqcup_{i=1}^{m} A_i)(\bigsqcup_{k=1}^{p} C_k)\circ\}' \quad . \tag{6.34}$$

Die analoge Zusatzbedingung für den anderen Teil $\{\circledast B_1, \ldots, \circledast B_n\}$ der zwiefachen Partition lautet ganz ähnlich:

$$\forall i,k \quad \exists j: \qquad \circledast B_jB_i \quad \supseteq \quad \circledast B_jC_k \quad . \tag{6.35}$$

Zusammen mit der alten Bedingung 6.4 folgt:

$$\{\circ(\bigsqcup_{j=1}^{n} B_j)(\bigsqcup_{i=1}^{n} B_i)\}' \quad = \quad \{\circ(\bigsqcup_{k=1}^{p} C_k)(\bigsqcup_{i=1}^{n} B_i)\}' \quad . \tag{6.36}$$

Daneben sind noch zwei weitere Zusatzbedingungen erforderlich. Sie lassen sich anschaulich so interpretieren, daß hinreichend lange Symbolsequenzen $T_{-d} \ldots T_{-1} \circledast T_0T_1 \ldots T_d$ der nicht zwiefachen Partition den Phasenraum mindestens ebenso fein unterteilen wie die zwiefache Partition. Formal bedeutet das, daß es ein $d > 0$ gibt, mit:

$$\{(\bigsqcup_{j=1}^{m} A_j)\circ\}' \quad \supseteq \quad \{(\bigsqcup_{i=1}^{p} C_i)^d \circ (\bigsqcup_{k=1}^{p} C_k)^d\}' \tag{6.37}$$

$$\{\circ(\bigsqcup_{j=1}^{n} B_j)\}^{\backprime} \supseteq \{(\bigsqcup_{i=1}^{p} C_i)^d \circ (\bigsqcup_{k=1}^{p} C_k)^d\}^{\backprime} \quad . \tag{6.38}$$

Diese Beziehungen sind notwendig, weil in den Graphen der zeitlichen Übergänge keine Ecke $A_j\circ$ oder $\circ B_j$ als Startecke ausgezeichnet ist. Sie lassen sich so interpretieren, daß durch hinreichend lange C_i–Symbolsequenzen eine Startecke festgelegt sein soll. Wenn diese Beziehungen für eine Sequenzlänge $d > 0$ erfüllt sind, dann natürlich auch für alle größeren Sequenzlängen $d^{\backprime} > d$. Da die Sequenzlänge d beliebig groß gewählt werden kann, kann die Zahl p der C_k–Symbole trotz dieser Zusatzbedingungen kleiner sein als die Zahl der A_i–Symbole und der B_j–Symbole.

Zusammen genügen diese vier Zusatzbedingungen 6.32, 6.35, 6.37 und 6.38 um zu garantieren, daß jedes expandierende bzw. kontrahierende Blatt der nicht zwiefachen Partition in einem mindestens ebenso großen expandierenden bzw. kontrahierenden Blatt der zwiefachen Partition enthalten ist. Die umgekehrte Beziehung wurde bereits im letzten Abschnitt bewiesen.

6.4.3 Beweis, daß die rekonstruierte Partition nicht zu große Blätter erzeugt

Es soll gezeigt werden, daß jedes expandierende Blatt $\ldots T_{-2}T_{-1}\circ$ der nicht zwiefachen Partition ($T_i \in \{C_1, \ldots C_p\}$) in einem mindestens ebenso großen expandierenden Blatt $\ldots S_{-2}S_{-1}\circ$ der zwiefachen Partition ($S_i \in \{A_1, \ldots A_m\}$) enthalten ist. Der Beweis stützt sich auf die Zusatzbedingung 6.37, die es erlaubt, hinreichend weit links stehende C_i–Symbole durch A_i–Symbole zu ersetzen. Dann wird die Zusatzbedingung 6.32 verwendet, um alle weiter rechts stehenden C_i–Symbole auch zu ersetzen.

Der Beweis läßt sich nur mit Bündeln in kompakter Form schreiben. Aus der Zusatzbedingung 6.37 folgt, daß für jedes $r \geq 0$:

$$\{(\bigsqcup_{i=1}^{p} C_i)^{2d} E^r \circ\}^{\backprime} \subseteq \{(\bigsqcup_{j=1}^{m} A_j) E^{r+d} \circ\}^{\backprime} \quad . \tag{6.39}$$

Für die expandierenden Blätter der nicht zwiefachen Partition folgt dann mit der bereits erwähnten Rechenregel B.29:

$$\{\mu \circ \nu\}^{\backprime} = \{\mu\circ\}^{\backprime} \cap \{\circ\nu\}^{\backprime} \tag{6.40}$$

aus Anhang B:

$$\{(\bigsqcup_{i=1}^{p} C_i)^{-\infty}\circ\}^{\backprime} \subseteq \{(\bigsqcup_{i=1}^{p} C_i)^{2d} E^r \circ\}^{\backprime} \cap \{(\bigsqcup_{i=1}^{p} C_i)^{r+d}\circ\}^{\backprime} \subseteq \{(\bigsqcup_{j=1}^{m} A_j)(\bigsqcup_{i=1}^{p} C_i)^{r+d}\circ\}^{\backprime} \quad . \tag{6.41}$$

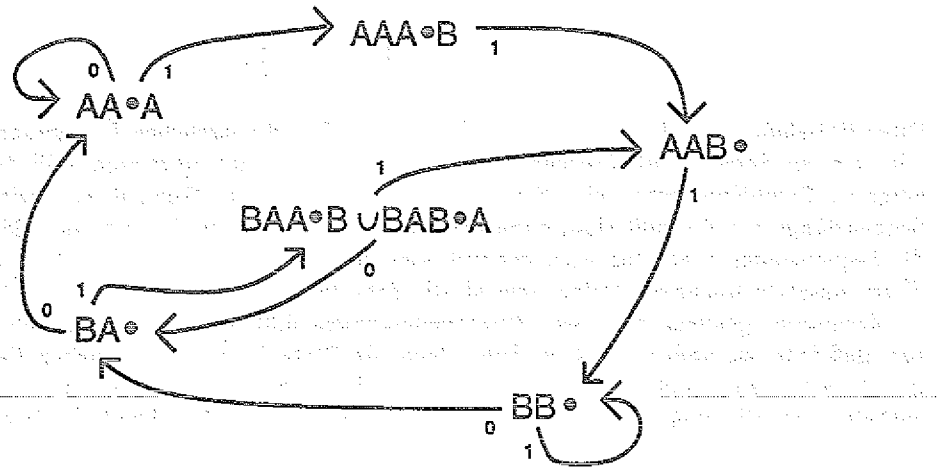
Die andere Bedingung 6.32 geht in der Form 6.33 ein. Indem man sie $(r+d)$–mal anwendet, erhält man:

$$\{(\bigsqcup_{j=1}^{m} A_j)(\bigsqcup_{i=1}^{p} C_i)^{r+d}\circ\}^{\backprime} \subseteq \{(\bigsqcup_{j=1}^{m} A_j)^2(\bigsqcup_{i=1}^{p} C_i)^{r+d-1}\circ\}^{\backprime} \subseteq \ldots \subseteq \{(\bigsqcup_{j=1}^{m} A_j)^{r+d+1}\circ\}^{\backprime} \quad . \tag{6.42}$$

Da man r beliebig groß wählen kann, folgt schließlich

$$\{(\bigsqcup_{i=1}^{p} C_i)^{-\infty}\circ\}^{\backprime} \subseteq \{(\bigsqcup_{j=1}^{m} A_j)^{-\infty}\circ\}^{\backprime} \quad . \tag{6.43}$$

Jedes expandierende Blatt der nicht zwiefachen Partition ist daher in einem mindestens ebenso großen expandierenden Blatt der zwiefachen Partition enthalten, was zu beweisen war. Ganz analog folgt aus den Zusatzbedingungen 6.35 und 6.38, daß jedes kontrahierende Blatt der nicht zwiefachen Partition in einem mindestens ebenso großen kontrahierenden Blatt der zwiefachen Partition enthalten ist.



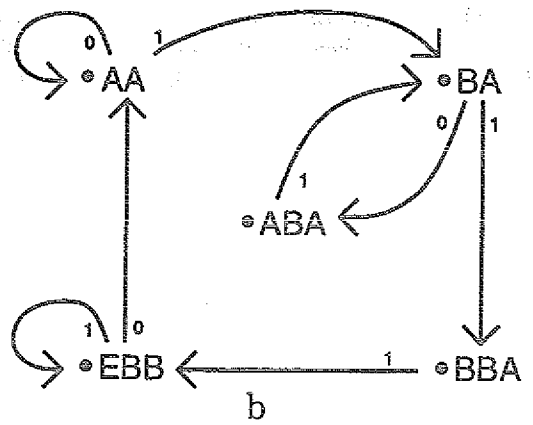


Abbildung 6.4: Zusammenfassung der Abbildungen 6.2 und 6.3. Die zeitlichen Übergänge zwischen den Symbolischen Gebieten der zwiefachen Partition sind mit den Symbolen $C_1 = 0$ und $C_2 = 1$ beschriftet.

6.4.4 Die 01–Partition erfüllt die Zusatzbedingungen

Die Vereinfachung der AB–Partition lieferte als Endergebnis die 01–Partition in Gl. 6.23 und 6.24. Abb. 6.4 zeigt noch einmal die beiden Graphen der zeitlichen Übergänge. Die Übergänge wurden in Kap. 6.3 so mit 0 und 1 beschriftet, daß die Symbolischen Gebiete $\circ 0$ und $\circ 1$ die beiden wichtigsten Bedingungen 6.3 und 6.4 erfüllen. Es stellt sich nun heraus, daß damit gleichzeitig auch die vier Zusatzbedingungen 6.32, 6.35, 6.37 und 6.38 erfüllt sind.

Die Zusatzbedingungen 6.32 und 6.35 sind erfüllt

Betrachten wir einen Übergang in Abb. 6.4.a von der Ecke $A_i\circ$ zu der Ecke $A_j\circ$. Wenn dieser Übergang mit dem Symbol $C_k \in \{0,1\}$ beschriftet ist, dann gilt laut Bed. 6.3:

$$A_i \circ A_j \quad = \quad A_i \circ C_k \quad . \tag{6.44}$$

Damit ist die Bedingung 6.32:

$$\exists j: \qquad A_i \circ A_j \quad \supseteq \quad A_i \circ C_k \tag{6.45}$$

für diese Wahl der Indices i und k trivialerweise erfüllt.

Es bleiben noch die Fälle zu betrachten, in denen von der Ecke $A_i\circledast$ kein mit C_k beschrifteter Übergang ausgeht. In Abb. 6.4.a betrifft dies die Ecken $A_i\circledast$, von denen nur ein einziger Übergang $A_i\circledast \to A_j\circledast$ ausgeht. Für diese Ecken gilt $A_i\circledast = A_i \circledast A_j$. Diese Wahl von j erfüllt dann auch trivialerweise die Bedingung

$$A_i \circledast C_k \quad \subseteq \quad A_i\circledast \quad = \quad A_i \circledast A_j \quad . \tag{6.46}$$

Damit ist die Zusatzbedingung 6.32 erfüllt. Analog zeigt man, daß die 01–Partition — oder jede andere Partition mit disjunkten Symbolischen Gebieten — auch die Zusatzbedingung 6.35 erfüllt.

Graphische Interpretation der Zusatzbedingungen 6.37 und 6.38

Interessanter sind die beiden anderen Zusatzbedingungen 6.37 und 6.38. Man beweist sie am besten in der verschärften Form:

$$\{(\bigsqcup_{i=1}^{p} C_i)^d\circ\}^{\cdot} \quad \subseteq \quad \{(\bigsqcup_{j=1}^{m} A_j)\circ\}^{\cdot} \tag{6.47}$$

$$\{\circ(\bigsqcup_{i=1}^{p} C_i)^d\}^{\cdot} \quad \subseteq \quad \{\circ(\bigsqcup_{j=1}^{n} B_j)\}^{\cdot} \quad . \tag{6.48}$$

Diese Bedingungen lassen sich nämlich graphisch interpretieren. Dazu betrachten wir eine Folge von d Übergängen in dem Graphen (das ist ein „Weg der Länge d"):

$$S_{-d-1}\circledast \stackrel{T_{-d}}{\to} S_{-d}\circledast \ldots \stackrel{T_{-1}}{\to} S_{-1}\circledast \tag{6.49}$$

(mit $S_i \in \{A_1, \ldots A_m\}$ und $T_i \in \{0,1\}$). Die Symbolischen Gebiete $\circledast 0$ und $\circledast 1$ wurden so gewählt, daß die Bedingung 6.3 erfüllt ist: $S_{i-1} \circledast S_i = S_{i-1} \circledast T_i$. Daraus folgt:

$$S_{-d-1}S_{-d}\ldots S_{-2}S_{-1}\circledast \quad = \quad S_{-d-1}T_{-d}\ldots T_{-2}T_{-1}\circledast \quad . \tag{6.50}$$

Die Startecke $S_{-d-1}\circledast$ und die Folge der Beschriftungen $T_{-d}\ldots T_{-2}T_{-1}$ entlang des Weges legen daher dasselbe Gebiet fest wie die Folge der Ecken $S_{-d-1}\ldots S_{-2}S_{-1}$.

Die Bed. 6.47 besagt, daß jedes Gebiet

$$T_{-d}\ldots T_{-2}T_{-1}\circledast \quad = \quad \bigcup_{S_{-d-1}} S_{-d-1}T_{-d}\ldots T_{-2}T_{-1}\circledast \tag{6.51}$$

als Ganzes in einem Symbolischen Gebiet $A_j\circledast$ enthalten ist. Die Vereinigung muß dabei über alle Startecken $S_{-d-1}\circledast \in \{A_1\circledast, \ldots A_m\circledast\}$ ausgeführt werden.

Doch nicht von jeder Startecke aus ist ein Weg möglich, der die Beschriftungen $T_{-d}\ldots T_{-2}T_{-1}$ trägt. Im allgemeinen Fall müßte man zusätzliche Beschriftungen einführen, um die folgende Argumentation zu retten. Dies will ich hier nicht darstellen. Bei der 01–Partition ist diese Schwierigkeit leicht zu überwinden, weil die Symbolischen Gebieten disjunkt sind: $\circledast 0 \cap \circledast 1 = \emptyset$. Daher kann ein Übergang $S_{i-1}\circledast \stackrel{T_i}{\to}$ nur dann fehlen, wenn $S_{i-1} \circledast T_i = \emptyset$. Und deshalb existiert ein Weg mit der Startecke $S_{-d-1}\circledast$ und den Beschriftungen $T_{-d}\ldots T_{-2}T_{-1}$ nur dann nicht, wenn

$$S_{-d-1}T_{-d}\ldots T_{-2}T_{-1}\circledast \quad = \quad \emptyset \quad . \tag{6.52}$$

In Gl. 6.51 genügt es daher, die Vereinigung nur über all die Startecken $S_{-d-1}\circledast$ zu bilden, von denen aus ein Weg mit den Beschriftungen $T_{-d}\ldots T_{-2}T_{-1}$ existiert. Dieser Weg führt laut Gl. 6.50 zu einer bestimmten Zielecke

$$S_{-1}\circledast \quad \supseteq \quad S_{-d-1}T_{-d}\ldots T_{-2}T_{-1}\circledast \quad . \tag{6.53}$$

(mit $S_{-1}\circ \in \{A_1\circ, \ldots A_m\circ\}$). Falls der so beschriftete Weg von allen möglichen Startecken $S_{-d-1}\circ$ zu derselben Zielecke $A_j\circ$ führt, dann gilt

$$T_{-d}\ldots T_{-2}T_{-1}\circ \;=\; \bigcup_{S_{-d-1}} S_{-d-1}T_{-d}\ldots T_{-2}T_{-1}\circ \;\subseteq\; A_j\circ \quad . \tag{6.54}$$

Die Bed. 6.47 ist dann für diese speziellle Folge von Beschriftungen $T_{-d}\ldots T_{-2}T_{-1}$ erfüllt. Um die verschärfte Zusatzbedingung 6.47 allgemein zu beweisen, muß man also nur nachweisen, daß jede Folge von d Beschriftungen $T_{-d}\ldots T_{-2}T_{-1}$ (mit $T_i \in \{0,1\}$) unabhängig von der Startecke $S_{-d-1}\circ$ zu einer bestimmten Zielecke $S_{-1}\circ$ führt.

Die Zusatzbedingungen 6.37 und 6.38 sind erfüllt

Im Falle der 01–Partition genügt es, die Länge $d = 3$ zu wählen. Denn in Abb. 6.4.a führt jede 01–Symbolsequenz, die mindestens die Länge 3 hat, von allen möglichen Startecken immer zu derselben Zielecke. Zum Beispiel sind die aufeinanderfolgenden Übergänge

$$S_{-3}\circ \xrightarrow{0} S_{-2}\circ \xrightarrow{0} S_{-1}\circ \tag{6.55}$$

von den Ecken $BB\circ, BA\circ, AA\circ A$ und $BAA\circ B \cup BAB\circ A$ aus möglich. Von allen diesen Startecken $S_{-3}\circ$ aus führen die beiden mit 0 bezeichneten Übergänge zu der Ecke $S_{-1}\circ = AA\circ A$. Kurz gesagt: $00\circ \subseteq AA\circ A$. Deshalb führen natürlich auch alle längeren Wege zu der Ecke $AA\circ A$, falls die letzten beiden Übergänge mit 0 beschriftet sind.

Ebenso führt die Sequenz 111 von jeder möglichen Ecke in Abb. 6.4.a immer zu der Ecke $BB\circ$, so daß $111\circ \subseteq BB\circ$ folgt. Betrachtung aller Ecken ergibt:

$$\begin{aligned}
111\circ &\subseteq BB\circ && (6.56)\\
011\circ &\subseteq AAB\circ && (6.57)\\
00\circ &\subseteq AA\circ A && (6.58)\\
10\circ &\subseteq BA\circ && (6.59)\\
001\circ &\subseteq AAA\circ B && (6.60)\\
101\circ &\subseteq BAA\circ B \;\cup\; BAB\circ A \quad . && (6.61)
\end{aligned}$$

Zusammen decken diese 01–Sequenzen alle möglichen Enden einer Symbolsequenz ab:

$$111\circ \;\cup\; 011\circ \;\cup\; 00\circ \;\cup\; 10\circ \;\cup\; 001\circ \;\cup\; 101\circ \;=\; E\circ \quad . \tag{6.62}$$

Es folgt die verschärfte Zusatzbdingung 6.47

$$\{(\bigsqcup_{i=1}^{m} A_i)\circ\}^{\omega} \;\supseteq\; \{(0 \sqcup 1)^3\circ\}^{\omega} \quad . \tag{6.63}$$

Laut Abschnitt 6.4.3 ist daher jedes expandierende Blatt der 01–Partition in einem expandierenden Blatt der zwiefachen Partition enthalten.

Analoges gilt für kontrahierende Blätter. Aus dem Graphen in Abb. 6.3 bzw. 6.4.b kann man ablesen, daß die folgenden, von rechts nach links gelesenen 01–Sequenzen zu einer eindeutigen Zielecke führen:

$$\begin{aligned}
\circ 00 &\subseteq \circ AA && (6.64)\\
\circ 011 &\subseteq \circ AA && (6.65)\\
\circ 010 &\subseteq \circ ABA && (6.66)\\
\circ 10 &\subseteq \circ BA && (6.67)\\
\circ 110 &\subseteq \circ BBA && (6.68)\\
\circ 111 &\subseteq \circ EBB \quad . && (6.69)
\end{aligned}$$

Zusammen decken diese 01–Sequenzen alle möglichen Anfänge einer Symbolsequenz ab:

$$(\circ 00 \cup \circ 011) \quad \cup \quad \circ 010 \quad \cup \quad \circ 10 \quad \cup \quad \circ 110 \quad \cup \quad \circ 111 \quad = \quad \circ E \tag{6.70}$$

Es folgt Bed. 6.48:

$$\{\circ(\bigsqcup_{i=1}^{n} B_i)\}' \quad \supseteq \quad \{\circ(0 \sqcup 1)^3\}' \quad . \tag{6.71}$$

Die 01–Partition, die in Kap. 6.3 nur aus den Bedingungen 6.3 und 6.4 abgeleitet wurde, erfüllt also auch die vier Zusatzbedingungen 6.32, 6.35, 6.37 und 6.38. Sie hat daher dieselben expandierenden und kontrahierenden Blätter wie die zwiefache Partition, die bei der Vereinfachung der AB–Partition entstand. Deshalb ist die 01–Partition einfacher als die AB–Partition und nach wie vor generierend.

6.4.5 Plausible Argumente für die Zusatzbedingungen

Im Falle dieser Beispielrechnung genügte es, bei der Rekonstruktion die beiden wichtigsten Bedingungen 6.3 und 6.4 zu berücksichtigen. Das Ergebnis, die 01–Partition, erfüllte dann auch gleichzeitig die beiden Zusatzbedingungen 6.37 und 6.38, und zwar sogar in ihrer verschärften Form 6.47 und 6.48: Das Ziel eines jeden hinreichend langen Weges hängt nur von den Beschriftungen der Übergänge entlang des Weges und nicht mehr von der Startecke ab. Dies ist kein Zufall, sondern eine plausible Folge davon, daß die gemäß Bed. 6.3 und 6.4 rekonstruierten Gebiete $\circ 0$ und $\circ 1$ disjunkt sind. Diese Plausibilitätsargumente sind für die späteren Kapitel nicht von Bedeutung. Sie können von dem Leser übersprungen werden. Sie können ihm aber auch ein tieferes Verständnis für die Hintergründe der Rekonstruktion geben. Es sind nämlich die bisher kaum betrachteten Schnitte zwischen expandierenden Blättern und den mehrfach iterierten Bildern von kontrahierenden Blättern, die dazu führen, daß an einer Ecke viele gleich beschriftete Übergänge enden.

Betrachten wir als Beispiel das expandierende Bündel $\{(A \sqcup B)^{-\infty} BB\circ\}'$ aus Abb. 6.2 und das kontrahierende Bündel $\{\circ EBB(A \sqcup B)^{+\infty}\}'$ aus Abb. 6.3. (Diese zeigt auch Abb. 6.4.) Wie man sieht, führt eine Folge von drei Übergängen in Abb. 6.4.a genau dann zu der Ecke $BB\circ$ und in Abb. 6.4.b genau dann zu der Ecke $\circ EBB$, wenn diese Übergänge mit 111 beschriftet sind. Wie kommt es zu dieser erstaunlichen Regelmäßigkeit?

Die Beschriftung mit Symbolen 0 und 1 unterlag den Bedingungen 6.3 und 6.4, so daß jedes expandierende oder kontrahierende Blatt der zwiefachen Partiton in einem mindestens ebenso großen expandierenden oder kontrahierenden Blatt der 01–Partition enthalten ist. Insbesondere ist jedes expandierende Blatt $\ldots S_{-4}S_{-3}BB\circ$ (mit $S_i \in \{A, B\}$) in einem Blatt $\ldots T_{-4}T_{-3}T_{-2}T_{-1}\circ$ (mit $T_i \in \{0, 1\}$) der 01–Partition enthalten. Wir suchen eine Erklärung dafür, warum alle Blätter des Bündels $\{(A \sqcup B)^{-\infty} BB\circ\}'$ in den Symbolen $T_{-3}T_{-2}T_{-1} = 111$ übereinstimmen.

Dazu betrachten wir die dreifach iterierten Bilder des kontrahierenden Bündels $\{\circ EBB(A \sqcup B)^{+\infty}\}'$. Diese Bilder haben die Gestalt $BB \circ S_2S_3 \ldots$ (mit $S_i \in \{A, B\}$). Sie sind in Blättern $f^3(\circ T_0T_1T_2T_3 \ldots) = T_0T_1T_2 \circ T_3 \ldots$ enthalten (mit $T_i \in \{0, 1\}$). Wenn eines dieser Bilder eines der obigen expandierenden Blätter schneidet, dann gilt

$$\ldots T_{-4}T_{-3}T_{-2}T_{-1} \circ \quad \cap \quad T_0T_1T_2 \circ T_3 \ldots \quad \neq \quad \emptyset \quad . \tag{6.72}$$

Da die Symbolischen Gebiete $\circ 0$ und $\circ 1$ disjunkt sind, folgt $T_{-3} = T_0$, $T_{-2} = T_1$ und $T_{-1} = T_2$.

Es gibt keine grammatikalischen Regeln, die Schnitte zwischen einem expandierenden Blatt $\ldots S_{-4}S_{-3}BB\circ$ und dem Bild $BB \circ S_2S_3 \ldots = f^3(\circ EBBS_2S_3 \ldots)$ eines kontrahierenden Blattes generell verbieten oder auf eine kleine Teilmenge von Blättern beschränken würden. Deshalb wird eines dieser Bilder normalerweise viele der expandierenden Blätter schneiden. Alle diese expandierenden Blätter müssen in ihren Symbolen $T_{-3}T_{-2}T_{-1}$ übereinstimmen. Umgekehrt schneiden auch diese vielen expandierenden Blätter wieder viele Bilder der kontrahierenden Blätter. Alle diese Bilder müssen in ihren Symbolen $T_0T_1T_2$ übereinstimmen. Da es sehr viele derartige Schnitte gibt, ist es zu erwarten, daß sogar alle Blätter des expandierenden Bündels $\{(A \sqcup B)^{-\infty} BB\circ\}'$

in ihren Symbolen $T_{-3}T_{-2}T_{-1}$ übereinstimmen und daß alle Blätter des kontrahierenden Bündels $\{\circ EBB(A \sqcup B)^{+\infty}\}'$ in ihren Symbolen $T_0T_1T_2$ übereinstimmen.

Es führen also nur ganz bestimmte Folgen von T_i–Beschriftungen zu den Ecken $BB\circ$ oder $\circ EBB$ in den beiden Graphen von Abb. 6.4. Von allen Startecken, von denen aus diese Ecken $BB\circ$ oder $\circ EBB$ in drei Zeitschritten zu erreichen sind, konvergieren die so beschrifteten Übergänge demnach alle zu dieser Ecke. Und wenn zwei gleichbeschriftete Wege an solch einer Ecke zusammenlaufen, dann nehmen sie ab dieser Stelle denselben Verlauf, führen also letztlich zu derselben Zielecke. Ähnliche Argumente kann man auch auf andere Ecken als $BB\circ$ und $\circ EBB$ anwenden. Damit wird es plausibel, daß jede hinreichend lange Folge von T_i–Beschriftungen unabhängig von der Startecke immer zu derselben Zielecke führt.

Dieses Plausibilitätsargument beruhte darauf, daß hinreichend viele Schnitte zwischen expandierenden Blättern und Bildern von kontrahierenden Blättern existieren und daß die Symbolischen Gebiete $\circ 0$ und $\circ 1$ disjunkt sind. Doch auch wenn die Symbolischen Gebiete der rekonstruierten Partition etwas überlappen, bleibt die Argumentation plausibel. Deshalb sollte es bei der Rekonstruktion einer nicht zwiefachen Partition normalerweise genügen, auf die beiden wichtigsten Bedingungen 6.3 und 6.4 zu achten, um gleichzeitig auch die verschärften Zusatzbedingungen 6.47 und 6.48 zu erfüllen.

6.5 Voraussetzungen für die Rekonstruierbarkeit nicht zwiefacher Partitionen

Im Überblick wurde die Rekonstruktion einer nicht zwiefachen Partition als ein Verfahren bezeichnet, das auf Versuch und Irrtum beruht. Man kann von vorneherein nicht wissen, welche Zahl p von Symbolen genügt, um eine nicht zwiefache Partition $\{\circ C_1, \ldots \circ C_p\}$ zu rekonstruieren, die dieselben expandierenden und kontrahierenden Blätter besitzt wie die zwiefache Partition $\{A_1\circ, \ldots A_m\circ\}$, $\{\circ B_1, \ldots \circ B_n\}$. Jedesmal, wenn die Rekonstruktion mit p Symbolen mißlungen ist, soll man einen neuen Versuch mit $p+1$ Symbolen starten.

Damit ist aber noch nicht geklärt, ob eine endliche Zahl von Symbolen überhaupt ausreicht. Dies soll nun untersucht werden.

6.5.1 Wahl der Symbole und Testen der Bedingungen

Wenn man die Zahl p der Symbole nicht klein, sondern nur endlich halten will, dann liegt es nahe, die Symbolischen Gebiete der nicht zwiefachen Partition einfach gleich den Symbolischen Gebieten der zwiefachen Partition zu setzen:

$$\{\circ C_1, \ldots \circ C_p\} \quad := \quad \{\circ A_1, \ldots \circ A_m, \circ B_1, \ldots \circ B_n\} \quad . \tag{6.73}$$

Damit sind die beiden wichtigsten Bedingungen 6.3 und 6.4 auf triviale Weise erfüllt. Es folgt sofort laut Kap. 6.4.1, daß alle expandierenden Blätter $\alpha \in \{(\bigsqcup_{i=1}^{m} A_i)^{-\infty}\circ\}'$ und alle kontrahierenden Blätter $\beta \in \{\circ(\bigsqcup_{i=1}^{n} B_i)^{+\infty}\}'$ der zwiefachen Partition in mindestens ebenso großen expandierenden oder kontrahierenden Blättern der nicht zwiefachen Partition enthalten sind.

Probleme können nur die Zusatzbedingungen 6.32, 6.35, 6.37 und 6.38 bereiten. Wenn man die obigen Symbolischen Gebiete $\circ C_i$ in die ersten beiden Zusatzbedingungen einsetzt, dann erhält man teils triviale Bedingungen und teils die nicht trivialen Bedingungen:

$$\forall i,k \quad \exists j: \qquad A_iA_j\circ \quad \supseteq \quad A_iB_k\circ \qquad \text{und} \tag{6.74}$$

$$\forall i,k \quad \exists j: \qquad \circ B_jB_i \quad \supseteq \quad \circ A_kB_i \quad . \tag{6.75}$$

Die erstere Bedingung läßt sich so interpretieren, daß das Symbolische Gebiet $B_k\circ$ keine größeren Teile aus dem Blatt $\ldots A_iE\circ$ herausschneidet als das Symbolische Gebiet $A_j\circ$. Dadurch wird die Größe der expandierenden Blätter begrenzt.

Die anderen beiden Zusatzbedingungen 6.37 und 6.38 werden am besten gleich in der verschärften Form von Gl. 6.47 und 6.48 gefordert:

$$\{(\bigsqcup_{i=1}^{p} C_i)^d \circ\}' \subseteq \{(\bigsqcup_{j=1}^{m} A_j)\circ\}' \tag{6.76}$$

$$\{\circ(\bigsqcup_{i=1}^{p} C_i)^d\}' \subseteq \{\circ(\bigsqcup_{j=1}^{n} B_j)\}' \quad . \tag{6.77}$$

Die erste dieser Beziehungen besagt, daß für hinreichend große Sequenzlänge d jedes Gebiet $T_{-d}\ldots T_{-2}T_{-1}\circledast$ (mit $T_i \in \{C_1,\ldots C_p\}$) ganz in einem Symbolischen Gebiet $A_j\circledast$ enthalten sein soll. Falls $T_{-1} \in \{A_1,\ldots A_m\}$, so ist dies trivialerweise erfüllt. Es bleibt also nur der Fall $T_{-1} \in \{B_1,\ldots B_n\}$ zu betrachten. Wenn die obige Bedingung 6.74 erfüllt ist, dann sind auch all die Fälle trivial, in denen die Symbole $T_{-d+1}\ldots T_{-1}$ aus $\{B_1,\ldots B_n\}$ und und das Symbol T_{-d} aus $\{A_1,\ldots A_m\}$ stammt. Denn wie schon so oft lassen sich dann die B_i–Symbole nacheinander durch A_i Symbole ersetzen. Und das letzte Symbol A_j, das den Platz von T_{-1} einnimmt, erfüllt die Beziehung:

$$A_j\circledast \supseteq T_{-d}\ldots T_{-2}T_{-1}\circledast \quad . \tag{6.78}$$

Dies garantiert die Beziehung 6.76. Der einzige Fall, in dem die Beziehung 6.76 verletzt sein könnte, ist also der Fall, in dem alle T_i–Symbole aus $\{B_1,\ldots B_n\}$ stammen. Es genügt daher, die Zusatzbedingungen 6.76 und 6.77 in der Form

$$\{(\bigsqcup_{i=1}^{p} B_i)^d \circ\}' \subseteq \{(\bigsqcup_{j=1}^{m} A_j)\circ\}' \qquad \text{und analog} \tag{6.79}$$

$$\{\circ(\bigsqcup_{i=1}^{p} A_i)^d\}' \subseteq \{\circ(\bigsqcup_{j=1}^{n} B_j)\}' \tag{6.80}$$

nachzuprüfen. Wenn diese beiden Voraussetzungen und die Voraussetzungen 6.74 und 6.75 erfüllt sind, dann erzeugt die in Gl. 6.73 definierte nicht zwiefache Partition dieselben expandierenden und kontrahierenden Blätter wie die zwiefache Partition.

6.5.2 Zerlegen von Bündeln

Damit die nicht zwiefache Partition $\{\circledast C_1,\ldots\circledast C_p\}$, die gleich $\{\circledast A_1,\ldots\circledast A_m,\circledast B_1,\ldots B_n\}$ gewählt wurde, dieselben expandierenden und kontrahierenden Bündel erzeugt wie die zwiefache Partition $\{\circledast A_1,\ldots\circledast A_m\}$, $\{\circledast B_1,\ldots B_n\}$, müssen die vier Voraussetzungen 6.74, 6.75, 6.79 und 6.80 erfüllt sein. Wenn eine von diesen Vorausetzungen nicht erfüllt ist, dann muß man eine andere Partition $\{\circledast C_1,\ldots\circledast C_p\}$ wählen.

Wie könnte solch eine andere Wahl aussehen? Die bisherige Wahl würde in dem Graphen der zeitlichen Übergänge bedeuten, daß alle Übergänge, die zu derselben Ecke $A_j\circledast$ führen, mit demselben Symbol C_k beschriftet sind, für welches gilt: $A_j\circledast = C_k\circledast$. Eine andere naheliegende Wahl besteht darin, alle Übergänge mit verschiedenen Symbolen C_k zu beschriften. Diese sind dann definiert durch $C_k\circledast = A_iA_j\circledast$. In dem anderen Graphen erhält man analog Symbole C_k, die durch $\circledast C_k = \circledast B_jB_i$ definiert sind. Die nicht zwiefache Partition lautet dann:

$$\{\circledast C_1,\ldots\circledast C_p\} := \{A_1\circledast A_1, A_2\circledast A_1,\ldots A_m\circledast A_1, A_1\circledast A_2,\ldots A_{m-1}\circledast A_m, A_m\circledast A_m, \circledast B_1B_1,\ldots B_nB_n\} \quad . \tag{6.81}$$

Was man mit dieser neuen Wahl erreicht hat, ist eine Zerlegung der expandierenden und kontrahierenden Bündel. Die Zerlegung von expandierenden und kontrahierenden Bündeln wurde bereits in Kap. 5.2.2 diskutiert. Indem man jedes Symbolische Gebiet $A_j\circledast$ in die m Symbolischen Gebiete $A_iA_j\circledast$ zerlegt, erhält man eine neue zwiefache Partition, die dieselben expandierenden Blätter erzeugt, aber kleinere expandierende Bündel. Analoges gilt für die Zerlegung eines Symbolischen Gebietes $\circledast B_j$ in die kleineren Symbolischen Gebiete $\circledast B_jB_i$.

Damit ist auch klar, welche Voraussetzungen die in Gl. 6.81 definierte Partition erfüllen muß, um die richtigen kontrahierenden und expandierenden Blätter zu erzeugen. Es sind die bekannten vier Voraussetzungen 6.74, 6.75, 6.79 und 6.80, in die nun statt der Symbolischen Gebiete $A_i\circ$ und $\circ B_l$ die kleineren Symbolischen Gebiete $A_hA_l\circ$ und $\circ B_lB_h$ eingesetzt werden müssen.

Betrachten wir zunächst den Fall, daß nur die expandierenden, aber nicht die kontrahierenden Bündel zerlegt wurden. In die vier Voraussetzungen muß man $A_hA_l\circ$ an Stelle von $A_l\circ$ einsetzen. Sie lauten dann:

$$\forall h,i,k \quad \exists j: \qquad A_hA_iA_j\circ \;\supseteq\; A_hA_iB_k\circ \tag{6.82}$$

$$\forall h,i,k \quad \exists j: \qquad \circ B_jB_i \;\supseteq\; A_h\circ A_kB_i \tag{6.83}$$

$$\{(\bigsqcup_{i=1}^{p} B_i)^d\circ\}^{\boldsymbol{\cdot}} \;\subseteq\; \{(\bigsqcup_{j=1}^{m} A_j)^2\circ\}^{\boldsymbol{\cdot}} \tag{6.84}$$

$$\{(\bigsqcup_{h=1}^{p} A_h)\circ(\bigsqcup_{i=1}^{p} A_i)^d\}^{\boldsymbol{\cdot}} \;\subseteq\; \{\circ(\bigsqcup_{j=1}^{n} B_j)\}^{\boldsymbol{\cdot}} \;. \tag{6.85}$$

Die erste und die dritte Voraussetzung werden bei dieser Zerlegung ein wenig schwächer. Die zweite und vierte Voraussetzung werden deutlich schwächer. Keine Voraussetzung wird schärfer, so daß die Zerlegung von expandierenden Bündeln schon erfüllte Voraussetzungen nicht verletzen kann. Analoges gilt für die Zerlegung von kontrahierenden Bündeln, nur daß es hierbei die erste und die dritte Voraussetzung sind, die deutlich schwächer werden.

Diese Zerlegungen können wiederholt werden. Wenn man die expandierenden Bündel r–mal zerlegt, dann werden die zweite und die vierte Voraussetzung zu:

$$\{\circ(\bigsqcup_{j=1}^{n} B_j)^2\}^{\boldsymbol{\cdot}} \;\supseteq\; \{(\bigsqcup_{h=1}^{p} A_h)^r\circ(\bigsqcup_{k=1}^{p} A_k)(\bigsqcup_{i=1}^{n} B_i)\}^{\boldsymbol{\cdot}} \qquad \text{und} \tag{6.86}$$

$$\{(\bigsqcup_{h=1}^{p} A_h)^r\circ(\bigsqcup_{i=1}^{p} A_i)^d\}^{\boldsymbol{\cdot}} \;\subseteq\; \{\circ(\bigsqcup_{j=1}^{n} B_j)\}^{\boldsymbol{\cdot}} \;. \tag{6.87}$$

Da die Sequenzlängen r und d beliebig groß gewählt werden können, besagt die Voraussetzung 6.87 in dieser abgeschwächten Form nichts anderes, als daß die durch endliche A_i–Sequenzen definierten Gebiete den Phasenraum mindestens ebenso fein unterteilen wie die Symbolischen Gebiete $B_i\circ$. Dann kann die A_i–Partition nicht mehr Orbits verwechseln als die B_j–Partition. Die andere Voraussetzung 6.86 werden wir weiter unten graphisch interpretieren. Dabei wird es zumindest plausibel werden, daß auch diese Voraussetzung sich erfüllen läßt.

Diese Plausibilitätsargumente ergeben noch keinen Beweis. Wer ganz sicher gehen will, letztlich eine nicht zwiefache Partition rekonstruieren zu können, muß daher die vier Voraussetzungen 6.74, 6.75, 6.79 und 6.80 schon während der Vereinfachung der zwiefachen Partition beachten. Die anfängliche zwiefache Partition $\{\tilde{C}_1\circ,\ldots\tilde{C}_{\tilde{p}}\circ\}$, $\{\circ\tilde{C}_1,\ldots\circ\tilde{C}_{\tilde{p}}\}$ erfüllt diese Voraussetzung trivialerweise wegen $\circ A_i = \circ\tilde{C}_i = \circ B_i$. Bei jedem Vereinfachungsschritt müssen die vier Voraussetzungen dann als Zusatzbedingungen beachtet werden. Gegebenenfalls müssen expandierende oder kontrahierende Bündel zerlegt werden, um diese Zusatzbedingungen zu erfüllen. Und wenn endlich viele Zerlegungen nicht genügen, dann muß man auf den Vereinfachungsschritt selbst verzichten.

In der Beispielrechnung, in der die AB–Partition vereinfacht wurde, treten diese Schwierigkeiten nicht auf; die Zusatzbedingungen sind jederzeit erfüllt. Und es gelang mir auch nicht, ein Beispiel zu finden, in dem man wegen dieser Zusatzbedingungen auf einen Vereinfachungsschritt verzichten müßte. Im Lichte der folgenden Plausibilitätsargumente überrascht dies nicht: Die Symbolischen Gebiete $A_i\circ$ und $\circ B_k$ müßten schon eine außergewöhnliche Gestalt haben, damit ein Vereinfachungsschritt durch die Zusatzbedingungen verhindert wird.

Wenn die vereinfachte zwiefache Partition dann schließlich alle vier Zusatzbedingungen erfüllt, dann kann aus ihr eine nicht zwiefache Partition mit endlich vielen Symbolen rekonstruiert werden. Wie viele Symbole notwendig sind, erkennt man aber erst bei der Rekonstruktion.

6.5.3 Geometrische Interpretation der Voraussetzungen

Die Voraussetzungen 6.83

$$\forall i,k,l \quad \exists j: \qquad \circledast B_j B_i \quad \supseteq \quad A_k \circledast A_l B_i \tag{6.88}$$

und die analoge Voraussetzung

$$\forall i,k,l \quad \exists j: \qquad A_i B_l \circledast B_k \quad \subseteq \quad A_i A_j \circledast \tag{6.89}$$

lassen sich geometrisch interpretieren. Dabei wird klar werden, wozu sie dienen. In Abb. 6.5 sind zwei expandierende Blätter α und β eingetragen, die mit den beiden kontrahierenden Blättern γ und δ Orbitpaare bilden. Die Bilder $f(\alpha)$ und $f(\beta)$ sind daher nicht selbst generierend. Sie bestehen aus den kleineren expandierenden Blättern α_1 und α_2 bzw. β_1 und β_2. Ebenso gilt $f^{-1}(\gamma) = \gamma_1 \cup \gamma_2$ und $f^{-1}(\delta) = \delta_1 \cup \delta_2$, wobei γ_1, γ_2, δ_1 und δ_2 kontrahierend sind. In Abb. 6.5.a sind diese Blätter auf eine natürliche Weise geordnet, so daß α_1 und β_1 nur $f(\gamma_1)$ und $f(\delta_1)$ schneiden und nicht $f(\gamma_2)$ und $f(\delta_2)$. Umgekehrt schneiden α_2 und β_2 nur $f(\gamma_2)$ und $f(\delta_2)$. In Abb. 6.5.b sind die Blätter dagegen auf eine so komplizierte Weise verschlungen, daß eine solche Ordnung wie in Abb. 6.5.a nicht erreicht werden kann, auch dann nicht, wenn die Indices 1 und 2 von manchen Blättern vertauscht werden.

Diese Verschlingungen stören bei der Rekonstruktion einer nicht zwiefachen Partition. In Abb. 6.5.a genügen zwei Symbole, die 1 und 2 genannt werden sollen, um die miteinander gepaarten Orbits zu unterscheiden. Man muß dazu die Symbolischen Gebiete $\circledast 1$ und $\circledast 2$ so wählen, daß $\circledast 1 \cap \circledast 2 = \emptyset$, $\alpha_1 \in 1\circledast$, $\alpha_2 \in 2\circledast$, $\gamma_1 \in \circledast 1$ und $\gamma_2 \in \circledast 2$. Die linken Schnittpunkte von Abb. 6.5.a liegen dann bei dem unteren Zeitpunkt alle in $\circledast 1$, die rechten Schnittpunkte alle in $\circledast 2$, so daß die gepaarten Orbits zu diesem Zeitpunkt verschiedene Symbole 1 und 2 haben. Dagegen ist in Abb. 6.5.b eine solch einfache Zuordnung nicht möglich.

Als Zweck der Bed. 6.89 erweist sich nun, Verschlingungen wie in Abb. 6.5.b auszuschließen, aber nicht unbedingt für alle Blätter, sondern nur für Blätter α und β, die in demselben expandierenden Bündel liegen, und Blätter γ und δ, die in demselben kontrahierenden Bündel liegen. Um dies zu sehen, muß man die Indices i, k, l und j so wählen, daß

$$\alpha,\beta \quad \in \quad \{(\bigcup_{h=1}^{m} A_h)^{-\infty} A_i \circledast\}' \tag{6.90}$$

$$\alpha_1,\beta_1 \quad \in \quad \{(\bigcup_{h=1}^{m} A_h)^{-\infty} A_i A_j \circledast\}' \tag{6.91}$$

$$\gamma,\delta \quad \in \quad \{\circledast B_k (\bigcup_{h=1}^{n} B_h)^{+\infty}\}' \tag{6.92}$$

$$f(\gamma_1), f(\delta_1) \quad \in \quad \{B_l \circledast B_k (\bigcup_{h=1}^{n} B_h)^{+\infty}\}' \quad . \tag{6.93}$$

Aus der Bed. 6.89 folgt dann, daß die Blätter $f(\gamma_1)$ und $f(\delta_1)$ keinen anderen Teil von $f(\alpha)$ und $f(\beta)$ außer α_1 und β_1 schneiden dürfen. Aus der anderen Bedingung 6.88 folgt analog, daß die Blätter α_1 und β_1 nur die Blätter $f(\gamma_1)$ und $f(\delta_1)$ und keinen anderen Teil von γ oder δ schneiden dürfen. Verwicklungen wie in Abb. 6.5.b sind damit ausgeschlossen.

Die Zerlegungen der expandierenden oder kontrahierenden Bündel dienen letztlich dazu, solche Verschlingungen wie in Abb. 6.5.b aufzulösen, das heißt dafür zu sorgen, daß α und β nicht mehr zu demselben expandierenden Bündel oder γ und δ nicht mehr zu demselben kontrahierenden Bündel gehören. Normalerweise sollten endlich viele Zerlegungen genügen, um diese Verschlingungen aufzulösen. Denn wenn die Symbolischen Gebiete $A_i\circledast$ und $\circledast B_k$ nicht eine außergewöhnliche Gestalt haben, dann entstehen durch die Zerlegung kleine expandierende und kontrahierende Bündel, die aus nahe benachbarten und ungefähr parallelen Blättern bestehen. Solche Blätter können sich nicht wie in Abb. 6.5.b verschlingen. Vor diesem Hintergrund wird es auch verständlich, warum in

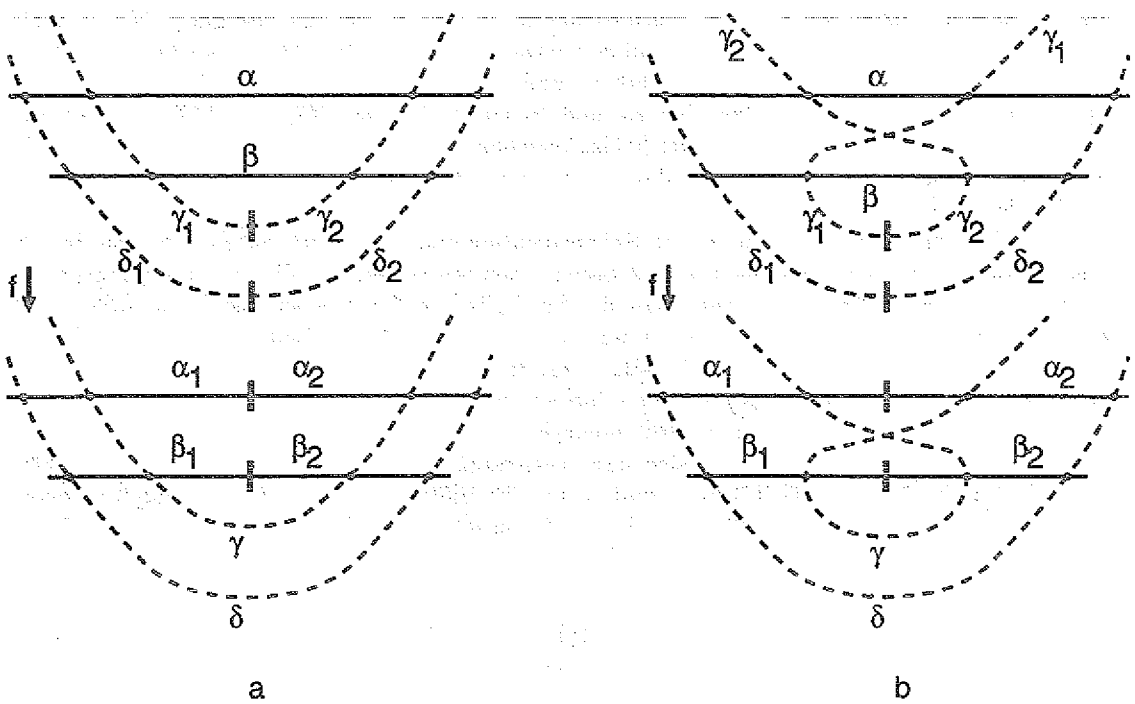


Abbildung 6.5: Schwierigkeit bei der Rekonstruktion einer nicht zwiefachen Partition. Die beiden Schemata sind ähnlich aufgebaut wie die in Abb. 5.8. Die expandierenden Blätter α, β, α_1, β_1, α_2 und β_2 sind als durchgezogene Linie eingetragen. Die kontrahierenden Blätter γ, δ, γ_1, δ_1, γ_2 und δ_2 sind als gestrichelte Linien eingetragen. Die kurzen senkrechten Striche begrenzen diese Blätter. In Abb. a schneidet das Blatt $f(\gamma_1)$ dieselben expandierenden Blätter α_1 und β_1 wie das Blatt $f(\delta_1)$. In Abb. b existiert keine derartige Ordnung, was die Rekonstruktion einer zwiefachen Partition erschwert.

der Beispielrechnung die Voraussetzungen 6.88 und 6.89 für die Rekonstruktion einer nicht zwiefachen Partition jederzeit erfüllt war, ohne daß bei den einzelnen Vereinfachungsschritten explizit darauf geachtet wurde.

6.6 Vereinigen von Bündeln

Im letzten Abschnitt wurden expandierende oder kontrahierende Bündel zerlegt. Auch das Gegenteil, das Vereinigen von expandierenden oder kontrahierenden Bündeln, ist eine mögliche Rechenmethode. Je größere Bündel man hat, desto mehr Blätter kann man auf einmal in der Zeit verschieben. Und desto weniger Bündel müssen bei der Vereinfachung der zwiefachen Partition verschoben werden. Durch das Vereinigen von expandierenden oder kontrahierenden Bündeln kann daher der Rechenaufwand verringert werden. Dabei ist aber die Warnung aus Kapitel 5.2.2 zu berücksichtigen: Wenn man zu große Bündel wählt, dann kann man mögliche Vereinfachungen übersehen. Auch Möglichkeiten, eine nicht zwiefache Partition zu rekonstruieren, können auf diese Weise übersehen werden. Ganz ausschließen kann ich diese Risiken bei der Vereinigung von Bündeln nicht. Aber durch bestimmte Kriterien können die Risiken zumindest klein gehalten werden. Diese heuristischen Kriterien will ich im folgenden definieren, begründen und an einem Beispiel vorführen.

6.6.1 Welche Bündel werden vereinigt?

Die Auswahlkriterien

Nach welchen Kriterien soll man die Bündel auswählen, die vereinigt werden? Betrachten wir hierzu zwei expandierende Bündel Λ und Γ. Zwei Eigenschaften von Γ waren für die Rechnung in den bisherigen Kapitel relevant:

- die kontrahierenden Bündel, mit denen Γ Orbitpaare bildet, und
- die Symbolischen Gebiete, in die das Gebiet $[\Gamma] = \bigcup_{\alpha \in \Gamma} \alpha$ durch die Dynamik f abgebildet wird.

Je mehr kontrahierende Bündel mit Γ Orbitpaare bilden und je mehr zeitliche Übergänge von $[\Gamma]$ ausgehen, desto schwieriger ist die Vereinfachung der zwiefachen Partition oder die Rekonstruktion der nicht zwiefachen Partition. Aus diesem Grund empfiehlt es sich, zwei Bündel Λ und Γ nur dann zu vereinigen, wenn die Vereinigung $\Lambda \cup \Gamma$ nicht mit mehr kontrahierenden Bündeln Orbitpaare bildet und nicht mehr Übergänge aufweist als eines der ursprünglichen Bündel, sagen wir Γ. Es ist aber durchaus sinnvoll, auch ein weniger komplexes Bündel Λ mit dem Bündel Γ zu vereinigen. Das heißt, das andere Bündel Λ darf mit weniger Bündeln Orbitpaare bilden und weniger Übergänge aufweisen als Γ. Die Auswahlkriterien für zu vereinigende Bündel Λ und Γ lauten deshalb:

- Das Symbolische Gebiet $[\Lambda]$ besitzt nur Übergänge zu den Symbolischen Gebieten $A_i \circ$, zu denen auch Übergänge von $[\Gamma]$ aus führen, und
- das Bündel Λ bildet Orbitpaare nur mit den kontrahierenden Bündeln, mit denen auch Γ Orbitpaare bildet.

Nach analogen Kriterien kann man auch die kontrahierenden Bündel aussuchen, die man vereinigt.

Beispiele

Bei der Vereinfachung der AB–Partition in Kap. 5 wurden zu Beginn expandierende und kontrahierende Bündel angemessener Größe gewählt. Diese Bündel erfüllen die eben postulierten Kriterien nur teilweise und können daher nicht vereinigt werden.

Zum Beispiel besitzen die expandierenden Bündel $\{(A \sqcup B)^{-\infty} BA\circ\}$' und $\{(A \sqcup B)^{-\infty} AA\circ\}$' in Abb. 5.4 Übergänge zu den gleichen Bündeln, nämlich $\{(A \sqcup B)^{-\infty} AB\circ\}$' und $\{(A \sqcup B)^{-\infty} AA\circ\}$'.

Sie bilden aber in Tabelle 5.3 Orbitpaare mit verschiedenen kontrahierenden Bündeln und sollten daher nicht zum expandierenden Bündel $\{(A \sqcup B)^{-\infty} A\circ\}$' vereinigt werden. In der Tat war es für die Vereinfachung der Partition wichtig, diese beiden Bündel zeitlich gegeneinander zu verschieben.

Ein weiteres Beispiel ist das expandierende Bündel $\{(A \sqcup B)^{-\infty} BA\circ\}$' in Tab. 5.3, das nur mit solchen kontrahierenden Bündeln Orbitpaare bildet, mit denen auch das expandierende Bündel $\{(A \sqcup B)^{-\infty} BB\circ\}$' Orbitpaare bildet. Aber auch diese Bündel sollten nicht miteinander vereinigt werden, denn in Abb. 5.4 (oder Abb. 5.5) führen die Übergänge von ihnen aus zu verschiedenen expandierenden Bündeln. Letztlich lassen sich keine der drei in Frage kommenden expandierenden Bündel $\{(A \sqcup B)^{-\infty} AA\circ\}$', $\{(A \sqcup B)^{-\infty} BA\circ\}$' und $\{(A \sqcup B)^{-\infty} BB\circ\}$' nach den postulierten Kriterien zu größeren Bündeln vereinigen. Dasselbe gilt für die kontrahierenden Bündel in Tabelle 5.3 und Abb. 5.6 (oder Abb. 5.7). In der Tat würde das Vereingen dieser Bündel auch dazu führen, daß die Beispielrechnung in Kap. 5 nicht mehr möglich wäre.

Anmerkungen

Dagegen erlauben es die Auswahlkriterien, die expandierenden Bündel $\{(A \sqcup B)^{-\infty} AB\circ\}$' und $\{(A \sqcup B)^{-\infty} BB\circ\}$' zu vereinigen. Dies würde den weiteren Verlauf der Beispielrechnung tatsächlich nicht blockieren, aber auch nicht wesentlich verkürzen.

Im allgemeinen Fall empfehle ich, noch zwei weitere Kriterien zu berücksichtigen. Das eine Kriterium betrifft den Fall, daß von dem Symbolischen Gebiet [A] nur ein Übergang ausgeht. Die expandierenden Blätter in A sind also alle Urbilder von expandierenden Blättern. Sie bilden keine Orbitpaare und werden im nächsten Zeitschritt nicht in kleinere expandierende Blätter unterteilt. Ein solches Bündel A, wie zum Beispiel das eben erwähnte Bündel $\{(A \sqcup B)^{-\infty} AB\circ\}$', läßt sich nach den obigen Auswahlkriterien besonders leicht mit anderen Bündeln Γ vereinigen. Doch dadurch wird die Rechnung nicht unbedingt einfacher. Denn die Methoden von Kap. 5 operieren im wesentlichen nur auf den anderen expandierenden Bündeln, die Blätter enthalten, welche im nächsten Zeitschritt geteilt werden. Es ist daher sinnvoll, zunächst die Bündel zu vereinigen, deren Blätter geteilt werden. Wie wir weiter unten sehen werden, wird durch diese Vereinigungen gleichzeitig festgelegt, was mit Bündeln geschieht, deren Blätter nicht geteilt werden.

Das andere Kriterium ist wichtig, wenn nach der Vereinfachung der zwiefachen Partition eine nicht zwiefache Partition rekonstruiert werden soll. Es ist sinnlos, Bündel zu vereinigen, wenn die Vereinigung wieder zerlegt werden muß, um die Voraussetzungen für die Rekonstruktion zu erfüllen. Man sollte daher bei jeder Vereinigung von Bündeln darauf achten, daß die vier Voraussetzungen 6.74, 6.75, 6.79 und 6.80 nicht verletzt werden. Bei den folgenden Beispielen sind diese Voraussetzungen für die Rekonstruktion einer zwiefachen Partition immer erfüllt.

6.6.2 Vereinigung von Bündeln während der Vereinfachung der AB–Partition

In der Beispielrechnung von Kap. 5 wurde die AB–Partition vereinfacht. Wie eben schon bemerkt, können die anfänglichen expandierenden oder kontrahierenden Bündel kaum vereinigt werden. Doch im weiteren Verlauf der Beispielrechnung entstanden Bündel, die alle Kriterien erfüllen, um vereinigt zu werden. Diese Beispiele sollen im folgenden skizziert werden.

