Nichtlineare Dynamik und Zeitreihenanalyse

— 3. Übung —

1 AFK von überlagerten Sinuszeitreihen

Betrachtet wird die Summe zweier Zeitreihen

$$f(t) = A_1 \sin(\omega_1 t) \qquad g(t) = A_2 \sin(\omega_2 t)$$

mit den Parametern

$$A_1 = 1 \qquad A_2 = 1$$
$$f_1 = 10 \text{ Hz} \qquad f_2 = 30 \text{ Hz}$$

Ferner gilt die Beziehung

$$\omega_n = 2\pi f_n \qquad n \in \{1, 2\}$$

Mittels der Funktion *Edit/Data sets/Edit/Create new/By formula* von Grace wurden diese Zeitreihen für eine Länge von 2 Sekunden und einer Abtastfrequenz (*sampling frequency*) $f_s = 5000$ Hz erzeugt. Für *Start load at* wurde der Wert 0, für *Stop load at* und *# of points* der Wert 10 000 eingetragen, in das Feld *X=* wurde '\$t' eingetragen, in das Feld *Y=* der Ausdruck 'sin$(2 * pi * 10 * \$t/5000)$'. Dieser Ausdruck erzeugt eine Zeitreihe mit 10000 Datenpunkten einer Sinusfunktion mit der Amplitude $A_1 = 1$, der Frequenz $f_1 = 10$ Hz

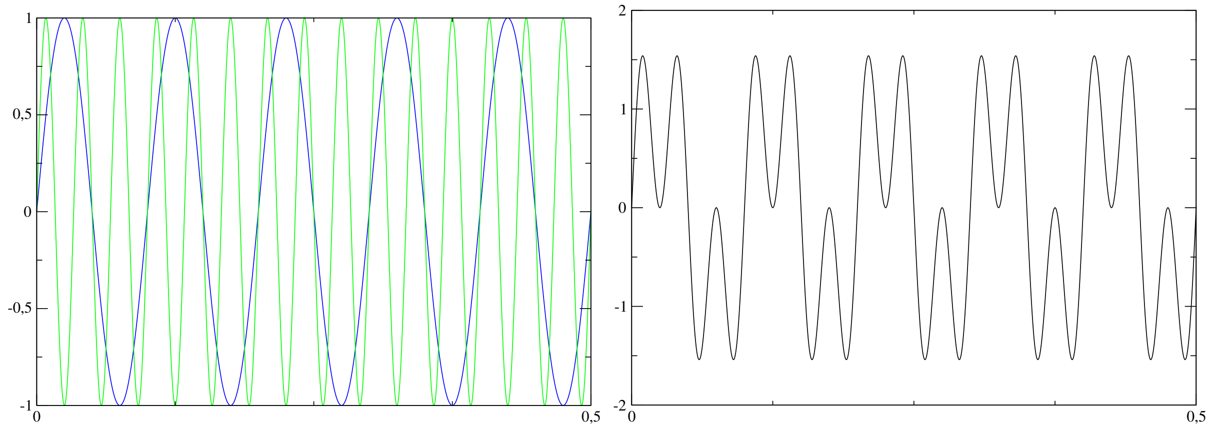

Abb. 1.1: Überlagerte Sinuszeitreihen. Auftragung der Amplitude gegen die Zeit (t/sec). Links sind die beiden Sinuszeitreihen gemeinsam in einen Graphen geplottet: $f(t) = \sin(2\pi \cdot 10t)$ und $g(t) = \sin(2\pi \cdot 30t)$. Die Zeitreihen haben eine Länge von 2 Sekunden bei einer Samplingfrequenz von $f_s = 5000$ Hz. Rechts die Summe der beiden Zeitreihen. Aufgetragen ist aus Gründen der Auflösung jeweils nur ein viertel der Zeitreihe.

