

**Seminar „Algorithmen in Akustik und Computermusik 02“**  
**Seminarleiter: Univ.-Prof. Dr. Robert Höldrich**  
**WS 2005/06**

**Seminararbeit**

***Empirical Mode Composition***

**von**  
**Daniel Hofer**

Institut für Elektronische Musik und Akustik  
Universität für Musik und Darstellende Kunst, Graz

## Inhaltsverzeichnis

• I. Einleitung	3
• II. Die Momentanfrequenz	3
• III. <i>Intrinsic Mode Function</i> (IMF)	8
• IV. Der Siebungsprozess	8
• V. Hilbert-Spektrum	13
• VI. Conclusio	18
• VII Quellennachweis	19

## I. Einleitung

Die „*empirical mode decomposition*“ (EMD) stellt eine Methode zur Analyse von nicht linearen und nicht stationären Daten dar. Sie wurde unter der Leitung von Dr. Norden E. Huang [1] entwickelt. Die Zerlegung solcher Daten wird mittels „*intrinsic mode functions*“ (IMF) im Zeitbereich adaptiv bewerkstelligt. Durch die Hilbert-Transformation der einzelnen IMFs und Bestimmung der Momentanfrequenz als Funktion der Zeit ergibt sich die Energie-Frequenz-Zeit Verteilung. Dies wird auch als Hilbert-Spektrum bezeichnet.

Das Neue an dieser Methode ist das Konzept der IMF, da sie die lokalen Eigenschaften der Daten analysieren.

## II. Die Momentanfrequenz

Die Begriff der Momentanfrequenz beinhaltet zwei grundlegende Schwierigkeiten. Die Erste rührt von der traditionellen Fourier-Spektralanalyse her, denn diese definiert die Frequenz als Sinus oder Kosinus mit konstanter Amplitude über die ganze Länge der Daten. Das heißt, die Momentanfrequenz muss entweder einem Sinus oder Kosinus entsprechen, und es braucht eine ganze Oszillation um die lokale Frequenz zu erhalten. Gemäß dieser Logik würde z.B. eine halbe Periodendauer nicht zur Bestimmung der Momentanfrequenz führen. Diese Definition macht also keinen Sinn für nicht stationäre Signale, in denen sich die Frequenz von Zeitpunkt zu Zeitpunkt ändern kann. Die zweite Schwierigkeit ergibt sich aus der nicht eindeutigen Definition der Momentanfrequenz. Seit der Einführung des analytischen Signals mittels Hilbert-Transformation besteht diese Schwierigkeit nicht mehr. Für ein reelles und kausales Zeitsignal  $x(t)$  gibt es immer eine Hilbert-Transformierte,  $y(t)$ .

Hilberttransformierte

$$y(t) = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{x(\tau)}{t-\tau} d\tau \quad (1)$$

Mit dieser Definition formen  $x(t)$  und  $y(t)$  ein konjugiert komplexes Paar, welche ein analytisches Signal  $z(t)$  bilden.  $x(t)$  und  $y(t)$  sind also über die Hilbert-Transformation miteinander verknüpft.

$$z(t) = x(t)+y(t)= a(t) \exp (j\theta(t)) \quad (2)$$

Aus dieser komplexen Darstellung lassen sich nun leicht Betrag und Phase bestimmen.

$$a(t)=[x(t)^2 + jy(t)^2]^{1/2} \quad \text{Betrag} \quad (3)$$

$$\theta = \tan^{-1} \left( \frac{y(t)}{x(t)} \right) \quad \text{Phase} \quad (4)$$

Was ist ein analytisches Signal?

Ein analytisches Signal hat per Definition die Eigenschaft, dass sein Fourier-Spektrum für negative Frequenzen den Wert Null hat. Das Fourier-Spektrum eines reellwertigen Signals  $x(t)$  hat immer negative Frequenzen; daraus lässt sich folgern, dass ein analytisches Signal immer komplex ist. Weiters ist bei einem analytischen Signal der Realteil  $x(t)$  über die Hilbert-Transformation mit dem Imaginärteil  $y(t)$  verknüpft.