Erstes Beispiel: Vereinigung kontrahierender Bündel

Der erste Vereinfachungsschritt in der Beispielrechnung führte zu den Orbitpaaren von Tab. 5.13 und den Übergängen von Abb. 5.12. Nun gibt es zwei kontrahierende Bündel, die sich vereinigen lassen, nämlich $\Lambda = \{\circ BA(A \sqcup B)^{+\infty}\}$' und $\Gamma = \{\circ EBB(A \sqcup B)^{+\infty}\}$'. Denn in Tab. 5.13 bildet Λ nur mit dem kontrahierenden Bündel $\{(A \sqcup B)^{-\infty} BB\circ\}$' Orbitpaare, mit dem auch Γ Orbitpaare bildet. Und in Abb. 5.12 (die in Abb. 6.4 enthalten ist) führen von $[\Lambda] = \circ BA$ zwei verschiedene Übergänge zu expandierenden Bündeln, deren Blätter geteilt werden, und zwar jeweils in zwei Zeitschritten:

$$\circ BA \to \circ BBA \to \circ EBB \qquad \text{und} \qquad \circ BA \to \circ ABA \to \circ BA \quad . \tag{6.94}$$

Beide Übergänge sollen mit Übergängen identifiziert werden, die von $[\Gamma] = \circledast EBB$ ausgehen. Der erste Übergang kann wegen seines Ziels nur mit

$$\circledast EBB \to \circledast EBB \to \circledast EBB \tag{6.95}$$

identifiziert werden, der zweite Übergang nur mit

$$\circledast EBB \to \circledast AA \to \circledast BA \quad . \tag{6.96}$$

Um die Bündel Λ und Γ vereinigen zu können, muß man deshalb zunächst die Zwischenstufen $\circledast BBA$ und $\circledast EBB$ bzw. $\circledast ABA$ und $\circledast AA$ miteinander identifizieren. Das heißt, man muß das Bündel $\{\circledast ABA(A \sqcup B)^{+\infty}\}$' mit $\{\circledast AA(A \sqcup B)^{+\infty}\}$' vereinigen, sowie das Bündel $\{\circledast BBA(A \sqcup B)^{+\infty}\}$' mit $\{\circledast EBB(A \sqcup B)^{+\infty}\}$'. Durch die Vereinigung der Bündel Λ und Γ, deren Blätter im nächsten Zeitschritt geteilt werden, ist also gleichzeitig festgelegt, was mit den Bündeln $\{\circledast ABA(A \sqcup B)^{+\infty}\}$' und $\{\circledast BBA(A \sqcup B)^{+\infty}\}$' geschieht, deren Blätter im nächsten Zeitschritt nicht geteilt werden. Nach dieser Vereinigung bleiben nur noch zwei kontrahierende Bündel übrig:

$$\{(\circledast EBB \sqcup \circledast BBA \sqcup \circledast BA(A \sqcup B))(A \sqcup B)^{+\infty}\}' \quad \text{und} \tag{6.97}$$

$$\{(\circledast AA(A \sqcup B) \sqcup \circledast ABA)(A \sqcup B)^{+\infty}\}' \quad . \tag{6.98}$$

Von den Beschriftungen 0 oder 1 der Übergänge in Abb. 6.4.b liest man leicht ab, daß diese beiden Bündel gleich $\{\circledast 1(0 \sqcup 1)^{+\infty}\}$' bzw. $\{\circledast 0(0 \sqcup 1)^{+\infty}\}$' sind. Daher kann die 01-Partition auch noch nach dieser Vereinigung expandierender Bündel rekonstruiert werden.

Zweites Beispiel: Vereinigung expandierender Bündel

Der zweite Vereinfachungsschritt in der Beispielrechnung führte zu den Orbitpaaren von Tab. 5.17 und den Übergängen von Abb. 5.16, die Abb. 6.4 noch einmal zeigte. Nun lassen sich auch expandierende Bündel vereinigen. Zuerst können die Bündel $\Lambda_1 = \{(A \sqcup B)^{-\infty}(BAA \circledast B \cup BAB \circledast A)\}$' und $\Gamma_1 = \{(A \sqcup B)^{-\infty} BB \circledast\}$' vereinigt werden. Dabei wird der Übergang $\Lambda_1 \to AAB \circledast \to BB \circledast$ mit $\Gamma_1 \to BB \circledast \to BB \circledast$ identifiziert und der Übergang $\Lambda_1 \to BA \circledast$ mit $\Gamma_1 \to BA \circledast$. Die Vereinigung ergibt das expandierende Bündel $\Sigma = \{(A \sqcup B)^{-\infty}((A \sqcup B)BB \circledast E \sqcup AAB \circledast E \sqcup (BAA \circledast B \cup BAB \circledast A))\}$'. Schließlich kann auch das Bündel $\Lambda_2 = \{(A \sqcup B)^{-\infty} AA \circledast A\}$' mit $\Gamma_2 = \{(A \sqcup B)^{-\infty} BA \circledast\}$' vereinigt werden, wobei $\{(A \sqcup B)^{-\infty} AAA \circledast B\}$' zu Σ hinzukommt. Als Endergebnis erhält man zwei expandierende Bündel:

$$\{(A \sqcup B)^{-\infty}(BA \circledast E \sqcup AA \circledast A)\}' \quad \text{und} \tag{6.99}$$

$$\{(A \sqcup B)^{-\infty}((A \sqcup B)BB \circledast E \sqcup AAB \circledast E \sqcup AAA \circledast B \sqcup (BAA \circledast B \cup BAB \circledast A))\}' \tag{6.100}$$

Diese beiden Bündel sind wiederum gleich $\{(0 \sqcup 1)^{-\infty} 0 \circledast\}$' bzw. $\{(0 \sqcup 1)^{-\infty} 1 \circledast\}$'. Nach diesen Vereinigungen der expandierenden und kontrahierenden Bündel ist es trivial, die 01-Partition zu rekonstruieren.

Mensch werde wesentlich,
denn wann die Welt vergeht,
so fällt der Zufall weg,
das Wesen das besteht.
aus Angelus Silesius:
„Cherubinischer Wandersmann"

Kapitel 7

Diskussion der schrittweisen Vereinfachung

7.1 Wie eindeutig ist das Ergebnis der schrittweisen Vereinfachung?

7.1.1 Welche Schritte der Rechnung sind zwangsläufig?

Die Suche nach einer einfachen generierenden Partition besteht, so wie sie hier vorgeschlagen wurde, aus drei aufeinanderfolgenden Abschnitten.

Die Wahl eines Ausgangspunktes ist, wie bei anderen Gradientensuchverfahren auch, ziemlich willkürlich. Zu dieser Wahl gehört nicht nur die Wahl der anfänglichen Partition, die die Größe der anfänglichen expandierenden und kontrahierenden Blätter bestimmt. Auch die Wahl der anfänglichen Bündelgrößen ist wichtig. Wenn diese Bündel zu groß gewählt werden, dann können mögliche Vereinfachungen übersehen werden.

Die schrittweise Vereinfachung der zwiefachen Partition läuft nach festen Regeln ab. Die Orbitpaare bestimmen, welche Bündel zeitlich verschoben werden können. Die Reihenfolge dieser Vereinfachungsschritte ist allerdings nicht völlig festgelegt. Wenn mehrere dieser Vereinfachungsschritte gleichzeitig möglich sind, besteht eine gewisse Willkür darin, einen der Vereinfachungsschritte auszuwählen und zu vollziehen.

Die Rekonstruktion einer nicht zwiefachen Partition, die nicht mehr Orbits verwechselt als die ursprüngliche Partition, läuft auch nach festen Regeln ab. Allerdings weiß man von vorneherein nicht, wie viele Symbole genügen, so daß man möglicherweise mehrere Versuche mit immer mehr Symbolen starten muß. Es kann dabei passieren, daß mehrere nicht zwiefache Partitionen gefunden werden, die dieselbe Zahl von Symbolen besitzen und dieselben expandierenden und kontrahierenden Blätter erzeugen. Solche Partitionen sind zum Bespiel die 01–Partition mit überlappenden Symbolischen Gebieten von Kap. 6.3.4 und die normale 01–Partition mit disjunkten Symbolischen Gebieten.

7.1.2 Eine weitere „einfachste" Partition

In der Beispielrechnung von Kap. 5 gab es nur eine Stelle, an der wir die Wahl zwischen verschiedenen Vereinfachungsschritten hatten. Die beiden Bündel $\{(A \sqcup B)^{-\infty} AA \circ\}$' und $\{\circ BB(A \sqcup B)^{+\infty}\}$', die im Abschnitt 5.4.7 einen Zeitschritt zurück verschoben wurden, können statt dessen auch einen Zeitschritt vor verschoben werden. Man erhält das neue expandierende Bündel $\{(A \sqcup B)^{-\infty} AAE \circ\}$' und das neue kontrahierende Bündel $\{B \circ B(A \sqcup B)^{+\infty}\}$'. Die weitere Rechnung verläuft ähnlich

wie in der Beispielrechnung von Kap. 5 und 6. Das Endergebnis ist eine binäre Partition $\{\circledast X, \circledast Y\}$.

$$\circledast X = \circledast A \cup AA \circledast B \tag{7.1}$$
$$\circledast Y = B \circledast B \cup BA \circledast B \tag{7.2}$$

Wegen $\circledast AABA = \circledast BABB = \emptyset$ erweisen sich die Symbolsequenzen $\circledast YXYY = \circledast YXXY = \emptyset$ der XY–Partition als unzulässig. Weil die XY–Partition nicht mehr Orbits verwechselt als die AB–Partition, können die XY–Symbolsequenzen auch rückübersetzt werden.

$$\circledast A = Y \circledast X \cup X \circledast XX \tag{7.3}$$
$$\circledast B = \circledast Y \cup X \circledast XY \tag{7.4}$$

Durch Einsetzen der Regeln 4.34, 4.35, 4.32 und 4.33, die 01–Sequenzen in AB–Sequenzen übersetzen und umgekehrt, kann man 01–Sequenzen auch in XY–Sequenenzen übersetzen und umgekehrt.

$$\circledast X = \circledast 0 \cup 01 \circledast 1 \cup 0 \circledast 11 \tag{7.5}$$
$$\circledast Y = 0 \circledast 10 \cup 11 \circledast 1 \tag{7.6}$$
$$\circledast 0 = Y \circledast X \cup X \circledast XXX \tag{7.7}$$
$$\circledast 1 = \circledast Y \cup X \circledast XY \cup X \circledast XXY \tag{7.8}$$

Abb. 7.1 zeigt diese XY–Partition. Die Partitionsgrenzen schneiden dieselben Orbits wie die Partitionsgrenzen der 01–Partition, nur zu anderen Zeiten. Auch sonst kenne ich kein koordinatenunabhängiges Kriterium der Einfachheit, dem zufolge die 01–Partition entweder einfacher oder komplizierter wäre als die XY–Partition. Sogar die intuititve Vorstellung von Strecken und Falten läßt sich auf die XY–Partition anwenden, wie Kap. 8.2.4 zeigen wird. Nur wegen der speziellen Wahl der Koordinatenachsen in Abb. 7.1 ist das Falten etwas weniger anschaulich als bei der 01–Partition.

Daß in vielen Veröffentlichungen über die Symbolische Dynamik des Hénon Attraktors bei Standardparameterwerten diese XY–Partition immer übersehen wurde und daß ich sie erst fand, nachdem ich nur noch mit Symbolen statt mit geometrischen Konzepten rechnete, mag als Warnung dienen, nicht zu sehr auf die geometrische Intuition zu vertrauen.

Die Übersetzungsregeln 7.5 und 7.6 sowie die grammatikalischen Regeln dieser XY–Partition des Hénon Attraktors bei Standardparameterwerten ähneln denen der XY–Partition des Hénon Attraktors bei schwacher Dissipation, die in Kap. 2.3.3, Gl. 2.10 und Gl. 2.11, eingeführt und in Kap. 4.2.1 und Anhang C.1 diskutiert wurden. Bei Standardparameterwerten ist allerdings die Sequenz $\circledast XXXXYX \neq \emptyset$ nicht verboten, so daß es Orbitpaare gibt, die von der XY–Partition, aber nicht von der 01–Partition erzeugt werden. Diese Orbitpaare werden am Ende von Anhang C.1 in Gl. C.51 berechnet. Abgesehen von der Lage des Punktes $\circledast$ haben sie die Form:

$$\overline{\ldots S_{-4}XXX} \circledast \overline{XXYXS_4\ldots} = \overline{\ldots T_{-4}000} \circledast \overline{11\bar{1}0T_4\ldots} \tag{7.9}$$
$$\underline{\ldots S_{-4}XXX} \circledast \underline{YXYXS_4\ldots} = \underline{\ldots T_{-4}011} \circledast \underline{10\bar{1}0T_4\ldots}\ , \tag{7.10}$$

wobei die horizontalen Striche wie in Gleichung 4.29 die Bereiche gleicher Symbole $T_i \in \{0,1\}$ und $S_i \in \{X,Y\}$ kennzeichnen. Umgekehrt gibt Anhang C.1 auch Orbitpaare der 01–Partition an, die nicht gleichzeitig Orbitpaare der XY–Partition sind. Keine der beiden Partitionen kann daher nach Kriterium 4.1.2 beanspruchen, die einfachere zu sein.

7.2 Mögliche Schwierigkeiten

Wenn man diese Methode auf andere dynamische Systeme anwenden will, dann können Schwierigkeiten auftauchen, die bei der Beispielrechnung von Kap. 5 und 6 nicht auftraten. Für die Schwierigkeiten, die mir auffielen, gibt es mehrere mögliche Lösungsansätze.

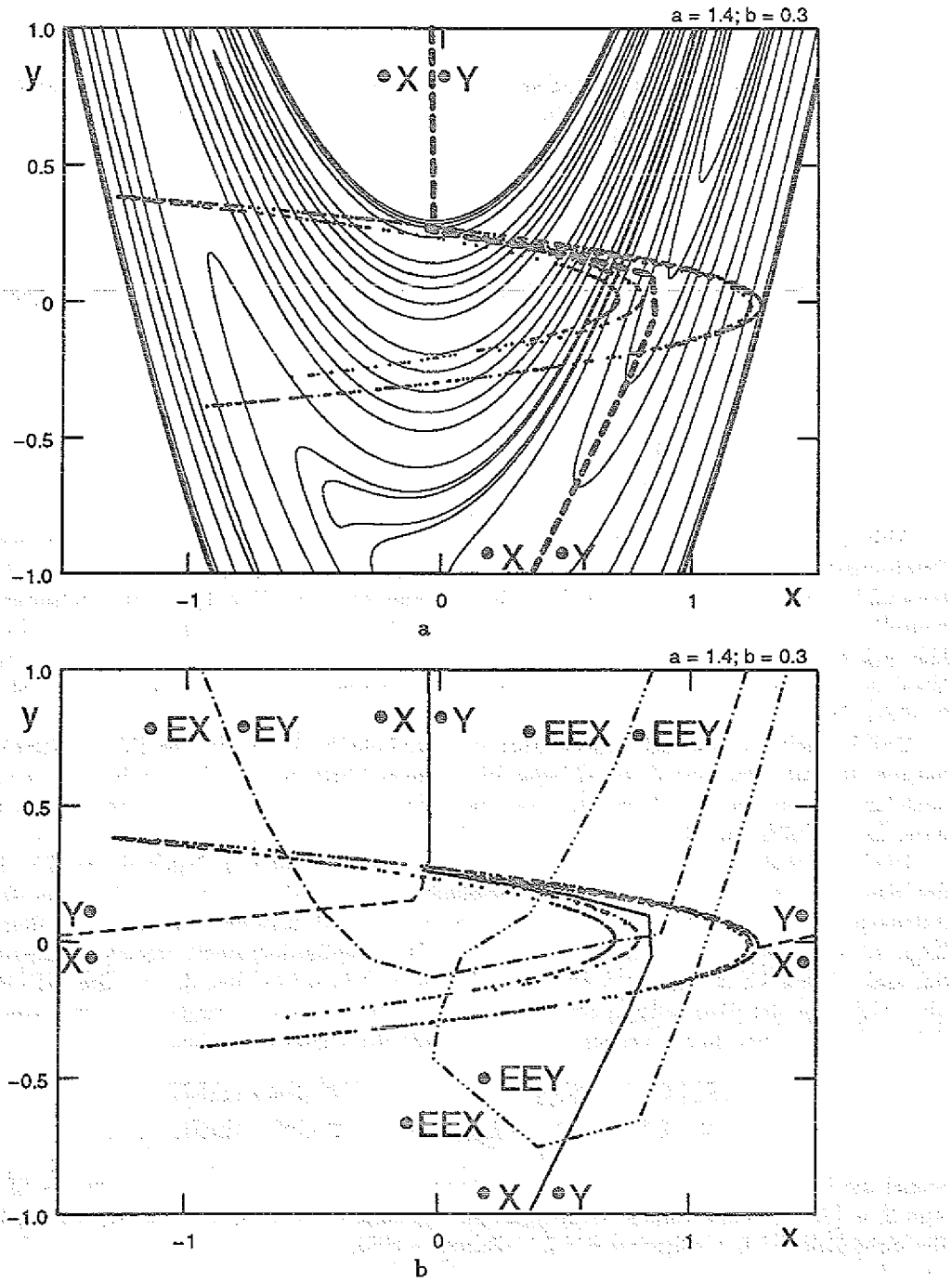


Abbildung 7.1: Die XY–Partition des Hénon Attraktors bei Standardparameterwerten. Sie ist anders, aber nicht unbedingt komplizierter als die 01–Partition in Abb. 2.4 und verwechselt auch nicht mehr Orbits als diese. In Abb. a sind die Attraktorpunkte und die Abschnitte der stabilen Mannigfaltigkeit wie in Abb. 2.4 eingetragen. Abb. b zeigt die Partitionsgrenzen bei drei aufeinanderfolgenden Zeiten. Sie trennen $\circ EX$ von $\circ EY$, $\circ X$ von $\circ Y$, bzw. $X\circ$ von $Y\circ$.

7.2.1 Anforderungen an einen Algorithmus

Der hauptsächliche Rechenaufwand liegt nicht in der Wahl der anfänglichen zwiefachen Partition, sondern in ihrer schrittweisen Vereinfachung und in der Rekonstruktion einer nicht zwiefachen Partition. Dieser Hauptteil der Rechnung läuft nach festen Regeln ab. Es liegt deshalb nahe, ihn einem Computer einzuprogrammieren. Wie Kap. 9 argumentieren wird, ist ein neuronales Netz für diese Aufgabe angemessener als ein traditionelles Computerprogramm. Trotzdem ist es interessant zu sehen, ob die hier vorgeschlagenen Rechenregeln den Anforderungen genügen, die man an Algorithmen stellt. Manche diese Anforderungen wurden bereits diskutiert, wie zum Beispiel die Forderung, daß endlich viele Symbole zur Rekonstruktion einer nicht zwiefachen Partition genügen. Es bleiben noch zwei wesentliche Anforderungen zu erfüllen: Die Reihenfolge der Vereinfachungsschritte muß eindeutig festgelegt sein und die Abfolge der Vereinfachungsschritte muß irgendwann an ein Ende führen.

Die Reihenfolge der Vereinfachungsschritte kann auf zwei Weisen festgelegt werden. Man kann durch recht willkürliche Zusatzregeln eine bestimmte Reihenfolge auswählen, auf die Gefahr hin, eine der „einfachsten" Partitionen zu übersehen. Oder man kann nacheinander alle möglichen Vereinfachungswege ausprobieren, was vor allem dann aufwendig wäre, wenn man mit einer feinen Partition aus vielen Symbolischen Gebieten beginnen würde. Diese Alternativen und ihre Nachteile sind von anderen numerischen Optimierungsverfahren her vertraut.

Das Ende der Vereinfachungsschritte ist nicht garantiert, wenn dieselben Bündel immer wieder in der Zeit vor und zurück verschoben werden dürfen, ohne daß neue Orbitpaare entstehen. Wenn man darauf achtet, daß auch wirklich neue Orbitpaare entstehen, dann sind solche trivialen Endlosschleifen zwar ausgeschlossen, ein Ende der Rechnung aber noch nicht unbedingt garantiert. Denn die Zahl der betrachteten Bündel kann zunehmen, wie zum Beispiel beim Übergang von Tab. 5.13 zu Tab. 5.17, so daß im Prinzip immer neue Orbitpaare durch immer kleinere Bündel gebildet werden könnten. Um ein Rechenende zu garantieren, sollte man deshalb zweitens die Zahl der betrachteten Bündel begrenzen, wodurch auch die Rechnung weniger aufwendig wird. Dies ist durch das Vereinigen von Bündeln möglich, das in Kap. 6.6 diskutiert wurde.

Man kann daher im Prinzip aus der hier vorgeschlagenen Methode einen Computeralgorithmus konstruieren. Dies ist aber vermutlich sehr aufwendig, da die verwendeten Konzepte neu sind und nicht leicht mit Hilfe der gängigen Programmiersprachen implementiert werden können.

7.2.2 Erratische expandierende und kontrahierende Blätter

Eine weitere Schwierigkeit kann auftreten, weil ich, um die vorgeschlagene Methode möglichst flexibel und leicht handhabbar zu machen, nur wenige Bedingungen an die expandierenden und kontrahierenden Blätter gestellt habe. Im Prinzip kann jede Teilmenge einer instabilen Mannigfaltigkeit als expandierendes Blatt deklariert werden, auch wenn sie völlig unzusammenhängend ist, in keiner klaren Beziehung zur Dynamik steht und deshalb bei Betrachtung des Phasenraumes als „erratisch" erscheinen würde. Dies birgt eine Gefahr: Solche erratischen expandierenden und kontrahierenden Blätter können ganz ungewöhnliche Orbitpaare bilden und dadurch die Vereinfachung der zwiefachen Partition stören.

Teilweise wurde diese Gefahr dadurch vermieden, daß bei den Vereinfachungsschritten nur ganz bestimmte Bündel als neue expandierende oder kontrahierende Bündel deklariert wurden. Normalerweise werden dabei keine erratischen expandierenden oder kontrahierenden Blätter entstehen. Wenn diese aber schon von der anfänglichen Partition erzeugt wurden, dann wird man sie auch nicht mehr los, es sei denn, sie können als Teilmenge eines größeren expandierenden oder kontrahierenden Blattes entfernt werden.

Wenn diese erratischen Blätter nicht entfernt werden, dann erschweren sie die Rekonstruktion einer nicht zwiefachen Partition. Denn von dieser nicht zwiefachen Partition wurde gefordert,

daß sie alle expandierenden und kontrahierenden Blätter der zwiefachen Partition erzeugt, also auch diejenigen, die man als erratisch bezeichnen würde. Und diese erratischen Blätter werden normalerweise nur mit zusätzlichen Symbolischen Gebieten erzeugt werden können.

Es kommt also zu einem Konflikt zwischen den beiden verwendeten Kriterien der Einfachheit. Nach dem neuen Kriterium der Einfachheit 4.1.2 (aus Kap. 4.1.2 soll die Zahl der Orbitpaare vergrößert werden, was durch erratische expandierende und kontrahierende Blätter erreicht werden kann. Das andere, triviale Kriterium wurde bei der Rekonstruktion der nicht zwiefachen Partition berücksichtigt: Die Zahl der Symbole soll möglichst klein sein. Deshalb sollten erratische Blätter möglichst entfernt werden.

Wie ein solcher Konflikt zu entscheiden ist, hängt von den Zielen ab, die man mit Hilfe der generierenden Partition erreichen will. Als Faustregel empfehle ich, während der Vereinfachung der zwiefachen Partition nur auf die Zahl der Orbitpaare zu achten, auch wenn dadurch die erratischen Blätter erhalten bleiben. Denn der Rechenaufwand ist geringer, wenn man nur auf dieses eine Kriterium der Einfachheit achtet. Nachdem man eine nicht zwiefache Partition rekonstruiert hat, kann man immer noch nachprüfen, welche Symbolischen Gebiete eigentlich nötig sind, um den Phasenraum zu überdecken, und ob man viele Orbitpaare verliert, wenn man die restlichen Symbolischen Gebiete entfernt. Allerdings kann es sich am Ende einer solchen Rechnung herausstellen, daß man für die Rekonstruktion einer nicht zwiefachen Partition mehr Symbole braucht, als die anfängliche nicht zwiefache Partition besaß.

7.2.3 Lokale Maxima der Einfachheit

Bei vielen Verfahren, die eine bestimmte „Güte" schrittweise optimieren, besteht eine Hauptgefahr darin, in einem lokalen, aber nicht globalen Maximum der Güte steckenzubleiben. Damit ist ein Ergebnis gemeint, das sich zwar noch verbessern läßt, aber nicht durch die kleinen Änderungen, aus denen das Optimierungsverfahren besteht.

Bei der Vereinfachung der AB–Partition zu der 01–Partition oder der XY–Partition bleibt man nur dann stecken, wenn man ganz besonders große Bündel wählt (siehe Kap. 5.2.2). Um dieses Risiko zu vermeiden, genügte es in diesen Beispielrechnungen, jeweils vier anfängliche expandierende und kontrahiernde Bündel zu wählen. Daß so wenige Bündel genügen, spricht für die Qualität dieser Methode.

Allerdings können auch hierbei die expandierenden oder kontrahierenden Blätter stören, die im letzten Abschnitt als „erratisch" bezeichnet wurden. Dies passiert zum Beispiel, wenn man ein expandierendes Bündel Γ^V in der Zeit vor verschieben und ein anderes expandierendes Bündel Γ^W unverändert lassen will. Wenn ein erratisches kontrahierendes Blatt mit beiden expandierenden Bündeln Orbitpaare bildet, dann muß es auch zusammen mit Γ^V vor verschoben werden, wobei aber die Orbitpaare verloren gehen können, die es mit Γ^W bildete. Was sonst ein normaler Vereinfachungsschritt sein könnte, der keine Orbitpaare entfernt, ist es nun wegen des erratischen kontrahierenden Blattes nicht mehr.

Normalerweise versucht man, solchen lokalen Maxima der Güte zu entkommen, indem man ab und zu auch einen Schritt erlaubt, der die Güte verschlechtert. Dies kann auf verschiedene Weisen geschehen, zum Beispiel bei dem „simulated annealing", wo kurzfristige Energieerhöhungen dazu dienen, das globale Energieminimum molekularer Konfigurationen zu finden. In unserem Fall würde das bedeuten, ab und zu Bündel auch dann in der Zeit zu verschieben, wenn dabei Orbitpaare verloren gehen. Es sollten natürlich möglichst wenige Orbitpaare verloren gehen. Wenn zum Beispiel ein kontrahierendes Bündel nur mit besonders wenigen expandierenden Bündeln Orbitpaare bildet, dann ist es ein Kandidat dafür, um einen Zeitschritt vor verschoben zu werden. Mit etwas Geschick kann man auf diese Weise von einer „einfachsten" Partition ausgehend die anderen „einfachsten" Partitionen finden. Zum Beispiel kann man von der 01–Partition des Hénon Attraktors bei Standardparameterwerten ausgehen und auf einem plausiblen Rechenweg die XY–Partition finden, was ich aber nicht vorführen will.

7.3 Vorteile der Methode

Letztlich wird eine Methode durch erfolgreiche Anwendungen gerechtfertigt. Der Schwerpunkt dieses Textes lag aber nicht bei Anwendungen, sondern bei der Diskussion des allgemeinen Falles. Wenn der Leser auf die Abschnitte zurückblickt, in denen die AB–Partition vereinfacht wurde, dann wird ihm auffallen, wie kurz man diese Beispielrechnung schreiben kann, nachdem man die grundlegende Methode und die besonderen Rechenzeichen verstanden hat. Ein paar Vorteile der allgemeinen Methode sind aber auch ohne weitere Beispielrechnungen zu erkennen.

7.3.1 Synthese bekannter Ansätze

Die Methode vereinigt Elemente aus mehreren bereits veröffentlichten Ansätzen, ist aber allgemeiner anwendbar als diese. Zum Beispiel wurde in dem neuen Kriterium 4.1.2 das Konzept der Tangentialpunkte nicht verwendet. Ein altes Kriterium für die Wahl der Partition forderte, die Partitionsgrenzen so zu legen, daß sie jeden Orbit von Tangentialpunkten genau einmal treffen [30]. Dieses Kriterium scheint durch die XY–Partition und die 01–Partition des Hénon Attraktors bei Standardparameterwerten erfüllt zu werden. Das neue Kriterium 4.1.2 der Einfachheit hatte also in diesen Fällen die Nebenwirkung, daß die ihm zufolge einfachsten Partitionen auch das alte Kriterium erfüllen und den Attraktor in möglichst wenig Tangentialpunkten schneiden.

Da das neue Kriterium der Einfachheit das Konzept der „Tangentialpunkte“ nicht verwendet, ist es auf viel allgemeinere Fälle anwendbar, sogar auf diejenigen von Kap. 2.3.3, wo die Partitionsgrenzen den Attraktor auch außerhalb von primären Tangentialpunkten schneiden mußten. Und manchmal kann das neue Kriterium der Einfachheit unter mehreren Partitionen selbst dann noch eine als die einfachste auszeichnen, wenn dies mit Hilfe der Tangentialpunkte nicht mehr möglich ist. Zum Beispiel erweist sich die 01–Partition des Hénon Attraktors bei schwacher Dissipation einfacher als die XY–Partition (siehe Anhang C.1), obwohl in beiden Fällen die Partitionsgrenze jeden Orbit von Tangentialpunkten nur einmal trifft.

Auch manche geometrischen Ansätze werden von der neuen Methode mit eingeschlossen und verallgemeinert. So wird in Kap. 8 das Konzept „Bündel von Blättern“ dazu verwendet werden, um räumliche Ordnungen zu konstruieren. Und obwohl der topologische Zusammenhang der Blätter für die hier vorgeschlagene Methode ohne Bedeutung ist, hat er sich wie von selbst eingestellt. Denn bei den Ergebnissen der Vereinfachung, nämlich der 01–Partition oder der XY–Partition des Hénon Attraktors, scheinen alle expandierenden Blätter einfach zusammenhängend zu sein.

7.3.2 Reine Symbolmanipulationen

Ein wesentlicher Vorteil der vorgeschlagenen Methode liegt darin, daß sie nur aus algebraischen Manipulationen von Symbolsequenzen besteht. Alles, was man als Ausgangspunkt für die Symbolmanipulationen brauchte, waren eine Partition des Phasenraumes und die grammatikalischen Regeln der entsprechenden Symbolischen Dynamik. In der weiteren Rechnung mußte die Geometrie des Phasenraumes nicht mehr betrachtet werden.

Deswegen ist die Methode auch unabhängig von der Wahl des Koordinatensystems. So konnte die XY–Partition des Hénon Attraktors bei Standardparameterwerten gefunden werden, obwohl sie im Standardkoordinatensystem recht ungewöhlich aussieht. Und deswegen ist die Methode, die hier nur für seltsame Attraktoren in zweidimensionalen dissipativen Systemen getestet wurde, sehr viel allgemeiner: Sie könnte im Prinzip auch auf Hamiltonsche Systeme oder chaotische Sattel und Repelloren angewendet werden, und zwar auch in höheren Dimensionen.

Es ist allerdings ungewiß, ob sich Partitionen dann immer noch so gut vereinfachen lassen, wie in den hier betrachteten zweidimensionalen Abbildungen. Denn auch in diesen einfachen dynamischen Systemen waren Rechnungen per Hand zunächst sehr aufwendig. Erst die besonderen Rechensymbole, die hier eingeführt wurden, machten diese Rechnungen handhabbar. Doch selbst mit diesen effizienten Rechenmethoden dürfte die Analyse höherdimensionaler Systeme sehr aufwendig sein.

7.3.3 Nur kurze Symbolsequenzen sind wichtig

In der Einleitung wurde gefordert, daß sich Methoden der Nichtlinearen Dynamik prinzipiell auch auf experimentelle Daten und nicht nur auf computergenerierte Daten anwenden lassen sollen. Aus verrauschten experimentellen Daten kann aber nicht abgelesen werden, welche langen Symbolsequenzen verboten oder erlaubt sind. Die Feinheiten der Symbolischen Dynamik bleiben dadurch verborgen. Dies beeinträchtigt die hier postulierte Methode wenig. Denn nur kurze Symbolsequenzen waren für die Berechnungen der Orbitpaare und die anderen Rechenschritte wichtig. Die längste Sequenz, deren Unzulässigkeit bei den Beispielrechnungen verwendet wurde, war nur sieben Symbole lang (Gl. C.18).

Weil das Kriterium 4.1.2 der Einfachheit kaum von der Zulässigkeit langer Symbolsequenzen abhängt, ist es zu vermuten, daß bei Parameteränderung „einfachste" Partitionen nur dann verschwinden oder neu erscheinen können, wenn sich die Zulässigkeit von kurzen Symbolsequenzen ändert. Diese ist experimentell durchaus meßbar. Im Falle des bereits erwähnten NMR-Lasers [23] konnte gemessen werden, bei welchen Parameterwerten die zunächst verbotene Symbolsequenz 00 zulässig wird [22]. Derartige Messungen könnten alle grammatikalischen Regeln finden, die man für die hier vorgeschlagene Methode benötigt.

7.4 Allgemeine Lehren

Die Einleitung benutzte das Vorbild des Gehirns, um zu argumentieren, daß es noch unbekannte mathematische Prinzipien geben muß, mit denen sich eine Dynamik strukturieren läßt. Diese Strukturierung der Dynamik sollte an die Dynamik selbst angepaßt sein. Von den einfachen chaotischen Systemen, die hier betrachtet wurden, bis zu den komplexen Sinnesdaten, die ein Säugetiergehirn verarbeitet, ist ein weiter Weg. Doch auch bei diesen einfachen chaotischen Systemen mußte ich ziemlich viele neue Konzepte einführen und mich zwischen ziemlich vielen alten Konzepten entscheiden, bevor ich die hier vorgestellte Methode fand. Dabei habe ich aus diversen Irrwegen einige Lehren gezogen, deren Bedeutung über die hier vorgeschlagene Methode hinaus gehen kann. Die meisten dieser Lehren habe ich bereits erwähnt, doch ich will sie hier noch einmal zusammenfassen und möglichst allgemein formulieren.

7.4.1 Getrennte Beschreibung von Zukunft und Vergangenheit

Eine normale, nicht zwiefache Partition beschreibt Zukunft und Vergangenheit mit denselben Symbolen. Eine zentrale Idee dieser Arbeit bestand darin, Zukunft und Vergangenheit durch verschiedene Symbole zu beschreiben. Die Symbolischen Gebiete einer solchen zwiefachen Partition lassen sich viel flexibler abändern, als die einer nicht zwiefachen Partition.

Man kann sich den Unterschied so vorstellen, daß eine Änderung der nicht zwiefachen Partition immer Auswirkungen auf alle Stellen der unendlichen Symbolsequenzen haben kann. Dagegen läßt sich eine zwiefache Partition durch das zeitliche Verschieben von Bündeln viel gezielter abändern. Zwar können sich dabei in den halbunendlichen Symbolseqeunzen der expandierenden Blätter $\ldots S_{-2}S_{-1}\circ$ und der kontrahierenden Blätter $\circ S_0 S_1 \ldots$ auch beliebig viele Symbole wandeln. Doch an den meisten Stellen bedeutet das nur eine Änderung der Schreibweise, die ohne Bedeutung für die Gestalt des Blattes ist. Nur wenn sich ein Symbol in der Nähe des Punktes $\circ$ ändert, ändert sich auch die Gestalt des Blattes.

In diesem Sinne kann die „Gegenwart" abgeändert werden, ohne daß die entfernte Zukunft und Vergangenheit betroffen werden. Dadurch lassen sich zwiefache Partitionen viel flexibler abändern als nicht zwiefache Partitionen.

7.4.2 Zusammenfassen von Orbits mit gemeinsamer Zukunft oder Vergangenheit

Die getrennte Beschreibung von Zukunft und Vergangenheit ist nicht völlig neu. Auch eine stabile Mannigfaltigkeit kann als Beschreibung der Zukunft und eine instabile Mannigfaltigkeit als Be-

schreibung der Vergangenheit aufgefaßt werden. Dieser Beschreibung fehlt allerdings ein wichtiges Element: Da die stabilen Mannigfaltigkeiten nur das asymptotische zukünftige Verhalten bestimmen, können zwei Orbits derselben stabilen Mannigfaltigkeit lange Zeit ihre eigenen Wege gehen, bevor sie gegeneinander konvergieren. Deshalb wurde das Konzept der kontrahierenden Blätter eingeführt. Ein kontrahierendes Blatt faßt Orbits zusammen, die sozusagen ab der Gegenwart eine gemeinsame Zukunft haben. Normalerweise werden diese Orbits in wenigen Zeitschritten gegeneinander konvergieren und dann in alle Zukunft beieinander bleiben. Durch dieses Konzept „gemeinsame Zukunft" und das analoge Konzept „gemeinsame Vergangenheit" werden die wichtigsten koordinatenunabhängigen Beziehungen zwischen den Orbits erfaßt.

7.4.3 Zusammenfassen der Blätter zu Bündeln

Es gibt unendlich viele expandierende und kontrahierende Blätter. Um auf ihnen operieren zu können, adaptierte ich die Methoden der Symbolischen Dynamik. So wie man sonst Orbits durch unendliche Symbolsequenzen beschreibt, beschrieb ich die expandierenden und kontrahierenden Blätter durch halbunendliche Symbolsequenzen. Durch diese Symbole wurden die unendlich vielen Blätter gleichzeitig zu einer endlichen Zahl von expandierenden und kontrahierenden Bündeln zusammengefaßt, mit denen sich rechnen ließ.

7.4.4 Symbolische Gebiete dürfen überlappen

Es hat für die praktische Rechnung mit Symbolen viele Vorteile, wenn die Symbolischen Gebiete disjunkt sind und jeder Orbit nur eine Symbolsequenz besitzt. Es hat aber auch einen gewaltigen Nachteil: Disjunkte Symbolische Gebiete können nur disjunkte expandierende Blätter erzeugen. Die Menge der expandierenden Blätter kann aber viel flexibler abgeändert werden, wenn bei dieser Änderung auch überlappende expandierende Blätter entstehen dürfen. Obwohl bei der Beispielrechnung von Kap. 5 und 6 disjunkte Symbolische Gebiete durchaus genügten, würde die allgemeine Methode sehr unflexibel werden, wenn man disjunkte Symbolische Gebiete fordern würde. Außerdem zeigte Anhang A.1 an einem eindimensionalen Beispiel, daß die Partitionsgrenzen nur dann durch die Dynamik und das Kriterium der Einfachheit eindeutig festgelegt sind, wenn die Symbolischen Gebiete überlappen dürfen.

Die Folgen von überlappenden Symbolischen Gebieten für die Orbits sind zwar ungewohnt, stören aber kaum: Jeder Orbit wird durch mehrere unendliche Symbolsequenzen beschrieben, von denen jede allein genügt, um ihn von anderen Orbits zu unterscheiden. Und wenn die beiden Orbits α und β auf einem gemeinsamen kontrahierenden Blatt liegen und auch die Orbits β und γ auf einem gemeinsamen kontrahierenden Blatt liegen, dann gilt, weil diese beiden Blätter verschieden sein können, dasselbe noch nicht unbedingt für die Orbits α und γ. In anderen Worten: Wenn α und β eine gemeinsame Zukunft haben und wenn β und γ ein gemeinsame Zukunft haben, dann brauchen α und γ nicht unbedingt eine gemeinsame Zukunft zu haben. Zwar müssen α und γ in der assymptotischen Zukunft gegeneinander konvergieren, doch die nahe Zukunft kann verschieden sein. Die durch die kontrahierenden Blätter, das heißt durch die „gemeinsame Zukunft", definierte Beziehung ist also nicht transitiv.

7.4.5 Topologischer Zusammenhang ist unwichtig

Manche erfolgreichen Methoden der Nichtlinearen Dynamik sind geometrischer Natur. Bei der hier vorgeschlagenen Methode haben sich geometrische Konzepte aber eher als irreführend erwiesen. Wer die Abbildungen von kontrahierenden und expandierenden Blättern in dieser Arbeit betrachtet, der wird zu der Vermutung verleitet, alle expandierenden Blätter könnten als einfach zusammenhängende Punktmengen gewählt werden. Eine solche geometrische Bedingung läßt sich aber nicht mit einfachen Symbolmanipulationen nachprüfen. Und in höherdimensionalen Phasenräumen läßt sich, wie Kap. 3.2.1 argumentierte, diese Bedingung wohl auch kaum erfüllen. Es war deshalb wichtig, auf diese Bedingung zu verzichten. Durch diesen Verzicht kann man sich zwar die

expandierenden und kontrahierenden Blätter einhandeln, die weiter oben als „erratisch" bezeichnet wurden. Doch ohne dieses Risiko ist es nicht möglich, die Änderungen der zwiefachen Partition flexibel zu gestalten.

7.4.6 Räumliche Ordnung ist unwichtig

Die räumliche Ordnung von Blättern läßt sich nicht nur geometrisch konstruieren, sondern, wie Kapitel 8 zeigen wird, auch durch reine Symbolmanipulationen. Dabei werden dieselben Konzepte verwendet wie bei der Vereinfachung generierender Partitionen. Die Konstruktion einer räumlichen Ordnung ließe sich deshalb auch mit der Vereinfachung einer generierenden Partition kombinieren. Trotzdem sind die Methoden voneinander unabhängig. Am leichtesten läßt sich eine zwiefache Partition dann vereinfachen, wenn man die räumliche Ordnung der Blätter ignoriert.

Dadurch unterscheidet sich die hier vorgeschlagene Methode zur Konstruktion generierender Partitionen von den Ansätzen in [17], [69] und den anderen Veröffentlichungen, die in Kap. 8.2.2 und 8.2.3 erwähnt werden. Diese Ansätze stützen sich auf die räumliche Ordnung im Phasenraum. Damit unterliegen sie auch der Einschränkung, daß eine räumliche Ordnung konstruierbar sein muß. In Kapitel 8.3.1 wird argumentieren werden, daß eine räumliche Ordnung nur auf eindimensionalen stabilen oder instabilen Mannigfaltigkeiten konstruiert werden kann. Die Methode zur Vereinfachung generierender Partitionen, die in dieser Arbeit vorgeschlagen wurde, kann dagegen prinzipiell auch auf höherdimensionale Systeme angewendet werden.

Jeder hat seinen Ort.
Man fragte Rabbi Abraham Jaakob: „Unsere Weisen sagen: 'Kein Ding, das seinen Ort nicht hätte.' Es hat also auch der Mensch seinen Ort. Warum ist dann den Leuten zuweilen so eng?" Er antwortete: „Weil jeder den Ort des anderen haben will."
M. Buber: „Die Erzählungen der Chassidim"

Kapitel 8

Räumliche Ordnung

8.1 Überblick

Es gibt noch andere Anwendungen für die hier eingeführten Rechenmethoden als nur die Vereinfachung von generierenden Partitionen. Eine solche Anwendung, die Ausdehnung von „räumlichen Ordnungen" auf immer größere Teile des Attraktors, wird in diesem Kapitel diskutiert. Über räumliche Ordnungen von chaotischen Attraktoren wurde viel veröffentlicht. Am einfachtsten sind räumliche Ordnungen zu verstehen, wenn die Abbildung f eindimensional ist. Abschnitt 8.2.1 stellt die Definition und den Nutzen dieser eindimensionalen Ordnung dar.

Bei zweidimensionalen invertierbaren Abbildungen f wurde auf eine analoge Weise versucht, zwei Ordnungen einzuführen, nämlich „rechts–links" entlang der expandierenden Richtung und „oben–unten" entlang der kontrahierenden Richtung. Damit läßt sich das intuitive Bild präzisieren, mit dem schon Hénon [40] die Wahl seiner Abbildung begründete, daß nämlich die Dynamik aus einem Strecken und Falten besteht. Tatsächlich gibt es sogar mehrere Weisen, wie die Dynamik des Hénon Attraktors bei Standardparameterwerten als ein Strecken und Falten interpretiert werden kann. Eine bekannte Weise, die auf der 01–Partition beruht, und eine neue Weise, die auf der XY–Partition beruht, werden in Kap. 8.2 dargestellt. Die durch diese Partitionen bestimmten räumlichen Ordnungen sind selbstkonsistent, das heißt, die sich kreuzenden Richtungen „rechts" und „oben" haben überall dieselbe Orientierung zueinander. Doch nicht jede Partition definiert eine selbstkonsistente räumliche Ordnung, die AB–Partition zum Beispiel nicht.

Abschnitt 8.3 schlägt eine neue und systematische Methode vor, um eine konsistente räumliche Ordnung zu finden. Die Methode setzt voraus, daß auf allen hinreichend kleinen expandierenden und kontrahierenden Bündeln bereits räumliche Ordnungen definiert werden konnten. Wenn man diese Ordnungen kennt und weiß, wie sie durch f aufeinander abgebildet werden, dann kann man die algebraischen Methoden der vorherigen Kapitel anwenden, um die räumlichen Ordnungen auf größere expandierende Bündel auszudehnen. Die Geometrie des Phasenraumes braucht dabei nicht weiter betrachtet zu werden. Als Beispiel werde ich die kleinen expandierenden Bündel betrachten, die in Kap. 5 bei der Vereinfachung der AB–Partition entstanden. Als Ergebnis dieser Beispielrechnung wird sich tatsächlich die senkrechte Ordnung der 01–Partition ergeben.

Ein Vorteil dieser algebraischen Rechenmethode ist wieder, daß sie sich auf Systeme anwenden läßt, die mehr als eine Dimension haben. Doch auch sie ist Einschränkungen unterworfen. Eine rechts–links Ordnung kann nur auf instabilen Mannigfaltigkeiten definiert werden, die eindimensional sind, und eine oben–unten Ordnung nur auf stabilen Mannigfaltigkeiten, die eindimensional sind. Abschnitt 8.3.1 argumentiert, daß jede sinnvolle räumliche Ordnung diesen Einschränkungen unterworfen ist.

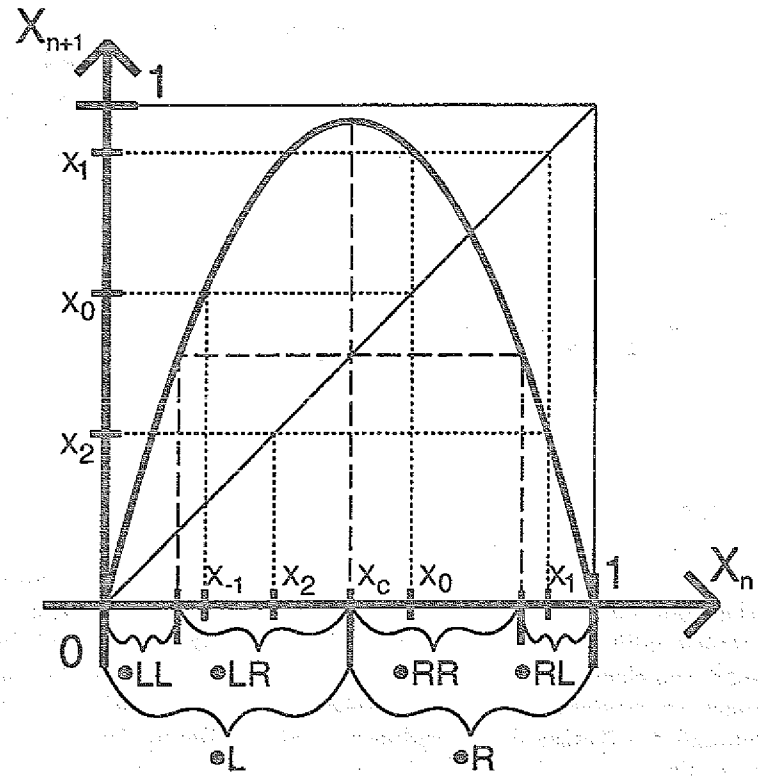


Abbildung 8.1: Wiederholung der eindimensionalen Funktion f in Abb. 2.2

8.2 Elementares zur räumlichen Ordnung

8.2.1 Räumliche Ordnung in einer Dimension

Abb. 8.1 zeigt noch einmal die eindimensionale Funktion f aus Kap. 2.2.2, Abb. 2.2. Wie man sieht, sind die Intervalle $\circ LL$, $\circ LR$, $\circ RR$ und $\circ RL$ in dieser Reihenfolge von links nach rechts entlang der x–Achse angeordnet. Auch Sequenzen mit mehr als zwei Symbolen können auf diese geometrische Weise angeordnet werden. Am meisten Information steckt in der Ordnung der halbunendlichen Sequenzen $\circ S_0 S_1 S_2 \ldots$ (mit $S_i \in \{R, L\}$). Mathematisch kann diese waagrechte Ordnung dadurch erfaßt werden, daß die Menge $\{\circ(R \sqcup L)^{+\infty}\}$ der halbunendlichen Symbolsequenzen durch eine Funktion r (d. h. „rechts") auf irgendeinen durch „$<$" geordneten Raum abgebildet wird, wie zum Beispiel den Raum der reellen Zahlen. Dabei müssen gemäß Abb. 8.1 die Ordnungsrelationen $r(\circ LL\ldots) < r(\circ LR\ldots) < r(\circ RR\ldots) < r(\circ RL\ldots)$ erfüllt werden. Weitere Ordnungsrelationen müssen erfüllt werden, um Symbolsequenzen zu ordnen, die in den ersten beiden Symbolen übereinstimmen. Alle solchen Ordnungsrelationen kann man am besten induktiv angeben: Für alle halbunendlichen Symbolsequenzen σ, τ der Gestalt $S_0 S_1 S_2 \ldots$ (mit $S_i \in \{R, L\}$) soll gelten:

$$r(\circ L\sigma) < r(\circ R\tau) \tag{8.1}$$

$$r(\circ L\sigma) < r(\circ L\tau) \quad \text{falls} \quad r(\circ\sigma) < r(\circ\tau) \tag{8.2}$$

$$r(\circ R\sigma) < r(\circ R\tau) \quad \text{falls} \quad r(\circ\sigma) > r(\circ\tau)\ . \tag{8.3}$$

Durch diese Ordnungsrelationen ist gleichzeitig die Lage r genau genug festgelegt. Denn es ist für das folgende ohne Bedeutung, in was für einem Raum die Funktionswerte $r(\circ\sigma)$ liegen und wie groß ihr Abstand in diesem Raum ist.

Wie man leicht sieht, stimmt die so definierte Ordnung r tatsächlich mit der Reihenfolge überein, in der die durch Symbolsequenzen $\circ S_0 S_1 \ldots$ beschriebenen Punkte auf dem Einheitsintervall liegen. Die Beziehungen 8.2 und 8.3 drücken nichts anderes aus, als daß die Funktion f auf dem

linken Teilintervall streng monoton wachsend und auf dem rechten Teilintervall streng monoton fallend ist.

Derartige räumliche Ordnungen von eindimensionalen Funktionen wurden in vielen Veröffentlichungen diskutiert. Ein Überblick findet sich in [70]. Worin besteht der Nutzen solcher räumlichen Ordnungen? Man kann mit ihrer Hilfe bestimmen, welche Symbolsequenzen erlaubt und welche verboten sind. Im Falle von Abb. 8.1 kann kein Attraktorpunkt weiter rechts liegen als das Abbild $f(x_c)$ des kritischen Punktes. Die Symbolsequenz $\circ\kappa$ des Abbildes $f(x_c)$, die Knetsequenz genannt wird [51], ist daher maximal: $\forall\tau : r(\circ\kappa) \geq r(\circ\tau)$. Wenn die Symbolsequenz $\circ S_0 S_1 S_2 \ldots$ möglich ist, dann gilt deshalb für alle ihre um t Symbole verkürzten Teilsequenzen:

$$r(\circ S_t S_{t+1} \ldots) < r(\circ\kappa) \quad . \tag{8.4}$$

Diese notwenige Bedingung ist sogar hinreichend und kann auf stetige Funktionen f mit mehr Maxima verallgemeinert werden (siehe z. B. [70]). Dieser scheinbare Umweg über die räumliche Ordnung ist tatsächlich der einfachste bekannte Weg, um alle möglichen Orbits eindimensionaler Abbildungen zu klassifizieren.

8.2.2 Die Standardordnung des Hénon Attraktors

In einer Dimension genügt die rechts–links Ordnung, um die Reihenfolge der Punkte im Phasenraum zu beschreiben. In zwei Dimensionen benötigt man zwei räumliche Ordnungen: die waagrechte Ordnung „rechts–links“ und die senkrechte Ordnung „oben–unten“. Diese Ordnungen wurden insbesondere für den Hénon Attraktor untersucht.

Abb. 8.2 zeigt noch einmal den Hénon Attraktor bei Standardparameterwerten $a = 1.4, b = 0.3$. Die durchgezogene Linie umrahmt ein Gebiet, das in sein Inneres abgebildet wird. Dieses Bild wird von der gestrichelten Linie umrahmt und hat die für die Hénon Abbildung typische Bogenform. Abb. 8.3 zeigt denselben Sachverhalt schematisch. Schon in Hénons ursprünglicher Arbeit [40] ist ein solches Schema zu finden. Es verdeutlicht die intuitive Vorstellung, daß die Dynamik f aus Strecken und Falten besteht. Damit sich diese Vorstellung möglichst bequem durch Symbole ausdrücken läßt, empfiehlt es sich, die Position des Punktes $\circ$ in Symbolsequenzen so zu wählen, daß die Symbolischen Gebiete $\circ A_i$ durch ungefähr senkrechte Partitionsgrenzen und ihre Abbilder $A_i\circ$ durch ungefähr waagrechte Partitionsgrenzen getrennt werden. Dies ist zum Beispiel bei der 01–Partition in Abb. 2.4 der Fall.

Damit ist nur die grobe Lage der Partitionsgrenzen festgelegt. Wie die genaue Lage der Partitionsgrenzen zu wählen ist, damit sie mit der intuitiven Vorstellung von Faltungen übereinstimmen, wurde bereits in Kap. 2.3.2 dargestellt und soll hier nur knapp wiederholt werden. Die Grenze zwischen den Symbolischen Gebieten $0\circ$ und $1\circ$ wurde so gewählt, daß der Attraktor in den „primären“ Tangentialpunkten [30] geschnitten wird, wo weder die Krümmung der instabilen noch die der stabilen Mannigfaltigkeit besonders groß ist. Bei den Bildern $f^t(x_c)$ dieser Tangentialpunkte x_c wächst die Krümmung der instabilen Mannigfaltigkeit exponentiell schnell mit t an. Umgekehrt finden sich große Krümmungen der stabilen Mannigfaltigkeit bei den Urbildern der Partitionsgrenze. Da Tangentialpunkte wahrscheinlich dicht auf dem Attraktor liegen (siehe Kap. 2.3.2), sind vermutlich in jeder noch so kleinen Umgebung eines Attraktorpunktes beliebig große Krümmungen der instabilen und stabilen Mannigfaltigkeit zu finden. Anhang A.2 zufolge sind diese Krümmungsradien so klein, daß der Eindruck, den Abbildungen wie 8.4 vermitteln, wohl kein numerisches Artefakt ist: Lange Abschnitte der stabilen Mannigfaltigkeit liegen in etwa parallel zueinander und strukturieren die Umgebung des Attraktors, indem sie die ungefähre Richtung der zukünftigen Kontraktion anzeigen. Umgekehrt sieht der Attraktor aus wie ein Stapel aus langen, parallelen Abschnitten der instabilen Mannigfaltigkeit, die eine ungefähre Richtung der vergangenen Expansion anzeigen.

Die Grundidee bei der Konstruktion einer zweidimensionalen räumlichen Ordnung besteht darin, diese Abschnitte als eine Art von Koordinatensystem zu betrachten. Das heißt, die Relation „oben“ ordnet Punkte, die auf demselben Abschnitt der stabilen Mannigfaltigkeit liegen; die Relation „rechts“ ordnet Punkte, die auf demselben Abschnitt der instabilen Mannigfaltigkeit liegen. Wie groß diese Abschnitte sind, ist durch die Symbolischen Gebiete $0\circ$ und $1\circ$ festgelegt.

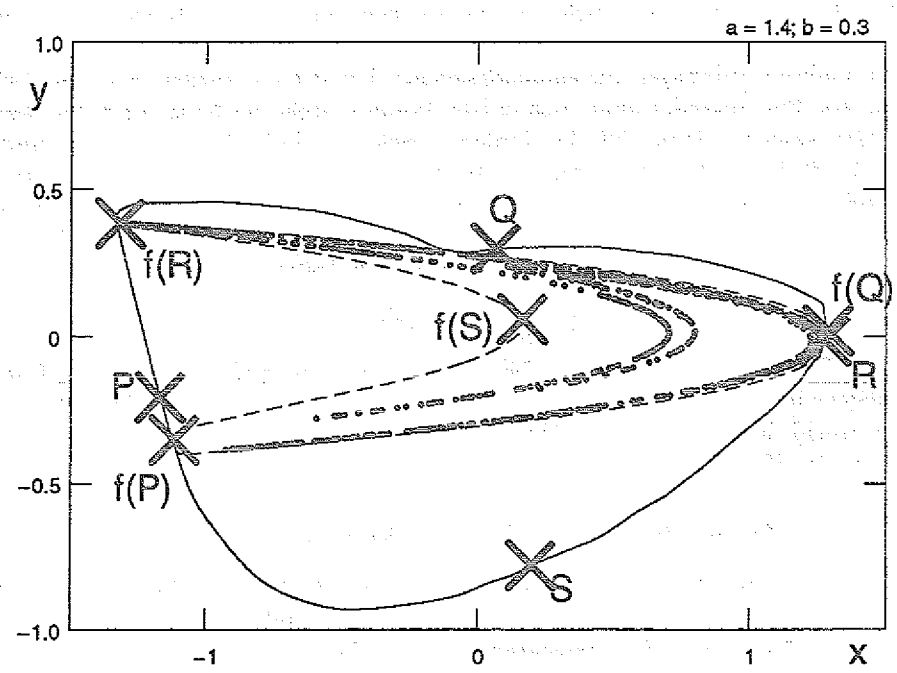


Abbildung 8.2: Attraktionsgebiete des Hénon Attraktors bei Standardparameterwerten. Das von der durchgezogenen Linie umrahmte Gebiet wurde so gewählt, daß es durch die Dynamik f in sein Inneres abgebildet wird. Die gestrichelten Linie ist das Bild der durchgezogenen Linie. Die Bilder der vier Punkte P, Q, R und S sollen verdeutlichen, wie das ursprünglich eher ovale Gebiet durch Strecken, Stauchen, Spiegeln und Falten bogenförmig wird. Die Attraktorpunkte und die Abschnitte der stabilen Mannigfaltigkeit sind wie in Abb. 2.4 eingetragen.

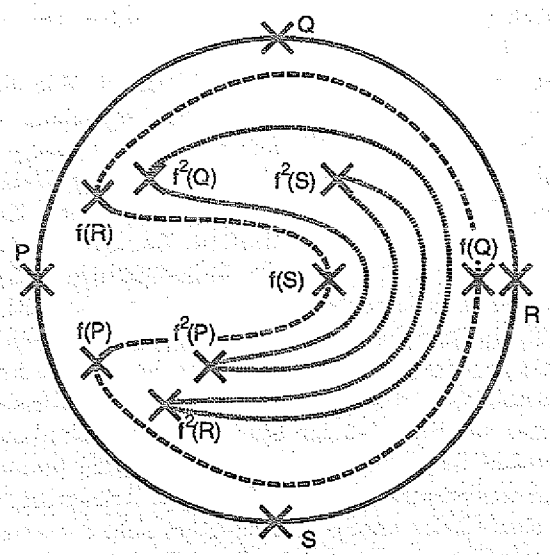


Abbildung 8.3: Schematische Wirkung der Hénon Abbildung. Wie in Abb. 8.2 ist eine durchgezogene Linie mit vier Punkten P, Q, R und S versehen. Die gestrichelte Linie soll ihr erstes Abbild illustrieren, die gepunktete Linie ihr zweites Abbild. Die Gestalt der gepunkteten Linie erinnert schon etwas an den Hénon Attraktor. Kontrahierende Blätter kann man sich schematisch als senkrechte Linien vorstellen und expandierende Blätter als waagrechte Linien.

Üblicherweise definiert man die senkrechte Lage U („up") so, daß sie nur von der vergangenen Symbolsequenz $\ldots S_{-2}S_{-1}\circ$ (mit $S_i \in \{0,1\}$) abhängt. Dann lassen sich die neuen Begriffe aus Kap. 3 anwenden: Die Abschnitte, entlang derer Punkte waagrecht oder senkrecht geordnet sind, sind die expandierenden oder kontrahierenden Blätter. Und alle Punkte eines expandierenden Blattes liegen auf gleicher Höhe. Umgekehrt hängt die waagrechte Lage R („right") nur von der zukünftigen Symbolsequenz $\circ S_0 S_1 \ldots$ ab und ist deshalb für alle Punkte eines kontrahierenden Blattes dieselbe.