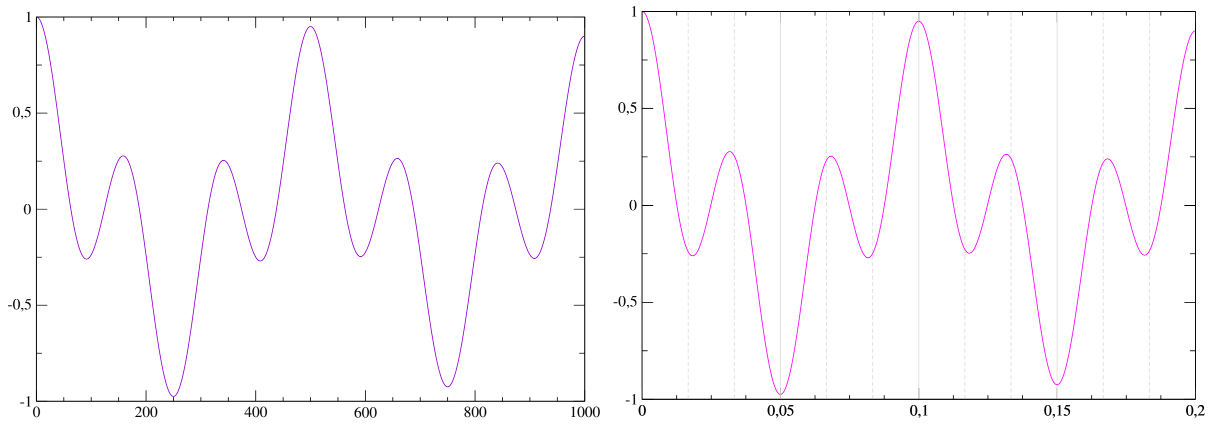

Abb. 1.2: Autokorrelationsfunktion (AKF) $C(k)$ der überlagerten Sinuszeitreihen. Für eine Darstellung der ursprünglichen Zeitreihe vgl. Abb. 1.1 (rechter Graph). Links die normale Darstellung der AKF, rechts mit auf Sekunden normierter Abszisse, um die Periodendauer besser ablesen zu können. Sehr gut erkennbar ist die langsamere Schwingung von 10 Hz und die überlagerte schnellere Schwingung von 30 Hz.

und der Abtastfrequenz $f_s = 5000$ Hz. Analog gilt für die zweite Sinuszeitreihe die Funktion '$\sin(2*pi*30*\$t/5000)$'. Die Abszisse wurde so skaliert, daß sie die Zeit in Sekunden anzeigt (*Data/Transformations/Evaluate expression*, '$X = 0.0002 * x$').

Beide Zeitreihen wurden addiert (*Data/Transformations/Evaluate expression*, '$y = sn.y+y$' mit n als Index der anderen Sinuszeitreihe) und das Ergebnis ebenfalls geplottet (vgl. Abb. 1.1).

Anschließend wurde die Autokorrelationsfunktio (AKF) $C(k)$ der überlagerten Sinuszeitreihen berechnet und ebenfalls dargestellt (vgl. Abb. 1.2). Hier wurde ebenfalls die Abszisse so skaliert, daß die Periodendauer in Sekunden ablesbar ist (Abb. 1.2, rechts). Sehr gut wiedererkennbar sind hier die beiden Periodendauern $T_1 = \frac{1}{f_1} = 0.1$ s und $T_2 = \frac{1}{f_2} \approx 0.033$ s. Für die Bestimmung von T_2 empfiehlt sich die Mittelung über mehrere Perioden.

2 Korrelationskoeffizienten für periodische Signale mit Rauschen

Zum besseren Verständnis der Berechnung der Autokorrelationsfunktion (AKF) $C(k)$ wurden für die Zeitreihe *sn1000.10.dat* (Sinuszeitreihe mit $f = 1000$ Hz, $f_s = 44\,100$ Hz und $\sigma_R = 1.0$) und *sn1000.001.dat* (Sinuszeitreihe mit $f = 1000$ Hz, $f_s = 44\,100$ Hz und $\sigma_R = 0.01$) jeweils die iterierten Abbildungen

$$x_{n+\frac{T}{4}} - x_n \qquad\qquad x_{n+\frac{T}{2}} - x_n \qquad\qquad x_{n+T} - x_n$$

erstellt (mit der Periodendauer T der Sinuszeitreihe in der Einheit *Samples*, $T \approx 44$) und davon die Korrelationskoeffizienten bestimmt. Für die Darstellung der Ausgangszeitreihen vgl. Abb. 2.1, für die iterierten Abbildungen Abb. 2.2. Anschließend wurden die so erhaltenen Korrelationskoeffizienten $C(x, y)$ mit den Werten aus der Berechnung der AKF für die Ausgangszeitreihen (Abb. 2.3) verglichen.

Aus den iterierten Abbildungen ergaben sich (durch lineare Regression) folgende Werte für die jeweiligen Korrelationskoeffizienten $C_{\sigma_\mathrm{R}}(t, t+\tau)$ (mit der Standardabweichung des