Trotz Anwendung der Hilbert-Transformation gibt es noch immer Kontroversen bei der Definition der Momentanfrequenz, wie folgt:

$$\omega = \frac{d\theta(t)}{dt} \quad (5)$$

Cohen (1995) [2] führt nun den Terminus „Monokomponente Funktion“ ein, denn die Gleichung (5) gibt zu jedem Zeitpunkt nur einen einzigen Frequenzwert aus, also nur eine Komponente, deshalb „Monokomponentes Signal“. Es gibt wiederum keine klare Definition, ob eine Funktion monokomponent ist oder nicht. Ein Ansatz um dieses Problem zu lösen wurde von Schwartz (1966) [3] beschrieben, indem er die „narrow band“ als Begrenzung der Daten verwendet, um somit die Momentanfrequenz in einem Band bestimmbar zu machen. Doch diese Methode stellte sich nur unter bestimmten Einschränkungen als nützlich heraus.

Um eine sinnvolle Momentanfrequenz von irgendeiner Funktion zu erhalten, muss die Fourier-Transformation des Realteils nur positive Frequenzen liefern, [siehe Titchmarsh (1948)[4]]. Diese Bedingung ist aber nur global anwendbar, das heißt es ist nicht möglich, lokal, zu einem bestimmten Zeitpunkt, die Momentanfrequenz zu bestimmen.

Es folgt nun ein Beispiel zur Veranschaulichung dieser Einschränkung.

Gegeben sei die Funktion,

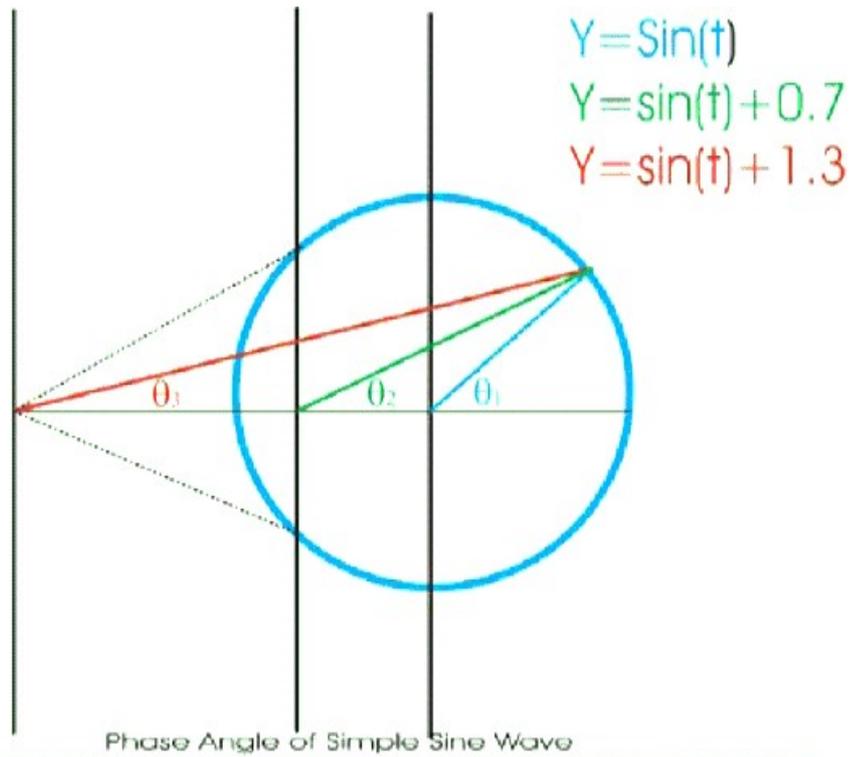
$$y(t) = \sin(t) \quad (6)$$

Die Hilberttransformierte dieser Funktion ist  $\cos(t)$ . Der Phasenplot in der x-y Ebene stellt den Einheitskreis dar, (Abb. 1a, S. 6). Die Phasenfunktion ist eine gerade Linie, Abb. 1b, und die Momentanfrequenz wird in Abb. 1c dargestellt, welche wie erwartet einen konstanten Wert annimmt. Zu der Funktion  $x(t)$  wird nun ein Gleichanteil dazu gegeben,

