Spektral- und Korrelationsanalyse von Zeitreihen Studienarbeit an der Technischen Universität Berlin

Jörg Arndt

1.Juli 1993

Inhaltsverzeichnis

1	Ziel	der Studienarbeit	2
2	Vor	bereitung der Daten	4
	2.1	Trend- und Mittelwert Elimination	4
	2.2	Filterung und Glättung	5
	2.3	Logarithmieren	5
	2.4	Erzeugung von Zufallsphasen-Surrogaten	6
3	Kla	ssische Methoden der Zeitreihenanalyse	7
	3.1	Histogramme	7
	3.2	Fouriertransformation und Spektrum	7
		3.2.1 Fast Fourier Transformation (FFT)	8
		3.2.2 Implementierung der Fouriertransformation	10
	3.3	Autokorrelationsfunktion (AKF)	10
		3.3.1 Allgemeines	10
		3.3.2 Direkte Berechnung der AKF	11
		3.3.3 Berechnung der AKF über FFT	12
		3.3.4 Implementierung der Berechnung der AKF	12
4	Met	thoden zur Attraktoranalyse	13
	4.1	Delay-Abbildung (Phasenraumportraits)	13
	4.2	Implementation der Delayabbildung	14
	4.3	Poincaré-Schnitte und Maximum-Abbildung	15
5	Vor	hersagen über Yule-Walker Gleichungen	16
6	Anv	vendungsbeispiele für die Programme	18
	6.1	Das Rösslersystem	18
	6.2	Simulationsdaten für einen Halbleiterlaser	26
٠	6.3	Analyse von Phonationszeitreihen	32
		6.3.1 Aufnahme einer pathologischen Stimme	32
		6.3.2 Computersimulation der menschlichen Stimmlinnen	38

Inhaltsverzeichnis

1	Ziel	der Studienarbeit	2
2	Vor	bereitung der Daten	4
	2.1	Trend- und Mittelwert Elimination	4
	2.2	Filterung und Glättung	5
	2.3	Logarithmieren	5
	2.4	Erzeugung von Zufallsphasen-Surrogaten	6
3	Kla	ssische Methoden der Zeitreihenanalyse	7
	3.1	Histogramme	7
	3.2	Fouriertransformation und Spektrum	7
		3.2.1 Fast Fourier Transformation (FFT)	8
		3.2.2 Implementierung der Fouriertransformation	10
	3.3	Autokorrelationsfunktion (AKF)	10
		3.3.1 Allgemeines	10
		3.3.2 Direkte Berechnung der AKF	11
		3.3.3 Berechnung der AKF über FFT	12
		3.3.4 Implementierung der Berechnung der AKF	12
4	Met	thoden zur Attraktoranalyse	13
	4.1	Delay-Abbildung (Phasenraumportraits)	13
	4.2	Implementation der Delayabbildung	14
	4.3	Poincaré-Schnitte und Maximum-Abbildung	15
5	Vor	hersagen über Yule-Walker Gleichungen	16
6	Anv	wendungsbeispiele für die Programme	18
	6.1	Das Rösslersystem	18
	6.2	Simulationsdaten für einen Halbleiterlaser	26
	6.3	Analyse von Phonationszeitreihen	32
		6.3.1 Aufnahme einer pathologischen Stimme	32
		6.3.2 Computersimulation der menschlichen Stimmlinnen	38

Ziel der Studienarbeit

Die vorliegende Studienarbeit soll eine Einarbeitung in die Thematik der Zeitreihenanalyse sein. Für einige klassische und moderne Methoden der Zeitreihenanalyse wurde ein Programmpaket erstellt. Diese werden in dieser Arbeit kurz erläutert. Fast alle Funktionen sind im Programm FT integriert.

Beschreibungen von Programmfunktionen - und Bedienung sind kursiv gesetzt Implementierte Funktionen:

- Datenmanipulation:
 - Mittelwertabziehen
 - Tiefpassfilterung: $(\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$ -Filter
 - Logarithmieren
 - Erzeugung von Random-Phase Zeitreihen
- Klassische Methoden:
 - Spektrum
 - Autokorrelationsfunktion
- Attraktoranalyse:
 - Maximum-Abbildung
 - Delay-Koordinaten (Programm 3D)
- Vorhersage:
 - Anpassen eines Autoregressiven Prozesses (Yule-Walker Gleichungen)

Die Programme sind in Gnu-C++ (MS-Dos-Portierung für IBM-PC) geschrieben, nur die Vorhersage über Yule-Walker Gleichungen wurde als Mathematica-Package implementiert. Es wurde Wert gelegt auf einfache Bedienbarkeit ("Eintasten-Funktionen") und hohe Interaktivität (zu jeder Zeit wird eine Graphik präsentiert). Angestrebt wurde, den Quellcode lesbar zu halten durch:

- Kommentierung im Quellcode selbst (später evtl. Einbindung ins Web-System)
- Modularität (je ein File für eine oder wenige zusammengehörige Funktionen)

- Konsistenter Programmierstil
- Benutzung von möglichst wenigen PC-spezifischen Funktionen
- so wenig ++ wie möglich (im wesentlichen nur für Ein-Ausgabe)

Es sollte also später möglich sein, die Software auf Unix-Maschinen zu übertragen, zumindest, falls eine Portierung des Gnu-C Compilers vorhanden ist und entsprechende Möglichkeiten der Graphikausgabe existieren.

Vorbereitung der Daten

Vielfach ist es notwendig, die Daten vor der (numerischen) Untersuchung vorzubearbeiten. Oft möchte man Trends und "langsame" Prozesse eliminieren, weil nur "schwingungsartige" Teilprozesse von Interesse sind. Die numerische Berechnung verschiedener Größen erfordert i.a. eine Vorverarbeitung, um Artefakte zu vermeiden. Beispielweise müssen verrauschte Daten geglättet werden bevor man versucht, lokale Maxima zu bestimmen.

Im Programm FT ist es zu (fast) allen Zeiten möglich, die gerade graphisch präsentierten Daten summarisch oder mit einstellbaren Grenzen und Schrittweite (z.B. ab Wert 250 bis Wert 1250, nimm jeden 10. Wert) als ASCII-File abzuspeichern.

2.1 Trend- und Mittelwert Elimination

Möchte man "langsame" Prozesse, Trends oder "Jahresgänge" in den Daten eliminieren, so gibt es eine Reihe von Möglichkeiten, dies zu erreichen. Es ist zweckmässig, sich vor Augen zu halten, daß all dies Spezialfälle von Frequenzfilterungen sind, Trendelimination entspricht Hochpassfilterung, Elimination von Jahresgängen entspricht einer Bandsperre. Also kann man diese Manipulationen im Fourierraum vornehmen, d.h. man Fouriertransformiert die Zeitreihe, multipliziert die Spektralfunktion mit der gewünschten Filtercharakteristik und transformiert dann zurück. Mittelwerte können natürlich auch direkt subtrahiert werden. Für Trends und Jahresgänge kann auch wie folgt verfahren werden:

- ermittele "Mittelwerte" für ein Fenster, das über die Zeitreihe bewegt wird
- lege eine Ausgleichsfunktion (Splines!) durch die Folge dieser Mittelwerte
- subtrahiere diese Ausgleichsfunktion von der Zeitreihe

Offensichtlich hängt das Ergebnis entscheidend von der Breite der Fenster und der Art der Interpolationsfunktion ab. Eine andere Möglichkeit um Trends zu eliminieren ist es, die Differenzfolge der gegebenen zu betrachten; davon wird hier kein Gebrauch gemacht.

Im Programm FT hat die BenutzerIn zu Beginn die Möglichkeit, den Mittelwert (über die gesamte Zeitreihe) subtrahieren zu lassen.

2.2 Filterung und Glättung

Besonders bei verrauschten Daten ist es oft notwendig eine Glättung durchzuführen, einerseits, um unerwünschte hochfrequente Anteile herauszufiltern, andererseits, um gewisse Weiterverarbeitung (wie z.B.Maximumssuche) überhaupt erst zu ermöglichen. Wieder könnte direkt im Fourierspektrum operiert werden, aber für die Glättung der Daten bietet sich ein anderes Verfahren an:

Man ersetzt den (zweiten bis vorletzten) Wert x_i durch $\left(\frac{x_{i-1}}{4} + \frac{x_i}{2} + \frac{x_{i+1}}{4}\right)$. Dies ist ein Spezialfall eines linearen Filters (siehe [1] Kap. 2.5.2), nämlich ein Tiefpaßfilter. Durch wiederholte Anwendung kann der Effekt verstärkt werden.

Im Programm FT wird mit der Taste 'i' ("Filterung") die Tiefpaßfilterung nach obigem Verfahren ausgelöst, man kann dann die Anzahl der hintereinander ausgeführten Filterdurchgänge eingeben. Danach liegen die Daten gefiltert für die Weiterverarbeitung im Programm bereit. Sollen diese in einem File gespeichert werden, so wähle man sofort (oder später im Menüpunkt 't', d.h. Ansicht der Zeitreihe) die Taste 's'. Hat man zuviel gefiltert, so muß das Programm neu gestartet werden, damit die ungefilterten Daten wieder vorliegen.

2.3 Logarithmieren

Beim Logarithmieren macht man sich zunutze, daß die Logarithmusfunktion Verhältnisse in Differenzen verwandelt (Funktionalgleichung des Logarithmus), also Größenordnungen in konstante Abstände.

So ist es üblich, Powerspektren zu logarithmieren, um in einem Plot die dominierenden Schwingungprozesse und gleichzeitig vielfach energieärmere (z.B. Rauschen) erkennen zu können. [Bild]

Ein exponentieller Trend in einer Funktion lässt sich nach Logarithmieren (als Gerade mit konstanter Steigung) einfacher identifizieren.

Ein Signal mit "peakartigem" Verhalten lässt sich manchmal durch Logarithmieren in eine relativ symmetrische Schwingung überführen [Bild Laserzeitreihe], was bei gewissen Untersuchungen von Vorteil ist, z.B. bei der Einbettung in einen Pseudophasenraum.

Vor dem Logarithmieren müssen die Daten linear auf ein Intervall abgebildet werden, dessen Grenzen größer Null sind. Man hat dabei die Freiheit, durch die Wahl der Grenzen den Grad der Nichtlinearität beim Logarithmieren zu bestimmen. Wählt man die Grenzen nahe bei 1, so ergibt sich eine geringe Verzerrung. Möchte man viele Größenordnungen (z.B. von Spektralanteilen) in einem einzigen Plot sehen, so ist es sinnvoll, das Intervall auf [1...große Zahl] abzubilden. Ist dabei die Obergrenze zu groß gewählt, läuft man Gefahr, alle Unterschiede "Auszubügeln" [Bild]. Möchte man einen Nullsymmetrischen Output, so sollte ein Intervall der Form [Wert, 1/Wert] gewählt werden.

Im Programm FT kann die Zeitreihe logarithmiert (und so abgespeichert) werden. Zunächst wird nach den Grenzen des linearen Intervalles gefragt, in das die Daten linear abgebildet werden sollen, bevor mit dem natürlichen Logarithmus logarithmiert wird. So gibt es die Möglichkeit, erst mit den Intervallgrenzen zu experimentieren, um das gewünschte Ergebniss zu finden. Möchte man mit den logarithmierten Daten weiterarbeiten, so muß (mit der Taste 's') in ein File abgespeichert werden, denn die logarithmierten Daten werden nach ver-

lassen der Plotfunktion wieder verworfen! Weiterhin bietet das Programm die Möglichkeit, Spektren zu logarithmieren.

2.4 Erzeugung von Zufallsphasen-Surrogaten

Zufallsphasen-Surrogate sind künstlich erzeugte Pseudo-Zeitreihen, die das gleiche Leistungsspektrum haben wie eine vorgegebene Zeitreihe, deren Phasen aber zufällig gewählt sind. Das Histogramm eines Zufallsphasen-Surrogats zeigt, daß die Daten Gaußverteilt sind, dies gilt jedoch nur, falls die Amplituden zu "vielen" Frequenzen von Null verschieden sind (das Histogramm einer reinen Sinusschwingung wird auch nach einer Phasenänderung keine Gaußverteilung aufweisen). Die Daten entsprechen also einer Realisierung eines stochastischen Prozesses mit gleichem Leistungsspektrum und gleicher Autokorrelationsfunktion wie die Originaldaten. Dies ist vor allem nützlich, um bei weitergehenden Untersuchungen einen Test für die Ergebnisse zu haben, d.h. Artefakte zu erkennen.