Die präzise mathematische Definition der senkrechten Lage U und der waagrechten Lage R verläuft wie im eindimensionalen Fall, der im letzten Abschnitt besprochen wurde. Wieder sind U und R Funktionen, die halbunendliche Symbolsequenzen in einen durch „$<$" geordneten Raum abbilden. Die Ordnungsrelationen werden wieder induktiv definiert [15]: Für alle halbunendlichen Symbolsequenzen σ, τ der Gestalt $S_0 S_1 S_2 \ldots$ (mit $S_i \in \{0,1\}$) soll gelten:

$$R(\circ 0\sigma) < R(\circ 1\tau) \tag{8.5}$$

$$R(\circ 0\sigma) < R(\circ 0\tau) \quad \text{falls } R(\circ\sigma) < R(\circ\tau) \tag{8.6}$$

$$R(\circ 1\sigma) < R(\circ 1\tau) \quad \text{falls } R(\circ\sigma) > R(\circ\tau)\ . \tag{8.7}$$

Und für alle σ, τ der Gestalt $\ldots S_{-3}S_{-2}S_{-1}$ soll gelten:

$$U(\sigma 0\circ) < U(\tau 1\circ) \tag{8.8}$$

$$U(\sigma 0\circ) < U(\tau 0\circ) \quad \text{falls } U(\sigma\circ) > U(\tau\circ) \tag{8.9}$$

$$U(\sigma 1\circ) < U(\tau 1\circ) \quad \text{falls } U(\sigma\circ) < U(\tau\circ)\ . \tag{8.10}$$

Wieder lassen sich diese Beziehungen geometrisch deuten. Die Ordnungsrelationen von R sind wie bei einer eindimensionalen Abbildung mit zentralem Maximum (Gleichungen 8.1 bis 8.3). Ergo behalten die linken kontrahierenden Blätter ihre waagrechte Reihenfolge unter der Abbildung f, während sich die waagrechte Reihenfolge der rechten kontrahierenden Blätter umkehrt. Dies kann man leicht in Abb. 8.2 oder 8.3 nachprüfen. Auch die durch Gl. 8.8 bis 8.10 ausgedrückte Wirkung der Abbildung f^{-1} auf die expandierenden Blätter kann geometrisch nachvollzogen werden. Die senkrechte Lage dieser Blätter ändert sich so wie die waagrechte Lage von Punkten, auf die eine eindimensionale Abbildung mit zentralem Minimum wirkt.

8.2.3 Andere Konstruktionen von räumlichen Ordnungen

Diese räumliche Ordnung des Hénon Attraktors wurde im Detail von Cvitanović, Gunaratne und Procaccia [15] untersucht. Sie interpretierten $R(\circ\sigma)$ und $U(\tau\circ)$ als reelle Zahlen. Als Koordinaten definieren diese Zahlen die „Symbolische Ebene". Nach dieser Koordinatentransformation ist die Abbildung f besonders einfach. Leider sind solche Koordinatentransformationen, genau wie ihre eindimensionalen Analoga, überall unstetig, falls irgendeine endliche Symbolsequenz verboten ist.

In der Symbolischen Ebene sind waagrechte Linien das, was expandierende Blätter im Phasenraum sind. Ebenso entsprechen sich senkrechte Linien und kontrahierende Blätter. Allerdings sind die Konzepte der expandierenden und kontrahierenden Blätter allgemeiner als die der waagrechten und senkrechten Linien, weil letztere nur in zweidimensionalen Systemen mit globaler räumlicher Ordnung definiert werden können und weil in allen mir bekannten Symbolischen Ebenen die Zukunft und die Vergangenheit durch dieselbe Partition beschrieben und die Symbolischen Gebiete als disjunkt angenommen wurden.

Leider kann man in zwei Dimensionen nur schwer aus der räumlichen Ordnung ablesen, welche Symbolsequenzen zulässig und welche verboten sind. In einer Dimension folgt aus dem einen kritischen Punkt x_c die eine Zulässigkeitsbedingung (8.4). In zwei Dimensionen folgen analog aus den unendlich vielen Tangentialpunkten unendlich viele Zulässigkeitsbedingungen (die in der Symbolischen Ebene als „Stutzfront" oder „pruning front" dargestellt werden). Es wird vermutet, daß diese Bedingungen auch hinreichend für die Existenz eines Orbits sind. Auf diese Weise wurden in [17] immerhin alle verbotenen Symbolsequenzen bis zur Länge 31 numerisch berechnet. In [69] wurden etwas weniger umfangreiche grammatikalische Regeln per Hand untersucht, ohne daß ein

einfaches Schema sichtbar wurde. Die Zahl der verbotenen Symbolsequenzen, die keine kürzeren verbotenen Sequenzen enthalten, wächst mit zunehmender Länge n der Sequenz rapide an, und zwar in zwei Dimensionen etwa wie $\exp(n \cdot \lambda_u \cdot (d_a - 1)/d_a)$, wobei d_a die fraktale Attraktordimension und λ_u der positive Lyapunov Exponent ist [16]. Diese Details der Grammatik brauchen hier nicht weiter betrachtet zu werden, da wir nur die kurzen verbotenen Symbolsequenzen zu kennen brauchen.

Übrigens wurden räumliche Ordnungen auch zur Analyse diverser anderer dynamischer Systeme verwendet. In hyperbolischen Systemen mit genau einer instabilen Richtung wurden ähnliche Konzepte schon vor langer Zeit verwendet, um höherdimensionale Attraktoren auf eine eindimensionale „verzweigte Mannigfaltigkeit" zu projizieren (siehe z. B. die Einführung in [33]). Aus diesem Ansatz entwickelten sich später Versuche, dynamische Systeme mit Hilfe topologischer Methoden zu behandeln, insbesondere mit Methoden der Knotentheorie. Ein knapper Überblick findet sich in [52]. Diese Methoden lassen sich auf Flüsse im Phasenraum noch besser anwenden als auf diskrete Abbildungen. Allerdings sind sie nur dann anwendbar, wenn es nicht mehr als eine instabile und eine stabile Richtung gibt. Die einzige mir bekannte Ausnahme [54] stellt eine Methode dar, die aus anderen Gründen auf ganz spezielle dynamische Systeme beschränkt ist.

Sogar in Hamiltonschen (d. h. nicht dissipativen) Systemen könnte man zumindest lokal eine räumliche Ordnung konstruieren. Der interessierte Leser sei an [49] verwiesen, wo eine räumliche Ordnung zwar nicht explizit erwähnt wird, aber in der Anordnung der „Resonanzen" doch offensichtlich ist (insbesondere in der dortigen Abbildung 57). All diese Veröffentlichungen zeigen, daß räumliche Ordnungen ein breites Anwendungsfeld haben.

8.2.4 Geometrische Illustration

Die Darstellung in der eben erwähnten Symbolischen Ebene ist wegen der unstetigen Koordinatentransformation wenig anschaulich. Im folgenden werde ich daher eine andere Darstellung verwenden. Dabei wird die Partitionsgrenze in der Zeit vor bzw. zurück iteriert, so daß sie die instabile bzw. stabile Mannigfaltigkeit nahe der starken Krümmungen schneidet. Zwischen diesen Schnitten kann man in Phasenraumportraits wie Abb. 8.4 bzw. 8.5 die expandierenden bzw. kontrahierenden Blätter erkennen. Die Richtungen „rechts" bzw. „oben" werden dann einfach durch Pfeile angedeutet, die ungefähr parallel zu den expandierenden bzw. kontrahierenden Blättern sind.

Wie werden diese Richtungen „oben" und „rechts" gewählt? Auf einem expandierenden Blatt kann die Richtung „rechts" frei gewählt werden. Da der zweidimensionale Phasenraum eine orientierbare Mannigfaltigkeit ist, wird dadurch die Richtung „oben" auf allen kontrahierenden Blättern festgelegt, die das expandierende Blatt schneiden. Diese kontrahierenden Blätter schneiden weitere expandierende Blätter und legen dadurch deren Richtungen „rechts" fest, und so weiter. Doch nicht immer erhält man auf diese Weise eine selbstkonsistente räumliche Ordnung. Die AB–Partition des Hénon Attraktors bei Standardparameterwerten erzeugt expandierende und kontrahierende Blätter, die sich so schneiden, daß die Richtungen „oben" und „rechts" nicht überall in Einklang gebracht werden können (Abb 8.6). Dagegen findet man für die XY–Partition eine räumliche Ordnung (Abb. 8.7 und 8.8). Die bekannte räumliche Ordnung der 01–Partition wurde bereits in Gleichung 8.5 bis 8.7 und 8.8 bis 8.10 definiert und in Abb. 8.4 und 8.5 dargestellt.

8.3 Allgemeine Konstruktion einer räumlichen Ordnung

Diese geometrische Konstruktion einer räumlichen Ordnung funktioniert nur dann gut, wenn übersichtliche Darstellungen des Phasenraumes vorliegen, komplett mit stabiler und instabiler Mannigfaltigkeit. Was tut man aber, wenn der Phasenraum hochdimensional ist und sich nicht mehr auf einen Blick übersehen läßt? Was, wenn der Verlauf der stabilen Mannigfaltigkeit nur in unmittelbarer Nähe des Attraktors bekannt ist? Die in den bisherigen Kapiteln eingeführten Methoden können helfen, auch in diesen schwierigeren Fällen eine senkrechte Ordnung zu finden.

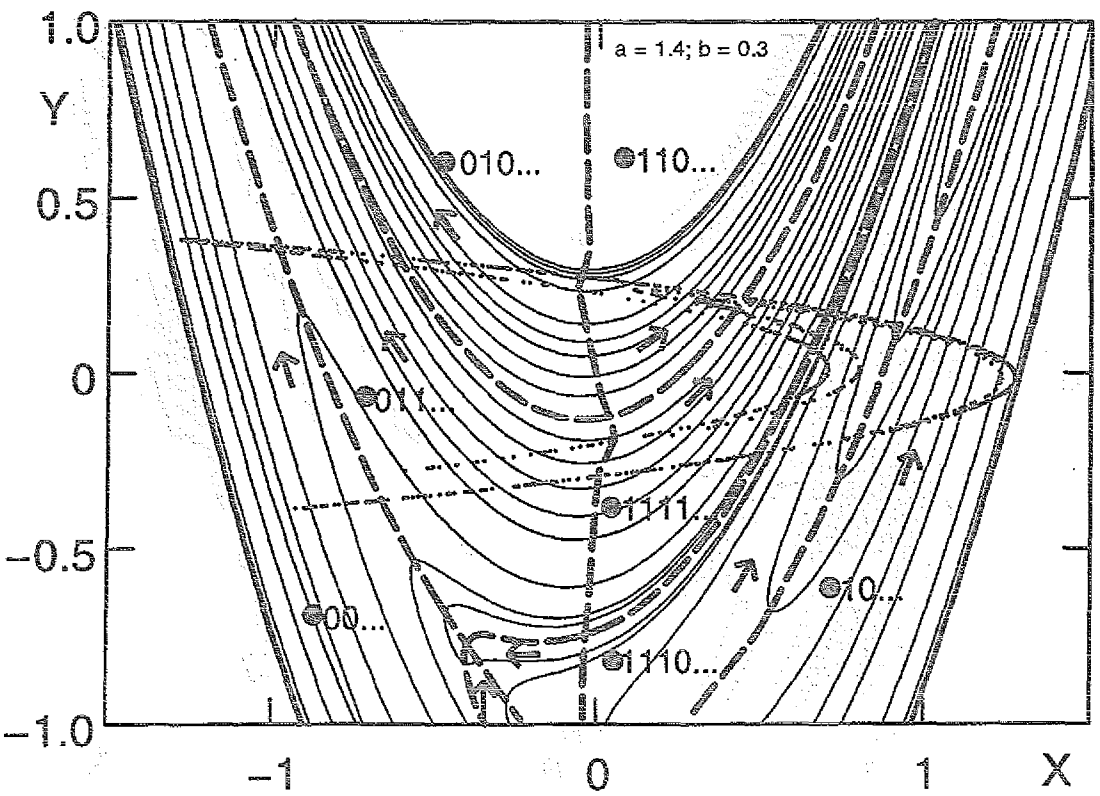


Abbildung 8.4: Die senkrechte Ordnung, die von der 01-Partition auf dem Hénon Attraktor bei Standardparameterwerten erzeugt wird. Die Attraktorpunkte und die Abschnitte der stabilen Mannigfaltigkeit sind wie in Abb. 2.4 eingetragen. Die gestrichelten Linien unterteilen wie in Abb. 3.3 die stabile Mannigfaltigkeit in kontrahierende Blätter. Die Pfeile deuten die Richtung „oben" auf diesen kontrahierenden Blättern an. Die Richtung „oben" zeigt immer entlang der stabilen Mannigfaltigkeit und wechselt abrupt an den gestrichelten Linien.

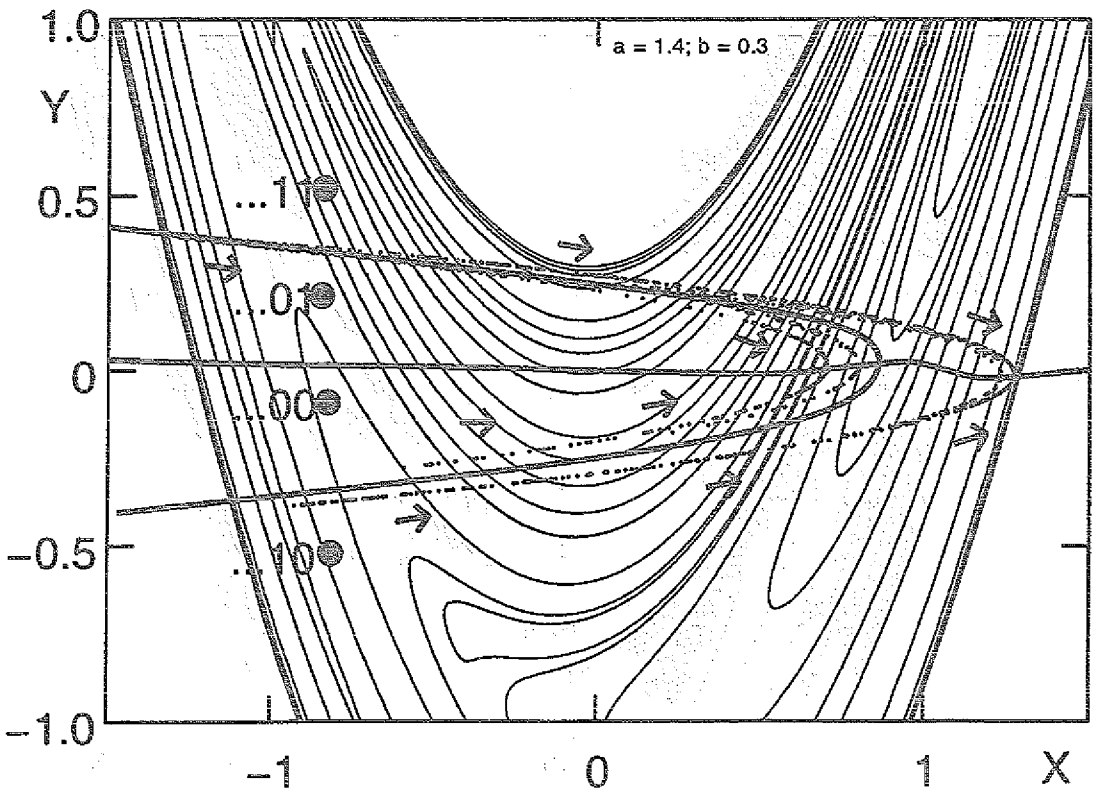


Abbildung 8.5: Die waagrechte Ordnung, die von der 01–Partition auf dem Hénon Attraktor bei Standardparameterwerten erzeugt wird. Die Attraktorpunkte und die Abschnitte der stabilen Mannigfaltigkeit sind wie in Abb. 2.4 eingetragen. Die dicken durchgezogenen Linien unterteilen wie in Abb. 3.2 die instabile Mannigfaltigkeit in expandierende Blätter. Die Pfeile deuten die Richtung „rechts“ auf diesen expandierenden Blättern an. Diese Richtung „rechts“ kreuzt die Richtung „oben“ aus Abb. 8.4 immer mit der richtigen Orientierung.

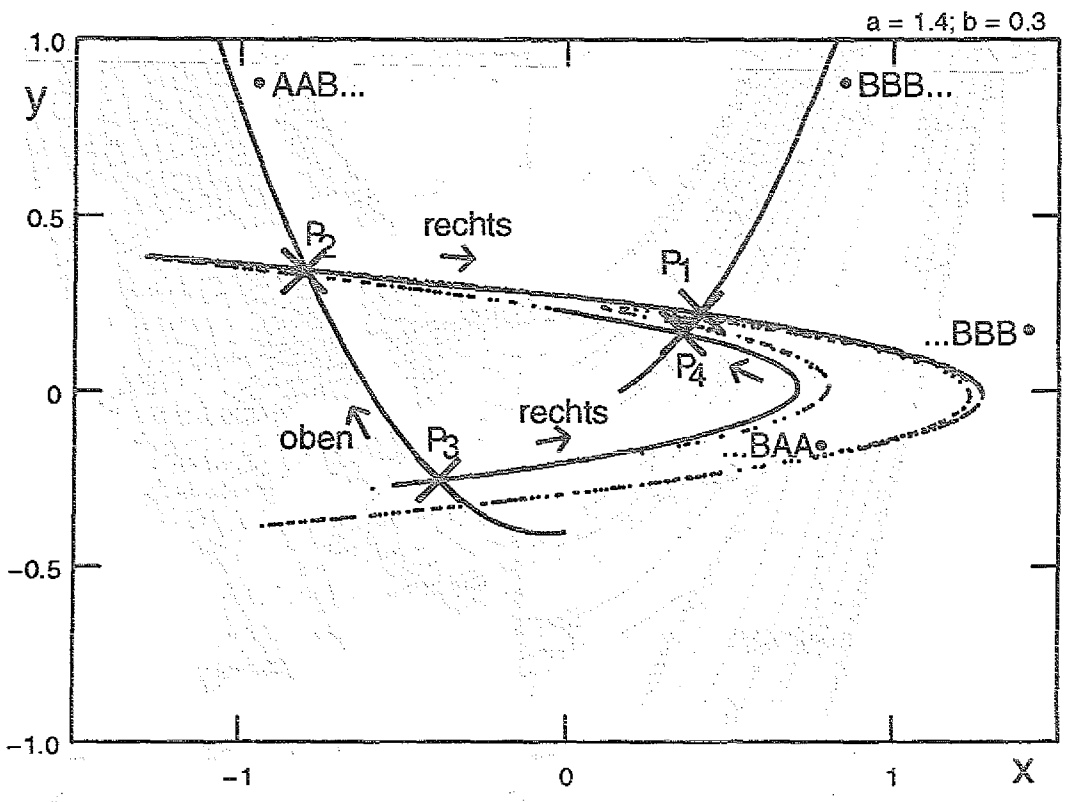


Abbildung 8.6: Die AB–Partiton des Hénon Attraktors bei Standardparameterwerten erzeugt keine selbstkonsistente räumliche Ordnung. Die Attraktorpunkte sind wie in Abb. 4.2 eingetragen. Die vier Attraktorpunkte P_1, P_2, P_3 und P_4 liegen jeweils paarweise auf kontrahierenden und expandierenden Blättern (dicke durchgezogene Linien), die von der AB–Partiton erzeugt werden. Wenn P_2 als oberhalb von P_3 definiert wird, dann liegt P_1 rechts von P_2 und P_4 rechts von P_3. Auf dem kontrahierenden Blatt von P_1 und P_4 kann dann aber keine Richtung „oben" gewählt werden, die mit diesen beiden Richtungen „rechts" in Einklang wäre.

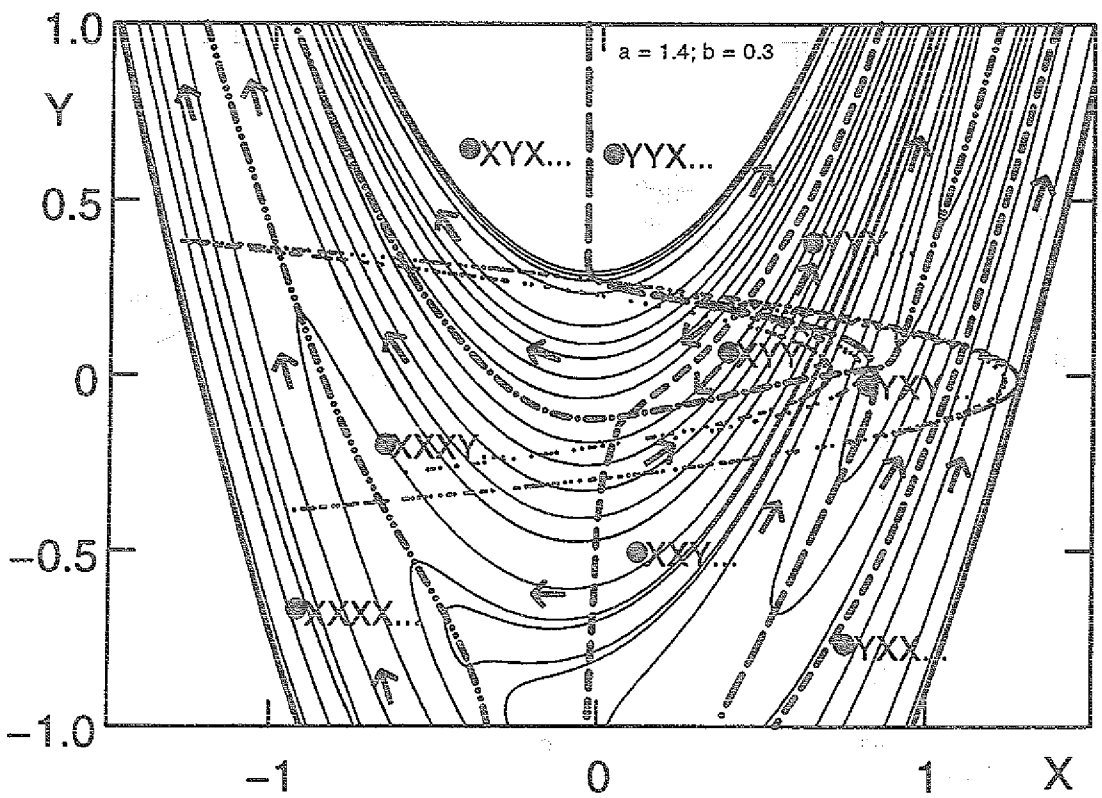


Abbildung 8.7: Die senkrechte Ordnung, die von der XY–Partition auf dem Hénon Attraktor bei Standardparameterwerten erzeugt wird. Die Attraktorpunkte und die Abschnitte der stabilen Mannigfaltigkeit sind wie in Abb. 7.1 eingetragen. Die gestrichelten Linien unterteilen die stabile Mannigfaltigkeit in kontrahierende Blätter $\circ S_0 S_1 \ldots$. Die Pfeile deuten die Richtung „oben" auf diesen kontrahierenden Blättern an. In $\circ XYY$ und in einem Teil von $\circ XYX$ ist nun dort „oben", wo im Falle der 01–Partition in Abb. 8.4 „unten" war.

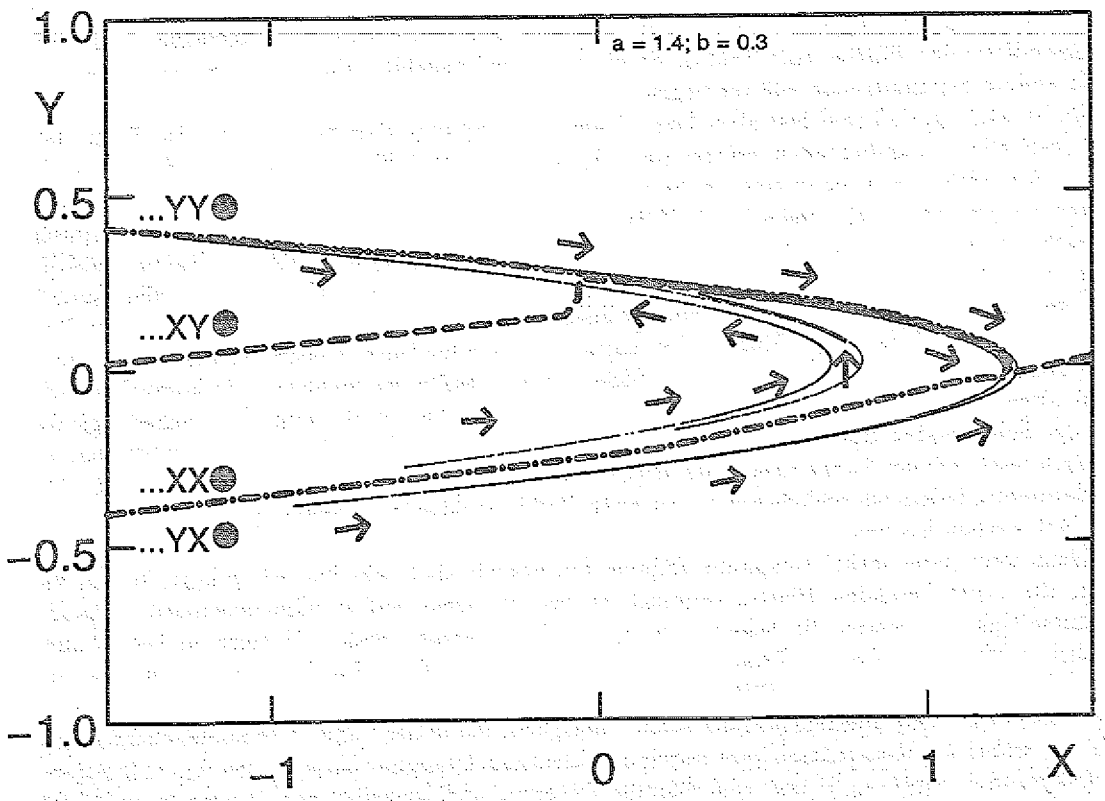


Abbildung 8.8: Die waagrechte Ordnung, die von der XY–Partition auf dem Hénon Attraktor bei Standardparameterwerten erzeugt wird. Die Attraktorpunkte sind wie in Abb. 7.1 eingetragen. Die gestrichelten Linien unterteilen die instabile Mannigfaltigkeit in expandierende Blätter $\ldots S_{-2}S_{-1}\circ$. Die Pfeile deuten die Richtung „rechts“ auf diesen expandierenden Blättern an. In dem Gebiet $XX\circ$ sind manche der expandierenden Blätter S–förmig gebogen. Trotzdem ist die Richtung „rechts“ überall in Einklang mit der Richtung „oben“ in Abb. 8.7.

8.3.1 Senkrechte Ordnung benachbarter expandierender Blätter

Dieser Abschnitt argumentiert, daß sich benachbarte expandierende Blätter dann und nur dann senkrecht ordnen lassen, wenn die stabile Mannigfaltigkeit eindimensional ist. Diese Ordnungsrelationen zwischen benachbarten Blättern sind Voraussetzung für die folgenden Abschnitte, in denen die Ordnungsrelationen schrittweise auf immer mehr Blätter ausgedehnt werden.

Beispiele für senkrechte Ordnungen in zwei Dimensionen wurden in Abb. 8.4 und 8.7 gegeben. Es gibt Plausibilitätsargumente, daß sich räumliche Ordnungen auch in höheren Dimensionen d_p konstruieren lassen, falls nur ein Lyapunov Exponent negativ ist, die Partition generierend ist, ihre Symbolischen Gebiete $A_i\circ$ disjunkt sind und deren Ränder $\delta(A_i\circ)$ alle Orbits von Tangentialpunkten schneiden. Denn dann wird die instabile Mannigfaltigkeit durch $f^t(\delta(A_i\circ))$, $t = 0, 1, 2\ldots$ bei allen scharfen Krümmungen geschnitten und zerfällt in expandierende Blätter, die wie in einem Stapel nebeneinander liegen, ohne sich zu schneiden. Auf einem seltsamen Attraktor bilden diese expandierenden Blätter eine fraktale Struktur, so daß beliebig nahe an jedem expandierenden Blatt andere expandierende Blätter liegen.

Es ist wichtig, daß (auf fast allen Orbits) nur ein Lyapunov Exponent negativ ist. Denn dann sind (fast alle) expandierenden Blätter $(d_p - 1)$–dimensionale Blätter in einem d_p–dimensionalen Raum. Als solche liegen sie in einer mehr oder weniger offensichtlichen Reihenfolge nebeneinander. Wären sie nur $(d_p - 2)$–dimensionale Blätter, dann wäre ihnen ebensowenig eine Reihenfolge anzusehen wie ungeordneten eindimensionalen Linien in einem dreidimensionalen Raum. Natürlich kann es auch bei $(d_p - 1)$ dimensionalen Blättern vorkommen, daß sie nicht ungefähr parallel sondern eher seitlich versetzt nebeneinander liegen. Dann ist es nicht offensichtlich, welches Blatt höher liegt. In solcher Lage könnten sie sogar von verschiedenen kontrahierenden Blättern in einer Reihenfolge geschnitten werden, die nicht immer dieselbe ist, sondern vom kontrahierenden Blatt abhängt. Doch weil die Blätter benachbart sind, hängen diese möglichen Schwierigkeiten von der Feinstruktur des Attraktors ab. Wie bei der Suche nach generierenden Partitionen will ich auch jetzt, bei der Suche nach einer senkrechten Ordnung, alle Probleme mit der Feinstruktur ausklammern, indem ich schlichtweg voraussetze, daß benachbarte expandierende Blätter senkrecht geordnet werden können.

Wenn zwei (oder mehr) Lyapunov Exponenten negativ sind, sehe ich, wie gesagt, keinen Weg mehr, die expandierenden Blätter räumlich zu ordnen. Denn auf zweidimensionalen instabilen Mannigfaltigkeiten wären die Schnittpunkte mit den kontrahierenden Blättern in keiner offensichtlichen Weise angeordnet. Zwar ist auch eine solche mehrdimensionale Mannigfaltigkeit durch niedrigdimensionalere Mannigfaltigkeiten strukturiert. Insbesondere geht laut [62] durch fast alle Attraktorpunkte eine eindimensionale Mannigfaltigkeit, die in der Zukunft besonders schnell kontrahiert, wobei die Kontraktionsrate durch den kleinsten Lyapunov Exponenten $\lambda_{d_p} < 0$ gegeben ist. (Abstand $\sim \exp(t \cdot \lambda_{d_p})$) Dies sind offenbar Untermannigfaltigkeiten der stabilen Mannigfaltigkeiten. Sie dürfen sich gegenseitig nicht kreuzen. Wenn diese eindimensionalen Untermannigfaltigkeiten in demselben Abschnitt der stabilen Mannigfaltigkeit liegen, dann variieren ihre Richtungen stetig mit dem Ort. Deshalb könnte man hoffen, auf solche Untermannigfaltigkeiten eine räumliche Ordnung aufzubauen. Doch auch Abschnitte verschiedener stabiler Mannigfaltigkeiten können im Phasenraum dicht nebeneinander liegen. Und da solche Abschnitte sich unter der Dynamik f in ihrer asymptotischen Zukunft unterscheiden, können die eindimensionalen Untermannigfaltigkeiten in ihnen ganz verschieden verlaufen. Sie können dann auch die expandierenden Blätter in praktisch beliebiger Reihenfolge schneiden. Aus diesem Grund sehe ich keinen Weg, auf diesen Untermannigfaltigkeiten eine räumliche Ordnung aufzubauen.

Aus ganz analogen Gründen sollte nur ein Lyapunov Exponent positiv sein, wenn man die kontrahierenden Blätter waagrecht ordnen will. Dabei macht es keinen wesentlichen Unterschied, daß die kontrahierenden Blätter, als Mengen von Attraktorpunkten betrachtet, fraktal sind. Denn die weiter unten stehenden Vorausetzungen für eine räumliche Ordnung lassen sich auch auf fraktale Blätter übertragen.

Wenn man sowohl eine waagrechte als auch eine senkrechte Ordnung konstruieren will, muß man sich daher auf dynamische Systeme mit nur einem positiven und nur einem negativen Lyapunov Exponenten beschränken, wozu insbesondere zweidimensionale chaotische Abbildun-

gen gehören. Bei höherdimensionalen Abbildungen kann man allenfalls eine senkrechte Ordnung oder eine waagrechte Ordnung konstruieren, aber nicht beide Ordnungen zugleich.

8.3.2 Wirkung der Dynamik auf die senkrechte Ordnung

Es wurde eben diskutiert, wann auf benachbarten expandierenden Blättern eine räumliche Ordnung definiert werden kann. Diese Diskussion war aber recht unpräzise. Welche Blätter sind benachbart? Und welche Bedingungen muß eine räumliche Ordnung erfüllen, um diesen Namen zu verdienen?

Zunächst zur ersten Frage. Als benachbart sollen zwei Blätter gelten, wenn sie in demselben expandierenden Bündel $\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} A_j \circledcirc\}$ liegen. (Dabei ist $\{A_1\circledcirc, \ldots A_m\circledcirc\}$ wieder eine Partition, welche die Vergangenheit beschreibt.) Nur zwischen solchen benachbarten Blättern sollen Ordnungsrelationen vorausgesetzt werden; wenn zwei Blätter nicht in demselben expandierenden Bündel liegen, braucht keines als das höhere definiert zu sein. Wie wir in Kap. 5.2 sahen, lassen sich expandierende Bündel verkleinern, ohne daß sich ihre expandierenden Blätter verkleinern. Dabei reduziert sich die Zahl der Nachbarn, die ein expandierendes Blatt hat, und es wird leichter, eine senkrechte Ordnung zu finden.

Nun zur anderen Frage. Ich werde ziemlich strenge Anforderungen an die senkrechte Ordnung stellen, nämlich die beiden folgenden Voraussetzungen 8.3.1 und 8.3.2. Aus denen folgt nicht nur, daß von jeweils zwei Blättern desselben Bündels $\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} A_j \circledcirc\}$ eines höher liegt als das andere, sondern auch, daß sich diese Ordnung leicht aus den Symbolen A_i bestimmen läßt. Außerdem darf die Abbildung f die senkrechte Ordnung nur auf ganz bestimmte Weisen verändern.

Voraussetzung 8.3.1 (Entweder Orientierungserhaltung oder Orientierungsumkehr) *Unter der Dynamik f ist jeder Übergang von einem expandierenden Bündel $\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} A_k \circledcirc\}$ zu einem expandierenden Bündel $\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} A_l \circledcirc\}$ entweder orientierungserhaltend oder orientierungsumkehrend. Wenn der Übergang die Orientierung erhält, dann gilt für alle expandierenden Blätter $\alpha A_k A_l \circledcirc$, $\beta A_k A_l \circledcirc \in \{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} A_k A_l \circledcirc\}$:*

$$U(\alpha A_k \circledcirc) < U(\beta A_k \circledcirc) \Leftrightarrow U(\alpha A_k A_l \circledcirc) < U(\beta A_k A_l \circledcirc) \quad . \tag{8.11}$$

Bei Orientierungsumkehr gilt für alle solchen Blätter:

$$U(\alpha A_k \circledcirc) < U(\beta A_k \circledcirc) \Leftrightarrow U(\alpha A_k A_l \circledcirc) > U(\beta A_k A_l \circledcirc) \quad . \tag{8.12}$$

Voraussetzung 8.3.2 (Senkrechte Ordnung der Übergänge) *Wenn unter der Dynamik f Übergänge von den expandierenden Bündeln $\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} A_j \circledcirc\}$ und $\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} A_k \circledcirc\}$ zu einem expandierenden Bündel $\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} A_l \circledcirc\}$ führen, dann liegen alle Blätter des einen Übergangs oberhalb der Blätter des anderen Übergangs. Entweder gilt also für alle expandierenden Blätter $\alpha A_j A_l \circledcirc$ und $\beta A_k A_l \circledcirc \in \{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} A_l \circledcirc\}$:*

$$U(\alpha A_j A_l \circledcirc) > U(\beta A_k A_l \circledcirc) \quad , \tag{8.13}$$

oder für alle solchen Blätter gilt:

$$U(\alpha A_j A_l \circledcirc) < U(\beta A_k A_l \circledcirc) \quad . \tag{8.14}$$

Beide Sachverhalte werden ausgedrückt, indem man in dem Graphen der zeitlichen Übergänge (wie z. B. Abb. 8.9) die Übergänge beschriftet. Das Pluszeichen „+" kennzeichnet Erhaltung der Orientierung, das Minuszeichen „-" kennzeichnet Umkehr der Orientierung. Die senkrechte Ordnung der Übergänge in Voraussetzung 8.3.2 wird durch die Beschriftungen $U_1, U_2, \ldots$ ausgedrückt. Dabei liegen die zu U_i gehörenden Blätter tiefer als die zu U_j gehörenden, falls $i < j$.

Wenn diese Voraussetzungen erfüllt sind, dann folgt die senkrechte Ordnung in jedem expandierenden Bündel $\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} A_l \circledcirc\}$ tatsächlich aus den Symbolen. Denn die Ordnungsrelation zwischen Teilbündeln $\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} A_j A_l \circledcirc\}$ und $\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} A_k A_l \circledcirc\}$ folgt aus der Beschriftung der beiden Übergänge $A_j \circledcirc \to A_l \circledcirc$ und $A_k \circledcirc \to A_l \circledcirc$ mit U_i. Und die Ordnungsrelationen in

einem Teilbündel $\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} A_k A_l \circledcirc\}$' folgen wegen Voraussetzung 8.3.1 aus denen im Bündel $\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} A_k \circledcirc\}$' und der Beschriftung des Übergangs $A_k \circledcirc \to A_l \circledcirc$ mit „+" oder „-".

Ein einfaches Beispiel für diese beiden Voraussetzungen bilden die beiden expandierenden Bündel $\{(0 \sqcup 1)^{-\infty} 0 \circledcirc\}$' und $\{(0 \sqcup 1)^{-\infty} 1 \circledcirc\}$' der 01–Partition. Die Gültigkeit der Voraussetzungen 8.3.1 und 8.3.2 folgt sofort aus den induktiven Regeln, mit denen die räumliche Ordnung in Gl. 8.8 bis 8.10 definiert wurde. Anschaulich bedeuten diese Voraussetzungen, daß nur die expandierenden Blätter im Gebiet $\circledcirc 0$ ihre senkrechte Reihenfolge ändern, wenn sie durch f abgebildet werden; die Reihenfolge der expandierenden Blätter in $\circledcirc 1$ bleibt erhalten. Außerdem liegen alle expandierenden Blätter in $1\circledcirc$ oberhalb derer in $0\circledcirc$.

Jetzt kann man auch klarer sehen, wie die Zerlegung von expandierenden Bündeln dabei helfen kann, die Voraussetzungen 8.3.1 und 8.3.2 zu erfüllen. Wenn diese Voraussetzungen für eine bestimmte Wahl der expandierenden Bündel verletzt sind, dann kann man versuchen, alle Blätter eines Überganges $A_k \circledcirc \to A_l \circledcirc$ zusammenzufassen, die den falschen Orientierungswechsel oder die falsche senkrechte Lage haben. Für diese Blätter kann man dann einen separaten Übergang erzeugen, indem man das größere expandierende Bündel $\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} A_k \circledcirc\}$' in die kleineren expandierenden Bündel $\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} A_k^1 \circledcirc\}$' und $\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} A_k^2 \circledcirc\}$' zerlegt (mit $A_k^1 \circledcirc \cup A_k^2 \circledcirc = A_k \circledcirc$). Es ist nicht garantiert, daß man durch endlich viele derartige Zerlegungen die beiden Voraussetzungen erfüllen kann. Aber zumindest werden die Voraussetzungen 8.3.1 bis 8.3.2, falls sie für die größeren Bündel $\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} A_k \circledcirc\}$' erfüllt sind, durch solche Zerlegungen nicht verletzt. Hierbei zeigt sich wieder der Vorteil bei der Benutzung von zwiefachen Partitionen: Man kann in getrennten Rechenschritten zunächst die expandierenden Bündel hinreichend klein wählen und dann schrittweise vergrößern, ohne daß man die expandierenden Blätter zu ändern braucht.

8.3.3 Ausdehnung der räumlichen Ordnung

Bisher wurden Ordnungsrelationen nur zwischen expandierenden Bündeln definiert, die in demselben kleinen expandierenden Bündel liegen. Wenn man eine umfassende senkrechte Ordnung sucht, muß man diese Lagebeziehungen auf möglichst viele expandierenden Blätter ausdehnen. Dies geschieht durch das Vereinigen von Bündeln, wobei die Ordnungen der einzelnen Bündel in geeigneter Weise kombiniert werden müssen. Dabei müssen manchmal die Richtungen „oben" und „unten" in einem der Bündel vertauscht werden. Bei diesen rein algebraischen Rechnungen braucht die Geometrie des Phasenraumes nicht mehr betrachtet zu werden, denn alle notwendige Information steckt in den Beschriftungen des Graphen in Abb. 8.9. Ich werde diese Rechenmethode allerdings nicht im allgemeinen Fall, sondern nur anhand eines Beispieles darstellen.

Die Vereinfachung der AB–Partition ergab in Kap. 5.4 als Endergebnis 5.51 die Partition

$$\{A_1\circledcirc, \ldots A_m\circledcirc\} = \{AA\circledcirc A, \quad AAA\circledcirc B, \quad AAB\circledcirc, \quad BB\circledcirc, \quad BA\circledcirc, \quad BAA\circledcirc B \cup BAB\circledcirc A\} \quad , \quad (8.15)$$

welche die Vergangenheit beschreibt. Die zeitlichen Übergänge zwischen diesen $m = 6$ Symbolischen Gebieten zeigt noch einmal Abb. 8.9. Die entsprechenden sechs expandierenden Bündel wurden bereits in Kap. 6.6 zu zwei größeren expandierenden Bündeln vereinigt. Dabei wurde jedoch nicht auf die senkrechte Ordnung geachtet.

Diese beiden größeren expandierenden Bündel erwiesen sich als die Bündel $\{(0 \sqcup 1)^{-\infty} 0 \circledcirc\}$' und $\{(0 \sqcup 1)^{-\infty} 1 \circledcirc\}$' der 01–Partition. Die Blätter dieser Partition wurden bereits durch Gl. 8.8 bis 8.10 geordnet. Doch auch wenn wir diese senkrechte Ordnung noch nicht kennen würden, könnten wir sie bei der Vereinigung der expandierenden Bündel schrittweise konstruieren. Dies ist der Inhalt der folgenden Beispielrechnung.

Zunächst muß jedes einzelne dieser sechs expandierenden Bündel senkrecht geordnet werden, so daß Voraussetzung 8.3.1 und 8.3.2 erfüllt sind. Bei diesem vorbereitenden Schritt muß man die Lage der expandierenden Blätter im Phasenraum betrachten. Wenn man von der unbekannten Feinstruktur des seltsamen Attraktors absieht, dann scheint man die senkrechten Blätter in den einzelnen Bündeln tatsächlich senkrecht ordnen zu können. Diese Ordnung kann durch die Beschriftungen „+", „−" und „U_i" ausgedrückt werden, die im letzten Abschnitt eingeführt wurden. Die Übergänge $A_i \circledcirc \to A_j \circledcirc$ in Abb. 8.9 sind auf diese Weise beschriftet.

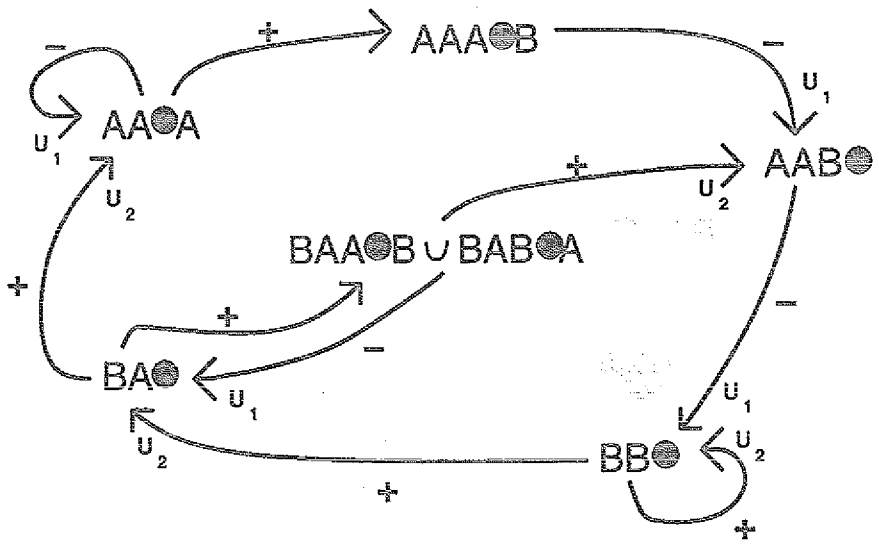


Abbildung 8.9: Die anfängliche senkrechte Ordnung auf kleinen expandierenden Bündeln. Der Graph ist derselbe wie in Abb. 5.16 und 6.2. Nur die Beschriftung der Übergänge ist anders. Ein Pluszeichen bedeutet, daß die senkrechte Reihenfolge der expandierenden Blätter bei diesem Übergang erhalten bleibt. Ein Minuszeichen bedeutet, daß sich diese Reihenfolge umkehrt. Wenn zwei Übergänge zu einem expandierenden Bündel führen, dann sind sie außerdem mit U_1 und U_2 beschriftet. Die expandierenden Blätter des mit U_2 beschrifteten Überganges kommen oberhalb der anderen expandierenden Blätter zu liegen.

Diese Beziehungen sind nicht die einzig möglichen, die von der Geometrie des Phasenraumes erlaubt werden. Sie hängen von der willkürlichen Wahl der Orientierung in jedem einzelnen Bündel ab und ändern sich, wenn man die Richtungen „oben" und „unten" in einem der Bündel vertauscht. In Abb. 8.9 wurde die Richtung „oben" so gewählt, daß in jedem der sechs Bündel $\{(A \sqcup B)^{-\infty}\alpha\}$' gilt: $U(B\alpha) > U(A\alpha)$. (Dabei ist α die Bezeichnung der entsprechenden Ecke in dem Graphen.)

Dadurch ist gleichzeitig festgelegt, welche der zu einem Bündel führenden Übergänge die höheren Blätter erzeugen: Betrachten wir als Beispiel die beiden Übergänge $BB\circ \to BB\circ$ und $AAB\circ \to BB\circ$ zu der Ecke $BB\circ$. Wegen $U(\ldots BBB\circ) > U(\ldots AABB\circ)$ ist der Übergang $BB\circ \to BB\circ$ mit U_2 und der andere mit U_1 beschriftet.

Die Aufgabe besteht nun darin, diese räumlichen Ordnungen von den einzelnen expandierenden Bündeln zu einer globalen Ordnung aller expandierenden Blätter auszudehnen. Dazu werden die expandierenden Bündel miteinander vereinigt. Die Auswahl der zu vereinigenden Bündel geschieht wie in Kap. 6.6. Auch in den Bündeln, die durch die Vereinigung entstehen, soll eine senkrechte Ordnung definiert sein. Deshalb muß man darauf achten, die beiden Voraussetzungen 8.3.1 und 8.3.2 bei der Vereinigung nicht zu verletzen.

Betrachten wir die erste Vereinigung, die in Kap. 6.6.2 vorgenommen wurde, und die erste Voraussetzung 8.3.1. Die Vereinigung wurde dadurch ausgedrückt, daß Übergänge miteinander identifiziert wurden. Als erstes wurde der Übergang $BB\circ \stackrel{+}{\to} BB\circ \stackrel{+}{\to} BB\circ$ mit dem Übergang $BAA\circ B \cup BAB\circ A \stackrel{+}{\to} AAB\circ \stackrel{-}{\to} BB\circ$ identifiziert, sowie der Übergang $BB\circ \stackrel{+}{\to} BA\circ$ mit $BAA\circ B \cup BAB\circ A \stackrel{-}{\to} BA\circ$. Offensichtlich unterscheiden sich diese Übergänge darin, ob sie die Orientierung erhalten oder umkehren. Um Voraussetzung 8.3.1 zu erhalten, müssen diese Unterschiede vor der Vereinigung beseitigt werden, indem die Richtungen „oben" und „unten" auf einem oder mehreren der Bündel vertauscht werden. In diesem Beispiel genügt es, die Orientierung von $\{(A \sqcup B)^{-\infty} BB\circ\}$' umzudrehen. Dabei ändern sich die Beschriftungen von Abb. 8.9 zu denen von Abb. 8.10. Wie man leicht sieht, folgen die neuen Beschriftungen aus den alten, ohne daß die Geometrie des Phasenraumes erneut betrachtet werden muß. Nun können die drei Bündel $\{(A \sqcup B)^{-\infty}(BAA\circ B \cup BAB\circ A)\}$', $\{(A \sqcup B)^{-\infty} AAB\circ\}$' und $\{(A \sqcup B)^{-\infty} BB\circ\}$' zu $\Sigma = \{(A \sqcup B)^{-\infty}((A \sqcup B)BB\circ E$

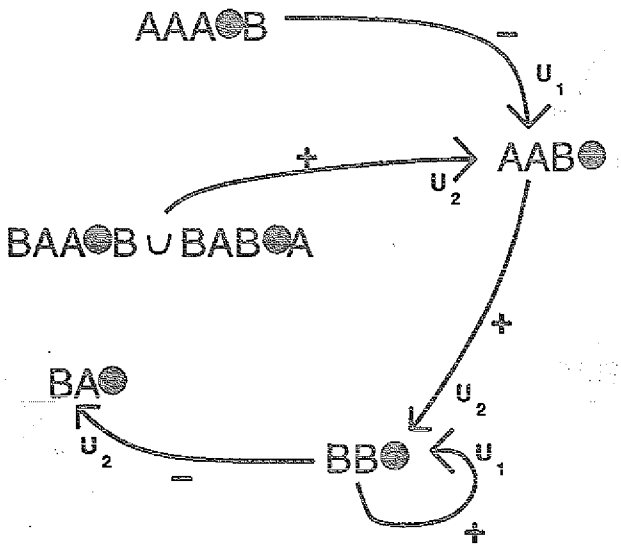


Abbildung 8.10: Die senkrechte Ordnung, nachdem die Orientierung im Gebiet $BB\circ$ umgedreht wurde. Nur ein Ausschnitt des Graphen 8.9 wird gezeigt. Geändert hat sich die Beschriftung „+" oder „−" nur auf den Übergängen, die entweder zu $BB\circ$ hin oder von $BB\circ$ weg führen, und außerdem die Beschriftung U_i der Übergänge, die zu $BB\circ$ hin führen.

$\sqcup AAB\circ E \sqcup (BAA\circ B \cup BAB\circ A))\}$' vereinigt werden, ohne daß Voraussetzung 8.3.1 verletzt wird. Abb. 8.11 zeigt das Teilergebnis, nachdem zunächst nur die beiden Bündel $\{(A \sqcup B)^{-\infty} AAB\circ\}$' und $\{(A \sqcup B)^{-\infty} BB\circ\}$' vereinigt wurden. Ihre Vereinigung wird durch die etwas ungewöhnlich bezeichnete Ecke $AAB \circ \sqcup BB\circ$ repräsentiert. Wieder folgen die Beschriftungen aller Übergänge aus denen der Abb. 8.10.

Weiterhin wurde in Kap. 6.6.2 der Übergang $BA\circ \xrightarrow{+} AA\circ A$ mit $AA\circ A \xrightarrow{-} AA\circ A$ identifiziert sowie der Übergang $BA\circ \xrightarrow{+} \Sigma \xrightarrow{+} \Sigma$ mit $AA\circ A \xrightarrow{+} AAA\circ B \xrightarrow{-} \Sigma$. Voraussetzung 8.3.1 wird dabei nicht verletzt, wenn zuvor die Orientierung in Σ und in $\{(A \sqcup B)^{-\infty} AAA\circ B\}$' umgedreht wird. Als Endergebnis ergeben sich die beiden Bündel in Gl. 6.99 und 6.100, die von den beiden Ecken in Abb. 8.12 repräsentiert werden.

Auch diese beiden Bündel könnte man noch trivialerweise vereinigen, indem man jeweils die mit „+" beschrifteten Übergänge und die mit „-" beschrifteten Übergänge identifiziert, so daß nur noch eine Ecke und zwei Übergänge übrigbleiben. Diese Vereinigung, die noch über die Rechenschritte von Kap. 6.6 hinausgeht, erzeugt dann schließlich eine globale senkrechte Ordnung aller expandierenden Blätter. Diese stimmt tatsächlich mit der senkrechten Ordnung der 01–Partition in Gl. 8.8 bis 8.10 überein. Allein aufgrund der Voraussetzung 8.3.1 wurden daher die Richtungen „oben" und „unten" in den einzelnen Bündeln so vertauscht, daß am Ende die richtige senkrechte Ordnung entstand.

Bleibt auch die andere Voraussetzung 8.3.2 erfüllt? Auch dies kann man durch bloße Symbolmanipulationen nachprüfen, ohne auf die Geometrie des Phasenraumes zurückzugreifen. Für die senkrechte Ordnung U', die von Abb. 8.10 angegeben wird, gilt zum Beispiel: $U'(\ldots BBB\circ) < U'(\ldots AAABB\circ) < U'(\ldots BAABB\circ)$. Nach der folgenden Vereinigung der Bündel $\{(A \sqcup B)^{-\infty} AAB\circ\}$' und $\{(A \sqcup B)^{-\infty} BB\circ\}$' werden daraus die in Abb. 8.11 gezeigten Lagebeziehungen $U_1 < U_2 < U_3$ der drei Übergänge zum vereinigten Bündel.

Die ausführliche Rechnung zeigt, daß bei einem Zwischenstand (jedoch nicht am Ende) der Rechnung die Voraussetzung 8.3.2 verletzt wird, und zwar nach der Identifizierung der beiden Übergänge $AAB\circ \sqcup BB\circ \to AAB\circ \sqcup BB\circ$ und $(BAA\circ B \cup BAB\circ A) \to AAB\circ \sqcup BB\circ$ in Abb. 8.11, die mit U_1 und U_3 beschriftet sind. Denn der Übergang, der durch diese Identifizierung entsteht, liegt weder höher noch tiefer als der mit U_2 beschriftete Übergang $AAA \circ B \to AAB \circ \sqcup BB\circ$.

Um die Voraussetzung 8.3.2 bei jedem Zwischenstand der Rechnung zu erhalten, muß man

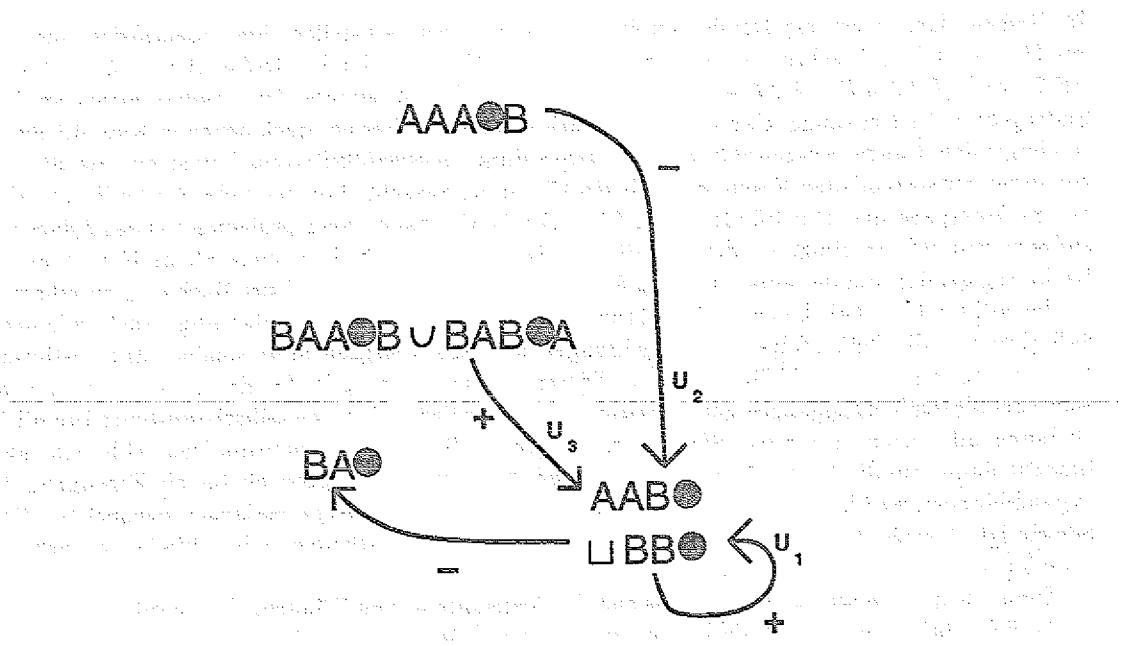


Abbildung 8.11: Die senkrechte Ordnung, nachdem die beiden expandierenden Bündel $\{(A \sqcup B)^{-\infty} AAB\circ\}$' und $\{(A \sqcup B)^{-\infty} BB\circ\}$' aus Abb. 8.10 zu $\{(A \sqcup B)^{-\infty}(AAB \sqcup (A \sqcup B)BB)\circ\}$' vereinigt wurden. Zu diesem neuen expandierenden Bündel führen drei Übergänge. Weil in Abb. 8.10 der Übergang $AAB\circ \to BB\circ$ die senkrechte Orientierung erhält, liegen auch in dem neuen Graphen die expandierenden Blätter, die von $BAA \circ B \cup BAB \circ A$ her stammen, oberhalb der Blätter, die von $AAA \circ B$ her stammen: $U_3 > U_2$.

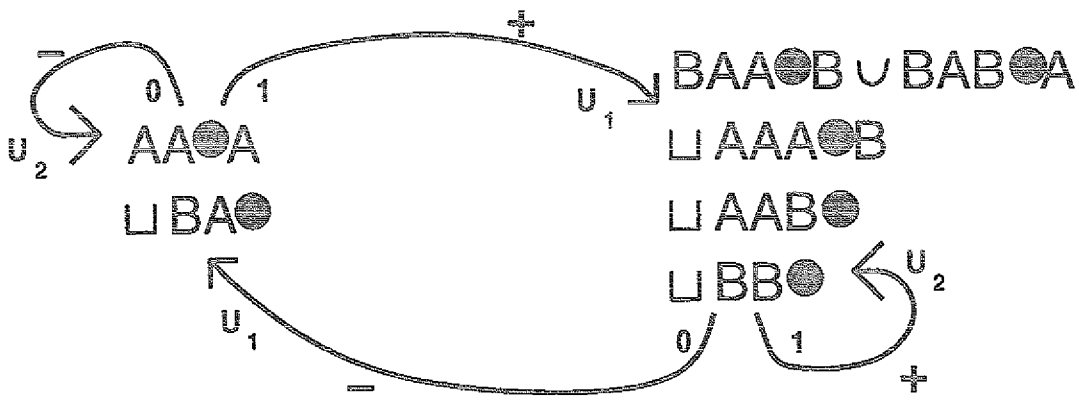


Abbildung 8.12: Die senkrechte Ordnung, nachdem alle expandierenden Blätter aus Abb. 8.9 zu den beiden Bündeln in Gl. 6.99 und 6.100 vereinigt wurden. Zusätzlich sind die Beschriftungen mit 0 und 1 aus Abb. 6.2 übernommen. Tatsächlich wurden während der Ausdehnung der senkrechten Ordnung nur Übergänge mit gleicher Beschriftung 0 oder 1 identifiziert. Das linke Bündel ergibt sich als $\{(0 \sqcup 1)^{-\infty} 0\circ\}$', das rechte Bündel als $\{(0 \sqcup 1)^{-\infty} 1\circ\}$'.

die Reihenfolge, in der die Bündel vereinigt werden, etwas umstellen: Man identifiziert die drei mit U_1, U_2 und U_3 beschrifteten Übergange $AAB \circ \sqcup BB\circ \rightarrow AAB \circ \sqcup BB\circ$, $AAA \circ B \rightarrow AAB \circ \sqcup BB\circ$ und $(BAA \circ B \cup BAB \circ A) \rightarrow AAB \circ \sqcup BB\circ$ alle auf einmal. Dies ändert nichts an dem Endergebnis der Rechnung. Und es verletzt auch nicht die Kriterien, nach denen in Kap. 6.6 die zu vereinigenden Bündel ausgewählt wurden. Denn dieses Auswahlkriterium betraf nur die Bündel, von deren entsprechenden Ecken mehr als ein Übergang ausgeht. Von der Ecke $AAA \circ B$ geht aber nur ein Übergang aus. Das Bündel $\{(A \sqcup B)^{-\infty} AAA \circ B\}'$ kann daher problemlos etwas früher mit anderen Bündeln vereinigt werden, als dies in Kap. 6.6 geschah. Und diese kleine Korrektur der Rechnung genügt, um die Voraussetzung 8.3.2 bei jedem Zwischenstand der Rechnung zu erhalten.