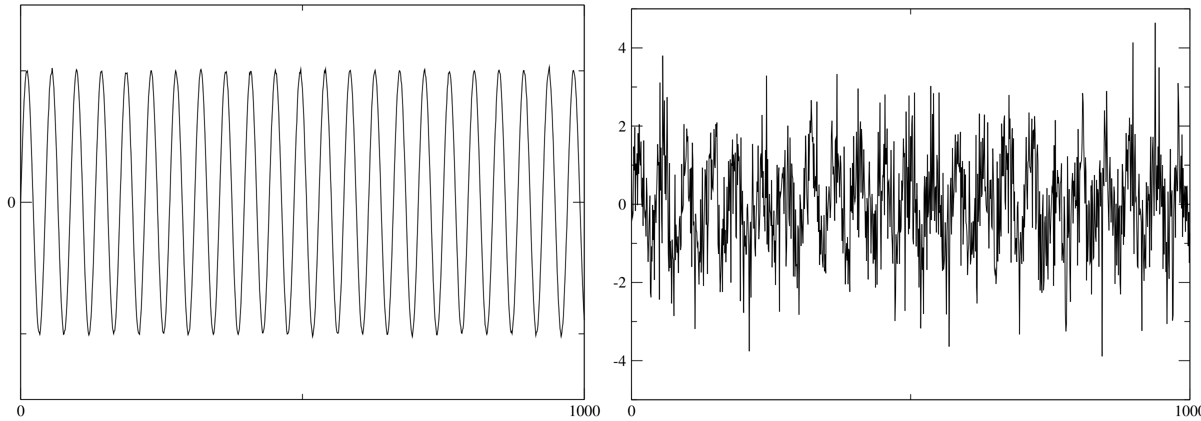

Abb. 2.1: Periodische Signale mit Rauschen. Beide Zeitreihen sind Sinuszeitreihen mit einer Frequenz von 1000 Hz und einer Abtastfrequenz von 44100 Hz mit verschieden starken Rauschanteilen. Die Standardabweichung σ_R der gaußverteilten Zufallszahlen des Rauschanteils betrug im linken Graphen $\sigma_\mathrm{R} = 0.01$, im rechten Graphen $\sigma_\mathrm{R} = 1.0$. Gezeigt ist aus Gründen der Auflösung nur das erste Zehntel der Zeitreihe (die ersten 1000 *samples*).

Rauschanteils σ_R und der Verschiebung τ):

$$\begin{aligned} C_{0.01}(t, t+0.25T) &= 0.0035508 & \qquad C_{1.0}(t, t+0.25T) &= 0.0020615 \\ C_{0.01}(t, t+0.5T) &= -0.9997782 & C_{1.0}(t, t+0.5T) &= -0.3378424 \\ C_{0.01}(t, t+T) &= 0.9997041 & C_{1.0}(t, t+T) &= 0.3439492 \end{aligned}$$

Sie wurden der Ausgabe des Programms Grace für die lineare Regression der jeweiligen iterierten Abbildung entnommen. Die entsprechenden korrespondierenden Werte für den *lag* ($k_T(\sigma_\mathrm{R})$) aus der Autokorrelationsfunktion der jeweiligen Zeitreihe ergaben sich zu:

$$\begin{aligned} k_{11}(0.01) &= 0.00354958 & \qquad k_{11}(1.0) &= 0.00197228 \\ k_{22}(0.01) &= -0.99757071 & k_{22}(1.0) &= -0.33716073 \\ k_{44}(0.01) &= 0.99530601 & k_{44}(1.0) &= 0.34270544 \end{aligned}$$

Ein Vergleich der jeweils korrespondierenden Werte zeigt, daß sie sich nicht signifikant voneinander unterscheiden. Das war zu erwarten, da die beiden Vorgehensweisen — Korrelationskoeffizient einer iterierten Abbildung und Autokorrelationsfunktion — einander entsprechen.

Die iterierten Abbildungen (Abb. 2.2) lassen mehrere Effekte sehr deutlich zu Tage treten. Nicht nur wird einmal mehr der Einfluß des Rauschens deutlich. Die iterierte Abbildung für ein periodisches Signal mit einer Verschiebung entsprechend einer viertel Periodenlänge ($\frac{T}{4}$) ergibt für ein reines oder nur wenig verrauschtes Sinussignal den Einheitskreis. Der Korrelationskoeffizient beträgt hier $C(x, y) = 0$, ist also identisch mit dem Korrelationskoeffizienten für ein völlig unkorreliertes Signal (z.B. weißes Rauschen). Trotzdem besteht hier sehrwohl eine Abhängigkeit zwischen den beiden Variablen x und y, die gegeneinander aufgetragen sind.

Die beiden anderen iterierten Abbildungen für die halbe bzw. volle Periodenlänge liefern sehr schöne Veranschaulichungen für maximale positive und maximale negative Korrelation, die Korrelationskoeffizienten liegen für das kaum verrauschte Signal bei nahezu ± 1.