Numerisch erstellt man dies einfach wie folgt:

• Originalzeitreihe Fouriertransformieren

- Realteil-Imaginärteil Darstellung in Betrag-Phase Darstellung umwandeln
- Phasen durch zufällige Phasen ersetzen (dabei ist zu beachten, daß bei N Datenpunkten, um ein reelles Ergebnis zu erhalten, für die Frequenzwerte von N/2...N die Phasen so gewählt werden, daß Phase(N-k) = -Phase(k) für k=0...N/2-1)
- Betrag-Phase Darstellung in Realteil-Imaginärteil Darstellung umwandeln
- Inverse Fouriertransformation ausführen

Klassische Methoden der Zeitreihenanalyse

3.1 Histogramme

Ein Histogramm ist eine "Häufigkeitsstatistik" der Zeitreihe, d.h. eine Funktion

$$h(y) := \sum_{\{t: x(t) \in [y-\epsilon, y+\epsilon]\}} 1$$

Histogramme erlauben erste Aussagen über Zeitreihen, so deutet eine Gaußverteilung auf einen stochastischen Prozess hin, eine "exponentielle" Flanke im Histogramm legt eine Logarithmierung der Daten vor der Weiterverarbeitung nahe.

3.2 Fouriertransformation und Spektrum

Die Gesamtheit der Beträge der Entwicklungskoeffizienten nach dem vollständigen Funktionen-Orthonormalsystem $e^{-i\,x\,\omega}$, $\omega\in R$, nennt man Spektrum. Für physikalische Systeme ist dies gleichbedeutend mit einer Zerlegung eines (oft periodischen) Zeitverhaltens in eine Summe aus harmonischen Schwingungen. Die Fouriertransformation ist eine unitäre Transformation, für eine quadratintegrierbare Funktion (genauer: Funktionen aus der Schwarzklasse) $f: R^n \to R^n$, $\vec{x} \mapsto f(\vec{x})$ ist sie definert durch:

$$F(\vec{\omega}) := \frac{1}{\sqrt{2\pi^n}} \int_{B^n} f(\vec{x}) \, e^{-i \, \vec{x} \, \vec{\omega}} d^n x$$

Die Rücktransformation ist einfach

$$f(\vec{x}) = \frac{1}{\sqrt{2\pi^n}} \int_{\mathbb{R}^n} F(\vec{\omega}) \, e^{+i \, \vec{x} \, \vec{\omega}} d^n \omega$$

d.h. gerade das konjugiert Komplexe der Hintransformation. Im 1-dimensionalen Fall gilt damit

$$F(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x)e^{-i\,x\,\omega} dx$$

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} F(\omega) e^{+ix\omega} d\omega$$

Für periodische Funktionen betrachtet man ein Periodizitätsintervall L und die Fouriertransformation geht über in die halbkontinuierliche Form:

$$\hat{c}_k := rac{1}{\sqrt{L}} \int_L f(x) \, e^{-i \, k \, x} dx$$

$$f(x) = \sum_{k=-\infty}^{k=+\infty} c_k e^{i kx}$$

Die letzte Formel gilt für Stellen x, an denen f(x) stetig ist, an Unstetigkeitsstellen konvergiert die Fourierreihe gegen $\frac{f(x+0)+f(x-0)}{2}$. Häufig wird die folgende Form der Fourierentwicklung angegeben:

$$a_k := \frac{1}{\sqrt{L}} \int_L f(x) \cos \frac{2\pi k x}{L} dx, \qquad k = 0, 1, 2, \dots$$

$$b_k := rac{1}{\sqrt{L}} \int_L f(x) \sin rac{2\pi k x}{L} dx, \qquad k = 1, 2, ...$$

und

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} \left(a_k \cos \frac{2\pi k x}{L} + b_k \sin \frac{2\pi k x}{L} \right)$$

Damit gilt

$$c_k = \begin{cases} \frac{a_0}{2} & (k=0) \\ \frac{1}{2}(a_k - ib_k) & (k>0) \\ \frac{1}{2}(a_k + ib_k) & (k<0) \end{cases}$$

Bei numerischen Berechnungen hat man es allerdings nicht mit einer kontinuierlichen Funktion f(x), sondern mit einem Satz von Datenpunkten f_j zu tun, beispielsweise mit Messwerten nach jeweils konstanten Zeiten. In diesem Fall gehen die obigen Fourierintegrale über in endliche Fouriersummen:

$$c_k := \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} f_j e^{-i j k}$$

Die Rücktransformation ist wieder ein "symmetrischer" Ausdruck:

$$f_j = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} c_k e^{+ijk}$$

Zur Fouriertransformation vgl. [5].

3.2.1 Fast Fourier Transformation (FFT)

Die programmiertechnische Realisierung der Fouriertransformation geschieht (für Datenfelder der Größe 2^n) mittels der sogenannten <u>Fast Fourier Transformation</u> (FFT). Der Operationsaufwand für die Fouriertransformation ist bei direkter Implementierung proportional N^2 , man berechnet nämlich N Summen mit je N Summanden. Die FFT reduziert den Aufwand auf proportional $N \log_2 N$, indem sie die Koeffizienten nach einem rekursiven Schema berechnet, das letztlich darauf beruht, daß die komplexen Zahlen $(\cos \frac{2\pi k}{N} + i \sin(\frac{2\pi k}{N}), k = 0...N-1)$ eine multiplikative Gruppe der Ordnung N darstellen, die im Falle $N=2^n$

vom Element mit k=1 erzeugt wird. Das Verfahren lässt sich für Datenlängen N mit anderer Primfaktorisierung verallgemeinern, praktisch ist es eher sinnvoll, die Daten durch angehängte Nullen bis zur nächsten Länge der Form 2ⁿ zu ergänzen. (zur FFT siehe z.B. [4]). Weitere Vereinfachungen ergeben sich für reelle Funktionen (Halbierung des Aufwandes) und durch Benutzung der Beziehung

$$e^{-2\pi i(k+1)/N} = e^{-2\pi i k/N} e^{-2\pi i/N}$$

So ist nur eine einzige sin-cos Berechnung notwendig und alle weiteren sin-cos Werte können rekursiv durch Multiplikation gewonnen werden.

Man sieht leicht, daß für Datenmengen, bei denen die direkte Implementation zu sehr langen Programmlaufzeiten führt, die FFT noch sehr schnell ist. Die mathematischen Eigenschaften der Fouriertransformation (Ableitung und Integration vereinfachen sich zu Multiplikation bzw. Division mit der unabhängigen Variablen, Faltung zweier Funktionen wird zur einfachen Multiplikation im Frequenzraum) in Verbindung mit den schnellen FFT-Algorithmen und der Tatsche, daß die Rücktransformation praktisch identisch zur Hintransformation ist, machen die Fouriertransformation zu einem universellen Werkzeug in vielen Bereichen der numerischen Analyse.

Bei der numerischen Auswertung von FFT-Spektren müssen allerdings einige Dinge beachtet werden, um richtig zu interpretieren und Artefakte zu erkennen. So darf man nicht erwarten, für einen gesampleten Ausschnitt eines Sinus-Signales etwa einen "Delta-Peak", d.h. ein Spektrum mit nur einem einzigen Fourierkoeffizienten ungleich Null zu finden. Im allgemeinen sind die Spektren aber "ähnlich" denen, die man durch die Integralvariante der Transformation bei idealisierten Signalen erhalten würde. Die charakteristischen Eigenschaften und Effekte der FFT sind:

- Die Maximalfrequenz liegt in der "Mitte" des Fourierspektrums ("Wrap-Around Order"), die Werte für $N/2 < k \le N$ müssen als negative Frequenzen angesehen werden; bei reellen Funktionen gilt stets: F(k) = F(N-k). Daher ist es sinnvoll, die Frequenzen N/2+1...N gar nicht darzustellen.
- das Abschneiden der Funktion an Anfang und Ende bewirkt sog. <u>Clip-Effekte</u>. (Dieser Effekt kann z.B. mittels "<u>Windowing</u>", d.h. Aufmultiplizieren von Fenstern vor der Transformation, vermindert werden)
- "Aliasing": Frequenzenwerte, die ausserhalb des eigentlichen Frequenzintervalles liegen, werden in das Frequenzintervall "hineingefaltet". Beispielweise werden Schwebungen von Frequenzen, die größer sind als jene, der der Wert N/2 entspricht, mit der Samplefrequenz im Spektrum als "unsinnige" Peaks auftauchen.

Gegen Fehlinterpretationen aufgrund von Aliasing hilft eigentlich nur eine (vorherige) Kenntnis der Signaleigenschaften. Ohne solche heuristischen Argumente ist es objektiv unmöglich, alias-Peaks von nichtartefaktischen zu unterscheiden. Versuche, solche Effekte nachträglich zu entfernen, stellen einen massiven und nichttrivialen Eingriff in die Daten dar, so daß nachfolgende Analysen mit besonderer Vorsicht interpretiert werden müssen.

Um etwas Vertrautheit im Umgang mit FFT-Spektren zu gewinnen, empfiehlt es sich, mit Testdatensätzen zu "spielen" und auf diese Weise die Effekte der verschiedenen möglichen Manipulationen zu beobachten.

3.2.2 Implementierung der Fouriertransformation

In Programm FT ist die Spektralanalyse als Fast Fourier Transformation implementiert. Als Implementation wurde ein Algol-Pseudoprogramm aus [6] (Kapitel 2.3 Trigonometric Interpolation, Seite 84 oben) in C umgeschrieben. Das Programm ist eine Cooley-Tukey-Implementation der FFT, mit der geringfügigen Veränderung, daß das Programm die symmetrische Form der FFT benutzt (Normerhaltung, Vorfaktor $\frac{1}{\sqrt{N}}$). Die Fourier Transformation kann über die gesamte Zeitreihe ausgeführt werden, wobei ein cos-Fenster (Welch-Fenster) benutzt werden kann, um Clip-Effekte zu vermeiden. Falls die Anzahl der Datenpunkte nicht einen Wert $N=2^n$ hat, werden die ersten 2^k Datenpunkte mit $2^k \leq N$ zur Transformation verwendet. Vor der Darstellung des Spektrums kann dieses logarithmiert werden.

Weiterhin besteht die Möglichkeit, ein A- Ω -t Diagramm (Amplitude-Frequenz-Zeit) zu erstellen, d.h. eine Serie von Spektren für jeweils einen Ausschnitt, der über die Zeitreihe bewegt wird. Die Breite dieses Ausschnittes (im Programm als "Fenster" bezeichnet) ist einstellbar, muß aber von der Form 2^n sein. Über jeden Auschnitt muß ein cos-Fenster gelegt werden, da sonst Clip-Artefakte unvermeidbar wären. Die Verschiebung der Fenster ist bei jedem Spektrum gerade die halbe Fensterbreite, es werden also bei N Datenpunkten rund 2N/Fensterbreite Spektren erstellt. Die Darstellung (auf Wunsch logarithmiert) kann als 3-dimensionaler Plot ("Wasserfall-Diagram") oder als Dichteplot (z.Zt. nur über den Umweg als Mathematica lesbares Outputfile) erfolgen. Die Anzahl der Frequenzen, die in der Darstellung unterschieden werden, kann auf einen Wert \leq Fensterbreite/2 eingestellt werden, was bei großen Fensterbreiten oft die Übersichtlichkeit verbessert und für Dichteplots praktisch immer notwendig ist. Im Programm ist dies so implementiert, daß verschiedene Frequenzen in gemeinsame "Boxen" aufsummiert werden.

Im Programm FT wählt man die Fouriertransformation über die gesamte Zeitreihe aus dem Hauptmenü mit der Taste 'f' an, danach kann Windowing (cos-Fenster) und Logarithmierung des Spektrums gewählt werden.

