



Bessere Marktvorhersagen
für Aktien- und Devisenmärkte
mit lokalen Prediktoren

von

Dr. rer. nat. Hans Uhlig, Wissenschaftsautor

Phone: +49(40)605 51 50 Fax: +49(40)947 90 932 E-Mail: Hans.Uhlig@hamburg.de

Copyright © 2010 - Dr. Hans Uhlig

Copyright Hinweis

Der Text, die Tabellen und Abbildungen dieses Beitrages sowie die dazugehörige Software unterliegen dem Urheberrechtsschutz. Wer diese Produkte erwirbt, darf sie für den eigenen Gebrauch nutzen. Kopien oder Wiedergaben in anderer Form ob vollständig oder nur teilweise, bedürfen meiner schriftlichen Zustimmung.

Hans Uhlig, Januar 2010

Vorwort 5

Was Sie von diesem Handbuch erwarten dürfen 6

Teil I – Die Welt des Chaos 7

Zufallsbewegungen (random walks) und chaotische Trajektorien 7

Chaotische Modellsysteme 7

- a) Der 'Schmetterling' von Lorenz 7
- b) Der 'Trichter' von Rössler 8

Rekonstruktion mehrdimensionaler Systeme 10

- Einbettung durch Zeitverzögerung 10
- Einbettung durch Differenzen 10
- Datensequenz des Lorenz-Systems, grafisch 11
- Datensequenz des Rössler Systems, grafisch 11

- Einbettung des Lorenz-Systems 12
 - a) durch Zeitverzögerung 12
 - b) durch Differenzen 12

- Einbettung des Rössler-Systems 13
 - a) durch Zeitverzögerung 13
 - b) durch Differenzen 13

Komplexe Populationsdynamik – Chaos in lebenden Systemen 14

- Datensequenz der Mackey-Glass Funktion MG17, grafisch 15
- Datensequenz der Mackey-Glass Funktion MG30, grafisch 15
- Einbettung der Mackey-Glass Daten (MG17) durch Zeitverzögerung 16
- Einbettung der Mackey-Glass Daten (MG30) durch Zeitverzögerung 16

Phasenraumdarstellung von Finanzmärkten 17

- Datenbeschaffung 17
- Grafische Darstellung der Marktdynamik 17

- Finanzmarktdaten – Einbettung durch Zeitverzögerung 18
 - amerikanischer Aktienindex S&P 500, Tagesschlusskurse 18
 - Währungspaar Euro/US-\$, Wochenschlusskurse 18
 - amerikanischer Aktienindex S&P500, 2-Tages-Volatilitäten 19

Teil II – Datenanalyse und Modellierung 20

Aktienmärkte im Zeitverlauf grafisch: 22ff

- Deutschland DAX 22
- Österreich ATX 23
- Schweiz SMI 24

Devisenmärkte im Zeitverlauf grafisch: 25ff

- Sfr/Euro 25
- US\$/Euro 25

Jap. Yen/Euro	26	
Brit. Pfund/Euro	26	
Ind. Rupie/Euro	27	
Bras. Real/Euro	27	
Aus.-\$/Euro	28	
Can.-\$/Euro	28	
<u>Besonderheiten der Devisenmärkte</u>	29	
<u>Güte von Marktvorhersagen - Messen der (Informations)-Entropie</u>		30
Das Entropieprofil	30	
Richtung, Stärke und Häufigkeit von Marktbewegungen		31
Weitere Analysen zu Ordnung und Zufall bei Marktentwicklungen		31
Bedingte Entropie	32	
Chi ² -Test auf Unabhängigkeit	32	
<u>Die Verwendung von Surrogat-Daten</u>	33	
Mit Surrogatdaten zur Nullhypothese	33	
Teil III – Marktvorhersagen mit lokalen Prediktoren	35	
<u>Vorhersage mit 'Nächsten Nachbarn' – Nearest Neighbour Prediktoren</u>		35
Vorhersage mit ungeglätteten Originaldaten	35	
Bessere Vorhersagen mit geglätteten Daten zu bestimmende Parameter	36	
Die kostenlosen 'Nearest Neighbour' Prediktoren (NNP)	36	
<u>Märkte als teils stochastische, teils determinierte komplexe Systeme</u>		38
Die Verknüpfung von Chaos und Statistik	38	
Das Programm Orbitor für die Chaos-Statistik-Prognose	39	
Das Programm OrbitVar_Kon auf der Konsole	40	
Die Ausgabedatei des Programms OrbitVar_Kon	41	
<u>Probleme der Datenglättung</u>	43	
Vorschläge für Glättungskriterien	43	
Teil IV – Testen von 'Nearest Neighbour' und Chaos-Statistik Prediktoren	45	
TradeStation(TM), MetaStock(TM) usw. taugen nur für globale Prediktoren		45
Auswertungsprogramme für lokale Prediktoren	45	
Auswertungsprogramm für 'Nearest Neighbour' Prediktoren	46	
Auswertungsprogramm für Chaos-Statistik Prediktoren	48	
Chaos-Statistik und Geldmanagement	49	
Literatur	50	

Vorwort

Was Sie von diesem Handbuch erwarten können

Hier werde ich zeigen, dass man aus der Chaostheorie praktischen Nutzen ziehen kann. Wenn man die Erkenntnisse der Chaosforschung auf Finanzmärkte anwendet, ermöglicht dies bessere und konsistentere Vorhersagen als es mit konventionellen Prognoseverfahren der linearen Statistik, wie Autoregression gleitender Durchschnitte (ARMA) möglich wäre. Dieses Verfahren (ARMA) wird in der sogenannten 'Technischen Analyse' häufig verwendet.

Das Buch ist in mehrere Abschnitte gegliedert. Der erste Abschnitt enthält eine knappe Einführung in die Welt des Chaos und seine Methoden. Ich werde erläutern, warum und wie die Entdeckung des deterministischen Chaos die modernen Analysemethoden für dynamische Systeme verändert hat. Warum die Konzepte zur Modellierung dynamischer Systeme neu erfunden werden mussten und woher man jetzt weiß, dass verlässliche Langzeitvorhersagen unmöglich sind. Es werden Beispiele für verschiedene Modellsysteme gezeigt, und die Begriffe, wie Einbettung und Phasenraumdarstellung praktisch vorgeführt. Neben den bekannten chaotischen Modellsystemen werden auch weniger bekannte, wie die Mackey-Glass Funktionen vorgestellt, die größere Ähnlichkeit zum Marktverhalten aufweisen.

Der zweite Abschnitt führt Sie durch die einzelnen Schritte der Datenanalyse. Begonnen wird mit den herkömmlichen grafischen Darstellungen von Kursverläufen einiger Aktienmärkte und Wechselkurse. Dann werden die Analyseprogramme vorgestellt, mit denen der Informationsgehalt von Finanzdaten bestimmt werden kann: Entropieprofil, bedingte Entropie und Chi-Quadrat-Test auf Unabhängigkeit. Auch die Verwendung von Surrogatdaten wird besprochen und gezeigt.

Im dritten Abschnitt wird die Marktprognose mit lokalen Prediktoren vorgestellt. Es wird erklärt, welche grundsätzlichen und wesentlichen Unterschiede zwischen den herkömmlichen globalen Prediktoren und den lokalen Prediktoren bestehen. Die 'Nearest Neighbour' Methode für lokale Vorhersagen wird nur kurz abgehandelt, da sie an anderer Stelle ausführlich beschrieben wird. Doch die von mir neu entwickelten Chaos-Statistik Prediktoren werden genauer erläutert und Programme dafür ebenfalls besprochen. Schließlich bespreche ich noch zwei Auswertungsprogramme für die lokalen Prediktoren. Mit den herkömmlichen Simulationsprogrammen, wie TradeStation^(TM) oder MetaStock^(TM), kann man lokale Prediktoren nicht testen, weil diese Programme nur für die Analyse globaler Strategien und Handelssysteme geeignet sind. Und ich weise auf den Nutzen der Chaos-Statistik für das Geldmanagement hin.

Ein ähnliches Handbuch habe ich vor etwa zehn Jahren einmal auf Englisch geschrieben, und zusammen mit Daten und Software in viele Teile der Welt verkauft. Neu in diesem Handbuch sind einmal die neuen Daten, auf die es sich bezieht, die Konzentration auf europäische Märkte, Hinzunahme von Devisenmarktdaten, Analyse von Tagesdaten sowie die neuen Programme für Analyse und Prognose und es ist in Deutscher Sprache geschrieben.

Teil I – Die Welt des Chaos

Zufallsbewegungen und chaotische Trajektorien

Statistische Vorhersagen beruhen nur auf Wahrscheinlichkeiten. Sie sind ideal angebracht, in den Fällen, in denen einzelne zukünftige Ereignisse unvorhersehbar sind und daher nur Wahrscheinlichkeiten für die jeweils möglichen Ereignisse angegeben werden können. Mit anderen Worten: Sie sind ideal zur Vorhersage zufälliger Ereignisse.

Über lange Zeit wurden Marktbewegungen als Zufallsereignisse angesehen und immer noch ist diese Ansicht recht populär. Daher rührt auch der Erfolg der 'random walk theory', der Theorie des Zufallspfades, die noch immer viele Anhänger im Bereich der Wirtschafts- und Finanzwissenschaften findet. Diese Theorie wurde in den sechziger Jahren des letzten Jahrhunderts aufgestellt, nachdem man viele Marktdaten statistisch ausgewertet hatte. Die Aussagen dieser Theorie wurden durch viele ähnliche Analysen im Laufe der Zeit mehrfach bestätigt.

Nun denn, was hat unsere Ansicht über diese Theorie gewandelt? Im Grunde war es die Entdeckung des deterministischen Chaos.

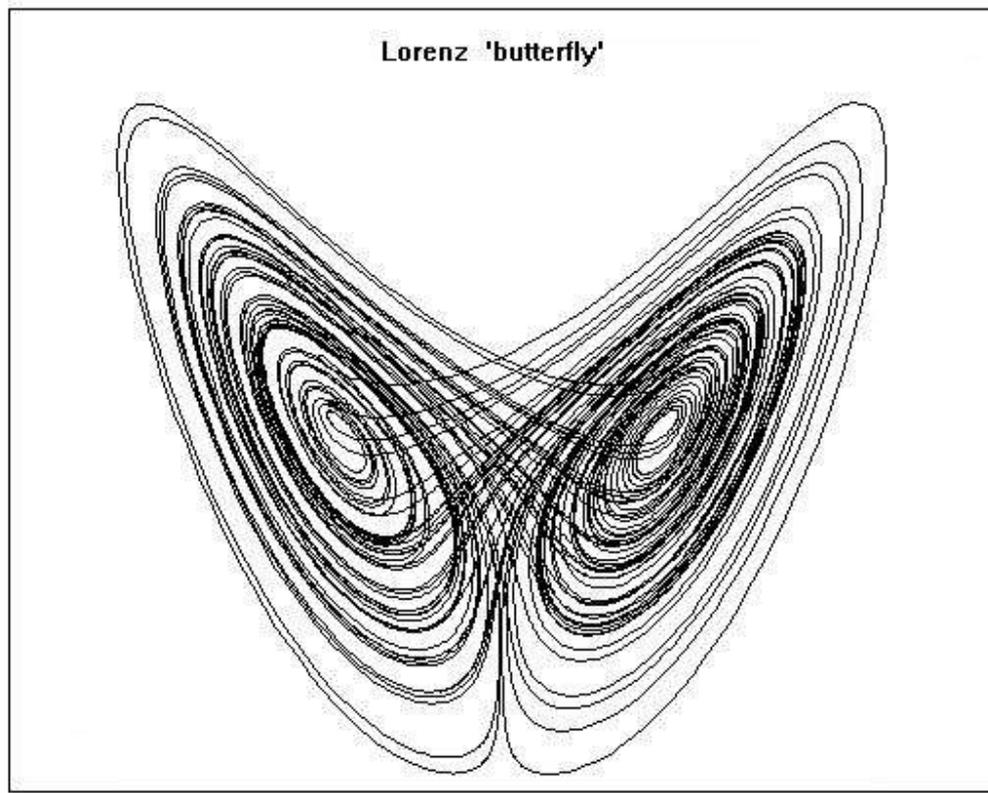
Chaotische Phänomene sind vielen seit langem bekannt und viele Wissenschaftler verschiedener Disziplinen haben sie beobachtet und beschrieben. Aber sie wurden als Sonderphänomene und Kuriositäten abgetan. Zum einen, weil man sie für unbedeutende Ausnahmen hielt und zum anderen, weil die Methodik fehlte, um chaotische Systeme sicher zu erkennen, sie genau zu beschreiben, zu analysieren und einzuordnen. Nicht vor der Mitte der achtziger Jahre wurden die ersten Verfahren entwickelt, die man auch heute noch benutzt, um chaotische Systeme zu identifizieren und qualitativ und quantitativ zu beschreiben. Und erst die detaillierte und systematische Untersuchung solcher Systeme durch Physiker und Mathematiker hat gezeigt, unter welchen Bedingungen überhaupt chaotisches Verhalten auftreten kann.

Inzwischen weiß man, dass recht einfache dynamische Systeme, die aus nur drei Komponenten bestehen, anscheinend unvorhersagbares Verhalten produzieren können: Chaos. Und dieses Verhalten kann auch dann auftreten, wenn die Wechselwirkungen zwischen den einzelnen Komponenten des Systems in exakten mathematischen Gleichungen beschrieben werden. Also: in diesen Gleichungen ist alles genau bestimmt. Einen Zufallseinfluss gibt es nicht und dennoch sieht das Verhalten dieser Systeme wie zufällig aus.

Die systematische Untersuchung chaotischer Modellsysteme erbrachte einige wichtige grundsätzliche Erkenntnisse: Nicht alle beliebigen Systeme aus drei Komponenten waren in der Lage, chaotisches Verhalten hervorzubringen. Als notwendige Bedingung, um Chaos zu erzeugen, erwies sich, dass mindestens eine Wechselbeziehung zwischen den Komponenten nichtlinear sein musste, d.h. ein Produkt aus mindestens zwei Variablen in der Funktion enthalten sein musste.

Das erste Modell eines chaotischen Systems wurde von dem amerikanischen Wetterforscher Edward Lorenz zufällig entdeckt. Es ist in der Öffentlichkeit bekannt als der sogenannte 'Schmetterling' von Lorenz. Dieses System besteht aus drei Komponenten. siehe Abbildung. Jeder Systemzustand entspricht einem Punkt im Raum. Da das System dynamisch ist, ändert es seinen Zustand fortwährend. Wenn man nun die aufeinander folgenden Zustände im Verlauf zeigen möchte, verbindet man die einzelnen Punkte zu Linien. Wenn man sich Linien in drei Dimensionen schwer vorstellen kann, ist vielleicht ein Modell aus Draht ein besseres Bild. Die dreidimensionale Figur, die man erhält, wenn man die Abfolge vieler Zustände so darstellt, ist ein Ausschnitt aus der Menge der möglichen Zustände des Systems. Um das System in zwei Dimensionen, also wie hier, auf dem Papier, darstellen zu können, lässt man die dritte Dimension fort und trägt nur zwei Dimensionen gegeneinander auf. Im folgenden Beispiel sind nur die Punkte auf der x,z Ebene aufgetragen worden. Dabei kommt es zwangsläufig zu Überschneidungen von Linien, weil die dritte Dimension fehlt. Es sieht dann so aus, als könnte

von einem Überschneidungspunkt ausgehend, das System zwei verschiedene Folgezustände annehmen. Das ist jedoch nur ein Darstellungsartefakt.



Lorenz model

$$\frac{dx}{dt} = \text{sigma} * (y - x)$$

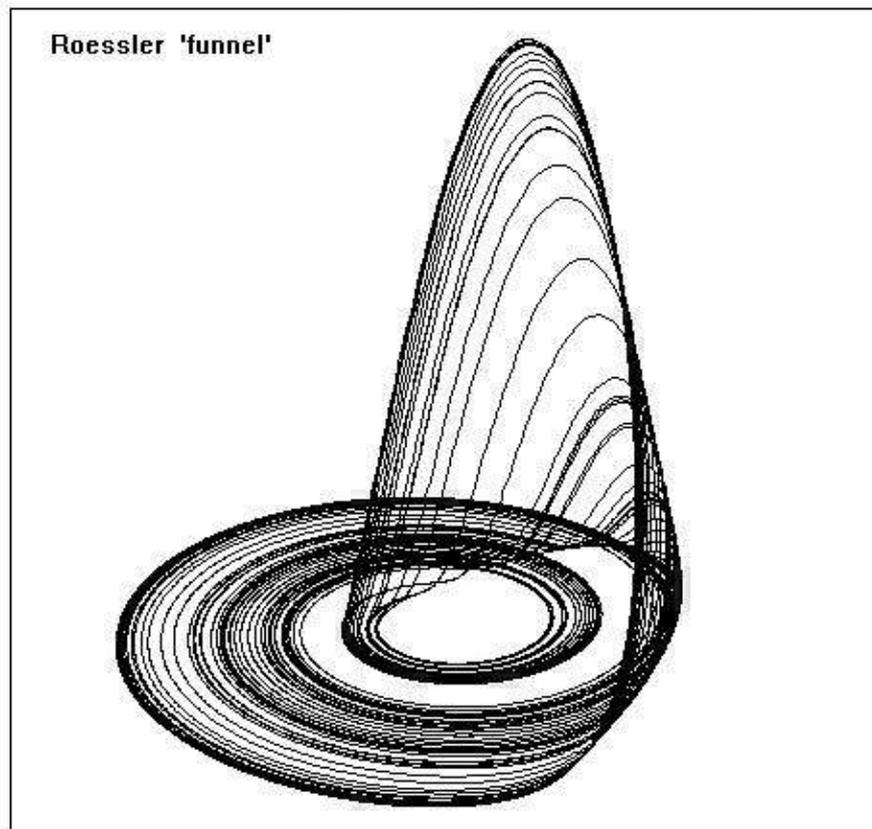
$$\left(\begin{array}{l} \frac{dy}{dt} = x * (r - z) - y \\ \frac{dz}{dt} = x * y - b * z \end{array} \right) \text{ non linear terms}$$

b, r, sigma = constants

Dieses dynamische System von Lorenz ist nicht das einfachste, chaotische System, denn es weist zwei nichtlineare Wechselwirkungen zwischen den Systemkomponenten auf. Die Formeln für das System aus drei Gleichungen ist unten angegeben. Sie beschreiben, wie sich die einzelnen Komponenten über die Zeit ändern. Bitte beachten Sie, dass b,r und sigma Konstanten sind. Nichtlinearität findet sich in der zweiten Komponente aus dem Produkt der Variablen (x*z) und in der dritten Komponente aus einem anderen Produkt der beiden Variablen (x*y).

Ein weniger bekanntes chaotisches Modellsystem, ist der sogenannte 'Trichter' von Rössler, den man in der nächsten Abbildung erkennen kann. Er ist ein Beispiel für das einfachste dynamische System mit chaotischen Eigenschaften. Und es war das erste überhaupt, das nur eine einzige nichtlineare Kopplung erforderte. Die Abbildung ist auch hier eine zweidimensionale Projektion der dreidimensionalen Bewegung des Systems. Unten sind die Formeln des Gleichungssystems angegeben. Sie zeigen, wie sich die drei Komponenten des

Systems über die Zeit ändern. Beachten Sie, dass a,b, und c Konstanten sind. Die Nichtlinearität entsteht in der Formel für die dritte Komponente aus dem Produkt von zwei Variablen (z*x).



Roessler model

$$\frac{dx}{dt} = -y - z$$

$$\frac{dy}{dt} = x + a * y$$

$$\left(\frac{dz}{dt} = b + z * (x - c) \right) \text{ non linear term}$$

a, b, c = constants

Chaotische Zeitreihen (Trajektorien) sehen wie Folgen von zufälligen Ereignissen aus, und das nicht nur für das bloße Auge. Lineare statistische Tests oder die Fourier Analyse können keine Unterschiede zwischen chaotischen Daten und Zufallsereignissen erkennen. Auch die Wavelet Analyse, eine Weiterentwicklung und Erweiterung der Fourier Analyse, versagt bei einem solchen Problem. Man erkannte, dass Tests, die sonst anscheinend zuverlässig geordnete, von zufälligen Daten unterscheiden konnten, deterministisch chaotische Systeme als zufällig erzeugt einstufen. Das stimmte die Statistiker nachdenklich. Mit der Zeit kamen immer mehr Zweifel an den Tests für Zufälligkeit auf, die mit solchen Methoden erhalten worden waren.

Weil aber diese Tests in der Vergangenheit als Universalinstrumente zur Datenanalyse eingesetzt worden waren, und eine ganze Reihe von Problemen damit untersucht worden sind, können damit Ereignisse fälschlich als zufällig klassifiziert worden sein, auf die dieses eigentlich nicht zutrifft.

Wenn Datenfolgen zufällig sind, dann können künftige Entwicklungen am besten mit Hilfe gleitender Durchschnitte (ARMA = auto-regressive moving averages) bzw. mit integrierten gleitenden Durchschnitten (ARIMA = auto-regressive integrated moving averages) vorhergesagt werden. Diese Methoden sind bei der sogenannten 'Technischen Analyse' weit verbreitet. Es ist ganz entscheidend, zu wissen, ob Datenfolgen zufällig sind oder nicht, denn wenn nachgewiesen wird, dass Marktbewegungen nicht zufällig sind, dann gibt es bessere Methoden als z.B. ARMA um die zukünftige Entwicklung vorherzusagen.

Zum Glück sind auch für so schwierige Probleme geeignete Tests verfügbar, die zweifelsfrei geordnete Ereignisse von zufälligen unterscheiden können. Das sind Tests, die zum Beispiel verwendet werden, um damit die Güte von Zufallszahlengeneratoren zu untersuchen. Zufallszahlengeneratoren sind Computerprogramme. Sie erzeugen Zahlenfolgen, die echten zufälligen Zahlenfolgen beinahe ideal entsprechen. Diese werden benutzt, um Simulationsexperimente durchzuführen. Die Tests, mit denen man die Zahlenfolgen überprüft, legen sehr strenge Maßstäbe an die 'Zufälligkeit' an. Sie können nicht nur lineare, sondern auch nichtlineare Abhängigkeiten zwischen Daten erkennen. Zu dieser Kategorie von Tests gehört z.B. der bekannte Chi²-Test auf Unabhängigkeit, oder der etwas weniger bekannte Test auf bedingte Informationsentropie, bzw. der Test auf Transinformation.

Mit Hilfe dieser strengen Tests habe ich die Daten vieler Märkte auf Zufälligkeit untersucht (Uhlig, H. 1999). Außerdem habe ich Marktdaten mit sogenannten Surrogatdaten verglichen. Das sind Datenreihen, die statistisch den Originaldaten entsprechen. Sie haben gleichen Mittelwert Standardabweichung, Schiefe und Spannweite. Nur die Reihenfolge der Zustandsänderungen wurde mit Hilfe von Zufallszahlengeneratoren vertauscht. Dabei zeigte sich, dass die Pseudozufallsdaten nicht von Zufallsereignissen unterscheidbar waren, während die echten Marktdaten zumindest signifikant ($p > 95\%$) von Zufallsdaten verschieden waren, meist waren die Daten hoch signifikant ($p > 99\%$) zu unterscheiden von Zufallsdaten. Diese Ergebnisse legen einmal mehr nahe, dass die 'random walk' Hypothese zurückgewiesen werden sollte. Märkte verhalten sich reflexiv, wie George Soros sagt, d.h. selbstbezogen. Mit anderen Worten: Sie reagieren auf ihre eigenen Veränderungen. Aber seien Sie vorsichtig, die Marktdynamik ist nicht so simpel, wie Anhänger der sogenannten 'Technischen Analyse' vermuten.

Die Erkenntnisse habe ich nicht allein gewonnen. Viele Forscher aus aller Welt haben die Wirtschaftsdaten neu überprüft und mehrere Hinweise auf nichtlineare Dynamik bei verschiedenen Kenndaten gefunden:

- Konjunkturzyklen
- Arbeitslosenzahlen
- Wirtschaftswachstum

und bei etlichen Finanzmärkten konnte man ebenfalls Hinweise auf nichtlineare Dynamik nachweisen:

- Aktien und Indekurse
- Devisenkurse und Volatilität
- Rohölpreise, Spotpreise und Futures
- Agrarpreise
- Immobilienpreise

Die Beispiele finden Sie in Publikationen, die im Literaturanhang aufgelistet sind. Viele davon kann man aus dem Internet als *.pdf Dateien erhalten.

Dass es in allen diesen Fällen bessere Vorhersagen geben muss, als mit gleitenden Durchschnitten, ist ermutigend und gleichzeitig Ansporn, nach entsprechenden Verfahren zu suchen. Wie man diese Verfahren findet, soll in den folgenden Abschnitten gezeigt werden.

Rekonstruktion mehrdimensionaler komplexer dynamischer Systeme

In den chaotischen Modellsystemen gibt es Wechselwirkungen zwischen den einzelnen Komponenten, wie man an den Formeln sehen kann. Jede Formel für eine der Einzelkomponenten enthält auch die beiden anderen Komponenten. Das bedeutet wiederum, dass jede der Komponenten auch Informationen über die jeweils anderen Komponenten enthält. Dies kann man sich zunutze machen, um aus den Daten für eine der Komponenten das ganze dynamische System zu rekonstruieren.

Für das Lorenz-System lassen wir uns eine Datenreihe von x-Werten ausdrucken und für das Roessler-System werden wir mit einer Datenreihe von y-Werten arbeiten. Grafische Darstellungen der beiden Zeitreihen lassen wenig Regelmäßigkeiten erkennen. Das Roessler-System sieht weniger unruhig aus als das Lorenz-System, doch dies scheint nur so, weil die Datenreihe des Roessler-Systems wesentlich kürzer ist. Die kürzere Zeitreihe wurde aus Gründen der Übersichtlichkeit gewählt.

Einbettung durch Zeitverzögerung

Eine Methode dafür ist die sogenannte Einbettung durch Zeitverzögerung. Ich werde sie für das Lorenz-System (Schmetterling) und das Roessler-System (Trichter) vorstellen. Es wird keine vollständige Einbettung werden, weil man dafür mindestens drei Dimensionen bräuchte. Auf dem Bildschirm oder auf Papier, die jeweils nur Darstellungsflächen sind, lassen sich drei Dimensionen jedoch nicht gut wiedergeben. Wir werden uns daher mit zwei Dimensionen begnügen.