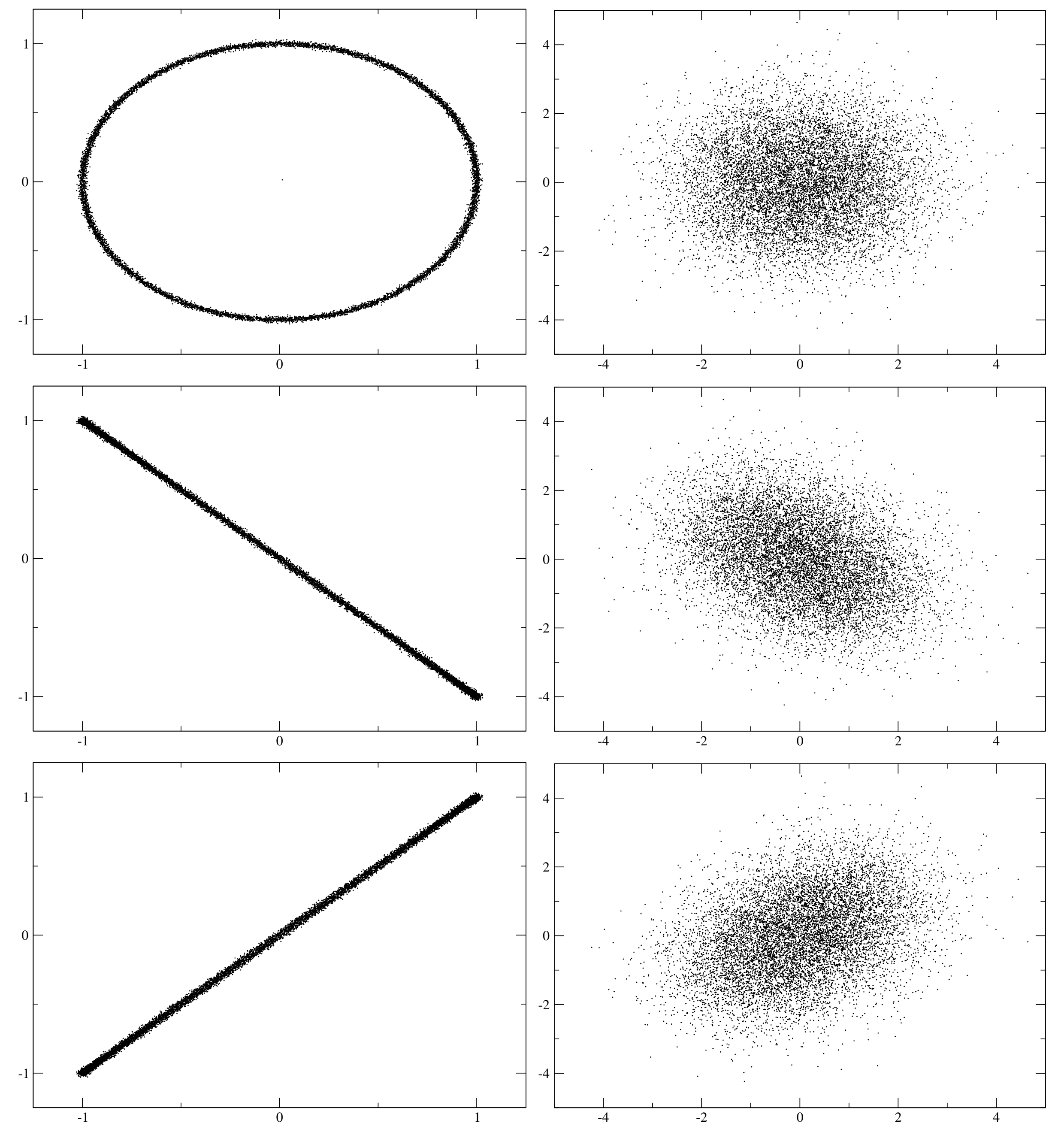

Abb. 2.2: Iterierte Abbildungen der periodischen Sinuszeitreihen mit Rauschen. Dargestellt sind im linken Teil der Abbildung die iterierten Abbildungen für die Zeitreihe *sn1000.001.dat* mit geringem Rauschanteil ($\sigma_{\mathrm{R}} = 0.01$), im rechten Teil für die Zeitreihe *sn1000.10.dat* mit hohem Rauschanteil ($\sigma_{\mathrm{R}} = 1.0$). Die Abbildungen entsprechen in ihrer Reihenfolge von oben nach unten folgenden Iterierten Abbildungen: $x_{n+\frac{T}{4}}$, $x_{n+\frac{T}{2}}$, x_{n+T} mit $T = 44$. Hinter dem Parameter T verbirgt sich die Periodendauer der Sinuszeitreihe in *samples*. Damit entsprechen die oberen Abbildungen der Iterierten Abbildung für ein viertel der Periodendauer (entspricht einer Phasenverschiebung um $\frac{\pi}{2}$), die mittleren für die halbe Periode (Phasenverschiebung von π) und die unteren für die Dauer einer Periode (Phasenverschiebung von 2π). Die Iterierte Abbildung eines reinen Sinussignals mit einer Verschiebung von $\frac{T}{4}$ entspricht dem Einheitskreis.