Die Berechnung eines Amplitude-Freqenz-Zeit Diagrammes wählt man aus dem Hauptmenü mit der 'x'-Taste. Danach entscheidet man über

- Logarithmieren des Spektrums
- Erstellung einer vom Programm "Mathematica" lesbaren Datei (Name: "fftfile.dat") um eventuell später einen Dichteplot erstellen zu können
- Anzahl der im Spektrum zu unterscheidenden Frequenzen
- Breite des über die Zeitreihe bewegten Fensters

3.3 Autokorrelationsfunktion (AKF)

3.3.1 Allgemeines

Allgemein ist eine (Kreuz-) Korrelationskoeffizient für zwei Zufallvariablen wie folgt definiert:

$$r(x,y) := \frac{C(x,y)}{\sqrt{C(x,x)\,C(y,y)}}$$

dabei ist

$$C(x,y) := \langle x \, y \rangle - \langle x \rangle \, \langle y \rangle$$

$$\langle x \rangle := \text{Zeitmittel, hier: } \langle x \rangle := \frac{1}{N} \sum_{i=1}^{N} x_i$$

(Abschätzung für Zeitmittel = Mittel über die Zeitreihe), C(x,x) ist die Varianz von x. Für zwei statistisch unabhängige Zufallsvariablen x,y gilt r(x,y)=0. Die Umkehrung gilt nicht, die Korrelationsfunktion erfaßt nur lineare Zusammenhänge, d.h. Zusammenhänge bei denen sich die beiden Zufallsvariablen um eine Gerade konzentrieren. Beispiel für eine Größe, die genau dann verschwindet, wenn zwei Zufallsvariablen x,y unabhängig sind, ist die sog. Transinformation:

$$H(x,y) := \int \int P(x,y) \log \frac{P(x,y)}{P(x) P(y)} dxdy$$

Die <u>Autokorrelationsfunktion</u> (AKF)

$$AKF(\tau) = r_{\tau}(x, x) := \frac{C_{\tau}(x, x)}{C_{0}(x, x)}, \text{ mit } C_{\tau}(x, x) := \langle x_{i} x_{i+\tau} \rangle - \langle x_{i} \rangle^{2}$$
(3.1)

ist ein Maß für das "Gedächtnis" einer Funktion. $C_{\tau}(x,x)$ heißt Kovarianz von x. Die AKF ist symmetrisch bzgl. der Null, hat dort den Wert 1 und überall sonst gilt $-1 \le AKF(\tau) \le +1$. Für Weißes Rauschen hat die AKF für alle $\tau \ne 0$ den Wert Null, allerdings auch etwa für eine Zeitreihe bestehend aus den sukzessiven Werten aus der logistischen Abbildung $x \mapsto r x (1-x)$ für r=4. Diese beiden Zeitreihen können mit der Delaykoordinaten-Methode unterschieden werden.

3.3.2 Direkte Berechnung der AKF

Ein Weg, eine Abschätzung für die Autokorrelationsfunktion zu gewinnen, ist die direkte Berechnung über 3.1. Zunächst werden $C_{\tau}(x,x) := \langle x_i x_{i+\tau} \rangle - \langle x \rangle^2$ für alle gewünschten τ berechnet und dann, um $r_{\tau}(x,x)$ zu gewinnen, die Werte durch $C_0(x,x)$ dividiert. Falls der Mittelwert der Zeitreihe Null ist, entfällt der rechte Term und die Berechnung wird so vereinfacht. Die Anzahl n der zur Berechnung benutzten Datenwerte muß durch Division der AKF-Abschätzung durch n ausgeglichen werden. So könnte für N Datenpunkte folgendes Verfahren verwendet werden (hier $n = N - \tau$):

$$C_{ au}(x,x) pprox rac{\sum_{k=1}^{N- au} x_k x_{k+ au}}{(N- au)}$$

damit dann:

$$AKF(\tau) = \frac{C_{\tau}(x,x)}{C_{0}(x,x)}$$

Dabei ist zu beachten, daß die statistischen Fluktuationen für Werte von τ in der Ordnung N sehr groß werden,allgemein sind die Schwankungen der Abschätzung proportional $\frac{1}{\sqrt{(N-\tau)}}$. Das obige Verfahren bringt einen Aufwand von proportional N^2 Operationen, was für längere Zeitreihen eine beträchtliche Berechnungszeit bedeuten kann. Die Benutzung von Intervallen konstanter Länge spart zwar Rechenzeit, vergrößert jedoch die statistischen Schwankungen der AKF-Abschätzung. Vorteilhafter ist die Berechnung über die Fast Fourier Transformation, die einen Berechnungsaufwand proportional $N \log_2(N)$ mit sich bringt, also für lange Zeitreihen prinzipiell vorteilhaft ist und außerdem stets alle Datenpunkte miteinbezieht.

3.3.3 Berechnung der AKF über FFT

Im Folgenden bezeichne $\mathcal{F}f$ die Fouriertransformierte von f(x).

$$\mathcal{F}C_{ au}(f(x),f(x))=\int\int f(x)\,f(x+ au)dx\,e^{-i\,\omega\, au}d au=$$

 $(\text{mit } z := \tau + x, dz = d\tau)$

$$\int \int f(x) f(z) dx e^{-i\omega(z-x)} dz = \int f(x) e^{+i\omega x} dx \int f(z) e^{-i\omega z} dz =$$
$$(\mathcal{F}^{-1}f) (\mathcal{F}f) = |\mathcal{F}f|^2$$

also (wegen $\mathcal{F}^{-1}f = cc(\mathcal{F}f)$ für reelle f):

$$C_{\tau}(f(x), f(x)) = \mathcal{F}^{-1}(|\mathcal{F}f|^2)$$
(3.2)

Die Autokovarianzfunktion ist also die inverse Fouriertransformation des Powerspektrums. Oft findet man diese Beziehung für Zeitreihen formuliert als

$$|\mathcal{F}f|^2 = \sum_{k=1}^N C_{\tau}(x,x) \cos \frac{2\pi \ \tau \ k}{N}$$

Dabei wurde berücksichtigt, daß wegen der Symmetrie der Autokovarianzfunktion der Imaginärteil verschwindet.

3.3.4 Implementierung der Berechnung der AKF

Die Berechnung der Abschätzung AKF wurde sowohl direkt als auch über FFT Implementiert. Zur direkten Berechnung wird die ganze Zeitreihe herangezogen, was gelegentlich zu langen Laufzeiten führt. Zu beachten ist, das der Mittelwert (bisher) nur summarisch abgezogen werden kann, also noch vorhandene Trends zu spürbaren Ungenauigkeiten führen können.

Bei der FFT-Implementation mußten im wesentlichen drei Dinge beachtet werden:

- Für die Berechnung von k Autokorrelationskoeffizienten müssen den Daten k Nullen angehängt werden, um ein "Herumfalten" bei der Faltungsberechnung via FFT zu vermeiden (siehe [3], Seite Kap. 12.4).
- Die AKF muß für endliche Ausschnitte unendlicher Signale korrigiert werden. Eine reine sin-Funktion soll eine (nichtabfallenden) cos-Funktion als AKF-Abschätzung ergeben, also muß der Wert k durch (N-k) dividiert werden, vgl. [Bild]. Danach werden alle Werte durch AKF(0) dividiert, so daß AKF(0)=1.
- Die Unsicherheit der Schätzwerte für die AKF steigt wie bei der direkten Berechnung, Werte in der Ordnung N sollten nicht benutzt werden.

Im Programm FT kann aus dem Hauptmenu mit der Taste 'c' die Berechnung der AKF ausgelöst werden, man kann zwischen direkter Berechnung (Taste 'd') und Berechnung über FFT (Taste 'f') wählen. Es werden jewiels nur die Korrelationskoeffizienten bis N/4 wiedergegeben.

Methoden zur Attraktoranalyse

4.1 Delay-Abbildung (Phasenraumportraits)

Stochastische und chaotische Systeme können die gleichen Spektren und Autokorrelationsfunktionen haben, also durch "klassische" Untersuchungmethoden ununterscheidbar sein. Dissipative deterministische Systeme besitzen aber sog. Attraktoren im Phasenraum, d.h. ihre Trajektorien nähern sich (nach einer Einschwingzeit) Strukturen im Phasenraum, die wichtige Aussagen über das System erlauben. Beispiele für Attraktoren sind Fixpunkt, Grenzzyklus, Torus, n-dimensionaler Torus und sog. seltsamer Attraktor. Seltsame Attraktoren sind kennzeichnend für chaotische Systeme und haben die für sie typischen fraktalen Eigenschaften; vgl. Bild [Bild] Rössler-Attraktor (a=.15, b=.20, c=10.). Gelingt es, für ein gegebenes System einen Attraktor zu finden, also seine Form (Topologie) zu bestimmen, so ist dieses System erfolgreich von den stochastischen Systemen unterschieden. Eine Methode, Attraktoren aufzuspüren, ist das Zeichnen von sog. Phasenportraits durch die Methode der Delay-Koordinaten. Benötigt wird hierbei nur eine einzige dynamische Variable, von der man eine Bahn im abstrakten m-dimensionalen sog. Pseudophasenraum gewinnt, indem man

$$\vec{X} = \{x(t), x(t+\tau), x(t+2\tau), ..., x(t+(m-1)\tau)\}$$

setzt. Diese Methode ist schon deshalb so wichtig, weil oft nur eine einzige dynamische Variable überhaupt zur Verfügung steht. Als Trivialbeispiel möge die Schwingung des harmonischen Oszillators $x(t) = \sin(\omega t)$ dienen. Trägt man im 2-dimensionalen Phasenraum

$$\vec{X} = \{\sin(\omega t), \sin(\omega (t + \tau))\}\$$

auf, so ergeben sich im allgemeinen Ellipsen, die nur in gewissen Sonderfällen (Delayzeit=Vielfaches der halben Periode) zu geraden Strecken entartet sind. Dies Beispiel zeigt schon, daß die richtige Wahl der Delayzeit τ wesentlich ist. Es erweist sich, daß im allgemeinen Fall die Delayzeit sinnvoll als ein Bruchteil der "typischen" (niedrigsten) Systemfrequenz gewählt wird. Für die Einbettungsdimension m gilt, daß sie hinreichend groß sein muß: Ein Satz von Takens sagt darüber aus, daß für $m > 2D_H + 1$ (wobei D_H die Hausdorff-Dimension des Attraktors ist) die Einbettung des Attraktors in den Pseudophasenraum eineindeutig ist. Für Zeitreihen ist man bei der Wahl der Delayzeit natürlich auf Vielfache der Samplezeit beschränkt und man erhält nur eine Punktmenge im Phasenraum. Als Bespiel für die Delaymethode bei Zeitreihen betrachten wir den Rössler-Attraktor für (a=.10, b=.10, c=8.): Bild [Bild] zeigt den Attraktor, der dargestellt wurde, indem aus den drei dynamischen Variablen x(t), y(t), z(t) ein dreidimensionales Bild gewonnen wurde. Die Bilder

Literaturverzeichnis

- [1] C. Chatfield: The Analysis of Time Series, Chapman and Hall, 1975
- [2] Ebeling, Engel, Herzel: Selbstorganisation in der Zeit, Akademie Verlag, 1990
- [3] Press u.a.: Numerical Recipes in C, Cambridge University Press, 1988
- [4] E. Oran Brigham: The fast Fourier Transform, Prentice-Hall 1974
- [5] Bronstein, Semendjajew: Taschenbuch der Mathematik, Verlag Harri Deutsch, 21. Auflage 1981
- [6] J. Stoer, R. Bulirsch: Introduction to Numerical Analysis, Springer Verlag 1980

[Bild] zeigen, wie nur aus der Variable x(t) allein der Attraktor rekonstruiert werden kann. (An der Punktstruktur kann man übrigens erkennen, daß die Attraktorrekonstruktionen aus Zeitreihen gewonnen wurden). Man sieht, daß die Rekonstruktionen des Attraktors typischerweise verzerrt sind, die interessierenden (globalen,topologischen) Eigenschaften wie Attraktor-Dimension, Lyapunov-Exponent und Kolmogorov-Entropie bleiben jedoch erhalten.

4.2 Implementation der Delayabbildung

Das Programm 3D stellt eine Attraktorrekonstruktion eines Teiles der Zeitreihe dreidimensional (Orthogonalperspektive) dar. Der Beginn des dargestellten Teils und seine Länge sind interaktiv wählbar. Der Delay ist einstellbar. Anfangs wird die Zeitreihe eingelesen und ein Teil der Daten wird mit dem Delay 1 dargestellt. Mit dem Delay 1 kann das Programm auch als primitiver 3d-Betrachter benutzt werden, falls man ein Datenfile hat, in dem Koordinatentripel von Punkten hintereinander als ascii-Zahlen stehen. Der erste Wert im file muß (wie auch für das Programm FT) die Anzahl der nachfolgenden Werte angeben. Das Programm wird mit "3D inputfile" aufgerufen.