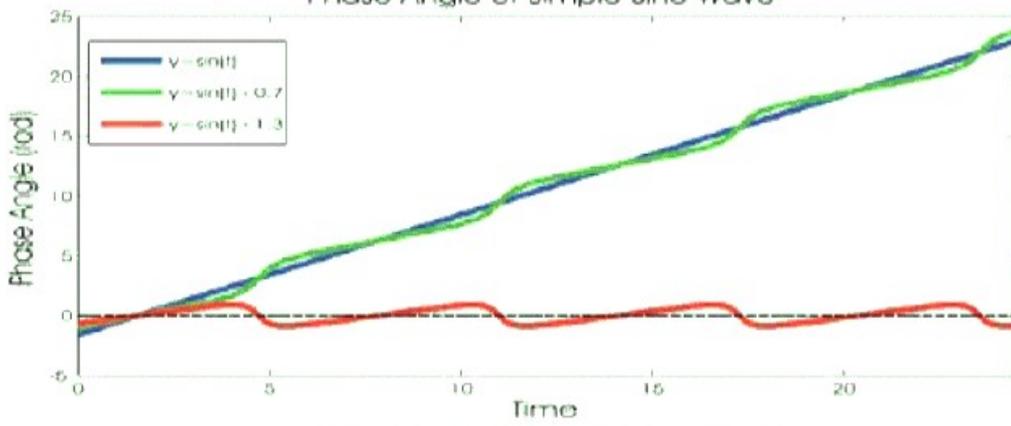
$$y(t) = \alpha + \sin(t) \quad (7)$$

Der Phasenplot in der x-y Ebene stellt noch immer den Einheitskreis dar, unabhängig welche Werte  $\alpha$  annimmt. Nur der Mittelpunkt des Kreises wird um den Wert  $\alpha$  verschoben (Abb. 1a, S. 6 ). Wenn  $\alpha < 1$  befindet sich der Mittelpunkt innerhalb des Kreises. Das Fourier-Spektrum weist einen Gleichanteil auf. Dennoch, die mittelwertsfreie Schwingung ist noch die Gleiche wie wenn  $\alpha=0$ , doch die Phasenfunktion und die Momentanfrequenz ändern sich (Abb. 1b,c., S. 6 ). Sei  $\alpha > 1$ , dann befindet sich der Mittelpunkt außerhalb des Einheitskreises und die Funktion erfüllt nicht mehr die Bedingungen. Die Momentanfrequenz und die Phasenfunktion nehmen dann negative Werte an, die bedeutungslos sind (Abb. 1b-c, S. 6).

a)



b)



c)

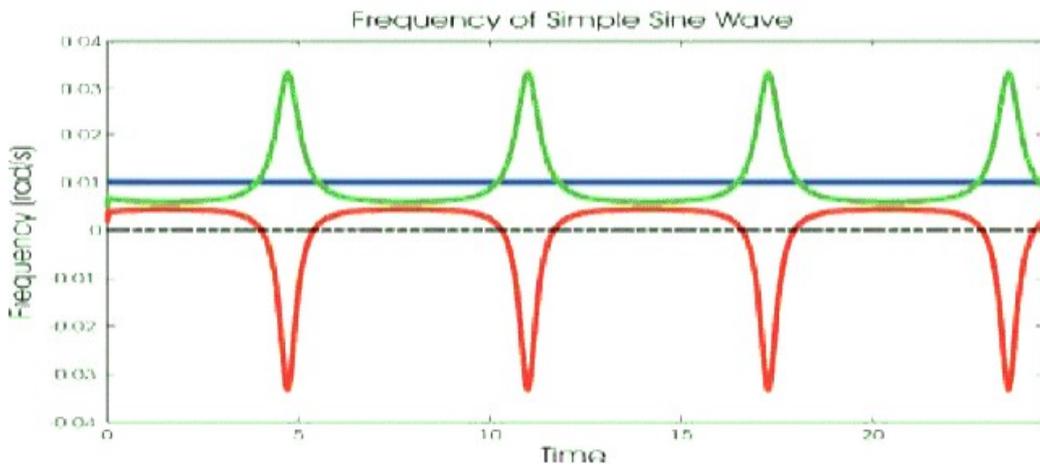


Abb. 1 [8]. Physikalische Interpretation der Momentanfrequenz; (a) Phasendarstellung der Funktion  $y(t) = \alpha + \sin(t)$  mit  $\alpha = 0,0.7,1.3$ , (b) Phasenwinkel nach der Zeit aufgetragen, (c) Momentanfrequenz bestimmt nach Gleichung (5).