Im allgemeinen Fall könnten die Voraussetzungen 8.3.1 und 8.3.2 allerdings auch erfordern, daß nicht nur die Reihenfolge der Vereinigungen umgestellt werden muß, sondern daß bestimmte Vereinigungen ganz unterbleiben müssen. Dieser Fall muß zwangsläufig eintreten, wenn man als Endergebnis der Vereinigungen eine Partition erhalten würde, die keine selbstkonsistente räumliche Ordnung aller expandierenden Blätter zuläßt, wie z. B. die AB–Partition von Abb. 8.6. Man braucht dann, um die räumliche Ordnung zu definieren, mehr Symbole, als für die Erzeugung der expandierenden und kontrahierenden Blätter nötig sind. Ein weiteres, einfaches Beispiel für diese Schwierigkeit zeigt Abb. 8.13, in der sich die Blätter zweier expandierender Bündel miteinander verzahnen.

Dieses Kapitel zeigte, daß man während der Vereinigung von Bündeln die räumliche Ordnung mitberücksichtigen und so die Ordungsrelationen auf möglichst viele Blätter ausdehnen kann. Mit etwas Mühe kann man die räumliche Ordnung auch während anderer Verfahren berücksichtigen, zum Beispiel während der Vereinfachung einer zwiefachen Partition oder während der Rekonstruktion einer nicht zwiefachen Partition. Vermutlich läßt sich jedes Verfahren, das auf expandierenden und kontrahierenden Bündeln operiert, durch eine solche Konstruktion einer räumlichen Ordnung ergänzen. Aber die räumliche Ordnung ist eben nur eine Ergänzung. Die Vereinfachung von generierenden Partitionen gelang in den vorherigen Kapiteln, ohne daß die räumliche Ordnung mitberücksichtigt wurde.

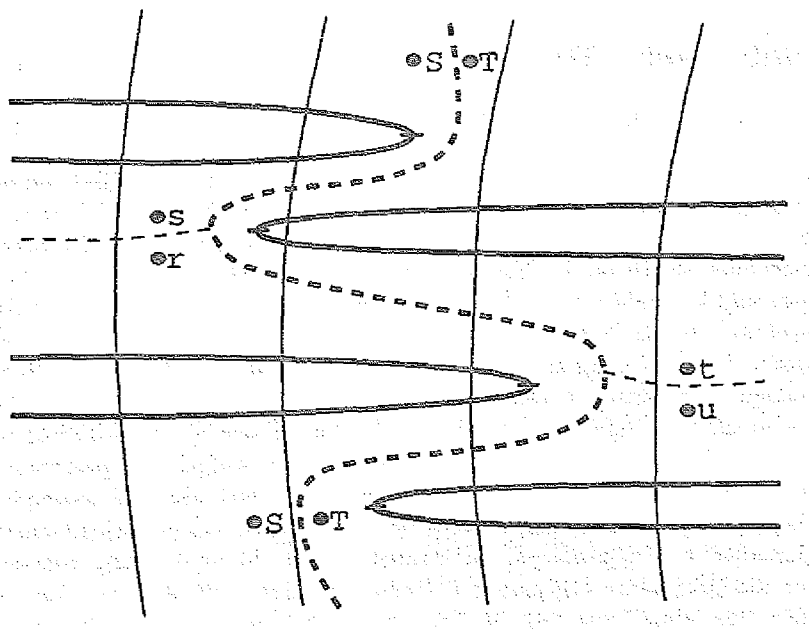


Abbildung 8.13: Schematisches Beispiel dafür, daß sich die senkrechte Ordnung nur mit zusätzlichen Symbolen beschreiben läßt. Gezeigt ist ein Auschnitt eines Attraktors mit expandierenden Blättern (dicke waagrechte Linien) und kontrahierenden Blättern (dünne senkrechte Linien). Selbst wenn in diesem Ausschnitt zwei Symbolische Gebiete $\circ S$ und $\circ T$ genügen, um eine generierende Partition zu konstruieren, so braucht man doch eine feinere Unterteilung in vier Symbolische Gebiete $\circ r$, $\circ s$, $\circ t$ und $\circ u$ (gestrichelte Linien), um die senkrechte Ordnung der expandierenden Blätter beschreiben zu können.

Alle Arbeiten, welche auf Nachbargebiete übergreifen, (...) sind mit dem resignierenden Bewußtsein belastet: daß man allenfalls dem Fachmann nützliche Fragestellungen liefert, auf die dieser von seinen Fachgesichtspunkten aus nicht so leicht verfällt, daß aber die eigene Arbeit unvermeidlich höchst unvollkommen bleiben muß.
aus Max Weber: „Wissenschaft als Beruf"; München 1919

Kapitel 9

Ausblick: Neuronale Netze

9.1 Symbolische Dynamik und Neuronale Netze

9.1.1 Das neuronale Netz soll eine generierende Partition finden

Damit ist der zentrale Teil dieser Doktorarbeit beendet. Wurden die Ziele erreicht, die in der Einleitung aufgestellt wurden? Nur teilweise: In einfachen, aus Veröffentlichungen bekannten dynamischen Systemen konnte ich tatsächlich eine an die Dynamik angepaßte Strukturierung der Dynamik ausrechnen. Doch die Rechnung wird umso schwieriger, je mehr Symbole oder Bündel expandierender und kontrahierender Blätter vorhanden sind. Da es etwa $(1/\epsilon)^{d_a}$ Symbolische Gebiete des Durchmessers ϵ erfordert, um einen Attraktor der Dimension d_a zu überdecken, wächst die Schwierigkeit der Rechnung rasch mit der Dimension d_a. Ich habe deshalb nicht einmal in dreidimensionalen Phasenräumen versucht, generierende Partitionen zu vereinfachen.

Weil der zentrale und aufwendigste Teil der Rechnung sich als Algorithmus formulieren läßt (siehe Kap. 7.2.1), läge es nahe, einen Computer für diese Aufgabe zu programmieren. Da dies ein sehr großes numerisches Projekt wäre, sollte man sich erst über die passende Rechnerarchitektur im Klaren sein. Die Rechnung besteht im wesentlichen aus der schrittweisen Veränderung von kontrahierenden und expandierenden Bündeln, die recht unabhängig voneinander sind und sich nur durch die Orbitpaare gegenseitig beeinflussen. Eine große Zahl von kontrahierenden und expandierenden Bündeln könnte deshalb am besten bewältigt werden, wenn jedes dieser Bündel von einem eigenen Prozessor eines Parallelrechners gespeichert und manipuliert werden würde.

Ein extrem paralleler „Rechner", der dynamische Eingaben analysiert und speichert, wurde bereits in der Einleitung erwähnt: Das Gehirn, und insbesondere der Neocortex der Säugetiere. Der Neocortex verarbeitet Sinnesdaten visuellen, auditorischen und somatosensorischen Ursprungs, also die besonders dynamischen Sinnesdaten. Er hat aber so erstaunliche Fähigkeiten und so komplexe Komponenten, daß man kaum erwarten sollte, ihn mit Hilfe der relativ elementaren mathematischen Konzepte der heutigen Nichtlinearen Dynamik verstehen zu können. Deshalb habe auch ich mich zunächst nicht für dieses neurobiologische Vorbild interessiert und einfach die groben Züge einer Rechnerarchitektur entworfen, die den mathematischen Prinzipien der Symbolischen Dynamik gemäß ist. Ich war äußerst überrascht, als diese Architektur der Anatomie des Neocortex ähnelte. Aus dieser Überraschung ist der Anhang D entstanden, der ein Modell für die Funktionsweise der Pyramidenzellen im Neocortex vorschlägt.

Nun ist die Architektur des Neocortex vielleicht eine effiziente, aber sicher nicht die einzig mögliche Weise, wie die Prinzipien der Symbolischen Dynamik implementiert werden können. Während ich im folgenden zeige, wie man Schritt für Schritt von den mathematischen Prinzipien zur neocortikalen Architektur kommt, werde ich deshalb auch alternative Vorgehensweisen erwähnen. Den bedeutenden, fundamentalen Fragen über den Neocortex und die Analyse dynamischer Zeitreihen werde ich hier nicht gerecht; die ganze Argumentation bleibt so oberflächlich und spekulativ, wie man es von einem „Ausblick" erwartet. Viele Argumente beruhen nicht nur auf mathematischer Notwendigkeit, sondern auch auf meiner persönlichen Meinung. Immerhin wird es zuletzt möglich

sein, aus der neocortikalen Anatomie etwas über die effiziente Analyse dynamischer Zeitreihen zu lernen. Eine chaotische Zeitreihe ist dabei wohl der schwierigste Fall, weil sie nur sehr begrenzt vorhersagbar ist. Doch auch nicht chaotische Zeitreihen sollten verarbeitet werden können. Auch dabei können die Konzepte der Symbolischen Dynamik hilfreich sein.

9.1.2 Neuronen sollen Symbolische Gebiete repräsentieren

Die Parallelprozessoren werden im folgenden einfach als Neuronen bezeichnet. Es gibt viele Weisen, wie Neuronen durch ihre Aktivität die Lage des Orbits im Phasenraum repräsentieren können. Im einfachsten Fall ist jedem Neuron ein „Aktivitätsgebiet" im Phasenraum zugeordnet. Während der Orbit durch dieses Aktivitätsgebiet läuft, wird das Neuron durch die sensorische Eingabe des Netzwerkes und durch die anderen Neuronen erregt und ist „aktiv", das heißt, es feuert schnell; ansonsten ist es „passiv". Wenn man zu einem Zeitpunkt alle aktiven Neuronen kennt, dann kann man aus dem Schnitt ihrer Aktivitätsgebiete die gegenwärtige Lage des Orbits folgern.

Diese einfache Repräsentation ist aus der Theorie der Neuronalen Netze vertraut, nur daß man dort von einem „Zustandsraum" statt von einem Phasenraum spricht. Die Analogie zur Symbolischen Dynamik ist offensichtlich: Wenn $\{\circledast A_1, \ldots \circledast A_m\}$ eine generierende Partition der invertierbaren Abbildung f ist und wenn jedes Neuron N_i^t $(i, t \in \mathbb{Z},\ 1 \leq i \leq m)$ das Aktivitätsgebiet $f^t(\circledast A_i)$ hat, dann bestimmt der Schnitt der Aktivitätsgebiete aller aktiven Neuronen die Lage des Orbits beliebig genau. So definierte Neuronen sind auf eine äußerst stereotype Weise aktiv. Ein Orbit $(x_t)_{t\in\mathbb{Z}}$ mit $x_0 \in \circledast A_k$ aktiviert zum Zeitpunkt 0 das Neuron N_k^0, zum Zeitpunkt 1 das Neuron N_k^1, dann das Neuron N_k^2, und so weiter. Solche Neuronen, die immer in der gleichen Reihenfolge aktiv sind, sollten sinnvollerweise zu einem Neuron zusammengefaßt werden, das über einen großen Zeitraum T aktiv ist. Dieses Neuron hat das Aktivitätsgebiet

$$\circledast C_k \quad := \quad \circledast A_k \quad \cup \quad A_k \circledast \quad \cup \quad A_k E \circledast \quad \cup \quad \ldots A_k E^{T-2} \circledast \quad . \tag{9.1}$$

Umgekehrt möchte man auch die Symbolischen Gebiete $\circledast A_k$ durch die neuronalen Aktivitätsgebiete ausdrücken können, damit letztere Gebiete die Lage des Orbits ebenso genau festlegen wie die ersteren Gebiete. Dies ist möglich, wenn es nicht nur Neuronen mit Aktivitätsgebiet $\circledast C_k$ gibt, sondern auch solche mit zeitverschobenen Aktivitätsgebieten $f^{-t}(\circledast C_k)$ (für $1 \leq t \leq T-1$). Denn dann gilt, sofern die auf der rechten Seite von Gl. 9.1 stehenden Gebiete nicht überlappen:

$$\circledast A_k = \bigcap_{t=0}^{T-1} f^{-t}(\circledast C_k) \quad . \tag{9.2}$$

Derartige Aktivitätsgebiete $\circledast C_k$ und $f^{-t}(\circledast C_k)$ sind dadurch ausgezeichnet, daß ein Orbit, der in sie eintritt, anschließend für eine lange Zeit T in ihnen bleibt. Bei kontinuierlicher Zeit sehen diese Aktivitätsgebiete aus wie lange, schmale Röhren, die sich in Richtung des Phasenraumflusses erstrecken (Abb. 9.1).

Somit wurde ein erstes grundlegendes Konzept der Symbolischen Dynamik für die Konstruktion des neuronalen Netzes verwendet: Um die Lage eines Orbits festzulegen, wird nicht nur das eine Symbolische Gebiet $\circledast A_i$ verwendet, sondern auch alle dessen Bilder $f^t(\circledast A_i)$. Hieraus folgen röhrenförmige Aktivitätsgebiete für die Neuronen. Der nächste Abschnitt wird andeuten, wie eine derartige Gestalt der Aktivitätsgebiete erlernt werden kann. Eigentlich werden auch die Aktivitätsgebiete $\circledast C_k$, $f(\circledast C_k)$, $f^2(\circledast C_k)$ etc. von Orbits in einer stereotypen Reihenfolge durchlaufen. Ich sehe aber keinen Grund, warum die Neuronen solche stereotypen Reihenfolgen besitzen sollten. Deshalb werde ich im folgenden zwar annehmen, daß die Neuronen lange röhrenförmige Aktivitätsgebiete haben, die sich gegenseitig stark überlappen und zusammen die Lage des Orbits charakterisieren, doch ich werde nicht annehmen, daß diese Aktivitätsgebiete alle dieselbe zeitliche Länge T haben, daß es unendlich viele Neuronen gibt oder daß es Neuronen gibt, die immer in einem festen zeitlichen Abstand Δt voneinander aktiv werden.

Über welchen Zeitraum T soll sich ein Aktivitätsgebiet $\circledast C_i$ erstrecken? Das hängt vom Rauschen ab. Jede sensorische Eingabe ist rauschbehaftet, so daß alle Strukturen der Aktivitätsgebiete,

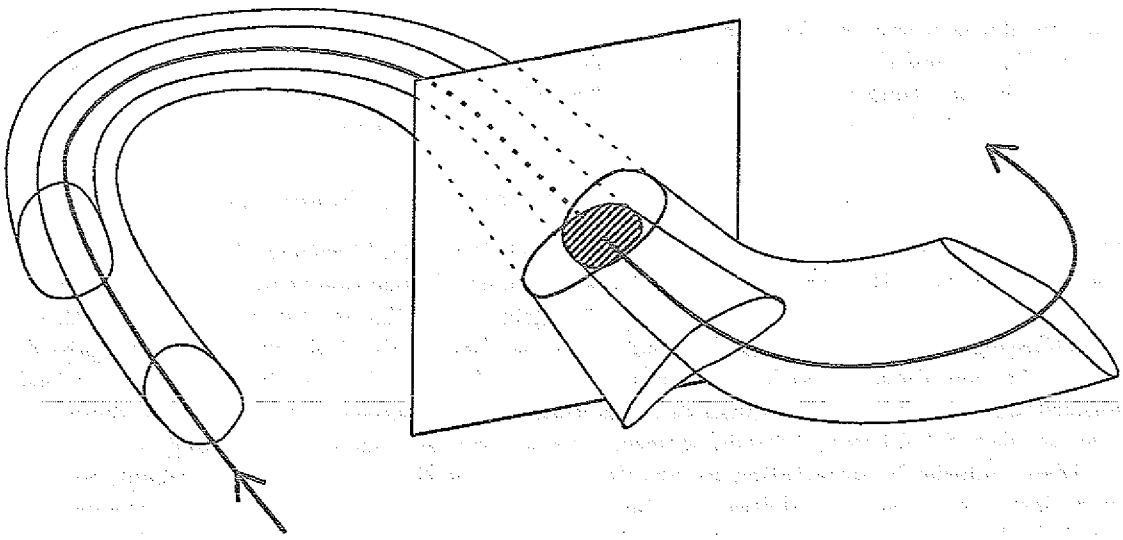


Abbildung 9.1: Aktivitätsgebiete von Beispielneuronen. Die sensorischen Daten bestimmen im hochdimensionalen Phasenraum einen Orbit $z(t)$; $t \in \mathbb{R}$. Dieser Orbit (dicke Linie) durchquert die überlappenden Aktivitätsgebiete zweier Neuronen (umrahmt von dünnen Linien). Diese Aktivitätsgebiete verlaufen ungefähr röhrenförmig in Flußrichtung, so daß jedes Neuron lange aktiv bleibt, nachdem es einmal aktiviert wurde. Wie der schraffierte Teil des Poincaré Schnittes zeigt, wird die Lage des Orbits durch die gemeinsame Aktivität beider Neuronen genauer repräsentiert als durch die Aktivität jedes einzelnen Neurons. Weitere aktive Neuronen könnten die Genauigkeit der Repräsentation verbessern, sofern die Aktivitätsgebiete wie bei einer generierenden Partition überlappen, also wie in Abb. 9.2.a, und nicht wie in Abb. 9.2.b.

die kleiner sind als die Rauschamplitude, nicht mehr unterschieden werden können. Auch wenn die Symbolischen Gebiete $\circ A_i$ selbst keine solchen filigranen Strukturen besitzen, so wird eine chaotische Dynamik f diese Gebiete im Laufe der Zeit doch so strecken und falten, daß die Gebiete $f^t(\circ A_i)$ für großes $|t|$ filigrane Strukturen besitzen. Solche filigranen Gebiete zu den Aktivitätsgebieten $\circ C_i$ hinzuzufügen ist überflüssig. Auch ohne sie ist die Lage des Orbits durch alle aktiven Neuronen so genau festgelegt, wie es das Rauschen erlaubt.

9.2 Lernprozesse

9.2.1 Überblick über die Lernprozesse

Zwei fundamentale Ziele müssen von den Lernprozessen erreicht werden: Erstens müssen die Neuronen die röhrenförmige Aktivitätsgebiete erlernen. In anderen Worten: Sie müssen lernen, nicht auf unspezifische und flüchtige, sondern auf spezielle, längerfristig relevante Sinnesreize zu reagieren. Und zweitens muß das ganze Netzwerk lernen, durch das Zusammenspiel der Neuronen sinnvolle Funktionen zu erfüllen. Die einfachste Funktion ist das Speichern der zeitlichen Reihenfolge von Sinnesreizen, was der Speicherung grammatikalischer Regeln der Symbolischen Dynamik entspricht. Dies geschieht in den meisten künstlichen neuronalen Netzen, wie vermutlich auch in biologischen Netzwerken, durch eine Variante der „Hebbschen Lernregel": Wenn die Aktivitätsgebiete zweier Neuronen so überlappen, daß das erste oft kurz vor dem zweiten aktiv ist, dann lernt das erste Neuron, das zweite zu erregen (siehe Abb. D.1). Durch solche internen Erregungen werden die Neuronen des Netzwerkes etwas unabhängiger von unzuverlässigen Sinneswahrnehmungen.

Doch auf welche Reize sollen die Neuronen reagieren? Manche Neuronen im Neocortex reagieren auf sehr komplexe Reize, zum Beispiel auf verschiedene Ansichten eines Gesichtes oder einer Hand. Die Bedeutung dieser Reize für das Verhalten ist offensichtlich. In der Tat besteht

ein zentrales Problem der künstlichen Intelligenz darin, abstrakte Begriffe zu finden, die relevante Informationen aus einer Flut von Daten zusammenfassen. Es wäre gut, wenn solche Begriffe nicht einprogrammmiert werden müßten, sondern vom Netzwerk eigenständig gelernt werden könnten.

In der Einleitung wurde dieses Lernziel als eine an die Dynamik angepaßte Strukturierung der Dynamik bezeichnet. In Anhang D unterscheide ich zwei Lernprozesse.

- Bei dem „Erlernen langfristiger Phänomene" ändern sich Aktivitätsgebiete der Neuronen, bis sie röhrenförmig und möglichst lang sind. „Langfristige Phänomene" können daher auch als langfristige Prozesse betrachtet werden, die immer auf die gleiche Weise sequentiell ablaufen.
- Bei dem „Lernen im Kontext" stimmen die Neuronen ihre Aktivitätsgebiete so aufeinander ab, daß sie zusammen die Lagen der Orbits möglichst genau beschreiben. Dieses Ziel ist dasselbe wie bei der Suche nach generierenden Partitionen.

Eine zentrale Lehre dieser Arbeit ist, daß Zukunft und Vergangenheit durch verschiedene Partitionen beschrieben werden sollten, damit sich die Aktivitätsgebiete flexibel ändern lassen. Dies impliziert eine Unterscheidung von zwei Neuronenmengen: Die einen Neuronen beschreiben die Vergangenheit und repräsentieren durch ihre gemeinsame Aktivität das expandierende Blatt, auf dem der Orbit liegt. Die anderen Neuronen beschreiben die Zukunft des Orbits, indem sie sein kontrahierendes Blatt repräsentieren. Es ist zweifellos gewagt, die Trennung von Zukunft und Vergangenheit, die sich bei der einfachen Hénon–Abbildung bewährt hat, deswegen auch bei komplexen Netzwerken zu erwarten. Doch tatsächlich werden auch die Pyramidalneuronen des Neocortex von Neurobiologen zu zwei großen Mengen zusammengefaßt, nämlich zu denen in den oberflächlichen neocortikalen Schichte 2, 3 und 4 und denen in den tiefen Schichten 5 und 6. Und Anhang D folgert aus den Verbindungen dieser Schichten, daß die Neuronen in den oberflächlichen Schichten die nahe Vergangenheit analysieren und daß die Neuronen in den tiefen Schichten die nahe Zukunft vorhersagen.

9.2.2 Erlernen langfristiger Phänomene

Zum ersten Lernziel: Wie können die Neuronen lernen, möglichst lange, röhrenförmige Aktivitätsgebiete zu haben? Der kritische Zeitpunkt ist der, an dem ein Orbit in ein Aktivitätsgebiet hineinläuft oder knapp an einem Aktivitätsgebiet vorbeiläuft und das Neuron sich entscheiden muß, ob es nun während der nächsten T Zeitschritte den Reiz signalisiert oder nicht. Diese „alles oder nichts"–Entscheidung hat ein Analogon in der Symbolischen Dynamik: Wenn ein Orbit eine Poincaré Ebene schneidet, die in Symbolische Gebiete unterteilt ist, dann folgt daraus ein Symbol, das bei einer unverrauschten Dynamik nicht nur während der nächsten T Zeitschritte, sondern sogar zu allen Zeiten verwendet wird, um die Lage des Orbits zu bestimmen.

Als Ort des Lernens gelten die „Synapsen", über die sich die Neuronen gegenseitig erregen. In Anhang D postuliere ich folgende Signalübertragungen und Lernprozesse: Gewisse Synapsen, nämlich die auf den „Basaldendriten", aktivieren das Neuron zu dem entscheidenden Zeitpunkt, wenn der Orbit in das Aktivitätsgebiet hineinläuft. Andere Synapsen, nämlich die auf dem „Apikaldendriten", sollen Signale in dem gesamten folgenden Zeitraum übertragen und dadurch die Dauer der Aktivität regulieren. Eine lang andauernde Aktivität wird wiederum von den basalen Synapsen als Bestätigung dafür aufgefaßt, im richtigen Augenblick aktiv gewesen zu sein. Durch das Wechselspiel von relativ einfachen, aber verschiedenen Lernprozessen in apikalen und basalen Synapsen kann das Neuron dann möglicherweise lernen, nur seltene, aber langandauernde Aktivität zu zeigen.

Ein Neuron wird ein ein möglichst langes röhrenförmigen Aktivitätsgebiet nur dann erlernen können, wenn es die Form dieser Röhre an die Expansion und Kontraktion des Phasenraumflusses anpaßt. Ideal ist es, wenn der Querschnitt der Röhre am Anfang in stabiler Richtung und am Ende der Röhre in instabiler Richtung ausgedehnt ist. Denn dann wird die Expansion und Kontraktion des Phasenraumflusses durch die Änderung des Querschnittes kompensiert und die Röhre kann besonders lang sein. Derartige Querschnitte sind uns bereits begegnet: Die Symbolischen Gebiete einer Partition, die schmale expandierende Bündel erzeugt, sind in instabiler Richtung ausgedehnt.

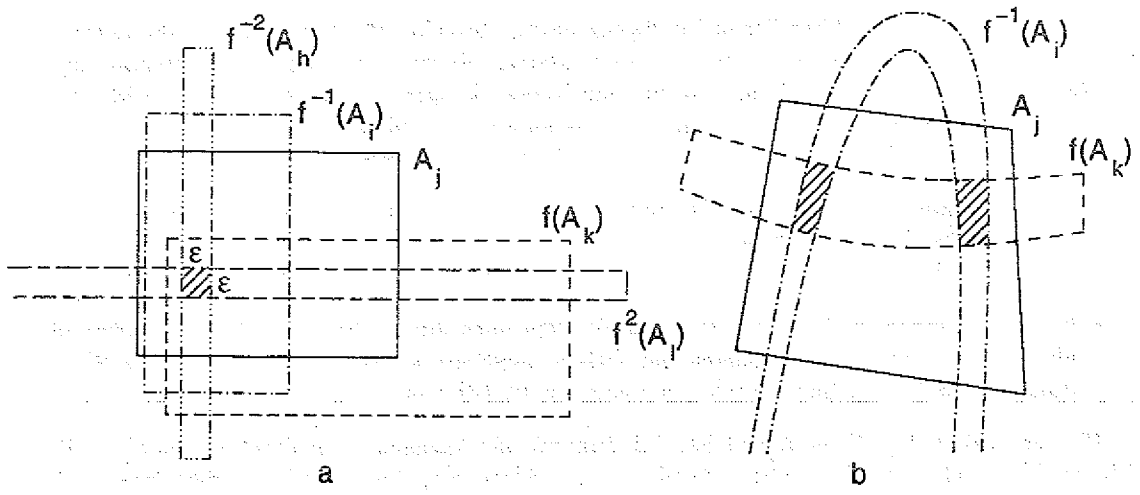


Abbildung 9.2: Wiederholung der Abb. 2.1. Die Orbits, welche die schraffierten Gebiete kreuzen, werden miteinander verwechselt.

Und die Symbolischen Gebiete einer Partition, die schmale kontrahierende Bündel erzeugt, sind in stabiler Richtung ausgedehnt. Insofern besteht eine gewisse Analogie zwischen Anfang und Ende des röhrenförmigen Aktivitätsgebietes und den beiden Teilpartitionen einer zwiefachen Partition.

Das hier vorgeschlagene Wechselspiel von apikalen und basalen Synapsen ist nicht die einzig mögliche Weise, wie langfristige Phänomene gelernt werden können. Eine andere Möglichkeit wäre, die Dynamik zu speichern, während man auf die Aktivität eines Neurons wartet. Nachdem das Neuron aktiv wurde, könnte man die Dynamik rückwärts abspielen, so daß das Neuron noch einmal die Signale erhält, die seiner Aktivitätsperiode vorausgehen, und lernen kann, auch auf diese zu reagieren. Eine solche Zeitinversion der Dynamik ist in technischen Systemen viel leichter zu realisieren als in biologischen Systemen. Die Zeitinversion hat allerdings den Nachteil, daß sie nur zum Erlernen langfristiger Phänomene taugt. Das Wechselspiel zwischen apikalen und basalen Synapsen in dem in Anhang D vorgeschlagenen Modell kann dagegen noch eine andere Funktion erfüllen, nämlich eine Art von motorischer Steuerung, durch die das Netzwerk die äußere Dynamik gezielt beeinflußt. Dabei teilt der Apikaldendrit dem Neuron mit, ob dessen vorausgegangene Aktivität die erwünschte steuernde Wirkung hat. Es werden in Anhang D.2 gleich zwei Varianten dieser Steuerung vorgeschlagen, die extrem vereinfachte Versionen des klassischen und des operanten Konditionierens sind.

9.2.3 Lernen im Kontext

Willkürlich gewählte röhrenförmige Aktivitätsgebiete würden die Lage eines Orbits ebenso schlecht beschreiben wie willürlich gewählte Symbolische Gebiete einer nicht generierenden Partition. Diese Problematik wurde bereits in Kap. 2.1 durch die Abb. 2.1 illustriert, die in Abb. 9.2 wiederholt wird. Die Neuronen sollen lernen, ihre Aktivitätsgebiete wie in Abb. 9.2.a, und nicht wie in Abb. 9.2.b zu legen, so daß nur ähnliche Orbits verwechselt werden. Doch ein Orbit der externen Dynamik wird die beiden schraffierten Gebiete, die in Abb. 9.2.b verwechselt werden, zu weit auseinander liegenden Zeitpunkten schneiden. Wie kann das Netzwerk überhaupt erkennen, daß an dieser Stelle im Phasenraum Orbits verwechselt werden können?

Eine Möglichkeit, solche Verwechslungen zu erkennen, besteht darin, ein zweites Netzwerk zu konstruieren, das sowohl die äußere Dynamik als auch die Dynamik des ersten Netzwerkes beobachtet und dabei kontrolliert, ob unähnliche Orbits verwechselt werden. Damit dieses zweite Netzwerk zuverlässig funktioniert, darf es selbst nur ähnliche Orbits verwechseln. Das Problem wurde also nur verlagert.

Eine zweite Möglichkeit besteht darin, zusätzliche Neuronen einzuführen, die so kleine Akti-

vitätsgebiete besitzen, daß sie nur sehr ähnliche Orbits miteinander verwechseln können. Die Zahl der notwendigen Aktivitätsgebiete wächst allerdings rapide mit abnehmendem Radius dieser Gebiete, wenn ein hochdimensionaler Phasenraum überdeckt werden soll. Man bräuchte deshalb viel mehr von diesen zusätzlichen Neuronen als von den ursprünglichen Neuronen, die die eigentliche Verarbeitung der Sinnesdaten leisten. Ein solches Netzwerk würde eher wie der Kleinhirncortex aussehen, der eine enorme Zahl kleiner granularer Neuronen enthält, und nicht wie der Neocortex mit seinen großen und komplexen Pyramidalneuronen.

Anhang D postuliert eine dritte, raffiniertere Möglichkeit, Verwechslungen zu erkennen. In dieser sind es nicht zusätzliche Neuronen, sondern die einzelnen Dendriten bereits vorhandener Neuronen, die so kleine Aktivitätsgebiete besitzen, daß sie nur auf sehr ähnliche Orbits reagieren. Die offensichtliche Schwierigkeit dieses Modells besteht darin, daß man der Ausgabe eines Neurons nicht ansehen kann, von welchem seiner Dendriten es zur Zeit erregt wird. Man muß deshalb aus der Aktivität der anderen Neuronen folgern können, welcher Dendrit zur Zeit aktiv ist. Anhang D nennt das Aktivitätsmuster der anderen Neuronen den „Kontext". Ideal wäre es, wenn keine zwei Dendriten desselben Neurons in demselben Kontext aktiv wären.

Durch Lernen im Kontext kann sich das Neuron diesem Idealfall nähern. Der Lernprozeß beruht darauf, daß der Kontext den Dendriten über hemmende „Shunt"-Synapsen signalisiert wird. Er erfordert keine besonders komplizierten oder biologisch unplausiblen Lernmechanismen und keine präzisen räumlichen Verschaltungen der Synapsen. Die Verwechslungen von Orbits werden letztlich dadurch vermieden, daß sich die Aktivitätsgebiete beim Lernen ein bißchen ändern. Dabei spielt die zeitliche Verschiebung der Aktivitätsgebiete, die in Kap. 5 benutzt wurde, keine besondere Rolle. Diese vagen Behauptungen werden in Anhang D.3 präzisiert und begründet.

Auch das Vereinigen oder Zerlegen von Symbolischen Gebieten, das in Kap. 7 wichtig war, hat wahrscheinlich kein neurobiologisches Analogon. Denn zwei reale Neuronen lassen sich nicht zu einem Neuron verschmelzen. Außerdem sind zu lange Lernphasen von Nachteil, da ein Neuron erst dann zuverlässig arbeiten kann, wenn die Signale, die es von anderen Neuronen erhält, nicht mehr fortwährend durch deren Lernprozesse verändert werden. Wegen dieser Einschränkungen wird ein biologisches Netzwerk wohl auch nicht die Strukturierung der Dynamik finden, die nach dem mathematischen Kriterium 4.1.2 die einfachste ist. Die Strukturierung der Dynamik kann aber nicht nur durch langandauernde Lernprozesse in einem Netzwerk optimiert werden, sondern auch dadurch, daß die Ausgabe des einen Netzwerkes in ein weiteres Netzwerk geleitet wird, dessen Neuronen dann eine etwas bessere Strukturierung der Dynamik erlernen. Das heißt, sie lernen noch etwas langfristigere Phänomene zu erkennen und noch etwas weniger Orbits zu verwechseln. Eine solche Aneinanderreihung von Netzwerken ist im Neocortex als Hierarchie von Arealen bekannt.

9.2.4 Spezialisierung

Typische Pyramidalneuronen im Neocortex haben etwa 30 Basaldendriten. Diesem Modell zufolge hätten also die Aktivitätsgebiete der Dendriten ein etwa dreißigmal kleineres Volumen im Phasenraum als die Aktivitätsgebiete der Neuronen. Wie ähnlich die verwechselten Orbits sind, hängt aber nicht vom Volumen, sondern vom Durchmesser der Aktivitätsgebiete ab. Da der Phasenraum, in dem die externe Dynamik verläuft, eine extrem hohe Dimension d_p haben kann, ergibt sich ein Problem: Weil das Volumen mit der d_p-ten Potenz des Durchmessers anwächst, kann ein großer Unterschied im Volumen durch einen kleinen Unterschied im Durchmesser erzeugt werden. In anderen Worten: Durch das Lernen im Kontext werden sehr viel weniger Orbits verwechselt, doch die verwechselten Orbits können sich immer noch ziemlich unähnlich sein.

Ein ganz ähnliches Problem tritt auch bei der Repräsentation statischer Muster auf: Wenn man jedes Muster durch die Aktivität eines einzelnen Neurons (oder eines einzelnen Dendriten) repräsentieren will, dann braucht man sehr viele Neuronen (oder Dendriten). In der Theorie der neuronalen Netze spricht man spaßeshalber von den „Großmutterzellen", die so stark spezialisiert sind, daß sie unter allen Mustern nur auf das Bild der Großmutter reagieren. Sehr viel weniger Neuronen werden für die „verteilten Repräsentationen" benötigt. In diesem Fall wird jedes Muster der Sinnesdaten durch ein Muster vieler gleichzeitig aktiver Neuronen repräsentiert. Leider ist ein solches Aktivitätsmuster schwer weiterzuverarbeiten: Neuronen, die nur von einem bestimmten

Muster aktiver Neuronen erregt werden, sind schwerer zu konstruieren als Neuronen, die nur von spezialisierten Einzelneuronen erregt werden. Für letztere braucht es nur ein paar starke Synapsen, für erstere eine komplexe Schaltung.

Wie verteilt die Repräsentation der Dynamik sein darf, hängt letztlich von der Dynamik selbst ab. Wenn langfristige Phänomene der Dynamik erlernt wurden und mehrere dieser Phänomene gleichzeitig erscheinen, dann werden sie auch durch mehrere gleichzeitig aktive Neuronen repräsentiert. Doch auch das Lernen im Kontext kann verteilte Repräsentationen ermöglichen: Als eine Motivation für die Suche nach generierenden Partitionen gab Kap. 2.3.3 an, daß man den „einfachsten" dieser Partitionen vielleicht ansehen kann, ob sich die Dynamik f in relativ unabhängige Teildynamiken zerlegen läßt. Im Neocortex werden solche Teildynamiken in verschiedenen Arealen repräsentiert, wodurch die Repräsentation der Gesamtdynamik verteilt wird. Zum Beispiel signalisieren manche visuellen Areale, was für ein Objekt gesehen wird, und andere visuelle Areale, wo das Objekt gesehen wird. Und manche Areale repräsentieren eher die Farbe eines Objektes und andere Areale seine Gestalt und Bewegung. Die beiden Dynamiken des Farbwechsels und der Bewegung können dann relativ unabhängig voneinander gelernt werden. Zusammen repräsentieren so spezialisierte Areale die Eigenschaften des Objektes genauer, als ein großes, unspezialisiertes Areal es könnte.

Kann eine solche Spezialisierung gelernt werden? Vermutlich ja: Dazu muß ein auf Farbe spezialisiertes Areal einem anderen Areal via dessen Kontext mitteilen, daß die Farbinformationen nicht mehr repräsentiert zu werden brauchen. Die Neuronen können dann in diesem Kontext lernen, die anderen visuellen Informationen, wie Form und Bewegung, zu repräsentieren. Und wie können die Informationen aus den Teildynamiken wieder zusammengesetzt werden? Die Antwort besteht in einer bestimmten Verbindung der Schichten verschiedener Areale. Im Rahmen des Modells von Anhang D.4 ergeben diese Verbindungen ein stimmiges Gesamtbild, dessen Einfachheit mich überraschte. Vorhersagen werden dabei sowohl aus den erlernten Phänomenen derselben Zeitskala als auch aus den längerfristigen Vorhersagen abgeleitet. Und die Unterscheidung verschiedener Zeitskalen beruht, ganz wie die Unterscheidung von Form und Farbe, auf Lernen im Kontext (siehe Anhang D.3). Ohne das neurobiologische Vorbild des Neocortex wäre ich trotz all der mathematischen Konzepte dieser Doktorarbeit wohl nie auf die Idee gekommen, neuronale Netze auf diese natürliche Weise miteinander zu verbinden.

„The time has come", the Walrus said,
„To talk of many things:
Of shoes — and ships — and sealing wax —
Of cabbages – and kings —
And why the sea is boiling hot —
And whether pigs have wings.
L. Carroll: „Through the Looking-Glass"; Kap. 4.

Anhang A

Kleine Ergänzungen zu Kap. 2

A.1 Das Beispiel einer Kreisabbildung

A.1.1 Motivation

Generierende Partitionen von eindimensionalen, stückweise invertierbaren Abbildungen werden auf dem Kreis in einer anderen Weise konstruiert als auf dem Einheitsintervall. Beide Fälle sind gut verstanden. Frühe Referenzen können in [1] gefunden werden, eine neuere Darstellung in [70].

Falls die Symbolischen Gebiete disjunkt gewählt werden, dann erweisen sich auf dem Kreis Partitionsgrenzen als notwendig, deren Lagen nicht eindeutig durch die Dynamik f festgelegt sind. Überlappende Symbolische Gebiete werden in der Literatur nicht diskutiert. Als Vorbereitung auf den höherdimensionalen Fall in Kap. 4 und 5 will ich in diesem Anhang anhand eines Beispiels darstellen, daß die Wahl generierender Partitionen weniger willkürlich ist, wenn überlappende Symbolische Gebiete erlaubt werden. Dabei werden die in Kap. 3 eingeführten Konzepte der expandierenden und kontrahierenden Blätter zwiefacher Partitionen verwendet. Die kontrahierenden Blätter sind bei eindimensionalen Abbildungen dadurch festgelegt, daß jedes kontrahierendes Blatt ein einzelner Punkt des Phasenraumes ist. Die neu eingeführten Konzepte drücken daher nicht mehr aus als die Weise, wie Teilintervalle aufeinander abgebildet werden.

A.1.2 Traditionelle Wahl einer generierenden Partition

In Kap. 2.2 wurden Abbildungen f auf dem Einheitsintervall betrachtet, die stückweise differenzierbar und stückweise invertierbar sind. Die Partitionsgrenzen verliefen durch die „kritischen Punkte", an denen die Ableitung f' ihr Vorzeichen wechselt. Die so definierten Partitionen sind zumindest dann generierend, wenn die Abbildung f an jeder Stelle x expandiert: $\inf_x |f'(x)| > 1$.

Dies gilt nicht mehr auf dem Kreis, der aus dem Einheitsintervall durch Identifizierung der Punkte $x = 0$ und $x = 1$ entsteht. Denn auf ihm könnte zum Beispiel $f(x) = 2 \cdot x$ an allen Stellen x sein, so daß es überhaupt keine kritischen Punkte gäbe. Ein anderes Beispiel zeigt Abb. A.1. Darin wird der Phasenraum durch die kritischen Punkte in vier Symbolische Gebiete $\circ G = (z_1, z_2)$, $\circ H = (z_2, z_4)$, $\circ K = (z_4, z_5)$ und $\circ L = (z_5, z_1 + 1)$ unterteilt. Anders als im Falle des Einheitsintervalles sind die Zweige der Abbildung f auf jedem so definierten Symbolischen Gebiet nicht unbedingt invertierbar. Ein Beispiel gibt Abb. A.1: Das Bild $f(\circ H)$ bedeckt den ganzen Kreis und überlappt an seinen beiden Enden.

Um eine generierende Partition zu erhalten, muß man einen Funktionswert $f_c \in [0, 1)$ wählen und die Symbolischen Gebiete an allen Stellen x mit diesem Funktionswert $f(x) = f_c$ sowie an der Stelle $x = f_c$ selbst unterteilen (siehe z. B. [70]). Üblicherweise wird $f_c = 0$ gewählt, doch auf dem Kreis ist die Stelle 0 natürlich durch nichts ausgezeichnet. Auf diese Weise wird das Symbolische Gebiet $\circ H$ in Abb. A.1 in zwei kleinere Gebiete $\circ H_1 = (z_2, z_3)$ und $\circ H_2 = (z_3, z_4)$ unterteilt sowie das Symbolische Gebiet $\circ L$ in $\circ L_1 = (z_5, 1)$ und $\circ L_2 = (0, z_1)$. Auf jedem der nun sechs Symbolischen Gebiete $\circ B_i \in \{\circ G, \circ H_1, \circ H_2, \circ K, \circ L_1, \circ L_2\}$ besitzt die Abbildung f ein

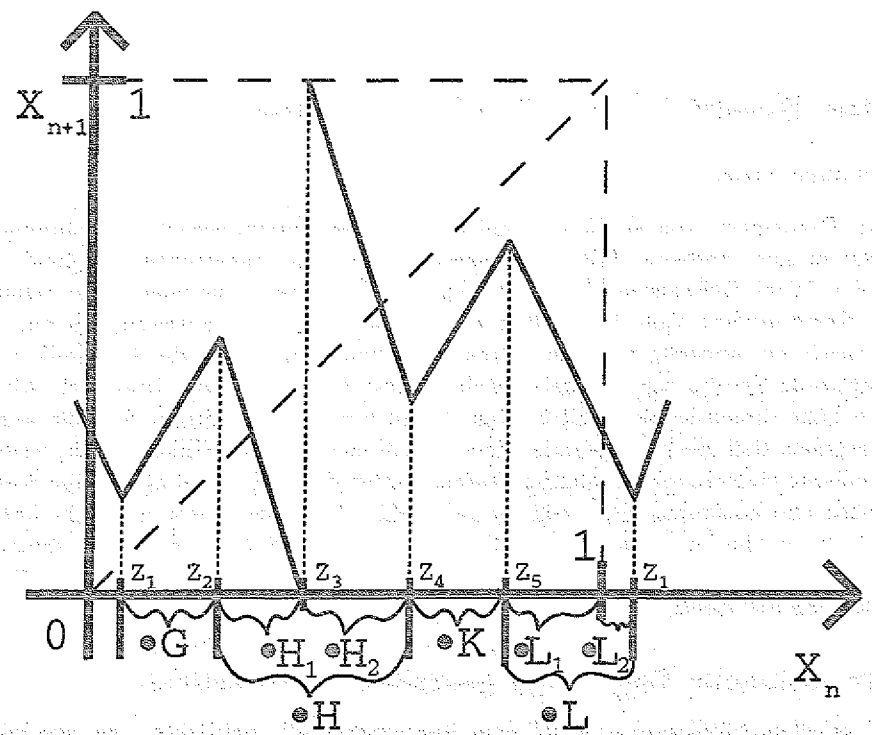


Abbildung A.1: Eine stückweis lineare Abbildung $f(z_n) = z_{n+1}$ des Kreises auf sich selbst. Auf den beiden Intervallen $\circ G$ und $\circ K$ ist die Ableitung $f' > 1$, auf den beiden Intervallen $\circ H$ und $\circ L$ ist die Ableitung $f' < -1$. Eine generierende Partition entsteht, wenn das Intervall $\circ H$ an der Stelle $x = z_3$ (mit $f(z_3) = 0$) in die kleineren Symbolischen Gebiete $\circ H_1$ und $\circ H_2$ unterteilt wird, sowie das Intervall $\circ L$ an der Stelle $x = 0$ in die kleineren Gebiete $\circ L_1$ und $\circ L_2$.

eindeutiges Inverses $f_{B_i}^{-1}$. Der Beweis, daß diese Partition generierend ist, verläuft dann analog zu dem Falle des Einheitsintervalles in Kap. 2.2.3. Die Symbolsequenz $\ldots S_{-1} \circ S_0 S_1 \ldots$ eines Orbits $(x_t)_{t\in\mathbb{Z}}$ legt durch $x_t \in \circ S_t$ die inversen Abbildungen fest, durch die die Punkte des Orbits ineinander überführt werden: $x_t = f_{S_t}^{-1}(x_{t+1})$. Diese inversen Abbildungen sind alle kontrahierend $\sup_x |df_{A_i}^{-1}(x)/dx| < 1$ und wegen der Partitionsgrenzen bei $x = f_c$ und $f(x) = f_c$ bilden sie jedes Teilintervall von $(f_c, f_c + 1)$ auf ein kleineres Teilintervall ab. Daher bestimmt die Folge der inversen Abbildungen

$$x_0 = f_{S_0}^{-1} \circ f_{S_1}^{-1} \circ \ldots f_{S_t}^{-1}(x_{t+1}) \tag{A.1}$$

die Lage des Orbits beliebig genau.

A.1.3 Kontrahierende und expandierende Blätter

Welche kontrahierenden und expandierenden Blätter hat die so gewählte generierende Partition? Die kontrahierenden Blätter sind leicht zu beschreiben: Alle Orbits $(x_t)_{t\in\mathbb{Z}}$, die in der Lage x_0 zur Zeit $t = 0$ übereinstimmen, bilden ein kontrahierendes Blatt. Denn wegen $x_t = f^t(x_0)$ haben diese Orbits zu allen Zeiten $t \geq 0$ gleiche Lage und deshalb auch gleiche Symbole S_t. Und umgekehrt ist die Lage x_0 in Gl. A.1 durch die zukünftige Symbolsequenz $\circ S_0 S_1 \ldots$ festgelegt. Während sich also in höherdimensionalen Phasenräumen die Größe der kontrahierenden Blätter ändern kann, wenn die Partition vereinfacht wird (Kap. 5), ist diese Größe bei eindimensionalen Abbildungen von vorneherein festgelegt.

Nur die Größe der expandierenden Blätter kann sich dann noch ändern. Wie groß sie sind, hängt davon ab, wie die Symbolischen Gebiete $\circ B_i \in \{\circ G, \circ H_1, \circ H_2, \circ K, \circ L_1, \circ L_2\}$ aufeinander abgebildet werden. Wenn jedes dieser Intervalle ganz in dem Bild eines anderen Intervalles enthalten ist und wenn der Attraktor den ganzen Kreis bedeckt, dann ist in jedem expandierenden Bündel $\{(\bigsqcup_i B_i)^{-\infty} B_j\}\circ$ ein expandierendes Blatt, das so breit ist wie das Intervall $B_j \circ = f(\circ B_j)$ selbst. Wie die Intervalle spezieller Abbildungen aufeinander abgebildet werden, ist eine Frage, die bei der Bifurkationsanalyse eindimensionaler Abbildungen detailliert untersucht wurde (siehe z. B. [42]).

Die Definition 2.2.1 für generierende Partitionen wurde in Bed. 3.2.1 dadurch ausgedrückt, daß kein expandierendes Blatt ein kontrahierendes Blatt in mehr als einem Orbit schneiden soll. Bei eindimensionalen Abbildungen läßt sich das noch einfacher sagen. Jedes kontrahierende Blatt vereinigt all die Orbits mit derselben Lage x_0. Deshalb darf kein expandierendes Blatt zwei verschiedene Orbits mit derselben Lage x_0 enthalten.

Diese Bedingung wird durch die Partition $\{\circ G, \circ H_1, \circ H_2, \circ K, \circ L_1, \circ L_2\}$ natürlich erfüllt, da diese Partition generierend ist. Sie folgt aber auch direkt aus der Invertierbarkeit der Abbildung f auf den Symbolischen Gebieten $\circ B_i$. Denn wenn zwei Orbits desselben expandierenden Blattes $\ldots S_{-2} S_{-1} \circ$ zur Zeit $t = 0$ dieselbe Lage x_0 haben, dann haben sie auch für $t < 0$ dieselbe Lage $x_t = f_{S_t}^{-1} \circ f_{S_{t+1}}^{-1} \circ \ldots f_{S_{-1}}^{-1}(x_0)$ und für $t > 0$ dieselbe Lage $x_t = f^t(x_0)$. Sie sind daher identisch.

A.1.4 Überlappende Symbolische Gebiete

Das Kriterium 4.1.2 der Einfachheit begünstigt große expandierende Blätter. Die Abbildung f läßt sich aber auch auf größeren Gebieten als $\circ H_1$ und $\circ H_2$ in Abb. A.1 invertieren, nämlich auf den Gebieten $\circ I \supset \circ H_1$ und $\circ J \supset \circ H_2$ in Abb. A.2. Während die erstere Partition von der willkürlichen Wahl $f_c = 0$ abhängt, sind alle Grenzen der letzteren Partition durch die Dynamik f selbst bestimmt. Denn $\circ I = (z_2, z_7)$ und $\circ J = (z_6, z_4)$ sind jeweils auf der einen Seite durch einen kritischen Punkt z_2 bzw. z_4 begrenzt und auf der anderen Seite durch einen Punkt z_7 bzw. z_6, dessen Bild $f(z_7) = f(z_2)$ bzw. $f(z_6) = f(z_4)$ mit dem Bild des kritischen Punktes übereinstimmt.

Erzeugt eine solche Partition in Symbolische Gebiete, auf denen die Abbildung f invertierbar ist, auch die richtigen kontrahierenden Blätter? Nicht unbedingt. In unserem Beispiel besitzt die Abbildung f bei $x = p_1$ und $x = p_2$ zwei instabile Fixpunkte (Abb. A.3). Da $f(p_2) > f(z_2)$, sind beide Fixpunkte in dem Symbolischen Gebiet $\circ I$ enthalten. Die beiden Orbits $\hat{x}_t := p_1$ und

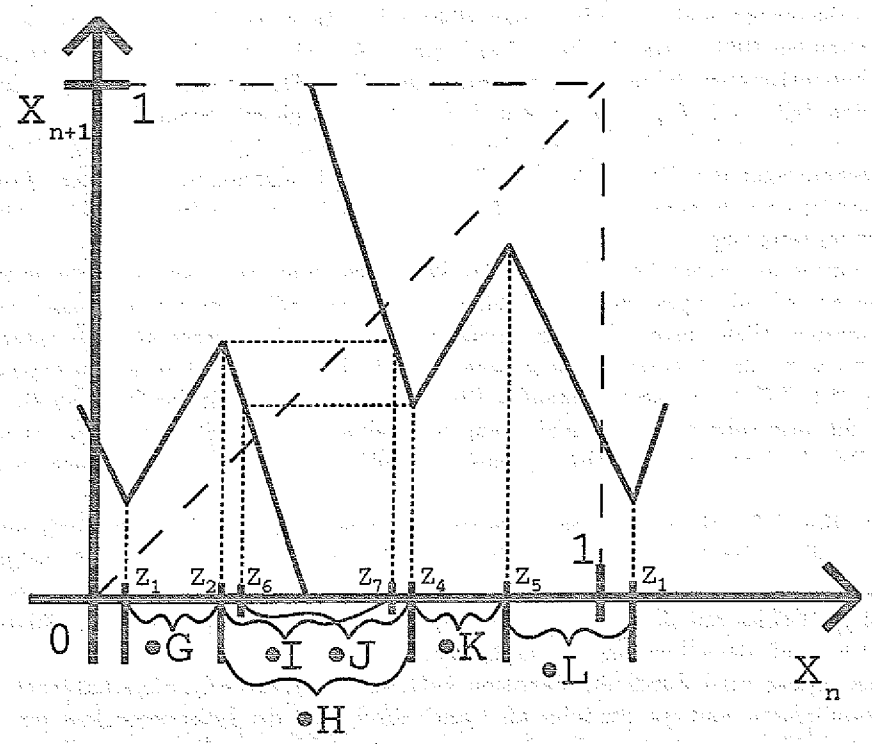


Abbildung A.2: Dieselbe stückweis lineare Abbildung $f(x_n) = x_{n+1}$ wie in Abb. A.1. Das Intervall $\circ H$ wird in die überlappenden Symbolischen Gebiete $\circ I = (z_2, z_7)$ und $\circ J = (z_6, z_4)$ unterteilt. Auf noch größeren Gebieten wäre die Abbildung f nicht invertierbar, da die Grenzen von $\circ I$ bzw. $\circ J$ auf denselben Punkt $f(z_2) = f(z_7)$ bzw. $f(z_6) = f(z_4)$ abgebildet werden.

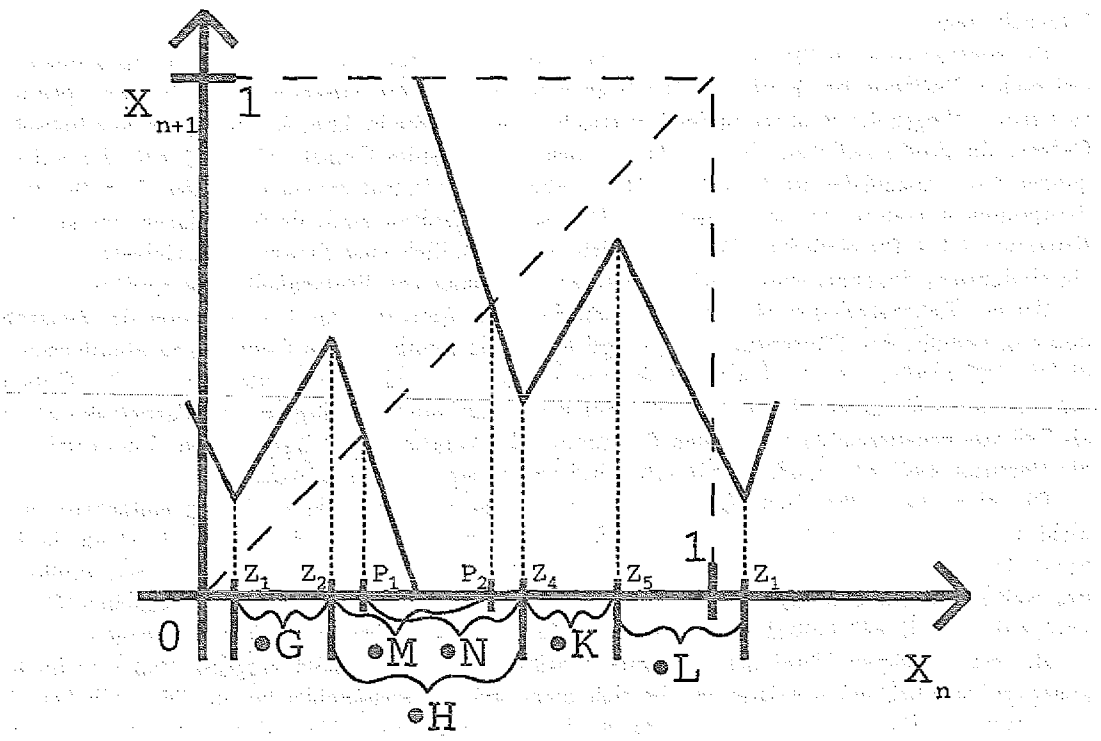


Abbildung A.3: Dieselbe stückweis lineare Abbildung wie in Abb. A.1. Die Größe des Symbolischen Gebietes $\circ M$ ist durch den instabilen Fixpunkt bei p_2 begrenzt. Und der instabile Fixpunkt bei p_1 begrenzt das Symbolische Gebiet $\circ N$. Im Gegensatz zu der Partition in Abb. A.2 ist diese Partition generierend.

$\check{z}_t := p_2$ $(t \in \mathbb{Z})$ sind daher beide in dem Blatt $\circ I^{+\infty}$ enthalten, welches daher kein kontrahierendes Blatt sein kann.

Die richtigen kontrahierenden Blätter lassen sich auf zwei Weisen erzeugen. Erstens kann man eine zwiefache Partition verwenden, das heißt man beschreibt nur die expandierenden Blätter durch die Partition in Abb. A.2 und die kontrahierenden Blätter durch eine andere Partition, zum Beispiel die in Abb. A.1. Auch für andere Kreisabbildungen f lassen sich auf diese Weise die größtmöglichen expandierenen Blätter leicht finden. Die Rekonstruktion einer nicht zwiefachen Partition aus einer solchen zwiefachen Partition ist dann aber nicht trivial (siehe Kap. 6).

Zweitens kann man versuchen eine nicht zwiefache Partition zu konstruieren, die generierend ist und trotzdem möglichst große expandierende Blätter besitzt. In unserem Beispiel darf dann kein Symbolisches Gebiet beide instabilen Fixpunkte enthalten. Die Symbolischen Gebiete $\{\circ G, \circ M, \circ N, \circ K, \circ L\}$ in Abb. A.3 ergeben sich als die größtmöglichen. Tatsächlich kann man zeigen, daß diese Partition generierend ist. Der Beweis beruht auf der Zerlegung des Gebietes $\circ H = \circ M \cup \circ N$ in kleinere Gebiete $\circ H = \circ MHK \cup \circ MHL \ldots$, die sich aufgrund ihrer Symbolsequenzen unterscheiden lassen, und auf der Tatsache, daß die Partition generierend ist, die man durch zusätzlichen Schnitte bei $f_c = f^2(z_2)$ erhält. Eine detaillierte Darstellung des Beweises wäre aber wenig instruktiv.

A.1.5 Zusammenfassung: Die einfachste Partition

Nach welchen Kriterien ergibt sich in unserem Beispiel oder im allgemeinen Fall eine eindeutig einfachste Partition? Das Kriterium der Einfachheit 4.1.2 begünstigt große Symbolische Gebiete. Die Bed. 3.2.1 erfordert, daß die Abbildung f auf allen diesen Gebieten invertierbar ist. Außerdem ist bei eindimensionalen Abbildungen f eine Zusatzbedingung sinnvoll, die in höheren Dimensionen unhandlich wird (Kap. 4.1.1): Die Symbolischen Gebiete sollen einfach zusammenhängende

Intervalle sein.

Die kontrahierenden Blätter sind von vorneherein festgelegt. Wenn wir die Dynamik f mit einer zwiefachen Partition beschreiben, dann brauchen nur noch die expandierenden Blätter optimiert zu werden. Wegen der Bedingung der Invertierbarkeit sind die in Abb. A.2 gezeigten Symbolischen Gebiete die größtmöglichen, doch nicht nur diese: Auch jedes Gebiet $(z^a, z^b) \subset \circ H$, das auf den ganzen Kreis abgebildet wird: $f(z^a) = f(z^b)$ (aber $z^a \neq z^b$), hat maximale Größe. Je mehr dieser Symbolischen Gebiete in der zwiefachen Partition enthalten sind, desto einfacher ist sie nach Kriterium 4.1.2. Die einfachste Partition hätte also unendlich viele Symbolische Gebiete, was zwar ein eindeutiges Ergebnis wäre, doch der naiven Auffasung von Einfachheit widerspräche.