Zur Bedienung des Programms:

- Mausbewegung ⇒ Drehung (falls keine Maus vorhanden, ist dies auch mit den Tasten 'x', 'X', 'y', 'Y', 'z', 'Z' möglich, wobei die geshiftete Taste jeweils zurückdreht)
- '+' und '-' ⇒ Vergrößerung und Verkleinerung.
- 'b' und 'w' \Rightarrow Hintergrundfarbe weiß bzw. schwarz
- 'c' ⇒ Menü "Färbemethode":
 - 'θ' ⇒ alle Punkte haben die gleiche Farbe
 - '1' ⇒ Färbung wie z-Koordinate der Attraktor-Rekonstruktion
 - '2' \Rightarrow Färbung wie z-Koordinate der Punkte in der aktuellen Darstellung
 - '3' \Rightarrow Färbung wie Reihenfolge der Punkte
- $'d' \Rightarrow Eingabe \ des \ Delays$
- 'D' ⇒ Eingabe von drei Delaywerten, die sukzessive angewandt werden (Bsp.: Mit 'D' 1 0 1 erhält man eine Einbettungsdimension zwei)
- ullet Pfeil nach oben \Rightarrow Vergrößerung des dargestellten Ausschnittes der Zeitreihe
- ullet Pfeil nach unten \Rightarrow Verkleinerung des dargestellten Ausschnittes der Zeitreihe
- Pfeil nach links ⇒ Bewegung des dargestellten Ausschnittes der Zeitreihe zum Anfang der Zeitreihe hin
- Pfeil nach rechts ⇒ Bewegung des dargestellten Ausschnittes der Zeitreihe zum Ende der Zeitreihe hin
- 'i' ⇒ Invertiere Bildschirmfarben und warte auf Tastendruck (um mit "pcxdump" ein schwarz/weiß-PCX file des Bildschirms produzieren zu können)

- 'F1' ⇒ kurzer Hilfsbildschirm
- 'q' ⇒ Verlassen des Programmes (nach Bestätigung mit 'y')

4.3 Poincaré-Schnitte und Maximum-Abbildung

Schneidet man den Attraktor im Phasenraum mit einer Ebene (allgemeiner: Fläche) und betrachtet die Menge der Durchstoßpunkte, so erhält man sog. <u>Poincaré-Schnitte</u> (Poincaré-Abbildung). Die kontinuierliche Dynamik wird so auf eine diskrete Dynamik abgebildet. Hat man nämlich ein deterministisches System vor sich, so läßt sich die Folge der Punkte \vec{P} in der Ebene allgemein beschreiben durch:

$$\vec{P}_{n+1} = \text{Nichtlineare Abbildung}(\vec{P}_n)$$

Man hat es also mit iterierten diskreten nichtlinearen Abbildungen zu tun. Die Eigenschaften solcher Abbildungen sind gut studiert (Bsp.: Henon-Abb., logistische Abb., Zelt-Abb., Bernoulli-Abb.) und die universellen Eigenschaften solcher Abbildungen, wie sie M. Feigenbaum entdeckt hat, erlauben Aussagen, über die Entwicklung von Attraktoren, wenn äußere Parameter durchfahren werden. Natürlich sind Poincaré-Schnitte immer sehr einfach darstellbar (2-dimensional!), sie sind häufig sehr ausssagekräftig und geben gute Anhaltspunkte für die Dimension des Attraktors.

Eine Modifikation der Poincaré-Schnitte ist die Maximum-Abbildung. Sie ist ganz besonders einfach (ohne Umweg über den Pseudophasenraum) zu gewinnen: Die Maxima der dynamischen Variablen werden ermittelt und dann 2-dimensional in der Form $\{\max_{n+1}, \max_n\}$ dargestellt. Bsp.: Rössler Attraktor (a=.15, b=.20 ,c=10.), Bild [Bild]. Die Maximum-Abbildung ist gewissermaßen eine Delay-Abbildung für den Durchschitt des Attraktors mit der (i.a. kompliziert gekrümmten) Fläche, die durch die Bedingung "Ableitung der dynamischen Variable = 0" bestimmt wird.

Vorhersagen über Yule-Walker Gleichungen

Aus den verschiedenen Möglichkeiten zur Vorhersage wurde das Verfahren der Anpassung an ein ein autoregressives Modell gewählt. Unter einen autoregressiven Modell der Ordnung p (AR(p)) versteht man ein stochastisches Modell der Form

$$x_t = \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \ldots + \alpha_p x_{t-p} + Z_t$$

dabei ist Z_n eine Zufallsgröße. Die sog. Yule-Walker Gleichungen sind das Gleichungssystem

$$\rho(1) = \alpha_1 \rho(0) + \alpha_2 \rho(-1) + \ldots + \alpha_p \rho(1-p)
\rho(2) = \alpha_1 \rho(1) + \alpha_2 \rho(0) + \ldots + \alpha_p \rho(2-p)
\vdots
\rho(p) = \alpha_1 \rho(p-1) + \alpha_2 \rho(p-2) + \ldots + \alpha_p \rho(0)$$

der Ordnung p, mit dem aus den Werten der Korrelationsfunktion ρ bei $\tau=0\ldots p$ die Parameter $\alpha_{1\ldots p}$ für ein AR(p) Modell ermittelt werden können (beachte, das $\rho(-\tau)=\rho(\tau)$). Über das AR(p)-Modell können aus jeweils p Werten Vorhersagen für die Zeitreihe gemacht werden. Damit die Vorhersage nicht i.a. eine einfache gedämpfte Schwingung ergibt, muß auch der Zufallsterm berücksichtigt werden. Dazu addiert man weißes Rauschen mit einer geeigneten Standardabweichung. Numerisch läßt sich weißes Rauschen durch Aufsummieren gaußverteilter Zufallszahlen simulieren. Je zwei unabhängige gaußverteilte Zufallszahlen z_1, z_2 können aus zwei gleichverteilten Zufallszahlen R_1, R_2 gewonnen werden durch die folgenden Beziehungen

$$z_1 = \sqrt{-2 \log(R_1)} \cos(2 \pi R_2)$$
$$z_2 = \sqrt{-2 \log(R_1)} \sin(2 \pi R_2)$$

(vgl. [2]).

Die Bild [5.1] zeigt eine Zeitreihe und die Vorhersage mit einem angepaßten AR(5)-Modell ohne Rauschen, [5.2] und [5.3] zwei Vorhersagen mit geeignetem Rauschen.

Dieses Programm ist als Mathematica-Package ausgeführt.

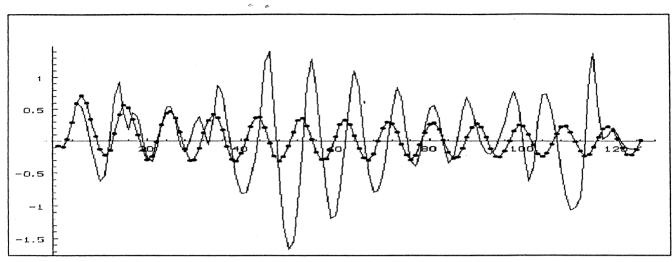


Bild [5.1] Zeitreihe und Vorhersage mittels AR(5)-Modell ohne Rauschen

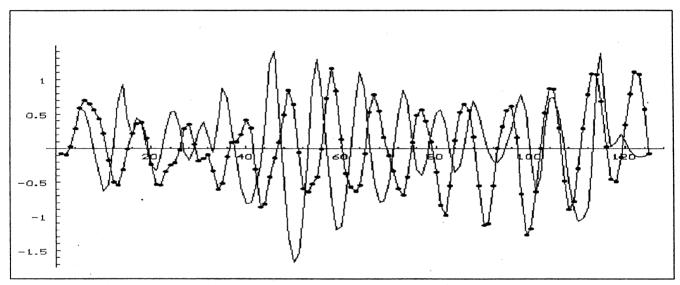


Bild [5.2] Zeitreihe und Vorhersage mittels AR(5)-Modell mit geeignetem Rauschen

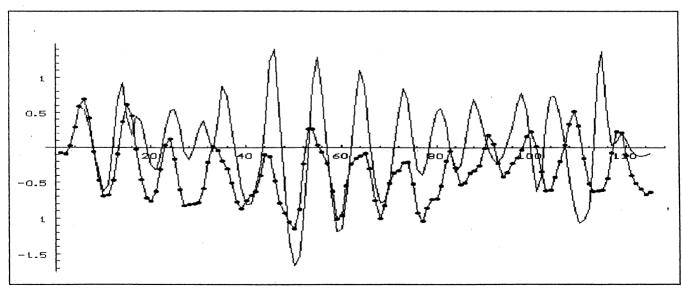


Bild [5.3] Zeitreihe und Vorhersage mittels AR(5)-Modell mit geeignetem Rauschen

Anwendungsbeispiele für die Programme

6.1 Das Rösslersystem

Das Rösslersystem ist ein dynamisches System, das durch die nichtlinearen gewöhnlichen gekoppelten Differentialgleichungen

$$\dot{x} = -y - z
\dot{y} = x + a y
\dot{z} = b + x z - c z$$

beschrieben wird. Für unterschiedliche Werte der Parameter a,b,c erhält man verschiedenes Verhalten. Das Rösslersystem ist ein "prototypisches" und ausgiebig untersuchtes chaotisches System. Es soll hier vor allem deshalb benutzt werden, weil man den Attraktor auch direkt durch Integration der Differentialgleichungen gewinnen kann und so einen Vergleich für die Attraktorrekonstruktionen aus Zeitreihen hat. Die Zeitreihen wurden mittels einer Runge-Kutta Integration 4.Ordnung produziert.

Bild [6.1.1] zeigt drei Ansichten des Attraktors für Parameterwerte a=.15, b=.20, c=10. Bild [6.1.2] zeigt die Histogramme für die drei dynamischen Variablen. Im folgenden wird die Zeitreihe der dynamischen Variablen x (vgl. Bild [6.1.3], Gesamtlänge 16384) betrachtet. Die "Abtastrate" war so gewählt, das pro Umlauf im Attraktor etwa 27 Werte für x gespeichert sind. Im Leistungsspektrum (Bild [6.1.4]) erkennt man deutlich die dominierende Frequenz ("Umlauffrequenz") und ihre Harmonischen. Die schwächeren Anteile jenseits der Grundfrequenz können durch logarithmieren hervorgehoben werden, vgl. Bild [6.1.5]. Die Autokorrelationskoeffizienten (Bild [6.1.6]) fallen nur äußerst langsam für größere τ . Selbst für 4096 Werte (entsprechend etwa 156 Schwingungen), bei der benutzten Zeitreihe vom Umfang 16384 die maximal sinnvoll mögliche Anzahl der Korrelationskoeffizienten, ist die "Amplitude" der AKF um weniger als zehn Prozent zurückgegangen. In Bild [6.1.7] sind drei Ansichten einer Attraktorrekonstruktion (Projektionen auf zueinander etwa senkrechte Ebebnen) mit Delay 4 abgebildet. Es läßt sich eine Ahnlichkeit mit dem wirklichen Attraktor (Bild [6.1.1]) erkennen. Die Verzerrung in des (rekonstruierten) z-"Ausschwingers" deutet auf eine zu niedrige Abtastrate hin. In diesem Bereich treten auch die größten Ungenauigkeiten bei der Integration der Differentialgleichungen auf. Man erkennt auch, daß es ein sinnloses Unterfangen wäre, die Punkte dieser Rekonstruktion verbinden zu wollen, auch

dazu wäre eine höhere Abtastrate vonnöten. Eine Maximum-Abbildung (nichtinterpolierte Maxima) zeigt Bild [6.1.8].

Bild [6.1.9] zeigt drei Ansichten des Attraktors für (a=0.10, b=0.10, c=8.0). Für diese Parameterwerte soll der Effekt der Delayzeit auf die Form des rekonstruierten Attraktors untersucht werden. Benutzt wird eine Zeitreihe der Variablen x (16384 Werte, ca. 27 Werte pro Umlauf im Attraktor). Bild [6.1.10] zeigt Attraktorrekonstruktionen für Delays 1,2,3. typisch ist, daß für kleine Delaywerte die Rekonstruktion zu einer länglichen Figur entartet (im Extremfall Delay=0 hätte man eine diagonale Strecke!). Bild [6.1.11] zeigt drei Ansichten der Rekonstruktion für Delay=4, die Ähnlichkeit mit dem wirklichen Attraktor ist unverkennbar. Hier wurde bewußt ein einfacher Attraktor gewählt, um die interesssierenden Effekte zu verdeutlichen, dadurch bedingt ist man hier bei fast allen Delaywerten geneigt, die topologische Äquivalenz mit dem Originalattraktor zu sehen. Bild [6.1.12] zeigt Rekonstruktionen für Delay=5,8,13. Im Falle Delay=5 gibt es eine "Schwebung" der Delayzeit mit der Umlaufzeit, so daß der Attraktor bruchstückhaft erscheint. Für Delay=8 erkennt man die Tendenz, daß sich die Rekonstruktionen für höhere Delaywerte zu "falten" beginnen. Der Fall Delay=13 zeigt, das eine ungünstige Delayzeit (Vielfaches der halben Umlaufszeit) die Rekonstruktion auf einen länglichen Bereich konzentriert. Zusätzlich ist hier die weitergehende Faltung zu bemerken. Schließlich sind in Bild [6.1.13] zwei Rekonstruktionen des Attraktors aus der z-Zeitreihe, die ein peakartiges Verhalten hat. Die typischen Verzerrungen sind gut zu erkennen (Delaywerte 1 und 4), verbesserte Rekonstruktionen ergeben sich nach Logarithmierung dieser Zeitreihe.