Dieses Beispiel zeigt sehr schön, dass die Momentanfrequenz nur bestimmbar ist, wenn die Bedingungen - mittelwertsfrei und lokal symmetrisch zum Nullpunkt - gegeben sind.

### III. Intrinsic Mode Functions

Der Name „*intrinsic mode function*“ wurde gewählt, da dieser die Schwingung der Mode innerhalb der Daten beschreibt. Dies bedeutet, dass eine IMF, die sich auch über die Nulldurchgänge definiert, nur eine Modeschwingung darstellt, die aber amplituden- und frequenzmoduliert sein kann.

Das Ziel bei der Entwicklung der IMF war es, eine Funktion zu finden, die es ermöglicht die Momentanfrequenz lokal, an jeder beliebigen Stelle, bestimmbar zu machen. Dies heißt, die Funktion muss symmetrisch sein bezogen auf den lokalen Mittelwert (=Null) und muss die gleiche Anzahl von Nulldurchgängen und Extremen besitzen. Dies lässt sich formal wie folgt definieren.

(1) Die Anzahl der Extreme und die Anzahl der Nulldurchgänge müssen gleich sein oder dürfen sich maximal nur um eins unterscheiden und (2) zu jedem Zeitpunkt muss der Mittelwert der Einhüllenden der Maxima und der Einhüllenden der Minima Null sein.

### IV. Die *empirical mode decomposition* Methode: Der Siebungsprozess

Diese Methode basiert auf der Annahme dass: (1) ein Signal hat zumindest zwei Extreme - ein Maximum und ein Minimum; (2) die charakteristische Zeitskala wird über die Zeitspanne zwischen zwei Extremen definiert; und (3) wenn die vorliegenden Daten keine Extreme aufweisen, sondern Wendepunkte kann ein- bis mehrere Male differenziert werden, um Extreme zu erhalten.

Die Essenz dieser Methode ist es, die innewohnenden Modeschwingungen über die charakteristische Zeitskala der Daten empirisch zu bestimmen und gemäß dieser, die Zerlegung der Daten zu vollziehen. Nach Drazin (1992) [5] ist der erste Schritt bei Datenanalyse die Kontrolle der Daten mit dem Auge. Somit lassen sich sofort die unterschiedlichen Zeitskalierungen auf zwei Arten bestimmen: über die Zeitspanne zwischen zwei aufeinander folgenden Minima und Maxima und über die Zeitspanne zwischen zwei Nulldurchgängen. Die Komplexität dieser Daten entsteht nun durch die

Verkettung von lokalen Extremen und der Nulldurchgänge: eine Schwingung wird durch eine andere überlagert, und diese überlagern eine weitere usw. Jede dieser Schwingungen definiert eine eigene charakteristische Skalierung der Daten. Bei dieser Methode wurde die Zeitspanne zwischen zwei Extremen als Definition der Zeitskalierung der intrinsischen Modenschwingung gewählt, da sich einerseits eine bessere Auflösung der Modenschwingungen ergibt und die Methode auch auf Daten angewandt werden kann, die keinen Mittelwert von Null besitzen, d.h. entweder alle positiven oder alle negativen Werte ohne Nulldurchgänge.