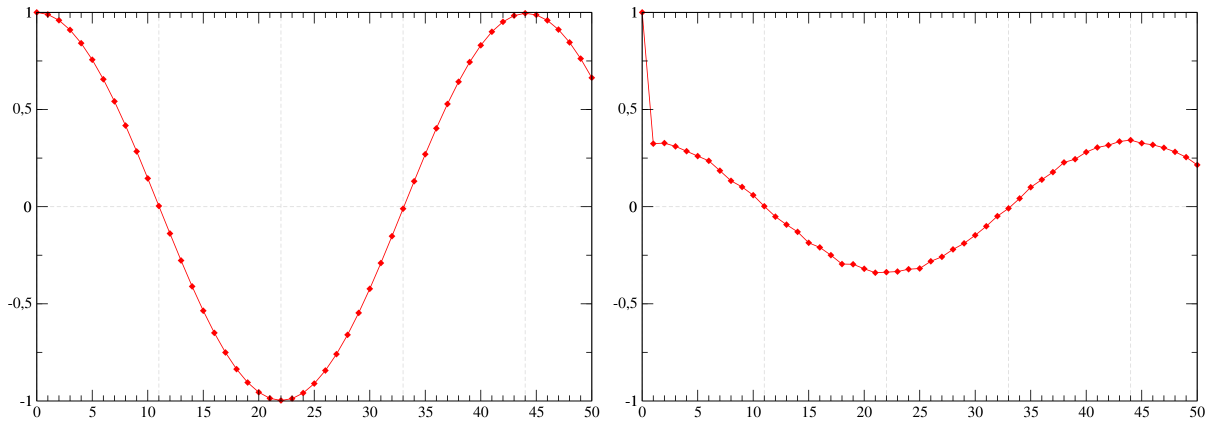

Abb. 2.3: Autokorrelationsfunktion (AKF) $C(k)$ der periodischen Signale mit Rauschen. Beide Zeitreihen sind Sinuszeitreihen mit einer Frequenz von 1000 Hz und einer Abtastfrequenz von 44100 Hz mit verschieden starken Rauschanteilen. Die Standardabweichung σ_{R} der gaußverteilten Zufallszahlen des Rauschanteils betrug im linken Graphen $\sigma_{\mathrm{R}} = 0.01$, im rechten Graphen $\sigma_{\mathrm{R}} = 1.0$. Gezeigt ist nur ein Ausschnitt der AKF, da zum Vergleich mit den Iterierten Abbildungen (Abb. 2.2) nur die Werte bis zu einem Lag $k = 44$ von Interesse sind.

3 AKF eines Stimmsignals

Die Datenreihe *vokale.dat* wurde in das Programm Grace geladen (vgl. Abb. 3.2). Sie besteht aus fünf Teilen: Rauschen, einem ersten, zweiten, dritten Vokal und wieder Rauschen. Das Stimmsignal wurde in seine fünf Teile zerlegt und die drei Vokalsignale einzeln analysiert.

Zur Zerlegung wurden fünf Regionen definiert (*Edit/Regions/Define*) und diese Regionen anschließend jeweils als eigenes *set* extrahiert (*Data/Transformation/Evaluate expression*, hier unter *Restriction* die jeweilige Region auswählen und ohne Eingabe einer Formel bestätigen) und im weiteren analysiert.

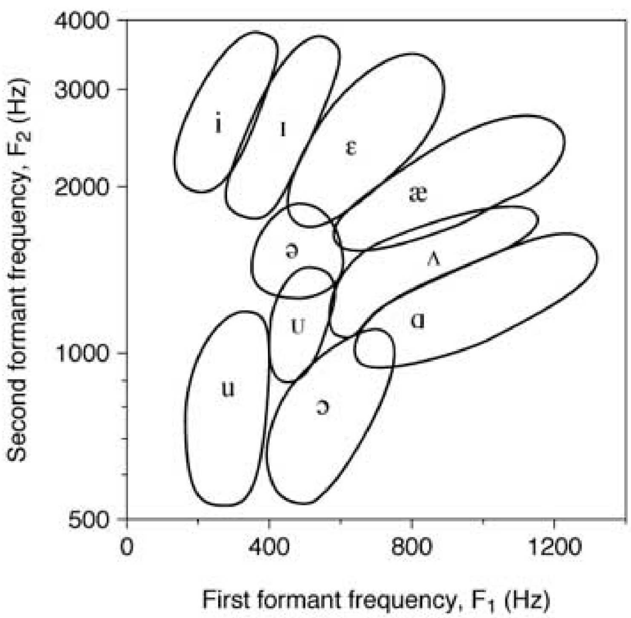

Abb. 3.1: Frequenzbereiche der ersten beiden Vokaltraktresonanzen (Formantfrequenzen) für verschiedene Vokale.