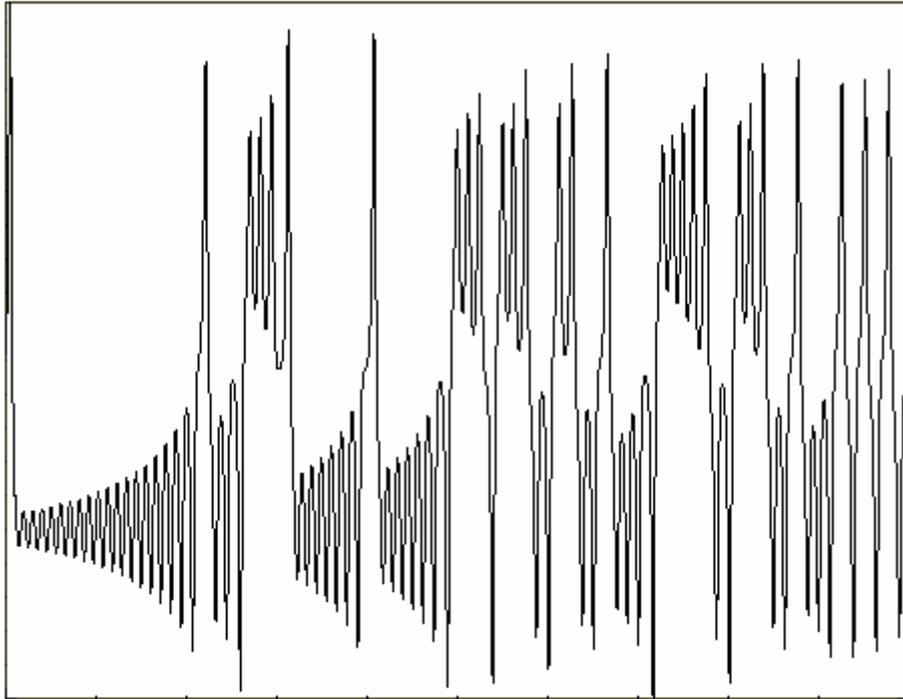
Die zweite Dimension erhalten wir aus der ersten Dimension, indem wir aus der ersten Reihe die oberen Daten weglassen und erst mit dem dritten, vierten, oder einem noch späteren Wert der ersten Reihe die zweite Reihe beginnen. Die zweite Reihe ist damit wie die erste, nur zeitlich gegen diese verschoben, bzw. zeitlich verzögert.

Wenn wir eine zweidimensionale Darstellung erhalten möchten, müssen wir jedes Wertepaar als einen Punkt in der Grafik betrachten, wobei der Wert aus der ersten Datenreihe die x-Koordinate angibt und der zugehörige Wert aus der zweiten Datenreihe die y-Koordinate. Wenn wir aufeinander folgende Punkte mit Linien verbinden, dann erhalten wir die folgenden Abbildungen. Für die zeitverzögerte Einbettung des Lorenz-Systems wurde die zweite Datenreihe um zwei Zeitschritte verschoben, beim Roessler-System sind es sieben Zeitschritte.

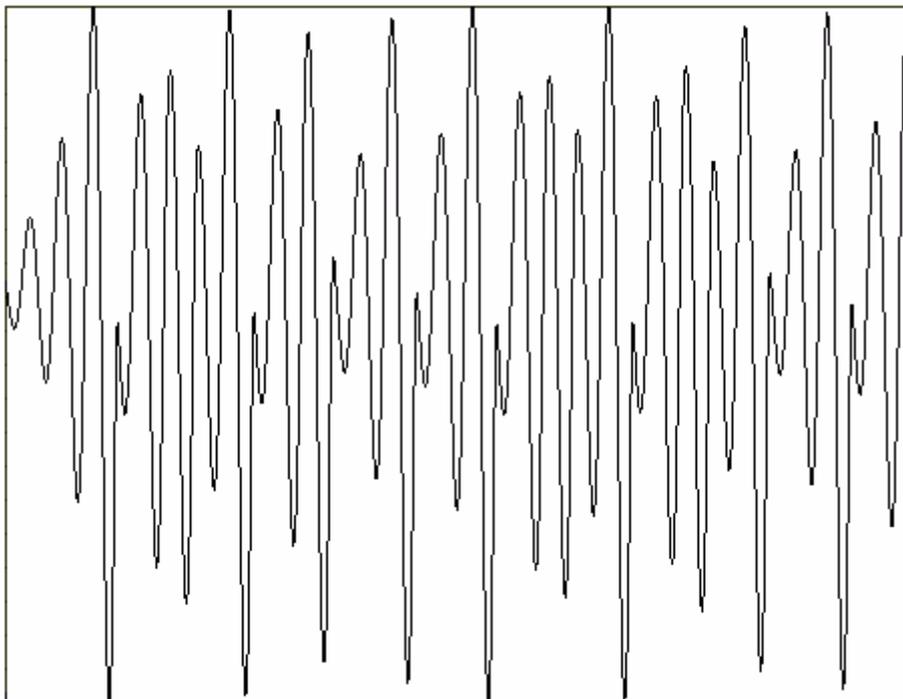
Die gerade beschriebene Form der Einbettung ist nicht die einzige, aber die am vielseitigsten einsetzbare. Die nachfolgend gezeigte Einbettung mit Differenzen setzt voraus, dass die erste Datenreihe keine größeren Sprünge enthält. Bei realen Daten kommt dies jedoch häufig vor, daher findet man diese Einbettung für reale Daten selten. Durch vorherige Glättung kann man auch reale Daten so transformieren, dass sie mit Differenzen eingebettet werden können, doch jede Glättung bringt auch Probleme mit sich, wie später noch erörtert werden wird.

Einbettung mit Differenzen

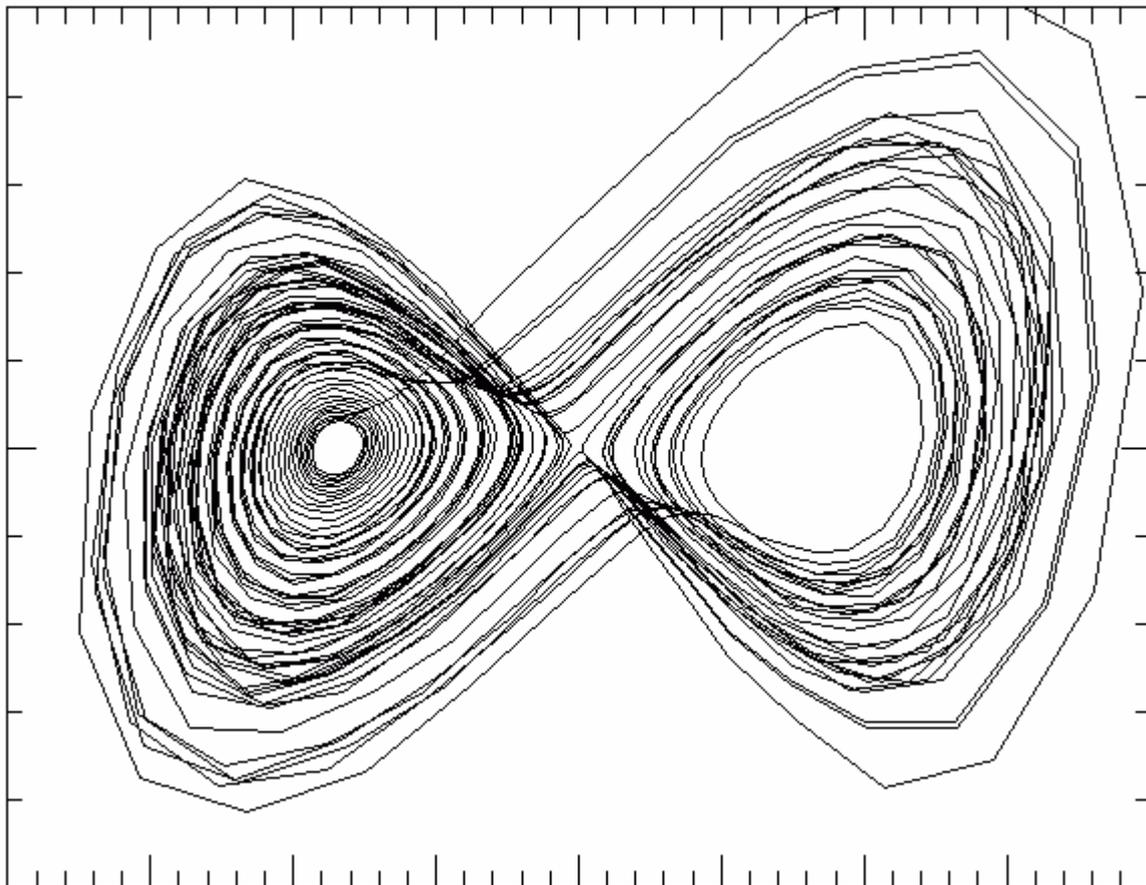
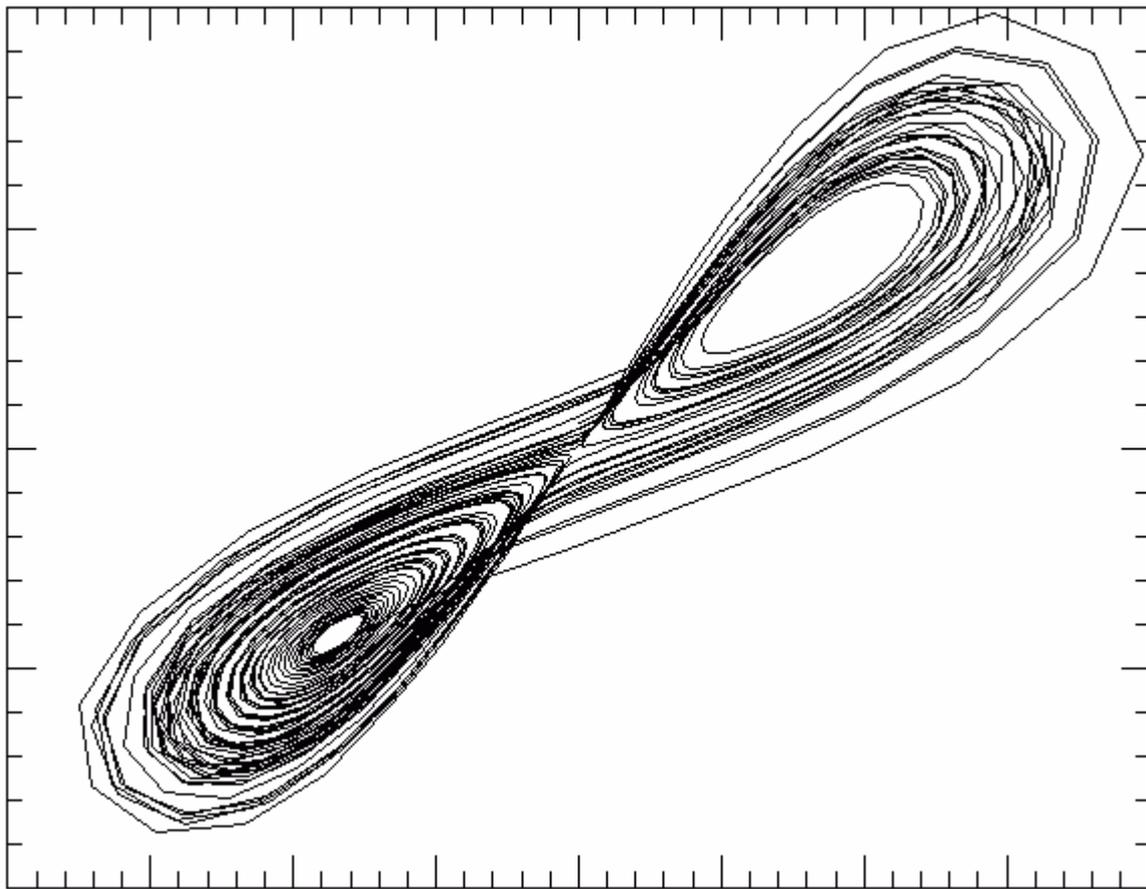
Bei dieser Einbettung verwendet man nicht direkt die Werte aus der ersten Zeitreihe für die zweite Dimension, sondern man erhält diese, indem man die Differenzen aufeinander folgender Werte berechnet und diese als die zweite Dimension verwendet. Der erste Wert der zweiten Dimension ist dann beispielsweise die Differenz der ersten beiden Werte der ersten Zeitreihe, der zweite Wert ist die Differenz zwischen dem zweiten und dem dritten Wert der ersten Zeitreihe usw.. Auf diese Weise wurden die nachfolgenden Abbildungen erhalten. In beiden Fällen beginnt die zweite Datenreihe mit der Differenz zwischen dem ersten und dem zweiten Wert der ersten Datenreihe.



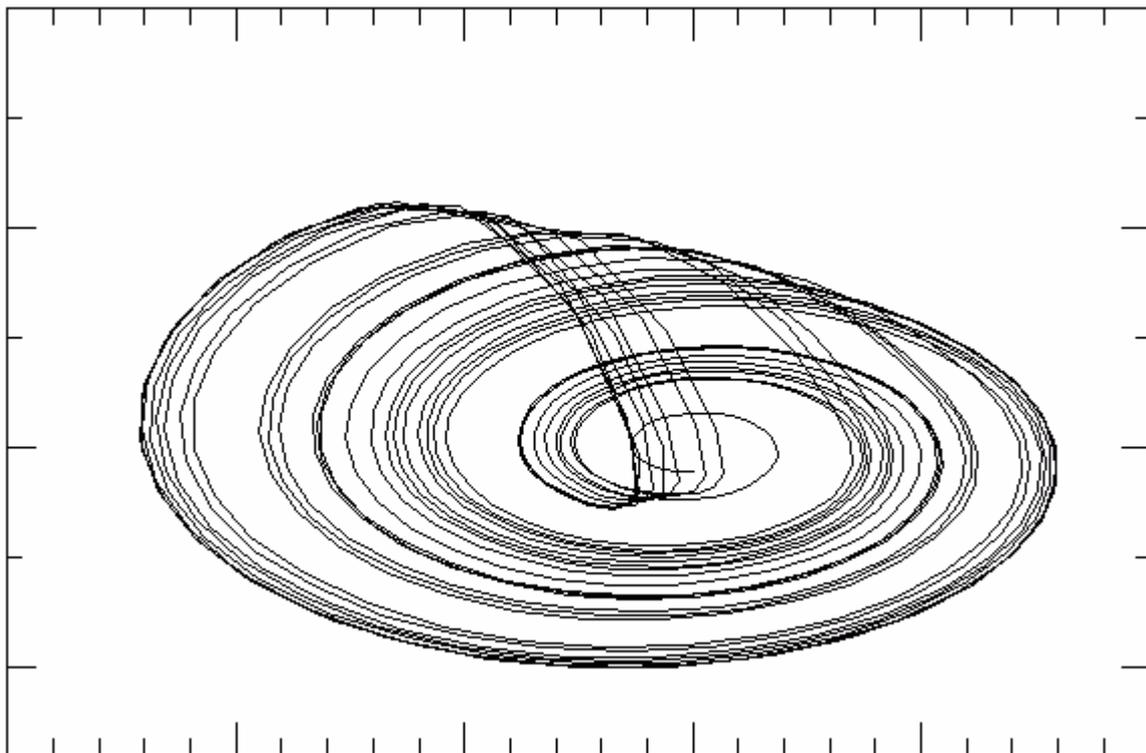
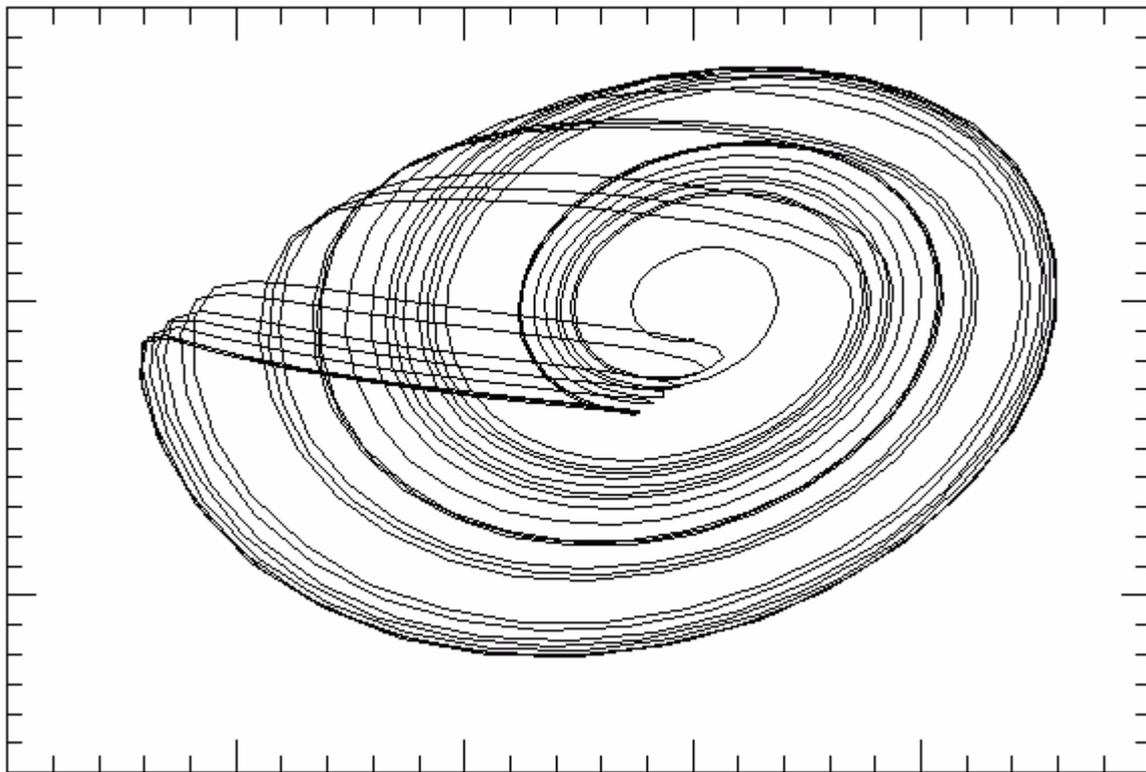
Die obere Grafik zeigt die Zeitreihe der x-Komponente des Lorenz-Systems, aus der die Rekonstruktion des Systems auf der nächsten Seite vorgenommen wurde.



Diese untere Grafik stellt die Zeitreihe der y-Komponente des Roessler-Systems dar. Aus dieser Zeitreihe wurde die Rekonstruktion des Systems vorgenommen, die man auf der übernächsten Seite sehen kann.



Die obere Grafik zeigt die Rekonstruktion des Lorenz-Systems aus der x-Komponente durch zeitverzögerte Einbettung. Unten sieht man die entsprechende Einbettung mit Differenzen.



Die obere Grafik zeigt die Rekonstruktion des Roessler-Systems aus der Zeitreihe der y-Komponente durch zeitverzögerte Einbettung. Die untere Grafik stellt die entsprechende Einbettung mit Differenzen dar.

Komplexe Populationsdynamik – Chaos in lebenden Systemen

Beispiele aus der Biologie, sind oft recht gute Modelle für Märkte, wie der Wirtschaftswissenschaftler Andrew Lo (1997,1999) von der amerikanischen Denkfabrik M.I.T. bemerkt. Darunter sind besonders die Modelle für Populationsdynamik bekannt geworden, die von James Mackey und Leon Glass im Jahr 1977 etabliert wurden.

Die beiden Forscher versuchten, ein Modell für die Regulierung der Konzentration bestimmter weißer Blutkörperchen (Granulozyten) in einem an Leukämie erkrankten Kind aufzustellen, die aperiodisch schwankten. Dabei fanden die Wissenschaftler heraus, dass mit verzögerten Differentialgleichungen gute Modelle für diese physiologische Kontrolle erstellt werden können.

Eine solche verzögerte Differentialgleichung hat die allgemeine Form

$$\frac{dx}{dt} = P^* (x_{(t-\tau)} - \alpha^*x_{(t)})$$

Der mathematische Term beschreibt, dass die zeitliche Änderung eines bestimmten Merkmals X nicht nur auf der gegenwärtigen Messwert für dieses Merkmal beruht, sondern auch davon abhängt, welchen Wert dieses Merkmal zu einem bestimmten Zeitpunkt in der Vergangenheit angenommen hatte. Dabei bezeichnet P die Neuproduktion, τ ist die Zeitverzögerung und α eine Zerfallskonstante. In dieser allgemeinen Form ist noch nichts über die Art der Funktion für die Neuproduktion ausgesagt. Mackey und Glass wählten dafür eine Buckelfunktion, die ein einzelnes Maximum bei mittleren Werten von X liefert.

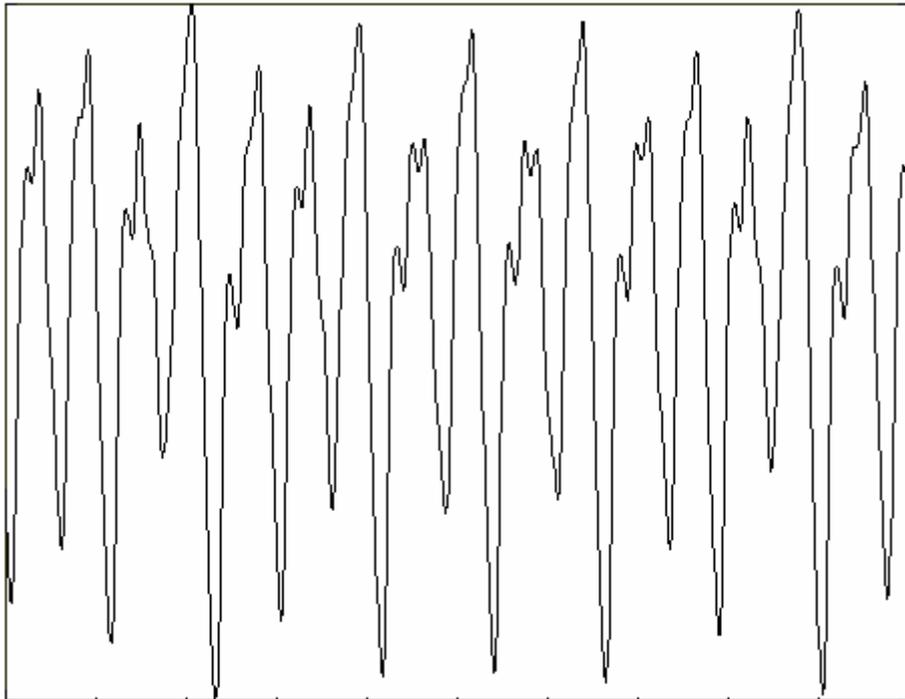
$$\frac{dx}{dt} = \frac{0,2^*x_{(t-20)}}{1+x_{(t-20)}^{10}} - 0,1^*x_{(t)}$$

Für Werte von $x_{(t-20)}$ die viel kleiner sind als 1, wird der Wert unter dem Bruchstrich nahe an eins, aber dann ist auch der Wert über dem Bruchstrich klein, so dass im vorderen Teil der Gleichung ein kleiner Wert entsteht. Wenn $x_{(t-20)}$ knapp unterhalb von 1 ist, dann wird der Wert unter dem Bruchstrich dicht an 1 liegen und der Wert über dem Bruchstrich nahe 0,2. Wenn aber $x_{(t-20)}$ viel größer ist als 1, dann wird der Wert unter dem Bruchstrich sehr viel größer als der Wert über dem Bruchstrich, denn im Nenner wird dieser Wert viele Male mit sich selbst multipliziert. Das heißt, der Bruch ergibt insgesamt wieder einen kleinen Wert. Wenn man verschiedene Zahlen für $x_{(t-20)}$ einsetzt, dann erkennt man, dass der Bruch im ersten Teil dieser Funktion ein Maximum liefert, wenn x zum Zeitpunkt t-20 etwa 0,8 war. Bei höheren oder niedrigeren Werten wird das Funktionsergebnis kleiner.

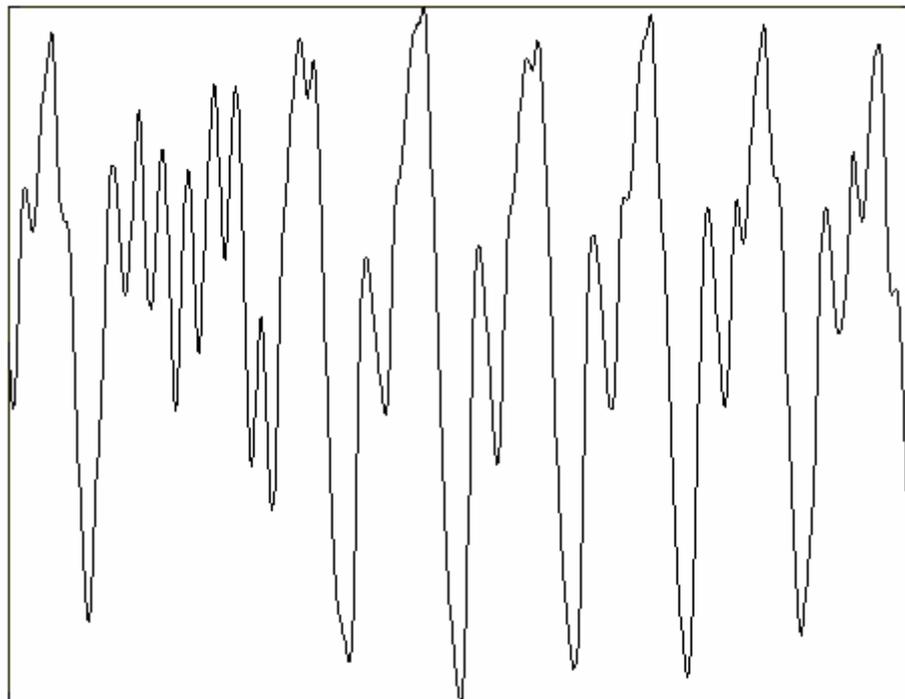
Eine solche Populationsdynamik ist nicht nur bei Blutzellen zu beobachten. Ähnliche Populationsschwankungen hatte bereits der australische Ökologe Alexander John Nicholson 20 Jahre vorher festgestellt. Er hatte über viele Generationen Goldfliegen (*Lucilia cuprina*) gezüchtet, metallisch glänzende grüne Fliegen die zu den Schmeißfliegen gehören. Man kann diese Insekten in landwirtschaftlichen Gebieten oft auf Kot von Kühen, Schafen oder Pferden sehen. Nicholson hatte festgestellt, dass die Populationen von Larven und ausgewachsenen Fliegen in seiner Zucht aperiodisch stark schwankten, trotz insgesamt stets ausreichendem Nahrungsangebot. Das Problem war offenbar die Verteilung. Nicholson nahm an, dass die Konkurrenz der Individuen untereinander diese Populationsschwankungen bewirkte. Diese These veröffentlichte er bereits im Jahr 1950.