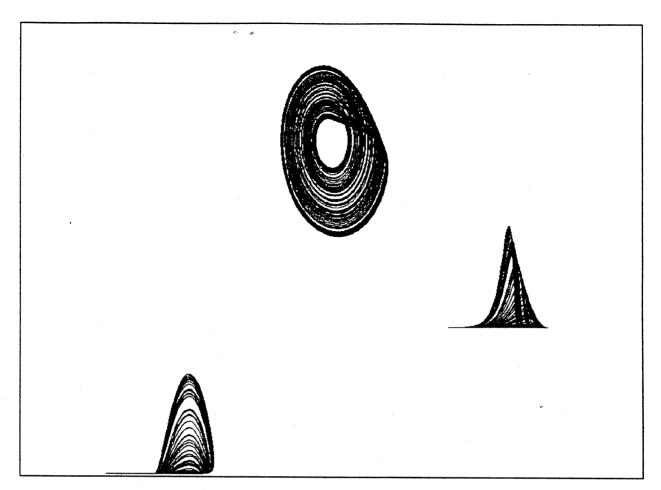


Bild [6.1.1] Rösslerattraktor (a=.15, b=.20, c=10.)

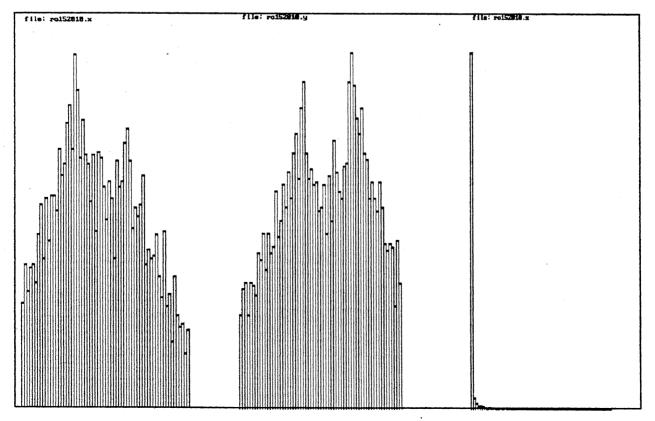


Bild [6.1.2] Histogramme für x,y,z

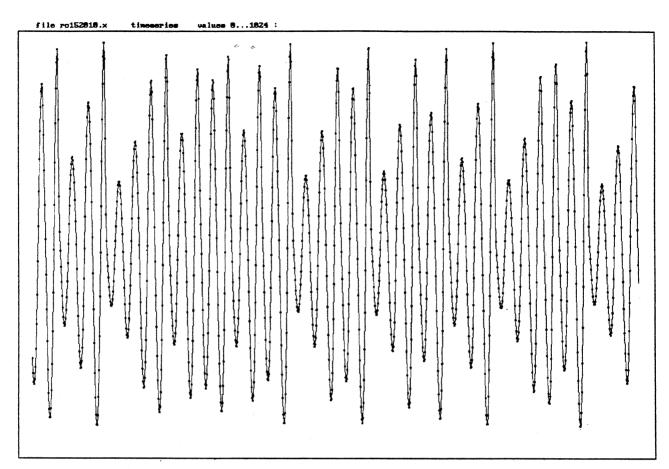


Bild [6.1.3] Ausschnitt aus der Zeitreihe für x

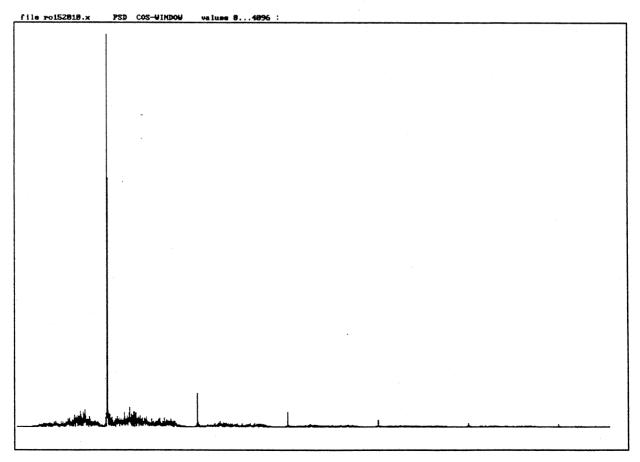


Bild [6.1.4] Leistungsspektrum

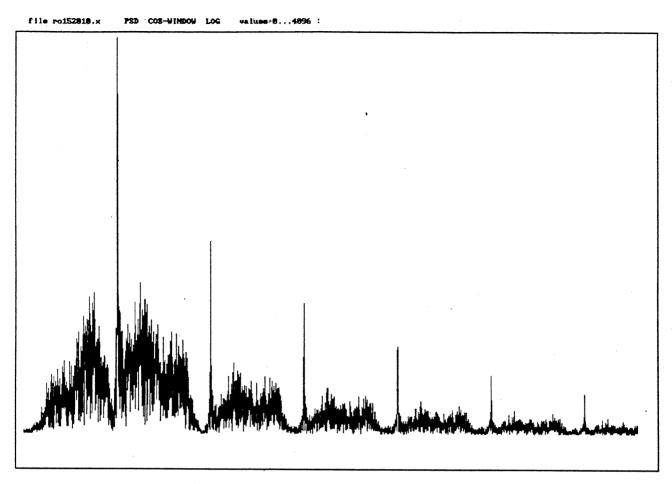
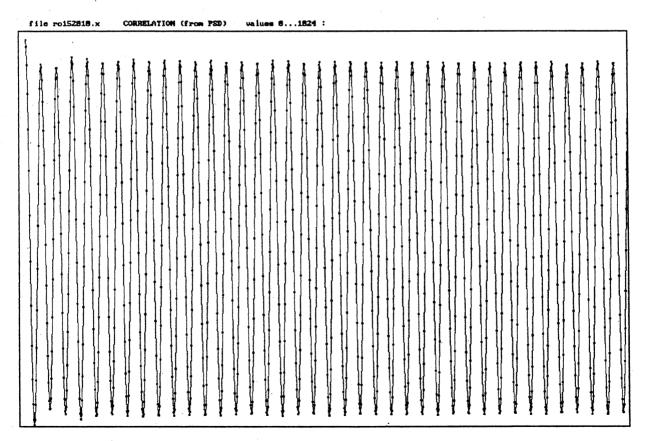


Bild [6.1.5] Leistungsspektrum (logarithmiert)



Bild~[6.1.6]~Autokorrelationkoeffizienten

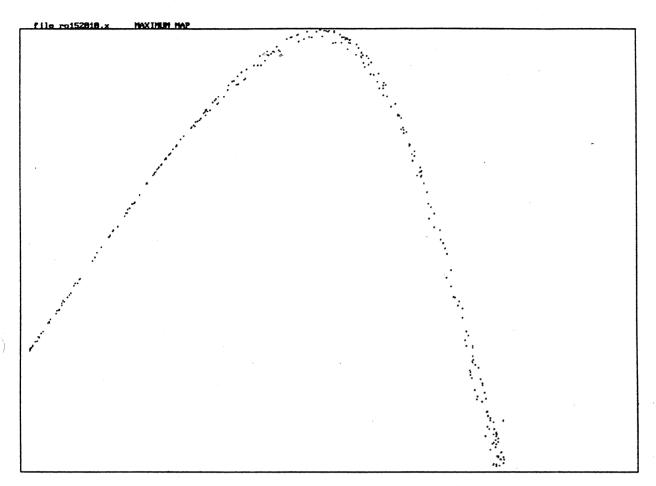


Bild [6.1.8] Maximumabbildung

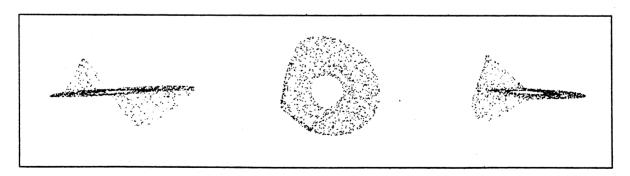


Bild [6.1.7] Drei Ansichten der Attraktorrekonstruktion für Delay=4

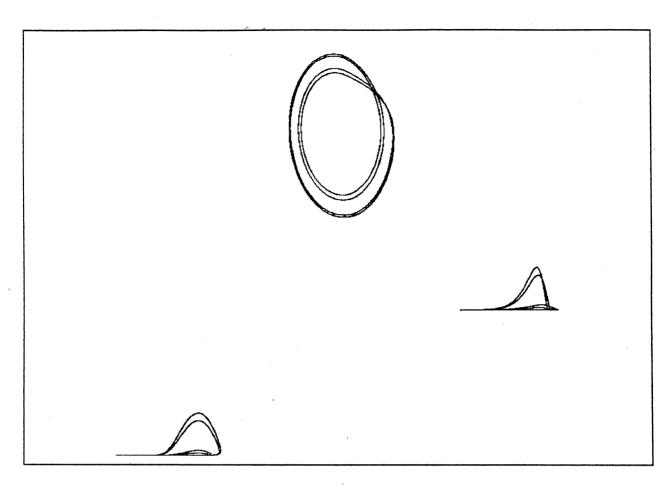


Bild [6.1.9] Rösslerattraktor (a=.10, b=.10, c=8.)

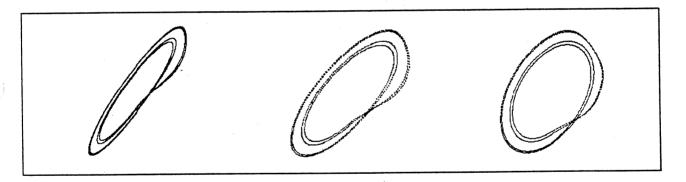


Bild [6.1.10] Attraktorrekonstruktionen für Delay=1,2,3 (a=.10, b=.10, c=8.)

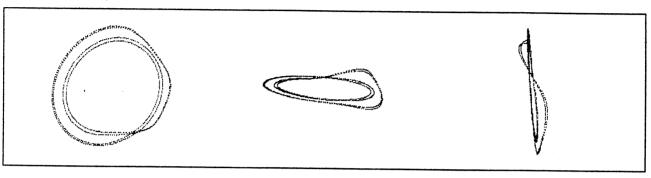


Bild [6.1.11] Drei Ansichten der Attraktorrekonstruktion für Delay=4

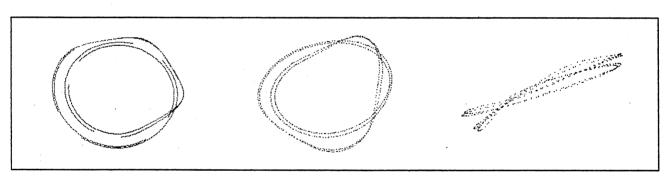


Bild [6.1.12] Attraktorrekonstruktionen für Delay=5,8,13

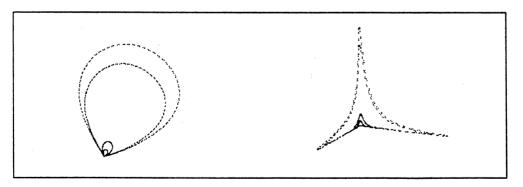


Bild [6.1.13] Attraktorrekonstruktionen aus der z-Zeitreihe

6.2 Simulationsdaten für einen Halbleiterlaser

Betrachtet wird hier eine Zeitreihe, die aus einer Computersimulation eines Zweistreifen-Halbleiterlasers gewonnen wurde. Die Zeitreihe besteht aus 4000 Werten, sie ist in Bild [6.2.1] wiedergegeben. Der Peakcharakter der Dynamik und die "exponentielle" Flanke im Histogramm (Bild [6.2.2]) legen eine Logarithmierung nahe. Tatsächlich läßt sich die Schwingung durch direktes Logarithmieren in eine ziemlich symmetrische Form überführen. Bild [6.2.3] zeigt einen Teil der logarithmierten Zeitreihe, Bild [6.2.4] das zugehörige Histogramm. Die Bilder [6.2.5] und [6.2.6] zeigen zwei Zufallsphasen-Surrogate der logarithmierten Zeitreihe. Im Leistungsspektrum (Bild [6.2.7]) würden durch Logarithmieren mit vorheriger Abbildung auf das Intervall [1 ... 10000] alle Unterschiede praktisch "ausgebügelt" (Bild [6.2.8]). Bild [6.2.9] zeigt die Autokorrelationsfunktion. Für eine Maximum-Abbildung gibt es zu wenig Extrema, man erhält keine erkennbare Struktur. Die Einbettung im Pseudophasenraum liefert für Delay=9 eine erkennbare flächenhafte Struktur, drei Ansichten davon sind in Bild [6.2.10] zu sehen.