Der Siebungsprozess:

Auf Grund der Definition der IMF werden bei dieser Methode die Einhüllende der Maxima und die Einhüllende der Minima separat verwendet. Als erstes werden alle Extreme detektiert. Weiters werden alle Maxima mittels kubischer *Spline Interpolation* verbunden und man erhält die Einhüllende der Maxima. Das Gleiche wird für die Minima wiederholt um die Einhüllende der Minima zu erhalten. Innerhalb dieser beiden Einhüllenden sollten sich alle Daten befinden. Der Mittelwert dieser beiden Einhüllenden wird hier als  $m1$  bezeichnet. Subtrahiert man nun diesen Mittelwert vom Originalsignal  $x(t)$  erhält man das so genannte erste 'Detail'  $h1$ .

$$h1(t) = x(t) - m1(t) \quad (8)$$

Idealer Weise sollte nun  $h1$  schon eine IMF sein, denn alle oben genannten Vorgaben wurden erfüllt. In der Realität treten so genannte Über- und Unterschwingungen auf, die neue Extreme erzeugen können und diese auch verschieben oder aufschaukeln können, (siehe Abb. 2, S. 10 bei Sekunde 4.6 und 4.7). Die Auswirkungen dieser Über- und Unterschwingen haben keinen direkten Einfluss auf den Siebungsprozess, da ja der Mittelwert verwendet wird. Dies heißt aber, dass eine kleine „Ausbuchtung“ in der Steigung der Amplitude eine Verstärkung erfahren kann und somit zu einem Extrem wird, indem es die lokale Nullstelle von einem rechteckigen in ein gekrümmtes Koordinatensystem ändert. Ein Beispiel dazu findet sich in Abb. 2 zwischen 4.5 und 4.6 Sekunden. Nach dem ersten Siebungsprozess wird die Ausbuchtung zu einem lokalen Maximum (siehe 2c zum gleichen Zeitpunkt). Neue Extreme, die so generiert werden, stellen eigentlich die richtigen Moden, die anfangs betrachtet wurden, wieder her. Daraus folgt, dass der Siebungsprozess kleine Amplitudenschwankungen durch wiederholtes Sieben wieder

herstellen kann.

Eine weitere Komplikation stellt sich ein, wenn der Mittelwert der beiden Einhüllenden sich vom lokalen Mittelwert bei nicht linearen Daten unterscheidet. Dann hat dies zur Folge, dass asymmetrische Wellenformen existieren, und zwar auch nach  $n$  Siebungen. Doch diese Approximation ist zu akzeptieren und auch zulässig.

Neben diesen theoretischen Schwierigkeiten gibt es auch noch praktische, die bei der Interpolation mittels kubischer Spline auftreten. Am Ende einer IMF kann es zu großen Überschwingungen kommen, weil der Interpolationsalgorithmus zukünftige Samples braucht, um den gegenwärtigen Verlauf der Einhüllenden zu bestimmen. Dies ist am Ende nicht möglich. Somit können sich diese Überschwinger unter gewissen Umständen auch innerhalb der Daten vom Ende aus ausbreiten und die Daten speziell im niederfrequenten Bereich korrumpieren. Norden E. Huang [1] hat sich diesbezüglich einen Lösungsansatz ausgedacht, indem er charakteristische Schwingungen am Ende dazu addiert, die über die aufeinander folgenden Extreme, Amplitude und Frequenz, bestimmt werden.

Der Siebungsprozess hat zwei grundlegende Eigenschaften, (1) so genannte *riding waves* zu eliminieren und (2) Amplitudenschwankungen zu glätten. Das Stopkriterium für die Berechnung einer IMF wird über die Standardabweichung definiert.

$$VAR = \sum_{t=0}^T \left[ \frac{(h_{1(k-1)}(t) - h_{1k}(t))^2}{(h_{1(k-1)}^2(t))} \right] \quad (9)$$

Typische Werte für die Standardabweichung liegen zwischen 0.2 und 0.3.

$$c_1 = h_{1k} \quad (10)$$

$$x(t) - c_1 = r_1 \quad (11)$$

Der Siebungsprozess ist beendet, wenn entweder  $c_n$  oder das Residuum  $r_n$  einen vorbestimmten Wert erreichen, oder wenn das Residuum  $r_n$  eine monotone Funktion darstellt, und somit keine IMF mehr extrahiert werden kann. Auch für Daten mit Mittelwert Null kann das letzte Residuum einen Trend darstellen, der sich von Null abhebt. Formell gilt nun der Zusammenhang zwischen den einzelnen IMFs, Residuum und Zeitsignal wie folgt:

$$\sum_{i=1}^n IMF(i) + r_n \quad (12)$$

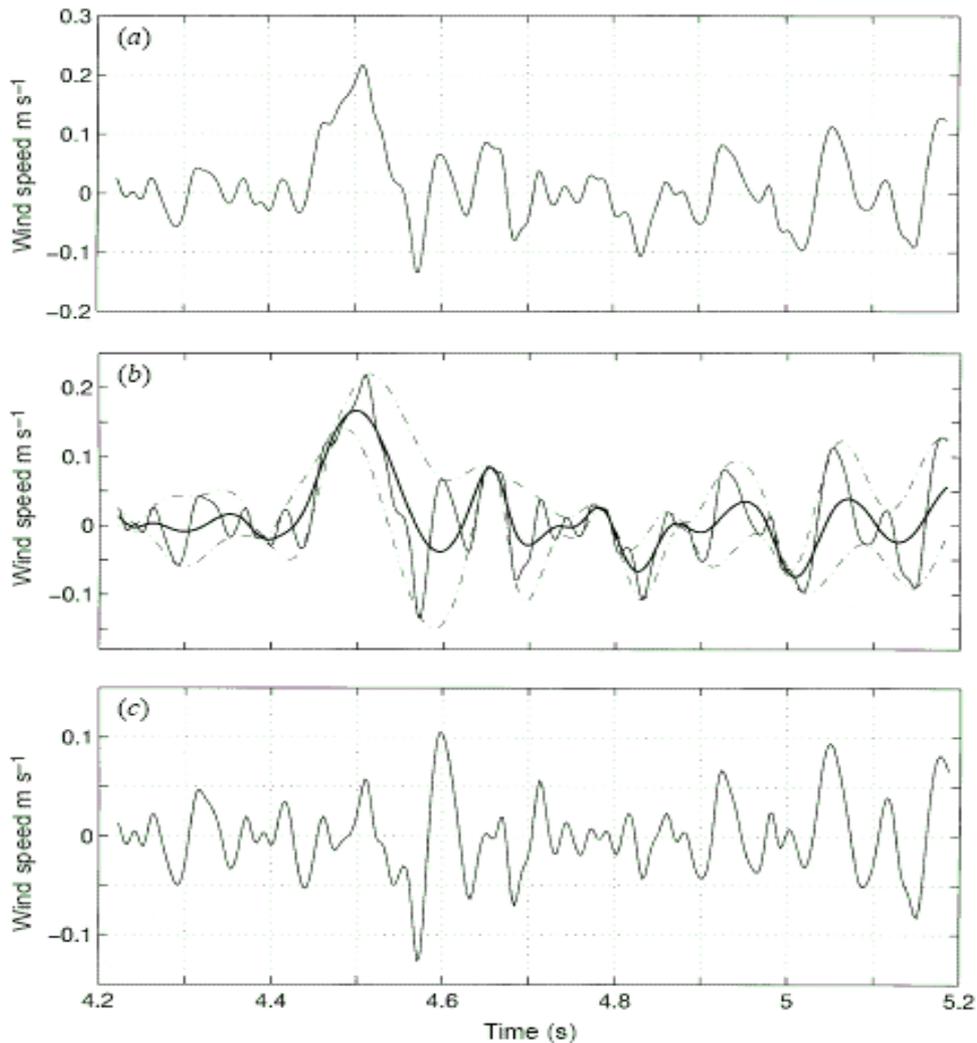


Abb. 2: Siebungsprozess (a) die originalen Daten, (b) die Daten in dünner Linie, obere und untere Einhüllende in strichlierter Linie und den Mittelwert mit dicker Linie dargestellt, (c) die Differenz zwischen Daten und  $m_1$ . Dies stellt noch keine IMF dar, da noch lokale negative Maxima und lokale positive Minima vorhanden sind.

Zur Illustration des Siebungsprozesses soll hier das Beispiel aus dem *DAFX-05 Proceedings* [6] aufgezeigt werden. Das Eingangssignal (Sig) stammt von einer Santur, ein iranisches Saiteninstrument mit trapezförmigem Resonanzkörper, welches mit Hämmerchen - ähnlich dem Hackbrett - gespielt wird. Aufgenommen wurde der Ton a' (440 Hz).