Die Formeln von Mackey und Glass haben sich als Standardobjekte der Chaosforschung etabliert. Besonders bekannt geworden sind die Gleichungen für Zeitverzögerungen von 17 bzw. 30 Zeitschritten. Diese sind Modelle für chaotische Systeme mit den fraktalen Dimensionen 2,1 und 2,9. Was gerade diese Systeme besonders interessant macht, ist nicht allein die Tatsache, dass auch Marktentwicklungen durch verzögerte Differentialgleichungen nachgebildet werden können, sondern dass für einige Devisen- und Aktienmärkte fraktale Systemdimensionen geschätzt wurden, die gerade im Bereich zwischen 2,0 und 3,0 liegen.

Mackey-Glass Funktionen – Verzögerte Differentialgleichungen
Grafische Darstellungen der Funktionswerte im Zeitverlauf

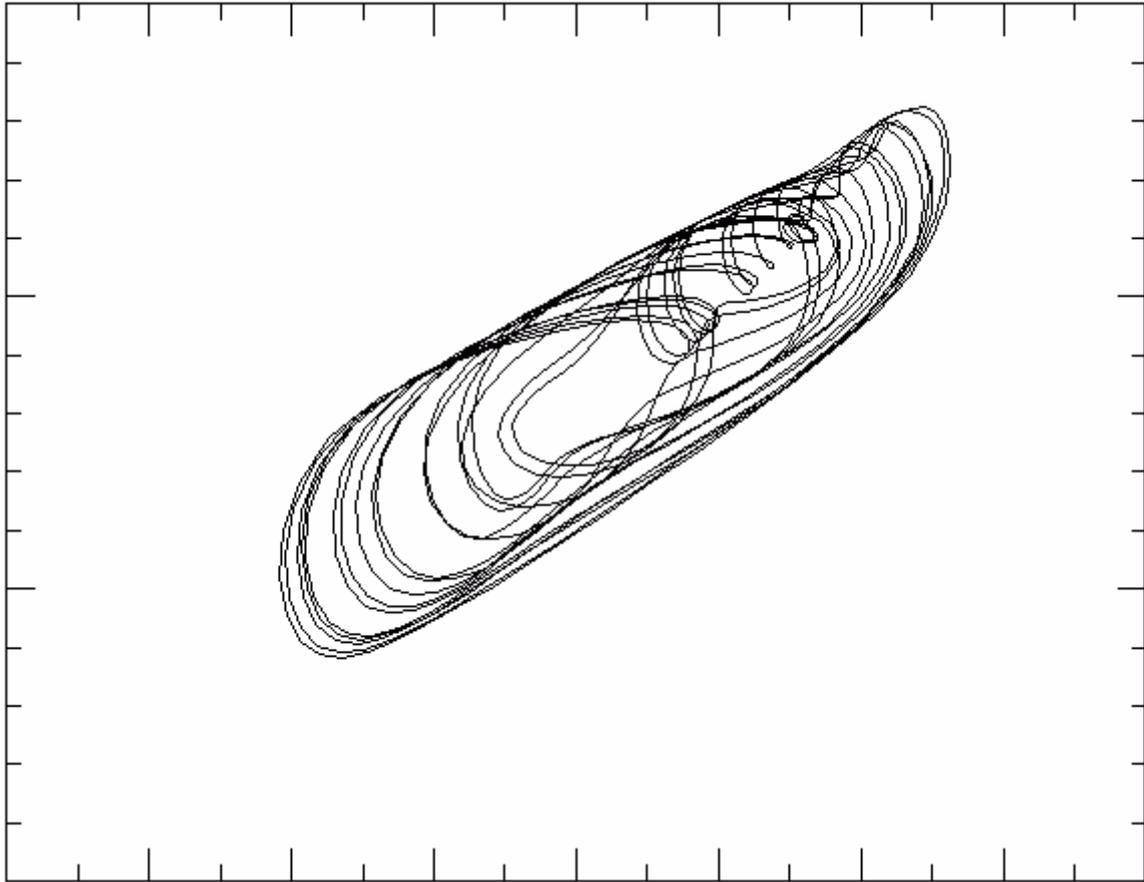


oben: MG17 – 17 Zeitschritte Verzögerung (fraktale Dimension = 2,1)

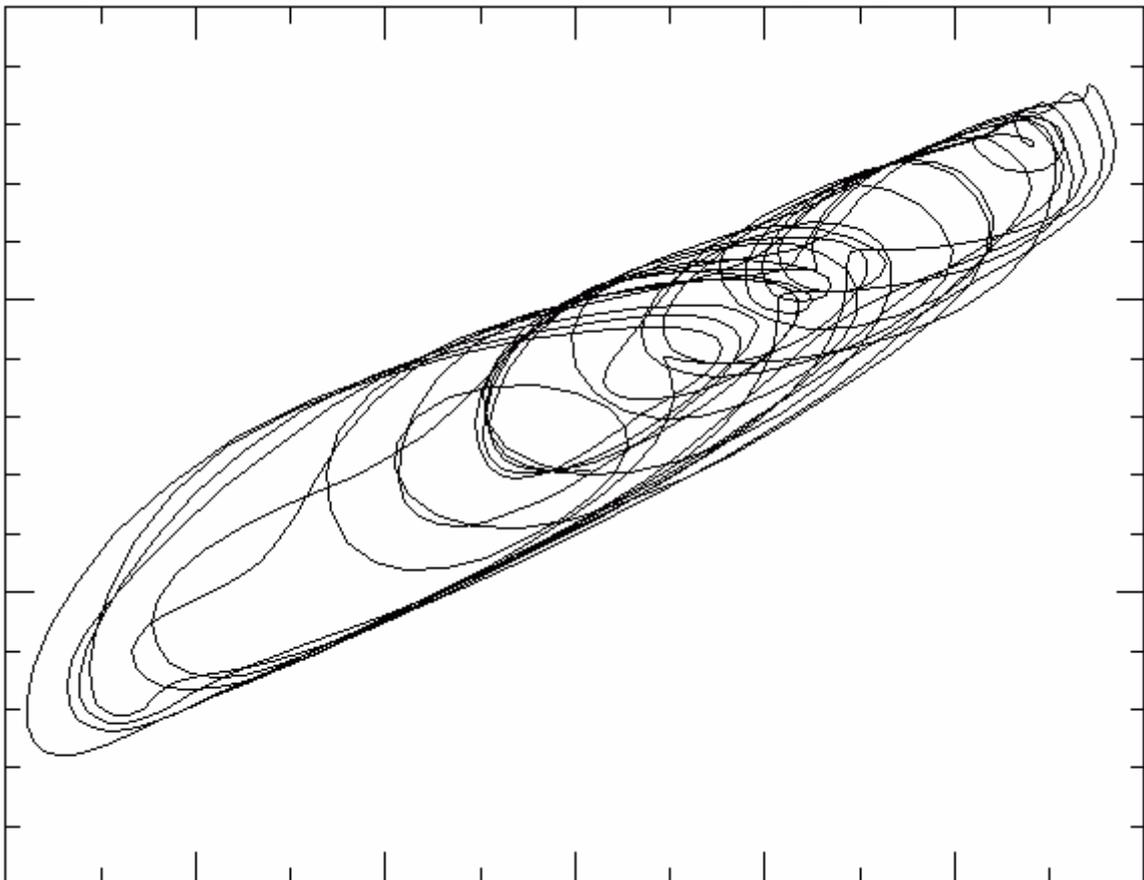


oben: MG30 – 30 Zeitschritte Verzögerung (fraktale Dimension = 2,9)

Mackey-Glass Funktion MG17 – Einbettung durch Zeitverzögerung



Mackey-Glass Funktion MG30 – Einbettung durch Zeitverzögerung



Phasenraumsdarstellungen von Finanzmärkten

Datenbeschaffung

Von den Modellsystemen kann man sich die Daten selbst herstellen. Wenn man die entsprechenden Formeln kennt, kann man beliebig lange Datenreihen erzeugen. Bei Finanzmarktdaten ist die Situation anders. Man braucht eine Datenquelle, möglichst eine kostenlose, bei der man auch historische Datenreihen abrufen kann. Früher war das bisweilen eine teure Angelegenheit, aber inzwischen kann man sich über das Internet kostenlos mit Finanzmarktdaten versorgen. Für aktuelle Daten gibt es eine Vielzahl von Quellen. Aber für lange historische Datenreihen ist das Angebot an kostenlosen Daten sehr überschaubar.

Eine gute Quelle für kostenlose Tages-, Wochen- und Monatsdaten für Aktienmärkte ist Yahoo.Finance. Dort findet man historische Daten zu vielen Aktienmärkten in Nord-, Mittel-, und Südamerika, Europa, Asien, Afrika, Australien und Neuseeland. Die Daten liegen als *.csv Dateien vor und können direkt in EXCEL^(TM) eingelesen werden.

Historische Daten zu Wechselkursen diverser Währungen findet man kostenlos bei der kanadischen Zentralbank, Bank of Canada. Als Vergleichswährung dient jeweils der Kanadische Dollar. Hier kann man zwischen verschiedenen Dateiformaten wählen, aber *.csv ist auch darunter. Möchte man den Wechselkurs von Euro zum Dollar wissen, dann muss man sich zwei Dateien herunterladen, nämlich Euro/Can.\$ und US-\$/Can.\$, jeweils in verschiedene Spalten eines Kalkulationsblatts. Daraus muss man nur noch in jeder Reihe den Quotienten bilden und schon hat man die gewünschten Wechselkurse.

Die hier verwendeten Daten stammen aus den oben angegebenen Quellen. Zu den Wochendaten für Aktienmärkte ist zu bemerken, dass die angegebenen Kurse Montagsschlusskurse sind, nicht die üblichen vom Freitag. Die von mir verwendeten Wochenkurse sind jedoch Freitagsschlusskurse. Ich habe sie aus den Tagesdaten selbst erstellt.

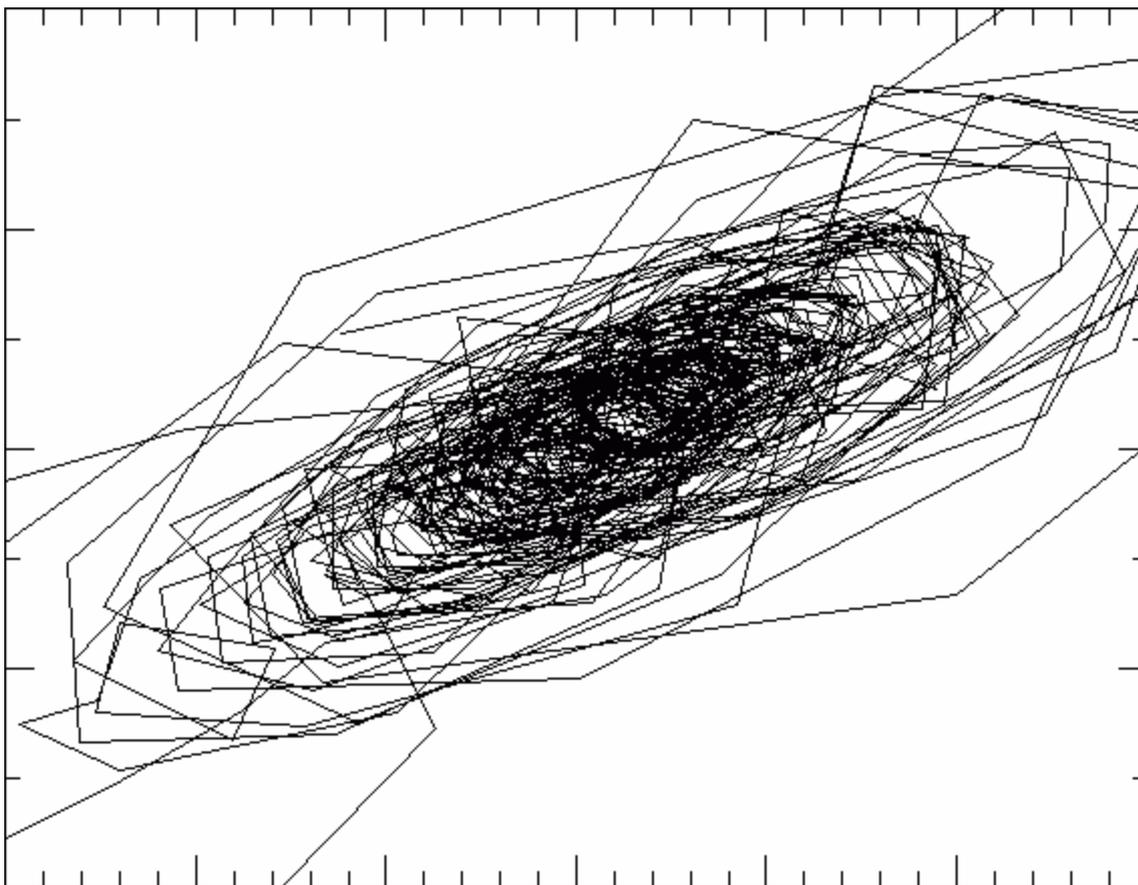
Grafische Darstellung der Marktdynamik

Auf der folgenden Seite sieht man jeweils eine Phasenraumdarstellung eines Aktienmarktes, hier der S&P 500 und eines Devisenmarktes, hier das Währungspaar Euro/US-\$. Beim Aktienindex handelt es sich um geglättete Tagesschlusskurse, bei den Devisen um Wochenschlusskurse. Ungeglättete Kurse würden sehr erratische Kurven ergeben, an denen man kaum Regelmäßigkeiten erkennen könnte.

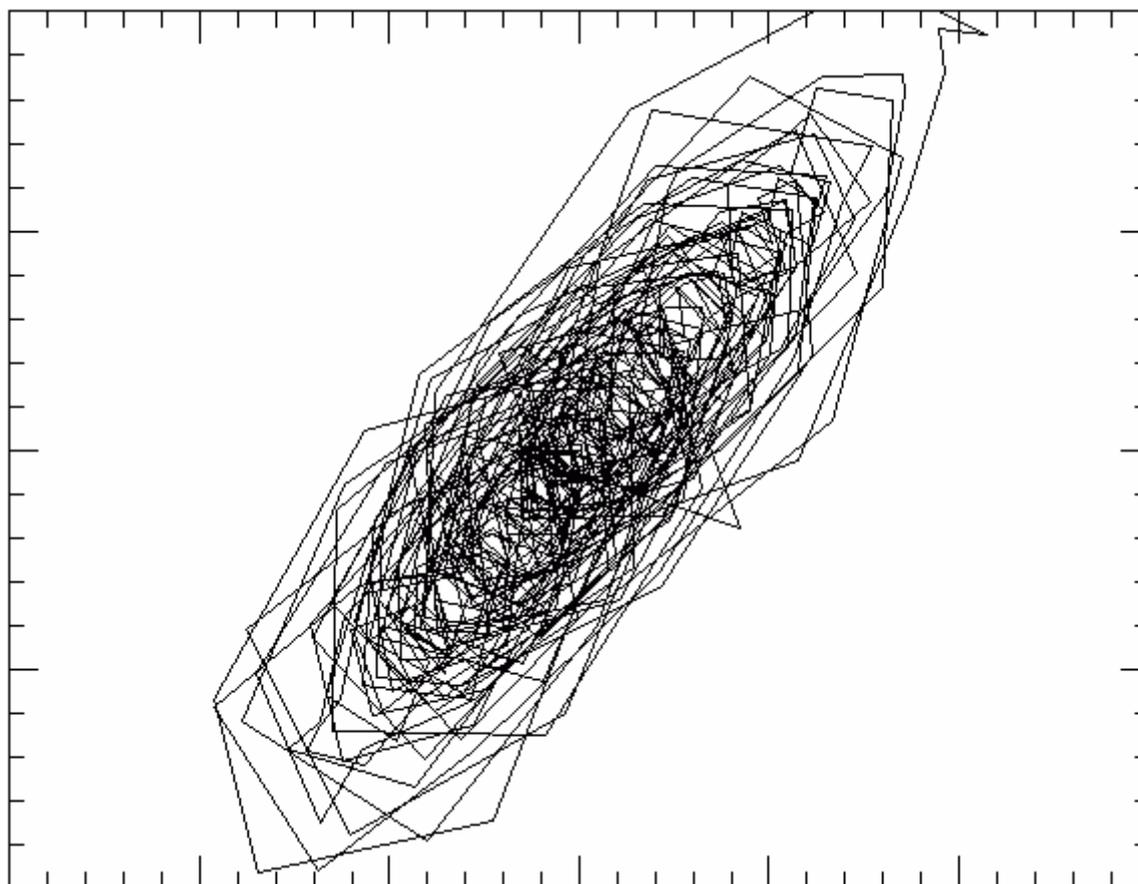
Mit den beiden folgenden Beispielen sollen mehrere Dinge gezeigt werden: erstens, dass man mit realen Marktdaten ähnliche Abbildungen erhalten kann, wie mit chaotischen Modellsystemen und zwar gilt dies für Aktien- und Devisenkurse. Zweitens sollte man erkennen können, dass gerade die verzögerten Differentialgleichungen, die in der breiten Öffentlichkeit weniger bekannt sind, Modellsysteme darstellen, die mit Märkten einige Gemeinsamkeiten haben. Drittens ist zu sehen, dass die Ähnlichkeit zur Funktion MG30 größer ist als zu MG17. Märkte reagieren auf sich selbst, und dies natürlich auch mit Verzögerung. Sie sind reflexiv, wie George Soros sagt. Aber die Reaktionen auf frühere Ereignisse sind zeitlich nicht so bestimmt, wie in den Mackey-Glass Funktionen. Darum sind die Kurven realer Daten zwangsläufig komplizierter, als die der Modellsysteme. Zudem sind Marktdaten auch noch Störungen unterworfen, die zu erratischen Ausschlägen führen.

Die Phasenraumdarstellung kann, wie man sehen wird, eine Marktdynamik veranschaulichen. Dazu sind nicht allein Wochendaten von Kursen geeignet. Auch Tageskurse kann man verwenden, um eine Marktdynamik sichtbar werden zu lassen. Es müssen nicht einmal die Kurse selbst, bzw. die Kursbewegungen, als Daten verwendet werden. Der Verlauf von Volatilitäten, d.h. von Standardabweichungen der Marktbewegungen, lässt sich ebenso mit diesem Mittel verdeutlichen.

Finanzmarktdaten – Einbettungen jeweils durch Zeitverzögerung



oben: amerikanischer Aktienindex (S&P 500), Tagesdaten, geglättet

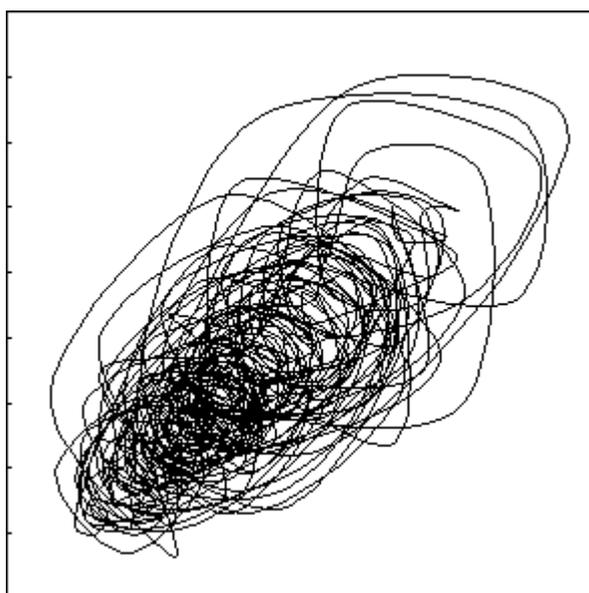


oben: Währungspaar Euro/US-\$, Wochendaten, geglättet

Durch Glättung der Daten kann man die Systemdynamik deutlicher hervor treten lassen, weil Störungen so zumindest teilweise unterdrückt werden können. Als Verhaltensmodell eignet sich eine Attraktorabbildung nur dann, wenn sie keine abrupten Richtungswechsel aufweist. Einige solcher Richtungswechsel sind unvermeidlich und das liegt daran, dass die Darstellung der Dynamik in einer Ebene nicht ausreicht, um ein System mit drei Variablen – das ist das Minimum für chaotische Systeme - korrekt abzubilden. Bahnkrümmungen in der dritten Dimension kann man auf einer Fläche nicht sehen, sondern nur ahnen, weil die dritte Dimension des Systems auf die Ebene projiziert wurde. Dieses Phänomen kann man bei der Darstellung des MG30 Attraktors auf der vorigen Seite gut erkennen, wo in der mittleren Region mehrfach zeltartige Strukturen auftreten.

Man könnte die Systemdynamik problemlos auch in drei Dimensionen darstellen, um die Projektionsartefakte zu vermeiden. Dabei würden auch die Überschneidungen der Bahnen verschwinden, die es eigentlich nicht geben dürfte, weil das System an den Kreuzungspunkten keine eindeutige Fortsetzung hat. Aber Überschneidungen lassen sich meist gut als Projektionsartefakte erkennen. Der Nachteil einer zusätzlichen Dimension wäre, dass eine Darstellung unübersichtlicher und dadurch weniger anschaulich wäre. Zudem sollte man versuchen, eine möglichst niedrige Dimension für ein Verhaltensmodell zu wählen, jedenfalls für die Darstellung von real beobachteten Systemen. Das hat seinen Grund darin, dass die unvermeidlich auftretenden Störungen und Messfehler auch Zufallsschwankungen bewirken. Bei Einbettungen in höheren Dimensionen werden diese immer mehr verstärkt. Die Hinzunahme einer weiteren Dimension bewirkt also keine bessere, sondern im Gegenteil eine undeutlichere Darstellung und Modellierung der Systemdynamik. Für die Modellierung von Märkten bedeutet es, dass man stets versuchen sollte, mit zwei Dimensionen auszukommen. Meine bisherigen Erfahrungen zeigen, dass diese meist ausreichen.

Für die Darstellung der Dynamik von Volatilitäten der Tagesschlusskurse beim amerikanischen S&P 500 Aktienindex, die unten auf der Seite zu sehen ist, erwies sich die zwei-Werte-Glättung der zwei-Tages- Volatilitäten als optimal für eine strukturierte Abbildung. Weder Volatilitäten von drei oder vier Tagen, noch eine Glättung über mehr als zwei Werte brachte eine Verbesserung.



In der obigen Grafik sind zwei-Tages-Volatilitäten des amerikanischen S&P 500 als Phasenraumdarstellung zu sehen. Die Grafik wurde in EXCEL erstellt und ist bemerkenswert glatt. Leider handelt es sich dabei um ein Artefakt, denn das Programm verwendet automatisch Bezier-Funktionen, um die Kurven zu glätten. Verbindet man die aufeinander folgenden Punkte stattdessen mit geraden Linien, dann ist die Darstellung bei weitem nicht so glatt.

Um eine Marktdynamik als Phasenraumdarstellung veranschaulichen zu können, ist ein Grafikprogramm, wie etwa das, was ich für die Abbildungen auf den vorigen Seiten jeweils benutzt habe, nicht erforderlich. Man kann den kompletten Vorgang vom Einlesen der Daten, über das Glätten, die Einbettung und schließlich die grafische Darstellung auch nur mit einem Kalkulationsprogramm vornehmen, wie z.B. mit dem bekannten EXCEL^(TM).

Den allermeisten Lesern ist die Phasenraumdarstellung der Dynamik von Märkten sicher neu und ungewohnt. Vielleicht hat mancher den Eindruck gewonnen, sie sei von mir als Ersatz für die herkömmlichen Darstellungen von Kursverläufen gedacht. Dieser Eindruck wäre falsch, denn wie Sie im Folgenden sehen werden, mache ich von der herkömmlichen Darstellung weiterhin intensiven Gebrauch. Ich verstehe die Phasenraumdarstellung als eine wichtige Ergänzung zu den üblichen 'Charts'. Diese können immer nur einen kleinen Ausschnitt der Marktbewegungen zeigen. Wenn man die Entwicklung über ein Jahr sieht, so ist das schon eine lange Zeitspanne für 'Charts' und die Auflösung wird immer schlechter. Bei längeren Zeitabschnitten kommt der Verzerrungseffekt durch die lineare Darstellung hinzu. Wählt man dagegen die Phasenraumdarstellung, dann kann man die Daten von fünfzig oder mehr Jahren in einer Grafik unterbringen, ohne dass die Auflösung schlechter wird. Und nur mit dieser Art der Darstellung kann man das sogenannte 'Langzeitgedächtnis' eines Marktes sichtbar machen. Nur so kann man auch zeigen, welche Marktbewegungen in der Vergangenheit bereits aufgetreten sind. Mit einer Wiederholung dieser, bzw. ähnlicher Marktentwicklungen ist irgendwann zu rechnen.