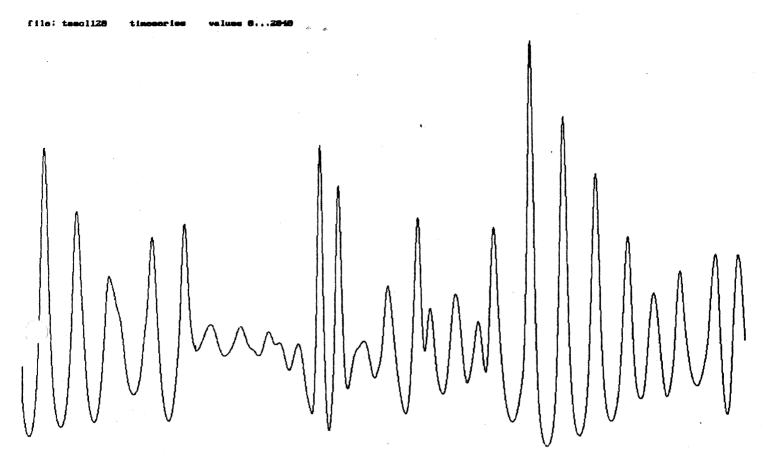
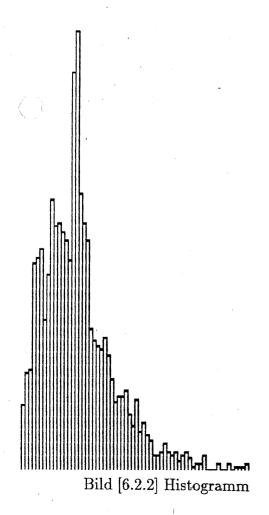


Bild [6.2.1] Ausschnitt aus der Zeitreihe





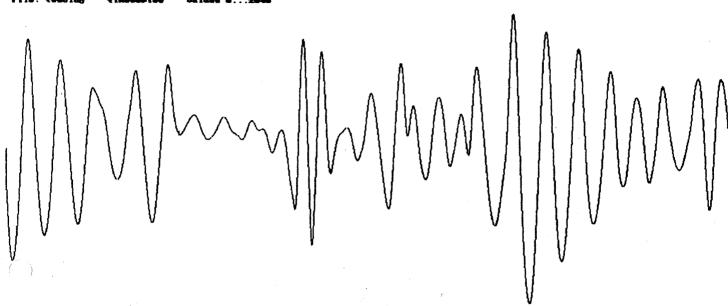


Bild [6.2.3] Ausschnitt aus der logarithmierten Zeitreihe

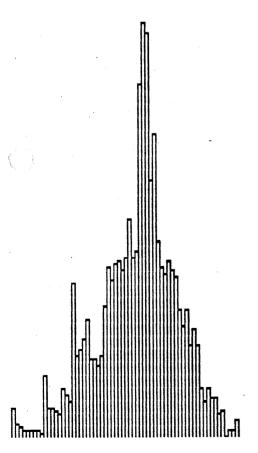
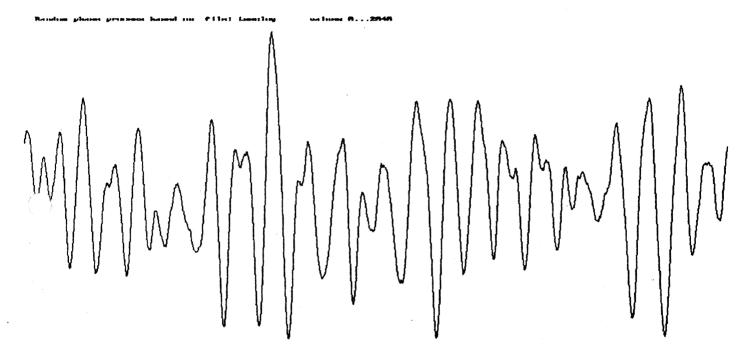
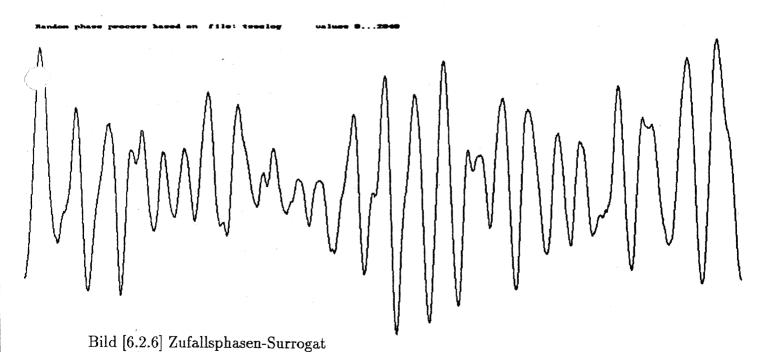


Bild [6.2.4] Histogramm der logarithmierten Zeitreihe



Bild~[6.2.5]~Zufallsphasen-Surrogat



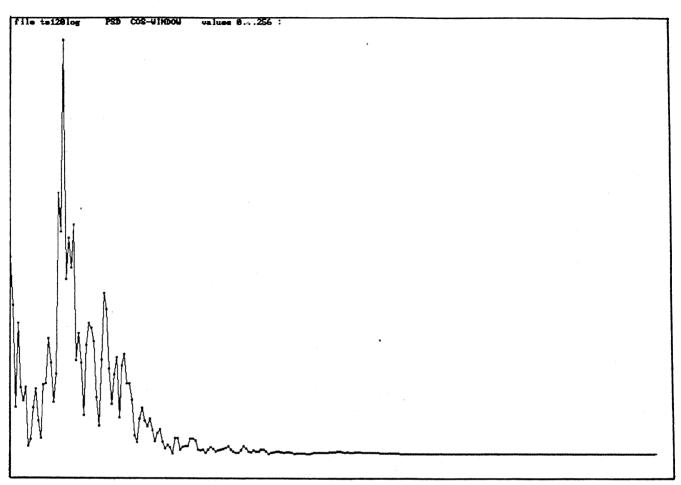


Bild [6.2.7] Leistungsspektrum, logarithmiert

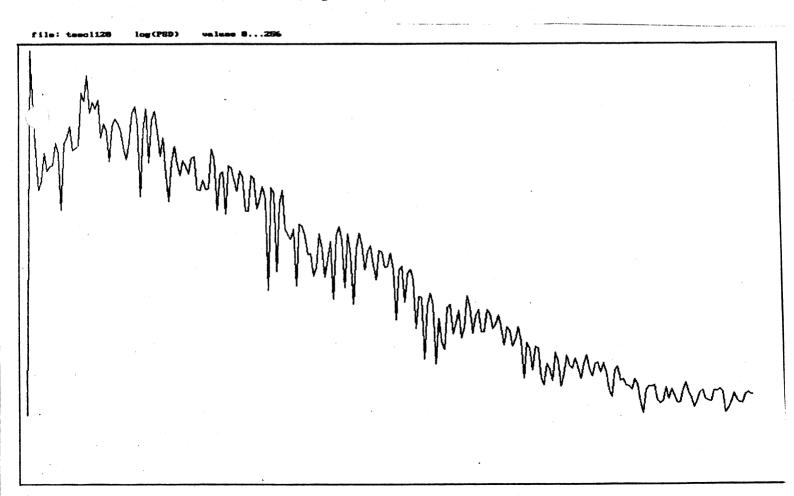


 Bild [6.2.8] Leistungsspektrum, das "zu stark" logarithmiert wurde

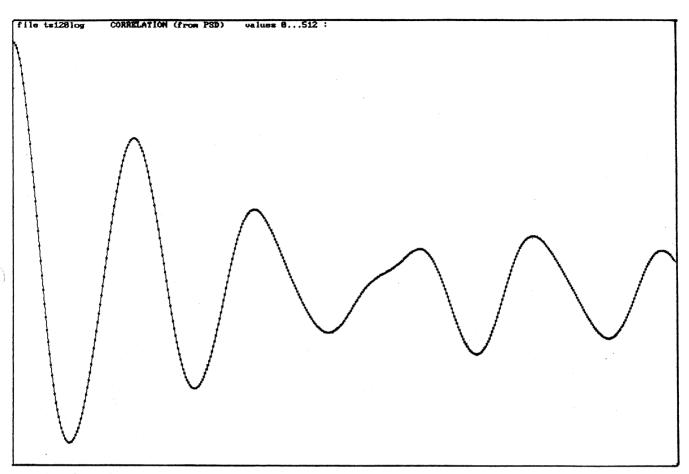


Bild [6.2.9] Autokorrelationskoeffizienten

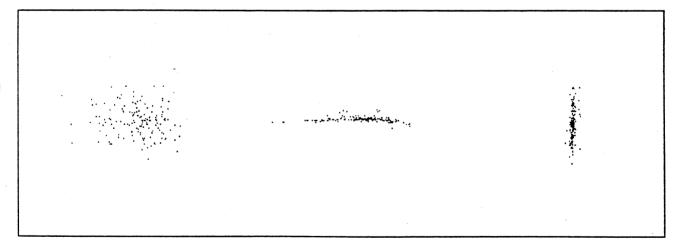


Bild [6.2.10] Drei Ansichten der Attraktorrekonstruktion für Delay=9

6.3 Analyse von Phonationszeitreihen

6.3.1 Aufnahme einer pathologischen Stimme

Es wird eine Zeitreihe betrachtet, die eine Stimmaufnahme des ausgehaltenen Vokales a (Lautschrift: [a:]) darstellt. Die Samplingfrequenz beträgt 44kHz, der Zeitraum von etwa 1.35 sec ergibt 60160 Datenwerte. Das Amplitude-Frequenz-Zeit Diagramm (Bild [6.3.1.1]) läßt erkennen, daß im letzten Drittel ein anderes Leistungsspektrum vorliegt als im Rest der Zeitreihe. Es soll nun der Unterschied des ersten Teils (etwa drei Fünftel am Anfang der Gesamtzeitreihe) und des letzten Teils (etwa zwei Fünftel am Ende der Zeitreihe) dargestellt werden. Auch die Histogramme der unterschiedlichen Teile legen nahe, die Zeitreihe grob in diese beiden Teile zu zerlegen (Bild [6.3.1.2]). Jeweils ein Ausschnitt der Teile der Zeitreihe ist in den Bildern [6.3.1.3] und [6.3.1.5] dargestellt. Die entsprechenden Autokorrelationskoeffizienten sind in den Bildern [6.3.1.4] und [6.3.1.6] wiedergegeben, besonders auffällig erscheint der extrem geringe Abfall der Korrelationskoeffizienten für den letzten Teil. Die Spektren sind in [6.3.1.7] bis [6.3.1.10] zu sehen, jeweils zusätzlich geglättete Spektren, in denen die Peaks zwischen der Grundfrequenz und ihren Harmonischen gut zu erkennen sind.

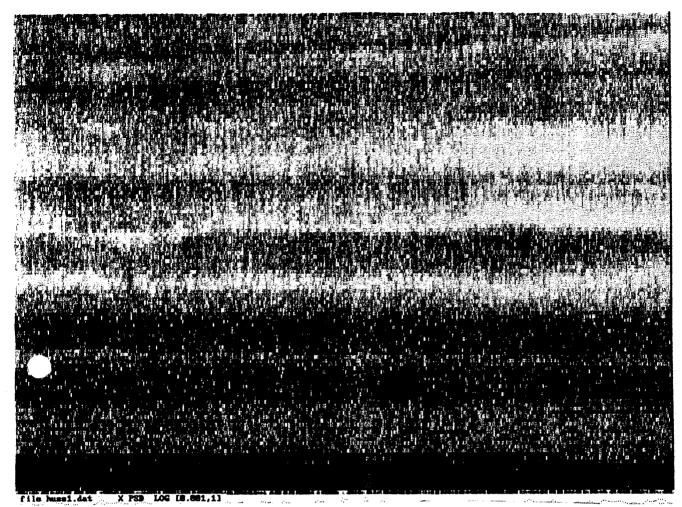


Bild [6.3.1.1] Amplitude-Frequenz-Zeit Diagramm (Dichteplot)