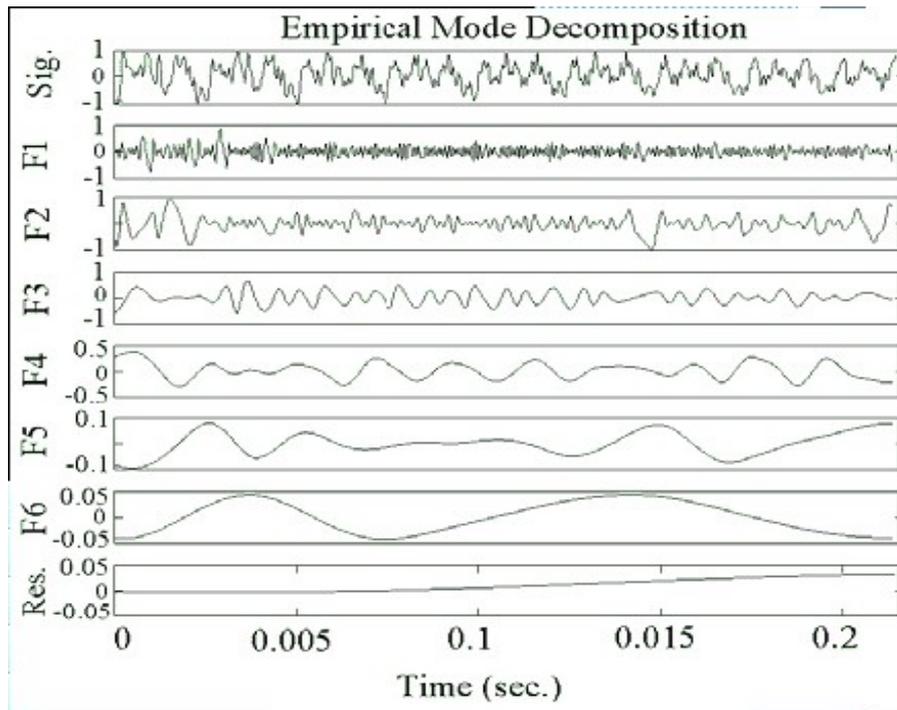


Abb. 3 Eingangssignal (Sig) zerlegt in 6 IMFs (F1-F6) und das Residuum (Res)

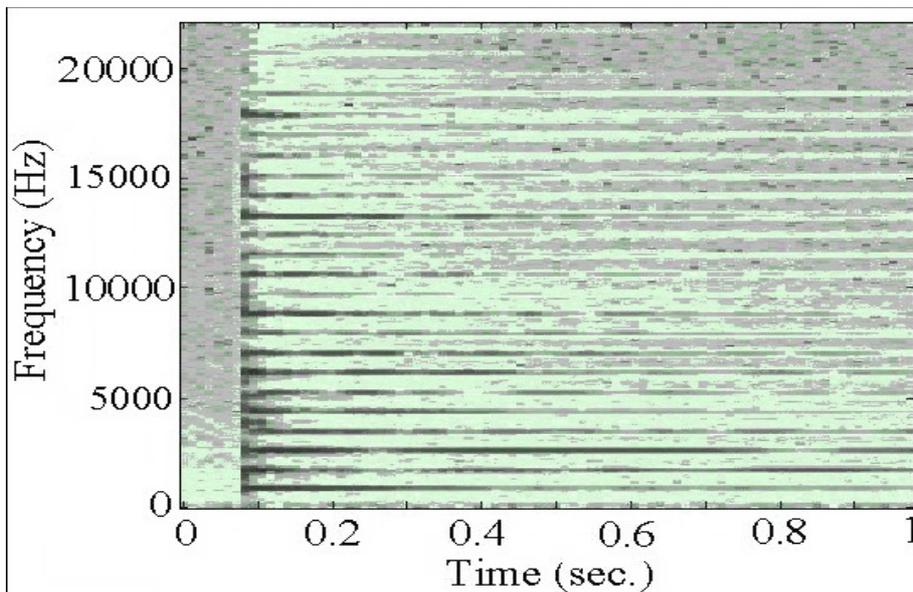


Abb. 4 zeigt das Spektrogramm (STFT) des Tons 'a'