Für die Prognoseverfahren, die ich weiter hinten vorstellen möchte und die der Hauptgegenstand dieses Handbuches sind, die lokalen Predikten nämlich, ist eine Phasenraumdarstellung zumindest hilfreich, wenn nicht unverzichtbar. Diese Verfahren sind: 'Nearest Neighbour' Prediktoren, d.h. Vorhersagen mit Hilfe 'Nächster Nachbarn' und als zweites die 'Chaos-Statistik' Prognose, die eine reine Eigenentwicklung von mir ist und daher später auch noch ausführlich erklärt wird. Es gibt noch mehr solcher lokaler Prediktoren. Bekannt sind vielleicht noch die topologischen Karten, auch als 'Kohonen maps' bekannt, denen das Verfahren der sogenannten 'Learning Vector Quantization' (LVQ) zugrunde liegt. Sie werden im allgemeinen den neuronalen Netzen hinzugerechnet, sind aber untypische Sonderformen, wengleich sie eine ganze Reihe vorteilhafter Eigenschaften besitzen. Ich habe sie selbst für die Vorhersage von Märkten ausprobiert und meine Erfahrungen dabei auch beschrieben (Uhlig, H. 1999).

Teil II – Datenanalyse und Modellierung

Ein Beispiel für die Anwendung der folgenden Analysemethoden auf den Deutschen Aktienindex (DAX) wurde von mir im Verlag der Versicherungswirtschaft, Dezemberheft 2001 veröffentlicht (Uhlig, H. 2001b).

Die erste Maßnahme bei der Analyse einer Datenreihe, ist die grafische Darstellung der Daten. Alle Werte werden gegen die Zeit aufgetragen, so dass man einen Eindruck von der Veränderung der Daten im Laufe der Zeit bekommt. Auf den folgenden Seiten sind 1500 Tagesschlusskurse verschiedener Aktienindizes bis Ende Januar 2009 aufgetragen. Es sind die Daten für den deutschen Leitindex DAX, für den österreichischen Leitindex ATX und für den schweizerischen Leitindex SMI. Die obere Grafik zeigt jeweils die Zeitreihe der Indexkurve in linearem Maßstab in hellgrau und in logarithmischem Maßstab in magenta, verwendet wurde der natürliche Logarithmus (e^x). An der linken senkrechten Skala kann man den Indexkurs in Punkten ablesen. Die horizontale Zeitskala reicht von 2003 bis 2009. Die Startpunkte der Skalen sind nicht genau gleich, weil es in den drei Ländern unterschiedlich viele Börsenhandelstage pro Jahr gab, die meisten waren es beim DAX, die wenigsten beim ATX. Alle Zeitreihen beginnen kurz vor dem Tiefpunkt der Indizes nach der Dot.com Krise im Frühjahr 2003.

Bei allen drei Indizes sieht man zunächst einen längeren Aufwärtstrend, der im Jahr 2007 den Höchststand erreicht und danach zuerst langsam, dann schneller zu einem starken Abwärtstrend wird. Als Anhaltspunkt für den langfristigen Trend ist der Durchschnittswert der letzten 200 Handelstage, bzw. der letzten 40 Wochen unter Börsianern recht beliebt. Wenn dieser Durchschnittswert nachhaltig durchbrochen wird, ist das für viele Börsianer das Signal, aus dem Markt zu gehen, d.h. ihre Aktien zu verkaufen. Das geschah bei den hier betrachteten Indizes DAX und ATX Mitte Januar 2008, beim SMI bereits Ende 2007, also lange vor der Pleite der Lehman Bros. Bank, die von vielen als Auslöser der weltweiten Finanz- und Wirtschaftskrise angesehen wird. Der 200 Tage Durchschnittswert von Aktien oder Aktienindizes wird schon seit vielen Jahrzehnten aufmerksam verfolgt und zu Recht, wie man sieht, denn wer sich danach richtet und rechtzeitig seine Engagements am Markt beendet hat, konnte sich den größten Teil der Marktgewinne sichern.

Bei Aktien war man es lange Zeit gewohnt, dass die Kurse ständig stiegen, gelegentlich war dieser Ablauf nur kurz gestört und dann wurde der Trend wieder fortgesetzt. Dies konnte man von 1980 bis zum Jahr 2000 beobachten und zwar weltweit. Damals war die Strategie, Aktien einmal zu kaufen und dann zu behalten beinahe ideal. Die Papiere wurden immer teurer, man bezahlte kaum Transaktionskosten und einen Anlageberater oder Portfoliomanager brauchte man auch nicht. Man betrachtete Aktien einfach als ein Wertpapier mit zwar variabler aber hoher Verzinsung. Der DAX wurde Anfang 1988 etabliert und hatte zum Start 1000 Punkte. Im Frühjahr 2000 stand er kurz bei 8000 Punkten. Wer damals gleich zu Anfang in die DAX-Werte eingestiegen war der hätte das ursprünglich eingesetzte Kapital verachtfacht, sogar ohne Dividenden einzurechnen. Das entsprach einer durchschnittlichen Verzinsung (mit Zinseszins), von über 17% pro Jahr.

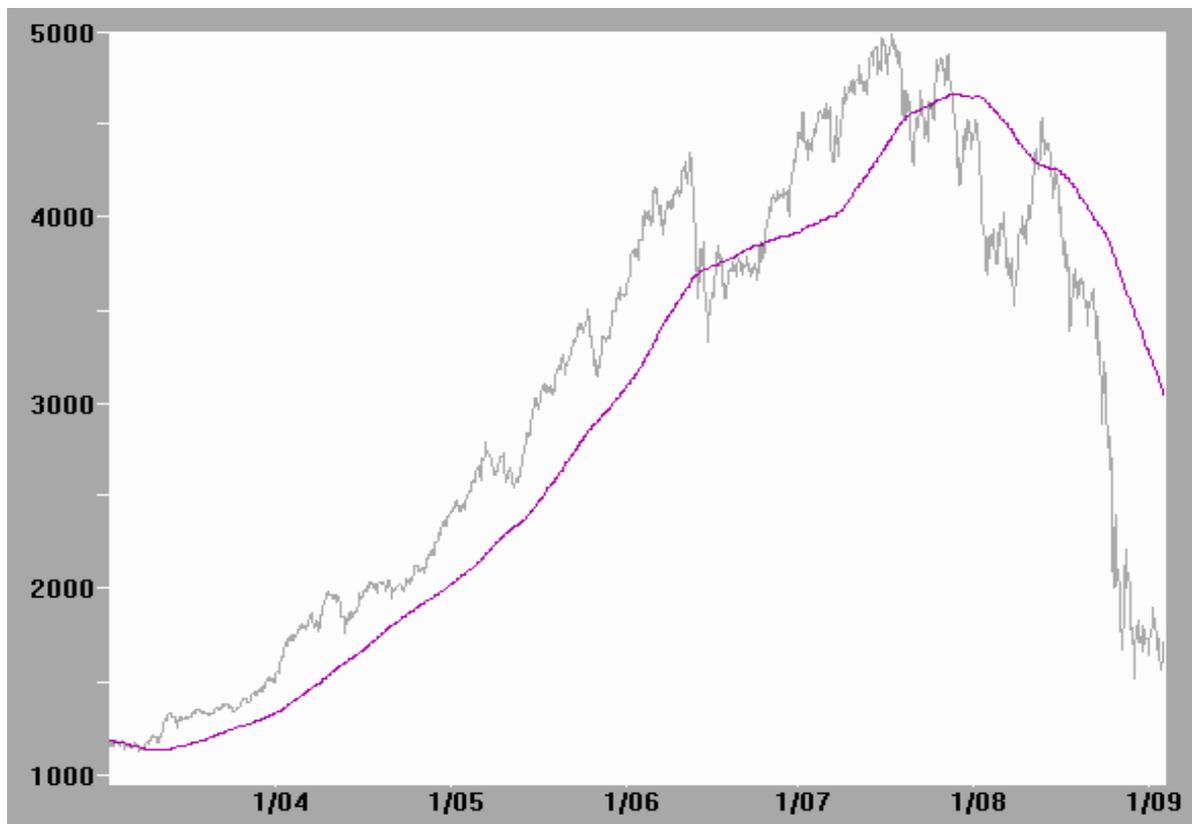
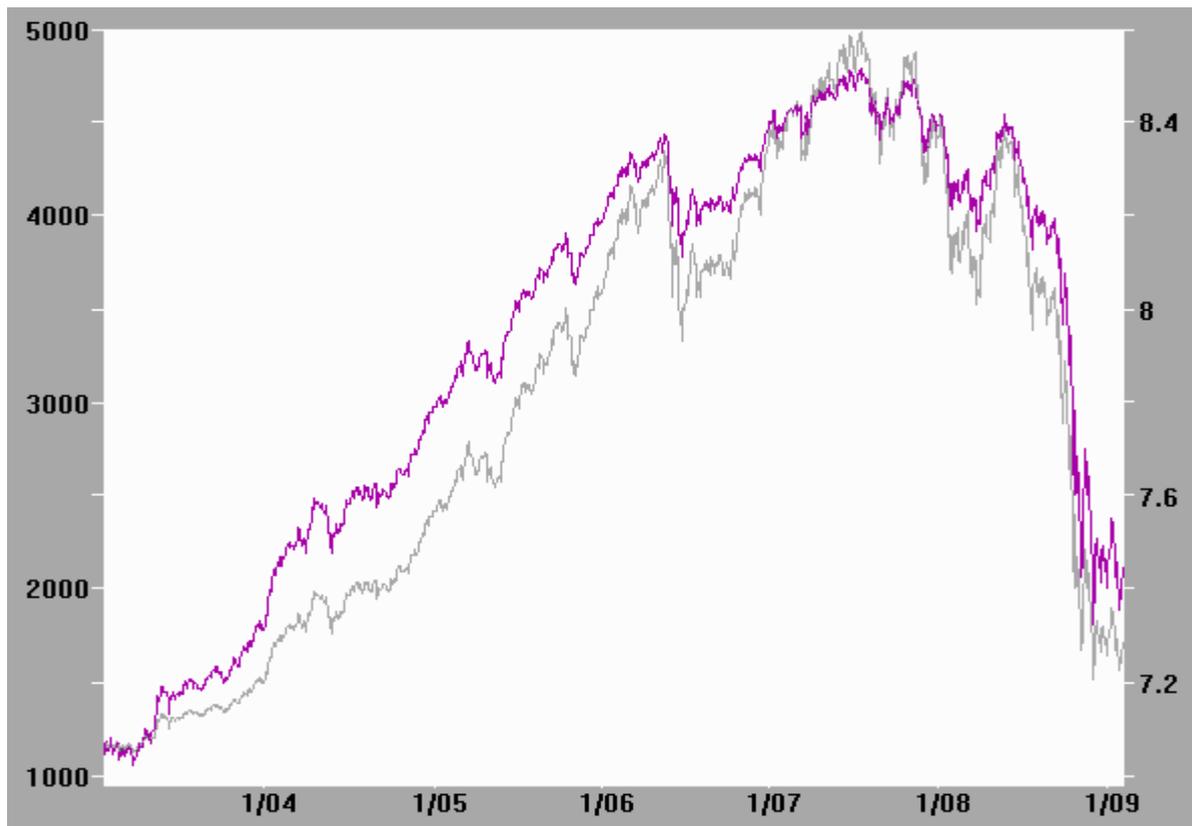
Der DAX verlor danach bis zum Frühjahr 2003 den größten Teil der vorher eingefahrenen Gewinne. Zwischen seinem Höchststand in 2000 und dem Tiefpunkt in 2003 hatte der DAX etwa 70% seines Kurswertes verloren. Ähnlich starke Einbußen musste der schweizerische SMI hinnehmen und den österreichischen ATX traf es noch härter. Derartig starke Schwankungen kannte man bis dahin nur von den Devisenmärkten. Tatsächlich kann man die jüngsten Grafiken der Kurszeitreihen (Charts) von Aktienmärkten kaum von solchen der Wechselkursverläufe unterscheiden. Um zu zeigen, wie ähnlich die Kursentwicklungen der Devisen zu denen von Aktienindizes verlaufen, habe ich einige der wichtigsten Währungen im Vergleich zum Euro für die letzten etwa sechs Jahre grafisch dargestellt. Angegeben ist jeweils der Kurs des Euro in der Fremdwährung, d.h. wie viel in der Fremdwährung für einen Euro bezahlt werden muss. Zudem habe ich den 200 Tage Durchschnitt ebenfalls mit abbilden lassen. Bei den Devisen ist dieser Durchschnitt jedoch zur besseren Unterscheidung in grün eingezeichnet.

Aktienmärkte – Deutschland (DAX)



Obere Grafik: DAX Tagesdaten linear (grau) und logarithmisch (magenta) aufgetragen.
Untere Grafik: DAX Tageskurse (grau) und 200 Tage Durchschnittswerte (magenta).

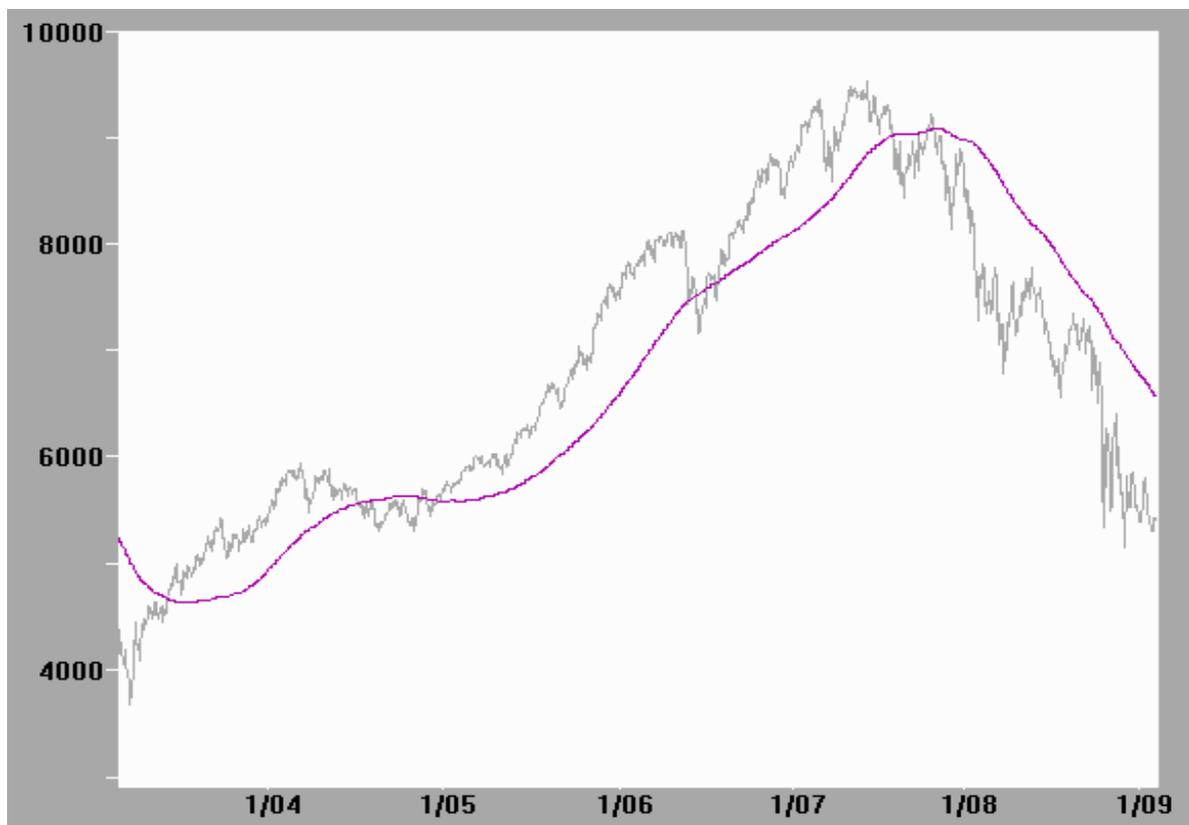
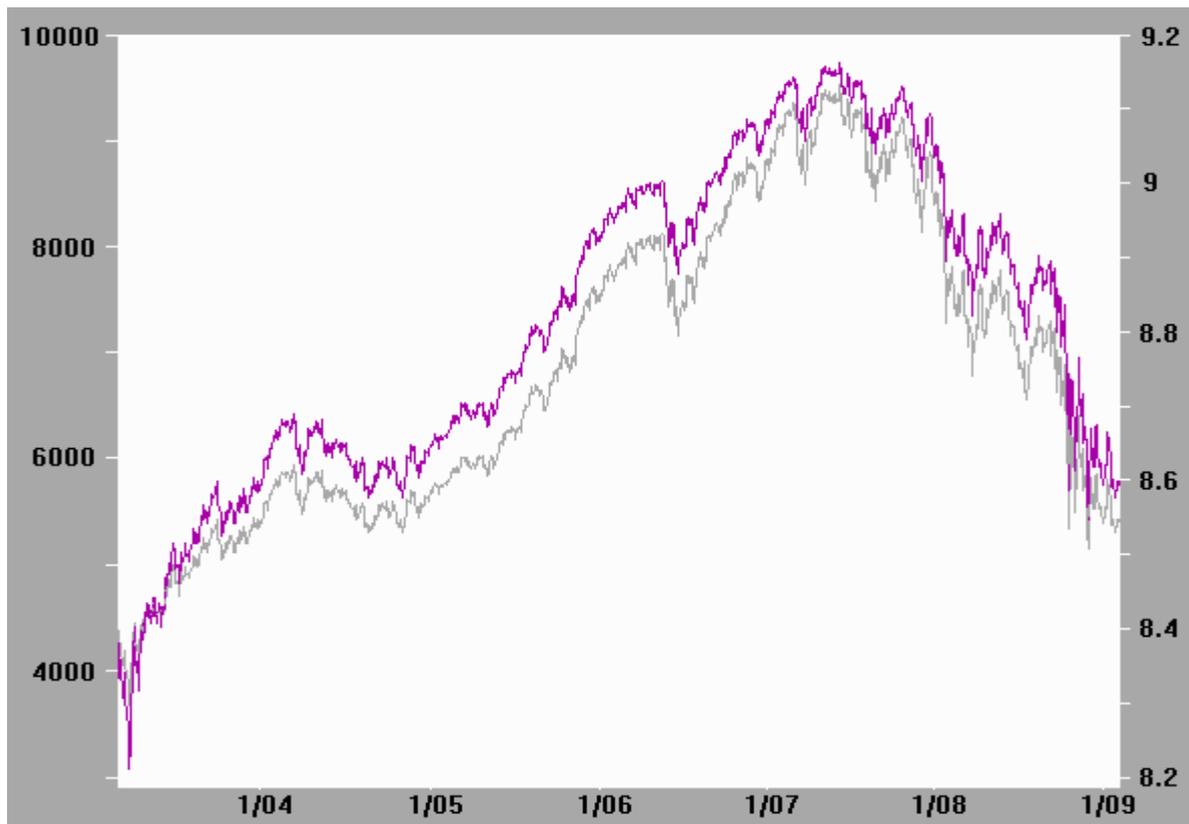
Aktienmärkte – Österreich (ATX)



Obere Grafik: ATX Tageskurse linear (grau) und logarithmisch (magenta).

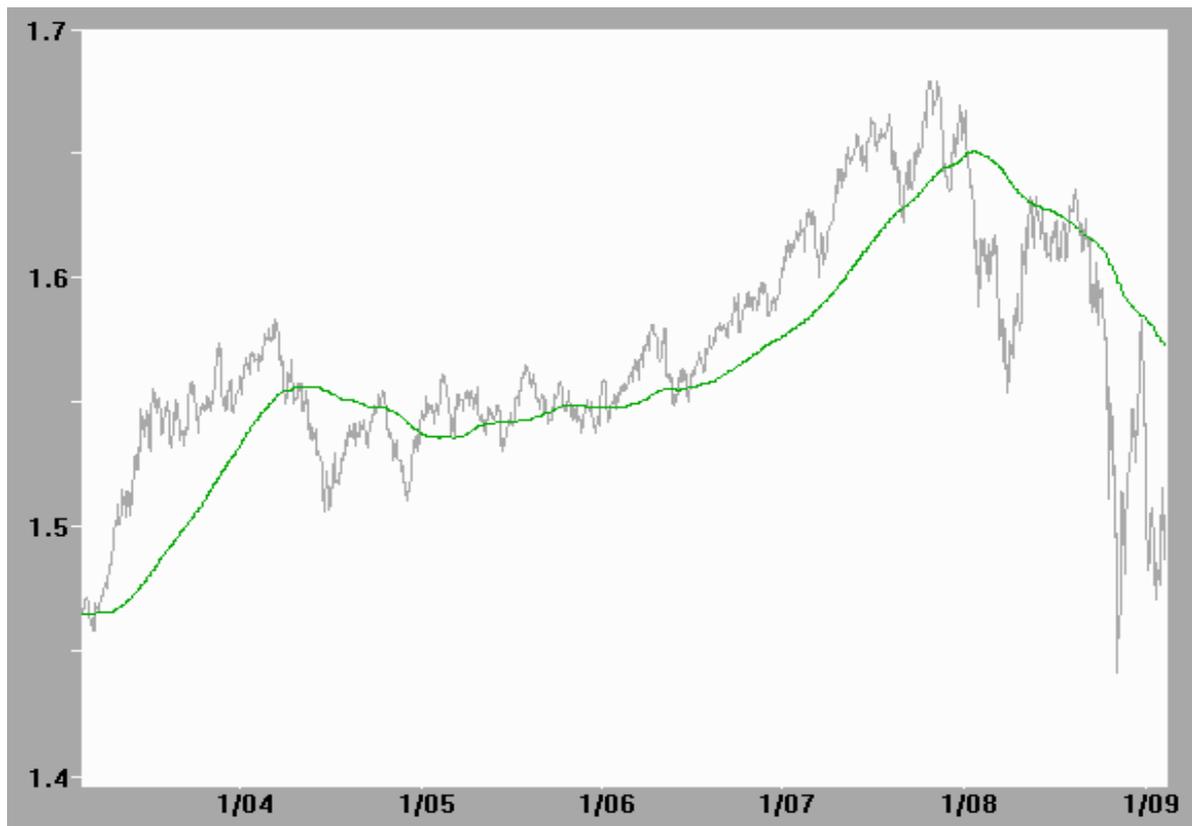
Untere Grafik: ATX Tageskurse linear (grau) und 200 Tage Durchschnittswerte (magenta).

Aktienmärkte – Schweiz (SMI)



Obere Grafik: SMI Tagesdaten linear (grau) und logarithmisch (magenta) aufgetragen.
Untere Grafik: SMI Tageskurse linear (grau) und 200 Tage Durchschnittswerte (magenta).

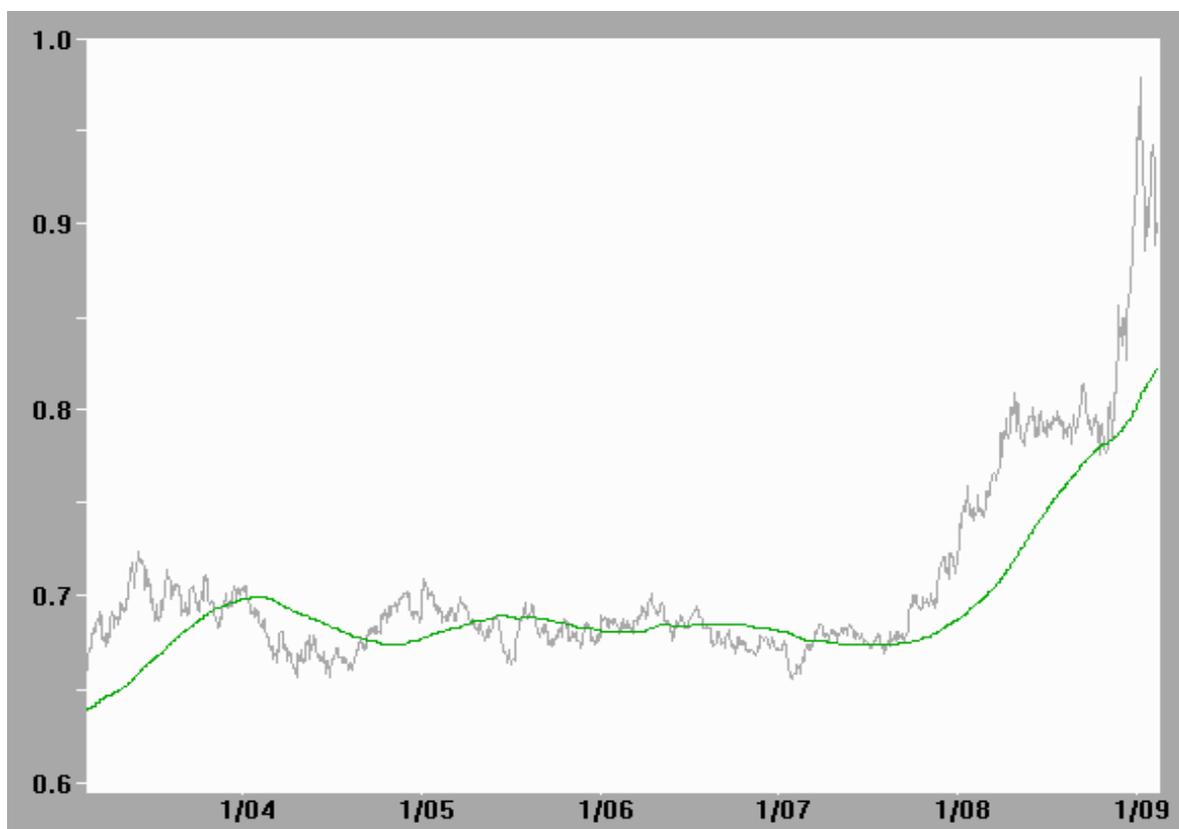
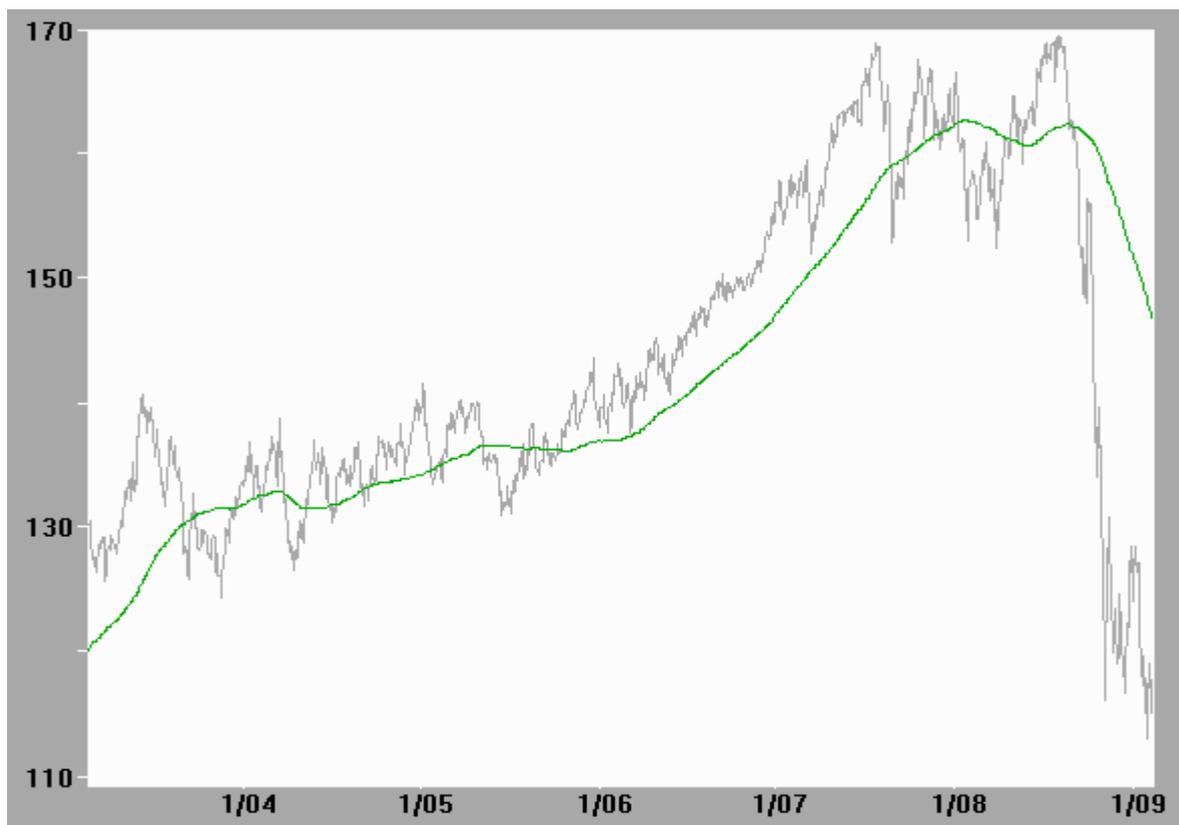
Devisenmärkte – Sfr/Euro, US-\$/Euro