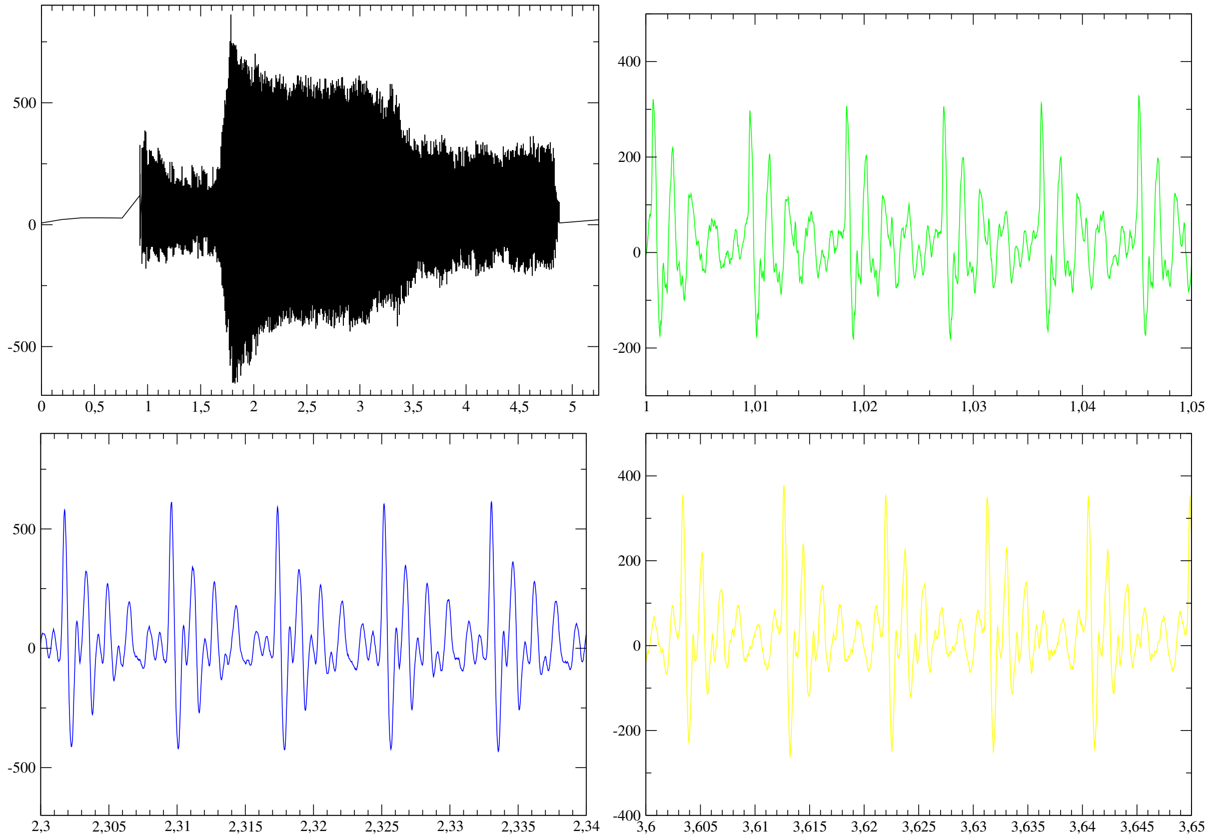

Abb. 3.2: Das Signal *vokale.dat* besteht aus fünf Teilen: Rauschen, einem ersten, zweiten, dritten Vokal und wieder Rauschen. Die Abbildungen zeigen der Reihe nach von links nach rechts und von oben nach unten: das Gesamtsignal gefolgt von den drei Vokalsignalen, wie sie aus dem Gesamtsignal extrahiert wurden. Die Abszisse ist in Sekunden skaliert. Für die drei Vokalsignale ist jeweils nur ein sehr kleiner Ausschnitt entsprechend gespreizt dargestellt, um die Bestimmung typischer Periodenlängen und daraus der Formantfrequenzen zu ermöglichen. Das Signal wurde mit einer Samplingfrequenz $f_s = 44\,444$ Hz aufgenommen.