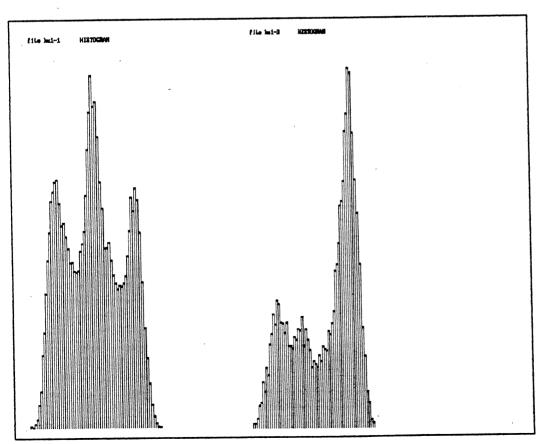


Bild [6.3.1.2] Histogramme für vorderen und hinteren Teil der Zeitreihe

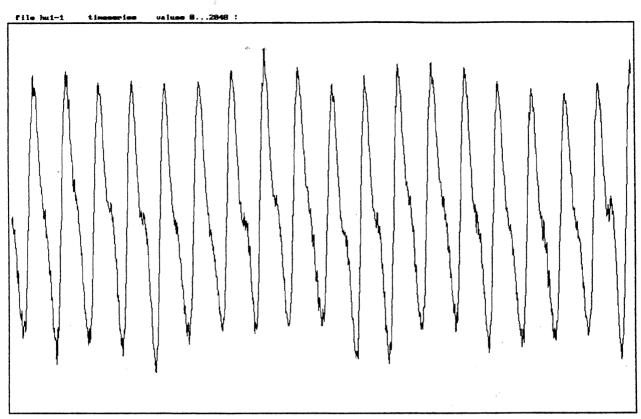


Bild [6.3.1.3] Ausschnitt aus dem vorderen Teil der Zeitreihe

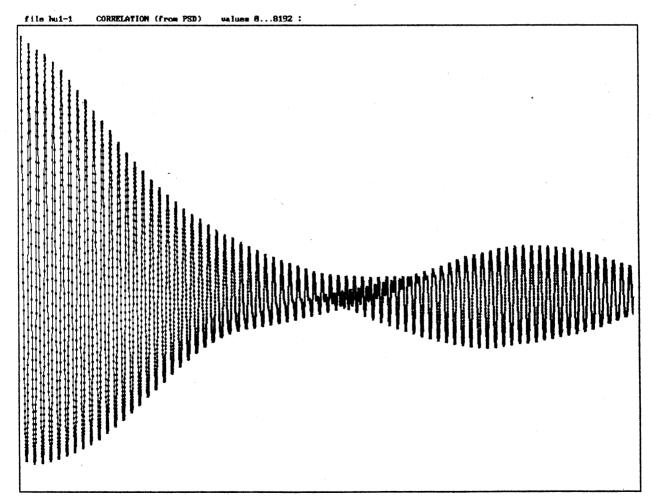


Bild [6.3.1.4] Autokorrelationskoeffizienten aus dem vorderen Teil der Zeitreihe

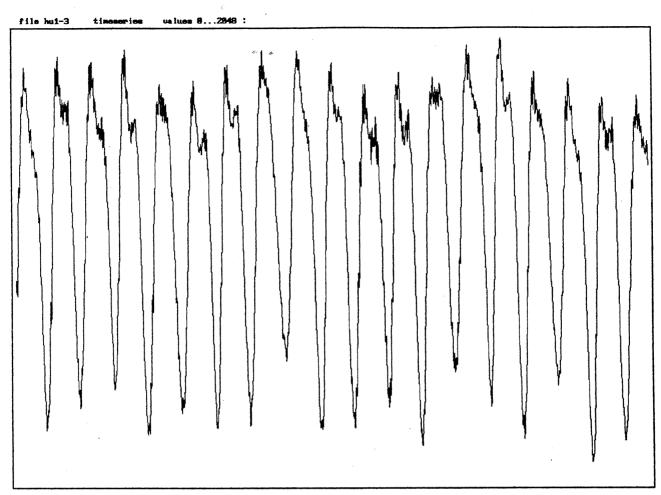


Bild [6.3.1.5] Ausschnitt aus dem hinteren Teil der Zeitreihe

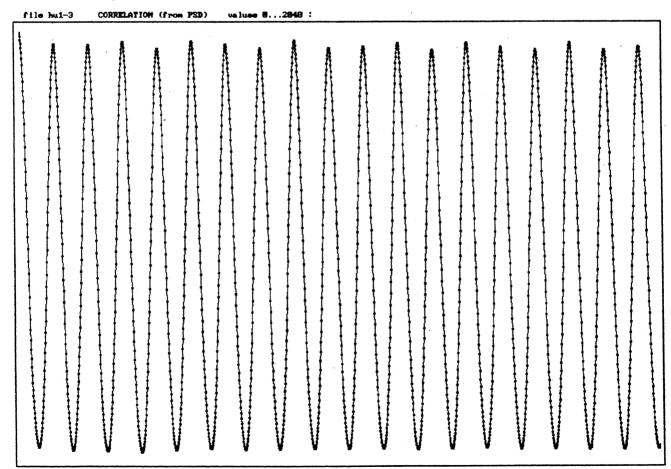


Bild [6.3.1.6] Autokorrelationskoeffizienten aus dem hinteren Teil der Zeitreihe

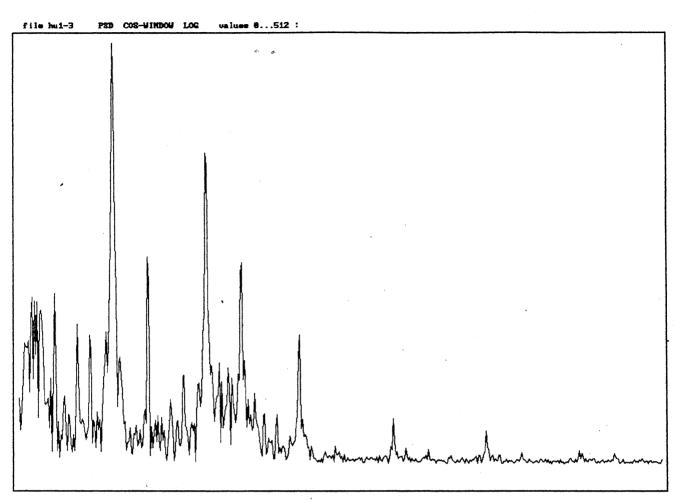


Bild [6.3.1.7] Spektrum aus dem vorderen Teil der Zeitreihe

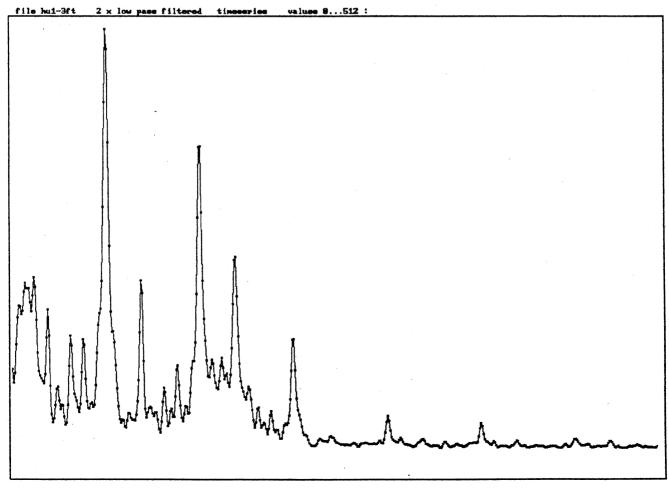


Bild [6.3.1.8] geglättetes Spektrum aus dem vorderen Teil der Zeitreihe

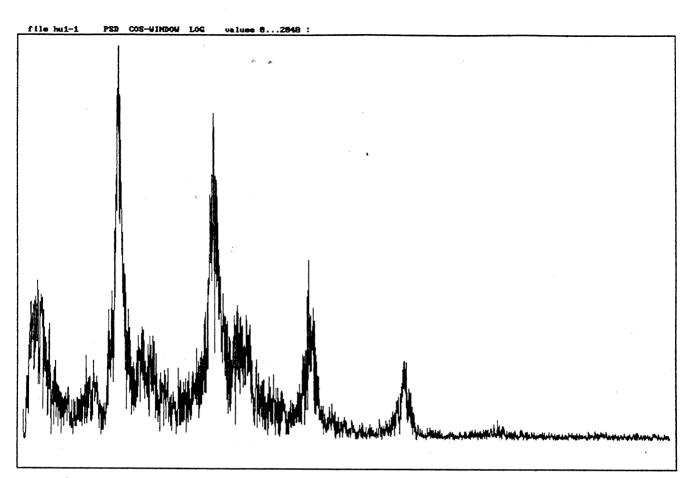
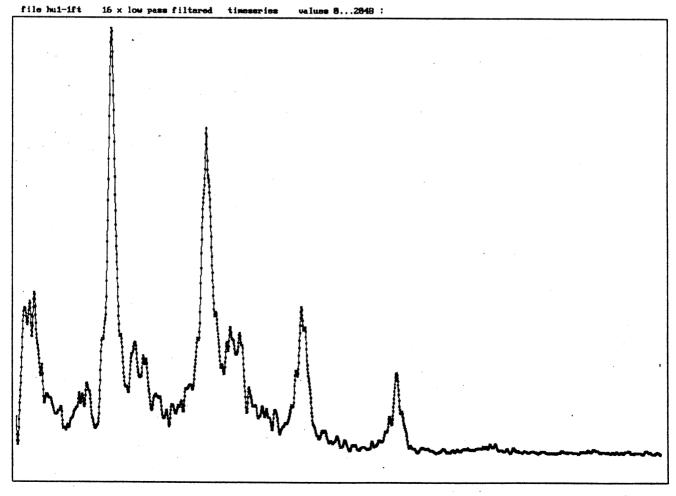


Bild [6.3.1.9] Spektrum aus dem hinteren Teil der Zeitreihe



 $\operatorname{Bild}\ [6.3.1.10]$ geglättetes Spektrum aus dem hinteren Teil der Zeitreihe

6.3.2 Computersimulation der menschlichen Stimmlippen

Bei der Zeitreihe handelt es sich um die Amplitude der wichtigsten "Normalschwingung" (empirische Orthogonalfunktion) eines Modells der menschlichen Stimmlippen, die aus einer Computersimulation gewonnen wurde. Dazu wurden Modelle der Stimmlippen triangularisiert und diejenigen partiellen (viskoelastischen) Differentialgleichungen numerisch integriert, die eine gedachte Anregung des Systems mit durchströmender Luft beschreiben. Dabei wurde das Elastizitätsmodul E der Stimmlippen schrittweise erhöht. Die Samplingrate ist 1 kHz, das E-Modul wurde beginnend mit 0.35 kPa in Schritten von 0.005 kPa (nach je 401 Datenpunkten, entsprechend einer Zeit von 0.4 sec) erhöht, der Endwert ist 0.5 kPa. Die Zeitreihe umfasst somit 12832 Werte, entsprechend einer Gesamtzeit von etwa 12.8 Sekunden. Im Bild [6.3.2.1] sieht man den Beginn der Zeitreihe, das Einschwingverhalten ist deutlich zu erkennen, es stellt sich eine ziemlich reguläre Schwingung ein. Bilder [6.3.2.2] und [6.3.2.3] zeigen spätere Ausschnitte der Zeitreihe, man erkennt die zunehmende Irregularität des Signals. Bild [6.3.2.4] zeigt das ungeglättete Leistungsspektrum; nach 64 (1/4,1/2,1/4)-Filterungen ergibt sich Bild [6.3.2.5]. Die Autokorrelationskoeffizienten sind in Bild [6.3.2.6] wiedergegeben. Das Histogram ist in Bild [6.3.2.7] zu sehen. Diese aus der gesamten Zeitreihe gewonnenen Größen sind natürlich wegen der Parameteränderung während der Zeitreihe nur von sehr begrenzter Aussagekraft. Hier bietet sich an, ein Amplitude-Frequenz-Zeit Diagramm zu erstellen. Ein solches ist, zum einen als Dichteplot in Bild [6.3.2.8], zum anderen als "Wasserfall-Diagramm" in Bild [6.3.2.9] dargestellt. Die Darstellungen wurden logarithmiert, die Fensterbreite ist 256, es sind alle 128 Frequenzwerte dargestellt. Eine aus diesem Signal gewonnene Attraktorrekonstruktion sollte, da die gesamte Dynamik eines Systems sich aus jeder beliebigen dynamischen Variable rekonstruieren lässt, die Gesamtdynamik beschreiben. Bild [6.3.2.10] zeigt Delay-Abbildungen für den Delay 1, der wegen der niedrigen Samplingrate schon gute Rekonstruktionen liefert. Es wurde jeweils ein Teil der Zeitreihe benutzt, spätere Teilbilder der Abbildung entsprechen Teilen der Zeitreihe, die weiter zum Ende liegen. In Bild [6.3.2.11] sind die Delay-Abbildungen für den Delay 2 zu sehen. Die beginnende Tendenz, für höhere Delaywerte einen gefalteten Attraktor zu produzieren, kann man in Bild [6.3.2.12] erkennen (Delay 4).