Obere Grafik: Wechselkurs des Schweizer Franken für 1 Euro

Untere Grafik: Wechselkurs des US-\$ für 1 Euro

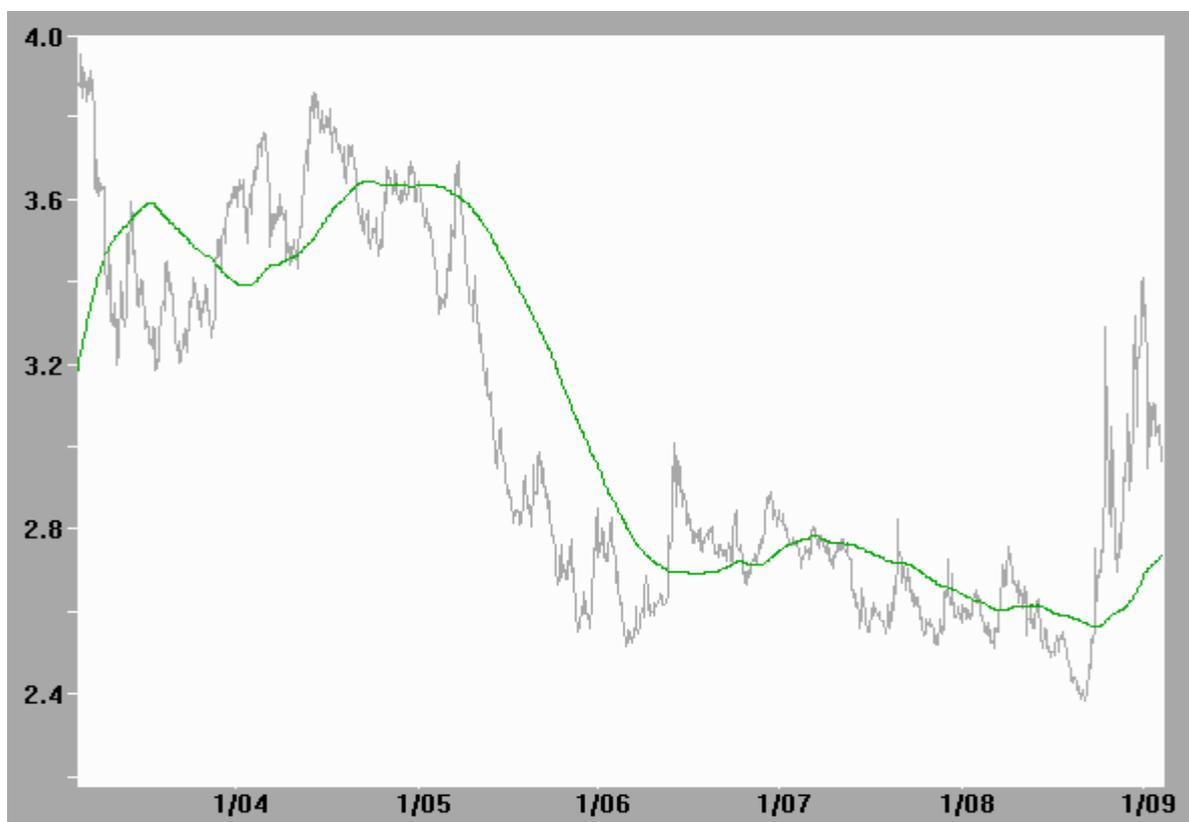
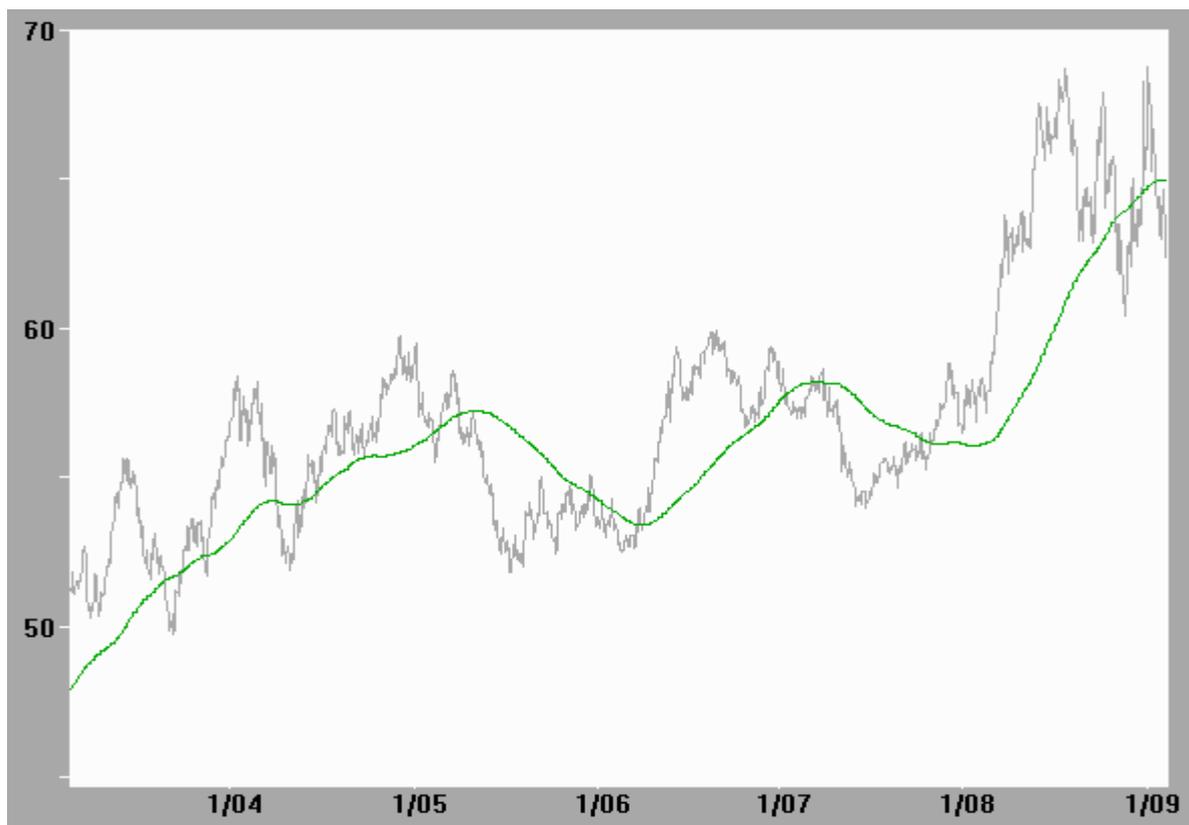
Devisenmärkte – Jap. Yen/Euro, Brit. Pfund/Euro



Obere Grafik: Wechselkurs in Yen für 1 Euro

Untere Grafik: Wechselkurs in Britischen Pfund Sterling für 1 Euro

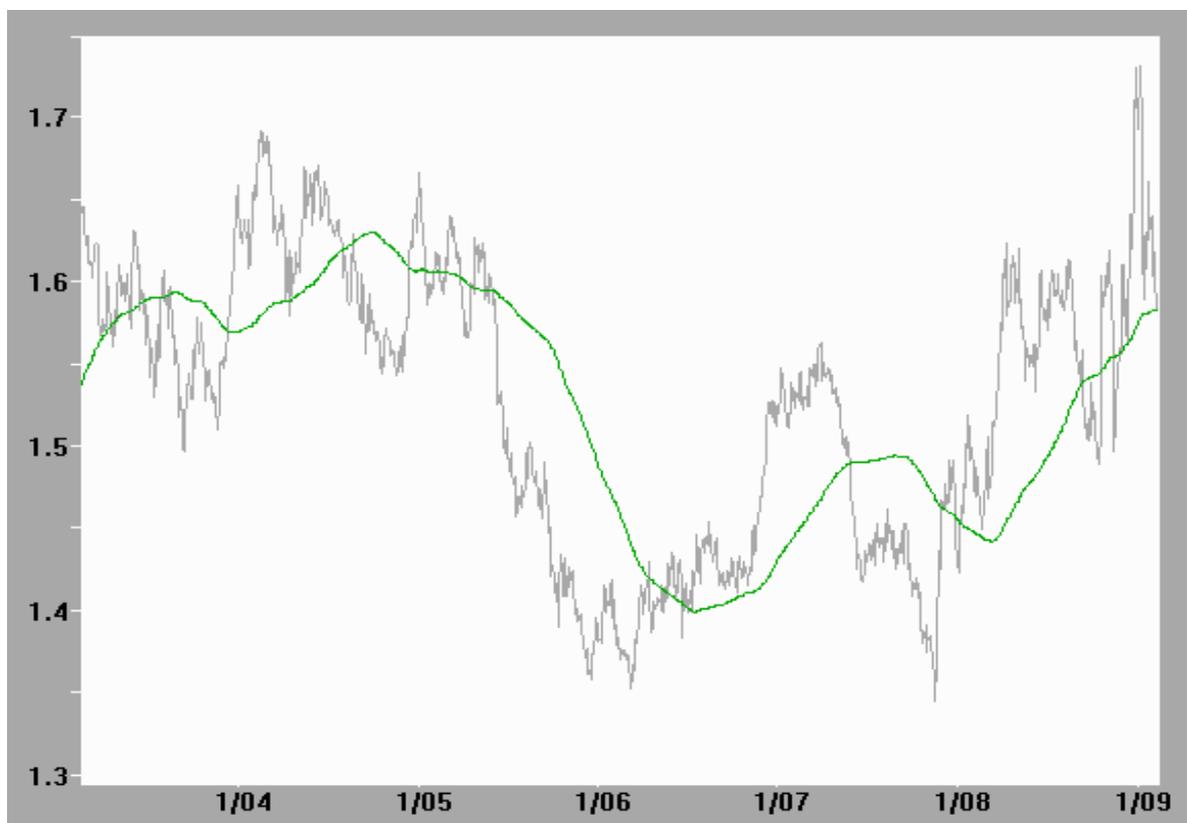
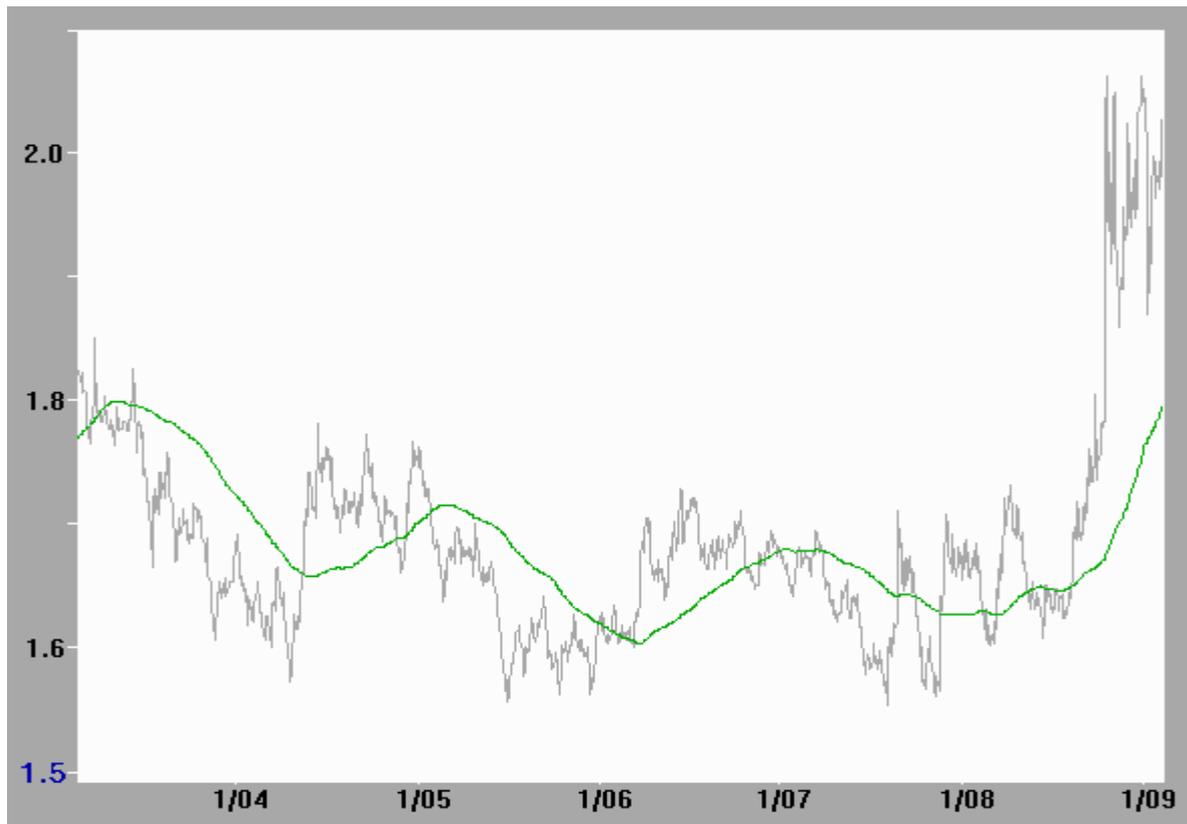
Devisenmärkte – Ind. Rupie/Euro, Bras. Real/Euro



Obere Grafik: Wechselkurs in Indischen Rupien für 1 Euro

Untere Grafik: Wechselkurs in Brasilianischen Real für 1 Euro

Devisenmärkte – Aus-\$/Euro, Kan-\$/Euro



Obere Grafik: Wechselkurs des Australischen \$ für 1 Euro

Untere Grafik: Wechselkurs des Kanadischen \$ für 1 Euro

Besonderheiten der Devisenmärkte

An den Devisenmärkten finden die meisten Umsätze in den Währungen US-Dollar, Euro und Japanischen Yen statt. Danach folgen der Schweizer Franken und das Britische Pfund sowie der Kanadische Dollar und der Australische Dollar. Die Wechselkursverläufe dieser Währungen zum Euro habe ich in verschiedenen Grafiken dargestellt. Von den Währungen der aufstrebenden Schwellenländer, wie China, Indien und Brasilien, sind nur die Indische Rupie und der Brasilianische Real frei handelbar. Der Kurs des Chinesischen Yuan wird von der Regierung des Landes festgelegt. Die Kurse von Rupie und Real zum Euro während der letzten sechs Jahre habe ich ebenfalls in Grafiken dargestellt.

Anders als bei Aktien handelt man bei Devisen immer Währungspaare. Während man die eine Währung verkauft, kauft man gleichzeitig die Gegenwährung. Der Handel ist also bidirektional. Man setzt einmal auf die Kurssteigerung der einen Währung und wenn man diese verkauft und dafür eine andere kauft, setzt man auf die Kurssteigerung bei dieser Währung. Prinzipiell sind beide Währungen gleichwertig. Man könnte auch die Notierung umgekehrt wählen, also statt des Verhältnisses von Währung A / Währung B den Quotienten von Währung B / Währung A. Dann würde man einen spiegelbildlichen Verlauf bekommen. Man ist praktisch immer im Markt und muss sich für eine Seite des Marktes entscheiden. Dieser Unterschied zum Aktienmarkt zeigt sich auch deutlich am Kursverlauf. Die Aufwärtsbewegungen und Abwärtsbewegungen der Märkte sind wesentlich symmetrischer als bei Aktien bzw. Aktienindizes. Anders als beim Aktienmarkt spielt der 200 Tage Mittelwert als Orientierung für den Zustand des Marktzyklus bei Devisen keine Rolle. Die Häufigkeitsverteilungen extremer Marktbewegungen sind gleich, ob der Kurs oberhalb oder unterhalb des 200 Tage Mittels liegt.

Auf den Devisenmärkten wird rund um die Uhr und überwiegend kurzfristig gehandelt. Die Engagements sind meist kürzer als 10 Tage. Die kurzen Erwartungshorizonte der Anleger bewirken auch, dass die wesentliche Dynamik im Bereich von 10 Tagen stattfindet. Darum kann man mit Monatsschlusskursen als Datenquelle praktisch keine Handelssysteme aufstellen. Man muss mindestens Wochendaten, besser Tagesdaten verwenden. An den Devisenmärkten wird pro Tag 40 mal mehr Kapital gehandelt als auf allen Aktienmärkten zusammen. Liquidität ist jedenfalls bei den Hauptwährungen kein Problem. Der Handel geschieht elektronisch und findet größtenteils zwischen Banken statt.

Wie man an den Grafiken erkennt, sind die Bewegungen viel erratischer, wenngleich die Schwankungsbreite je Zeiteinheit deutlich geringer ist, als bei Aktien. Es sieht zwar so aus, als gäbe es bei den Devisen mehr und stärkere Ausschläge in die eine oder die andere Richtung, doch das ist nur ein Darstellungsartefakt, weil bei den Devisen im Vergleich zu den Aktien der Kursverlauf praktisch durch ein Vergrößerungsglas betrachtet wird.

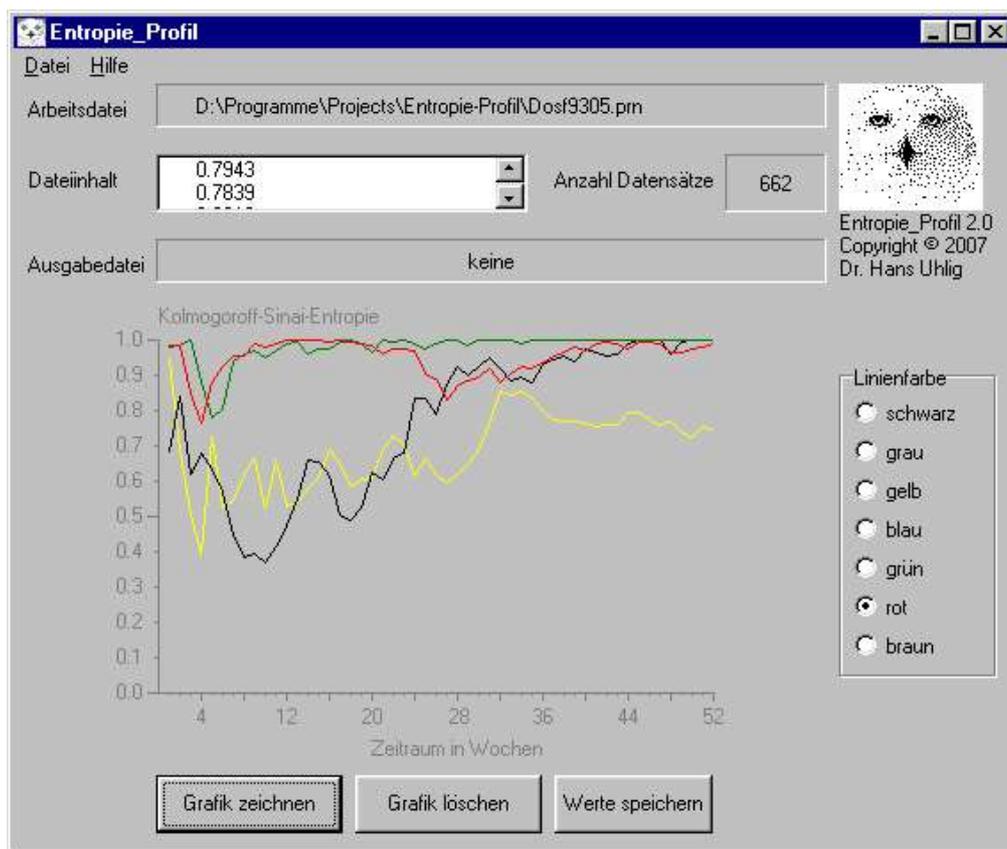
Die Kursnotierungen der Währungspaare sind auf vier signifikante Dezimalstellen genau. Beim Kurs US-\$/Euro sind es vier Nachkommastellen und bei Yen / US-\$ und Yen /Euro sind es zwei Nachkommastellen. Die letzte Dezimalstelle legt fest, welches die kleinste feststellbare Kursdifferenz, genannt 1 BIP, ist. Es gibt Kauf- und Verkaufskurse und diese differieren um einige BIPs. Bei den Hauptwährungen sind es oft nur zwei oder drei BIPs, aber beim Handel von kleinen Nebenwährungen kann die Differenz (bid/ask spread) ein Vielfaches davon sein. Für Privatinvestoren sind die Bedingungen schlechter, als für die mächtigen Marktteilnehmer. Berücksichtigt man dazu die eher kurzfristige Orientierung, die häufiges Handeln erzwingt, dann bedeutet dies insgesamt relativ hohe Transaktionskosten. Wenn man also auf dem Devisenmarkt nicht nur sein Fremdwährungsrisiko absichern will, sondern beabsichtigt dort richtig Geld zu verdienen, der muss über ein wirklich gutes Handelssystem verfügen und einen der wenigen realen Broker in diesem Markt finden. Nach Ansicht eines Kenners des Devisenmarktes ist das Angebot an geeigneten Plattformen auf diesem Gebiet sehr überschaubar

Die Güte von Marktvorhersagen - Messen der (Informations)-Entropie

Eine wichtige, allgemeine Information über einen Markt kann uns die Entropie liefern. In den Börsenzeitschriften und den Wirtschaftsteilen der Zeitungen finden Sie dazu nichts, dafür jedoch um so mehr nichtssagende, aber bunte 'Charts'.

Entropie ist ein Maß für die Unsicherheit von Informationen. Um sie von der physikalischen Entropie zu unterscheiden, die man aus dem zweiten Hauptsatz der Thermodynamik kennt, wird sie auch Informationsentropie genannt. Der Begriff stammt aus der Informationstheorie und hat mit der relativen Wahrscheinlichkeit von Ereignissen zu tun. Je gleicher die Wahrscheinlichkeiten für verschiedene mögliche Ereignisse sind, um so größer ist unsere Unsicherheit. Außerdem wächst unsere Unsicherheit mit der Anzahl der Möglichkeiten. Gibt es nur zwei mögliche, aber gleich wahrscheinliche Ereignisse, dann ist die Wahrscheinlichkeit, zufällig das richtige zu erraten 50%. Gibt es aber beispielsweise vier gleich wahrscheinliche Möglichkeiten, dann ist die Wahrscheinlichkeit, die richtige zu erraten, nur 25%. Unsere Unsicherheit ist also doppelt so hoch, wie im vorherigen Beispiel.

Auf die Vorhersage von Marktbewegungen übertragen, heißt das: Es ist einfacher die richtige Aussage zu treffen, wenn gefragt wird, in welchem von zwei Wertebereichen die nächste Marktbewegung liegen wird, was einer 1 bit (**b**inary **d**igit) Entscheidung entspricht, oder ob man zwischen vier verschiedenen möglichen Ausgängen entscheiden müsste, beispielsweise indem man noch die Stärke der Bewegung in jeweils zwei Bereiche einteilt. Denn dieses wäre eine 2 bit Entscheidung. Die Entropie wird daher üblicherweise in Einheiten pro bit angegeben. Sie kann Werte zwischen 0 und 1 annehmen.



Die 1 bit Entscheidung kann sehr verschiedene Werte annehmen, je nachdem, wie man die Bereiche definiert. Ich gebe ein paar Beispiele: 1.) steigt der Markt oder fällt er 2.) Ist die kommende Bewegung größer oder kleiner als die durchschnittliche Marktbewegung 3.) Ist die Marktbewegung dichter an der stärksten Aufwärtsbewegung als an der stärksten Abwärtsbewegung.

Die Entropie ist minimal, (null) wenn ein Ereignis sicher eintreten wird und alle anderen

unmöglich sind. Die Entropie ist maximal, (eins) wenn alle möglichen Ereignisse gleich wahrscheinlich sind.

Ich habe Computerprogramme geschrieben, die ein Entropieprofil erstellen, d.h. Entropien/Bit für verschiedene Zeitabschnitte zwischen einem und 52 Zeitschritten (ein Jahr) berechnen. Die nicht kommerzielle Version ist ein Konsolenprogramm. Es zeigt die Werte tabellarisch zunächst nur am Bildschirm. Doch man kann sich die Tabelle auch in eine Datei ausdrucken lassen. Die kommerzielle Version ist ein MS-Windows^(R) Programm. Sie gibt den zeitlichen Verlauf grafisch aus. Ausführliche Informationen zum kommerziellen Programm können Sie dem dazu gehörigen Begleittext entnehmen. Beide Programme arbeiten mit einer Auflösung von 1 bit, d.h., sie unterscheiden jeweils zwei verschiedene Zustände, aber sie können drei verschiedene Kriterien für die 1 bit Entscheidung anlegen. Kriterium I : Kurs steigt , bzw. bleibt gleich / Kurs fällt. Kriterium II: Kursänderung ist näher am Maximum / Minimum. Kriterium III: Kursänderung ist größer / kleiner als der Mittelwert.