Bei der Rekonstruktion einer nicht zwiefachen Partition in Kap. 6 wird daher die Zusatzbedingung gestellt, den Phasenraum mit möglichst wenig Symbolischen Gebieten zu überdecken. In diesem Fall genügen die beiden Symbolischen Gebiete $\circ I$ und $\circ J$, und nur diese beiden Gebiete, um das Intervall $\circ H$ zu überdecken. Wendet man diese Zusatzbedingung auf Partitionen an, die als Teil von generierenden zwiefachen Partitionen die Vergangenheit beschreiben, dann ergibt sich die Partition $\{\circ G, \circ I, \circ J, \circ K, \circ L\}$ in Abb. A.2 als die einfachste Partition.

Die nicht zwiefache Partition in Abb. A.3 ist schon dann die eindeutig einfachste, wenn nicht zusätzlich eine minimale Zahl von Symbolischen Gebieten gefordert wird. Denn da kein Symbolisches Gebiet einen kritischen Punkt oder zwei Fixpunkte enthalten darf, müßte jedes weitere zusammenhängende Symbolische Gebiet ganz in einem der Symbolischen Gebiete $\{\circ G, \circ M, \circ N, \circ K, \circ L\}$ enthalten sein, um die Bedinung „generierend“ nicht zu verletzen.

Bei einer anderen Wahl der Dynamik f wäre es aber durchaus möglich, daß verschiedene generierende Partitionen existieren, die sich nicht weiter vereinfachen lassen. Die einfachste Beschreibung der Dynamik mit Hilfe von expandierenden und kontrahierenden Blättern ist in diesem Fall leichter zu finden und weniger mehrdeutig als die einfachste Beschreibung mit Hilfe von nicht zwiefachen Partitionen.

In den einfachsten Partitionen von Abb. A.2 und A.3 liegen die Partitionsgrenzen nicht nur bei kritischen Punkten, sondern auch bei Fixpunkten oder bei Punkten, deren Bilder mit Bildern von kritischen Punkten zusammenfallen. Bei anderer Wahl der Dynamik f könnten zum Beispiel auch periodische Orbits Partitionsgrenzen bilden. Jedenfalls sind die Partitionsgrenzen durch die Dynamik f festgelegt, in Einklang mit der Aufgabenstellung von Kap. 1.1.

A.2 Wie stark sind stabile und instabile Mannigfaltigkeit gekrümmt?

In Kap. 2.3.2 wurden die scharfen Krümmungen der stabilen und der instabilen Mannigfaltigkeit diskutiert, die man in typischen, nicht hyperbolischen Systemen findet. Abb. 2.8 illustrierte den dynamischen Ursprung dieser Krümmungen: Krümmungen der instabilen Mannigfaltigkeit entstehen nach Kontraktion in instabiler Richtung. Und Krümmungen der stabilen Mannigfaltigkeit existieren vor Expansion in stabiler Richtung. In demselben Kap. 2.3.2 wurde aber auch argumentiert, daß diese scharfen Krümmungen dicht auf dem Attraktor liegen. Es ist deshalb gar nicht sicher, ob man den numerisch berechneten Abbildungen der stabilen und der instabilen Mannigfaltigkeit trauen soll. Dort sieht man zwischen den scharfen Krümmungen der Mannigfaltigkeiten lange, ungefähr parallele Abschnitte, die kaum gekrümmt sind.

Um den Fall auszuschließen, daß die stabile und instabile Mannigfaltigkeiten wirr gekrümmt sind, soll in diesem Anhang die Häufigkeit scharfer Krümmungen abgeschätzt werden. Diese Abschätzungen sind sehr grob. Es wird unter Vernachlässigung aller multifraktalen Korrekturen angenommen, daß sich die Expansion und Kontraktion in dem Szenario von Abb. 2.8 durch die Lyapunov Exponenten beschreiben läßt. Dabei ergibt sich eine sehr schnelle Schrumpfung des Krümmungsradius der stabilen Mannigfaltigkeit unter der Dynamik f^{-t}, so daß diese scharfen Krümmungen nur in einem exponentiell kleinen Teil des Phasenraums zu erwarten sind.

Wie verändert sich die Krümmung unter der Abbildung f? Wir betrachten zunächst eine eindimensionale stabile Mannigfaltigkeit W^s bzw. instabile Mannigfaltigkeit W^u in einem zweidimensionalen Phasenraum. Wenn die Mannigfaltigkeit durch $\underline{x}(s)$ parametrisiert wird, dann ist ihre Richtung durch $\dot{\underline{x}}(s) := \frac{d}{ds}\underline{x}(s)$, ihre Krümmung c und der Krümmungsradius r_c durch

$$c = 1/r_c := |\dot{x}_1\ddot{x}_2 - \dot{x}_2\ddot{x}_1|/|\dot{\underline{x}}|^3 \tag{A.2}$$

gegeben. Bei der Abbildung nach $\tilde{\underline{x}}(s) = \underline{f}(\underline{x}(s))$ wird der Richtungsvektor linear gestreckt:

$$\dot{\tilde{x}}_i := \frac{d}{ds}\tilde{x}_i(s) = \sum_j \frac{\partial f_i}{\partial x_j}(\underline{x}(s)) \cdot \dot{x}_j \quad . \tag{A.3}$$

Für die Änderung der Krümmung ist die zweite Ableitung wichtig:

$$\ddot{\tilde{x}}_i = \sum_j \left(\frac{\partial^2 f_i}{\partial x_j \partial x_k}\dot{x}_j\dot{x}_k + \frac{\partial f_i}{\partial x_j} \cdot \ddot{x}_j(s) \right) \quad , \tag{A.4}$$

die in einen nichtlinearen und einen linearen Anteil aufspaltet.

Bei scharfer Krümmung c einer Mannigfaltigkeit ist $|\ddot{\underline{x}}|$ groß, verglichen mit $(\dot{\underline{x}})^2$ und der Vektor $\ddot{\underline{x}}$ transformiert sich in Gl. A.4 etwa wie ein Vektor des Tangentialraumes. Über große Zeitintervalle t kann man diese Expansion mit den Lyapunov Exponenten $\lambda_1 > 0 > \lambda_2$ beschreiben.

Uns interessieren auf Abschnitten der instabilen Mannigfaltigkeit nur die Punkte, wo die Krümmung lokal am schärfsten ist. Nach Gleichung A.2 ist die Krümmung dort scharf, wo der Vektor $|\ddot{\underline{x}}|$ groß ist, der Vektor $|\dot{\underline{x}}|$ klein ist und beide Vektoren keinen allzu kleinen Winkel einschließen. Wenn bei einer Zeit 0 die Krümmung auf einem Abschnitt der instabilen Mannigfaltigkeit überall gering ist, dann ist sie auf dem t–fach iterierten Abschnitt dort am schärfsten, wo der Vektor $|\ddot{\underline{x}}|$ exponentiell expandierte ($|\ddot{\underline{x}}| \sim \exp(\lambda_1 t)$) und der Vektor $|\dot{\underline{x}}|$ exponentiell kontrahierte ($|\dot{\underline{x}} \sim \exp(\lambda_2 t)$).

Nach Gleichung A.2 ist diese Krümmung dann von der Größenordnung:

$$c_t \sim \exp((\lambda_1 - 2\lambda_2) \cdot t) \tag{A.5}$$

Wenn umgekehrt ein Abschnitt der stabilen Mannigfaltigkeit erst ab dem Zeitpunkt 0 geringe Krümmung zeigt, dann ist seine maximale Krümmung zum Zeitpunkt $-t$ von der Größenordnung

$$c_{-t} \sim \exp((2\lambda_1 - \lambda_2) \cdot t) \quad . \tag{A.6}$$

Wie man in Abb. 2.8.a erkennt, liegen die so erzeugten scharfen Krümmungen der stabilen Mannigfaltigkeit normalerweise nicht isoliert, sondern entlang von Linien. Es ist zu vermuten, daß die Länge dieser Linien für $t \to -\infty$ so wächst wie die Länge typischer Linien, nämlich wie $\exp(-\lambda_2 t)$. Entlang dieser Linien kann man auf einer Breite von ungefähren $1/c_{-t}$ die scharfe Krümmung der stabilen Mannigfaltigkeit erkennen. Die gesamte Fläche ist dann ungefähr das Produkt $\exp(-\lambda_2 t)/c_{-t} = \exp(-2\lambda_1 t)$, also exponentiell klein. Diese Abschätzung stützt die numerisch berechneten Bilder, die lange ungekrümmte Abschnitte stabiler Mannigfaltigkeit zeigen, und zwar gerade bei starker Vergrößerung des Phasenraumes, wie zum Beispiel in Abb. 2.11. Für die instabile Mannigfaltigkeit und für Mannigfaltigkeiten in höherdimensionalen Phasenräumen führen analoge Abschätzungen zu einem ähnlichen Ergebnis.

Anhang B

Rechenregeln

Dieser Anhang enthält genaue Definitionen und Rechenregeln für den Punkt $\circ$, die Klammern $\{\}'$, die binären Operatoren $\sqcup$ und $\sqcap$ sowie die Klammern $[\,]$, $[,\,[$ und $],]$.

B.1 Definition des Punktes $\circ$

Der schwarze Punkt $\bullet$ ist bereits aus Kap. 2 oder z. B. [70, Kap.9] bekannt: Symbolsequenzen wie $\ldots S_{-1} \bullet S_0 S_1 \ldots$ kennzeichnen Mengen von Orbits (oder auch Punktmengen, wenn die Dynamik f invertierbar ist). Der weiße Punkt $\circ$ wurde in Kap. 3.2.3, Gl. 3.16 für den Fall:

$$\{\nu \circ \mu\}' := \{\alpha \subseteq \bullet E \mid \exists \beta \in \{\nu \bullet \mu\}' : \alpha \subseteq \beta\} \tag{B.1}$$

definiert, daß er zwischen geschweiften Klammern $\{\}'$ steht. Dabei dürfen die Sequenzen ν und μ außer Symbolen auch noch die Operatoren $\cup$, $\cap$, $\sqcup$ und $\sqcap$ enthalten, die im folgenden behandelt werden.

Man kann den Ausdrücken $\nu \circ \mu$ und $\nu \bullet \mu$ auch ohne die Klammern $\{\}'$ einen mathematischen Sinn verleihen. Jeder solcher Ausdruck ist Element einer abstrakten Algebra, das durch die Klammern $\{\}'$ auf eine Menge von Orbitmengen abgebildet wird. Die Algebra wird aus den Symbolsequenzen durch die noch zu definierenden Operatoren $\cup$, $\cap$, $\sqcup$ und $\sqcap$ erzeugt. Sie ist durch bestimmte Assoziativ- und Distributivgesetze charakterisiert. Diese Regeln werden weiter unten aus den Regeln der Booleschen Mengenalgebra abgeleitet.

Am einfachsten läßt sich die Algebra verstehen, die von den Operatoren $\cap$ und $\cup$ aus einstelligen Symbolsequenzen erzeugt wird. Betreachten wir zwei Symbole S und T, deren Symbolische Gebiete $\bullet S$ und $\bullet T$ definiert sind. Der Schnitt $S \cap T$ und die Vereinigung $S \cup T$ waren bisher nicht definiert und können nun als neue Symbole interpretiert werden. Die zu diesen Symbolen gehörenden Symbolischen Gebiete sind natürlich definiert als:

$$\bullet(S \cap T) := (\bullet S) \cap (\bullet T) \qquad \text{und} \qquad \bullet(S \cup T) := (\bullet S) \cup (\bullet T) \tag{B.2}$$

Die Klammern () auf der rechten Seite werden in Zukunft weggelassen. In dieser Schreibweise hat die Aneinanderreihung von Symbolen und Punkten $\bullet$ und $\circ$ eine höhere Priorität als die binären Operatoren $\cap$ oder $\cup$ (oder die unten definierten Operatoren $\sqcup$ und $\sqcap$).

Auf dieselbe natürliche Weise kann man auch den Schnitt $\alpha \cap \beta$ und die Vereinigung $\alpha \cup \beta$ von längeren Symbolsequenzen wieder als Symbolsequenzen interpretieren. Dabei müssen allerdings α und β dieselbe Länge haben, da sonst die Stellung der Symbole unklar wäre. Die Länge der Sequenz $\alpha \cap \beta$ oder $\alpha \cup \beta$ wird so definiert, daß sie mit der Länge von α (und nicht etwa mit der summierten Länge von α und β) übereinstimmt. Die so entstandenen Symbolsequenzen $\alpha \cup \beta$ oder $\alpha \cap \beta$ können mit weiteren Sequenzen vereinigt oder geschnitten werden, wenn diese dieselbe Länge haben. Die Regeln der Mengenalgebra, also die Assoziativität, die Kommutativität und die beiden Distributivgesetze, können auf diese Symbolsequenzen übertragen werden, wodurch eine Boolesche Algebra entsteht.

Der Vorteil dieser Schreibweise besteht darin, daß Symbole, die zwei Sequenzen gemeinsam sind, nun ausgeklammert werden können. Zum Beispiel ergibt sich aus diesen Definitionen sofort für alle Symbole S_1, S_2 und T:

$$\begin{aligned} S_1 \circledcirc T \cup S_2 \circledcirc T &= (S_1 \circledcirc \cap \circ T) \cup (S_2 \circledcirc \cap \circ T) = (S_1 \circledcirc \cup S_2 \circledcirc) \cap \circ T \quad \text{(B.3)} \\ &= (S_1 \cup S_2) \circledcirc \cap \circ T = (S_1 \cup S_2) \circledcirc T \ . \quad \text{(B.4)} \end{aligned}$$

Dieses Ausklammern verkürzt die Rechnungen dieser Arbeit wesentlich, weil sich darin viele lange Sequenzen nur in wenigen Symbolen unterscheiden.

Auf demselben abstrakten Niveau kann auch eine Boolesche Algebra auf den Symbolsequenzen definiert werden, die den weißen Kreis $\circ$ enthalten. Die Operatoren $\cap$ und $\cup$ sind in dieser Algebra definiert durch:

$$\begin{aligned} \mu_1 \circ \nu_1 \ \cap \ \mu_2 \circ \nu_2 &= \mu \circ \nu \Leftrightarrow \mu_1 \circledcirc \nu_1 \ \cap \ \mu_2 \circledcirc \nu_2 = \mu \circledcirc \nu \quad \text{(B.5)} \\ \mu_1 \circ \nu_1 \ \cup \ \mu_2 \circ \nu_2 &= \mu \circ \nu \Leftrightarrow \mu_1 \circledcirc \nu_1 \ \cup \ \mu_2 \circledcirc \nu_2 = \mu \circledcirc \nu \ . \quad \text{(B.6)} \end{aligned}$$

Dabei müssen μ_1, μ_2 oder μ (bzw. ν_1, ν_2 oder ν) nicht unbedingt selbst Symbolsequenzen sein, sondern können auch durch die Operatoren $\cap$ und $\cup$ (oder auch $\sqcup$ und $\sqcap$) aus Symbolsequenzen gleicher Länge erzeugt werden. In dieser Booleschen Algebra gelten per Definition die Assoziativ–, Kommutativ– und auch die Distributivgesetze:

$$\begin{aligned} (\sigma \cap \tau) \ \cup \ \rho &= (\sigma \cup \rho) \ \cap \ (\tau \cup \rho) \quad \text{(B.7)} \\ (\sigma \cup \tau) \ \cap \ \rho &= (\sigma \cap \rho) \ \cup \ (\tau \cap \rho) \ . \quad \text{(B.8)} \end{aligned}$$

Diese Elemente der abstrakten Algebra können nun auf Mengen von Orbitmengen abgebildet werden. Die Abbildung wird nicht, wie üblich, durch einen Buchstaben ausgedrückt, sondern durch das Einklammern des Elementes μ. Das eingeklammerte Element $\{\mu\}$‘ bezeichnet daher eine Menge von Orbitmengen. Um diese Abbildung $\{\}$‘ zu definieren, genügt es (wie bei allen auf Algebren definierten Abbildungen) das Bild jedes erzeugenden Elementes und die Wirkung der Abbildung auf die Operatoren $\cup$ und $\cap$ anzugeben. Die erzeugenden Elemente der Algebra, nämlich die Sequenzen $\ldots S_{-1} \circ S_0 S_1 \ldots$, wurden schon in Gl. B.1 abgebildet. Das Bild aller durch die Operatoren $\cup$ und $\cap$ erzeugten Elemente ist dann definiert durch:

$$\begin{aligned} \{\mu \cup \nu\}‘ &:= \{\rho_1 \cup \rho_2 \mid \rho_1 \in \{\mu\}‘ \wedge \rho_2 \in \{\nu\}‘\} \quad \text{(B.9)} \\ \{\mu \cap \nu\}‘ &:= \{\rho_1 \cap \rho_2 \mid \rho_1 \in \{\mu\}‘ \wedge \rho_2 \in \{\nu\}‘\} \ . \quad \text{(B.10)} \end{aligned}$$

Dies soll gelten für alle Sequenzen μ and ν, die gemäß den Rechenregeln aus Symbolen, einem Punkt $\circ$ oder $\circledcirc$ und den Operatoren $\cup$ und $\cap$ (und später auch $\sqcup$ und $\sqcap$) zusammengesetzt sind. Die Wirkung von $\cup$ bzw. $\cap$ in der Booleschen Algebra von Symbolsequenzen entspricht also der paarweisen Vereinigung bzw. dem paarweisen Schnitt von Orbitmengen ρ_i. Wenn $\circledcirc N = \emptyset$, dann läßt sich das Nullelement der Booleschen Algebra als $\circ N$ schreiben, durch $\{\}$‘ wird es laut B.1 auf das Nullelement $\{\emptyset\}$ abgebildet. Das Einselement $\circ E$ wird auf die Potenzmenge $\{\circ E\}$‘ abgebildet, die alle Orbitmengen enthält.

B.2 Definition der binären Operatoren $\sqcup$ und $\sqcap$

Nun können aber nicht nur die Orbitmengen selbst, sondern auch Mengen von Orbitmengen geschnitten und vereinigt werden. Dies induziert eine weitere Boolesche Algebra. Um in ihr effizient rechnen zu können, bedarf es einer Schreibweise, bei der die Symbole nicht durch Klammern }‘ { von den neuen Operatoren $\sqcup$ und $\sqcap$ getrennt werden.

$$\begin{aligned} \{\mu \sqcup \nu\}‘ &:= \{\mu\}‘ \cup \{\nu\}‘ = \{\rho \mid \rho \in \{\mu\}‘ \vee \rho \in \{\nu\}‘\} \quad \text{(B.11)} \\ \{\mu \sqcap \nu\}‘ &:= \{\mu\}‘ \cap \{\nu\}‘ = \{\rho \mid \rho \in \{\mu\}‘ \wedge \rho \in \{\nu\}‘\} \quad \text{(B.12)} \end{aligned}$$

Diese Definition gilt wieder für alle Sequenzen μ and ν, die gemäß den Rechenregeln aus Symbolen, einem Punkt $\circ$ oder $\circledast$ und den Operatoren $\cup, \cap, \sqcup$ und $\sqcap$ zusammengesetzt sind.

Wie die Operatoren $\cup$ und $\cap$ des letzten Abschnittes erzeugen auch die Operatoren $\sqcup$ und $\sqcap$ eine Boolesche Algebra auf den Symbolsequenzen. Die Distributivgesetze lauten:

$$(\sigma \sqcap \tau) \quad \sqcup \quad \rho \quad = \quad (\sigma \sqcup \rho) \quad \sqcap \quad (\tau \sqcup \rho) \tag{B.13}$$
$$(\sigma \sqcup \tau) \quad \sqcap \quad \rho \quad = \quad (\sigma \sqcap \rho) \quad \sqcup \quad (\tau \sqcap \rho) \tag{B.14}$$

Das Einselement dieser Boolschen Algebra ist wiederum $\circ E$. Das Nullelement ist ein anderes als im letzten Abschnitt, da es durch $\{\}'$ auf $\emptyset$ abgebildet wird.

Bisher wurden die durch $\cap$ und $\cup$ bzw. $\sqcap$ und $\sqcup$ erzeugten Boolschen Algebren getrennt betrachtet. Aus den Definitionen B.9, B.10, B.11 und B.12 der Abbildung $\{\}'$ folgen zwei Beziehungen, in denen die beiden Operatorentypen gemischt sind:

$$\begin{aligned} \{(\sigma \sqcup \tau) \cup \rho\}' &= \{\rho_1 \cup \rho_2 \mid \rho_1 \in \{\sigma \sqcup \tau\}' \wedge \rho_2 \in \{\rho\}'\} && \text{(B.15)} \\ &= \{\rho_1 \cup \rho_2 \mid \rho_1 \in \{\sigma\}' \wedge \rho_2 \in \{\rho\}'\} \cup \{\rho_1 \cup \rho_2 \mid \rho_1 \in \{\tau\}' \wedge \rho_2 \in \{\rho\}'\} \\ &= \{\sigma \cup \rho\}' \quad \cup \quad \{\tau \cup \rho\}' && \text{(B.16)} \end{aligned}$$

$$\begin{aligned} \{(\sigma \sqcup \tau) \cap \rho\}' &= \{\rho_1 \cap \rho_2 \mid \rho_1 \in \{\sigma\}' \wedge \rho_2 \in \{\rho\}'\} \cup \{\rho_1 \cap \rho_2 \mid \rho_1 \in \{\tau\}' \wedge \rho_2 \in \{\rho\}'\} \\ &= \{\sigma \cap \rho\}' \quad \cup \quad \{\tau \cap \rho\}' && \text{(B.17)} \end{aligned}$$

Es ist deshalb sinnvoll und in Einklang mit der Definition der Abbildung $\{\}'$ möglich, beide Boolschen Algebren zu einer zusammenzufassen, in der zwei weitere Distributivgesetze gelten:

$$(\sigma \sqcup \tau) \quad \cup \quad \rho \quad = \quad (\sigma \cup \rho) \quad \sqcup \quad (\tau \cup \rho) \tag{B.18}$$
$$(\sigma \sqcup \tau) \quad \cap \quad \rho \quad = \quad (\sigma \cap \rho) \quad \sqcup \quad (\tau \cap \rho) \tag{B.19}$$

Die Operatoren $\cap$ und $\sqcap$ tauchen in den Rechnungen viel seltener auf als die Operatoren $\cup$ und $\sqcup$. Statt $\cap$ zu verwenden, kann man nämlich meistens die Symbole wie in

$$\sigma \circ \tau = \sigma \circ \quad \cap \quad \circ \tau \tag{B.20}$$

aneinanderreihen. Die Distributivgesetze B.8 und B.19 lauten dann:

$$\sigma \circ (\tau \cup \rho) \quad = \quad \sigma \circ \tau \quad \cup \quad \sigma \circ \rho \tag{B.21}$$
$$\sigma \circ (\tau \sqcup \rho) \quad = \quad \sigma \circ \tau \quad \sqcup \quad \sigma \circ \rho \tag{B.22}$$

Analoge Gleichungen folgen für eine rechts, nicht links stehende Sequenz $\circ\sigma$ oder für einen Punkt $\circ$ bzw. $\circledast$ an anderer Stelle. Die Regel (B.22) macht übrigens auch dann noch Sinn, wenn kein Punkt $\circ$ vorhanden ist. So kann man die Menge aller 01–Sequenzen der Länge n mit $\{(0 \sqcup 1)^n\}'$ bezeichnen.

Der Operator $\sqcap$ kann immer dann durch den Operator $\cap$ ersetzt werden, wenn alle Symbolsequenzen mit dem Punkt $\circ$ und nicht mit dem Punkt $\circledast$ geschrieben werden. Denn das Bild $\{\Psi\}'$ einer solchen Symbolsequenz Ψ ist laut Def. B.1 eine Menge, die mit einer Orbitmenge ρ auch jede ihrer Teilmengen σ enthält:

$$\forall \rho \in \Psi, \quad \forall \sigma \subseteq \rho: \quad \sigma \in \Psi \quad . \tag{B.23}$$

Durch Einsetzen in die Definitionen B.9, B.10, B.11 und B.12 ergibt sich, daß die Bedingung B.23 auch für $\{\mu \cup \nu\}'$, $\{\mu \cap \nu\}'$ und $\{\mu \sqcup \nu\}'$ gilt, falls sie für $\{\mu\}'$ und $\{\nu\}'$ gilt. Der vierte Operator $\sqcap$ wirkt in diesem Fall wie der Operator $\cap$:

$$\begin{aligned} \{\mu \cap \nu\}' &= \{\rho_1 \cap \rho_2 \mid \rho_1 \in \{\mu\}' \wedge \rho_2 \in \{\nu\}'\} && \text{(B.24)} \\ &= \{\rho_1 \cap \rho_2 \mid \rho_1 \cap \rho_2 \in \{\mu\}' \wedge \rho_1 \cap \rho_2 \in \{\nu\}'\} && \text{(B.25)} \\ &= \{\mu\}' \quad \cap \quad \{\nu\}' \quad = \quad \{\mu \sqcap \nu\}' && \text{(B.26)} \end{aligned}$$

Insbesondere gilt

$$\{\nu \circ \mu\}' = \{(\nu \circ E) \cap (E \circ \mu\}' \quad (B.27)$$
$$= \{(\nu \circ E) \sqcap (E \circ \mu\}' \quad (B.28)$$
$$= \{\nu\circ\}' \cap \{\circ\mu\}' \ . \quad (B.29)$$

B.3 Die eckigen Klammern [ ] und [, "]

Bisher wurden einzelne Orbits zu Orbitmengen zusammengefaßt und diese wiederum zu Mengen von Orbitmengen. Beide Schritte können mit geschweiften Klammern {} oder {}' geschrieben werden. In den Rechnungen in Kap. 5, Anhang C und anderswo muß der umgekehrte Weg begangen werden: Von Mengen von Orbitmengen Γ und Λ zu Mengen von Orbits $[\Gamma]$ und $[\Lambda]$ oder zu Mengen von Orbitpaaren $[\Gamma, "]$ oder $[\Gamma, \Lambda]$. Diese drei verschiedenen Schreibweisen mit eckigen Klammern bezeichnen verschiedene Rechenoperationen.

Die einfache Klammer [ ] vereinigt alle enthaltenen Orbitmengen:

$$[\Gamma] := \bigcup_{\alpha\in\Gamma} \alpha \ . \quad (B.30)$$

Insbesondere gilt wegen Def. B.1 und B.11 für alle Gebiete $\circledast\alpha, \circledast\beta \in \circledast E$:

$$[\{\circledast\alpha\}'] = [\{\circ\alpha\}'] = \circledast\alpha \quad (B.31)$$
$$[\{\circledast(\alpha \cup \beta)\}'] = [\{\circledast(\alpha \cup \beta)\}'] \quad (B.32)$$

In diesen einfachen eckigen Klammern können also alle Rechenzeichen $\sqcup$ durch $\cup$ ersetzt werden und alle $\circledast$ durch $\circ$, sofern kein Operator $\sqcap$ in ihnen auftritt.

Ein Paar verschiedener Orbits x und y soll als $[x, y]$ geschrieben werden. Dann ist

$$[\Gamma, "] := \{[x, y] \mid x \neq y \wedge \exists\alpha \in \Gamma : \ x, y \in \alpha\} \quad (B.33)$$

definiert als die Menge aller Orbitpaare $[x, y]$, die in derselben Operator α enthalten sind. Es folgt

$$[\Gamma \cup \Lambda, "] = [\Gamma, "] \cup [\Lambda, "] \quad (B.34)$$

und, wenn man $\{\mu \sqcup \nu\}'$ für $\Gamma \cup \Lambda$ einsetzt,

$$[\{\mu \sqcup \nu\}', "] = [\{\mu\}', "] \cup [\{\nu\}', "] \ . \quad (B.35)$$

B.4 Die eckige Klammer [,]

Der doppelte Strich " in $[\Gamma, "]$ kennzeichnet den Fall, in dem Orbitpaare innerhalb eines Bündels gebildet werden. Andererseits können auch zwei verschiedene Bündel miteinander Orbitpaare bilden. Für alle Mengen Γ, Λ von Orbitmengen ist

$$[\Gamma, \Lambda] := \{[x, y] \mid x \neq y \wedge x \in [\Gamma] \wedge y \in [\Lambda]\} \quad (B.36)$$

definiert als die Menge der Orbitpaare, die aus einem Orbit in $[\Gamma]$ und einem Orbit in $[\Lambda]$ bestehen.

Diese Schreibweise [,] sollte nicht mit [, "] verwechselt werden. Zwar gilt immer $[\Gamma, "] \subseteq [\Gamma, \Gamma]$, und wenn die Sequenz $\alpha \circ \beta$ kein $\sqcup$ enthält, dann kann man sogar zeigen, daß $[\{\alpha \circ \beta, "\}'] = [\{\alpha \circ \beta\}', \{\alpha \circ \beta\}']$. Doch im allgemeinen ist $[\Gamma, "] \subset [\Gamma, \Gamma]$ möglich. Denn $[\Gamma, \Gamma]$ hängt ja nur vom Gebiet $[\Gamma]$ ab und nicht von der genauen Form der Orbitmengen in Γ.

Nach dem im letzten Abschnitt über die Klammern [ ] Gesagten folgt, daß auch in den Klammern [,] jedes Rechenzeichen $\sqcup$ durch $\cup$ ersetzt werden darf und jedes $\circledast$ durch $\circ$, sofern kein Operator $\sqcap$ in ihnen auftritt. Außerdem folgt aus der Def. B.36:

$$[\Gamma] = [\Gamma_1] \cup [\Gamma_2] \ \Rightarrow \ [\Gamma, \Lambda] = [\Gamma_1, \Lambda] \cup [\Gamma_2, \Lambda] \ . \quad (B.37)$$

Um effizient mit eckigen Klammern [,] rechnen zu können, muß man einen Weg haben, geschweifte Klammern und alle Sequenzen gemeinsamer Symbole aus den Klammern [,] herausziehen zu können, so daß zuletzt in den Klammern nur noch die Symbole verbleiben, in denen sich die gepaarten Orbits unterscheiden. Dies erlaubt die Definition:

$$\{\mu \circ [\sigma,\tau]\nu\}' := [\{\mu\circ\}', \text{“}] \cap [\{\circ\sigma\}', \{\circ\tau\}'] \cap [\{\circ E^n\nu\}', \text{“}] \tag{B.38}$$

und analoge Definitionen, bei denen die Lage des Punktes $\circ$ verschoben ist. Die beiden Sequenzen σ und τ in der Klammer $[\sigma,\tau]$ müssen die gleiche Länge n haben und den Punkt $\circ$, falls er überhaupt in der Klammer auftritt, an derselben Stelle enthalten. Alle vier Sequenzen μ, ν, σ und τ dürfen nach den algebraischen Regeln mit den Operatoren $\sqcup$, $\sqcap$, $\cup$ und $\cap$ gebildet werden.

B.5 Orbitpaare einer zwiefachen Partition

In Kap. 3.2.2 wurde eine zwiefache Partition $\{A_1\circledast, \ldots A_m\circledast\}, \{\circledast B_1, \ldots \circledast B_n\}$ betrachtet. Das Bündel all ihrer expandierenden Blätter wurde geschrieben als:

$$\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty}\circledast\}' \tag{B.39}$$

(Gl. 3.11) und das Bündel all ihrer kontrahierenden Blätter als:

$$\{\circledast(\bigsqcup_{i=1}^{n} B_i)^{+\infty}\}' \quad . \tag{B.40}$$

(Gl. 3.12). Diese Schreibweise wurde durch Def. B.11 präzisiert. Wie man leicht sieht, kann die Menge der Orbitpaare dieser Partition nun mit Def. B.38 geschrieben werden als:

$$\bigcup_{t=-\infty}^{+\infty} f^t\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} \circ [E,E](\bigsqcup_{j=1}^{n} B_j)^{+\infty}\}' \quad . \tag{B.41}$$

Den Operator $\bigcup_{t=-\infty}^{+\infty} f^t$, der nichts anderes tut als den Punkt $\circ$ zu verschieben, werde ich, wo immer möglich, vermeiden.

Die Bedingung 3.2.1, daß keine Orbits verwechselt werden, wurde in Kap. 3.2.3 schon so ausgedrückt, daß die Menge $\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} \circledast (\bigsqcup_{j=1}^{n} B_j)^{+\infty}\}'$ keine Orbitmengen mit mehr als einem Orbit enthält. Da in Def. B.36 nur verschiedene Orbits $x \neq y$ gepaart werden, folgt aus Bed. 3.2.1, daß für jedes k:

$$\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} \circ [B_k,B_k](\bigsqcup_{j=1}^{n} B_j)^{+\infty}\}' = \emptyset \quad . \tag{B.42}$$

Bei einer binären Partition ist $\circledast E = \circledast B_1 \cup \circledast B_2$. Die Menge der Orbitpaare in B.41 kann dann umgeschrieben werden. Dabei wird Gl. B.37 verwendet, um $[\circledast B_1 \cup \circledast B_2, \circledast B_1 \cup \circledast B_2]$ zu ersetzen durch die vier Terme $[\circledast B_1, \circledast B_1]$, $[\circledast B_1, \circledast B_2]$, $[\circledast B_2, \circledast B_1]$ und $[\circledast B_2, \circledast B_2]$. Der erste und der letzte Term verschwinden wegen Gl. B.42. Die Menge aller Orbitpaare ist dann, abgesehen von anderen möglichen Lagen des Punktes $\circ$, gleich

$$\{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} \circ [B_1,B_2](\bigsqcup_{j=1}^{n} B_j)^{+\infty}\}' \quad \cup \quad \{(\bigsqcup_{i=1}^{m} A_i)^{-\infty} \circ [B_2,B_1](\bigsqcup_{j=1}^{n} B_j)^{+\infty}\}' \quad . \tag{B.43}$$

In dieser Form werden Orbitpaare in Anhang C berechnet, wobei der zweite Term aus dem ersten einfach durch Vertauschen der Seiten in der Klammer [,] folgt.

B.6 Ausklammern von Symbolen

Bei der Berechnung von Orbitpaaren ist es sehr praktisch, daß einzelne Symbole S aus der Klammer $[,]$ ausgeklammert werden dürfen. Denn eine kurze Rechnung, die auf die Def. B.36 zurückgreift, ergibt für beliebige Sequenzen μ, σ und τ:

$$\{\mu[S\circ\sigma, S\circ\tau]\}' = [\{\mu E\circ\}', ''] \cap [\{S\circ\}', ''] \cap [\{\circ\sigma\}', \{\circ\tau\}'] \quad \text{(B.44)}$$

$$= \{\mu S\circ[\sigma,\tau]\}' \ . \quad \text{(B.45)}$$

Durch Anwendung der Abbildung f erhält man analoge Formeln, in denen nur der Punkt $\circ$ an einer anderen Stelle liegt. Auch rechts darf man gemeinsame Symbole ausklammern:

$$\{[\sigma\circ S, \tau\circ S]\mu\}' = \{[\sigma\circ,\tau\circ]S\mu\}' \ . \quad \text{(B.46)}$$

Demnach dürfen auch längere Symbolsequenzen ausgeklammert werden, sofern sie nicht die Rechenzeichen $\sqcup$ oder $\sqcap$ enthalten.

Eine andere, oft verwendete Rechenregel folgt direkt aus Gl. B.35:

$$\{(\mu\sqcup\nu)\circ[\sigma,\tau]\}' = \{\mu\circ[\sigma,\tau]\}' \cup \{\nu\circ[\sigma,\tau]\}' \quad \text{(B.47)}$$

Diese Rechenregel ist besonders dann nützlich, wenn die Orbitmengen $[\{\nu\circ\}']$ und $[\{\circ\sigma\}']$ nicht überlappen: $[\{\nu\circ\}'] \cap [\{\circ\sigma\}'] = \emptyset$. Dann ist $\{\nu\circ\sigma\}' = \{\emptyset\}$ und aus der Def. B.38 folgt sofort:

$$\{\nu\circ[\sigma,\tau]\}' = \emptyset. \quad \text{(B.48)}$$

Wenn sich die Sequenz ν nicht sowohl durch σ als auch τ fortsetzen läßt, dann kann man sie demnach in Gl. B.47 streichen.

$$\{\nu\circ\sigma\}' = \{\emptyset\} \vee \{\nu\circ\tau\}' = \{\emptyset\} \Rightarrow \{(\mu\sqcup\nu)\circ[\sigma,\tau]\}' = \{\mu\circ[\sigma,\tau]\}' \quad \text{(B.49)}$$

An den Rechenregeln dieses Abschnittes sind die äußeren geschweiften Klammern $\{\}'$ wieder unbeteiligt. In Anhang C werden sie manchmal weggelassen. Dies ist erlaubt, wenn man ganz analog zu dem Vorgehen in Kap. B.1 eine abstrakte Algebra definiert, die die Rechenzeichen $[,]$ enthält und erst durch die Abbildung $\{\}'$ auf Mengen von Orbitpaaren abgebildet wird. Das Entfernen der geschweiften Klammern $\{\}'$ führt aber, anders als in Kap. B.1, zu keinen weiteren Vereinfachungen, so daß es hier nicht weiter diskutiert werden soll.

Anhang C

Orbitpaare von Partitionen des Hénon Attraktors

In diesem Anhang werden jeweils zwei Partitionen von zwei seltsamen Attraktoren miteinander verglichen. Alle Orbitpaare werden berechnet, woraus sich nach dem Kriterium 4.1.2 die Einfachheit der Partitionen ergibt. Die Rechenregeln aus Anhang B werden intensiv verwendet. Die Ergebnisse wurden bereits in Kap. 4.2 zusammengefaßt.

C.1 Vergleich zweier Partitionen bei schwacher Dissipation

Die Dynamik f des Hénon Attraktors bei $a = 1.0$ und bei der schwachen Dissipation $b = 0.54$ wurde in Gl. 4.2 und 4.3 angegeben.

$$f(x_n) = x_{n+1} = 1 - x_n^2 + y_n \tag{C.1}$$
$$f(y_n) = y_{n+1} = 0.54 \cdot x_n \tag{C.2}$$

In Kap. 2.3.3 wurden zwei Partitionen auf diesem Attraktor vorgestellt, nämlich die 01–Partition (Abb. 2.6) und die XY–Partition (Abb. 2.12). Abb. C.1 vergleicht diese beiden Partitionen. Es ist nicht offensichtlich, welche von beiden die einfachere ist.

C.1.1 Übersetzungsregeln

Die Symbole X und Y wurden in Gl. 2.10 und Gl. 2.11 durch 01–Symbolsequenzen ausgedrückt. Diese Gleichungen lassen sich wegen Gl. C.5 umschreiben zu:

$$\circledast X := \circledast 0 \cup 01 \circledast 11 \cup 0 \circledast 111 \tag{C.3}$$
$$\circledast Y := \circledast 10 \cup 11 \circledast 11 \ . \tag{C.4}$$

Die beiden Teile der XY–Partition sind disjunkt $\circledast X \cap \circledast Y = \emptyset$. Der Abb. 4.1 entnahmen wir, daß die Sequenzen 0110 und 0010 verboten sind.

$$\circledast 0E10 = \emptyset \tag{C.5}$$

Deshalb überdeckt die Vereinigung $\circledast X \cup \circledast Y = \circledast E \setminus 0 \circledast 110$ den Attraktor. Hieraus und aus der weiteren grammatikalischen Regel 4.19 werden in diesem Anhang alle Orbitpaare beider Partitionen berechnet. Dabei werden die Schreibweisen und Rechenregeln aus Anhang B verwendet. Das Ergebnis wurde bereits in Abschnitt 4.2.1 dargestellt.

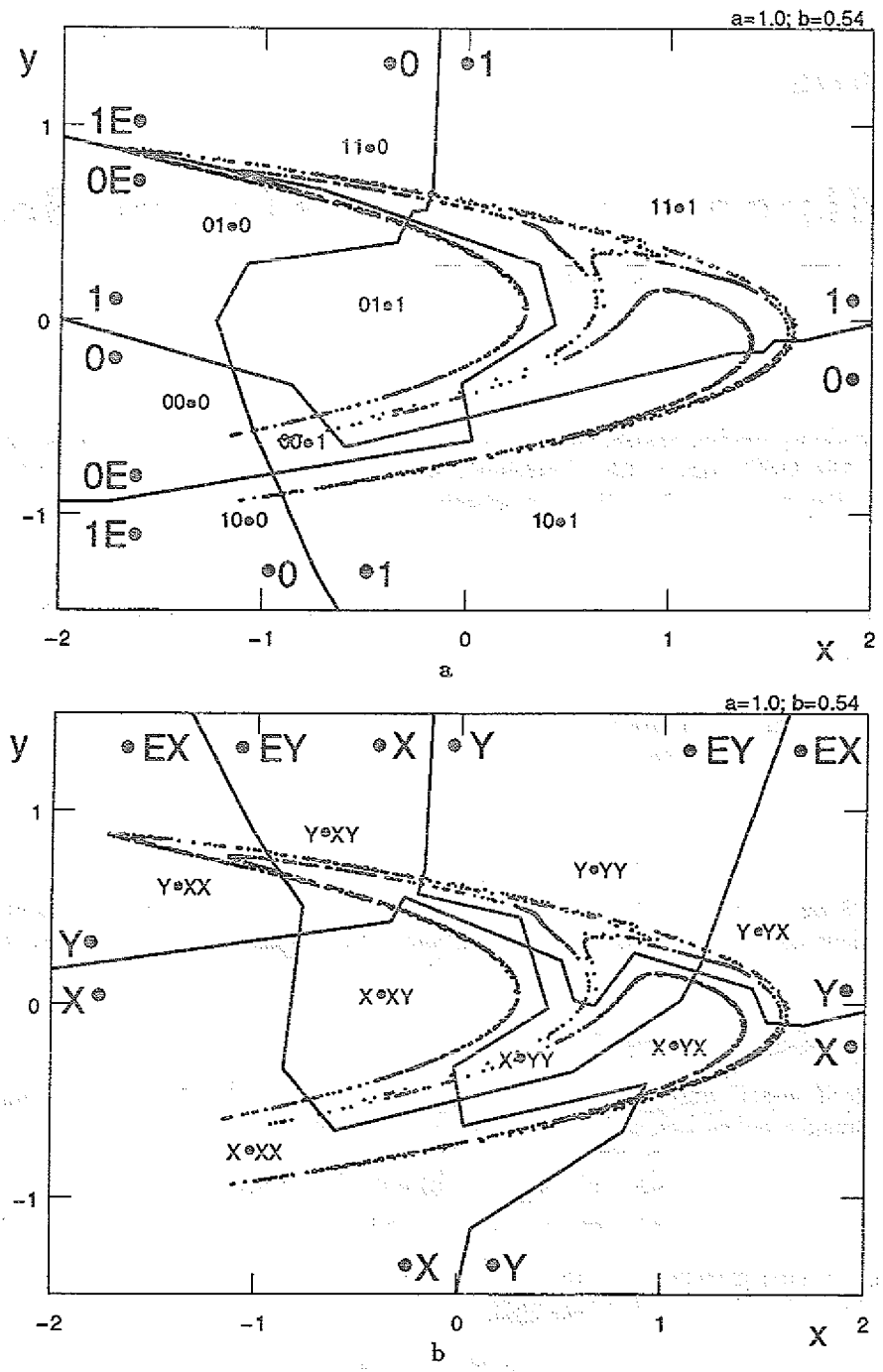


Abbildung C.1: Wiederholung der Abb. 4.1, in der die 01–Partition und die XY–Partition des Hénon Attraktors bei schwacher Dissipation verglichen wurden.

Zunächst sollen die Symbole 0 und 1 durch die Symbole X und Y ausgedrückt werden. Dazu bestimmen wir ein paar kurze Symbolsequenzen. Einsetzen von C.3 und C.4 ergibt nach kurzen Rechnungen:

$$\circledast YX = \circledast 10 \tag{C.6}$$
$$\circledast YY = 11 \circledast 11 \tag{C.7}$$
$$\circledast XX = \circledast 00 \cup 0 \circledast 111 \cup \circledast 0111 \ . \tag{C.8}$$

Zweimaliges Einsetzen der letzten Gleichung ergibt wegen des Verbotes (C.5):

$$\circledast XXXX = \circledast 00 \ . \tag{C.9}$$

Zusammen mit der Definition (C.4) von Y ergibt (C.8):

$$\circledast XXY = 0 \circledast 111 \ , \tag{C.10}$$

was mit dem Verbot (C.5) und mit $\circledast 0 \subseteq \circledast X$ umgeschrieben werden kann:

$$\circledast XXY = X \circledast XXY = 0 \circledast 11 = 0 \circledast 111 \ . \tag{C.11}$$

Deshalb ist die Sequenz $\circledast YXXY = \emptyset$ verboten. Aus (C.6) und (C.7) folgt gleichermaßen das Verbot von $YXYY$.

$$\circledast YXEY = \emptyset \tag{C.12}$$

Mit den Regeln C.6 und C.9 können wir die Symbole $\circledast 0 = 1 \circledast 0 \cup 0 \circledast 0$ und $\circledast 1 = \circledast E \setminus \circledast 0$ durch X und Y ausdrücken:

$$\circledast 0 = Y \circledast X \cup X \circledast XXX \tag{C.13}$$
$$\circledast 1 = \circledast Y \cup X \circledast XXY \cup XX \circledast XY \ . \tag{C.14}$$

Da wir alle unendlichen 01-Symbolsequenzen in unendliche XY-Sequenzen und umgekehrt übersetzen können, sind beide Partitionen generierend, wenn eine von beiden generierend ist, das heißt, wenn in

$$\{(0 \sqcup 1)^{-\infty} \circ (0 \sqcup 1)^{+\infty}\}' = \{(X \sqcup Y)^{-\infty} \circ (X \sqcup Y)^{+\infty}\}' \tag{C.15}$$

keine Menge mit mehr als einem Orbit enthalten ist.

Die bisherigen Gleichungen haben eine Symmetrie, die man der Abb. C.1 des seltsamen Attraktors nicht ansieht: Wenn man die Zeit umkehrt, also rechts und links in allen Symbolsequenzen vertauscht, und außerdem die Symbole gemäß $X \leftrightarrow 1$ und $Y \leftrightarrow 0$ paarweise vertauscht, dann bleiben alle bisherigen Gleichungen wahr. Deshalb kann auch kein Kriterium für Einfachheit, durch das keine Zeitrichtung ausgezeichnet wird, entscheiden, ob die XY– oder die 01–Partition einfacher ist. Unser Kriterium zeichnet keine Zeitrichtung aus. Die weitere Rechnung ist nur deshalb interessant, weil weitere Symbolsequenzen verboten sind, die diese Symmetrie nicht erfüllen. Aus der Abb. 2.6 bzw. 2.12 ersieht man:

$$\circledast XXXXXX = \emptyset = \circledast 0000 \tag{C.16}$$

$$\circledast XXXXY = \circledast XXXXYYY \quad \Leftrightarrow \quad \circledast 001 = \circledast 0011111 \ . \tag{C.17}$$

In die folgende Rechnung gehen nur die zwei Beziehungen in Gl. 4.19 ein:

$$\circledast XXXXXXX = \emptyset \tag{C.18}$$
$$\circledast XXXXYX = \emptyset \ , \tag{C.19}$$

Die Definitionen (C.3) und (C.4) der Symbole X und Y lassen sich mit dem Verbot (C.5) umschreiben.

$$\circledast X = \circledast 0 \cup 01 \circledast 1 \cup 0 \circledast 11 \tag{C.20}$$
$$\circledast Y = \circledast 1 \setminus (01 \circledast 1 \cup 0 \circledast 11) \tag{C.21}$$

Die dabei auftretenden Sequenzen reichen nur um zwei Zeitschritte auf beiden Seiten über den Punkt $\circ$ hinaus. Deshalb liegt jedes Gebiet $S_{-2}S_{-1} \circ S_0S_1$; $S_i \in \{0,1\}$ ganz in einem Gebiet $\circ T_0$; $(T_0 \in \{X,Y\})$.

$$\{(0 \sqcup 1)^2 \circ (0 \sqcup 1)^2\}' \quad \subseteq \quad \{\circ(X \sqcup Y)\}' \tag{C.22}$$

Indem wir diese Beziehung mit f iterieren und den Schnitt

$$\{\circ(X \sqcup Y)^n\}' \quad \cap \quad \{\circ E^n(X \sqcup Y)\}' \quad = \quad \{\circ(X \sqcup Y)^{n+1}\}' \tag{C.23}$$

bilden, können wir diese Beziehung auf halbunendliche Sequenzen verallgemeinern:

$$\{\circ(0 \sqcup 1)^{+\infty}\}' \quad \subseteq \quad \{\circ EE(X \sqcup Y)^{+\infty}\}' \tag{C.24}$$

$$\{(0 \sqcup 1)^{-\infty} \circ\}' \quad \subseteq \quad \{(X \sqcup Y)^{-\infty} E \circ\}' \quad . \tag{C.25}$$

Da in diese Ableitung nur Regeln eingingen, die unter der erwähnten Zeitumkehr symmetrisch sind, folgen ganz analog aus den zu (C.20) und (C.21) symmetrischen Ausdrücken für die Symbole 0 und 1:

$$\circledast 0 \quad = \quad \circledast X \setminus (\circledast XXY \quad \cup \quad X \circledast XY) \tag{C.26}$$

$$\circledast 1 \quad = \quad \circledast Y \quad \cup \quad \circledast XXY \quad \cup \quad X \circledast XY \tag{C.27}$$

die zu (C.24) und (C.25) symmetrischen Beziehungen:

$$\{\circ(X \sqcup Y)^{+\infty}\}' \quad \subseteq \quad \{\circ E(0 \sqcup 1)^{+\infty}\}' \tag{C.28}$$

$$\{(X \sqcup Y)^{-\infty} \circ\}' \quad \subseteq \quad \{(0 \sqcup 1)^{-\infty} EE \circ\}' \quad . \tag{C.29}$$

C.1.2 Orbitpaare der XY–Partition

Wir berechnen nun zunächst die Menge aller Orbitpaare in der XY–Partition. Laut Gl. B.43 genügt es, die Menge

$$\Phi_{XY} \quad = \quad \{(X \sqcup Y)^{-\infty} \circ [X,Y](X \sqcup Y)^{+\infty}\}' \tag{C.30}$$

$$= \quad [\{(X \sqcup Y)^{-\infty} \circ\}', \text{“}] \quad \cap \quad [\circ E\{(X \sqcup Y)^{+\infty}\}', \text{“}] \quad \cap$$
$$\cap \quad \{(X \sqcup Y)^4 \circ [X,Y](X \sqcup Y)^4\}' \tag{C.31}$$

zu berechnen, da alle anderen Orbitpaare durch Verschieben des Punktes $\circ$ oder Vertauschen von X und Y folgen. Dabei hätten wir auch größere Exponenten als die 4 wählen können, aber dies ist für die weitere Rechnung nicht erforderlich.

Wir wollen Φ_{XY} mit 01-Sequenzen ausdrücken. Aus den Beziehungen (C.28) und (C.29) folgt sofort:

$$\Phi_{XY} \quad \subseteq \quad [\{(0 \sqcup 1)^{-\infty} EE \circ\}', \text{“}] \quad \cap \quad [\{\circ EE(0 \sqcup 1)^{+\infty}\}', \text{“}] \tag{C.32}$$

Es bleibt also noch der letzte Term in (C.31) als 01–Sequenz umzuschreiben:

$$\Phi_{XY} \quad \subseteq \quad \bigcup_{T_i} \{T_1T_2T_3T_4 \circ [X,Y]T_5T_6T_7T_8\}' \tag{C.33}$$

(mit $T_i \in \{X,Y\}$). Dazu bedarf es einer Fallunterscheidung:

1. Fall: Kein XXY in $T_3T_4XT_5T_6$ oder $T_3T_4YT_5T_6$
   Nach den Übersetzungsregeln (C.26) und (C.27) wird in diesem Fall das zentrale X in eine 0 und das Y in eine 1 übersetzt. An allen anderen Plätzen hängen die Übersetzungsregeln nicht von diesem X oder Y ab. Für jede Wahl der $T_1 \ldots T_8$ gibt es daher $S_3 \ldots S_6 \in \{0,1\}$, so daß:

$$\{T_1T_2T_3T_4 \circ [X,Y]T_5T_6T_7T_8\}' \quad \subseteq \quad \{S_3S_4 \circ [0,1]S_5S_6\}' \tag{C.34}$$

   Die mit zentralem X bzw. Y übersetzten Sequenzen unterscheiden sich daher auch nur an einer Stelle. Die nun folgenden Fälle werden XXY in $T_3T_4XT_5T_6$ oder $T_3T_4YT_5T_6$ enthalten, und zwar jeweils an einer anderen Stelle.

2. Fall: $T_4 \circ [X,Y]T_5 \quad = \quad X \circ [X,Y]Y$

$$\begin{aligned} \{(X \sqcup Y)^2 X \circ [X,Y]Y\}' &= \{(X \sqcup Y)XX \circ [X,Y]Y\}' \quad \text{da } YX \circledast XY = \emptyset & \text{(C.35)} \\ &= \{XXX \circ [X,Y]Y\}' \qquad \text{da } YXX \circledast Y = \emptyset & \text{(C.36)} \\ &= \{0[0,1]1 \circ 11\}' & \text{(C.37)} \end{aligned}$$

Die 01–Partition unterscheidet die beiden Orbits also um zwei Zeitschritte früher als die XY–Partition.

3. Fall: $\circ[X,Y]T_5T_6 \quad = \quad \circ[X,Y]XY$

$$\begin{aligned} \{(X \sqcup Y)^2 \circ [X,Y]XY(X \sqcup Y)\}' &= \{(X \sqcup Y)^2 \circ [X,Y]XYX\}' \text{ da } \circledast YXYY = \emptyset \\ &= \{(X \sqcup Y)X \circ [X,Y]XYX\}' \qquad \text{da } Y \circledast XXY = \emptyset & \text{(C.38)} \\ &= \{YX \circ [X,Y]XYX\}' \qquad \text{da } XX \circledast XXYX = \emptyset & \text{(C.39)} \\ &= \{10 \circ 1[1,0]10\}' & \text{(C.40)} \end{aligned}$$

Die 01–Partition unterscheidet die beiden Orbits also um einen Zeitschritt später als die XY–Partition.

4. Fall: $T_3T_4 \circ [X,Y] \quad = \quad XX \circ [X,Y]$

Der einzige noch nicht behandelte Fall ist der mit $T_5T_6 = XX$.

$$\begin{aligned} \{(X \sqcup Y)XX \circ [X,Y]XX(X \sqcup Y)\}' &= \\ &= \{XXX \circ [X,Y]XXX\}' \text{ da } YXX \circledast Y = \circledast YXXY = \emptyset & \text{(C.41)} \\ &= \emptyset \qquad \text{da } XXX \circledast XXXX = \emptyset & \text{(C.42)} \end{aligned}$$

Alle Fälle zusammen ergeben:

$$\begin{aligned} \{(X \sqcup Y)^4 \circ [X,Y](X \sqcup Y)^4\}' \subseteq \; & \{[0,1](0 \sqcup 1) \circ (0 \sqcup 1)^2\}' \quad \cap \\ & \{(0 \sqcup 1)^2 \circ (0 \sqcup 1)[1,0]\}' \quad \cap \\ & \{(0 \sqcup 1)^2 \circ [0,1](0 \sqcup 1)\}' & \text{(C.43)} \end{aligned}$$

Wenn wir diese Gleichung mit (C.31) und (C.32) verknüpfen, dann sehen wir, daß alle Orbitpaare von der XY–Partition auch Orbitpaare von der 01–Partition sind.

$$\begin{aligned} \Phi_{XY} \subseteq \; & \{(0 \sqcup 1)^{-\infty}[0,1](0 \sqcup 1) \circ (0 \sqcup 1)^{+\infty}\}' \quad \cap & \text{(C.44)} \\ & \cap \; \{(0 \sqcup 1)^{-\infty} \circ (0 \sqcup 1)[1,0](0 \sqcup 1)^{+\infty}\}' \quad \cap \quad \{(0 \sqcup 1)^{-\infty} \circ [0,1](0 \sqcup 1)^{+\infty}\}' \end{aligned}$$

Die 01–Partition ist nach unserem Kriterium 4.1.2 also mindestens so einfach wie die XY–Partition.

C.1.3 Orbitpaare der 01–Partition

Die Menge aller Orbitpaare in der 01–Partition können wir durch ganz analoge Rechenschritte mit den Symbolen X und Y ausdrücken. Die Analogie endet erst dann, wenn wir in der Fallunterscheidung die Beziehung $XX \circledast XXYX = \emptyset$ oder $XXX \circledast XXXX = \emptyset$ verwenden. Denn die durch Zeitinversion und Symbolvertauschung $X \leftrightarrow 1$ und $Y \leftrightarrow 0$ erzeugten Sequenzen $1011 \circledast 11$ und $1111 \circledast 111$ sind im allgemeinen $\neq \emptyset$. Da wir diese beiden Beziehungen nicht verwenden können, müssen wir die Rechnung an den entsprechenden Stellen der Fallunterscheidung auf andere Weise fortsetzen.

1. und 2. Fall: Keine Änderung der Rechnung.

2. Fall: Wegen $1011 \circledast 11 \neq \emptyset$ ändert sich die Rechnung.

$$\begin{aligned}\{(0 \sqcup 1)^{-\infty} 101[0,1] \circ 11(0 \sqcup 1)^{+\infty}\}' &= \\ &= \{(0 \sqcup 1)^{-\infty} 101[0,1] \circ 111(0 \sqcup 1)^{+\infty}\}' \text{ da } 0 \circledast 110 = \emptyset \quad \text{(C.45)} \\ &= \{(X \sqcup Y)^{-\infty} YX[YX \circ XX, XX \circ YY]Y(X \sqcup Y)^{+\infty}\}' \text{(C.46)}\end{aligned}$$

Diese Orbitpaare der 01–Partition haben zu drei verschiedenen Zeiten verschiedene Symbole X oder Y.

3. Fall: Wegen $1111 \circledast 111 \neq \emptyset$ ändert sich die Rechnung. Wir unterteilen die Orbitpaare mit zentralem $111[0,1] \circ 111$ in zwei Mengen:

$$\begin{aligned}\{(0 \sqcup 1)^{-\infty} 0111[0,1] \circ 111(0 \sqcup 1)^{+\infty}\}' &= \\ &= \{(X \sqcup Y)^{-\infty} XXXY[X \circ XX, Y \circ YY]Y(X \sqcup Y)^{+\infty}\} \text{(C.47)} \\ \{(0 \sqcup 1)^{-\infty} 1111[0,1] \circ 111(0 \sqcup 1)^{+\infty}\}' &= \\ &= \{(X \sqcup Y)^{-\infty} YY[X \circ XX, Y \circ YY]Y(X \sqcup Y)^{+\infty}\}' \quad \text{(C.48)}\end{aligned}$$

Diese Orbitpaare der 01–Partition haben zu drei aufeinanderfolgenden Zeiten verschiedene Symbole X oder Y.

Somit haben wir drei Mengen von Orbitpaaren der 01–Partition, die nicht Orbitpaare der XY–Partition sind. Es gibt keinen Grund anzunehmen, daß diese drei Mengen alle leer seien. Deshalb ist die 01–Partition einfacher als die XY–Partition.