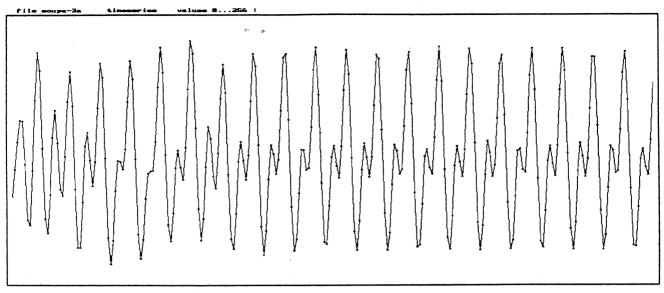


Bild [6.3.2.1] Ausschnitt aus der Zeitreihe

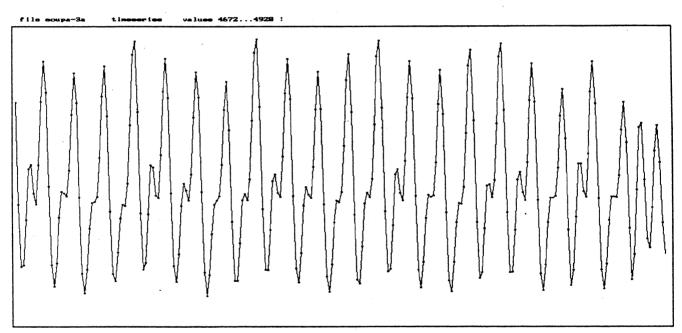


Bild [6.3.2.2] Ausschnitt aus der Zeitreihe

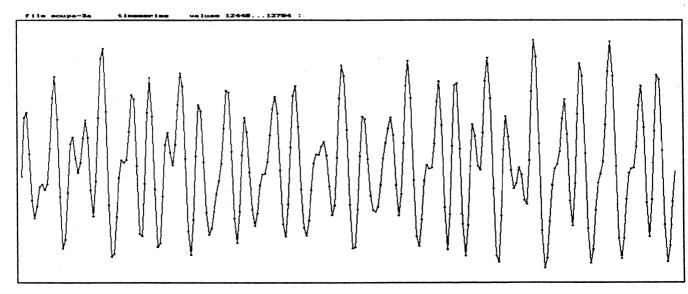


Bild [6.3.2.3] Ausschnitt aus der Zeitreihe

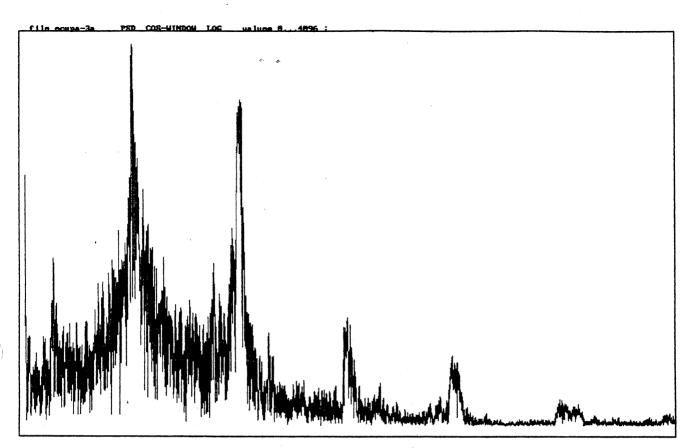


Bild [6.3.2.4] Leistungsspektrum, logarithmiert

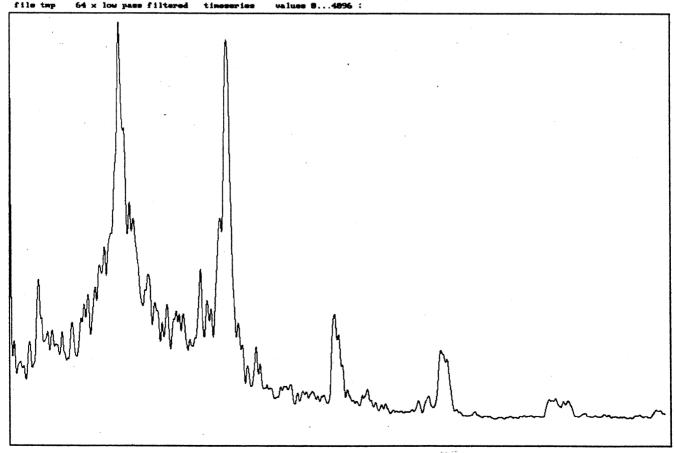


Bild [6.3.2.5] Leistungsspektrum, logarithmiert und geglättet

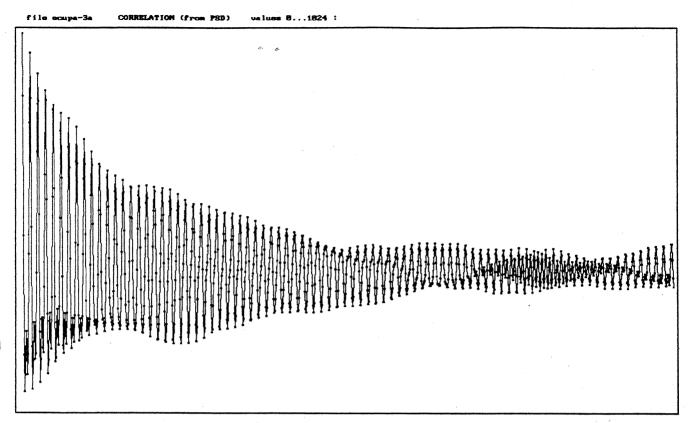


Bild [6.3.2.6] Autokorrelationskoeffizienten

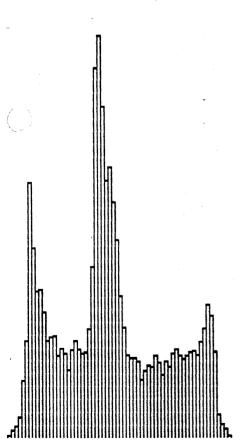


Bild [6.3.2.7] Histogramm

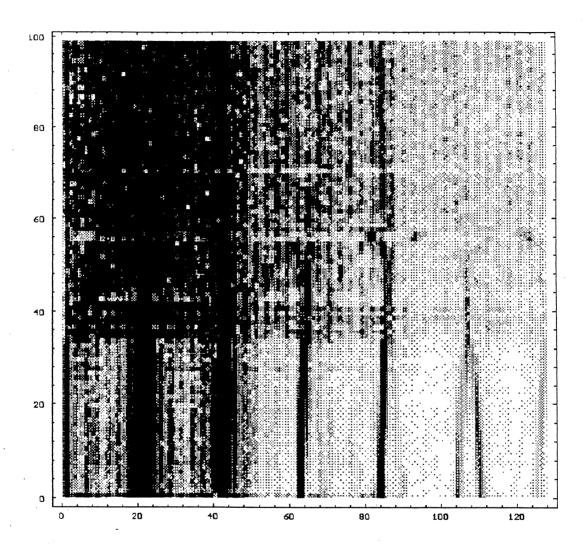


Bild [6.3.2.8] Amplitude-Frequenz-Zeit Diagramm (Dichteplot)

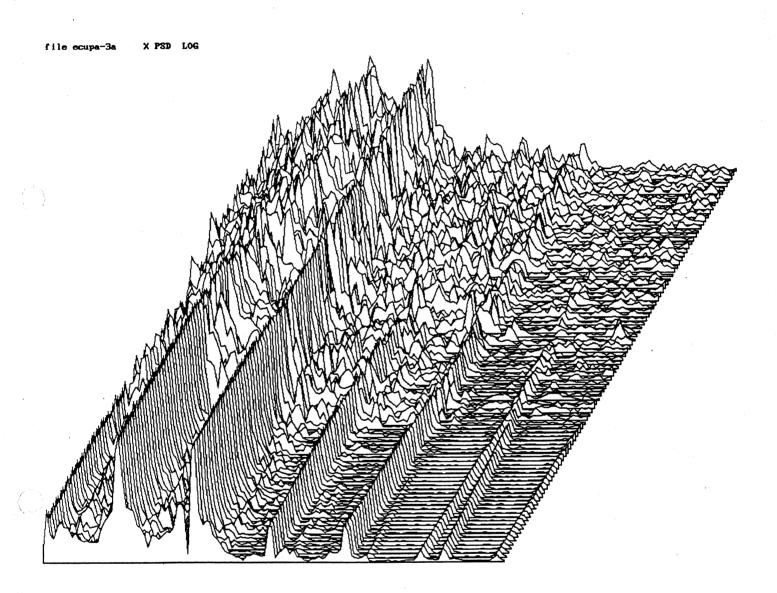


Bild [6.3.2.9] Amplitude-Frequenz-Zeit Diagramm ("Wasserfall" -Darstellung)

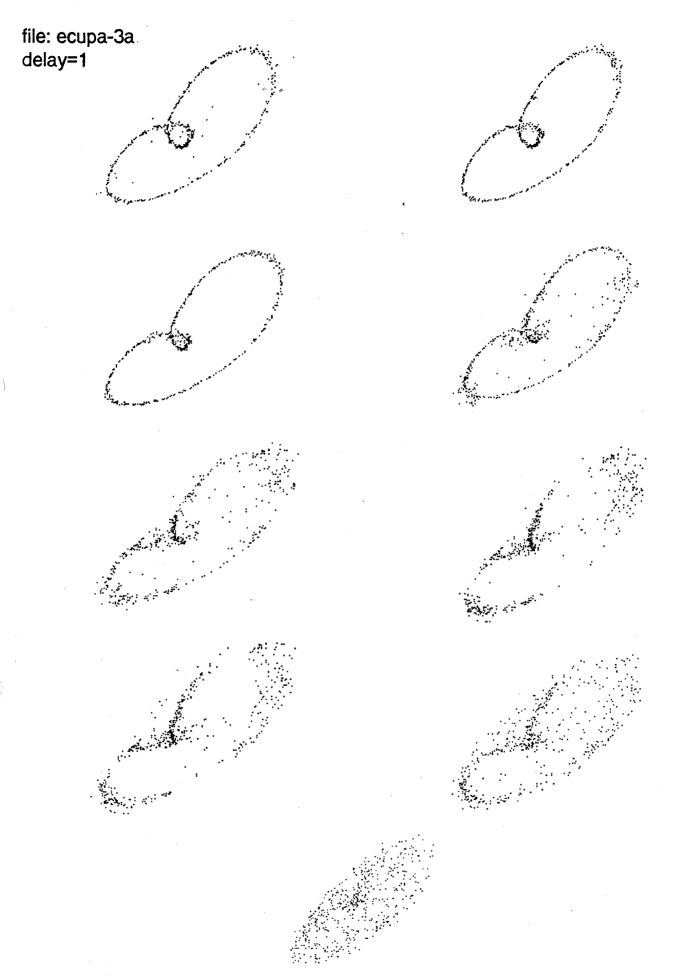


Bild [6.3.2.10] Attraktorrekonstruktionen für Delay=1 (aus verschiedenen Teilen der Zeitreihe)

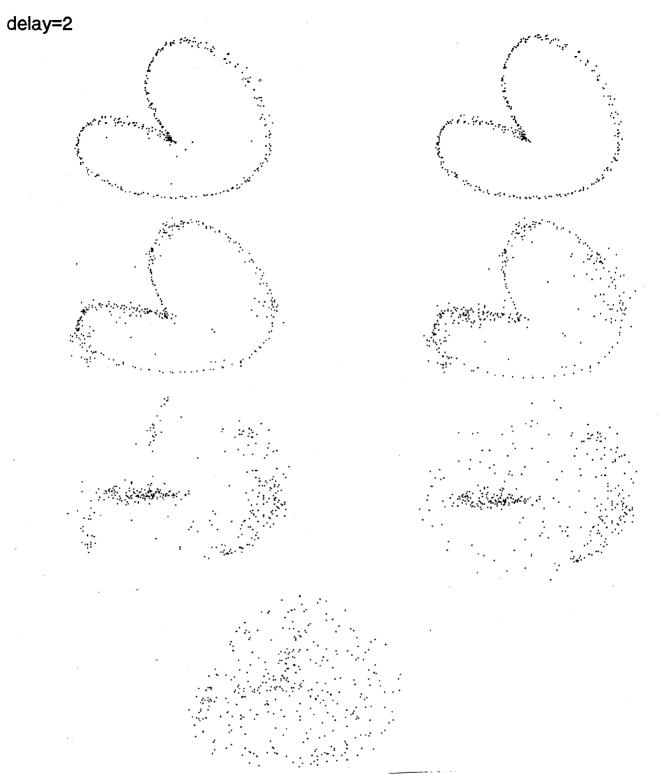


Bild [6.3.2.11] Attraktorrekonstruktionen für Delay=2 (aus verschiedenen Teilen der Zeitreihe)



Bild [6.3.2.12] Attraktorrekonstruktion für Delay=4