Die Ergebnisse sind meistens sehr ernüchternd. Wenn man fragt, ob die Kurse steigen oder fallen werden, gibt es nur eine schwache langfristige Tendenz bei den darauf hin untersuchten Aktienmärkten (DAX, SMI, ATX). Ähnlich ist es bei der Frage, ob die Kursänderungen über oder unter dem Mittelwert liegen. Bei den Devisenmärkten (Euro/US-\$, Sfr/US-\$) ist die Unsicherheit praktisch vollkommen. Es gibt keine kurzfristige oder langfristige Tendenz. In diesem Zusammenhang ist es auch unerheblich, ob man mit Wochendaten oder mit Tagesdaten arbeitet, denn man erhält sehr ähnliche Muster.

Richtung, Stärke und Häufigkeit von Marktbewegungen

Nur wenn man fragt, ob die kommende Marktänderung dichter am Höchstwert als am Tiefstwert liegt, erhält man für gewisse Zeiträume deutlich vom Zufall verschiedene Ergebnisse. Nun, was macht diese Frage so verschieden von den anderen, dass es so unterschiedliche Antworten darauf gibt und wichtiger noch ist: was kann man mit dieser Information anfangen? Bevor man eine schnelle Antwort darauf gibt, sollte man fragen, wie die Marktbewegungen qualitativ und quantitativ verlaufen. Die üblichen 'Charts' zeigen immer nur deutlich die Summe der qualitativen Marktbewegungen. In diese Summe gehen aber nicht nur die Richtung, sondern auch die Stärke der Bewegung und die Häufigkeit in das Ergebnis ein. Oft ist aber die Richtung gar nicht so entscheidend, sondern die Stärke ist es. Wenn die Ausschläge nach oben deutlich höher sind, als die Ausschläge nach unten, dann kann die Häufigkeit der Bewegungen in die eine oder andere Richtung gleich sein und dennoch gibt es eine klare Vorzugsrichtung. Genau das beobachtet man in kurzfristigen bis mittelfristigen Trendphasen. Die richtungsgebenden Bewegungen sind stark und die Korrekturen sind schwach. In Aufwärtsphasen sind die einzelnen Ausschläge nach oben stärker und in Abwärtsphasen sind die einzelnen Ausschläge nach unten stärker. Die weiter oben gestellten Fragen kann man daher so beantworten: Der Unterschied besteht darin, dass die Stärke der Bewegungen berücksichtigt wird und diese Information ist wichtig, um die durchschnittliche Stärke und die durchschnittliche Dauer von kurz- bis mittelfristigen Trends am Markt zu erkennen.

Weitere Analysen zu Ordnung und Zufall bei Marktentwicklungen

Neben dem Programm für das Entropieprofil habe ich noch weitere Programme geschrieben, Auskünfte über die Unsicherheit von Prognosen auf Finanzmärkten geben können. Es sind die Programme für bedingte Entropie und für den Chi-Quadrat-Test auf Abhängigkeit. Sie liefern Auskünfte anderer Art. Das Entropie-Profil gibt an, in welchem Ausmaß die Summe der vorherigen Daten Aussagen über die künftige Entwicklung zulässt. Die anderen Programme geben Auskunft darüber, welche einzelnen Daten, in welchem Zeitabstand, wie wichtige Informationen zur künftigen Entwicklung liefern können. Diese Angaben sind bedeutsam, wie wir sehen werden, für die Auswahl von Datenmustern, die eine Marktsituation charakterisieren können. Solche Muster werden für die 'Nearest Neighbour' – Prognose, also die Vorhersage aufgrund von sogenannten 'nächsten Nachbarn' benötigt.

Bedingte Entropie

Wie sehr verringert sich unsere Unsicherheit über die kommende Marktbewegung, wenn wir wissen, wie sich der Markt vor 1..43 Zeitschritten bewegt hat.

Dateien 1: dax9305.prn 2: dax9305.prn 618-661 Daten.

delay	entropy	cond.entrp.	mut.inf.	delay	entropy	cond.entrp.	mut.inf.
0	0.6845	0.0000	1.0000	1	0.6846	0.6846	0.0001
2	0.6848	0.6816	0.0046	3	0.6846	0.6836	0.0014
4	0.6848	0.6845	0.0003	5	0.6845	0.6832	0.0020
6	0.6847	0.6788	0.0086	7	0.6849	0.6847	0.0003
8	0.6850	0.6849	0.0002	9	0.6848	0.6846	0.0004
10	0.6850	0.6834	0.0023	11	0.6852	0.6850	0.0002
12	0.6849	0.6845	0.0007	13	0.6847	0.6837	0.0015
14	0.6849	0.6849	0.0000	15	0.6847	0.6844	0.0004
16	0.6848	0.6842	0.0009	17	0.6850	0.6842	0.0012
18	0.6848	0.6848	0.0000	19	0.6846	0.6844	0.0002
20	0.6843	0.6843	0.0000	21	0.6845	0.6841	0.0007
22	0.6843	0.6800	0.0062	23	0.6845	0.6839	0.0008
24	0.6846	0.6846	0.0000	25	0.6848	0.6848	0.0000
26	0.6850	0.6847	0.0004	27	0.6852	0.6850	0.0003
28	0.6849	0.6812	0.0055	29	0.6851	0.6844	0.0010
30	0.6853	0.6852	0.0001	31	0.6855	0.6849	0.0009
32	0.6852	0.6849	0.0005	33	0.6854	0.6852	0.0003
34	0.6856	0.6856	0.0000	35	0.6854	0.6853	0.0001
36	0.6855	0.6852	0.0004	37	0.6853	0.6809	0.0065
38	0.6851	0.6851	0.0000	39	0.6853	0.6842	0.0015
40	0.6850	0.6849	0.0001	41	0.6852	0.6848	0.0006
42	0.6854	0.6853	0.0001	43	0.6856	0.6856	0.0000

Chi-Quadrat-Test

Ist die Abfolge der Marktbewegungen zufällig? Je höher der Wert von Chi-Quadrat ist und je niedriger der Fehler, desto geordneter sind die Daten. Fehler < 0.05 = nicht zufällig.

Datei: dax9305.prn

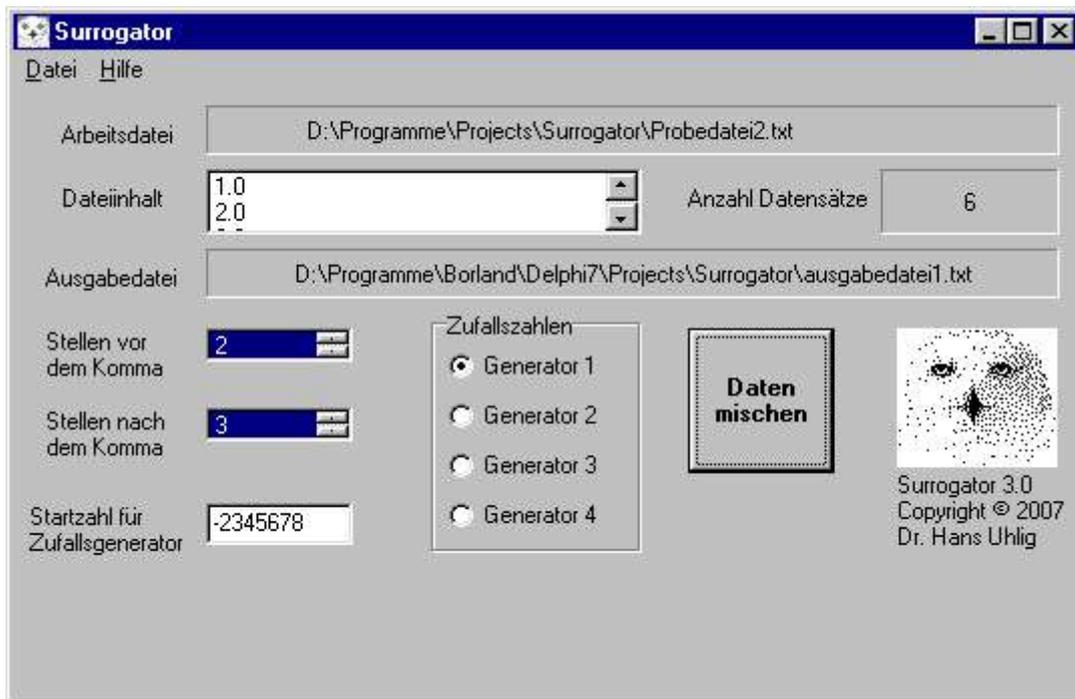
tau	10*ChiQ	Freigr.	Fehler	Cramer-U	tau	10*ChiQ	Freigr.	Fehler	Cramer-U
0	6610	1	0.0000	1.0000	1	41	1	0.0420	0.0792
2	55	1	0.0190	0.0914	3	29	1	0.0859	0.0669
4	37	1	0.0542	0.0751	5	2	1	0.6578	0.0173
6	74	1	0.0066	0.1062	7	3	1	0.6054	0.0202
8	3	1	0.6086	0.0200	9	22	1	0.1398	0.0578
10	48	1	0.0286	0.0858	11	2	1	0.6261	0.0191
12	1	1	0.7198	0.0141	13	30	1	0.0812	0.0685
14	0	1	0.9691	0.0015	15	12	1	0.2764	0.0428
16	33	1	0.0714	0.0710	17	2	1	0.6662	0.0170
18	11	1	0.2979	0.0411	19	1	1	0.7647	0.0118
20	8	1	0.3625	0.0360	21	1	1	0.7047	0.0150
22	159	1	0.0001	0.1575	23	1	1	0.8211	0.0090
24	4	1	0.5447	0.0240	25	1	1	0.7597	0.0121
26	17	1	0.1926	0.0517	27	0	1	0.9354	0.0032
28	91	1	0.0025	0.1202	29	6	1	0.4557	0.0297
30	12	1	0.2696	0.0440	31	16	1	0.2110	0.0498
32	8	1	0.3753	0.0353	33	6	1	0.4489	0.0302
34	8	1	0.3660	0.0361	35	0	1	0.9276	0.0036
36	1	1	0.7836	0.0110	37	63	1	0.0124	0.1001
38	0	1	0.8830	0.0059	39	2	1	0.6415	0.0187
40	10	1	0.3077	0.0409	41	0	1	0.9744	0.0013
42	2	1	0.6899	0.0160	43	0	1	0.9359	0.0032

Die Verwendung von Surrogatdaten

Da der Informationsgehalt der Daten, wie wir oben gesehen haben mit der Zeit abnimmt und irgendwann ein Plateau erreicht wird, auf dem sich die Werte einpendeln, kann man daraus schließen, dass dieses Plateau den Informationsgehalt des ungeordneten Zustands wiedergibt. Interessant wäre jedoch ein Vergleich mit Daten, von denen man weiß, dass sie ungeordnet sind. In der Statistik wird oft die Normalverteilung als Vergleich herangezogen, um festzustellen, ob ein Ereignis zufällig stattgefunden hat, oder ob es dafür einen besonderen Grund gegeben haben muss. Dieser Vergleich mit normalverteilten Zufallswerten ist aber nur dann zulässig und sinnvoll, wenn unsere zu prüfenden Daten tatsächlich normalverteilt sind. Wenn diese Voraussetzung nicht gegeben ist, dann kann der Vergleich wertlos sein. Für eine Reihe von Marktdaten habe ich gezeigt, dass diese nicht normalverteilt sind (Uhlig, H. 1999). Somit sind normalverteilte Daten nicht geeignet für die Prüfung der Nullhypothese (Die Nullhypothese bezeichnet die Annahme, dass für die Abfolge der Daten kein weiterer Einfluss als der Zufall angenommen werden muss und zwar mit mindestens 95% Wahrscheinlichkeit).

Mit Surrogatdaten zur Nullhypothese

In diesem Falle bieten sich sogenannte Surrogatdaten an. Wenn man die Differenzen aufeinander folgender Daten bestimmt und nur die Reihenfolge dieser Änderungen anders ordnet, bzw. mischt, dann ändert man an der Verteilung der Daten nichts: Mittelwert, Schiefe, Streuung, Standardabweichung, alles statistische Kenngrößen, bleiben gleich. Anfangs- und Endwert bleiben gleich und ebenso ein eventuell vorhandener Trend in den Daten. Mit Hilfe guter Zufallszahlen-Generatoren kann man die Daten so mischen, dass sie von zufällig erhaltenen nicht zu unterscheiden sind. Von den Surrogatdaten erzeugt man vorzugsweise mehrere.



Die Abbildung zeigt das von mir entwickelte Programm SURROGATOR, das beliebige Surrogatdaten erzeugen kann. Die ausführliche Beschreibung findet man im Begleittext zum Programm. Es verwendet viel bessere Zufallszahlengeneratoren als die, die man in Kalkulationsblättern oder in den Programmiersprachen standardmäßig verwendet. Drei der vier Generatoren sind portierbar, d.h. sie liefern auf verschiedenen Computern die gleichen

Surrogatreihen. Die Startzahl (seed-Wert) für den Generator kann man vorgeben und so reproduzierbar verschiedene Surrogatreihen erstellen.

Der Begriff Surrogat bedeutet so viel wie Ersatz. Ein solcher Ersatz entspricht nicht genau dem Vorbild, doch er besitzt möglichst viele ähnliche Eigenschaften. Auch auf Surrogatdaten soll dies zutreffen. Die angestrebte Ähnlichkeit kann man auf verschiedene Weisen erreichen und so kennt man auch unterschiedliche Verfahren, diese zu erzeugen. Die Resultate sind ebenfalls verschieden. Gemeinsam ist den Surrogatdaten, dass sie die ursprüngliche Reihenfolge der Daten nicht bewahren, aber dass die statistischen Eigenschaften der Daten erhalten bleiben. Es gibt Verfahren zur Herstellung von Surrogatdaten, die zunächst eine Fouriertransformation der Daten durchführen, dann die Phaseninformation tauschen und schließlich eine inverse Fouriertransformation durchführen. Bei anderen Verfahren wird ebenfalls auf Grundlage der Fouriertransformation ein Phasenspektrum erzeugt und dann die Wahrscheinlichkeit für die jeweiligen Frequenzen bestimmt. Anschließend erzeugt man zufällige Frequenzwerte mit den gleichen Wahrscheinlichkeiten und führt schließlich die inverse Fouriertransformation durch.

Teil III – Marktvorhersagen mit lokalen Prediktoren

Vorhersage mit 'Nächsten Nachbarn' - Nearest Neighbour Prediktoren

Wir haben weiter vorn gesehen, dass die Marktdaten, d.h. die Originaldaten Information enthalten. Durch unvorsichtiges, willkürliches Glätten könnte man die vorhandene Information verwischen und somit zumindest teilweise verlieren. Wir wissen zwar aus den Testergebnissen, dass die Daten untereinander verknüpft sind, doch wir haben keinen Hinweis darauf, wie diese Verknüpfung aussieht.

Vorhersagen mit ungeglätteten Originaldaten

Ein Verfahren zu verwenden, das mit ungeglätteten Daten umgehen kann, ist daher vorteilhaft. Das Verfahren müsste die Daten auswerten können und sie für weitergehende Entscheidungen beurteilen, das heißt daraus konkrete Aussagen über deren Bedeutung ableiten. Die im folgenden beschriebene Methode der 'Nearest Neighbours' Prognose ist ein solches Verfahren.

Das Rauschen erhöht jedoch, wie man festgestellt hat, die Dimension des Systems, was sich in einem zusätzlichen Bedarf an Daten für jedes Datenmuster äußert. Dadurch wird das System komplexer und schwieriger zu verstehen. Dies ist ein Grund dafür, möglichst Modelle mit niedriger Dimension anzustreben. Ich habe festgestellt, dass für die Einbettung des S&P 500 mit ungeglätteten Daten vier Werte erforderlich sind, um eine optimale Einbettung zu erzielen. Aufgrund der ungeglätteten Werte war mit Hilfe eines 'Nearest Neighbour' Prediktors eine wesentlich bessere Vorhersage des S&P 500 möglich, wenn man als Vergleich die Strategie (buy and hold) bzw. die naive Annahme, der Markt bewege sich immer nur nach oben, heranzieht. Die Ergebnisse wurden zum Teil in 'Technical Analysis of STOCKS & COMMODITIES', Septemberheft 2001 publiziert (Uhlig, H. 200a).

Bessere Vorhersagen mit geglätteten Daten

Nach einer geeigneten Glättung der Daten war es sogar möglich, mit drei Werten auszukommen und einen geringeren Prognosefehler zu erzielen. Doch dazu musste ich auch noch zwei Einbettungen miteinander kombinieren. Fazit ist: die Glättung hat zur Reduktion der Dimension geführt und sie hat wichtige Information dabei erhalten. Daran erkennt man, dass die Glättung zumindest teilweise und offenbar spezifisch das Rauschen aus den Originaldaten entfernt hat.

Die Vorhersage mit 'Nächsten Nachbarn' gehört zu den lokalen Prediktoren. In der Sprache der 'Technischen Analyse' würde man sie als Chartprognose bezeichnen. Sie vergleicht die jeweils aktuelle Marktsituation mit allen bis dahin bekannten Situationen aus der Vergangenheit und sortiert sie nach ihrer Ähnlichkeit zum aktuellen Vergleichsmuster. Sie verwendet nur die zur aktuellen Marktsituation ähnlichsten Marktzustände der Vergangenheit, um auf die künftige Entwicklung zu schließen. Die Entwicklung der drei bis sechs ähnlichsten Situationen wird gemittelt und liefert die Schätzung für die künftige Entwicklung der aktuellen Marktsituation. Doch die Vorhersage mit 'Nächsten Nachbarn' ist im Gegensatz zu herkömmlichen Chartprognosen nicht subjektiv, sondern tatsächlich objektiv. Die oben angesprochenen Ähnlichkeiten werden nicht vom Betrachter von Fall zu Fall beurteilt, sondern vom Programm nach einheitlichen und klar definierten Kriterien gemessen. Anders als Neuronale Netze sind diese Prediktoren keine 'black boxes', sondern vollständig transparent. Man benötigt dazu eigentlich keine speziellen Programme, sondern man kann den Prediktor mit Hilfe eines Kalkulationsblattes (z.B. EXCEL oder OpenOffice.Calc) erstellen. Wie man dabei vorgeht, hat der Autor bereits vor einigen Jahren in dem weiter oben erwähnten Artikel (Uhlig, H. 2001a) beschrieben.

Um eine Vorhersage mit 'Nächsten Nachbarn' auch mit Aussicht auf Erfolg durchführen zu können, muss man zuerst geeignete Datenmuster finden, also herausfinden, wie viele und

welche Daten, eine Marktsituation hinreichend definieren. Die geeigneten Datenmuster sind marktspezifisch, d.h. sie lassen sich nicht einfach von einem Markt auf andere Märkte übertragen. Bei der Definition der Muster kann man sich von den Programmen zur bedingten Entropie und dem Chi-Quadrat-Test Hilfe holen. Diese zeigen, welche Daten die meiste Information enthalten. Märkte, das ist inzwischen vielfach nachgewiesen, besitzen ein Langzeitgedächtnis. Sie sind dynamische wiederkehrende Systeme. Nur darum kann man aus der Kenntnis vergangener Marktentwicklungen einen Nutzen ziehen. Die Phasenplots, siehe dazu auch weiter vorn, können uns die Marktdynamik enthüllen. Für die Erstellung eines 'Nearest Neighbour' Prediktors sind sie nicht nötig, falls man ungeglättete Daten verwendet.

Ungeglättete Daten kann man nur durch Zeitverzögerung einbetten. Jeder einzelne Marktzustand wird durch ein Datenmuster ausgedrückt, das mathematisch gesprochen der Ortsvektor der Marktsituation ist. Jeder Einzelwert in dem Datenmuster gilt für eine einzelne Dimension. Wird das Muster aus drei Werten gebildet, dann ist das Marktmodell dreidimensional, sind es vier Werte, arbeitet man mit vier Dimensionen. Möchte man mit weniger Dimensionen auskommen oder eine Einbettung mit Differenzen versuchen, kommt man um eine Glättung nicht herum. Manchmal kann man für die Prognose mit 'Nächsten Nachbarn' auch mehrere Einbettungen vorteilhaft miteinander kombinieren, um einen Marktzustand zu definieren. Darüber gibt es in der Literatur keine Information. Ich habe es selbst herausgefunden. Siehe dazu auch die Anmerkungen beim 'Nearest Neighbour' Prediktor für den S&P 500. Die optimalen Bedingungen muss man durch geeignetes Probieren herausfinden. Auch für die Abstandsberechnungen von 'Nächsten Nachbarn' gibt es verschiedene Metriken, wie die Euklidische Metrik und die Manhattan Metrik. Die Ergebnisse unterscheiden sich manchmal nur geringfügig, aber das muss man erst herausfinden. Es ist auch zu bestimmen, wie viele 'Nächste Nachbarn' man als Entscheidungsgrundlage heran zieht und außerdem muss man bestimmen, ob die Ausgabewerte über alle Nachbarn einfach gemittelt werden sollen, oder ob man besser die einzelnen Nachbarn umgekehrt proportional zu ihrem Abstand vom aktuellen Referenzmuster gewichten sollte. Letzteres würde bedeuten, dass man den ähnlicheren Mustern mehr Gewicht beimessen würde: je ähnlicher, desto mehr Gewicht.

Folgende Parameter müssen bestimmt werden:

Datenmuster: Größe und Zusammensetzung
Glättung: wenn ja, welche
Einbettung: welche
Abstandsmetrik: welche
Nächste Nachbarn: Anzahl und Gewichtung

Damit die Auswertungen der jeweils gewählten Parameter nicht nur für kurze Marktperioden gelten, sondern allgemeine Gültigkeit besitzen, muss man einen repräsentativen Anteil von Marktsituationen mit den verschiedenen Möglichkeiten durchtesten. Prinzipiell wäre das sogar mit einem Kalkulationsblatt möglich, doch würde es sehr lange dauern, selbst wenn man, wie bei EXCEL, Makroanweisungen verwenden könnte. Mit Hilfe der von mir entwickelten Test-Programme, siehe weiter hinten, kann man innerhalb von Sekunden die Auswertung für hunderte oder tausende von Daten erhalten. Die meiste Zeit verbringt man dabei noch mit dem Eingeben der Parameter.

Die kostenlosen 'Nearest Neighbour' Prediktoren (NNP)

Es stehen Ihnen zwei Versionen von NNP kostenlos zur Verfügung. Die eine besteht aus einem Kalkulationsblatt mit Wochendaten des S&P 500 und heißt *S&PKalkdeutsch.xls*. Das ist ein fertiger Prediktor als Datei im *.xls Format von EXCEL. Falls Sie EXCEL nicht besitzen, können Sie stattdessen auch das kostenlose OpenOffice.calc dafür verwenden.

Zu dem Kalukulationsblatt gehört eine Datei die *nnp-deutsch.pdf* heißt. Darin ist die deutsche Übersetzung der Urversion des englischsprachigen Artikels enthalten, den ich seinerzeit zur Publikation bei 'Technical Analysis of STOCKS & COMMODITIES' eingereicht habe und der dort im September 2001 erschienen ist, siehe Literaturanhang.

Das zweite Programm ist etwas komfortabler. Es ist ein JAVA^(TM) Applet dem fünf Dateien mit Finanzmarktdaten beigefügt sind. Es sind Wochenschlusskurse verschiedener Aktienindizes, des deutschen DAX, des österreichischen ATX und des schweizerischen SMI sowie die wöchentlichen Wechselkursangaben der Währungspaare Euro/US-\$ und Schweizer Franken/US-\$.

Märkte als teils stochastische, teils determinierte komplexe Systeme

Die vorigen Abschnitte haben gezeigt, dass Marktdaten, für eine gewisse Zeit, nachweisbare Abhängigkeiten untereinander aufweisen. Entropiemessung und der χ^2 -Test haben den Beweis dafür erbracht. Aber man konnte auch sehen, dass der Informationsgehalt der Daten mit der Zeit schnell abnimmt. Selbst nach einem Zeitschritt ist nur ein Bruchteil der ursprünglichen Information noch in den Daten enthalten. Zieht man den Informationsgehalt ab, so ist der Rest nur Zufallsrauschen oder Signalstörung.

Der Markt ist also ein gemischtes System, das zwar eine gewisse Ordnung besitzt, aber auch ständigen Störungen ausgesetzt ist. Er ist damit kein reines deterministisch chaotisches System, denn er enthält zweifellos Zufallselemente, die in jenen Systemen fehlen. Aber er ist auch kein stochastisches System, das ausschließlich vom Zufall bestimmt ist. Gerade derartige gemischte Systeme erscheinen uns besonders komplex und darum sind sie für uns schwierig zu begreifen. Auch von der mathematischen Forschung werden solche Systeme bisher stiefmütterlich behandelt, weil eine Theorie zu ihrer Behandlung bisher nicht existiert. Die Forschergemeinde hat sich in zwei Lager aufgeteilt: Auf der einen Seite stehen die Chaosforscher, die sich um die deterministisch chaotischen Systeme kümmern und versuchen, sie zu modellieren und eventuell einmal steuern und regeln zu können. Daneben gibt es die Stochastiker, die sich um die Beschreibung und Berechnung zufälliger Ereignisse verdient machen.

Um sicher zu gehen, dass sie bei ihrer Forschung auch zu verwertbaren Ergebnissen kommen, beschäftigen sich die Forscher der beiden Gruppen auch nur entweder mit chaotischen oder mit stochastischen Problemen. Diese Haltung ist in gewisser Weise verständlich, denn in beiden Fällen kann man inzwischen auf etablierten Theorien aufbauen. Wenn man dagegen Neuland beschreitet, sind verwertbare Ergebnisse nicht garantiert. Da man im Bereich der Grundlagenforschung stets um Fördermittel kämpfen und bangen muss, kann man es sich als Forscher kaum leisten, Fehlschläge zu riskieren.