Ziel der Analyse war, durch Extraktion typischer Frequenzen aus den Stimmsignalen mittels des Graphen mit der Darstellung der ersten und zweiten Formantfrequenzen (*first, second formant frequency*, F_1, F_2) auf die Vokale zu schließen.

Zunächst wurden per Hand aus dem Signal aus der Zeitauftragung (Abb. 3.2) typische Periodenlängen (Grundfrequenz und Resonanzfrequenzen des Vokaltraktes) versucht zu extrahieren. Das Signal wurde mit einer Samplingfrequenz $f_s = 44\,444$ Hz aufgenommen, die Abszisse ist in Sekunden skaliert. Als typische Frequenzen f_n ergaben sich für die einzelne Vokale voc1 bis voc3

$$
\begin{array}{lll}
f_1(\text{voc1}) = 105\ \text{Hz} & f_1(\text{voc2}) = 128\ \text{Hz} & f_1(\text{voc3}) = 108\ \text{Hz} \\
f_2(\text{voc1}) = 526\ \text{Hz} & f_2(\text{voc2}) = 518\ \text{Hz} & f_2(\text{voc3}) = 571\ \text{Hz} \\
f_3(\text{voc1}) = ?\ \text{Hz} & f_3(\text{voc2}) = 1075\ \text{Hz} & f_3(\text{voc3}) = 980\ \text{Hz}
\end{array}
$$

Bezeichnend ist hier, daß sich für den ersten Vokal nicht sinnvoll eine dritte Frequenz finden ließ. Das wird noch deutlicher bei der Auftragung der Autokorrelationsfunktion im nächsten Schritt.

Anschließend an die Analyse auf der Basis der Darstellung des ursprünglichen Signals (Abb. 3.2) wurde die Autokorrelationsfunktion (AKF) $C(k)$ jedes der drei Vokalsignale berechnet und ebenfalls dargestellt (Abb. 3.3, links). Die Abszisse wurde in Sekunden umskaliert, um die Perioden direkt ablesen und einfach in Frequenzen umrechnen zu können. Als typische Frequenzen f_n ergaben sich für die einzelne Vokale voc1 bis voc3 aus der Darstellung der AKF:

$$\begin{array}{lll} f_2(\text{voc1}) = 525 \text{ Hz} & f_2(\text{voc2}) = 639 \text{ Hz} & f_2(\text{voc3}) = 533 \text{ Hz} \\ f_3(\text{voc1}) = ? \text{ Hz} & f_3(\text{voc2}) = 1278 \text{ Hz} & f_3(\text{voc3}) = 1173 \text{ Hz} \end{array}$$

Dargestellt sind hier nur die extrahierten Frequenzen f_2 und f_3, da die Frequenz f_1 (die Grundfrequenz) keine Vokaltraktfrequenz (Formantfrequenz, *formant frequency*) ist und daher für die Bestimmung des Vokals aus den Daten keine Rolle spielt. Wie schon bei der Analyse auf der Basis der Darstellung der ursprünglichen Zeitreihe angesprochen, ließ sich für den ersten Vokal keine Frequenz deutlich jenseits der 500 Hz bestimmen.

Erst die sich anschließende zusätzliche Fourier–Analyse der AKF (Abb. 3.3, rechts) brachte die gewünschte Klarheit: Beim ersten und dritten Vokalsignal liegen die beiden Formantfrequenzen nahezu aufeinander, daher konnte beim ersten Signal auch keine weitere Frequenz extrahiert werden. Dagegen sind beim zweiten Vokalsignal sehr deutlich die beiden Formantfrequenzen erkennbar. Für die drei Vokale ergaben sich aus der durch Fourier–Transformation erzeugten Darstellung des Leistungsspektrums (*power spectrum*) folgende Frequenzen für die erste (F_1) und zweite (F_2) Formantfrequenz:

$$\begin{array}{lll} F_1(\text{voc1}) = 535 \text{ Hz} & F_1(\text{voc2}) = 640 \text{ Hz} & F_1(\text{voc3}) = 526 \text{ Hz} \\ F_2(\text{voc1}) = 561 \text{ Hz} & F_2(\text{voc2}) = 1276 \text{ Hz} & F_2(\text{voc3}) = 635 \text{ Hz} \end{array}$$

Demzufolge handelt es sich nach Vergleich mit der Darstellung der ersten und zweiten Formantfrequenz (Abb. 3.1) beim ersten und dritten Vokalsignal um den Vokal [ɔ], beim zweiten Vokalsignal dagegen um den Vokal [ɑ].