Wenn man aber die Parameter a, b der Abbildung f nur ein wenig verändert, dann könnte es sein, daß in der auch nur wenig veränderten XY–Partition die Symbolsequenz $\circledast XXXXXXX \neq \emptyset$ erlaubt ist. Dann gibt es Orbitpaare der XY–Partition, die nicht Orbitpaare der 01–Partition sind. Sie ergeben sich aus Gl. C.47 und C.48 durch Zeitinversion und Symbolvertauschung $X \leftrightarrow 1$ und $Y \leftrightarrow 0$.

$$\begin{aligned}\{(X \sqcup Y)^{-\infty} XXX \circ [X,Y]XXXX(X \sqcup Y)^{+\infty}\}' &= \\ &= \{(0 \sqcup 1)^{-\infty} 0[00 \circ 0, 11 \circ 1]00(0 \sqcup 1)^{+\infty}\}' \quad \text{(C.49)} \\ \{(X \sqcup Y)^{-\infty} XXX \circ [X,Y]XXXY(X \sqcup Y)^{+\infty}\}' &= \\ &= \{(0 \sqcup 1)^{-\infty} 0[00 \circ 0, 11 \circ 1]0111(0 \sqcup 1)^{+\infty}\}' \quad \text{(C.50)}\end{aligned}$$

In diesem Fall könnte weder die XY–Partition noch die 01–Partition als die einfachere bezeichnet werden.

Auch in einem anderen Fall gibt es zwei gleich einfache Partitionen: Für den Hénon Attraktor bei Standardparameterwerten wird in Kap. 7.1.2 eine XY–Partition definiert, die ganz ähnliche Übersetzungsregeln und grammatikalische Regeln hat wie die hier betrachtete XY–Partition. Ein Unterschied besteht darin, daß bei Standardparameterwerten die Sequenz $\circledast XXXXYX \neq \emptyset$ nicht verboten ist. Deshalb gibt es bei Standardparameterwerten Orbitpaare der XY–Partition, die nicht Orbitpaare der 01–Partition sind. Sie ergeben sich aus der Gl. C.45 durch Zeitinversion und Symbolvertauschung $X \leftrightarrow 1$ und $Y \leftrightarrow 0$.

$$\begin{aligned}\{(X \sqcup Y)^{-\infty} XXX \circ [X,Y]XYX(X \sqcup Y)^{+\infty}\}' &= \\ &= \{(0 \sqcup 1)^{-\infty} 0[00 \circ 11, 11 \circ 10]10(0 \sqcup 1)^{+\infty}\}' \quad \text{(C.51)}\end{aligned}$$

Anders als bei schwacher Dissipation ist deshalb bei Standardparameterwerten die 01–Partition nicht einfacher als die XY–Partition.

C.2 Vergleich zweier Partitionen bei Standardparameterwerten

Die Dynamik f des Hénon Attraktors bei Standardparameterwerten $a = 1.4$ und $b = 0.3$ wurde in Gl. 4.30 angegeben.

$$f(x_n) = x_{n+1} = 1 - 1.4 \cdot x_n^2 + y_n \tag{C.52}$$
$$f(y_n) = y_{n+1} = 0.3 \cdot x_n \tag{C.53}$$

In Kap. 4.2.2 wurden zwei Partitionen auf diesem Attraktor verglichen, nämlich die 01–Partition (Abb. 2.4) und die AB–Partition (Abb. 4.2). Abb. C.2 vergleicht beide Partitionen. Die altbekannte 01–Partition [30] erscheint einfacher, weil die AB–Partition manche Orbits von Tangentialpunkten nicht nur einmal, sondern zweimal schneidet. Dies stimmt mit dem neuen Kriterium der Einfachheit 4.1.2 überein, wie ich hier durch Berechnung aller Orbitpaare zeigen werde.

C.2.1 Übersetzungsregeln

Die Symbole A und B sind durch 01–Symbolsequenzen definiert (Gl. 4.34 und 4.35):

$$\circledast A := \circledast 0 \cup 0 \circledast 11 \tag{C.54}$$
$$\circledast B := 1 \circledast 1 \cup 0 \circledast 10 \ . \tag{C.55}$$

Die beiden Teile der AB–Partition sind disjunkt $\circledast A \cap \circledast B = \emptyset$, und ihre Vereinigung

$$\circledast A \cup \circledast B = \circledast 0 \cup 0 \circledast 10 \cup 0 \circledast 11 \cup 1 \circledast 1 = \circledast E \tag{C.56}$$

überdeckt den Attraktor.

Wären alle 01–Sequenzen erlaubt, so wäre die so definierte AB–Partition nicht generierend. In der Dynamik C.52 sind jedoch einige Sequenzen verboten, zum Beispiel die Sequenzen 0010 und 0110. Mit etwas Übung kann man tatsächlich schon der Abbildung 2.4 ansehen, daß:

$$\circledast 0E10 = \emptyset \ . \tag{C.57}$$

Damit endet die Aufgabenstellung. Welche der beiden Partitionen einfacher ist, folgt nun allein aus Manipulationen der Symbolsequenzen.

Weil die Sequenz 0010 verboten ist, folgt aus den Definitionen C.54 und C.55 von A und B:

$$A \circledast AB = (0 \circledast \cup 01 \circledast 1) \cap (\circledast 0 \cup 0 \circledast 11) \cap (\circledast 11 \cup \circledast 010) \tag{C.58}$$
$$= 0 \circledast 010 \cup 0 \circledast 11 = 0 \circledast 11 \ . \tag{C.59}$$

Damit können wir die Definition C.54 von A umdrehen:

$$\circledast 0 = \circledast A \setminus A \circledast AB \tag{C.60}$$
$$= B \circledast A \cup A \circledast AA \tag{C.61}$$
$$\circledast 1 = \circledast E \setminus \circledast 0 \tag{C.62}$$
$$= \circledast B \cup A \circledast AB \tag{C.63}$$

und erhalten die Symbole 0 und 1 ausgedrückt durch A und B. Wenn 01–Sequenzen verboten sind, so sind natürlich auch AB–Sequenzen verboten. Wir erhalten sie erstens durch Übersetzen von Gl. C.57 nach den Regeln C.61 und C.63:

$$\circledast 0110 = \emptyset = \circledast AABA \tag{C.64}$$

und zweitens durch Rückübersetzen der Definition C.55 von B:

$$\circledast B = 1 \circledast 1 \cup 0 \circledast 10 = B \circledast B \cup AA \circledast B \cup BA \circledast BA = \circledast B \setminus BA \circledast BB \tag{C.65}$$

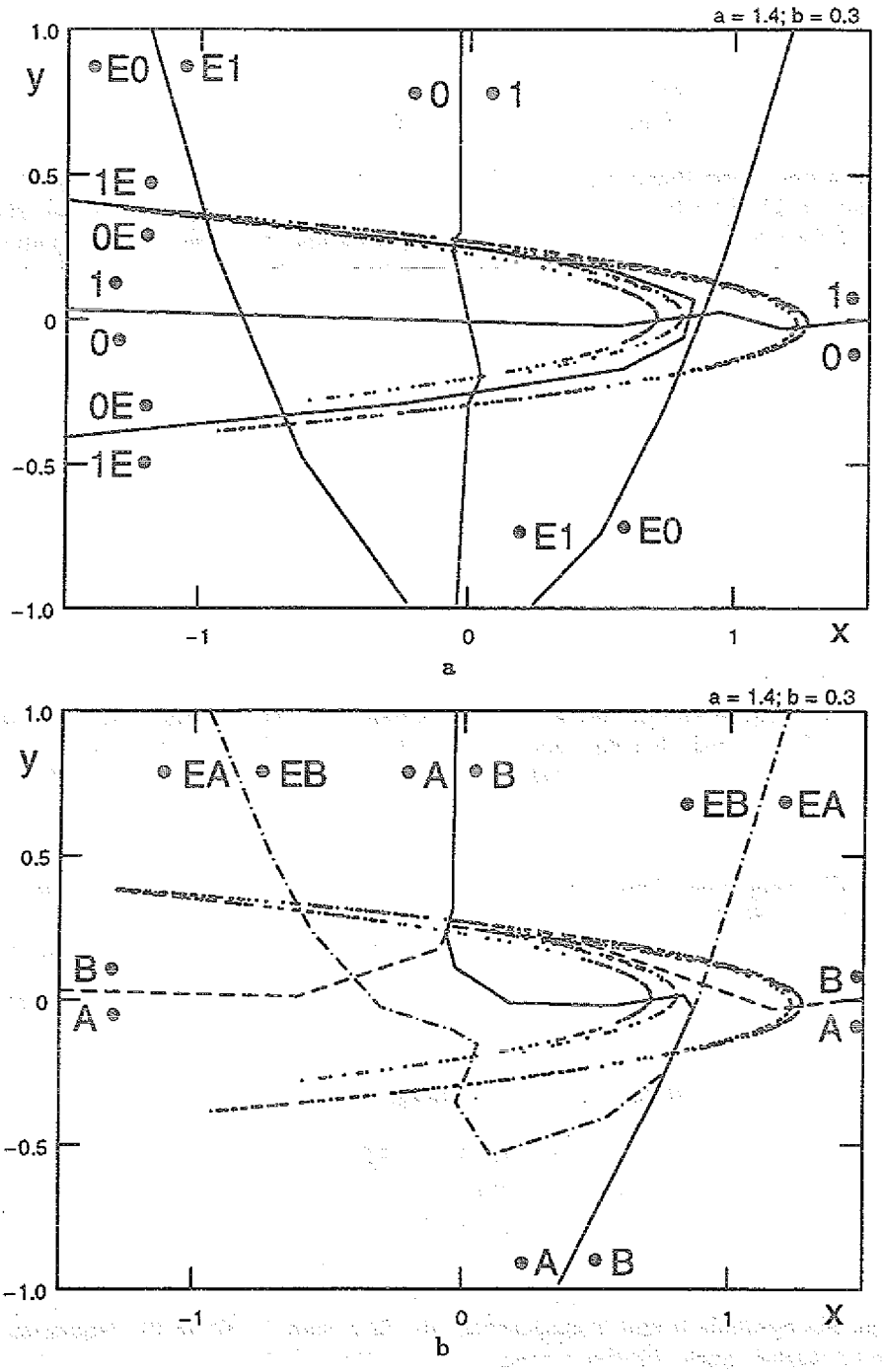


Abbildung C.2: Wiederholung der Abb. 5.2, in der die 01–Partition und die AB–Partition des Hénon Attraktors bei Standardparameterwerten verglichen wurden.

$$\Rightarrow \quad \circledast BABB \;=\; \emptyset \tag{C.66}$$

Da wir alle unendlichen 01-Symbolsequenzen in unendliche AB-Sequenzen und umgekehrt übersetzen können, sind beide Partitionen generierend, wenn eine von beiden generierend ist, das heißt wenn in

$$\{(0 \sqcup 1)^{-\infty} \circ (0 \sqcup 1)^{+\infty}\}' \;=\; \{(A \sqcup B)^{-\infty} \circ (A \sqcup B)^{+\infty}\}' \tag{C.67}$$

keine Menge mit mehr als einem Orbit enthalten ist.

In der folgenden Rechnung werden wir 01–Sequenzen durch AB–Sequenzen ausdrücken und umgekehrt. Dafür sind zusätzlich zu C.59 die folgenden Beziehungen hilfreich, die man unmittelbar durch Einsetzen der Ausdrücke C.61 und C.63 für 0 und 1 oder C.54 und C.55 für A und B erhält:

$$\circledast 00 \;=\; \circledast AAA \tag{C.68}$$
$$\circledast 10 \;=\; \circledast BA \tag{C.69}$$
$$1 \circledast 11 \;=\; \circledast BB \;. \tag{C.70}$$

Die Definitionen C.54 und C.55 enthalten nur Sequenzen bis zur Länge drei. Aus ihnen folgt, daß jedes Gebiet $\circledast S_0 S_1 S_2$; $S_i \in 0, 1$ ganz in einem Gebiet $\circledast E T_1$; ($T_1 \in \{A, B\}$) liegt.

$$\{\circ(0 \sqcup 1)^3\}' \;\subseteq\; \{\circ E(A \sqcup B)\}' \tag{C.71}$$

Indem wir diese Beziehung mehrmals mit f iterieren und den Schnitt

$$\{\circ(A \sqcup B)^n\}' \;\cap\; \{\circ E^n(A \sqcup B)\}' \;=\; \{\circ(A \sqcup B)^{n+1}\}' \tag{C.72}$$

bilden, können wir diese Beziehung auf Sequenzen beliebiger Länge $n \geq 3$ verallgemeinern:

$$\{\circ(0 \sqcup 1)^n\}' \;\subseteq\; \{\circ E(A \sqcup B)^{n-2}\}' \;, \tag{C.73}$$

auch auf halbunendliche:

$$\{\circ(0 \sqcup 1)^{+\infty}\}' \;\subseteq\; \{\circ E(A \sqcup B)^{+\infty}\}' \tag{C.74}$$
$$\{(0 \sqcup 1)^{-\infty} \circ\}' \;\subseteq\; \{(A \sqcup B)^{-\infty} E \circ\}' \;. \tag{C.75}$$

Da auch in den Ausdrücken C.61 und C.63 für 0 und 1 nur Symbolsequenzen bis zur Länge drei auftreten, folgt analog:

$$\{\circ(A \sqcup B)^n\}' \;\subseteq\; \{\circ E(0 \sqcup 1)^{n-2}\}' \tag{C.76}$$
$$\{\circ(A \sqcup B)^{+\infty}\}' \;\subseteq\; \{\circ E(0 \sqcup 1)^{+\infty}\}' \tag{C.77}$$
$$\{(A \sqcup B)^{-\infty} \circ\}' \;\subseteq\; \{(0 \sqcup 1)^{-\infty} E \circ\}' \;. \tag{C.78}$$

C.2.2 Orbitpaare der AB–Partition

Wir berechnen nun zunächst die Menge aller Orbitpaare der AB–Partition. Laut Gl. B.43 genügt es,

$$\Phi_{AB} \;=\; \{(A \sqcup B)^{-\infty} \circ [A, B](A \sqcup B)^{+\infty}\}' \tag{C.79}$$
$$\;=\; [\{(A \sqcup B)^{-\infty} \circ\}', \text{“}] \cap [\circ E\{(A \sqcup B)^{+\infty}\}', \text{“}] \cap \{(A \sqcup B)^2 \circ [A, B](A \sqcup B)^3\}' \tag{C.80}$$

zu berechnen, da alle anderen Orbitpaare durch Verschieben des Punktes $\circ$ oder Vertauschen von A und B folgen. Dabei hätten wir auch größere Exponenten als 2 und 3 wählen können, aber diese genügen für die folgende Rechnung.

Wir wollen Φ_{AB} mit 01-Sequenzen ausdrücken. Aus den Beziehungen C.77 und C.78 folgt sofort:

$$\Phi_{AB} \;\subseteq\; [\{(0 \sqcup 1)^{-\infty} E \circ\}', \text{“}] \;\cap\; [\{\circ E^2(0 \sqcup 1)^{+\infty}\}', \text{“}] \tag{C.81}$$

Es bleibt also noch der letzte Term in C.80 als 01–Sequenz umzuschreiben:

$$\Phi_{AB} \quad \subseteq \quad \bigcup_{\substack{T_1T_2\\T_3T_4}} \{T_1T_2 \circ [A,B]T_3T_4(A \sqcup B)\}^{\circ} \tag{C.82}$$

(mit $T_i \in \{A,B\}$). Dazu bedarf es einer Fallunterscheidung:

1. Fall: Kein AAB in $T_1T_2AT_3T_4$ oder $T_1T_2BT_3T_4$
   Nach den Übersetzungsregeln C.61 und C.63 wird dann jedes A von T_2AT_3 und T_2BT_3 in eine 0 und jedes B in eine 1 übersetzt. Es gibt also Symbole $S_2, S_3 \in \{0,1\}$ mit:

$$T_1T_2 \circledast AT_3T_4 \quad \subseteq \quad S_2 \circledast 0S_3 \tag{C.83}$$
$$T_1T_2 \circledast BT_3T_4 \quad \subseteq \quad S_2 \circledast 1S_3 \quad . \tag{C.84}$$

   In diesem Fall gilt also:

$$\{T_1T_2 \circ [A,B]T_3T_4\}^{\circ} \quad \subseteq \quad \{S_2 \circ [0,1]S_3\}^{\circ} \tag{C.85}$$

   Die mit A bzw. B übersetzten Sequenzen unterscheiden sich daher auch nur an einer Stelle.

2. Fall: $T_1T_2 \circ [A,B] \quad = \quad AA \circ [A,B]$
   Da $AA \circledast BA = \emptyset$, folgt:

$$\{AA \circ [A,B]T_3\}^{\circ} \quad = \quad \{AA \circ [A,B]B\}^{\circ} \tag{C.86}$$
$$= \quad \{[00 \circ 11, 01 \circ 11]\}^{\circ} \quad = \quad \{0[0,1] \circ 11\}^{\circ} \tag{C.87}$$

3. Fall: $T_2 \circ [A,B]T_3 \quad = \quad A \circ [A,B]B$
   Da $BA \circledast BB = \emptyset$, folgt:

$$\{T_1A \circ [A,B]B\}^{\circ} \quad = \quad \{AA \circ [A,B]B\}^{\circ} \tag{C.88}$$

   Diesen Fall haben wir eben schon behandelt.

4. Fall: $\circ[A,B]T_3T_4 \quad = \quad \circ[A,B]AB$
   Da $\circledast AABA = \circledast BABB = \emptyset$, folgt:

$$\{\circ[A,B]AB(A \sqcup B)\}^{\circ} \quad = \quad \emptyset \tag{C.89}$$

   Alle Fälle zusammen ergeben:

$$\{(A \sqcup B)^2 \circ [A,B](A \sqcup B)^3\}^{\circ} \quad \subseteq \quad \{[0,1] \circ (0 \sqcup 1)^2\}^{\circ} \quad \cap \quad \{(0 \sqcup 1) \circ [0,1](0 \sqcup 1)\}^{\circ} \tag{C.90}$$

Wenn wir diese Gleichung mit C.80 und C.81 verknüpfen, dann sehen wir, daß alle Orbitpaare der AB–Partition auch Orbitpaare der 01–Partition sind.

$$\Phi_{AB} \quad \subseteq \quad \{(0 \sqcup 1)^{-\infty}[0,1] \circ (0 \sqcup 1)^{+\infty}\}^{\circ} \quad \cap \quad \{(0 \sqcup 1)^{-\infty} \circ [0,1](0 \sqcup 1)^{+\infty}\}^{\circ} \tag{C.91}$$

Die 01–Partition ist nach dem Kriterium 4.1.2 der Einfachheit also mindestens so einfach wie die AB–Partition.

C.2.3 Orbitpaare der 01–Partition

Die Menge aller Orbitpaare der 01–Partition können wir durch ganz analoge Rechenschritte mit den Symbolen A und B ausdrücken. Sie entsteht aus

$$\Phi_{01} \quad = \quad \{(0 \sqcup 1)^{-\infty} \circ [0,1](0 \sqcup 1)^{+\infty}\}^{\circ} \tag{C.92}$$
$$= \quad [\{(0 \sqcup 1)^{-\infty} \circ\}^{\circ}, \text{“}] \quad \cap \quad [\circ E\{(0 \sqcup 1)^{+\infty}\}^{\circ}, \text{“}] \quad \cap \quad \{(0 \sqcup 1)^3 \circ [0,1](0 \sqcup 1)^2\} \tag{C.93}$$

durch Verschieben des Punktes $\circ$ oder Vertauschen der Symbole 0 und 1.

Wir wollen Φ_{01} mit 01-Sequenzen ausdrücken. Aus den Beziehungen C.74 und C.75 folgt:

$$\Phi_{01} \quad \subseteq \quad [\{(A \sqcup B)^{-\infty} E \circ\}', ''] \quad \cap \quad [\{\circ E^2 (A \sqcup B)^{+\infty}\}', ''] \tag{C.94}$$

Der letzte Term in C.93

$$\Phi_{01} \quad \subseteq \quad \bigcup_{\substack{S_1 S_2 \\ S_3 S_4}} \{(0 \sqcup 1) S_1 S_2 \circ [0,1] S_3 S_4 (0 \sqcup 1)\}' \tag{C.95}$$

(mit $S_i \in \{0,1\}$) kann erst nach einer Fallunterscheidung als AB–Sequenz geschrieben werden:

1. Fall: Kein 011 in $S_1 S_2 0 S_3 S_4$ oder $S_1 S_2 1 S_3 S_4$
   Nach den Übersetzungsregeln C.54 und C.55 wird dann jede 0 von $S_2 0 S_3$ und $S_2 1 S_3$ in ein A und jede 1 in ein B übersetzt. Es gibt also Symbole $T_2, T_3 \in \{A, B\}$ mit:

$$S_1 S_2 \circ 0 S_3 S_4 \quad \subseteq \quad T_2 \circ A T_3 \tag{C.96}$$
$$S_1 S_2 \circ 1 S_3 S_4 \quad \subseteq \quad T_2 \circ B T_3 \quad . \tag{C.97}$$

   In diesem Fall gilt also:

$$\{S_1 S_2 \circ [0,1] S_3 S_4\}' \quad \subseteq \quad \{T_2 \circ [A,B] T_3\}' \quad . \tag{C.98}$$

   Die beiden übersetzten Sequenzen unterscheiden sich daher auch nur an einer Stelle.

2. Fall: $\circ[0,1]S_3 \quad = \quad \circ[0,1]0$
   Den Fall ohne ein 011 in $S_1 S_2 0 S_3 S_4$ oder $S_1 S_2 1 S_3 S_4$ haben wir soeben schon behandelt. Es bleibt nur noch der Fall $S_1 S_2 \circ [0,1]0 = 01 \circ [0,1]0$ übrig. Wegen $01 \circ 10 = \emptyset$ gilt aber

$$\{01 \circ [0,1]0\}' \quad = \quad \emptyset \quad . \tag{C.99}$$

3. Fall: $S_2 \circ [0,1] S_3 \quad = \quad 0 \circ [0,1]1$
   Da $0 \circ 110 = \emptyset$ folgt:

$$\{0 \circ [0,1]1(0 \sqcup 1)\}' \quad = \quad \{0 \circ [0,1]11\}' \tag{C.100}$$
$$= \quad \{[A \circ AAB, A \circ ABB]\}' \tag{C.101}$$
$$= \quad \{A \circ A[A,B]B\}' \quad . \tag{C.102}$$

4. Fall: $S_2 \circ [0,1] S_3 \quad = \quad 1 \circ [0,1]1$
   Bisher ließen sich alle Orbitpaare der 01–Partition auf Orbitpaare der AB–Partition zurückführen. In diesem, dem letzten Fall ist dies nicht mehr möglich.

   Wegen $0E1 \circ 0 = \emptyset$ läßt sich dieser Fall schreiben als:

$$\{(0 \sqcup 1)^2 1 \circ [0,1]1(0 \sqcup 1)\}' \quad = \quad \{1(0 \sqcup 1)1 \circ [0,1]1(0 \sqcup 1)\}' \tag{C.103}$$

   Da außerdem eine Symbolsequenz 011 auftreten muß, erhält man drei Mengen von Orbitpaaren der 01–Partition, die nicht Orbitpaare der AB–Partition sind:

$$\{(0 \sqcup 1)^{-\infty} 101 \circ [0,1]10(0 \sqcup 1)^{+\infty}\}' \quad = \quad \{(B \sqcup A)^{-\infty} BA[B \circ A, A \circ B]BA(B \sqcup A)^{+\infty}\}' \tag{C.104}$$
$$\{(0 \sqcup 1)^{-\infty} 101 \circ [0,1]11(0 \sqcup 1)^{+\infty}\}' \quad = \quad \{(B \sqcup A)^{-\infty} BA[B \circ AA, A \circ BB]B(B \sqcup A)^{+\infty}\}' \tag{C.105}$$
$$\{(0 \sqcup 1)^{-\infty} 111 \circ [0,1]11(0 \sqcup 1)^{+\infty}\}' \quad = \quad \{(B \sqcup A)^{-\infty} BB \circ [AA, BB]B(B \sqcup A)^{+\infty}\}' \tag{C.106}$$

   Nicht alle verbotenen Symbolsequenzen sind bekannt. Wegen der vielen Möglichkeiten, eine Symbolsequenz ins Unendliche fortzusetzen, ist es aber zu vermuten, daß diese drei Mengen nicht alle leer seien. Dann ist die 01–Partition einfacher als die AB–Partition.

Anhang D

Neurobiologischer Anhang

Dieser Anhang schlägt einige spekulative Funktionsprinzipien für den Neocortex im allgemeinen und seine Pyramidenzellen im besonderen vor. Elementare neurobiologische Kenntnisse sind für das Verständnis des Folgenden von Nutzen. Dieser Anhang wurde deshalb an diese physikalische Arbeit angehängt, weil die Anatomie des Neocortex uns vermutlich einiges über die effiziente Analyse von nichtlinearen Dynamischen Systemen offenbaren kann (siehe Kap. 1). Umgekehrt beruht auch das Modell, das hier für die Pyramidenzellen des Neocortex vorgeschlagen wird, auf den Konzepten der Symbolischen Dynamik und auf den Erfahrungen, die ich bei der Suche nach generierenden Partitionen gewonnen habe. Einige dieser Beziehungen zwischen den Neuronalen Netzen und der Symbolischen Dynamik wurden bereits in Kap. 9 zusammengefaßt.

In diesem Text wird eine Erklärung für die erstaunliche Homogenität des Neocortex vorgeschlagen: Ein gemeinsames Funktionsprinzip für alle neocortikalen Areale. Dabei werden drei Eingaben zu Pyramidenzellen unterschieden. Basale Erregungen sollen den Reiz bestimmen, auf den das Neuron reagiert. Apikale Erregungen sollen die Dauer der Reaktion regulieren. Shunt–Synapsen sollen signalisieren, welche Aspekte der Reize das Neuron nicht zu unterscheiden braucht. Diese Funktionen werden durch geeignete Verallgemeinerungen der Hebbschen Lernregeln implementiert. Sowohl die Signalverarbeitung als auch die Lernprozesse werden nur qualitativ diskutiert. Mehrere Lernziele scheinen für ein Netzwerk solcher Neuronen erreichbar zu sein. Es kann sich auf bestimmte Aspekte der Sinnesdaten spezialisieren, langfristige Erscheinungen darin erkennen und Verhalten steuern. Die hierfür erforderliche Architektur des Netzwerkes stimmt mit bekannten anatomischen Tatsachen überein: mit der Gestalt und Lage der Pyramidenzellen, mit den Unterschieden zwischen dornigen und glatten Neuronen, mit der neuronalen Zusammensetzung neocortikaler Schichten, mit der Verschaltung der Schichten und mit der hierarchischen Ordnung neocortikaler Areale. Mögliche Funktionen des Thalamus, des Hippocampus, der Bipolarzellen und des REM–Schlafes werden nur am Rande erwähnt. Seine charakteristischten experimentellen Vorhersagen macht das Modell darüber, welche Axone eher dornige Neuronen und welche eher bestimmte Interneuronen erregen. Das Modell beruht weniger auf der Theorie neuronaler Netze als vielmehr auf den Methoden, mit denen in der Nichtlinearen Dynamik, insbesondere der Symbolischen Dynamik, generische zeitliche Muster analysiert werden. Dieser interdisziplinäre Zusammenhang wird allerdings nicht erläutert, da die Darstellung des Modells auf Mathematik verzichtet.

D.1 Einleitung

Der Neocortex zeigt die erstaunliche Fähigkeit, trotz eines überall recht ähnlichen Aufbaus ganz verschiedene Funktionen zu erfüllen. Dazu gehört die Verarbeitung von visuellen, auditiven und somatosensorischen Sinnesdaten durch den sensorischen Cortex, die Verknüpfung dieser Sinnesdaten durch den Assoziationscortex und die Kontrolle von Körperbewegungen durch den Motorcortex. Die Breite der sechs neocortikalen Schichten schwankt zwar von Ort zu Ort, und in manchen

Arealen können Schichten ganz verschwinden, ununterscheidbar verschmelzen oder in schmälere Schichten unterteilt sein, aber die neuronale Zusammensetzung und die Verbindung der Schichten hängt nur wenig vom Areal ab. Dieser grundlegende Schaltkreis wurde in den letzten Jahren in mehreren Übersichten dargestellt: [130], [84], [152], [170, Kap. 1–3], [89], [71, Kap. 1]. In dieser Arbeit wird eine Funktionsweise für ihn vorgeschlagen.

Die meisten anatomischen Messungen lassen die Frage offen, ob die synaptischen Verschaltungen einzelner Neuronen ebensowenig vom Areal abhängen wie die Verbindungen der Schichten. Denn leicht zu messen ist nur das Ausmaß, in dem die Axonkollaterale eines Neuronentyps mit den Dendritenbäumen eines anderen Neuronentypes überlappen, aber nicht die Zahl der dabei gebildeten Synapsen. Die unregelmäßige Gestalt der Neuronen, die große Divergenz und Konvergenz der neuronalen Verbindungen und andere Gründe [80, 71] sprechen dafür, daß genetisch nur die Größe der Überlappungen festgelegt ist, die tatsächliche Ausbildung von Synapsen aber dem Zufall überlassen bleibt. Inwieweit bestimmte Typen von Axonen und Dendriten dabei bevorzugt Kontakte bilden, ist eine offene Frage [170]. Das hier vorgeschlagene Modell fordert, daß Dendriten bestimmter Interneuronen mit bestimmten Axonen bevorzugt Synapsen bilden. Ansonsten können die Verknüpfungen innerhalb einer Schicht aber rein zufällig sein.

Die erstaunliche Gleichförmigkeit des Neocortex ließe sich erklären, wenn seine Areale nicht mit genetisch fixierten, spezialisierten Schaltkreisen, sondern alle nach dem gleichen Prinzip arbeiten würden. Dazu müßten sich außer der Anatomie der Areale auch die Lernprinzipien ähneln, nach denen sich die Stärke von Synapsen ändert. Die Funktion eines Areals hinge dann vor allem von seiner Lernerfahrung ab. Tatsächlich bleibt auch der Neocortex des erwachsenen Tieres sehr flexibel [135], und in seiner Phylogenese und Ontogenese deutet einiges darauf, daß Afferenzen einen großen Einfluß auf den zunächst eher unspezifischen Cortex haben [119]. Zum Beispiel können die Neuronen der primären Hörrinde lernen, so zu reagieren wie die Neuronen einer typischen primären Sehrinde, nachdem das Faserbündel, das sonst Sinnesdaten zur Sehrinde überträgt, operativ zur Hörrinde umgeleitet wurde [160].

Doch welche Funktion mag eine solch universelle Rolle spielen, daß alle Areale des Neocortex von ihr geprägt sind? Bei vielen Funktionen des Gehirns wird eine Beteiligung des Neocortex vermutet. Philosophisch interessante Funktionen wie das logische Denken oder das moralische und ästhetische Urteilen werden hier nicht diskutiert, ebenso das Speichern und Abrufen von deklarativem Wissen (Episoden, Fakten), die Lenkung der Aufmerksamkeit oder das Klassifizieren von statischen Mustern. Die Funktion, die hier als prägend für die Anatomie vorgeschlagen wird, ist dynamischer Natur: Der Neocortex lernt in den schnell veränderlichen Sinnesdaten Erscheinungen zu erkennen, die über längere Zeit existieren, sagt damit Sinnesdaten voraus und steuert sie. Ein Beispiel für dieses Lernziel ist das Erkennen vertrauter Laute inmitten von Lärm. Oder das Wahrnehmen eines Körpers als Einheit, auch wenn er sich bewegt, verformt, oder teilweise verdeckt ist. Oder die Steuerung der eigenen Bewegung um Hindernisse herum. Dies sind Aufgaben, bei denen Säugetiergehirne den Computern weit überlegen sind, doch mag dies an der überlegenen „Hardware", nämlich der enormen Dichte der Synapsen liegen (ca. $8 \cdot 10^8$ pro mm^3 [80, Kap.5], [71, Kap. 1.5]). Die „Software", nämlich die zugrundeliegenden mathematischen Prinzipien, könnten Mathematikern bereits bekannt sein.

Universelle Prinzipien, nach denen sich die zeitliche Struktur von Datensequenzen unbekannten Ursprungs analysieren läßt, werden in zwei Feldern der Wissenschaft untersucht: in der Theorie neuronaler Netze und in der Nichtlinearen Dynamik. Aus der Theorie der neuronalen Netze werden im folgenden nur elementare Konzepte verwendet, die bereits in die Lehrbücher der Neurobiologie Eingang gefunden haben. Denn die meisten neuronalen Netze wurden für statische Eingabemuster oder für zeitliche Folgen zufälliger Muster entworfen. Das hier vorgeschlagenen Modell ist vermutlich weniger effizient, da es nicht von der präzisen Verschaltung neuronaler Einheiten Gebrauch macht, die in numerischen Simulationen möglich ist.

Es soll auch nur die zeitlichen Muster bearbeiten können, die in der Nichtlinearen Dynamik als typisch für physikalische oder chemische Systeme betrachtet werden: Muster, die sich stetig in der Zeit ändern, und zwar nach deterministischen Gesetzen, die selbst zeitinvariant sind. Die Muster der fernen Zukunft können dabei unvorhersagbar sein, selbst wenn die Dynamik von keinem Zufall gestört wird. Die Theorie der Nichtlinearen Dynamik, die mit der Verbreitung billiger Computer

als Chaos–Theorie bekannt wurde, hat bisher die Neurobiologie kaum beeinflußt (e. g. [110, 120]). Doch aus ihr hat sich das Modell entwickelt, das hier für den Neocortex vorgeschlagen wird. Besonders wichtig waren dabei die Grundlagen der Symbolischen Dynamik. (Trotz des Namens besteht kein Bezug zu den Symbolen der kognitiven Psychologie und der künstlichen Intelligenz.) Diese Grundlagen werden zum Beispiel in [33] erklärt, oder detaillierter in [36, 70]. Im Gegensatz zu einem bekannten Modell für chaotische Dynamik im olfaktorischen Cortex [156] soll das hier postulierte Modell durch seine interne Dynamik einfach nur die äußere Dynamik der Umwelt widerspiegeln. Es soll allerdings auch eine chaotische äußere Dynamik noch analysieren können. Und nur in diesem Fall ist die interne Dynamik chaotisch.

Ein Brückenschlag zwischen zwei so verschiedenen wissenschaftlichen Disziplinen wie Neuroanatomie und Symbolischer Dynamik bleibt unvermeidlich höchst unvollständig. Viele Referenzen zu diesem Text sind nicht Originalarbeiten, sondern Übersichten der letzten Jahre (insbesondere [170, 80, 101, 153, 142]), in denen der Leser weitere Angaben finden kann. Die mathematischen Methoden und die Konzepte der Nichtlinearen Dynamik werden nur implizit verwendet. Nur Kap. D.4.7 erwähnt explizit einige dieser Beziehungen. So kann dieses Modell auch nicht mit anderen möglichen Implementierungen von mathematischen Prinzipien der Nichtlinearen Dynamik verglichen werden. Ein effizienter arbeitendes Alternativmodell ist aber schon deshalb nicht in Sicht, weil das hier vorgeschlagene Modell an die Grenzen dessen geht, was die Anatomie des Neocortex erlaubt. Nicht nur werden fünf verschiedene Funktionen für verschiedene Schichten postuliert, auch das Modell für Pyramidenzellen ist sehr komplex: Der Apikaldendrit muß nach anderen Prinzipien funktionieren als die Basaldendriten. Und die meisten Zweige dieser dornigen Dendriten müssen auch während der Aktivität des Neurons durch hemmende Shunt–Synapsen auf ihrem Ruhepotential gehalten werden. Diese Stichpunkte deuten schon an, daß ein detaillierter Vergleich des Modells mit der Anatomie des Neocortex möglich ist.

Die beiden Kapitel D.2 und D.3 postulieren die beiden fundamentalen Lernprinzipien und das Kapitel D.4 die Architektur des Modells. Die Lernprinzipien besagen nur, unter welchen Umständen Synapsen tendentiell stärker oder schwächer werden. Quantitative Lernregeln werden nicht vorgeschlagen. So bleibt auch die Diskussion der Lernfolgen auf einem qualitativen Niveau.

In Kapitel D.2.1 wird zunächst ganz allgemein die Information diskutiert, die Neuronen durch ihre Aktivität repräsentieren. Kapitel D.2.2 präzisiert das Lernziel der Pyramidenzellen, auf solche Reize zu reagieren, die erfahrungsgemäß über lange Zeit relevant sind. Dieses Ziel soll mit den in Kapitel D.2.3 postulierten Lernprinzipien erreicht werden. Sie bewirken, daß eine Pyramidenzelle am Anfang einer Periode starken Feuerns vor allem von ihren Basaldendriten aktiviert wird, während ihr Apikaldendrit wichtige modulierende Einflüsse ausübt. Nicht nur die Gestalt der Dendriten (Kap. D.2.4) rechtfertigt diese Annahme, sondern auch die weiteren Folgen des Wechselspiels von Apikal- und Basaldendriten: Gruppen von Neuronen können lernen, verschiedene Möglichkeiten der zukünftigen Entwicklung früher zu unterscheiden (Kapitel D.2.5) und sogar eine gewisse motorische Kontrolle darüber auszuüben. Dabei sind Basaldendriten eher am klassischen Konditionieren (Kapitel D.2.6) und Apikaldendriten am operanten Konditionieren (Kapitel D.2.7 und Kapitel D.4.8) beteiligt.

Kapitel D.3.1 setzt es Mengen benachbarter Neuronen zum Ziel, durch ihre gemeinsame Aktivität spezielle Aspekte der Sinnesdaten, wie zum Beispiel Form oder Farbe, möglichst genau zu beschreiben. Es wird postuliert, daß diese Spezialisierung der Neuronen durch eine Spezialisierung ihrer Dendritenzweige erreicht wird. Dazu müssen die Shunt–Synapsen lernen, ihre Dendritenzweige vor Störungen durch die anderen Dendritenzweige desselben Neurons zu schützen (Kap. D.3.2). Die geshunteten Neuronen können dann lernen, die Informationen besonders genau zu repräsentieren, die den hemmenden Interneuronen unbekannt sind, von denen die Shunt–Synapsen gebildet werden (Kap. D.3.3). Dieser Vorgang wird durch ein schematisches Beispiel illustriert (Kap. D.3.4).

Kapitel D.4 vergleicht das Modell mit der Anatomie des Neocortex, wodurch die Lernprinzipien erst plausibel werden. Dabei wird die Funktion der Schichten aus den bekannten Eingaben zu den Basaldendriten und zu den Apikaldendriten der Pyramidenzellen abgeleitet. Die Prinzipien aus Kapitel D.2 lassen die bekannte Hierarchie kortikaler Areale (Kap. D.4.2) in neuem Licht erscheinen (Kap. D.4.3). Auf jeder Stufe der Hierarchie lernt die Schicht 2+3 (Kapitel D.4.7), unterstützt

durch die Schicht 4 (Kapitel D.4.5), längerfristige Erscheinungen in den veränderlichen Sinnesdaten zu identifizieren. Die tiefen Schichten 5 und 6 dagegen integrieren längerfristige Vorhersagen von höheren Hierarchiestufen (Kapitel D.4.3). Mögliche Funktionen des Thalamus (Kap. D.2.6), der Hippocampusformation (Kap. D.4.8), der Bipolarzellen (Kap. D.4.4) und des REM-Schlafes (Kap. D.4.4) werden am Rande erwähnt. Die abschließende Diskussion in Kapitel D.5 faßt anatomische Tatsachen, die das Modell rechtfertigen, und einige experimentelle Vorhersagen zusammen, betont aber auch die Lücken des Modells.

D.2 Apikaldendriten und Basaldendriten

D.2.1 Hebbsches Lernen von Erwartungen

Wie kann ein Netzwerk von Neuronen lernen, Sinnesdaten vorherzusagen? Indem die Neuronen von den Sinnesdaten erregt werden und gleichzeitig lernen, ihre Aktivitäten gegenseitig vorherzusagen. Dies geschieht am einfachsten nach der wohlbekannten Hebbschen Lernregel für erregende Synapsen: Wenn über das präsynaptische Axon ein Aktionspotential ankommt, kurz vor oder während der postsynaptische Dendrit depolarisiert ist, dann wird die Synapse gestärkt. Die Stärkung soll aber so gering sein, daß erst nach häufigem Zusammentreffen der beiden Aktivitäten eine starke Erregung vom Axon auf den Dendriten ausgeübt wird.

Als Beispiel können zwei komplexe Zellen des visuellen Cortex betrachtet werden (Abb. D.1). Sie sollen beide auf Lichtbalken gleicher Orientierung und Bewegungsrichtung mit erhöhter Feuerrate reagieren, doch ihre rezeptiven Felder sollen so versetzt sein, daß ein stetig durch das Blickfeld wandernder Lichtbalken erst das Neuron 1 und dann das Neuron 2 erregt. Wenn solche stetigen Bewegungen häufiger sind als abrupte Sprünge, dann wird Neuron 2 nach der Hebbschen Regel lernen, von Neuron 1 erregt zu werden. Je sicherer die Vorhersage ist, desto stärker wird die erregende Synapse. Wegen dieser internen Erregung kann das Neuron 2 die erfahrungsgemäß richtige Position des Lichtbalkens auch dann signalisieren, wenn es keine eindeutigen Sinnesdaten erhält.

Diese Beispielneuronen reagieren auf recht spezielle Reize. Es wurde viel darüber diskutiert, ob solche spezifischen Reaktionen für neocortikale Neuronen typisch sind (siehe z. B. [77, 114, 151, 99]). Wenn jedes Neuron auf viele verschiedene Reize reagiert, und nur die Reaktion von großen Neuronenmengen eindeutig ist, dann besagt die Aktivität eines Neurons wenig über die gegenwärtige oder zukünftige Aktivität des restlichen Netzwerkes. Um das einfache Lernschema zu retten, wird im folgenden der umgekehrte Fall betrachtet: Jedes Neuron lernt zuverlässig auf eine Menge ähnlicher Reize zu reagieren, so wie die „Kardinalzellen" von Barlow [77]. Ein starker Einfluß des Zufalls auf die interne Dynamik des Netzes, zum Beispiel bei der Signalübertragung an Synapsen, ist deshalb in diesem Modell, im Gegensatz zu den Attraktor–Neuronalen–Netzen [73], nur schädlich. Die spezifischen Reize, auf die ein Neuron reagiert, müssen dabei nicht unbedingt Sinnesreize sein. Im folgenden werden zwar vor allem Daten aus dem visuellen Cortex und Beispiele aus der Welt der Farben und Formen diskutiert. Doch analoge Modellvorstellungen gelten für Reize, die vom Nervensystem oder vom Neocortex selbst erzeugt werden.

Über Zeiträume von mehr als ein paar Millisekunden können auf diese einfache Weise nur dann Erwartungen aufgebaut werden, wenn entweder sehr viele kurzfristige Vorhersagen aneinandergereiht werden oder wenn die Neuronen genügend lange Perioden erhöhter Aktivität zeigen. Nur in dem zweiten Fall kann der Apikaldendrit durch Verlängerung der neuronalen Aktivität die Basaldendriten so beeinflussen, wie es hier in Kapitel D.2.3 nötig sein wird. Zwar werden solche Perioden erhöhter Aktivität in den gemessenen Autokorrelationen neocortikaler Neuronen nicht gesehen [71], vermutlich weil darin die langen Zeiträume spontaner Aktivität dominieren. Aber bei der Messung der spezifischen Reize, auf die einzelne Neuronen reagieren, zeigen sich solche Perioden erhöhter Aktivität. Im präfrontalen Cortex, zum Beispiel, sind bei Verzögerte–Reaktionsaufgaben viele Zellen während der ganzen Wartezeit aktiv, auch wenn sie mehr als eine Minute dauert [103, 102].

Gerade im visuellen Cortex deutet vieles darauf hin, daß Areale auf höheren Hierarchiestufen abstraktere Eigenschaften repräsentieren, wie zum Beispiel beleuchtungsunabhängige Farbtöne,

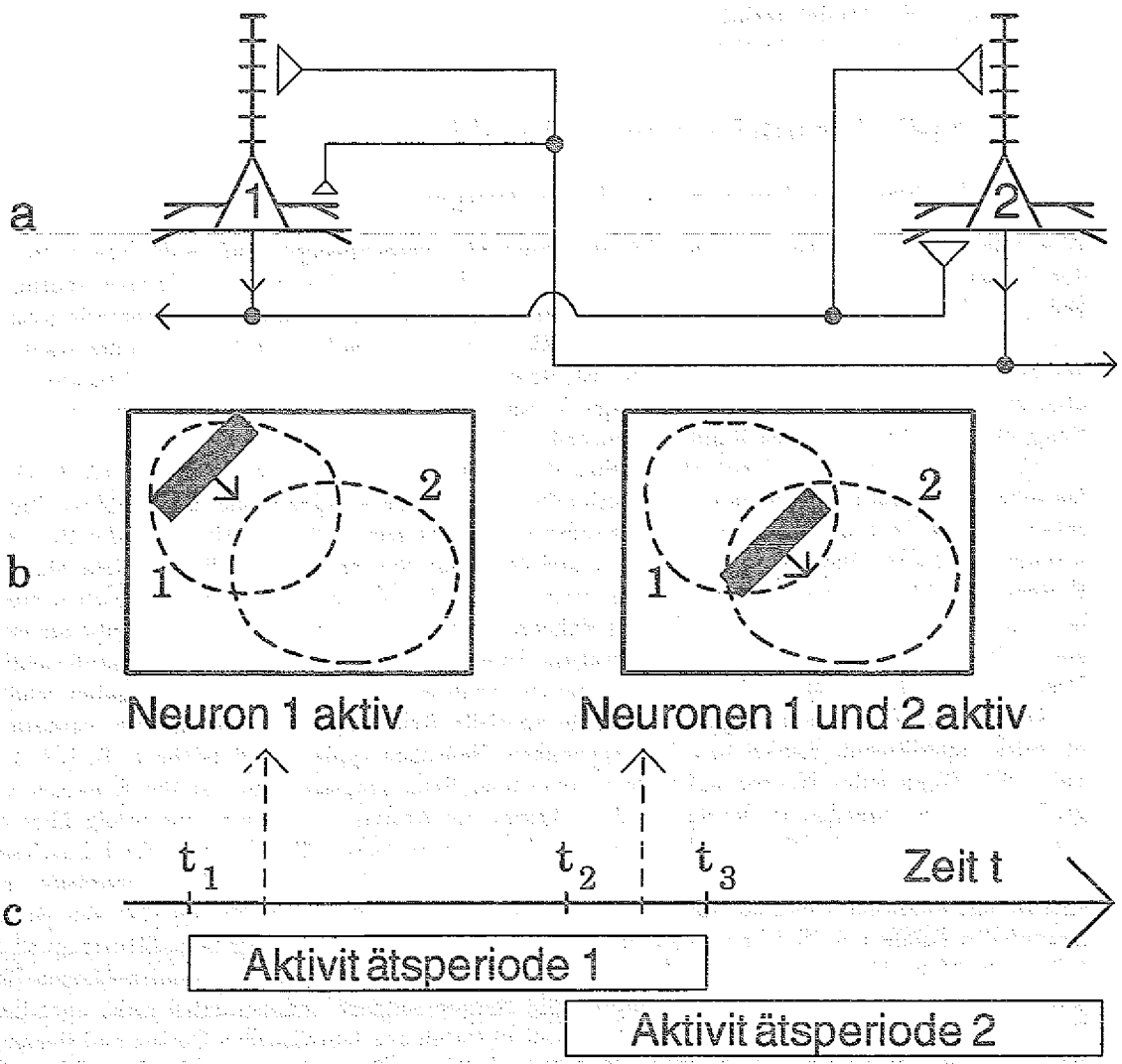


Abbildung D.1: Beispiel für die gegenseitige Erregung von Pyramidenzellen. Die Basaldendriten sind seitlich, der Apikaldendrit ist oberhalb und das Axon ist unterhalb des Zellkörpers eingetragen. Die kleinen Dreiecke stellen erregende Synapsen dar (a). Beide Pyramidenzellen reagieren auf die visuelle Präsentation bewegter Lichtbalken gleicher Orientierung und Bewegungsrichtung (b), das heißt, ihre Aktivitäten sind korreliert. Ihre rezeptiven Felder (gestrichelte Kreise) sind so gegeneinander versetzt, daß bei gleichmäßiger Bewegung das Neuron 1 früher als das Neuron 2 erregt wird (c). Nach den im Text postulierten Lernprinzipien bleibt die Erregung der Basaldendriten des Neurons 1 durch das Neuron 2 schwach, was durch eine kleine Synapse gekennzeichnet wird (a).

Farbsättigung, Helligkeit, visuelle Textur, Reflektivität, Transparenz, Form, absolute Größe, Formbeständigkeit, sowie Orientierung, Position, Bewegungsrichtung und Geschwindigkeit im dreidimensionalen Raum [88]. Dagegen repräsentieren niedrigere Hierarchiestufen einfache sensorische Zeichen in kleinen rezeptiven Feldern, wie zum Beispiel Wellenlänge oder Position auf der zweidimensionalen Netzhaut eines Auges zu einer bestimmten Zeit. Da die letzteren Eigenschaften sich normalerweise schneller ändern, sollten auf tieferen Hierarchiestufen die Neuronen im Mittel kürzere Perioden erhöhter Aktivität haben als auf höheren Hierarchiestufen.

Über das zeitliche Muster der Impulse während der Aktivitätsperiode sagt dieses Modell nichts aus. Die genaue Feuerrate ist weniger wichtig als die einfache Tatsache, daß die Feuerrate erhöht ist („value unit" statt „variable unit" in der Sprechweise von [76]). Das Model erfordert nicht, daß die einzelnen Aktionspotentiale einer Aktivitätsperiode ein spezielles zeitliches Muster bilden, und steht daher im Einklang mit experimentellen Hinweisen, daß die ersten 20–50 Millisekunden einer Aktivitätsperiode schon einen deutlichen Anteil der Gesamtinformation enthalten [164]. Und zeitlich synchrones Feuern mehrerer Neuronen ist möglich, aber nicht notwendig („gradual transmission between nodes" statt „synfire chains" in der Sprechweise von [71]).

D.2.2 Lernziel von Pyramidenzellen

Dieses einfache Hebbsche Lernschema läßt eine zentrale Frage offen: Auf welchen spezifischen Reiz reagiert ein Neuron? In den primären sensorischen Arealen sind Neuronen durch ihre kleinen rezeptiven Felder auf relativ einfache Reize beschränkt, die auf viele plausible Weisen festgelegt werden können. Doch Neuronen im Assoziationscortex können zwischen vielen Reizkombinationen wählen. Zum Beispiel die Neuronen, die auf verschiedene Ansichten derselben Gesichter reagieren [139, 87], können sicher nützliche Erwartungen aufbauen. Denn wenn ein Gesicht erschienen ist, dann mag es sich drehen und wenden, aber es wird normalerweise nicht so schnell wieder verschwinden. Doch wie lernen Neuronen solche Reaktionen? Bei einer willkürlichen Auswahl von Reizen wären die neuronalen Aktivitäten praktisch unkorreliert und Vorhersagen unmöglich. Und eine genetische Festlegung wäre nicht nur aufwendig, sondern auch unflexibel.

Vielleicht besteht die Antwort darin, die bisherige Argumentationskette umzudrehen. In diesem Modell ist daher das Lernziel einer Pyramidenzelle: auf solche Reize zu reagieren, auf denen sich langfristige Erwartungen aufbauen lassen. Jedes Neuron versucht also für sich, die Aktivitätsperioden anderer Neuronen möglichst sicher und möglichst früh anzukündigen. Als Gruppe kann das den Neuronen nur gelingen, wenn sie lernen, dauerhafte Erscheinungen in den flüchtigen Sinnesdaten zu erkennen. All die erwähnten Abstraktionsleistungen des Neocortex, wie Farbkonstanz, Erkennen von Gesichtern etc., wären dann auf die Suche nach immer dauerhafteren Erscheinungen zurückzuführen. Dieses Lernschema ist offensichtlich nur dann möglich, wenn sich die beobachteten Muster dynamisch ändern. Wie weit es tatsächlich genügt, um die Abstraktionsfähigkeit des Neocortex zu erklären, kann der Leser nur aus eigener Erfahrung beurteilen, solange kein numerisches neuronales Netz bei einer Suche in natürlichen Sinnesdaten dieselben dauerhaften Erscheinungen gefunden hat.

D.2.3 Lernprinzip von Pyramidenzellen

Wie kann ein Neuron erfahren, ob es durch seine Aktivitätsperiode die spätere Aktivität anderer Neuronen angekündigt hat? Chemische Signale, die sein Axon retrograd transportiert, können dem Neuron zwar mitteilen, ob es überhaupt starke Synapsen bildet und nicht überflüssig ist. Aber sie kommen zu spät, um verschiedene Aktivitätsperioden desselben Neurons zu unterscheiden. Dies können nur elektrische Signale. Das nächste Kapitel D.2.4 enthält Argumente, warum die Pyramidenzellen, die etwa 60 bis 90% der neocortikalen Neuronen ausmachen [141], geeignet sind, um solche Signale mit ihren Apikaldendriten zu empfangen. Doch zunächst werden die Modellannahmen über ihre Signalverarbeitung und ihre Lernprozesse postuliert.

Eine Pyramidenzelle erhält apikale und basale Erregungen (Abb. D.1 und D.2).

- Durch die Erregung der Basaldendriten sollen alle Aktivitätsperioden des Neurons gestartet

werden, das heißt, die Stärke der basalen Synapsen bestimmt, auf welche Reize das Neuron reagiert.

- Dagegen soll die Erregung des Apikaldendriten dem Neuron signalisieren, wie lange seine Aktivität benutzt wird, um andere Neuronen zu aktivieren. Das heißt, die Stärke der apikalen Synapsen bestimmt, wie lange das Neuron das schnelle Feuern fortsetzt.

Die dafür richtigen Stärken der erregenden Synapsen lassen sich nach Prinzipien erlernen, die sowohl in Apikal- als auch in Basaldendriten vom Hebbschen Typ sind, sich aber in der Berücksichtigung zeitlicher Reihenfolgen unterscheiden.

1. Auf Apikaldendriten sollen alle Synapsen gestärkt werden, bei denen prä- und postsynaptische Aktivität (positiv) korreliert sind, auch dann, wenn die postsynaptische Membran bereits depolarisiert ist, bevor das präsynaptische Aktionspotential ankommt. Im Fall von Abbildung D.1, in dem die Aktivitätsperioden von Neuron 1 und Neuron 2 für viele Muster überlappen, kann somit jedes Neuron lernen, den Apikaldendriten des anderen zu erregen, sofern synaptische Kontakte existieren.

2. Auf Basaldendriten sollen dagegen die Synapsen gestärkt werden, bei denen die präsynaptische der postsynaptischen Aktivität vorausgeht, und zwar um so mehr, je stärker und länger die postsynaptische Depolarisation dauert. Auf diese Weise lernt Neuron 1, da es früher aktiv ist, die Basaldendriten des Neurons 2 zu erregen, aber nicht umgekehrt Neuron 2 die des Neurons 1.

Die Reaktion der Neuronen 1 und 2 auf den bewegten Lichtbalken in Abbildung D.1 verläuft nach erfolgreichem Lernen so: Zuerst, zur Zeit t_1, fängt Neuron 1 an zu feuern. Da es gelernt hat, die Basaldendriten von Neuron 2 zu erregen, bewirkt es etwas später, zur Zeit t_2, zusammen mit anderen Neuronen die Aktivierung von Neuron 2. Auch die apikale Erregung des Neurons 2 beginnt etwa zu dieser Zeit t_2. Umgekehrt hat Neuron 2 gelernt, den Apikaldendriten von Neuron 1 zu depolarisieren. Durch solche apikale Erregungen wird die Aktivität von Neuron 1 aufrechterhalten.

Da aber die Zukunft nur begrenzt vorhersagbar ist, wird sich das Neuron 1 irgendwann an der Aktivierung weiterer Neuronen nur noch über schwache Synapsen beteiligen können. Zur selben Zeit werden nur schwache Synapsen auf dem Apikaldendriten von Neuron 1 ihre Erregung starten. Mit dieser Stagnation der apikalen Erregung soll die Aktivitätsperiode von Neuron 1 enden (Zeit t_3).

Auf diese Weise ist das Neuron gerade über den Zeitraum aktiv, in dem es sich am Aufbau von Erwartungen beteiligt. Je länger dieser Zeitraum ist, desto mehr werden die basalen Synapsen gestärkt, die das Neuron aktiviert haben. Dadurch lernt das Neuron auf die Reize zu reagieren, auf denen sich langfristige Erwartungen aufbauen lassen, das Lernziel von Kapitel D.2.2 wird erreicht.

D.2.4 Anatomische Kommentare

Diese Skizze der Aktivitäts- und Lernprozesse läßt zwar viele Details offen, aber zeigt doch zumindest, daß an allen Synapsen und Neuronen zur rechten Zeit die nötigen Signale vorhanden sind. So kann sich jede Pyramidenzelle auf ihrem eigenen Weg dem Lernziel nähern. Sie braucht dazu nur synaptische Kontakte zu anderen Neuronen, die auf ähnliche Reize reagieren. Besonders viele solcher Synapsen existieren in einem zufälligen Netzwerk dann, wenn diese Neuronen in der Nähe liegen. Durch solche Nachbarschaften von ein paar hundert Mikrometer Durchmesser, in denen Neuronen auf ähnliche Reize reagieren, ist die „modulare Struktur" des Neocortex definiert. Sie bildet daher einen nützlichen, wenn auch keinen notwendigen Bestandteil dieses Modells. Anatomische Untersuchungen konnten bisher nicht klären, ob Module hochentwickelte Schaltkreise mit wichtiger Funktion [115, 144] oder nur unwichtige Nebeneffekte der Lernprozesse [162] sind. Im Vergleich der Arten zeigt sich kein deutlicher Zusammmenhang zwischen Fähigkeiten der Tiere und der Existenz bestimmter modularer Strukturen [145].

Eine reziproke Verknüpfung der Neuronen, wie in Abbildung D.1, ist wohl selten, da selbst nahe Pyramidenzellen nur mit etwa 10% Wahrscheinlichkeit über eine Synapse verbunden sind [80]. Die

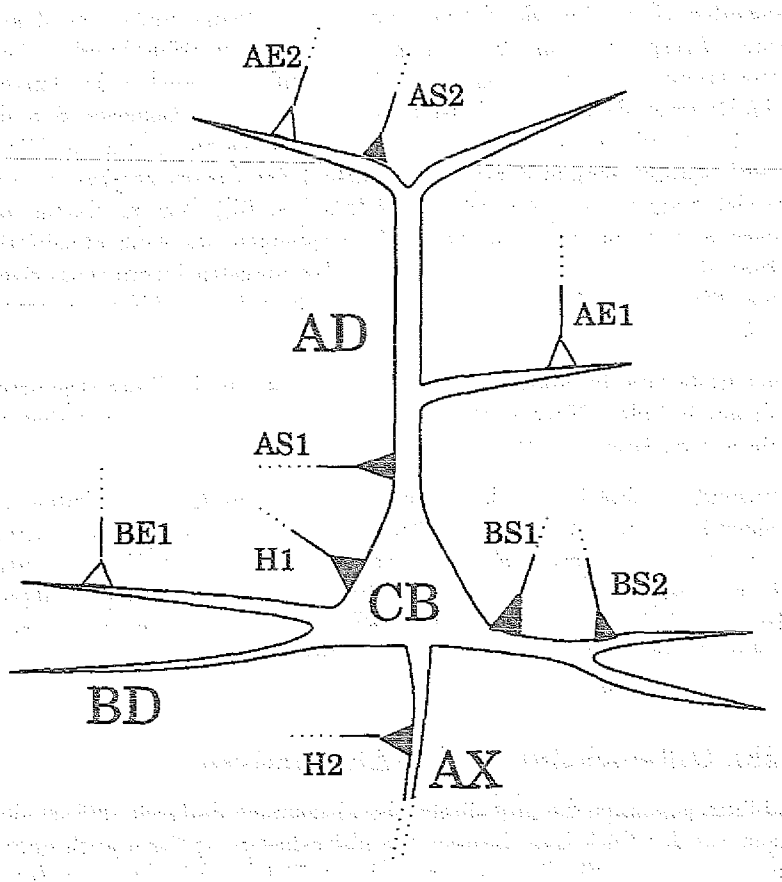


Abbildung D.2: Synapsen auf einer Pyramidenzelle. Die apikalen Dendritenzweige (oben) vereinigen sich zu einem Apikaldendriten (AD). Die basalen Dendritenzweige (seitlich) vereinigen sich zu mehreren Basaldendriten (BD). Das Axon (AX) beginnt an der Basis des Zellkörpers (CB). Erregende Synapsen sind als offene Dreiecke und hemmende Synapsen, das heißt hyperpolarisierende oder Shunt–Synapsen als schwarze Dreiecke dargestellt. Durch Shunt–Synapsen können einzelne Dendritenzweige (AS2, BS2) oder auch ganze Dendriten (AS1, BS1) vom restlichen Neuron elektrisch isoliert werden. Basale Erregungen (BE1) wirken wie klassisch konditionierte Reize. Apikale Erregungen bestimmen im Modell die Dauer der Reaktion (AE1), melden unkonditionierte oder konditioniere Reaktionen oder signalisieren diskriminative Reize beim operanten Lernen (AE2). Die neuronale Aktivität kann durch axo–somatale (H1) oder axo–axonale (H2) Synapsen gehemmt werden.

reziproke Verknüpfung ist für das Modell auch nicht notwendig, wenn ein Reiz in der Nachbarschaft nicht nur eine (Papst– oder Großmutter–) Zelle sondern eine ganze Schar spezialisierter (Kardinal–) Zellen aktiviert. Ein Teil dieser Schar kann von Neuron 1 basal erregt werden, während ein anderer Teil die apikale Rückmeldung zu Neuron 1 sendet.