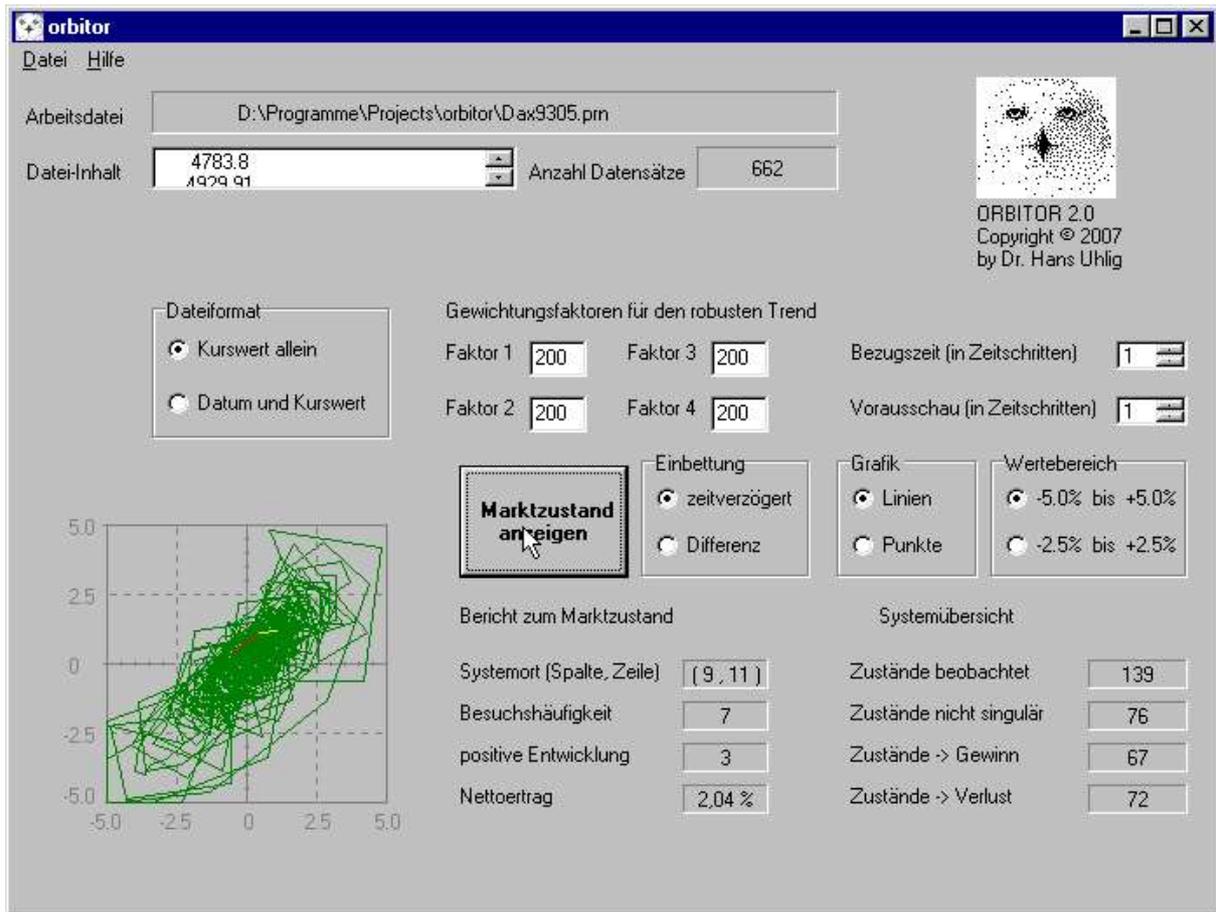
Die mangelnde Bereitschaft seitens der Forscher, sich um die komplexen gemischten Systeme intensiv zu kümmern, wird von einigen namhaften Wissenschaftlern bitter beklagt und sie rufen dazu auf, sich um eine 'stochastische Chaostheorie' zu bemühen. Doch leider ist gegenwärtig in dieser Richtung kein Fortschritt in Sicht. Von der Wissenschaft ist daher zur Zeit nicht viel Hilfe zu erwarten und so ist man auf der Suche nach Lösungen, die zu einem besseren Verständnis von Märkten führen könnten, weitgehend auf sich gestellt.

Die Verknüpfung von Chaos und Statistik

Ein Versuch, mit gemischten Systemen umzugehen, ist auf der folgenden Seite zu betrachten. Er geht zurück auf eine Forschungsarbeit von John Nese am Lorenz-Attraktor. Nese wies nach, dass die Vorhersagbarkeit der Systementwicklung in verschiedenen Bereichen des Attraktors große Unterschiede aufweist. Diese Unterschiede sind offensichtlich gewesen, weil es Bereiche auf dem Attraktor gibt, die stark divergieren und andere die nicht nur parallele, sondern sogar konvergierende Entwicklungen nehmen. John Nese hat sich verdient gemacht, indem er die verschiedenen Bereiche genau benannt und die Vorhersagbarkeit mit Hilfe sogenannter lokaler Lyapunov Exponenten auch quantifiziert hat.

Beim Versuch, Attraktoren von Märkten zu rekonstruieren, konnte auch ich erkennen, dass es Bereiche gab, in denen deutlich parallele Bahnen auftraten, während in anderen Bereichen beinahe beliebige Entwicklungsverläufe der Kursänderungen möglich waren. Mit Hilfe eines Rasters, das ich über die Phasenebene gelegt habe, ist es mir gelungen, einzelne Bereiche eines Attraktors isoliert zu betrachten und die Wahrscheinlichkeiten für Entwicklungen in die eine oder andere Richtung zu quantifizieren. Ich habe ein Programm geschrieben, mit dem ich Raster in nahezu beliebiger Auflösung, zentriert im Nullpunkt anlegen und auswerten kann.

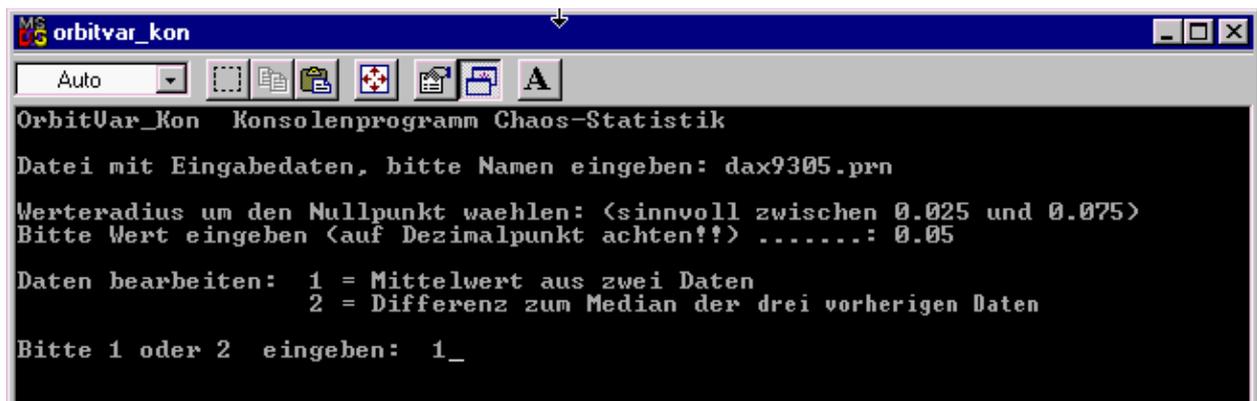
Das Programm ORBITOR für die Chaos-Statistik-Prognose



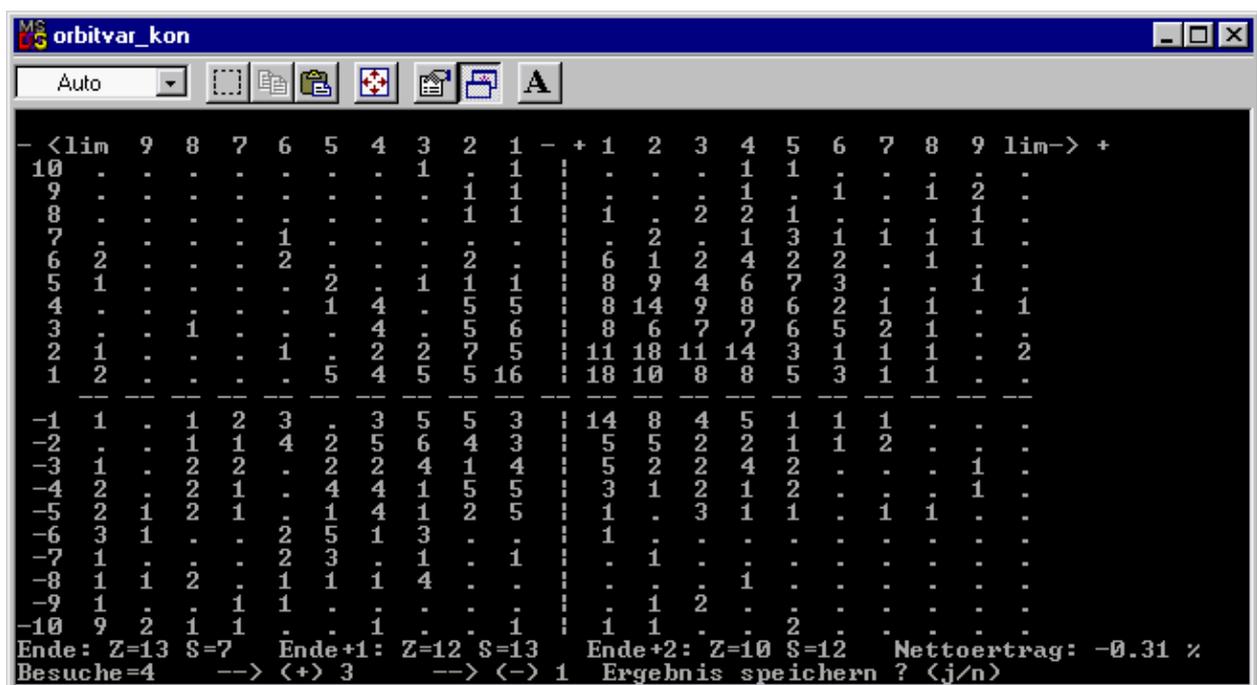
Das oben abgebildete Programm zeichnet die Dynamik des Marktes nach den Vorgaben des Benutzers. Man kann wählen zwischen verschiedenen Glättungen, Einbettungen und Wertebereichen. Der Wertebereich des Zustandsraum wird in 20×20 mögliche Zustandsklassen gerastert. Im Bericht zum Marktzustand wird angezeigt, in welcher der Zustandsklassen das System sich aktuell aufhält und wie es sich davon ausgehend in der Vergangenheit entwickelt hat. Der Systemort wird einmal grafisch angezeigt, als freier Endpunkt der roten Linie. Der Systemort wird auch noch schriftlich angezeigt, mit den Spalten- und Zeilenkoordinaten. Diese werden von links nach rechts und von unten nach oben gezählt. Für den genannten Systemort wird die Häufigkeit der Besuche und die Zahl der positiven Entwicklungen und der Saldo der Erträge des nächsten Entwicklungsschrittes angegeben. Dies ist ein kommerzielles Programm, das nicht kostenlos mitgeliefert wird.

Ein anderes Programm, das dem ORBITOR weitgehend entspricht und bis auf die grafische Darstellung prinzipiell die gleichen Möglichkeiten bietet, ist dagegen unentgeltlich nutzbar. Es wird auf der folgenden Seite vorgestellt. Das Programm ORBITVAR_KON ist, im Gegensatz zum grafisch orientierten ORBITOR ein Konsolenprogramm, das im Zeilenmodus arbeitet. Doch es bietet eine erweiterte Funktionalität in dem Sinne, dass der Wertebereich frei wählbar ist, während im ORBITOR nur zwischen zwei Wertebereichen gewählt werden kann. Dies wurde nötig, weil im Devisenmarkt die kurzfristigen Schwankungen deutlich kleiner sind als beim Aktienmarkt. Um beide mit etwa gleicher Auflösung präsentieren zu können, musste dieser zweite Wertebereich ermöglicht werden. Mit dem variablen Wertebereich von ORBITVAR_KON kann man die Auflösung besser an die individuellen Anforderungen anpassen.

Das Erscheinungsbild des Programms OrbitVar_Kon auf der Bildschirmkonsole



Das Erscheinungsbild kann etwas abweichen, doch betrifft das nur den Konsolenrahmen. Der sieht bei den verschiedenen Windows-Versionen unterschiedlich aus. Das Programm ist interaktiv: Zuerst müssen Sie eine Datei nennen, von der die Daten eingelesen werden sollen. Dann werden Sie aufgefordert, einen Wertebereich einzugeben und schließlich müssen Sie noch die Datenbearbeitung auswählen. Die Eingaben sind jeweils mit der Eingabetaste zu bestätigen. Anschließend erhalten Sie die Bildschirmausgabe, siehe unten.



Der Bildschirm wird angezeigt, bis Sie die Frage beantwortet haben, ob das Ergebnis gespeichert werden soll. Geben Sie < n > ein, dann wird der Bildschirm gelöscht. Geben Sie < j > ein, - Bitte nur den jeweiligen Buchstaben eingeben, nicht die kleiner/größer Zeichen. Dann werden Sie aufgefordert, einen Namen für die Ausgabedatei zu nennen, in die das Ergebnis gedruckt werden soll. Wiederum müssen Sie alle Eingaben mit der Eingabetaste bestätigen.

Die Ausgabedatei wird in das gleiche Verzeichnis geschrieben, in dem sich auch das Programm und die Datei mit den eingelesenen Daten befindet. Achten Sie darauf, dass Einlesedatei und Ausgabedatei nicht den gleichen Namen haben sollen, ansonsten wird die Einlesedatei von der Ausgabedatei überschrieben.

Auf der folgenden Seite sehen Sie den Inhalt der, zu den Abbildungen oben, dazugehörigen Ausgabedatei.

Chaos-Statistik mit dem Programm OrbitVar_Kon

Das Programm gibt die Häufigkeit der Aufenthalte des Marktes in bestimmten Bereichen des Phasenraumes an. Diese Bereiche stehen für definierte Marktsituationen. Der Bereich kann kleiner oder größer gewählt werden, bleibt aber immer im Nullpunkt (Mittelpunkt) zentriert.

```

OrbitVar_Kon: Konsolenprogramm Dateiausgabe
-----
Datei: dax9305.prn  Daten: 662
Wertebereich: 0.0500
Datenbearbeitung: 1 = Mittelwert aus zwei Daten
                  2 = Differenz zum Median der drei vorherigen Daten
Datenbearbeitung: 1 wurde gewaehlt.
Mehrfach besuchte Felder: 120

\ S 1  2  3  4  5  6  7  8  9 10  | 11 12 13 14 15 16 17 18 19 20
Z \ -----+-----
20| .  .  .  .  .  .  .  1  .  1  | .  .  .  1  1  .  .  .  .  .
19| .  .  .  .  .  .  .  .  1  1  | .  .  .  1  .  1  .  1  2  .
18| .  .  .  .  .  .  .  .  1  1  | 1  .  2  2  1  .  .  .  1  .
17| .  .  .  .  1  .  .  .  .  .  | .  2  .  1  3  1  1  1  1  .
16| 2  .  .  .  2  .  .  .  2  .  | 6  1  2  4  2  2  .  1  .  .
15| 1  .  .  .  .  2  .  1  1  1  | 8  9  4  6  7  3  .  .  1  .
14| .  .  .  .  .  1  4  .  5  5  | 8 14  9  8  6  2  1  1  .  1
13| .  .  1  .  .  .  4  .  5  6  | 8  6  7  7  6  5  2  1  .  .
12| 1  .  .  .  1  .  2  2  7  5  | 11 18 11 14  3  1  1  1  .  2
11| 2  .  .  .  .  5  4  5  5 16  | 18 10  8  8  5  3  1  1  .  .
+-----+-----
10| 1  .  1  2  3  .  3  5  5  3  | 14  8  4  5  1  1  1  .  .  .
 9| .  .  1  1  4  2  5  6  4  3  | 5  5  2  2  1  1  2  .  .  .
 8| 1  .  2  2  .  2  2  4  1  4  | 5  2  2  4  2  .  .  .  1  .
 7| 2  .  2  1  .  4  4  1  5  5  | 3  1  2  1  2  .  .  .  1  .
 6| 2  1  2  1  .  1  4  1  2  5  | 1  .  3  1  1  .  1  1  .  .
 5| 3  1  .  .  2  5  1  3  .  .  | 1  .  .  .  .  .  .  .  .  .
 4| 1  .  .  .  2  3  .  1  .  1  | .  1  .  .  .  .  .  .  .  .
 3| 1  1  2  .  1  1  1  4  .  .  | .  .  .  1  .  .  .  .  .  .
 2| 1  .  .  1  1  .  .  .  .  .  | .  1  2  .  .  .  .  .  .  .
 1| 9  2  1  1  .  .  1  .  .  1  | 1  1  .  .  2  .  .  .  .  .

Ende+2: Z=10 S=12
Ende+1: Z=12 S=13
Ende:   Z=13 S=7

Nettoertrag: -0.31 %
Besuche    : 4    --> (+) 3    --> (-) 1
    
```

Oben sehen Sie die Abbildung eines Dateiausdrucks der von einem meiner Programme erzeugt wurde. Sie zeigt ein Raster von 20 * 20 (400) prinzipiell möglichen Marktzuständen. In den einzelnen Feldern eines jeden Rasters ist manchmal ein Punkt zu sehen. Dieser steht für einen bis dahin nicht aufgetretenen Marktzustand. Oder im Feld steht eine Zahl, die angibt, wie oft ein so klassifizierter Marktzustand aufgetreten ist. Anders als bei einer Liniengrafik kann man hier die Abfolge der verschiedenen Zustände nicht erkennen. Um die aktuelle Dynamik des Systems sichtbar zu machen, sind die letzten drei Zustände bezeichnet. In den drei Zeilen unter dem Raster sind die Koordinaten der Punkte angegeben. 'Ende' bezeichnet den letzten beobachteten Zustand, 'Ende + 1' den vorletzten und 'Ende + 2' den Zustand davor. Der Buchstabe Z kennzeichnet jeweils die Zeilenkoordinate und der Buchstabe S die Spaltenkoordinate. Der letzte Zustand befindet sich somit am Schnittpunkt von Zeile 13 und Spalte 7. Man sieht, dass dort die Zahl 4 steht. Das ist die Anzahl der Besuche des Systems in diesem Feld. In den letzten beiden Zeilen der Abbildung kann man noch weitere Angaben zu diesem Zustand finden. Die auf diesen Zustand folgenden Marktbewegungen erbrachten per

Saldo einen Verlust von -0.31%. Von den sechs Besuchen entwickelten sich drei positiv und einer negativ.

Wie man sehen kann, wurden im Programm ORBITOR und in dem vorliegenden Programm die gleichen Daten verwendet. Die Marktdynamik wird aber verschieden dargestellt. So ist auch der aktuelle Marktzustand der angezeigt wird, nicht der gleiche. Diese Unterschiede sind Ausdruck der verschiedenen Datenglättungen und der verschiedenen Wertebereiche.

Ein Marktzustand wird um so genauer charakterisiert, je größer die Auflösung ist, doch damit verbunden ist der Nachteil, dass weniger Zustände mit den gleichen Merkmalen bis dahin aufgetreten sind und somit eventuell keine statistische Aussage über den wahrscheinlichen künftigen Verlauf möglich ist. Darum muss man im Einzelfall abwägen, welche Auflösung man wählen sollte.

Vielleicht fragen Sie sich, ob man ein solches Programm auch mit einem Kalkulationsblatt realisieren kann. Nun, mit einem reinen Kalkulationsblatt geht es nicht, weil man für die Umsetzung verschachtelte Schleifen programmieren muss, die nicht zum Funktionsumfang normaler Kalkulationsblätter gehören. Das Programm EXCEL von Microsoft hat jedoch zusätzlich eine Version der Programmiersprache BASIC integriert, mit der man ein solches Programm durchaus realisieren könnte. Wer will, kann es also versuchen.

Probleme der Datenglättung

In den vorigen Abschnitten war öfter von Datenglättung die Rede wenn es um die Verwendung von realen Marktdaten ging. Wir haben angenommen, dass ein Markt ein niedrigdimensionales, chaotisches System ist, das ständig durch Zufallsereignisse gestört wird. Daraus haben wir geschlossen, dass die Störungen zu Schwankungen um den wahren Marktpreis führen. Durch Datenglättung sollte man eigentlich die Störungen jedenfalls teilweise herausfiltern können. Tatsächlich ist dies der übliche Ansatz. Aber, Datenglättung, so heißt es in einem bekannten Buch, ist eine Kunst, keine Wissenschaft. Es gibt viele Möglichkeiten, Daten zu glätten, und keine davon ist perfekt.

Wie könnte in diesem Fall eine vorsichtige Glättung aussehen? Offensichtlich ist eine Glättung um so vorsichtiger, je weniger Werte in die Glättung mit einbezogen werden. Die geringste Glättung wird durch zwei aufeinander folgende Werte erreicht. Dann sind der Mittelwert und der Median gleich. Bei drei Werten und mehr sind Median und Mittelwert meist verschieden. Je mehr Daten in eine Glättung einbezogen werden, um so stärker ist sie. Aber nach welchen Kriterien sollen wir uns richten, wenn wir uns für eine bestimmte Art der Glättung entscheiden wollen?

Vorschläge für Glättungskriterien

Zunächst muss klar sein, was genau geglättet wird. Sinnvoll ist es in jedem Fall, trendbereinigte Daten zu glätten. Außerdem ist zu berücksichtigen, dass eine logarithmische Datenskala eine viel bessere Anpassung an eine Trendgerade liefert, als eine lineare Skala. Es spricht also einiges für die Annahme, dass die Differenzen der logarithmierten Kurse die gesuchten trendbereinigten Daten sind, die geglättet werden sollen.

Erster Vorschlag: Die Glättung sollte so gewählt werden, dass die Summe der Schwankungen zwischen aufeinander folgenden geglätteten Werten minimal wird. Wenn man so glättet, dann erhält man eine möglichst glatte Kurve.

Zweiter Vorschlag: Die Glättung sollte zu einer möglichst monotonen Kurve führen, d.h. möglichst wenige Richtungswechsel enthalten. Wenn ein Wert größer ist als sein Vorgänger, dann soll er auch möglichst kleiner sein, oder wenigstens nicht größer als sein Nachfolger. Entsprechendes gilt für die Bewegung in die Gegenrichtung.

Dritter Vorschlag: Mit der Glättung sollten möglichst wenige Vorzeichenwechsel angestrebt werden. Dies unterscheidet sich vom zweiten Vorschlag dadurch, dass die geglätteten Werte nicht möglichst monoton sein sollen, es reicht, wenn sie das gleiche Vorzeichen besitzen, wie ihre Vorgänger bzw. ihre Nachfolger.

Derartig optimierte Glättungen kann man nicht mit herkömmlichen gleitenden Durchschnitten erreichen. Die Daten müssen auf jeden Fall unterschiedlich gewichtet werden. Auch die beliebten exponentiell gewichteten gleitenden Durchschnitte liefern nicht die angestrebten Ergebnisse. Man benötigt zu diesem Zweck spezielle Optimierungsprogramme. Lineare Optimierung, 'simulated annealing' oder Optimierung mit 'genetischen Algorithmen' sind dafür geeignete Verfahren. Ich habe die Optimierungen mit dem Programm GENESIS V von John Grefenstette vorgenommen. Es ist ein Programm, das genetische Algorithmen verwendet und es ist im Internet für UNIX und MS-DOS frei erhältlich. Man braucht aber einen 'C'-Compiler oder einen 'C++'-Compiler und Kenntnisse der Programmiersprache 'C', um das Programm nutzen zu können. Denn die zu optimierenden Funktionen müssen in 'C' (nicht in C++) kodiert und dann in den weiteren Quelltext des Programm eingebunden werden. Anschließend kann man daraus mit Hilfe des C/C++-Compilers ein ausführbares Programm erstellen.

Obwohl dies nicht das Ziel der Optimierung ist, kann die Glättung allein bereits für bessere Vorhersagen von Marktbewegungen ausgenutzt werden. Ich habe dies an anderer Stelle am

Beispiel des deutschen Aktienindex DAX gezeigt, siehe mein Artikel im Verlag der Versicherungswirtschaft November 2001. Auch für den französischen CAC40 kann dies gezeigt werden. Die entsprechenden Daten hierzu wurden von mir noch nicht in einer Zeitschrift veröffentlicht, aber ich stelle sie als *.pdf Datei zur Verfügung. Dies Vorhersagesystem gehört zu den sogenannten globalen Prediktoren. Das besondere an der von mir dafür vorgenommenen Glättung der DAX-Daten ist, dass die Glättungsfaktoren unverändert über viele Jahre gute Ergebnisse liefern, so dass auf ihrer Grundlage profitable Handelssysteme etabliert werden können. Dies funktioniert so nicht bei allen Aktienmärkten. Im Gegensatz zum Beispiel des deutschen DAX und des französischen CAC40 kann der globale Prediktor für den robusten Trend beim S&P in Hausse-Phasen keine besseren Ergebnisse erzielen, als die Strategie 'buy and hold'.

Idealerweise sollte die Datenglättung bewirken, dass die in den Originaldaten enthaltenen Störungen bzw. Zufallselemente beseitigt werden und nur die Information erhalten bleibt. Dieser Idealfall wird nie erreicht. Die Datenglättung geschieht immer auch auf Kosten der Informationserhaltung. Eine gute Glättung ist eine, die viel Zufallsrauschen aus den Daten entfernt und nur wenig Information dafür opfert. Glättungen mit nur zwei Faktoren erhalten zwar die Information besser, entfernen jedoch nicht ausreichend das Zufallsrauschen. Eine starke Glättung über viele Werte (sechs und mehr) kann das Rauschen besser entfernen, doch dafür wird dann zu viel Information geopfert. So werden, im Falle von Wochendaten, wie sie hier betrachtet werden, mit Glättungen über drei bis fünf Werte meistens die besten Resultate erzielt.