Dieses Ergebnis wurde dadurch bestätigt, daß das Signal in eine *wav*–Datei rückumgewandelt und akustisch überprüft wurde. Dazu wurden aus der Datenreihe *vokale.dat* mittels perl–Skript die Frequenzwerte extrahiert und diese dann in das Programm scilab (scilab-2.6, Copyright (C) 1989-2001 INRIA) eingelesen und über die Funktion *savewave* mit entsprechender *sampling rate* ($f_s = 44\,444$ Hz) in eine *wav*–Datei gespeichert und abgespielt.

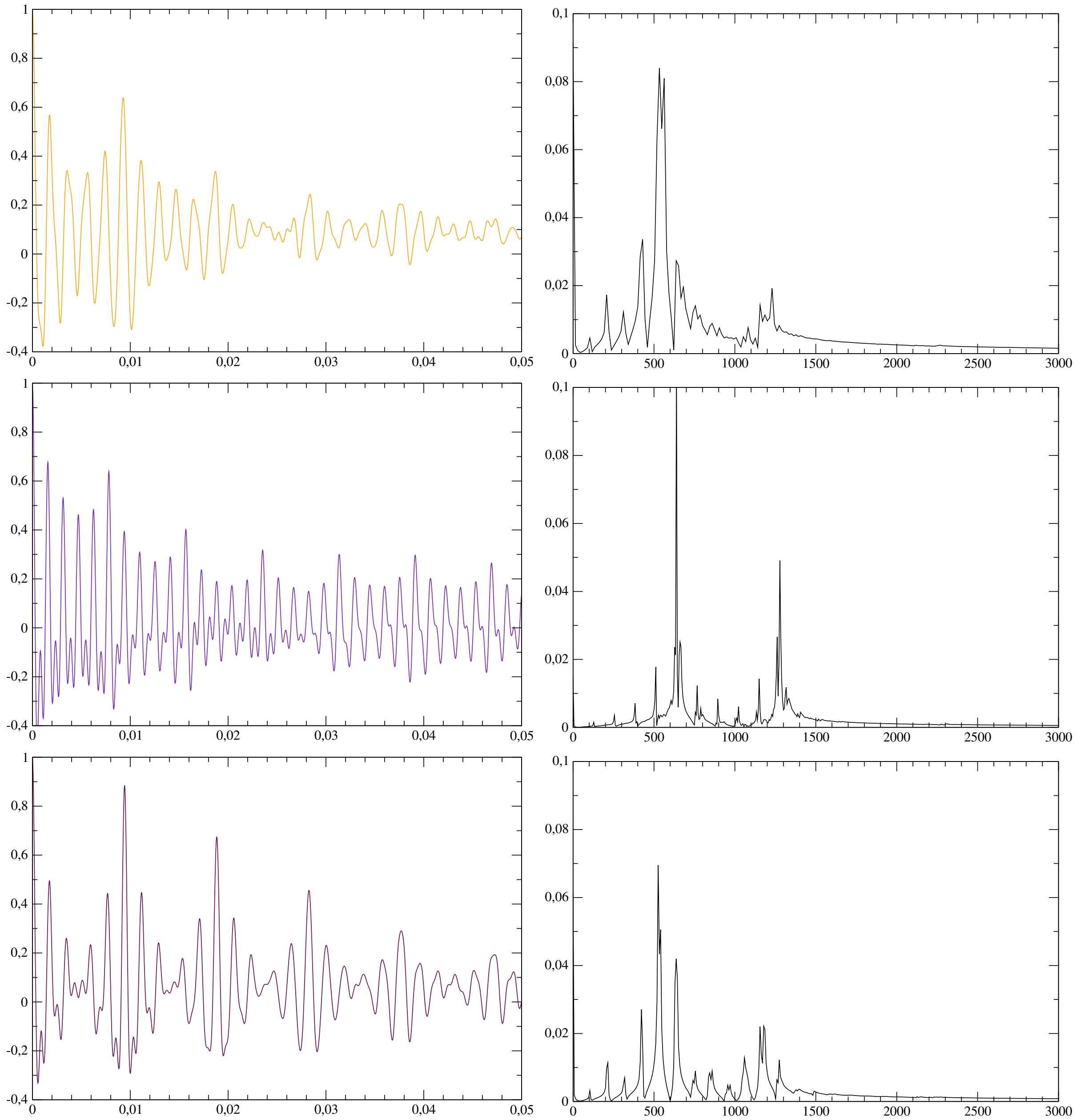

Abb. 3.3: AKF eines Stimmsignals. Links ist für jeden der drei Vokale ein Ausschnitt aus der AKF geplottet, rechts das korrespondierende Leistungsspektrum (*power spectrum*), also die zugehörige Fourier–Transformation (hier DFT). Erst in der Fourier–Transformation lassen sich die erste und zweite Formant–Frequenz deutlich erkennen und so der zugehörige Vokal bestimmen.