Es gibt mehrere anatomische Gründe, warum die aktivierende Eingabe den Basaldendriten und die modulierende Eingabe den Apikaldendriten zugewiesen wurde. Manche hängen mit der Funktion der Schichten zusammen (siehe Kap. D.4). Ein offensichtlicher Grund ist die größere Länge der Apikaldendriten. Aus ihr folgern manche Neuroanatomen eine deutlich langsamere Leitung entfernter Erregungen zum Soma (e.g. ca. 40 statt 5 Millisekunden [159]) und deshalb nur einen modulierenden Einfluß des Apikaldendriten auf alle schnellen Denkprozesse.

Wichtiger im Rahmen dieses Modells ist die charakteristische Eigenschaft apikaler Dendritenzweige, sich vor dem Soma zu einem einzigen Apikaldendriten zu vereinigen (Die Ausnahmen der „extraverted cells" wurden bloß in der oberen Schicht 2 des Cortex gesehen, und zwar, abgesehen vom Hörcortex der Katzen [107], nur ziemlich selten [98, 86]). Nur in diesem, meist dornarmen Segment zwischen dem Soma und den ersten Verzweigungen des Apikaldendriten (Position AS1 in Abb. D.2) kann die Summe der Erregungen von den apikalen Dendritenzweigen auf das Soma begrenzt werden. Ein solcher Mechanismus mag zum Beispiel auf Shunt–Synapsen beruhen und dient zwei Zwecken:

- Er soll am Ende von der Aktivitätsperiode des Neurons die dann stagnierende apikale Erregung abrupt beenden. Noch vorhandene basale Erregung kann zum Beispiel durch apikale Hyperpolarisation kompensiert werden.
- Er soll verhindern, daß Pyramidenzellen nur durch den Apikaldendriten, ohne Beteiligung der Basaldendriten aktiviert werden. Zwar mag eine solche apikale Aktivierung dann geschehen, wenn der Neocortex als Teil des deklarativen Wissensgedächtnisses funktioniert („Gewußt was" im Gegensatz zum „Gewußt wie" des prozeduralen Habitgedächtnisses). Bei der bloßen Verarbeitung von Sinnesdaten wäre es jedoch verwirrend, wenn Reize repräsentiert würden, die zwar erfahrungsgemäß mit diesen Sinnesdaten assoziiert, zur Zeit aber nicht wahrgenommen werden.

D.2.5 Frühe Unterscheidung von Alternativen

Weitere Plausibilität gewinnen die postulierten Lernprinzipien dadurch, daß sie nicht nur dauerhafte Erscheinungen aus den flüchtigen Sinnesdaten abstrahieren, sondern auch noch andere wichtige Funktionen erfüllen können. Ein Beobachter wird zum Beispiel oft vor die Aufgabe gestellt, alternative Entwicklungen in der beobachteten Dynamik zu unterscheiden: Ein Objekt bewegt sich auf ein Hindernis zu und kann es entweder unsichtbar dahinter oder sichtbar davor passieren. Oder eine Lautfolge wird durch die ähnlichen Laute „p" oder „b" fortgesetzt, wobei der richtige Laut aus dem späteren Zusammenhang klar wird.

Nach dem Prinzip D.2.3.1 kann ein Apikaldendrit lernen, die Alternative zu signalisieren, auf die sein Neuron bevorzugt reagiert. Zum Beispiel die Neuronen, die „Objekt vor dem Hindernis" bevorzugen, zeigen korrelierte Aktivitäten und lernen deshalb, gegenseitig ihre Apikaldendriten zu erregen. Diese apikalen Erregungen erhöhen die Aktivität dieser Neuronen, wenn das Objekt tatsächlich vor dem Hindernis ist, was wiederum basale Synapsen stärkt. Die Neuronen lernen dadurch, ihre ursprüngliche Bevorzugung einer Alternative noch zu steigern. An ihrer Reaktion lassen sich die Alternativen dann deutlicher unterscheiden. Dies gilt auch für Neuronen, deren Aktivitätsperiode zu einem frühen Zeitpunkt beginnt, wenn sich die Alternativen noch ähneln. Denn nach dem Prinzip D.2.3.2 stärkt auch spät eintreffende postsynaptische Erregung eine dann nicht mehr aktive basale Synapse. Somit lernt das Netzwerk der Pyramidenzellen, häufig beobachtete Alternativen nicht nur deutlicher, sondern auch früher zu unterscheiden. Nur die Neuronen bleiben unbeeinflußt, die so früh aktiv werden, daß sich in ihren basal empfangenen Signalen die Alternativen noch nicht unterscheiden lassen.

Da sich Lernprinzipien von depolarisierende Synapsen durch Vorzeichenwechsel auf hyperpolarisierende Synapsen übertragen lassen (siehe Kap. D.5), kann derselbe Lernvorgang durch

hemmende Interneuronen ausgelöst werden, die von den Neuronen der einen Alternative aktiviert werden und die Neuronen der anderen Alternative hyperpolarisieren. Dieses Lernschema unterscheidet sich von lateraler Hemmung dadurch, daß die Neuronen der beiden Alternativen nicht räumlich voneinander getrennt sein müssen. Nur hemmende Synapsen, die die Feuerrate reduzieren (z. B. H1 in Abb. D.2), lösen solche Lerprozesse aus, nicht aber Synapsen (wie H2 in Abb. D.2), die Aktionspotentiale nur am Weiterlaufen entlang des Axons hindern.

D.2.6 Klassisches Konditionieren

So wie ein Apikaldendrit seinem Neuron signalisiert, welche neocortikalen Neuronen später aktiv werden, kann es auch den späteren Verlauf der Dynamik außerhalb des Neocortex melden. Dazu sind keine weiteren Lernprinzipien nötig. Diese Rückmeldung ist dann wichtig, wenn das Neuron die äußere Dynamik (durch motorische Steuerung) beeinflußt.

Neuronen vieler neocortikaler Areale projizieren zu denselben subcortikalen Strukturen. In der Entwicklung des Gehirns werden zunächst mehr dieser Efferenzen gebildet, als letztlich bestehen bleiben. Wahrscheinlich wird daher der Einfluß eines Neurons nicht nur genetisch, sondern auch durch die Erfahrung festgelegt [158]. Wie lernen die Neuronen zusammenzuwirken, die einen ähnlichen Einfluß auf die Sinnesdaten haben, die zum Beispiel alle mitwirken, dasselbe Augenlid zu schließen? Durch kein offensichtliches Merkmal von den Neuronen unterschieden, die das Augenlid öffnen, müssen sie die Folgen ihrer eigenen Aktivität erst selbst erkennen. Dazu müssen sie Erregungen erhalten, die das tatsächliche Schließen des Augenlides signalisieren. Nach den postulierten Lernprinzipien werden apikale Synapsen gestärkt, die solche Erregungen übertragen (e. g. AE1 in Abb. D.2). Denn auch wenn der Einfluß jedes einzelnen Neurons gering ist, wird doch eine positive Korrelation zwischen seiner Aktivität und dem späteren Schließen des Augenlides bestehen.

Durch diese apikalen Erregungen werden wiederum die basalen Synapsen gestärkt, die das Neuron aktivieren, sobald das Schließen des Auges vorherzusehen ist (e. g. BE1 in Abb. D.2). Wenn zum Beispiel ein Glockenton (zu konditionierender Reiz) erfahrungsgemäß dem Schließen des Auges (unkonditionierte Reaktion) vorausgeht, dann lernt das Neuron beim Glockenton basal erregt zu werden und wirkt mit, das Auge zu schließen (konditionierte Reaktion). Dies unterscheidet sich von den einfachsten Modellen klassischen Konditionierens dadurch, daß der unkonditionierte Reiz, zum Beispiel ein Windstoß, das lernende Neuron nicht direkt zu erregen braucht, er muß nur das Auge schließen.

Die meisten Efferenzen des Neocortex werden von den tiefen Schichten 5 und 6 gebildet. Tatsächlich liegen deren Apikaldendriten güstig, um von Neuronen erregt zu werden, die Sinnesreize repräsentieren (siehe Abb. D.5 und Kap. D.4.3). Die Neuronen der Schicht 5 projizieren zu vielen subcortikalen, insbesondere motorischen Kernen. Sie sollten deshalb Körperbewegungen und deren verzögerte, aber vielfältige Einflüsse auf die Sinnesdaten vorhersagen können. Ihre langen Apikaldendriten erhalten tatsächlich vielfältige Erregungen. Sie durchqueren all die oberen Schichten 4, 3 und 2, um schließlich in Schicht 1 Synapsen mit vielen langreichweitigen Axonen zu bilden.

Die Apikaldendriten aus Schicht 6 dagegen enden oft schon in Schicht 4 [98], in der auch die meisten thalamocortikalen Afferenzen enden. Auf diesem Weg werden Neuronen der Schicht 6 vermutlich dafür belohnt, Signale vom Thalamus richtig vorherzusagen. Tatsächlich ähneln sich die neuronalen Reaktionen der Schichten 4 und 6 des primären visuellen Cortex der Katze darin, daß einfache Zellen hier besonders häufig sind (e. g. [109]). Da Schicht 6 selbst wieder zum Thalamus projiziert, können dort Vorhersagen aller neocortikalen Areale miteinander und mit den Sinnesdaten verglichen werden.

D.2.7 Operantes Konditionieren

Durch operantes Konditionieren können sehr komplexe Verhaltensweisen gelernt werden, die dann wahrscheinlich von vielen Teilen des Gehirns gemeinsam gesteuert werden. Erstaunlicherweise kann im Rahmen dieses Modells auch jede einzelne Pyramidenzelle operant lernen, wenn die Lernprinzipien nur ein wenig modifiziert werden.

Zwei Voraussetzungen sind nötig, damit Pyramidenzellen operantes Verhalten steuern können:

- Die Situation muß alternative Verhaltensweisen gestatten, zum Beispiel Strecken und Beugen desselben Fingers.
- Jede Verhaltensweise kann dadurch ausgelöst werden, daß bestimmte Pyramidenzellen aktiv und die anderen passiv sind.

Im letzten Kapitel D.2.6 wurde beschrieben, wie Neuronen mit ähnlicher Wirkung lernen, zusammen zu agieren. Wenn beide Verhaltensweisen gleichermaßen möglich sind, dann werden die „beugenden" ebenso wie die „streckenden" Pyramidenzellen basal erregt. Die Neuronenmenge mit der früheren oder höheren Aktivität wird sich gegen die alternative Menge durchsetzen, indem sie diese wie in Kapitel D.2.5 hemmt. Dieser Wettbewerb kann durch apikale Erregungen entschieden werden. Zwar könnten auch basale Erregungen die Entscheidung bringen, doch die Steuerung operanten Verhaltens durch Apikaldendriten hat zwei Vorteile:

- In den basalen Synapsen ist das Wissen darüber gespeichert, welche zeitlichen Abläufe wahrscheinlich sind. Wenn operantes Lernen basale Synapsen modifizieren würde, könnte das Wissen darüber, welche Abläufe angenehm sind, mit dem Wissen interferieren, welche Abläufe wahrscheinlich sind.
- Die vielen Synapsen (wie AE2 in Abb.D.2), die ein Apikaldendrit mit langreichweitigen Axonen in Schicht 1 bildet, sind geeignet, um eine große Vielfalt diskriminativer Reize zu signalisieren.

Dazu muß das Hebbsche Lernprinzip D.2.3.1, daß aktive erregende Synapsen auf Apikaldendriten aktiver Neuronen gestärkt werden, etwas modifiziert werden: Die Stärkung dieser Synapsen soll dann besonders deutlich ausfallen, wenn die Schicht 1 einen positiv verstärkenden Reiz mitgeteilt bekommt. Da die Stärkung der Synapsen sicher ein langsamer Prozeß ist, eilt auch die Mitteilung des positiven Verstärkers nicht. Sie könnte zum Beispiel über unspezifische Afferenzen, von denen viele in Schicht 1 enden, und weiter über chemische Signalstoffe erfolgen. Umgekehrt soll ein negativ verstärkender Strafreiz die Stärkung der Synapsen gering ausfallen lassen, so daß neutrale Signale, die weder mit positiven noch mit negativen Verstärkern korreliert sind, im zeitlichen Mittel die unmodifizierten Lernprozesse auslösen.

Nachdem mehrmals in ähnlichen Situationen das Verhalten „Beugen des Fingers" positiv verstärkt wurde, haben die „beugenden" Pyramidenzellen also gelernt, durch erneutes Auftreten der diskriminativen Reize apikal erregt zu werden und dadurch den Wettbewerb gegen die „streckenden" Pyramidenzellen zu gewinnen. Frühe apikale Erregungen (oder Hemmungen) sind im Wettbewerb besonders wirksam, so daß die von Schicht 6 vorhergesagten Sinnesreize (Kap. D.2.6) für Schicht 1 relevant wären. Entsprechende Verbindungen existieren tatsächlich: Die meisten Martinotti-Zellen sitzen in den tiefen Schichten 5 und 6 und projizieren zur Schicht 1 [96, 129, 80]. Und tiefe Schichten projizieren auch relativ stark zur Schicht 1 anderer Areale (Feedback [101, 175]).

D.3 Lernen im Kontext

D.3.1 Lernziel von Neuronenmengen

Nach den bisherigen Lernprinzipien würden die Neuronen ihre Aktivität nicht aufeinander abstimmen. Im schlimmsten Fall würden alle auf ähnliche Reize reagieren, so daß sie fast alle redundant wären. Im Neocortex sind die Neuronen spezialisierter. Viele neocortikale Areale stehen auf ähnlichen Hierarchiestufen und scheinen auf verschiedene Aspekte der Sinnesdaten, wie zum Beispiel Form, Farbe oder Bewegung spezialisiert zu sein. Diese Spezialisierung ist notwendig: In einer komplexen äußeren Dynamik, die sich niemals exakt wiederholt, können sich Erwartungen nur auf einzelne Aspekte stützen, die sich wiederholen. Doch wie kann die Spezialisierung gelernt werden?

Und wie kann es vermieden werden, daß diese Spezialisierung zu weit geht? Wenn das eine Areal nur die Farbe und gar nicht mehr die Form, und das andere Areal nur die Form und gar nicht mehr die Farbe repräsentiert, dann läßt sich bei gleichzeitiger Repräsentation mehrerer Figuren nicht mehr feststellen, welche Form zu welcher Farbe gehört.

Für dieses wohlbekannte Problem („binding problem" [151, 94, 155] oder „cross-talk" [99, 76] bei „coarse-fine coding" [100, 99] oder „distributed representations" [114]) soll in diesem Kapitel ein Lösungsansatz vorgeschlagen werden, der aus mehreren Schritten besteht. Die Spezialisierung beginnt in den basalen Dendritenzweigen. In Kapitel D.3.2 werden Lernprinzipien für Shunt-Synapsen postuliert, die die Dendritenzweige so gegeneinander isolieren, daß sie auf viel spezifischere Reize reagieren als das ganze Neuron. Die Ausgabe des Neurons unterscheidet nicht, welche Dendritenzweige das Neuron erregen, so daß hier ein Teil der Information verlorengeht, die durch die Dendriten repräsentiert wird. Es wird in Kapitel D.3.3 vorgeschlagen, daß die Shunt-Synapsen gerade die Information übertragen, die verlorengehen darf. Das Neuron lernt dann, speziell die Information zu repräsentieren, die nur von den erregenden Synapsen übertragen wird. Daß dieser Lernprozeß möglich ist, ohne daß weitere Lernprinzipien postuliert werden müssen, illustriert Kapitel D.3.4 an einem Beispiel.

D.3.2 Lernprinzipien von Shunt-Synapsen

Wie lernen nun basale Dendriten, oder gar Dendritenzweige, auf ganz spezifische Reize zu reagieren? Ohne besondere Vorkehrungen würde das Aktionspotential des Neurons, das ja nicht nur Axon sondern auch Soma und zumindest die nahen Dendritenzweige polarisiert, alle Dendriten gleichermaßen beeinflussen und deshalb in Synapsen aller Dendriten analoge Hebbsche Lernprozesse auslösen. Dies kann durch Shunt-Synapsen verhindert werden.

Eine Shunt-Synapse (siehe z. B. [89]) öffnet Ionenkanäle in der postsynaptischen Membran, um deren Leitfähigkeit zu erhöhen. Wenn sie auf Dendritenschäften sitzt, dann dämpft sie jeden elektrischen Strom, der vom Soma an der Shunt-Synapse vorbei in den Dendriten oder in umgekehrter Richtung fließt, beeinflußt die anderen Dendriten aber kaum. Einzelne Dendriten oder Dendritenzweige können dadurch auf dem Ruhepotential gehalten und elektrisch vom restlichen Neuron isoliert werden (BS1 und BS2 in Abb. D.2). Im Neocortex wirken auf diese Weise die $GABA_A$-Rezeptoren, die schnell und häufig sind und durch Öffnen von Kanälen für Chlorionen die postsynaptische Membran nur schwach de- oder hyperpolarisieren [89, 83]. Es gibt Hinweise, daß diese $GABA_A$-Rezeptoren nur von einer Teilpopulation der hemmenden Interneuronen gesteuert werden [79] Die Mehrheit der GABAergischen Synapsen sitzt auf Dendritenschäften [78]. Dagegen werden die meisten erregenden Synapsen auf Dornen der Dendriten vermutet (Ausnahme in [74]). Und die allermeisten Dornen der Basaldendriten von Pyramidenzellen sitzen auf ihren entferntesten, nicht weiter verzweigten Zweigen [122]. Diese Synapsenverteilung erlaubt, daß erregende Synapsen auf verschiedenen Dendritenzweigen durch Shunt-Synapsen voneinander getrennt werden. Dagegen können sich erregende Synapsen desselben Dendriten gegenseitig in ihrer Wirkung steigern, wenn dort Flecken aktiver dendritischer Membran [146] oder hohe Konzentrationen von NMDA-Rezeptoren [134] sind. Und Messungen im visuellen Cortex deuten auf etwa 8–10 verschiedene erregbare Untereinheiten pro Zelle [93].

Dies alles weist darauf hin, daß die basalen Dendritenzweige mehr können, als nur Erregungen zu integrieren [150]. Komplexe Signalverarbeitung in Dendritenzweigen erfordert aber, daß die Synapsen entweder in bestimmter Ordnung auf die verschiedenen Dendriten verteilt (e. g. [100]) oder an einem Dendriten aneinandergereiht (e. g. [163]) werden. Es gibt aber keine experimentellen Hinweise auf solche exakten Positionierungen von Synapsen [90].

In diesem Modell dagegen dienen die Shunt-Synapsen in erster Linie den Lernprozessen, nicht der Signalverarbeitung. Diese Funktionen können sie auch dann lernen, wenn sie zufällig auf den Dendritenschäften verteilt sind. Der erste Lernschritt von Kapitel D.3.1 wäre abgeschlossen, wenn das Neuron nie mehr durch schwache Erregung vieler Dendritenzweige, sondern nur noch durch starke Erregung weniger Dendritenzweige aktiviert werden würde. Drei Zustände eines Dendriten wären dann noch möglich (Abb. D.3):

a) Der Dendrit wird so stark erregt, daß das Neuron an einem schwachen Shunt vorbei aktiviert wird.

b) Der Dendrit reagiert nicht oder wenig, aber das Neuron wird an anderen Teilen erregt und feuert. Ein starker Shunt schützt den Dendriten davor zu lernen, in dieser Situation zu reagieren.

c) Weder das Neuron noch sein Dendrit sind erregt.

Dieser Idealfall legt die Richtung des Lernens fest. Wenn aktive Synapsen noch nicht die in Abb. D.3 gezeigten Stärken haben, dann sollen sie sich ihnen zumindest nähern:

- Erregende Synapsen (Nummer 1, 3, 5 in Abb. D.3) lernen nach der üblichen Hebbschen Regel. In diese Regel geht eine postsynaptische Depolarisation ein, die sowohl von anderen Erregungen desselben Dendriten als auch vom Aktionspotential des postsynaptischen Neurons abhängen darf, aber jedenfalls von Shunt–Synapsen gedämpft werden kann.

- Shunt–Synapsen (Nummer 2, 4, 6 in Abb. D.3) dagegen müssen zwischen Aktionspotentialen und anderen Erregungen unterscheiden. Nahe am Soma gelegen, sollten sie dazu in der Lage sein. Sie ändern sich deutlich nur dann, wenn sie Aktionspotentiale spüren. Sie werden stärker bei starker und schwächer bei schwacher Erregung des Dendriten. Der dazwischen liegende Schwellenwert, wie auch all die anderen Lernparameter, darf in großen Zeiträumen variieren und dadurch die mittlere Aktivität von Dendriten und Neuron regulieren.

Nur numerische Simulationen werden zeigen können, ob ein solches Wechselspiel von erregenden und Shunt–Synapsen es tatsächlich erreichen kann, daß dornige Dendritenzweige ihr Neuron selten, dann aber stark erregen. man beachte, daß ein starker Shunt nur während der neuronalen Aktivitätsperiode wichtig ist (Abb. D.3.b). Tatsächlich konnten experimentelle und theoretische Untersuchungen an neocortikalen Pyramidenzellen außerhalb der Aktivitätsperioden nur schwache Shunts feststellen (siehe [81]).

D.3.3 Kontextuelle Interneuronen

Die zwei verschiedenen Eingaben zu dornigen Dendriten, die über erregende und die über Shunt–Synapsen, kommen von zwei verschiedenen Neuronentypen. Die meisten erregenden Synapsen werden von nahen oder fernen dornigen Neuronen gebildet. Die hemmenden Synapsen werden von glatten (d. h. dornlosen oder dornarmen) Interneuronen gebildet, die höchstens ein paar Millimeter entfernt liegen. Wenn Interneuronen auf dornigen Dendriten Shunt–Synapsen bilden, die nach den im letzten Abschnitt postulierten Prinzipien funktionieren, dann sollen sie „kontextuelle Interneuronen“ genannt werden. Eine Menge kontextueller Interneuronen, zum Beispiel alle in einer bestimmten Schicht eines bestimmten kortikalen Areals, repräsentieren durch ihre Aktivität Information, die „Kontext“ genannt werden soll. Der Kontext bestimmt, welche dornigen Dendriten geshuntet sind, und somit, wie das gegenwärtige Netzwerk aussieht, über das sich dornige Neuronen gegenseitig erregen.

Die Schichtstruktur des Neocortex wird in diesem Modell damit erklärt, daß verschiedene Schichten oder Teilschichten eines Areals verschiedene Kontexte repräsentieren. Die kontextuellen Interneuronen wirken dann wie lokale Relaystationen. Sie empfangen Signale von Axonen in ihrer Schicht, repräsentieren diese Information und geben sie an die dornigen Dendriten desselben Gebietes weiter. Tatsächlich enden Axonkollaterale und glatte Dendritensysteme relativ abrupt an Schichtgrenzen (im primären visuellen Cortex von Affen noch deutlicher als in dem von Katzen), im Gegensatz zu dornigen Dendritensystemen [127, 75]. Und die Axonkollaterale eines glatten Neurons überlappen meist stark mit dem Gebiet seiner Dendriten. Da sich die glatten Neuronen aber darin unterscheiden, über wieviele Schichten oder Teilschichten sich ihre Dendriten erstrecken, ist das Prinzip „Eine Schicht – Ein Kontext“ wohl nur eine erste Näherung der Wirklichkeit.

Auch andere anatomische Kennzeichen hemmender Neuronen machen in diesem Modell Sinn. Die gleichmäßige Verteilung hemmender und erregender Synapsen auf ihren Dendriten (und Somata) wäre kaum mit den komplizierten Lernprozessen zu vereinbaren, die für dornige Neuronen

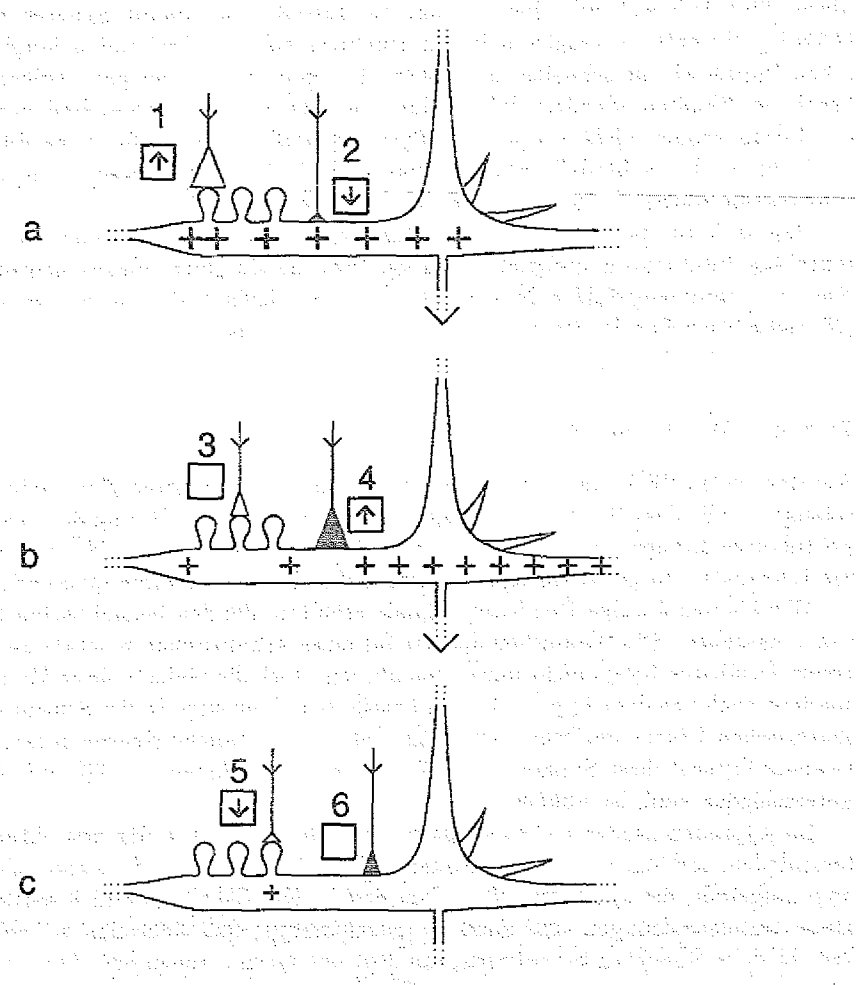


Abbildung D.3: Die drei idealen Erregungszustände eines Basaldendriten dieses Modells. Der Dendrit ist entweder stark erregt (a), geshuntet (b) oder passiv (c). Depolarisierung des Dendriten oder des Zellkörpers wird durch Pluszeichen symbolisiert. In (a) und (b) feuert die Pyramidenzelle. Nur aktive Synapsen sind eingezeichnet. Die ideale Stärke einer aktiven erregenden Synapse (offenes Dreieck auf Dorn) und einer aktiven Shunt–Synapse (schwarzes Dreieck auf Schaft) sind durch die Größe (groß–mittel–klein) der Dreiecke angedeutet. Die Richtung der Lernprozesse, die solche synaptischen Stärken erzeugen, wird durch den eingerahmten Pfeil symbolisiert: Pfeil nach oben bedeutet Stärkung, Pfeil nach unten Schwächung, ein leeres Kästchen geringe Änderung der Synapse.

gefordert wurden. Aber einfache Hebbregeln genügen, damit kontextuelle Interneuronen ihre Funktion lernen können. Um die Möglichkeit zu haben, jeden Dendriten jederzeit zu shunten, sollten die kontextuellen Interneuronen von Anfang an genug aktive Shunt–Synapsen bilden. Die beobachteten dichten lokalen Verzweigungnen ihrer Axone können vermutlich verschiedene Dendriten des gleichen Neurons gleichzeitig hemmen. Daß erregende Synapsen so viel häufiger sind (85–90% aller Synapsen im primären visuellen Cortex der Katze [141], ca. 90% in diversen Arealen der Maus [80]), läßt sich mit einer häufigeren Aktivität der Shunt–Synpsen erklären. Die schnellen $GABA_A$–Rezeptoren reagieren länger (mehrere zehn Millisekunden lang) auf einen präsynaptischen Impuls als die schnellen erregenden Rezeptoren (nur ein paar Millisekunden)[89]. Ununterbrochenes Shunten erfordert daher keine hohen Feuerraten, und außerdem adaptiert die Feuerrate der Interneuronen nicht wie die der Pyramidenzellen [131, 118]. Auch dürfen die Interneuronen im Rahmen dieses Modells weniger spezifische Reize signalisieren (d. h. die Repräsentation der Information durch die Interneuronen ist besonders verteilt).

Demnach ist die Gestalt der meisten hemmenden Interneuronen für die Funktion als kontextuelles Interneuron geeignet. Dazu gehören die als „local plexus neurons“ [143] oder anderen Namen zusammengefaßten Neuronentypen, einschließlich der „neurogliaform or spiderweb cells“ (Neroglioformzellen [116]), und vielleicht auch die Korbzellen. Welche Typen von Interneuronen tatsächlich den Kontext repräsentieren, bleibt unbekannt.

D.3.4 Wirkung des Kontexts

Auf den ersten Blick mag es irrelevant erscheinen, ob Dendriten gleich oder verschieden reagieren, solange sie alle dasselbe Neuron erregen. Bei genauerer Betrachtung des Zusammenspiels der bisher postulierten Lernprinzipien zeigt sich aber, daß sich die dornigen Neuronen spezialisieren können, die Informationen genau zu repräsentieren, die sie nur von ihren erregenden Synapsen erhalten.

Wie können dornige Dendriten Signale erhalten, die den benachbarten kontextuellen Interneuronen entgehen? Die Wahrscheinlichkeit für einen synaptischen Kontakt zwischen einem Axon und einem Dendriten hängt nicht nur davon ab, wie stark die Gebiete ihrer Verzweigungen überlappen, sondern auch von dem Typ der beiden beteiligten Neuronen. In der Schicht 4 des primären somatosensorischen Cortex der Maus, zum Beispiel, bilden manche Neuronen fast keine und andere über zwanzig Prozent ihrer Synapsen mit Afferenzen vom Thalamus [170]. Wie fein die Neuronentypen unterschieden sind, ist unklar.

Im folgenden werden nicht nur die dornigen Neuronen von den kontextuellen Interneuronen unterschieden, sondern auch die Neuronen in jeder Schicht von jedem kortikalen Areal als besonderer Typ aufgefaßt, der spezifische Eingaben erhält. Nur Schicht 2 und 3 werden nicht unterschieden. Diese Unterscheidungen sind dadurch gerechtfertigt, daß Neuronen verschiedener Schichten auch verschiedene Schichten bevorzugen, um dort mit ihren Axonen oder Dendriten Synapsen zu bilden (siehe z. B. [130] für intraareale und [165] für interareale Verbindungen).

Es soll nun als Beispiel die Aufgabe betrachtet werden, Form und Farbe geometrischer Figuren zu erkennen. Die kontextuellen Interneuronen sollen von einem anderen kortikalen Areal erregt werden, das bereits darauf spezialisiert ist, die Form zu erkennen. Die einzelnen Dendriten eines Neurons sind also nur dann nicht geshuntet, wenn zum Beispiel ein Dreieck allein zu sehen ist, ein Kreis und ein Quadrat zusammen, und so weiter. Innerhalb seines Kontexts hat ein Dendrit gelernt, nur auf Figuren zu reagieren, die sich in Form und Farbe ähneln, also zum Beispiel ein einzelnes rotes Dreieck, ein einzelnes blaues Quadrat, oder eine blaue runde und eine grüne eckige Figur zusammen. Ein Neuron mit solchen Dendriten reagiert auf ganz verschiedene Reize. Dies illustriert Abb. D.4.a mit vier, anstatt der circa 30 basalen Dendritenzweigen typischer Pyramidenzellen [124]. Wie kann dieses hypothetische Neuron lernen, speziell die Farbe zu repräsentieren?

In dem Grenzfall extrem starker Shunt–Synapsen spüren die erregenden Synapsen bei Reizen außerhalb ihres Kontexts keine Depolarisation und bleiben daher unverändert. An einem Neuron beeinflussen sich dann nur die Dendriten gegenseitig, die in demselben Kontext reagieren, zum Beispiel ein Dendrit 1, der auf rote Dreiecke reagiert, und ein Dendrit 2, der auf blaue Dreiecke reagiert (Abb. D.4.a). Wenn die Erregung des einen Dendriten das Neuron aktiviert, dann ist der andere Dendrit ungeshuntet und lernt nach der Hebbschen Regel von Abbildung D.3 auf denselben

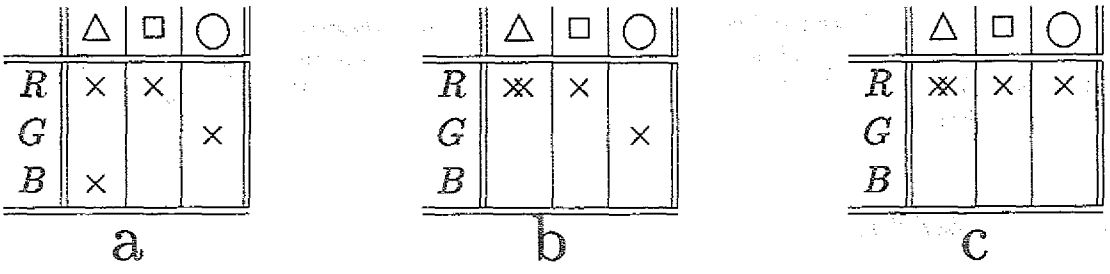


Abbildung D.4: Schematische Beispiel für das Lernen im Kontext. In (a–c) sind die Reaktionen eines hypothetischen Neurons in aufeinanderfolgenden Lernphasen dargestellt. R, G, und B bezeichnet eine rote, blaue oder grüne Figur (Dreieck, Quadrat oder Kreis). Jedes Kreuz in der Tabelle bedeutet, daß bei Präsentation der entsprechenden farbigen Figur ein Dendrit stark erregt wird und das Neuron aktiviert. Jede Spalte bezeichnet einen anderen Kontext, so daß zum Beispiel bei Präsentation eines Dreiecks die auf Quadrate und Kreise reagierenden Dendriten geshuntet sind. In (a) besteht noch keine Beziehung zwischen den Reaktionen der vier Dendriten. In (b) haben die beiden Dendriten innerhalb desselben Kontexts „Dreieck" gelernt, auf denselben Reiz zu reagieren. In (c) haben alle Dendriten ihre Reaktionen soweit aneinander angepaßt, wie es ihnen innerhalb ihrer ursprünglichen Kontexte möglich ist. Das Neuron reagiert dann nur noch auf eine Farbe.

Reiz zu reagieren. So können beide Dendriten Reaktionen auf sowohl rote als auch blaue Dreiecke entwickeln, doch dieser Zustand ist nicht stabil unter den Lerprozessen. Denn wenn das Neuron stärker auf einen der beiden Reize, zum Beispiel auf das rote Dreieck, reagiert, dann wird diese Bevorzugung durch die Hebbschen Lernprozesse noch weiter verstärkt. Jeder Mechanismus, der die mittlere dendritische Aktivität beschränkt, wird dabei die Reaktion auf das blaue Dreieck schwächen, bis sie ganz verschwunden ist (Abb. D.4.b). Auf diese Weise verlernt das Beispielneuron die Reaktion auf gleiche Formen verschiedener Farben.

Die angestrebte Reaktion auf verschiedene Formen gleicher Farbe kann allerdings nur dann gelernt werden, wenn die Shunt–Synapsen außerhalb des richtigen Kontexts doch ein wenig von dem Aktionspotential in ihre Dendriten hineinlassen. Die Dendriten werden dann lernen, auf ähnliche Reize zu reagieren, so weit das innerhalb ihrer verschiedenen Kontexte möglich ist, also zum Beispiel alle auf rote Formen (Abb. D.4.c).

Ob das Zusammenspiel der Lernprozesse wirklich dieses Ziel erreicht, hängt natürlich von der quantitativen Form der Lernregeln ab. Hier sollte nur gezeigt werden, wie eine Spezialisierung auf der Ebene von Einzelneuronen möglich ist, ohne daß unrealistische Wechselwirkungen zwischen Dendriten angenommen werden müssen. Das Beispiel ist recht künstlich, weil statische, nicht dynamische Muster betrachtet wurden und weil die wichtige Trennung von Farbe und Form schon in der Eingabe zum primären visuellen Cortex angelegt ist (e. g. [175]). Doch die weitere funktionelle Spezialisierung auf höheren Hierarchiestufen kann in analoger Weise erlernt werden. In dem obigen Beispiel war ein kortikales Areal bereits darauf spezialisiert, die Form zu repräsentieren. Aber auf dieselbe Weise könnten mehrere kortikale Areale auch gleichzeitig ihre Spezialisierung auf verschiedene Aspekte der Sinnesdaten steigern, indem sie gegenseitig ihre kontextuellen Interneuronen erregen. Man beachte, daß kontextuelle Interneuronen spezifischere Erregungen erhalten sollten als dornige Neuronen. Denn weitere, unspezifische Eingaben zu kontextuellen Interneuronen könnten die Spezialisierung und ergo die Funktionsweise des Netzwerkes verschlechtern, im Gegensatz zu unspezifischen Erregungen der dornigen Neuronen, die nur überflüssige Synapsen erfordern.

Wie Abbildung D.4.b exemplarisch zeigt, kann ein äußerer Beobachter an der Aktivität der dornigen Neuronen die Sinnesdaten so genau ablesen wie an der Aktivität ihrer Dendriten, sofern er den Kontext, das heißt die Aktivität der Neuronen im anderen kortikalen Areal, kennt. Die Aktivitäten in beiden Arealen sind also so aufeinander abgestimmt, daß sie gemeinsam eine genaue Repräsentation der Sinnesdaten geben. Das in Kapitel D.3.1 erwähnte Problem zu starker Spezialisierung wird vermieden. Wenn zum Beispiel der Kontext ein Dreieck und einen Kreis signalisiert

und eine Figur rot, die andere blau ist, dann können Dendriten reagieren, die auf rote eckige und blaue runde Formen ansprechen, oder solche, die umgekehrt auf rote runde und blaue eckige Formen ansprechen. Beide Dendritentypen an demselben Neuron würden Verwechslungen der beiden Fälle ermöglichen. Aber zwei solche Dendriten konkurrieren dann ebenso wie die Dendriten 1 und 2 des obigen Beispiels. Das Neuron lernt dadurch, nur noch auf eine Färbung zu reagieren.

D.4 Schichten

D.4.1 Vertikale Gliederung der Areale

Die vertikale Unterteilung der kortikalen Areale in Schichten ist auffälliger und regelmäßiger als die horizontale Unterteilung in Module, die in Kapitel D.2.4 erwähnt wurde. Auf Untersuchungen, wie die Neuronentypen verschiedener Schichten miteinander verbunden sind, stützen sich die Vorschläge, wie die grundlegenden Schaltkreise des Neocortex aussehen könnten [130, 170, 89, 152]. Der hier verwendetet Schaltkreis (Abb. D.5 und Abb. D.6) ähnelt dem Schaltkreis in [152, 154]. Nur die Rolle der hemmenden Interneuronen und der cortikocortikalen Verbindungen zu und von tiefen Schichten werden stärker betont.

Die Funktionen der neocortikalen Schichten wird hier aus der Lage ihrer Apikal- und Basaldendriten gefolgert, ausgehend von den in Kapitel D.2 postulierten Funktionen dieser Dendriten. Daß diese Funktionen zusammen mit den Afferenzen und Efferenzen der Areale ein sinnvolles Gesamtbild ergeben, rechtfertigt die Postulate im nachhinein. Durch Lernen im Kontext (Kap. D.3) können diese Funktionen verbessert werden, falls die kontextuellen Interneuronen spezifische Erregungen erhalten. Da es schwer zu messen ist, ob ein erregendes Axon mehr Synapsen mit dornigen Neuronen oder mit kontextuellen Interneuronen bildet, können nur wenige dieser Vorhersagen mit experimentellen Daten verglichen werden.

D.4.2 Basale Verbindungen

Unter Vernachlässigung von Verbindungen, die aus relativ wenigen Synapsen bestehen, wurde ein serieller Schaltkreis vorgeschlagen (z. B. [115, 109, 108]): Die Eingabeschicht 4 empfängt Daten, insbesondere vom Thalamus, und leitet sie an die Schichten 2 und 3 weiter. Diese beiden Schichten ähneln sich so sehr in Aufbau und Funktion (siehe [80, Kap. 28] für Ausnahmen), daß sie im folgenden als eine Schicht 2+3 behandelt werden. Von dieser oberflächlichen Schicht werden Signale zu den tiefen Ausgabeschichten 5 und 6 gesendet. Diese Verbindung besteht entweder aus den Apikaldendriten tiefer Schichten oder aus tiefen Kollateralen der Axone oberflächlicher Neuronen (Abb. D.5). In dem Modell von Kapitel D.2 haben die Apikaldendriten nur einen sekundären Einfluß auf die neuronale Aktivität. Deshalb dominieren die Kollaterale, die Basaldendriten erregen. Ebensolche Kollaterale existieren aber auch von tiefen zu oberflächlichen Schichten (Übersicht in [170, Kap. 2]).

Auch die schon in Kapitel D.2.1 erwähnten hierarchischen Verbindungen zwischen den Arealen modifizieren das Bild einer seriellen Verarbeitung [165]. Anzeichen für solche Hierarchien wurden im sensorischen Cortex und Assoziationscortex mehrerer Tierarten gefunden [101]. Auf der untersten Stufe stehen die primären sensorischen Cortices. Die aufsteigenden „Feedforward–Verbindungen" gehen vor allem von Schicht 2+3 der niedrigeren zu Schicht 4 der höheren Stufe. Die absteigenden „Feedback–Verbindungen" starten bevorzugt in den Schichten 5 und 6 der höheren Stufe und vermeiden Schicht 4 der niedrigeren Stufe [165]. Die Feedback–Verbindungen zu Schicht 2+3 sind im Rahmen dieses Modells überflüssig und anscheinend auch relativ schwach ([165, Tab. 2], [175], Ausnahmen in [101]). Abbildung D.7 faßt die wichtigsten Verbindungen über Basaldendriten zusammen. Mehrere spezialisierte Areale können auf der gleichen hierarchischen Stufe stehen, Verbindungen können Stufen überspringen, und schwache Verbindungen können überall auftreten und die Schaltung komplizieren. „Laterale Verbindungen" zwischen ähnlichen Hierarchiestufen sind weniger schichtspezifisch.

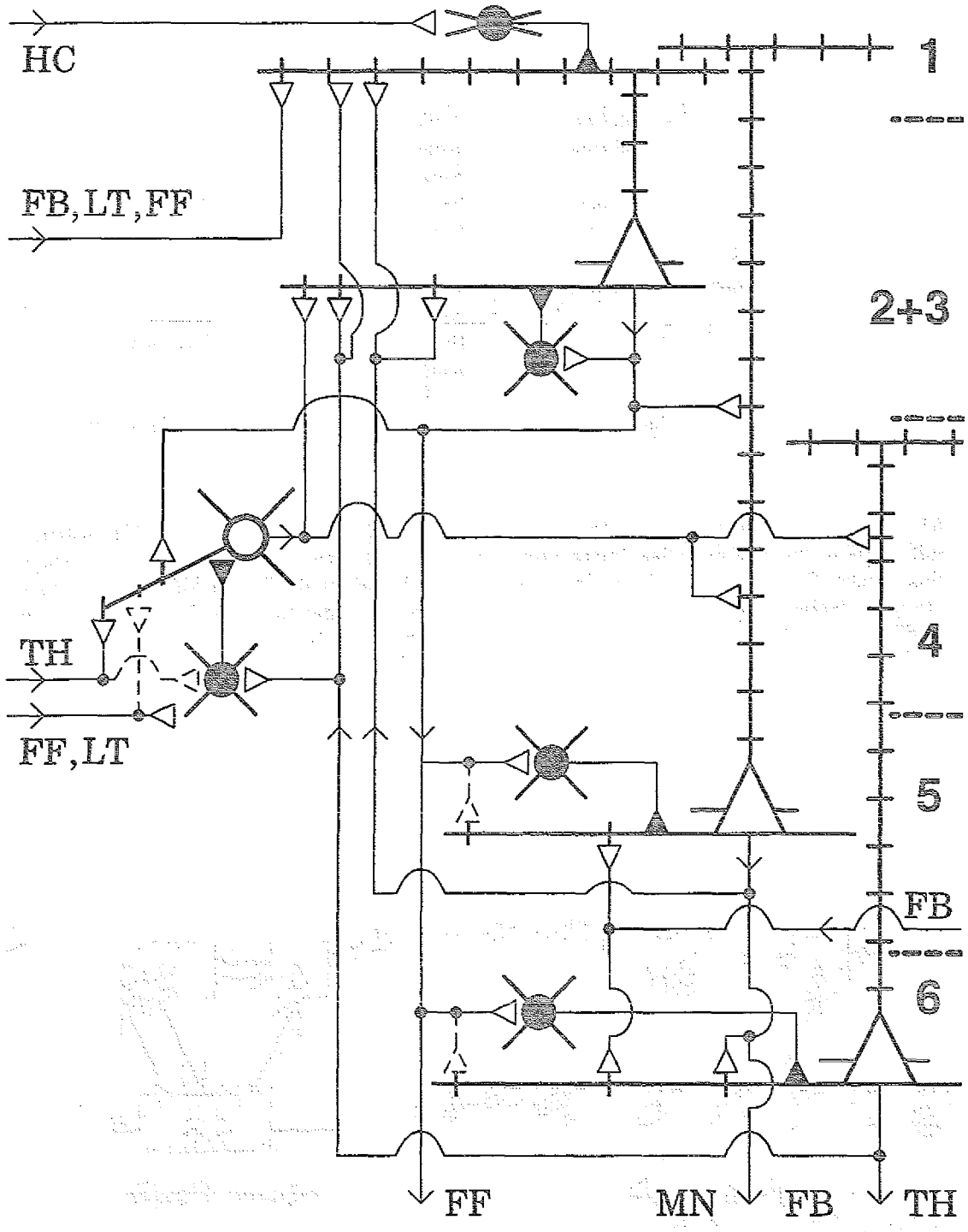


Abbildung D.5: Verbindungen der neocortikalen Schichten, die in diesem Modell wichtig sind. Es ist jeweils ein hemmendes Neuron und ein erregendes Neuron mit einem Basaldendriten stellvertretend für alle neuronalen Elemente einer Schicht eingezeichnet. Hemmende kontextuelle Interneuronen und Shunt–Synapsen sind schwarz ausgefüllt. Relativ seltenen synaptische Kontakte eines Axons sind durch eine gestrichelte Linie gekennzeichnet. Abkürzungen: TH: Thalamus; MN: Motorkerne; HC: Hippocampusformation; FF: Feedforward; FB: Feedback; LT: Laterale Verbindungen (von Arealen auf ähnlicher Hierarchiestufe).

..., Feedback, Feedforward, Lateral, 2+3, 5, 6	→	Apikal	1		
..., Hippocampusformation	→	Kontext			
..., 1, 2+3	→	Apikal	2+3	→	Feedforward, ...
..., 2+3, 4, 5, 6	→	Basal			
..., Lateral, 2+3	→	Kontext			
..., Thalamus, Lateral, Feedforward, 2+3, 4	→	Basal	4		
..., Lateral, Feedforward, 6	→	Kontext			
..., 1, 2+3, 4, 5	→	Apikal	5	→	Feedback, Subcortikal, ...
..., Feedback, 2+3, 5, 6	→	Basal			
..., Lateral, 2+3	→	Kontext			
..., 4, 5, 6	→	Apikal	6	→	Feedback, Thalamus, ...
..., Feedback, 2+3, 5, 6	→	Basal			
..., Lateral, 2+3	→	Kontext			

Abbildung D.6: Schaltkreis eines Areals in Tabellenform. Die Eingaben zu Apikaldendriten, Basaldendriten und kontextuellen Interneuronen sowie die Efferenzen der dornigen Neuronen sind eingetragen. Die Punkte und die Verbindungen, die zusätzlich zu denen von Abbildung 5 eingetragen sind, deuten an, daß eine scharfe Trennung zwischen wichtigen und unwichtigen Verbindungen in diesem Modell nicht existiert. Auch die Unterteilung in verschiedene Kontexte, insbesondere die Unterscheidung von Schicht 5 und Schicht 6, ist nicht die einzig mögliche.

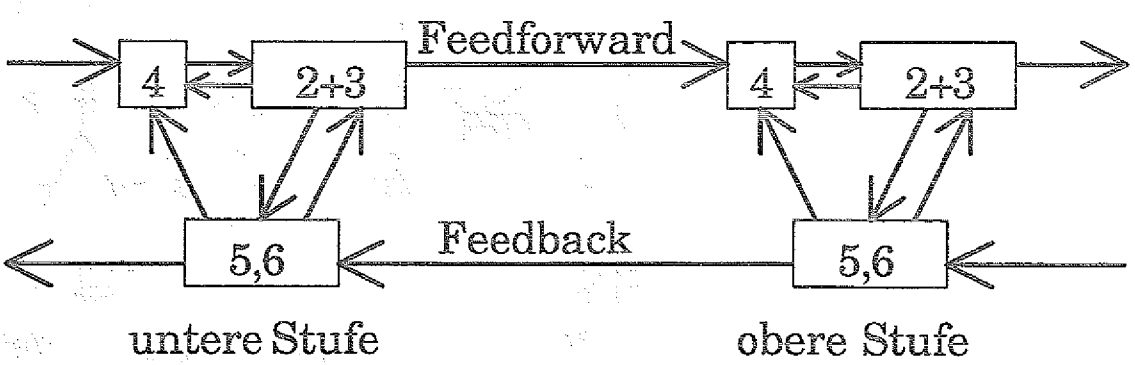


Abbildung D.7: Die wichtigsten intracortikale Verbindungen über Basaldendriten. Zwei Areale auf verschiedenen Stufen der neocortikalen Hierarchie sind über Feedforward– und Feedback Verbindungen verbunden. In diesem Modell überträt der Feedforward–Kanal exakte, aber veraltete Daten und der Feedback–Kanal unsichere Prognosen.

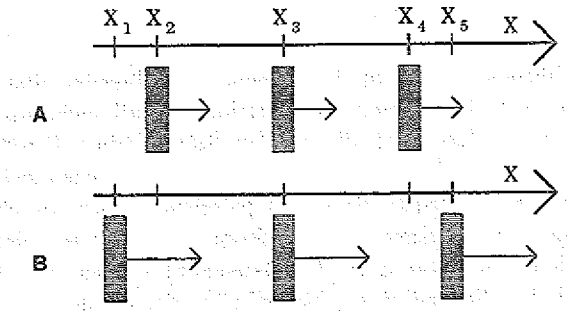


Abbildung D.8: Möglicher Test für den funktionellen Unterschied zwischen tiefen und oberflächlichen Schichten. Sensorische Reize (Balken) legen denselben Weg mit verschiedenen Geschwindigkeiten zurück. A: kleine Geschwindigkeit. B: große Geschwindigkeit. Die Position x des Reizes beim Beginn der Aktivitätsperiode eines Neurons kann von der Geschwindigkeit abhängen. Wenn sich die Reize bei den Positionen x_4 bzw.x_5 befinden, dann haben sie eine ähnliche Vergangenheit (Position x_3), aber eine unähnliche Zukunft. Im Rahmen des Modells sollten oberflächliche Neuronen auf solche Reize mit ähnlicher Vergangenheit reagieren. Dagegen sollten tiefe Neuronen ihre Aktivität schon dann starten, wenn die Reize eine ähnliche Zukunft haben (Positionen x_1, x_2).

D.4.3 Funktion der Hierarchie

Nach den in Kapitel D.2.1 angeführten Beispielen zeigen Neuronen auf höheren Hierarchiestufen normalerweise längere Aktivitätsperioden als solche auf niedrigen Hierarchiestufen.

In diesem Modell wurde dem Neocortex die grundlegende Aufgabe zugewiesen, längerfristige Erscheinungen in den flüchtigen Sinnesdaten zu erkennen (Kap. D.1). Dieser Übergang von kurzen zu langen Zeitskalen kann in einem Schritt geschehen, etwa durch die direkte Eingabe vom Thalamus zu multimodalen Assoziationsarealen, oder in mehreren Stufen. Die Signalverarbeitung eines cortikalen Areals wird wenig von Schädigungen anderer Areale beeinträchtigt [174]. Sie kann anscheinend recht gut in einem Schritt geschehen, sofern die Neuronen bereits gelernt haben, die zu derselben langfristigen Erscheinung gehörenden Sinneseindrücke zusammenzufasssen. Dieses können ganz verschiedene Muster sein, wie zum Beispiel die verschiedenen Ansichten desselben Gesichtes. Die Lernprozesse finden daher am besten in mehreren Stufen statt, bei denen jeweils aus etwas kürzerfristigen Erscheinungen etwas längerfristige abstrahiert werden. Die hierarchische Unterteilung des Neocortex in Areale wird damit erklärt, daß jeder dieser Lernstufen ein cortikales Areal entspricht.

Wie bereits in Kapitel D.2.6 argumentiert, müssen die tiefen Schichten 5 und 6 die nahe Zukunft vorhersagen, um motorische Bewegungen (wie z. B. Balancieren) steuern oder Sinnesdaten ankündigen zu können. Damit wird der Nutzen der Feedback–Verbindungen in diesem Modell klar: Längerfristige Vorhersagen auf höheren Stufen steuern kürzerfristige Vorhersagen auf niedrigeren Stufen. Durch die beiden entgegengesetzten Kanäle in Abbildung D.7 werden demnach Informationen übertragen, die sich nur allzu leicht verwechseln lassen: Die Feedforward–Verbindungen übertragen das, was das Netzwerk tatsächlich sieht, und die Feedback–Verbindungen das, was es erwartet beziehungsweise sehen will. Dieser Unterschied sollte im Prinzip meßbar sein. Wenn ein Neuron auf bewegte Stimuli verschiedener Geschwindigkeit reagiert, dann kann für alle die mittlere Position am Beginn der Aktivitätsperiode gemessen werden. Oberflächliche Neuronen sollten im Mittel erst reagieren, nachdem die Stimuli gleiche Position hatten, tiefe Neuronen aber bereits davor (siehe Abb. D.8). Solche Messungen sind im somatosensorischen Cortex vermutlich leichter als im visuellen Cortex, da die Position eines Reizes auf der Netzhaut nur schwer zu messen ist [136]. Bei Mutanten mit abnormaler Reihenfolge der Schichten [138] können vielleicht Folgen davon gemessen werden, daß Feedforward– und Feedback–Kanäle mangelhaft getrennt sind.

D.4.4 Bipolarzellen und REM–Schlaf

In dieses hierarchische Modell passen Spekulationen über Bipolarzellen und den Traumschlaf. Die Vielfalt der neocortikalen Funktionen läßt vermuten, daß nicht nur die mittlere Aktivität eines neocortikalen Areals, sondern auch die relative Stärke seiner Feedforward– und Feedback–Verbindungen der momentanen Aufgabe angepaßt wird. Diskrepanzen zwischen Feedforward– und Feedback–Signalen können zum Beispiel bei einem plötzlichen Szenenwechsel auftreten, nach dem ganz unerwartete Sinnesdaten erscheinen. Bei Diskrepanzen während der Verarbeitung sehr dynamischer, aber deutlicher Sinnesdaten (z. B. Fernsehen) sollten sich die Feedforward–Signale durchsetzen. Bei undeutlichen Sinnesdaten (Dunkelheit) oder bei der Visualisierung gar nicht vorhandener Objekte sind dagegen die Feedback–Verbindungen gefordert.

Die Bipolarzellen [140] wären geeignet, das Erkennen solcher Diskrepanzen zu beschleunigen. Manche dieser weitverbreiteten Neuronen (Ratte, Kaninchen, Hund, Affe, Mensch, ...) sind erregend und auch die hemmenden bilden Synapsen bevorzugt auf Dornen [97], so daß sie als kontextuelle Interneuronen ungeeignet sind. Sie sind charakterisiert durch einen Zellkörper, der in mittleren Schichten sitzt, und zwei, selten drei, nach oben und unten gerichtet glatte Dendritenbäume, die mehrere Schichten durchqueren. In Analogie zu anderen glatten Neuronen dieses Modells (Kap. D.3.3) würden sie lernen, auf korrelierte Signale zu reagieren. In Analogie zu Apikaldendriten (Kap. D.2.4) würden sie weder vom oberen noch vom unteren Dendriten alleine, sondern nur von der korrelierten Aktivität tiefer und oberflächlicher Schichten zusammen erregt werden, also nur dann, wenn die Sinnesdaten den Erwartungen entsprechen.

Variable Stärken der Feedforward– und Feedback–Verbindungen stellen neue Anforderungen an die Lernprozesse. Das Erlernen der richtigen Verbindungsstärken hängt in allen postulierten Lernprinzipien davon ab, daß das Neuron auch wirklich durch diese Verbindungen aktiviert wird. Andere aktive Verbindungen können stören. Es wäre möglich, daß besondere Lernphasen existieren, in denen solche Störungen ausgeschaltet sind. Die Lernphase der Feedback–Verbindungen ist dann ungestört, wenn trotz aktivem Cortex keine Sinnesdaten mehr verarbeitet werden. Dieser Zustand ähnelt Schlafphasen mit desynchronisiertem EEG, insbesondere den phasischen REM–Perioden, die alle Säugetiere, (während 15–25% ihrer Schlafzeit) allerdings auch Vögel (während ca. 0.5% ihrer Schlafzeit) zeigen [95]. Der bizarre Ablauf der Träume [128] könnte mit der spontanen Erzeugung der Feedback–Signale in Arealen auf hohen Hierarchiestufen zusammenhängen.

D.4.5 Kontext der Schicht 4

Das Lernen im Kontext wurde eingeführt, damit sich die Neuronen spezialisieren und nicht alle dieselben Erwartungen bilden. Da die Areale auf niedrigeren Stufen sich früher spezialisieren und am stärksten zur Schicht 4 höherer Stufen projizieren, ist es in jedem Areal die Schicht 4, die sich am frühesten spezialisieren kann. Und im primären und sekundären Cortex, in dem die Feedforward–Verbindungen besonders stark auffächern [101], in dem daher Spezialisierung besonders wichtig scheint, ist die Schicht 4 besonders breit.