Glättung der Originaldaten soll nicht nur das störende Rauschen vermindern. Geglättete Daten sind auch die Voraussetzung, um den sogenannten Attraktor, siehe die Beispiele weiter vorn, für das dynamische System des Marktes aus der Datenreihe rekonstruieren zu können. Der Attraktor ist eine grafische Darstellung der Abfolge von Zuständen des Systems über längere Zeit. Er kann als Vorlage für ein Modell des Marktes dienen und er soll die Bahnen der Systemdynamik beschreiben, auf die das System nach einer Störung immer wieder zurück findet.

Teil IV - Testen von 'Nearest Neighbour' und Chaos-Statistik Prediktoren.

Mit Auswertungsprogrammen kann man den Wertpapier- oder Devisenhandel nach vorgegebenen Regeln simulieren und diese Simulation über einen gewissen Zeitraum laufen lassen, um zu testen, wie sich das System bewährt hätte. Ein System, das in der Vergangenheit funktioniert hat bietet zwar keine Sicherheit, dass es auch weiterhin funktionieren wird, doch es ist der einzige Anhaltspunkt, den man hat.

TradeStation^(R), MetaStock^(R) usw. taugen nur für globale Prediktoren

Institutionelle und unabhängige selbständige Börsenhändler können derzeit zwischen mehreren kommerziellen Programmen wählen, die in der Lage sind, Handelsstrategien zu prüfen und Handelsstatistiken auszugeben. An erster Stelle wäre da die TradeStation^(R) von Firma OMEGA zu nennen, aber auch andere Produkte, wie beispielsweise MetaStock^(R) von Firma EQUIS und weitere mehr bieten derartige Möglichkeiten. Diese Programme sind recht flexibel. Neben Varianten vorgegebener Strategien können sie auch neu programmierte Regeln auswerten. Eine wesentliche Einschränkung gibt es jedoch für alle diese Programme: sie können nur globale Strategien auswerten; lokale Strategien gehören nicht zu ihrem Repertoire.

Globale Handelsstrategien sind solche, die bestimmte Kriterien für Marktsituationen definieren, die günstig sind für Marktengagements. Alle Daten werden geprüft, ob sie diese Kriterien erfüllen und wenn das der Fall ist, wird gehandelt. Der Erfolg der Strategie wird dadurch bestimmt, wie gut diese Kriterien die günstigen Marktbedingungen beschreiben und wie häufig solche Bedingungen vorliegen. Als Beispiel für eine einfache globale Handelsstrategie mag die von mir vorgestellte Methode des robusten Trends (Uhlig, H. 2001b) dienen.

Die Auswertungsprogramme für lokale Prediktoren

Da die allermeisten Börsenhändler globale Prediktoren verwenden, gibt es auf diesem Gebiet viel Konkurrenz und das schmälert den möglichen Gewinn. Schon aus dem Grund wäre es sinnvoll, andere Ansätze zu prüfen. So hat man eventuell die Chance, profitable Gelegenheiten zu finden, die bei herkömmlicher Betrachtung übersehen werden. Wirklich gute lokale Prediktoren werden nicht veröffentlicht, sonst wäre ihr Nutzen schnell dahin. Wer davon profitieren möchte, muss sie selbst finden, dann braucht er den Ertrag auch nicht mit anderen zu teilen. Die Suche nach geeigneten oder gar optimalen Parametern ist ohne schnelle Auswertungsprogramme eine mühsame und zeitaufwendige Angelegenheit. Mit Hilfe der hier vorgestellten Programme kann man mögliche Kandidaten für profitable Strategien schnell, ausgiebig und konsistent, d.h. widerspruchsfrei prüfen.

Ich habe mehrere Auswertungsprogramme für die 'Nearest Neighbour' Prediktoren und für die Chaos-Statistik Vorhersagen geschrieben, von denen ich hier zwei kurz vorstellen möchte. Wie auch für die vorher besprochenen Programme, habe ich hier wieder die wöchentlichen Schlusskurse des Deutschen Aktienindex DAX von Anfang 1993 bis Mitte 2005 verwendet. Diese Zeitreihe umfasst verschiedene Börsenphasen mit steigenden und fallenden Kursen und schien mir daher gut geeignet für Testzwecke. Beide Programme sind für die Konsole geschrieben, ohne den grafischen Überbau für die GUI von MS-Windows^(TM). Kommerzielle Versionen, die für den Umgang mit Tagesdaten vorgesehen sind und über eine MS-Windows^(TM) GUI verfügen, ähnlich den Programmen ORBITOR und SURROGATOR, sind bereits in Arbeit. Die Konsolen sehen unter den verschiedenen Betriebssystemen von MS-Windows^(TM) jeweils unterschiedlich aus. Damit man die Bildschirminhalte auch ohne Screenshot-Programm speichern kann, habe ich jeweils die Ausgabe in eine Datei mit vorgesehen. Falls Interesse besteht, kann ich die Programme auch für das Betriebssystem LINUX zugänglich machen.

Auswertungsprogramm für 'Nearest Neighbour' Prediktoren

```
MS-DOS-Eingabeaufforderung - G032
7 x 12
+++++++ direkte Nearest Neighbour Prognose ++++++
Das Programm simuliert Vorhersagen fuer 1 Zeitschritt(e)
Nun 4 Glaettungsfaktoren mit Leerzeichen getrennt eingeben:

Einbettungen einheitlich oder kombiniert:
  1 = mit 4 Zeitverzoegerungen
  2 = mit 3 Zeitverzoegerungen + 1 Differenz
  3 = mit 2 Zeitverzoegerungen + 2 Differenzen
  4 = mit 1 Zeitverzoegerung + 3 Differenzen
  5 = mit 4 Differenzen

Bitte eine Auswahl treffen (1..5): 2

Eingabe: 3 Zeitschritte und 1 Differenz (Beispiel: 1 2 4 2): 1 1 4 1

Bitte geben Sie 4 Gewichtungsfaktoren ein (Dezimalbrueche erlaubt),
mit Leerzeichen dazwischen (Beispiel: 1 1.2 1.3 1.4): 1 1 1 1

NN Auswertung waehlen (1=mit3, 2=gmit3, 3=mit4, 4=gmit4, 5=mit5, 6=gmit5): 4

Daten einlesen (max 5000), Bitte Dateinamen eingeben: dax9305.prn
Referenzdatensatz von dem begonnen werden soll !! muss > 0 sein !!: 1

Wie viele, Testdaten sollen ausgewertet werden: 640
```

oben sehen Sie den Eingabedialog für ein Auswertungsprogramm. Nicht alle Eingaben sind lesbar, denn mein Know-how möchte ich natürlich nicht verschleiern. Wichtig ist, dass man sehen kann, welche Möglichkeiten angeboten werden und wie die Eingaben erfolgen.

Die Zahl der Vorhersageschritte ist hier voreingestellt. Es ist immer ein Zeitschritt - für ein professionelles Programm werde ich die Zeitschritte wählbar machen. Als nächstes wird nach vier Glättungsfaktoren gefragt. Diese beziehen sich auf vier aufeinander folgende Werte in der eingelesenen Datenreihe. Sie müssen nicht mit geglätteten Werten arbeiten. Falls Sie mit den Rohdaten arbeiten möchten, dann geben Sie einfach nur 1 0 0 0 ein. Dadurch wird der erste Wert mit 1 multipliziert und die nachfolgenden mit null. Falls Sie nur über zwei Werte glätten wollen, geben Sie 1 1 0 0 ein. Denken Sie, dass der erste Wert mehr Gewicht haben sollte, können Sie ihn auch mit 1.2 oder 1.5 oder einfach beliebig gewichten. Entsprechendes gilt für drei und mehr Gewichtungsfaktoren.

Die Einbettung kann einheitlich oder eine Kombination aus zwei Verfahren sein. Im Beispiel habe ich eine Kombination gewählt. Die ersten drei Werte sind zeitverzögert eingebettet und der vierte mit Differenz. Sie sehen, dass der erste Wert doppelt eingegeben ist, dazu der vierte und die erste Differenz.

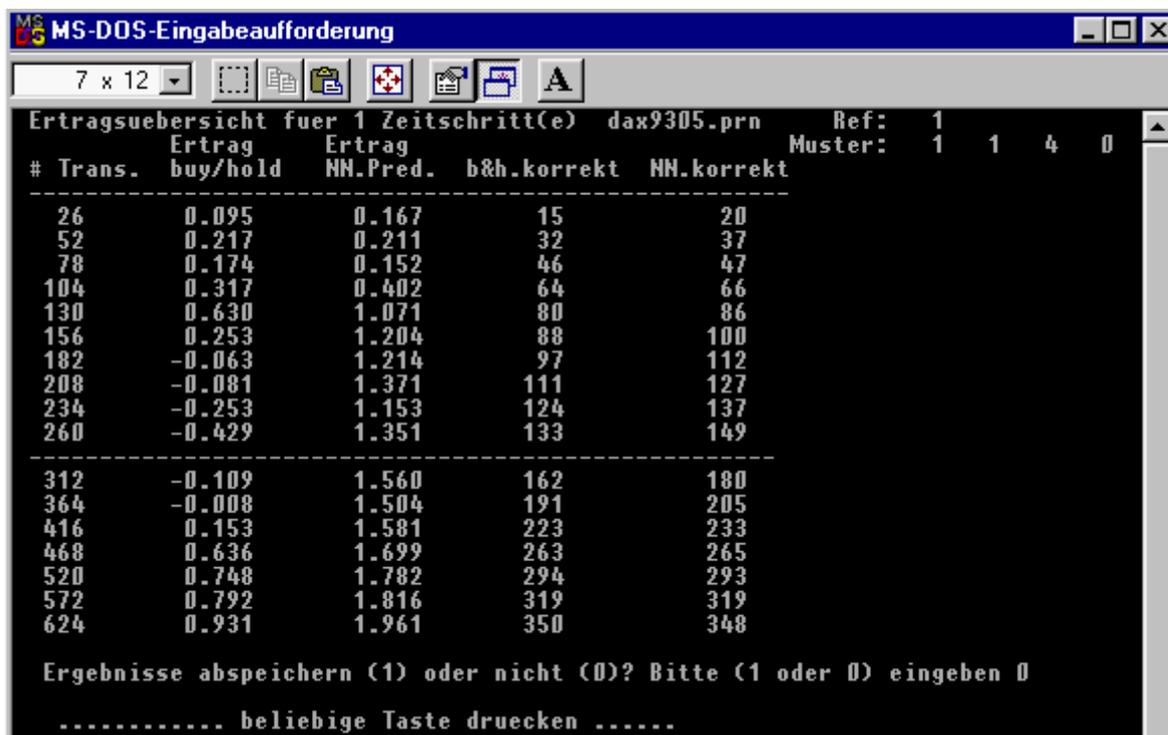
Als nächstes kann man die Eingaben gewichten. Ich habe allen den Wert 1 gegeben. Dadurch wird der erste Wert doppelt gewichtet, weil er zwei mal auftritt. Ich hätte auch als Gewichtung 0.5 0.5 1 1 eingeben können, dann wären alle drei Werte gleich gewichtet worden.

Die anschließende Frage beriff die Zahl der 'nächsten Nachbarn', über die das Ergebnis gemittelt werden soll : mit3, mit4 usw. sind Mittel über drei bzw. vier 'nächste Nachbarn'. Steht ein 'g' davor, also gmit5 beispielsweise, dann bedeutet dies, dass über fünf nächste Nachbarn gemittelt werden soll. Diese sollen aber gewichtet werden und zwar umgekehrt proportional zu ihrem Abstand vom Referenzmuster. Das heißt, die ähnlichsten Werte werden am stärksten

gewichtet, die unähnlichsten am geringsten. Im Beispiel wurde Modus 4 gewählt, das sind vier gewichtete 'nächste Nachbarn'.

Schließlich muss man noch die Datei eingeben, von der die Daten gelesen werden und festlegen, von welchem Datensatz begonnen und wieviele insgesamt ausgewertet werden sollen. Im Beispiel beginnt man bei 1 und lässt 640 Daten auswerten.

Sobald man die letzte Eingabe durch Drücken der Eingabetaste bestätigt hat, sieht man das Ergebnis.



# Trans.	Ertrag buy/hold	Ertrag NN.Pred.	b&h.korrekt	NN.korrekt
26	0.095	0.167	15	20
52	0.217	0.211	32	37
78	0.174	0.152	46	47
104	0.317	0.402	64	66
130	0.630	1.071	80	86
156	0.253	1.204	88	100
182	-0.063	1.214	97	112
208	-0.081	1.371	111	127
234	-0.253	1.153	124	137
260	-0.429	1.351	133	149
312	-0.109	1.560	162	180
364	-0.008	1.504	191	205
416	0.153	1.581	223	233
468	0.636	1.699	263	265
520	0.748	1.782	294	293
572	0.792	1.816	319	319
624	0.931	1.961	350	348

Ergebnisse abspeichern (1) oder nicht (0)? Bitte (1 oder 0) eingeben 0
 beliebige Taste druecken

Oben sieht man die Bildschirmausgabe des Auswertungsprogramms für die 'Nearest Neighbour' Vorhersage. In der linken Spalte ist die Anzahl der Transaktionen angegeben, d.h. die Anzahl der Wochen, für die der Test durchgeführt wurde. Die nächsten beiden Spalten zeigen die Erträge 'buy+hold', also 'kaufen und halten' und des 'Nearest Neighbour' Prediktors. Die beiden rechten Spalten geben an, wie oft die Investitionsentscheidung günstig war. Oberhalb der gestrichelten Linie sind die Erträge im Halbjahresabstand angegeben, darunter im Jahresabstand. Es werden also insgesamt die Ergebnisse von 12 Jahren ausgegeben, von Mitte 1993 bis Mitte 2005.

Das Programm zählt rückwärts vom aktuellsten Wert bis zum ältesten. Die obere Zeile ist also das Ergebnis der letzten 26 Wochen, darunter stehen die letzten 52 Wochen ... usw. Die Methode 'kaufen und halten' hat sehr schwankende Erträge, während das 'Nearest Neighbour' Programm einen nahezu kontinuierlichen Zuwachs verzeichnet. In den letzten 12 Jahren hätte es beim DAX etwa den doppelten Ertrag gebracht, jedenfalls unter den hier betrachteten Bedingungen des konstanten Investitionsvolumens. Das trifft aber bei der Methode 'kaufen und halten' nicht zu, bei der ständig die Erträge reinvestiert werden. Eine solche Strategie ist jedoch problematisch, wie ich gleich erläutern werde.

Wenn man die Erträge reinvestiert, dann verändert man auch das Anlagerisiko, denn angenommen, die Erträge sind positiv, riskiert man mit der Zeit immer höhere Verluste. In den Zeiten, als die Finanzmärkte nicht so turbulent waren, hat man das Verlustrisiko vernachlässigt. Doch die Schönwetterperiode ist vorbei. Inzwischen sollte jedem klar sein, dass man die Erträge riskanter und weniger riskanter Anlagen nicht 1 : 1 vergleichen darf. Man muss auch das Risiko berücksichtigen und das ist es, was der oben angeführte Vergleich zeigen soll, denn er bezieht sich auf ein einheitliches Risiko.

Auswertungsprogramm für Chaos-Statistik

```

MS-DOS ORBVARD5
7 x 12
***** Vorhersagequote *****
-----
Wertebereich-Ausdehnung um den Nullpunkt waehlen (Vorschlag 0.05): 0.05
Datenbearbeitung 2 = Mittelwert aus jeweils zwei Daten.
Modus auswaehlen: 3 = Mittelwert aus jeweils drei Daten.
ausgewaehlter Modus = 3
Einbettung durch 1 = Zeitverzoegerung von einem Zeitschritt
Zeitverzoegerung: 2 = Zeitverzoegerung von zwei Zeitschritten
Wahl der Einbettung = 1
Datei mit Eingabedaten, bitte Namen eingeben: dax9305.prn
Auswertung StatChaos und minimaler Trend
-----
Datei: dax9305.prn Datenbearbeitung: 3 Einbettung: 1 Wertebereich: 0.050
Daten: 658 positive Wochen: 370 negative Wochen: 286
Quadratsumme der Besuche : 6230 besuchte Felder: 171 mehrfach besucht: 93
Marktengagement und absolute Ertraege
-----
1. Woche StatChaos mit Anzahl + >= - engagiert: 39 % Ertrag: 86
1. " " mit Vorwert >= 0 engagiert: 37 % Ertrag: 111
2. Woche StatChaos mit Anzahl + >= - engagiert: 40 % Ertrag: 36
2. " " mit Vorwert >= 0 engagiert: 41 % Ertrag: 92
1. Woche minimaler Trend + engagiert: 61 % Ertrag: 110
2. Woche minimaler Trend + engagiert: 61 % Ertrag: 105
1. Woche buy + hold engagiert: 100 % Ertrag: 111
StatChaos - Vorhersagen und Ergebnisse im Vergleich:
-----
1. Woche: Anzahl + Entw.: 254 korrekt: 150 = 59 % Ertrag/Risiko: 34
1. Woche: Wert + Entw.: 245 korrekt: 151 = 62 % " / " 45
2. Woche: Anzahl + Entw.: 263 korrekt: 151 = 57 % " / " 14
2. Woche: Wert + Entw.: 270 korrekt: 164 = 61 % " / " 34
1. Woche min.Trend + Entw.: 403 korrekt: 239 = 59 % " / " 27
2. Woche " " + Entw.: 403 korrekt: 235 = 58 % " / " 26
1. Woche buy+hold + Entw.: 658 korrekt: 370 = 56 % " / " 17
Ergebnis in eine Datei schreiben? (j/n)

```

oben sehen Sie den Bildschirmausdruck eines der Auswertungsprogramme zur Chaos-Statistik. Es ist, wie die vorher gezeigten ein Konsolenprogramm, das seine Auswertung auch in eine Datei schreiben kann. Interessant ist vor allem die untere Hälfte der Abbildung, denn sie zeigt die Auswertung eines StatChaos-Programms, das mit einer bestimmten Parameterwahl getestet wurde. Zudem wurde ein Vergleich mit einem minimalen Trend und der Methode 'buy+hold', also 'kaufen und halten' durchgeführt.

Der obere Teil der unteren Hälfte zeigt jeweils die Dauer des Marktengagements an. Die Methode 'kaufen und halten' ist ständig am Markt engagiert (100%), während die anderen Strategien zwischen 37% und 61% der Zeit am Markt sind. Die Dauer des Marktengagements ist nach meiner Erfahrung praktisch immer proportional zur Standardabweichung der Erträge. Der angegebene Ertrag ist nicht risikobereinigt, ausgedrückt in Prozenten des eingesetzten Kapitals, bezogen auf konstantes Investitionsvolumen, also ohne Reinvestition der Erträge.

Der untere Bereich zeigt den spezifischen Ertrag der jeweiligen Methode, d.h. den Ertrag bezogen auf einheitliches Risiko, also Ertrag / (Standardabweichung der Erträge). Weiterhin wird gezeigt, wie oft die Investitionsentscheidung korrekt war. Für 'kaufen und halten' war die Entscheidung in 56% der Fälle richtig. Der minimale Trend war da schon besser, mit 58% und 59% vorteilhaften Entscheidungen. Die Chaos-Statistik zeigt noch bessere Ergebnisse, zwischen 57% und 62% je nach Parameterwahl. Der spezifische Ertrag ist bis zum Faktor 2,64 besser als 'kaufen und halten'.

Man sieht, dass es Prognosen für die erste und die zweite Woche gibt. Das sind tatsächlich die

Vorhersagen einmal für die nächste und getrennt davon für die übernächste Woche. Also 2. Woche bedeutet nicht, um es noch einmal klar zu sagen, die erste und die zweite Woche, sondern nur die zweite Woche. Das gilt für die Chaos-Statistik und für den minimalen Trend in gleicher Weise.

Für die Chaos-Statistik sieht man für die erste und die zweite Woche jeweils zwei Angaben. Bei einer dieser Angaben steht *Anzahl* und bei der anderen steht *Wert*. Diese Bezeichnungen sollen darauf hinweisen, wie die Statistik erstellt wurde. *Anzahl* heißt, dass die Marktsituation, in der die Investitionsentscheidung gefallen ist, in der Vergangenheit mehr positive oder mindestens gleich viel positive, wie negative Entwicklungen genommen hat. *Wert* bedeutet, dass der Saldo der Erträge in der Vergangenheit nicht negativ sein durfte, wenn die Investitionsentscheidung zugunsten eines Marktengagements getroffen wurde. Es sieht so aus, als wäre der Saldo der vorangegangenen Erträge der bessere Ratgeber für die Investitionsentscheidung.

Chaos-Statistik und Geldmanagement

Die Chaos-Statistik soll nicht nur bessere Vorhersagen ermöglichen, sondern sie soll auch ein Hilfswerkzeug für das Geldmanagement sein. Wenn man die verschiedenen Marktsituationen identifiziert und die jeweilige Wahrscheinlichkeit für einen positiven Ertrag berechnet hat, kann man diese Information nutzen, um seine Investitionen auch quantitativ danach auszurichten. Man könnte also in den Situationen, die mit größerer Wahrscheinlichkeit einen positiven Ertrag erwarten lassen, das Investitionskapital erhöhen. Entsprechend gilt für weniger profitable Marktsituationen, dass weniger Kapital investiert werden sollte.

Ein gezieltes, situationsabhängiges Geldmanagement ist in jedem Fall wünschenswert und mit diesem Hilfsmittel hat man Anlass zur Hoffnung, dass ein Versuch in dieser Richtung erfolgreich sein könnte.

Literatur

- Antoniou, Antonios, Kyrtsov, Catherine und Vorlow, Constantinos: Surrogate Data Analysis and Stochastic Chaotic Modelling: Application to Stock Exchange Return Series. 10th International Conference on Computing in Economics and Finance, 2004. Internet Publikation SCE2004-27[1].pdf
- Caballe, Jordi, Jarue, Xavier und Michetti, Elisabetta: Chaotic dynamics in credit constrained emerging economies. *Journal of Economic Dynamics & Control*, 2006 vol 30, s.1261-1275.
- Chen, Ping: Empirical and theoretical evidence of economic chaos. *System Dynamics Review* Vol 4, Nr. 1-2, 1988. Internet Publikation SDR88p.pdf
- Copeland Laurence: Nonlinear Dynamics in Real-Time Equity Market Indices: Evidence from the UK. Social Science Research Network, 1998 SSRN.com/abstract=5356.
- De Grauwe, Paul, Dewachter, Hans and Embrechts, Mark: Exchange Rate Theory - Chaotic Models of Foreign Exchange Markets. Blackwell Publishers, Oxford, United Kingdom, 1993.
- Fan, Chongjun, und Xu, hui: The Review on Non-Linear Analysis and Forecasting Methods of the Real Estate Market in China. *Asian Social Science*, Vol 4, Nr. 11, Nov. 2008.
- Finkenstädt, Bärbel: Nonlinear Dynamics in Economics - A Theoretical and Statistical Approach to Agricultural Markets. *Lecture Notes in Economics and Mathematical Systems* Vol. 426, Springer Verlag Berlin, Heidelberg, 1995.
- Jian, Wang und Guangxue, Cui: The Chaos and its Complexity for Petroleum Price Sequence. Society of Petroleum Engineers, Paper Number 104409-MS 2006.
- Lo, Andrew, W.: On The Fundamentals Of Technical Analysis. Interview in *Technical Analysis of STOCKS & COMMODITIES*, Dezember 1997, S. 50-67.
- Lo, Andrew, W. & MacKinlay, A. Craig: *A Non-Random Walk Down Wall Street*. Princeton University Press, Princeton, N.J., U.S.A., 1999.
- Maheu, John M. und McCurdy, Thomas H.: Nonlinear Features of Realized FX Volatility. 2001 Scientific series ISSN 1198-8177 Internet Publikation 2001s-42.pdf
- Mihail, Ossama, Eberwein, Curtis J. und Handa, Jagdish: On the Evidence of Non-Linear Structure in Canadian Unemployment. 3. November 2003, Internet Publikation, 03-34Mikhail.pdf
- Moshiri, Saeed und Faezeh, Foroutan: Forecasting nonlinear crude oil futures prices. *The Energy Journal*, Oct. 2006.
- Nese, John M.: Quantifying Local Predictability in Phase Space. *Physica D* 35, p. 237-250, 1989.
- Prokhorov, Artem: Nonlinear Dynamics and Chaos Theory in Economics: a Historical Perspective. Dez. 2001, Internet Publikation papereco.pdf
- Rao, M.J. Manohar: *Chaotic Economic Dynamics and Business Cycles*. 1995, University of Bombay, Department of Economics, Working Paper 95/2 Internet Publikation dwp47.pdf
- Roa, Maria Jose, Saura, Dulce und Vazquez, Francisco J.: A simple chaotic model of economic growth and unemployment. Internet Publikation LX5TLBMD-CDUX-5MVP-XYBDYW963EBVPXKM.pdf
- Tan, Clarence N.W.: A Hybrid Financial Trading System Incorporating Chaos Theory, Statistical and Artificial Intelligence/Soft Computing Methods. Invited Paper - Queensland Finance Conference 1999. Neural29.pdf

Uhlig, Hans: Finanzprognosen mit Neuronalen Netzen., 1995, Verlag Franz Vahlen, ISBN 3-8006-1958-X.

Uhlig, Hans: Finanzmarktanalyse - Neue Ansätze aus der Chaosforschung., 1999, Verlag Franz Vahlen, ISBN 3-8006-2363-3.

Uhlig, Hans: Nearest Neighbour Prediction, "Technical Analysis of STOCKS & COMMODITIES", Seattle, WA, U.S.A., November 2001a,

Uhlig, Hans: Risikoarme Kapitalanlage am Deutschen Aktienmarkt., Dezember 2001b, Versicherungswirtschaft. Verlag Versicherungswirtschaft, Karlsruhe.

Uhlig, Hans: Aktienanlage am deutschen Markt – Wenig Risiko bei gutem Ertrag. Versicherungswirtschaft, Heft 21, S. 1661-1664, November 2005, Verlag Versicherungswirtschaft, Karlsruhe.

Weston, Rae und Premachandran, Prem: Could exchange rates just be chaotic? 2004, Internet Publikation 274-weston_paper.pdf