Es wird daher hier angenommen, daß die Spezialisierung des Areals (z. B. auf Farbe, dynamische Form etc.) vor allem in Schicht 4 gelernt und von dort an die anderen Schichten übertragen wird. Diese Annahme paßt auch zu der Gestalt vieler dorniger Neuronen in Schicht 4 des primären sensorischen Cortex: Die Sternpyramiden haben nur einen relativ kleinen Apikaldendriten, und die häufigen dornigen Sternzellen haben gar keinen. Lernen im Kontext ist noch möglich, wenn ihre Dendriten wie Basaldendriten von Pyramidenzellen funktionieren. Aber das Abstrahieren langfristiger Erscheinungen aus den Feedforward–Signalen bleibt der Schicht 2+3 überlassen.

Der Übergang von einer kürzeren zu einer längeren Zeitskala entspricht dann in etwa dem Übergang von Schicht 4 zu Schicht 2+3. Messungen der Reaktionen einzelner dorniger Zellen im primären visuellen Cortex (Übersichten z. B. in [170, Kap.4] oder [137]) passen zu dieser Modellvorstellung: Schicht 4 besteht fast nur aus einfachen Zellen, die nur von Lichtreizen an bestimmten Stellen („ON–region") ihres rezeptiven Feldes, normalerweise also nur für kurze Zeit, erregt werden, wogegen die in Schicht 2+3 häufigen komplexen Zellen von Reizen überall in ihren größeren rezeptiven Feldern erregt werden. Nur auf den ersten Blick widersprechen dem Messungen

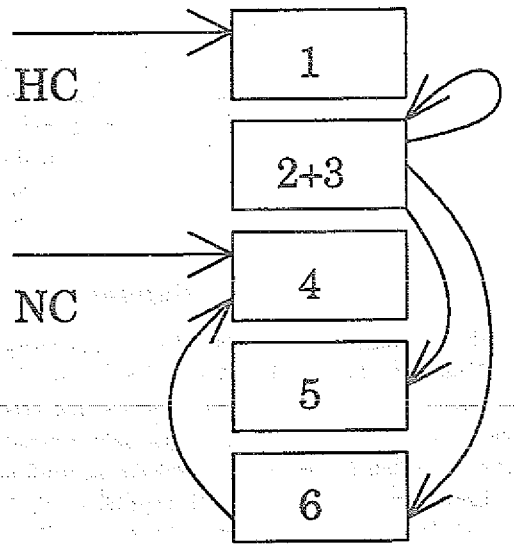


Abbildung D.9: Axone, die dem Modell zufolge kontextuelle Interneuronen relativ stärker erregen als dornige Dendriten. Nur für wenige dieser spezifischen Erregungen gibt es experimentelle Hinweise. Dies ist nur eine Übersicht. Welche spezifischen Erregungen nötig, welche möglich und welche unwahrscheinlich sind, beschreibt der Text im Detail. Abkürzungen: NC: Laterale und Feedforward Verbindungen (von tieferen oder gleichen Hierarchiestufen); HC: Hippocampusformation.

im primären somatosensorischen Cortex [161]. Dort finden sich die langsam adaptierenden Zellen, die auf statische Reize mit langen Aktivitätsperioden reagieren, vor allem in der Schicht 4. Solche statischen Reize erlauben aber wohl kaum, spätere Reize vorherzusagen. Nach den Lernprinzipien von Kapitel D.2.3 sollten Neuronen in Schicht 2+3 daher eher auf bewegte Reize reagieren, bei statischen Reizen dürfen sie schnell adaptieren.

Nach dem Prinzip von Kapitel D.3.4 sollen alle Signale, auf deren Repräsentation sich die Schicht 4 nicht spezialisieren soll, die kontextuellen Interneuronen dort relativ stärker erregen als die dornigen Dendriten. Das sind in diesem Modell Signale von vier Quellen (Abb. D.9). Erstens Signale von Arealen derselben Hierarchiestufe (Kap. D.3.4). Wenn die Schicht 4 wie im primären visuellen Cortex in spezialisierte Teilschichten unterteilt ist (4Ca+4B bzw. 4Cb+4A, siehe z. B. [175]), dann sollen zweitens diese Teilschichten gegenseitig ihre kontextuellen Interneuronen erregen.

Aber die Eigenschaften, die in verschiedenen Arealen repräsentiert werden, unterscheiden sich nicht nur durch physikalische Qualitäten, wie zum Beispiel Form oder Farbe. Manche Eigenschaften können nur auf langen und manche nur auf kurzen Zeitskalen erkannt werden. Wenn zum Beispiel ein bewegtes Objekt für kurze Zeit von einem Hindernis verdeckt wird, dann sind die Alternativen „Objekt hinter dem Hindernis“ und „Objekt vor dem Hindernis“ am besten von niedrigen Hierarchiestufen anhand der augenblicklichen Sinnesdaten zu unterscheiden. Die Alternativen „Objekt hinter dem Hindernis“ und „Kein Objekt hinter dem Hindernis“ dagegen sind nur von Neuronen auf hohen Hierarchiestufen zu unterscheiden, die über ausreichend lange Zeitperioden aktiv sind.

Jede Schicht 4 sollte sich auf eine ihrer Hierarchiestufe angemessene, mittlere Zeitskala spezialisieren. Das heißt, daß ihre kontextuellen Interneuronen die Eigenschaften kürzerer und längerer Zeitskalen signalisiert bekommen sollen. Sie sollen also drittens von neocortikalen Arealen tieferer Hierarchiestufen erregt werden. Der Thalamus würde nur dann als tiefste Hierarchiestufe gelten, falls er wie neocortikale Areale Erwartungen bilden würde. Direkte Signale von höheren neocortikalen Arealen zu den kontextuellen Interneuronen von Schicht 4 sind überflüssig, da die Vorhersagen der Sinnesdaten bereits von den tiefen Schichten zusammengefaßt werden. Insbesondere die zum Thalamus projizierenden Zellen der Schicht 6, aber vermutlich auch die in Schicht 5 sollen daher viertens die kontextuellen Interneuronen der Schicht 4 erregen.

Es gibt dagegen keinen Grund, warum die kontextuellen Interneuronen der Schicht 4 aus der Schicht 2+3 erregt werden sollten. Wenn aber die dornigen Neuronen der Schicht 4 auf diesem Weg erregt werden, dann können sie lernen, schon auf der kürzeren Zeitskala Eigenschaften zu repräsentieren, die erst auf der etwas längeren Zeitskala der Schicht 2+3 wichtig werden. Zum Beispiel könnte es in Schicht 4 Neuronen geben, die, auch wenn ihr rezeptives Feld von einem Sichthindernis verstellt ist, trotzdem auf das sich dahinter bewegende Objekt reagieren, weil es in den größeren rezeptiven Feldern von Schicht 2+3 sichtbar ist. Dies wurde bisher nicht gemessen.

D.4.6 Messen von spezifischen Verbindungen

Im Prinzip ist es meßbar, ob bestimmte eingefärbte Axone bevorzugt synaptische Kontakte mit dornigen oder mit glatten Neuronen bilden. Unter der Annahme, daß beide Synapsentypen von ähnlicher Stärke sind, kann man folgern, welche Ziele wie stark erregt werden. Aber schon das Zählen von Synapsen unter dem Elektronenmikroskop ist sehr zeitaufwendig. Und um den Anteil der Synapsen auf glatten, hemmenden Neuronen abzuschätzen, muß man außerdem zwischen Synapsen auf Dornen und Synapsen auf Schäften glatter Dendriten [171, 133, 169, 92, 132, 167, 85] oder zwischen Dendriten GABAergischer und nicht–GABAergischer Neuronen [121, 126] unterschieden. Bisher liegen vor allem Ergebnisse über den primären somatosensorischen Cortex Sm1 der Maus und den primären visuellen Cortex V1 der Katze vor. Sie wurden in [170] und auch in [86] zusammengefaßt.

Ein generelles Ergebnis folgt mit großer Gewißheit aus diesen Untersuchungen: Axone bevorzugen bestimmte Ziele, um Synapsen zu bilden. Die Wahrscheinlichkeit, daß eine Synapse gebildet wird, hängt nicht nur von dem Überlapp zwischen dem Gebiet der Axonkollaterale und dem Gebiet des Dendritenbaumes ab, sondern auch davon, wo das postsynaptische Neuron liegt und wohin es projiziert. Dagegen scheinen benachbarte Dendriten desselben Neurons oder benachbarte Pyramidalneuronen mit demselben Projektionsgebiet gleichwahrscheinliche Ziele zu sein [170]. Die Ziele werden im Neocortex daher etwa genauso fein unterschieden, wie es im Rahmen dieses Modells nötig ist, in dem Neuronen verschiedener Schichten und verschiedenden morphologischen Typs als Ziele unterschieden werden.

Eine der Vorhersagen aus Abb. D.9 wird vom Experiment teilweise bestätigt. Die Verbindung von Pyramidenzellen der Schicht 6 zu kontextuellen Interneuronen der Schicht 4 paßt zu den bisherigen Messungen: In Schicht 4 bildeten erregende Axonkollaterale im Durchschnitt etwa 30% ihrer Synapsen auf Dendritenschäften. Zwei Pyramidenzellen der Schicht 6 bildeten dagegen 72% ihrer 151 in Schicht 4 untersuchten Synpasen auf Dendritenschäften, von denen viele zu glatten Dendriten gehörten (Katze, V1 [133]). In Sm1 der Maus zeigen zum Thalamus projizierende Neuronen der oberen Schicht 6 und auch der tiefen Schicht 5 eine analoge, noch spezifischere Projektion zu Dendritenschäften der Schicht 4, allerdings auch der Schicht 5 [169]. Allerdings ist es ungewiß, ob dies wirklich Dendritenschäfte von hemmenden Neuronen oder nicht vielmehr Dendritenschäfte von dornigen Sternzellen sind. Denn dornige Sternzellen in Schicht 4 erhalten sehr viele Erregungen aus Schicht 6 über Synapsen, die auf Schäften sitzen [72].

Nur wenige andere Messungen zeigten eine Bevorzugung der Dendritenschäfte als Ziele der Axone: Man sah dies für Afferente vom Klaustrum [125] und — in einigen kleinen Studien — für Verbindungen in oberflächlichen Schichten: Von Schicht 4 nach 4 [157], von Schicht 3 nach 4 [168] und von Schicht 3 nach 2+3 [171]. Auf der anderen Seite zeigten viele ausführliche Messungen eine deutliche Bevorzugung der Dornen als Ziele der folgenden Axone: Kollaterale von Schicht 5b zu Schicht 4 (und auch zu Schicht 5 und 6) in V1 der Katze [104]; Afferenzen vom Thalamus zu Schicht 4 in V1 der Katze [106, 172, 125] und in Sm1 der Maus (Zählungen von Elhanay und White in [170]); extrinsische und intrinische Kollaterale von Neuronen, die durch den Balken projizieren, in Sm1 der Maus [167] und in V1 der Maus [85]; Kollaterale von Schicht 3 zu Schicht 3 und zu Schicht 5 in V1 der Katze [121] und in Sm1 der Maus [92] (aber weniger deutlich in V1 des Affen [132]); Feedforward Verbindungen aus dem Areal V1 der Katze zu dem posteromedialen lateralen suprasylvianen visuellen Areal [126]. Es stellt sich daher die Frage, wo all die Synapsen auf Dendritenschäften herkommen, die man im Neocortex sieht. Diesem Modell zufolge sollten insbesondere die lateralen Verbindungen von Arealen derselben Hierarchiestufen (und vielleicht auch entfernten

Teilen desselben Areals), die zu allen Schichten projizieren, dort relativ viele Synapsen auf den Dendritenschäften kontextueller Interneuronen bilden. Untersuchungen derartiger Verbindungen sind noch nicht bekannt.

D.4.7 Kontexte der Schichten 2+3, 5 und 6

Die oberflächliche Schicht 2+3 und die tiefen Schichten 5 und 6 lernen trotz ihrer ähnlichen neuronalen Zusammensetzung verschiedene Funktionen, weil sie verschiedene basale Eingaben erhalten (Abb. D.7) und weil sie verschiedene apikale Eingaben erhalten: Im Feedforward–Kanal, also in den oberflächlichen Schichten, sollen Apikaldendriten Reaktionen auf langfristige Erscheinungen belohnen (Kap. D.2.2). Im Feedback–Kanal, also in den tiefen Schichten, sollen Apikaldendriten korrekte Vorhersagen und erfolgreiche Steuerungen belohnen (Kap. D.2.6).

Wie früh sich ein Ereignis vorhersagen läßt, dürfte stark von der Art des Ereignisses abhängen. Optimal funktionierende tiefe Schichten würden daher manche Ereignisse langfristig und manche nur kurzfristig vorhersagen. Nach den Prinzipien von Kapitel D.2.3 werden insbesondere die langfristigen Vorhersagen belohnt. Durch Lernen im Kontext jedoch könnten gleichermaßen kurz– wie langfristige Vorhersagen gelernt werden.

Die tiefen Schichten erhalten laut Abbildung D.7 alle Signale, die für dieses Lernziel erforderlich sind. Die relativ zuverlässigen Informationen von Schicht 2+3 beschreiben die nahe Vergangenheit, die Informationen von höheren Hierarchiestufen stützen sich mehr auf Erfahrung, beschreiben aber neben der Vergangenheit auch die Zukunft. Die Information, die die Zukunft beschreibt, ist also die Information, die in den Signalen von Schicht 2+3 noch nicht enthalten ist. Nach dem Prinzip von Kapitel D.3 spezialisieren sich die tiefen Schichten dann auf die Repräsentation dieser Vorhersagen, falls ihre kontextuellen Interneuronen relativ stärker von Schicht 2+3 als von höheren Hierarchiestufen erregt werden (Abb. D.9).

Ein ähnlicher Lernprozess kann auch dann ablaufen, wenn die kontextuellen Interneuronen stärker durch Kollateralen von ihrer eigenen Schicht als von anderen Schichten oder Arealen erregt werden und wenn diese synaptische Erregung ein wenig verzögert ist. Denn dann repräsentiert der Kontext die Information, die der Schicht bisher bekannt war. Und die neu aktivierten dornigen Neuronen lernen die neuen Informationen zu repräsentieren. Dies mag auf Schicht 2+3 zutreffen (Abb. D.9). Denn dort sind die tendenziell nach unten gerichteten Basaldendriten in einer geeigneten Lage, stärker als die Dendriten multipolarer Interneuronen durch thalamocortikale Afferenzen erregt zu werden, da diese Afferenzen vor allem in Schicht 4 und der tiefen Schicht 3 enden.

Diese Modellannahmen passen allerdings nicht zu Messungen von Pyramidenzellen der Schicht 2+3 in V1 der Katze. Diese Pyramidenzellen bilden bevorzugt Synapsen auf Dornen, sowohl in der eigenen Schicht wie auch in den tiefen Schichten [121]. Ein ähnliches Bild ergibt sich für bestimmte Zelltypen in der Maus [92, 167, 85]; in V1 des Affen ist es weniger klar (Schäfte [171] und Dornen [132]). Daten aus nicht primärem sensorischen Cortex wären hier aufschlußreich. Weitere spezifische Verbindungen sind möglich, ohne notwendig aus dem Modell zu folgen: Wie stark erregen die tiefen Schichten 5 und 6 ihre eigenen kontextuellen Interneuronen? Wie viele verschiedene Teilschichten besitzen einen eigenen Kontext? Wie wirken die anderen Teile des Gehirns, zu denen die tiefen Schichten projizieren, auf deren kontextuelle Interneuronen zurück?

Noch einen Vorteil hat ein Kontext, der aus den Signalen benachbarter, schon längere Zeit aktiver Neuronen gebildet wird. Laut Kap. D.3.4. lernen die Neuronen, die den Kontext signalisieren, und die in diesem Kontext neu aktivierten Neuronen, so zu kooperieren, daß die Sinnesdaten genau repräsentiert werden. In diesem Fall kooperieren also alle aktiven dornigen Neuronen einer Schicht. Diese genaue Repräsentation der Sinnesdaten eignet sich als Eingabe für andere kortikale Areale, und tatsächlich starten die meisten corticocortikalen Verbindungen von Schicht 2+3.

Wenn die Neuronen einer neocortikalen Schicht tatsächlich auf diese Weise kooperieren, dann bestehen enge Analogien zu den mathematischen Grundlagen der Symbolischen Dynamik (siehe auch Kap. D.1) [33, 36, 70]: In dieser Theorie werden dynamische Prozesse durch Sequenzen von Symbolen beschrieben, wie hier durch die aufeinanderfolgenden Aktivitätsperioden von Neuronen. Dort werden die Symbole von vornherein so gewählt, daß Symbole aller Zeiten zu einer exakten Beschreibung zusammenwirken („generierende Partition"). Hier wurden Prinzipien postuliert,

nach denen Neuronen lernen, zu einer möglichst genauen Beschreibung zusammenzuwirken, auch wenn ihre Aktivitätsperioden nicht gleichzeitig beginnen. Und solche vagen Konzepte wie „Informationen über Zukunft bzw. Vergangenheit" können in der Mathematik durch die „stabilen bzw. instabilen Mannigfaltigkeiten" präzisiert werden.

D.4.8 Kontext der Schicht 1

Schicht 1 enthält nur wenige Neuronen. Zum größten Teil besteht sie aus Axonen diversen Ursprungs und den oberflächlichen Verzweigungen der Apikaldendriten von tiefer liegenden Pyramidenzellen. Trotzdem wird die einfache Funktion beim operanten Konditionieren (Kap. D.2.7) der Komplexität dieser Schicht nicht gerecht. Zum Beispiel haben große Pyramidenzellen der Schicht 5 etwa 10% ihrer Dornen in der Schicht 1 [123], so daß diese apikalen Zweige wie ein zweites, den basalen Dendritenzweigen ähnelndes Dendritensystem erscheinen (siehe [166] für Analogien und [122] für Unterschiede). Und von den Neuronen der Schicht 1 sind 95–100% GABAergisch [113, 117] und können deshalb die apikalen Zweige shunten.

Wenn nun angenommen wird, daß in Analogie zu den anderen Schichten auch in Schicht 1 ein Kontext repräsentiert wird, dann wird das operante Lernen flexibler. In jedem Kontext lernen nur wenige Zweige der Apikaldendriten, während die anderen Zweige geshuntet sind (AS2 in Abb. D.2). In verschiedenen Kontexten der Schicht 1 können verschiedene Verhaltensweisen gelernt werden. Der Teil des Gehirns, der den Kontext der Schicht 1 steuert, müßte also Situationen miteinander assoziieren, in denen dasselbe Verhalten ähnliche Folgen erwarten läßt, und von Situationen unterscheiden, in denen die Folgen anders sind. Diese Aufgabe erscheint lösbar: Assoziieren und Kategorisieren relativ statischer Muster wird der Hippocampusformation zugetraut und auch numerisch in neuronalen Netzen simuliert. Für das hochentwickelte Modell der Adaptiven Resonanz [82, 112] sprechen Vergleiche mit den Auswirkungen von Schädigungen des Hippocampus. Zum Beispiel die Durchtrennung des Fornix, des wichtigsten Bündel von Ausgabefasern, behindert die Fähigkeit, ein gewohntes Verhalten zu ändern, wenn veränderte Umstände auftreten, die den gewohnten Umständen ähneln und deshalb leicht mit ihnen verwechselt werden können [105]. Weitere Hinweise können hier nur skizziert werden: Dem entorhinalen Cortex wird von Feedforward–Verbindungen des Neocortex [101] die gegenwärtige Situation signalisiert, der eigentliche Hippocampus könnte lernen, von verschiedenen Situationen auf ähnliche Weise erregt zu werden („Autoassoziation", siehe z. B. [147]). Am Ende dieses Schaltkreises steht wieder der entorhinale Cortex, der im Neocortex vor allem zu Schicht 1 projiziert [173, Kap. 4.1.1.]. Außerdem ist die Hippocampusformation in der Angsttheorie von Gray [111] die zentrale Instanz des Gehirns, die feststellt, ob Erwartungen erfüllt oder verletzt werden.

Nun, da allen Schichten des Modells eine Funktion zugewiesen wurde, wird ihre Reihenfolge verständlich. Die Schichten 2 und 3, sowie 5 und 6, sollten wegen ihrer ähnlichen Funktion benachbart sein. Von der Schicht 1 sollten bei operantem Verhalten Apikaldendriten aus Schicht 2, 3 und 5 de– oder hyperpolarisiert werden. Die Apikaldendriten aus Schicht 5 sollten außerdem von den Schichten 2+3 und 4, die aus Schicht 6 von Schicht 4 kontaktiert werden (Kap. D.2.7). Schicht 4 erregt die Basaldendriten von Schicht 2+3 (Kap. D.4.2). Wenn außerdem noch gefordert wird, daß die Verbindungen über Apikal– oder Basaldendriten möglichst kurz sein sollen, dann ergibt sich als beste Lösung die tatsächliche Reihenfolge im Neocortex: 1, 2+3, 4, 5 und 6.

D.5 Diskussion

Das vorgestellte Modell gibt zwar keine zwingenden, aber doch plausible Antworten auf eine lange Liste offener Fragen: Warum sitzen hemmende Synapsen auf den Schäften dorniger Dendriten (Kap. D.3.2)? Wie unterscheiden sich Lernprozesse in dornigen und glatten Dendriten (Kap. D.3.3)? Warum vereinigen sich die apikalen Dendritenzweige einer Pyramidenzelle zu einem einzigen Apikaldendriten (Kap. D.2.4)? Welche Funktionen unterscheiden die Schichten 1, 2+3, 4, 5 und 6 (Kap. D.4)? Weshalb sind die Schichten gerade in dieser Weise angeordnet (Kap. D.4.8) und durch Dendriten und Axonkollaterale verknüpft (Kap. D.4.3)? Warum brauchen in Schicht

1 keine dornigen Neuronen zu sein (Kap. D.4.8)? Weshalb besitzen die dornigen Neuronen der Schicht 4 häufig kein Apikaldendrit (Kap. D.4.5)? Warum projizieren nur die dornigen Neuronen zu anderen Arealen (Kap. D.3.3)? Wieso sind kortikale Areale hierarchisch angeordnet und wieso bevorzugen die Feedforward– und Feedback–Verbindungen bestimmte Schichten (Kap. D.4.3)? Warum bestehen reziproke Verbindungen mit Thalamus (Kap. D.2.6) und Hippocampusformation (Kap. D.4.8)? Eignet sich der Schlaf für Lernprozesse (Kap. D.4.4)? Es werden also Erklärungen für gerade die Eigenschaften des Neocortex angeboten, die den meisten Säugetierarten und neocortikalen Arealen gemeinsam sind.

Das vorgestellte Modell ist aber auch höchst unvollständig. Es wird nur eine Möglichkeit aufgezeigt, wie trotz weitgehend zufälliger Verknüpfungen innerhalb einer Schicht präzise Analyse von Daten erlernt werden kann. Aus dieser Aufgabenstellung konnte nur die prinzipielle Richtung der Lernprozesse abgeleitet werden, jedoch keine quantitativen Lernregeln. Denn viele verschiedenen Lernregeln können ähnliche Funktionen erfüllen. Das hier zumeist diskutierte Hebbsche Lernen, bei dem auf aktiven Neuronen die aktiven depolarisierenden Synapsen gestärkt werden, ähnelt in seiner Wirkung zum Beispiel den verallgemeinerten Hebbschen Lernregeln, bei denen passive depolarisierende Synapsen geschwächt, passive hyperpolarisierende Synapsen gestärkt oder aktive hyperpolarisierende Synapsen geschwächt werden. Ebenso gibt es mehrere, hier gar nicht diskutierte, Möglichkeiten, wie eine mittlere Aktivität in Dendriten und Neuronen gehalten und ewige Ruhe oder epileptische Ausbrüche vermieden werden können. Auch bleibt offen, wie lange und wie stark sich die Synapsen ändern. Ohne Simulation mit quantitativen Lernregeln kann es letztlich auch keine Gewißheit geben, daß das Modell die geforderte Funktion überhaupt erfüllen kann. Aber man kann dem Modell eine gewisse Robustheit zutrauen, weil in ihm verschiedene Synapsen sich gegenseitig in ihrer Funktion bestärken. So lernen Dendriten, nur dann geshuntet zu sein, wenn sie nicht stark erregt werden, und umgekehrt. Auch lernen Apikaldendrit und Basaldendriten auf denselben Reiz, nur zu verschiedenen Zeiten zu reagieren.

Physiologische Experimente werden bloß qualitative Lernprinzipien nur schwer testen können. Aber sie können vielleicht die Größe der Schwellenwerte, Zeitkonstanten und der anderen offen gelassenen Lernparameter einschränken. Die Kenntnis der Reize, auf die einzelne Neuronen reagieren, hilft wenig, wenn diese Reaktionen aus Erfahrung gelernt werden. Nur nachdem realistische Daten nach den postulierten Lernprinzipien in aufwendigen numerischen Simulationen verarbeitet wurden, wäre ein Vergleich möglich. Der postulierte Unterschied zwischen oberflächlichen Schichten (zuverlässige Sinnesdaten) und tiefen Schichten (Vorhersagen) läßt sich aber vielleicht messen (Kap. D.4.3).

Da sich das Modell bisher vor allem auf anatomische Fakten stützt, läßt es sich wohl auch am besten durch anatomische Untersuchungen testen. Wenn man in der Anordnung der Synapsen Gründe fände, daß die Erregung des Somas durch den Apikaldendriten nicht beschränkt werden kann oder daß die Basaldendriten nicht durch Shunt–Synpasen vor dem Aktionspotential des Somas geschützt werden können, dann wären die in Kapitel D.2 beziehungsweise Kapitel D.3 postulierten Lernprozesse ausgeschlossen. Die Vorhersagen darüber, mit welchen Dendriten bestimmte Axone bevorzugt synaptische Kontakte bilden (Abb. D.9), sollten sich durch Messungen noch relativ leicht überprüfen lassen (Kap. D.4.6). Nicht nur einzelne Schichten können bevorzugte Ziele sein, sondern auch Basaldendriten, Apikaldendriten und insbesondere die glatten Dendriten kontextueller Interneuronen: Über diese macht das Modell besonders charakteristische Aussagen.

Abbildungsverzeichnis

Index

Index von Rechenzeichen

Literaturverzeichnis

[1] Alekseev, V. M.; Yakobson, M. V. (1981). Symbolic Dynamics and Hyperbolic Dynamic Systems. *Phys. Rep.* 75, S. 287–325.

[2] Arnold,V. I.; Avez, A. (1986). *Ergodic Problems of Classical Mechanics.* Addison Wesley, New York.

[3] Artuso, R.; Aurell, E.; Cvitanovic, P. (1990). Recycling of Strange Sets: 1. Cycle Expansions. *Nonlinearity* 3, S. 325–359.

[4] Artuso, R.; Aurell, E.; Cvitanovic, P. (1990). Recycling of Strange Sets: 2. Applications. *Nonlinearity* 3, S. 361–386.

[5] Badii, R. (1990). Complexity as Unpredictability of the Scaling Dynamics. *Europhys. Lett.* 13, S. 599-604.

[6] Badii, R.; Finardi, M.; Broggi, G.; Sepulveda, M. A. (1992). Hierarchical Resolution of Power Spectra. *Physica* 58 D, S. 304–324.

[7] Benedicks, M.; Carleson, L. (1991). The Dynamics of the Henon map. *Ann. Math.* 133, S. 73–169.

[8] Benedicks, M.; Young, L. S. (1993). Sinai–Bowen–Ruelle Measures for Certain Henon Maps. *Inven. Math.* 112, S. 541–576.

[9] Biham, O.; Wenzel, W. (1989). Characterization of Unstable Periodic Orbits in Chaotic Attractors and Repellers. *Phys. Rev. Lett.* 63, S. 819–822.

[10] Birkhoff, G. D. (1927). *Dynamical Systems.* A. M. S. Publications, Providence, reprinted 1966.

[11] Bowen, R. (1973). Symbolic Dynamics for Hyperbolic Flows. *Amer. J. Math.* 95, S. 429–459.

[12] Bowen, R. (1978). *Proc. Amer. Math. Soc.* 71, S. 130.

[13] Collet, P.; Eckmann, J. P. (1980). *Iterated Maps on the Interval as Dynamical Systems.* Birkhäuser, Boston.

[14] Curry, J. H. (1979). On the Hénon Transformation. *Commun. Math. Phys.* 68, S. 129–140.

[15] Cvitanović, P.; Gunaratne, G. H.; Procaccia, I. (1988). Topological and Metric Properties of Henon–Type Strange Attractors. *Phys. Rev. A* 38, S. 1503–1520.

[16] D'Alessandro, G.; Politi, A. (1990). Hierarchical Approach to Complexity with Applications to Dynamical Systems. *Phys. Rev. Lett.* 64, S. 1609–1612.

[17] D'Alessandro, G.; Grassberger, P.; Isola, S.; Politi, A. (1990). On the Topology of the Henon Map. *J. Phys. A* 23, S. 5285–5294.

[18] de Melo, W.; van Strien, S. (1993). *One–Dimensional Dynamics.* Springer, Berlin.

[19] Derrida, B.; Gervois, A.; Pomeau, Y. (1979). Iteration of Endomorphisms on the Real Axis and Representation of Numbers. *Ann. Inst. Henri Poincaré* 29 A, S. 305.

[20] Eckmann, J.-P.; Ruelle, D. (1985). Ergodic Theory of Chaos and Strange Attractors. *Reviews of Modern Physics* 57, S. 617–656 und S. 1115.

[21] Feit, S. D. (1978). Characteristic Exponents and Strange Attractors. *Commun. math. Phys.* 61, S. 249–260.

[22] Finardi, M.; Badii, R.; Flepp, L.; Parisi, J.; Reyl, C.; Holzner, R.; Brun, E. (1992). Symbolic Ananlysis of a Heteroclinic Crisis in the NMR–Laser. *Helvetica Physica Acta* 65, S. 826–827.

[23] Flepp, L.; Holzner, R.; Brun, E.; Finardi, M.; Badii, R. (1991). Model Identification by Periodic–Orbit Analysis for NMR–Laser Chaos. *Phys. Rev. Lett.* 67, S. 2244–2247.

[24] Fornaess, J. E.; Gavosto, E. A. (1991). Existence d'Attracteurs Étranges pour Certaines Applications de Hénon. *C. R. Acad. Sci. Paris* 313, S. 495–498.

[25] Gallas, J. A. C. (1993). Structure of the Parameter Space of the Henon Map. *Phys. Rev. Lett.* 70, S. 2714–2717.

[26] Giovannini, F.; Politi, A. (1991). Homoclinic Tangencies, Generating Partitions and Curvature of Invariant Manifolds. *J. Phys. A.* 24, S. 1837–1887.

[27] Giovannini, F.; Politi, A. (1992). Generating Partitions in Henon–Type Maps. *Phys. Lett.* 161 A, S. 332–336.

[28] Gonchenko, S. V.; Turaev, D. V.; Shilnikov, S. P. (1993). On models with non-rough Poincaré homoclinic curve. *Physica* D 62, S. 1–14.

[29] Grassberger, P.; Badii, R.; Politi, A. (1988). Scaling Laws for Invariant Measures on Hyperbolic and Nonhyperbolic Attractors. *J. Stat. Phys.* 51, S. 135–178.

[30] Grassberger, P.; Kantz, H. (1985). Generating Partitions for the Dissipative Henon Map. *Phys. Lett.* 113 A, S. 235–238.

[31] Grassberger, P.; Kantz, H.; Mönig, U. (1989). On the Symbolic Dynamics of the Henon Map. *J. Phys. A* 22, S. 5217–5230.

[32] Grebogi, C.; Ott, E.; Yorke, J. A. (1987). Basin Boundary Metamorphoses: Changes in Accessible Boundary Orbits. *Physica* 24D, S. 243–262.

[33] Guckenheimer, J.; Holmes, P. (1983). *Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields.* Springer, Berlin.

[34] Hadamard, J. (1901). Sur l'itération et les solutions asymptotiques des équations différentielles. *Bull. Soc. Math. France* 29, S. 224–228.

[35] Hansen, K. T. (1992). Remarks on the Symbolic Dynamics for the Henon Map. *Phys. Lett.* 165 A, S. 100–104.

[36] Hao B.–L. (1989). *Elementary Symbolic Dynamics and Chaos in Dissipative Systems.* World Scientific, Singapur.

[37] Hao B.–L. (1991). Symbolic Dynamics and Characterization of Complexity. *Physica D* 51, S. 161–176.

[38] Hao B.–L.; Zheng W.–M. (1989). Symbolic Dynamics of Unimodal Maps Revisited. *Int. J. Mod. Phys.* B3, S. 235.

[39] Heagy, J. F. (1992). A Physical Interpretation of the Henon Map. *Physica* 57 D, S. 436–446.

[40] Hénon, M. (1976). A Two–Dimensional Mapping with a Strange Attractor. *Commun. math. Phys.* 50, S. 69–77.

[41] Hobson, D. (1993). An Efficient Method for Computing Invariant–Manifolds of Planar Map. *J. Comp. Phys.* 104, S. 14–22.

[42] Holmes, P.; Whitley, D. (1984). Bifurcations of One– and Two–Dimensional Maps. *Phil. Trans. R. Soc. Lond. A* 311, S. 43–102.

[43] Hopcroft, J. E.; Ullman, J. D. (1979). *Introduction to Automata Theory, Languages and Computation.* Addison–Wesley, Reading MA.

[44] Ikeda, K. (1979). Multiple–valued stationary state and its instability of the transmitted light by a ring cavity system. *Opt. Commun.* 30, S. 257.

[45] Jackson, E. A. (1990). *Perspectives of Nonlinear Dynamics.* Cambridge University Press, Cambridge.

[46] Krieger, W. (1970). On Entropy and Generators of Measure–Preserving Transformations. *Trans. Amer. Math. Soc.* 149, S. 453–464, siehe auch 168, S. 519.

[47] Lai Y.–C.; Grebogi, C.; Yorke, J. A. (1993). How Often are Chaotic Saddles Nonhyperbolic? *Nonlinearity* 6, S. 779–798.

[48] Levinson, N. (1950). Small Periodic Perturbations of an Autonomous System with a Stable Orbit. *Ann. Math.* 52, S. 727–738.

[49] Meiss, J. D. (1992). Symplectic Maps, Variational Principles, and Transport. *Rev. Mod. Phys.* 64, S. 795–848.

[50] Metropolis, N.; Stein, M. L.; Stein, P. R. (1973). On Finite Limit Sets for Transformations on the Unit Interval. *J. Comb. Theor.* 15, S. 25–44.

[51] Milnor, J.; Thurston, W., Preprint (1977). In: On Iterated Maps of the Interval. *Lecture Notes in Mathematics Vol. 1342*, S. 465, Springer, Berlin.

[52] Mindlin, G. B.; Gilmore, R. (1992). Topological Analysis and Synthesis of Chaotic Time Series. *Physica D* 58, S. 229–242.

[53] Mira, C. (1987). *Chaotic Dynamics.* World Scientific, Singapur.

[54] Moore, C. (1993). Braids in Classical Dynamics. *Phys. Rev. Lett.* 70, S. 3675–3679.

[55] Morse, M. (1921). *Trans. Am. Math. Soc.* 22, S. 84.

[56] Morse, M.; Hedlund, G. A. (1938). *Am. J. Math.* 60, S. 815, nachgedruckt in *Collected Papers* von M. Morse, Vol. 2, World Scientific, Singapur, 1986.

[57] Oseledec, V. I. (1968). A Multiplicative Ergodic Theorem. Ljapunov Characteristic Numbers for Dynamical Systems. *Trans. Moscow Math. Soc.* 19, S. 197–231.

[58] Ott, E.; Grebogi, C.; Yorke, J. A. (1989). Theory of First Order Phase Transitions For Chaotic Attractors of Nonlinear Dynamical Systems. *Phys. Lett. A* 135, S. 343–348.

[59] Politi, A.; Badii, R.; Grassberger, P. (1988). On the Geometric Structure of Non–Hyperbolic Attractors. *J. Phys. A* 21, S. L763–L769.

[60] Prouhet, E. (1851). *C. R. Acad. Sci.* 33, S. 225.

[61] Roessler, O. E. (1983). The Chaotic Hierarchy. *Z. Naturforsch.* 38 a, S. 788–801.

[62] Ruelle, D. (1979). Ergodic Theory of Differentiable Dynamical Systems. *Publ. IHES* 50, S. 27.

[63] Sarkovskij, A. N. (1964). Coexistence of Cycles of a Continuous Map of a Line into Itself. *Ukranian Math. J.* 16, S. 61 (in Russisch).

[64] Shaw, R. (1981). Strange Attractors, Chaotic Behavior and Information Flow. *Z. Naturforsch.* 36 a, S. 80–112.

[65] Simó, C. (1979). On the Hénon–Pomeau Attractor. *J. Stat. Phys.* 21, S. 465–494.

[66] Smale, S. (1963). Diffeomorphisms with Many Periodic Points. In *Differential and Combinatorial Topology*, S. S. Cairns (Ed.), S. 63–80, Princeton University Press, Princeton.

[67] Wiggins, S. (1988). *Global Bifurcations and Chaos: Analytical Methods.* Springer, Berlin.

[68] Zhao H.; Zheng W.–M.; Gu Y. (1992). Determination of Partition Lines from Dynamic Foliations for the Henon Map. *Commun. Theor. Phys.* 17, S. 263–268.

[69] Zhao H.; Zheng W.–M. (1993). Symbolic Analysis of the Henon map at a = 1.4 and b = 0.3. *Commun. Theor. Phys.* 19, S. 21–26.

[70] Zheng W.–M.; Hao B.–L. (1990). Applied Symbolic Dynamics. In *Experimental Study and Characterization of Chaos*, S. 363–459 in *Directions in Chaos, Vol 3.* Hao B.–L. (Ed.), World Scientific, Singapur.

Neurobiologisches Literaturverzeichnis

[71] Abeles, M. (1991). *Corticonics: Neural Circuits of the Cerebral Cortex.* Cambridge University Press, Cambridge.

[72] Ahmed, B., Anderson, J. C., Douglas, R. J., Martin, K. A. C., Nelson, J. C. (1994). *Polyneuronal Innervation of Spiny Stellate Neurons in Cat Visual Cortex. J. Comp. Neurol.* **341**, pp. 39–49.

[73] Amit, D. J. (1989). *Modeling Brain Function: The World of Attractor Neural Networks.* Cambridge University Press, Cambridge.

[74] Anderson, J. C., Douglas, R. J., Martin, K. A. C., Nelson, J. C. (1993). Map of the Synapses Formed With the Dendrites of Spiny Stellate Neurons of Cat Visual Cortex. *J. Comp. Neurol.* **341**, pp. 25–38.

[75] Anderson, J. C.; Martin, K. A. C.; Whitteridge, D. (1993). Form, Function and Intracortical Projections of Neurons in the Striate Cortex of the Monkey Macacus Nemestrinus. *Cerebral Cortex* **3**, pp. 412–420.

[76] Ballard, D. H. (1986). Cortical Connections and Parallel Processing: Structure and Function. *Behavioural and Brain Science* **9**, pp. 67–120.

[77] Barlow, H. B. (1972). Single Units and Sensation: A Neuron Doctrine for Perceptual Psychology. *Perception* **1**, pp. 371–394.

[78] Beaulieu, C.; Kisvarday, Z.; Somogyi, P.; Cyander, M.; Cowey, A. (1992). Quantitative Distribution of GABA–immunopositive and –immunonegative Neurons and Synapses in the Monkey Striate Cortex (Area 17). *Cereb. Cortex* **2**, pp. 295–309.

[79] Benardo, L. S. (1994). Separate Activation of Fast and Slow Inhibitory Postsynaptic Potentials in Rat Neocortex in Vitro. *J. Physiol.* **476**, pp. 203–215.

[80] Braitenberg, V., Schüz, A. (1991). *Anatomy of the Cortex: Statistics and Geometry.* Springer, Berlin.

[81] Bush, P. C.; Sejnowski, T. J. (1994). Effects of Inhibition and Dendritic Saturation in Simulated Neocortical Pyramidal Cells. *J. Neurosci.* **71**, pp. 2183–2193.

[82] Carpenter, G. A.; Grossberg, S. (1993). Normal and Amnesic Learning, Recognition and Memory by a Neural Model of Cortico–Hippocampal Interactions. *Trends in Neurosciences* **16**, No. 4, pp. 131–137.

[83] Connors, B. W. (1992). GABA-A-Mediated and GABA-B-Mediated Processes in Visual Cortex. *Progr. in Brain Res.* **90**, pp. 335–348.

[84] Crick, F. H. C.; Asanuma, C. (1986). Certain Aspects of the Anatomy and Physiology of the Cerebral Cortex. In [148], Vol. 2, pp. 333–371.

[85] Czeiger, D.; White, E. L. (1993). Synapses of Extrinsic and Intrinsic Origin Made by Callosal Projection Neurons in Mouse Visual Cortex. *J. Comp. Neurol.* **330**, pp. 502–513.

[86] DeFelipe, J.; Farinas, I. (1992). The Pyramidal Neuron of the Cerebral Cortex: Morphological and Chemical Characteristics of the Synaptic Inputs. *Progr. Neurobiol.* **39**, pp. 563–607.

[87] Desimone, R. (1991). Face–Selective Cells in the Temporal Cortex of Monkeys. *J. Cog. Neruosci.* **3** pp. 1–8.

[88] DeYoe, E. A.; Van Essen, D. C. (1988). Concurrent Processing Streams in Monkey Visual Cortex. Trends in Neurosciences **11**, pp. 219–226.

[89] Douglas, R. J.; Martin, K. A. C. (1990). Neocortex. In [153], pp. 389–438.

[90] Douglas, R. J.; Martin, K. A. C. (1991). Opening the Grey Box. *Trends in Neurosciences* **14**, No. 7, pp. 287–293.

[91] Durbin, R.; Miall, C.; Mitchison, G. (Eds.) (1989). *The Computing Neuron.* Addison-Wesley, New York.

[92] Elhanany, E.; White, E. L. (1990). Intrinsic Circuity: Synapses Involving the Local Axon Collaterals of Corticocortical Projection Neurons in the Mouse Primary Somatosensory Cortex. *J. Comp. Neurol.* **291**, pp. 43–54.

[93] Emerson, R. C.; Citron, M. C.; Vaughn, W. J.; Klein, S. A. (1987). Nonlinear Directionally Selective Subunits in Complex Cells of Cat Striate Cortex. *J. Neurophys.* Vol. 58, pp. 33–65.

[94] Engel, A. K.; König, P.; Kreiter, A. K.; Schillen, T. B.; Singer, W. (1992). Temporal Coding in the Visual Cortex: New Vistas on Intergration in the Nervous System. *Trends in Neurosciences* **15**, No. 6, pp. 218–226.

[95] Erwin, C. W.; Somerville, E. R.; Radtke, R. A. (1984). A Review of Electroencephalographic Features of Normal Sleep. *J. Clin. Neurophys.* **1**, pp. 253–274.

[96] Fairén, A.; DeFelipe, J.; Regidov, J. (1984). Nonpyramidal Neurons. In [142], Vol. 1, pp. 201–254.

[97] Fairén, A.; Smith–Fernández, A. (1992). Electron Microscopy of Golgi–Impregnated Interneurons: Notes on the Intrinsic Connectivity of the Cerebral Cortex. *Micr. Res. Tech.* **23**, pp. 289–305.

[98] Feldman, M. L. (1984). Morphology of the Neocortical Pyramidal Neuron. In [142], Vol. 1, pp. 123–200.

[99] Feldman, J. A. (1990). Computational Constraints on Higher Neural Representations. In [149], pp. 163–178.

[100] Feldman, J. A.; Ballard, D. H. (1982). Connectionist Models and their Properties. *Cogn. Sci.* **6**, pp. 205–254.

[101] Felleman, D. J.; Van Essen, D. C. (1991). Distributed Hierarchical Processing in the Primate Cerebral Cortex. *Cereb Cortex* **1**, pp. 1–47.

[102] Fuster, J. M. (1984). Behavioural Electrophysiology of the Prefrontal Cortex. *Trends in Neurosciences* **7**, pp. 408–414.

[103] Fuster, J. M.; Alexander, G. E. (1971). Neuron Activity Related to Short–Term Memory. *Science* **173**, pp. 652–654.

[104] Gabbott, P. L. A.; Martin, K. A. C.; Whitteridge, D. (1987). Connections between Pyramidal Neurons in Layer 5 of Cat Visual Cortex (Area 17). *J. Comp. Neurol.* **259**, pp. 364–381.

[105] Gaffan, D. (1985). Hippocampus: Memory, Habit and Voluntary Movement. *Philos. Trans. Roy. Soc. Lond. Ser. B* **308**, pp. 87–99.

[106] Garey, L. J.; Powell, T. P. S. (1971). An Experimental Study of the Termination of the Lateral Geniculo–Cortical Pathway in the Cat and Monkey. *Proc. Roy. Soc. Lond. Ser. B* **179**, pp. 21–40.

[107] Ghosh, S.; Fyffe, R. E. W.; Porter, R. (1988). Morphology of Neurons in Area 4gamma of the Cat's Cortex Studied with Intracellular Injection of HRP. *J. Comp. Neurol.* **269**, S. 290–312.

[108] Gilbert, C. D. (1993). Circuity, Architecture, and Functional Dynamics of Visual Cortex. *Cereb. Cortex* 3, pp. 373–386.

[109] Gilbert, C. D.; Wiesel, T. N. (1979). Morphology and Intracortical Projections of Functionally Characterised Neurones in the Cat Visual Cortex. *Nature* 280, pp. 120–125.

[110] Globus, G. G. (1992). Toward a Noncomputational Cognitive Neuroscience. *J. Cog. Neurosci.* 4, pp. 299–310.

[111] Gray, G. (1982). *Neuropsychology of Anxiety.* Oxford University Press, Oxford.

[112] Grossberg, S.; Merrill, J. W. L. (1992). A Neural Network Model of Adaptively Timed Reinforcement Learning and Hippocampal Dynamics. *Cog. Brain Res.* 1; pp. 3–38.

[113] Hendry, S. H. C.; Schwark, H. D.; Jones, E. G.; Yan, J. (1987). Numbers and Proportions of GABA-Immunoreactive Neurons in Different Areas of Monkey Cerebral Cortex. *J. Neurosci.* 7, pp. 1503-1519.

[114] Hinton, G. E.; McClelland, J. L.; Rumelhart, D. E. (1986). Distributed Representations. In [148], Vol. 1, pp. 77–109.

[115] Hubel, D. H. ; Wiesel, T. N. (1977). Functional Architecture of Macaque Monkey Visual Cortex. *Proc. Roy. Soc. Lond. Ser. B.* 198, pp. 1–59.

[116] Jones, E. G. (1984). Neuroglioform or Spiderweb Cells. In [142], Vol. 1, pp. 409–418.

[117] Jones, E. G. (1993). GABAergic Neurons and their Role in Cortical Plasticity in Primates. *Cereb. Cortex* 3, pp. 361-372.

[118] Kawaguchi, Y. (1993). Groupings of Nonpyramidal and Pyramidal Cells With Specific Physiological and Morphological Characteristics in Rat Frontal Cortex. *J. Neurosci.* 69, pp. 416. Pyr. Neuronen adaptieren, nicht Pyr. neuronen nicht.

[119] Killackey, H. P. (1990). Neocortical Expansion: An Attempt toward Relating Phylogeny and Ontogeny. *J. Cog. Neurosci.* 2, pp. 1–17.

[120] King, C. C. (1991). Fractal and Chaotic Dynamics in Nervous Systems. *Prog. Neurobiol.* 36, pp. 279–308.

[121] Kisvarday, S. F.; Martin, K. A. C.; Freund, T. F.; Magloczky, Z. F.; Whitteridge, D.; Somogyi, P. (1986). Synaptic Targets of HRP-Filled Layer 3 Pyramidal Cells in the Cat Striate Cortex. *Exp. Brain Res.* 64, pp. 541–552.

[122] Larkman, A. U. (1991). Dendritic Morphology of Pyramidal Neurones in the Visual Cortex of the Rat: 1. Branching Patterns. *J. Comp. Neurol.* 306, pp. 307–319.

[123] Larkman, A. U. (1991). Dendritic Morphology of Pyramidal Neurones in the Visual Cortex of the Rat: 3. Spine Distributions. *J. Comp. Neurol.* 306, pp. 332–343.

[124] Larkman, A. U.; Mason, A. (1990). Correlations Between Morphology and Electrophysiology of Pyramidal Neurons in Slices of Rat Visual Cortex. 1. Establishment of Cell Classes. *J. Neurosci.* 10, pp. 1407–1414.

[125] LeVay, S. (1986). Synaptic Organization of Claustral and Geniculate Afferents to the Visual Cortex of the Cat. *J. Neurosci.* 6, pp. 3564–3575.

[126] Lowenstein, P. R.; Somogyi, P. (1991). Synaptic Organization of Cortico–Cortical Connections From the Primary Visual Cortex to the Posteromedial Lateral Suprasylvian Visual Area in the Cat. *J. Comp. Neurol.* 310, pp. 253–266.

[127] Lund, J. S. (1990). Excitatory and Inhibitory Circuity and Laminar Mapping Strategies in the Primary Visual Cortex of the Monkey. In *Signal and Sense: Local and Global Order in Perceptual Maps*, Edelman, G. M.; Gall, W. E.; Cowan, W. M. (Eds.), Wiley, New York, pp. 51–66.

[128] Mamelak, A. N.; Hobson, J. A. (1989). Dream Bizarreness as the Cognitive Correlate of Altered Neuronal Behavior in REM Sleep. *J. Cog. Neurosci.* 1, pp. 201–222.

[129] Marin–Padilla, M. (1984). Neurons of Layer 1: A Developmental Analysis. In [142], Vol. 1, pp. 447–478.

[130] Martin, K. A. C. (1984). Neuronal Circuits in Cat Striate Cortex. In [142], Vol. 2, pp. 241–284.

[131] McCormick, D. A.; Wang, Z.; Huguenard, J. (1993). Neurotransmitter Control of Neocortical Neuronal Activity and Excitability. *Cereb. Cortex* 3, pp. 387–398.

[132] McGuire, B. A.; Gilbert, C. D.; Rivlin, P. K.; Wiesel, T. N. (1991). Targets of Horizontal Connections in Macaque Primary Visual Cortex. *J. Comp. Neurol.* 305, pp. 370–392.

[133] McGuire, B. A.; Hornung, J.-P.; Gilbert, C. D.; Wiesel, T. N. (1984). Patterns of Synaptic Input to Layer 4 of Cat Striate Cortex. *J. Neurosci.* 4, pp. 3021–3033.

[134] Mel, B. W. (1993). Synaptic Integration in an Excitable Dendritic Tree. *J. Neurophysi,* 70, pp. 1086–1101.

[135] Merzenich, M. M.; Kaas, J. H. (1982). Reorganization of Mammalian Somatosensory Cortex Following Peripheral Nerve Injury. *Trends in Neurosciences* 5, pp. 434–436.

[136] Orban, G. A. (1991). Quantitative Electrophysiology of Visual Cortical Neurones. In *The Neural Basis of Visual Function.* Leventhal, A. G. (Ed.), Macmillan Press, pp. 173–222.

[137] Parnavelas, J. G. (1985). Physiological Properties of Identified Neurons. In [142], Vol. 2, pp. 205-239.

[138] Pearlman, A. L. (1985). The Visual Cortex of the Normal Mouse and the Reeler Mutant. In [142], Vol. 3, pp. 1–18.

[139] Perret, D. I., Mistlin A. J., Chitty A. J. (1987). Visual Neurones Responsive to Faces. *Trends in Neurosciences* 10, pp. 358–364.

[140] Peters, A. (1984). Bipolar Cells. In [142], Vol. 1, pp. 381–408.

[141] Peters, A. (1987) Number of Neurons and Synapses in Primary Visual Cortex. in [142], Vol. 6, pp. 267—294.

[142] Peters, A.; Jones, E. G. (1984–1991) *Cerebral Cortex. Vol. 1 Cellular Components of the Cerebral Cortex; Vol. 2 Functional Properties of Cortical Cells; Vol. 3 Visual Cortex; Vol. 4 Association and Auditory Cortices; Vol. 5 Sensory-Motor Areas and Aspects of Cortical Connectivity; Vol. 6 Further Aspects of Cortical Function, Including Hippocampus; Vol. 7 Development and Maturation of Cerebral Cortex; Vol 8.1, 8.2 Comparative Structure and Evolution of Cerebral Cortex. Vol. 9 Normal and Altered States of Function.* Plenum, New York.

[143] Peters, A.; Saint Marie, R. L. (1984). Smooth and Sparsely Spinous Nonpyramidal Cells Forming Local Axon Plexuses. In [142], Vol. 1, pp. 419–445.

[144] Peters, A.; Sethares, C. (1991). Organization of Pyramidal Neurons in Area 17 of Monkey Visual Cortex. *J. Comp. Neurol.* 306, pp. 1–23.

[145] Purves, D.; R. Riddle, D.; LaMantia, A.-S. (1992). Iterated Patterns of Brain Circuity (Or How the Cortex Gets its Spots). *Trends in Neurosciences* **15**, No. 10, pp. 362–368.

[146] Rall, W.; Segdev, I. (1990). Dendritic Branches, Spines, Synapses and Excitable Spine Clusters. In [149], pp. 69–81.

[147] Rolls, E. (1989). The Representation and Storage of Information in Neuronal Networks in the Primate Cerebral Cortex and Hippocampus. In [91], pp. 125–159.

[148] Rumelhart, D. E.; McClelland, J. L. et al. (1986). *Parallel Distributed Processing. Volume 1: Foundations. Volume 2: Psychological and Biological Models.* MIT Press, Cambridge, MA.

[149] Schwartz, E. L. (Ed.) (1990). *Computational Neuroscience.* MIT Press, Cambridge, MA.

[150] Segev, I. (1992). Single Neurone Models: Oversimple, Complex and Reduced. *Trends in Neurosciences* **15**, No. 11, pp. 414–421.

[151] Sejnowski, T. J. (1986). Open Questions about Computation in Cerebral Cortex. In [148], Vol. 2, pp. 372–389.

[152] Shepherd, G. M. (1988). A Basic Circuit of Cortical Organization. In *Perspective of Memory Research*, Gazzaniga, M. S. (Ed.), MIT Press, Cambridge, MA, pp. 93–134.

[153] Shepherd, G. M. (Ed.) (1990^3). *The Synaptic Organization of the Brain.* Oxford University Press, Oxford.

[154] Shepherd, G. M.; Koch, C. (1990). Introduction to Synaptic Circuits. In [153], pp. 3–31.

[155] Singer, W. (1993). Neuronal Representations, Assemblies and Temporal Coherence. *Progr. in Brain Res.* **95**, pp. 461–474.

[156] Skarda, C. A.; Freeman, W. J. (1987). How the Brain Makes Chaos in Order to Make Sense of the World. *Behavioural and Brain Science* **10**, pp. 161-195. olfactory bulb, in Neurocomputing 2

[157] Somogyi, P. (1978). The Study of Golgi Stained Cells and of Experimental Degeneration under the Electron Microscope: A Direct Method for the Identification in the Visual Cortex of Three Successive Links in a Neuron Chain. *Neuroscience* **3**, pp. 167–180.

[158] Stanfield, B. B. (1992). The Development of the Corticospinal Projection. *Progr. Neurobiol.* **38**, pp. 169–202.

[159] Stratford, K.; Mason, A.; Larkman, A.; Major, G.; Jack, J. (1989). The Modelling of Pyramidal Neurones in the Visual Cortex. In [91], pp. 296–321.

[160] Sur, M.; Pallas, S. L.; Roe, A. W. (1990). Cross–Modal Plasticity in Cortical Development: Differentiation and Specification of Sensory Neocortex. *Trends in Neurosciences* **13**, No. 6, pp. 227–233.

[161] Sur, M.; Wall, J. T.; Kaas, J. H. (1981). Modular Segregation of Functional Cell Classes within Postcentral Somatosensory Cortex of Monkeys. *Science* **212**, pp. 1059–1061.

[162] Swindale, N. V. (1990). Is the Cerebral Cortex Modular? *Trends in Neurosciences* **13**, No. 12, pp. 487–492.

[163] Torre, V.; Poggio, T. (1978). A Synaptic Mechanism Possibly Underlying Directional Selctivity to Motion. *Proc. Roy. Soc. Lond. Ser. B.*, Vol. 202, pp. 409–416.

[164] Tovee, M. J.; Rolls, E. T.; Treves, A.; Bellis, R. P. (1993). Information Encoding and the Responses of Single Neurons in the Primate Temporal Visual–Cortex. *J. Neurophysiol.* **70**, pp. 640–654.

[165] Van Essen, D. C. (1985). Functional Organization of Primate Visual Cortex. In [142], Vol. 3, pp. 259–330.

[166] Vogt, B. A. (1991). The Role of Layer 1 in Cortical Function. In [142], Vol. 9, pp. 49–80.

[167] White, E. L.; Czeiger, D. (1991). Synapses Made by Axons of Callosal Projection Neurons in Mouse Somatosensory Cortex: Emphasis on Intrinsic Connections. *J. Comp. Neurol.* **303**, pp. 233–244.

[168] White, E. L.; Hersch, S. M. (1981). Thalamocortical Synapses of Pyramidal Cells Which Project from Sm1 to Ms1 Cortex in the Mouse. *J. Comp. Neurol.* **198**, pp. 167–181.

[169] White, E. L.; Keller, A. (1987). Intrinsic Circuity Involving the Local Axonal Collaterals of Corticothalamic Projection Cells in Mouse Sm1 Cortex. *J. Comp. Neurol.* **262**, pp. 13–26.

[170] White, E. L.; Keller, A. (1989). *Cortical Circuits: Synaptic Organization of the Cerebral Cortex — Structure, Function and Theory.* Birkhäuser, Boston.

[171] Winfield, D. A.; Brooke, R. N. L.; Sloper, J. J.; Powell, T. P. S. (1981). A Combined Golgi-Electron Microscopic Study of the Synapses Made by the Proximal Axon and Recurrent Collaterals of a Pyramidal Cell in the Somatic Sensory Cortex of the Monkey. *Neuroscience* **6**, 1217–1230.

[172] Winfield, D. A.; Powell, T. P. S. (1983). Laminar Cell Counts and Geniculocortical Boutons in Area 17 of Cat and Monkey. *Brain Res.* **277**, pp. 223–229.

[173] Witter, M. P.; Groenewegen, H. J.; DaSilva, F. H. L.; Lohman, A. H. M. (1989). Functional Organization of the Extrinsic and Intrinsic Circuity of the Parahippocampal Region. *Progr. Neurobiol.* **33** pp. 161–253.

[174] Zarzecki, P. (1986). Functions of Corticocortical Neurons of Somatosensory, Motor, and Parietal Cortex. In [142], Vol. 5, pp. 185–215.

[175] Zeki, S.; Shipp, S. (1988). The Functional Logic of Cortical Connections. *Nature* **335**, pp. 311–317.

Danksagung

Vor allem danke ich Prof. G. Eilenberger für die große wissenschaftliche Freiheit, die er mir gewährte, und für das Interesse, das er dieser Arbeit trotzdem entgegenbrachte. Seine Ratschläge haben die Darstellung viel verständlicher werden lassen. Auch von Prof. T. Tél und Prof. D. Mayer habe ich einiges über Nichtlineare Dynamik gelernt. Viele Mitarbeiter und Gäste des Institutes für Festkörperforschung haben diese Arbeit durch ihr Interesse oder ihre Hilfsbereitschaft gefördert. Meiner Mutter danke ich für die gründliche Suche nach Rechtschreibfehlern. Andreas Körner und Christian Schroer danke ich für gute Kameradschaft.