Die Analyse der IMFs zeigt, dass die erste IMF (F1) den 6., 16. und 20. Oberton enthält, die zweite IMF (F2) den 4. Oberton, die dritte IMF (F3) den 2. Oberton, die vierte IMF (F4) die Grundschiwingung, die fünfte IMF (F5) die halbe Grundschiwingung und die sechste IMF (F6) ein Viertel der Grundschiwingung enthält. Das Residuum (Res) zeigt einen ansteigenden Trend des Eingangssignals.

## V. Hilbert-Spektrum

Zu jeder IMF kann nun die Hilbert-Transformierte und daraus die Momentanfrequenz gemäß Gleichung (5) berechnet werden. Nach Anwendung der Hilbert-Transformation für jede IMF kann das Eingangssignal wie folgt beschrieben werden:

$$x(t) = \sum_{j=1}^n a_j(t) e^{i \int w_j(t) dt} \quad (13)$$

Das Residuum wird bewusst weggelassen, da es entweder eine monotone Funktion oder eine Konstante ist. Obwohl die Hilbert-Transformation im Stande ist monotone Funktionen als einen Teil einer langen Oszillation darstellen zu können. Gleichung (13) ermöglicht es die Amplitude und die Frequenz als Funktion der Zeit dreidimensional darzustellen. Die Frequenz-Zeit Verteilung der Amplitude wird als Hilbert-Amplitudenspektrum,  $H(\omega, t)$ , oder auch nur Hilbert-Spektrum bezeichnet.

Beispiele:

- 1) Zur Illustration der von Huang benannten „*intra-wave frequency modulation*“ wird das Beispiel angeführt, das auch im Seminar besprochen wurde. Es wird ein amplitudenmoduliertes Signal analysiert

$$x = \exp(-0.01 \cdot t) \cos(2/32 \cdot \pi \cdot t) \quad (14)$$

Die Wellenform ist in Abb. 5, S. 13, dargestellt. Die EMD muss für diese Schwingung nicht mehr angewandt werden, da es sich schon um eine IMF handelt.

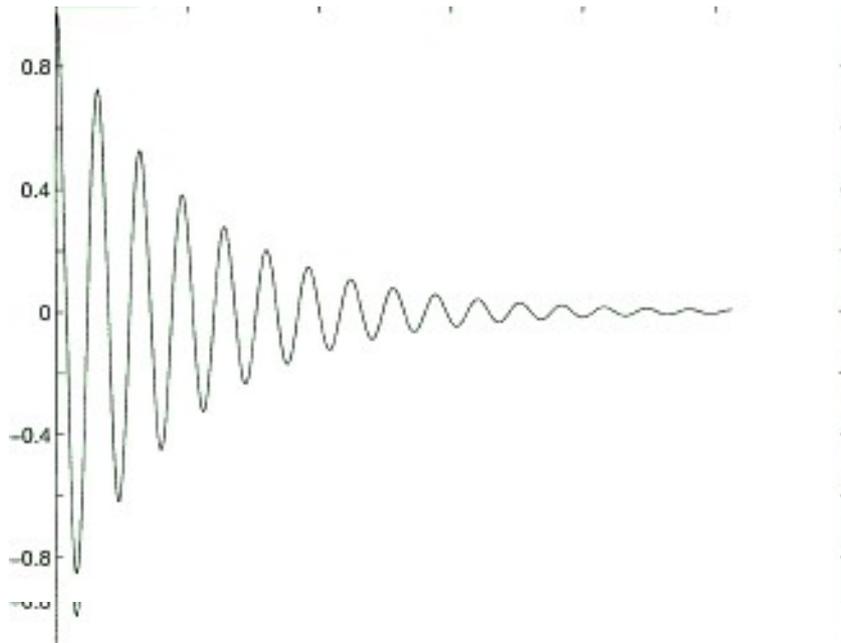


Abb. 5 Schwingungsform von Gleichung (14)

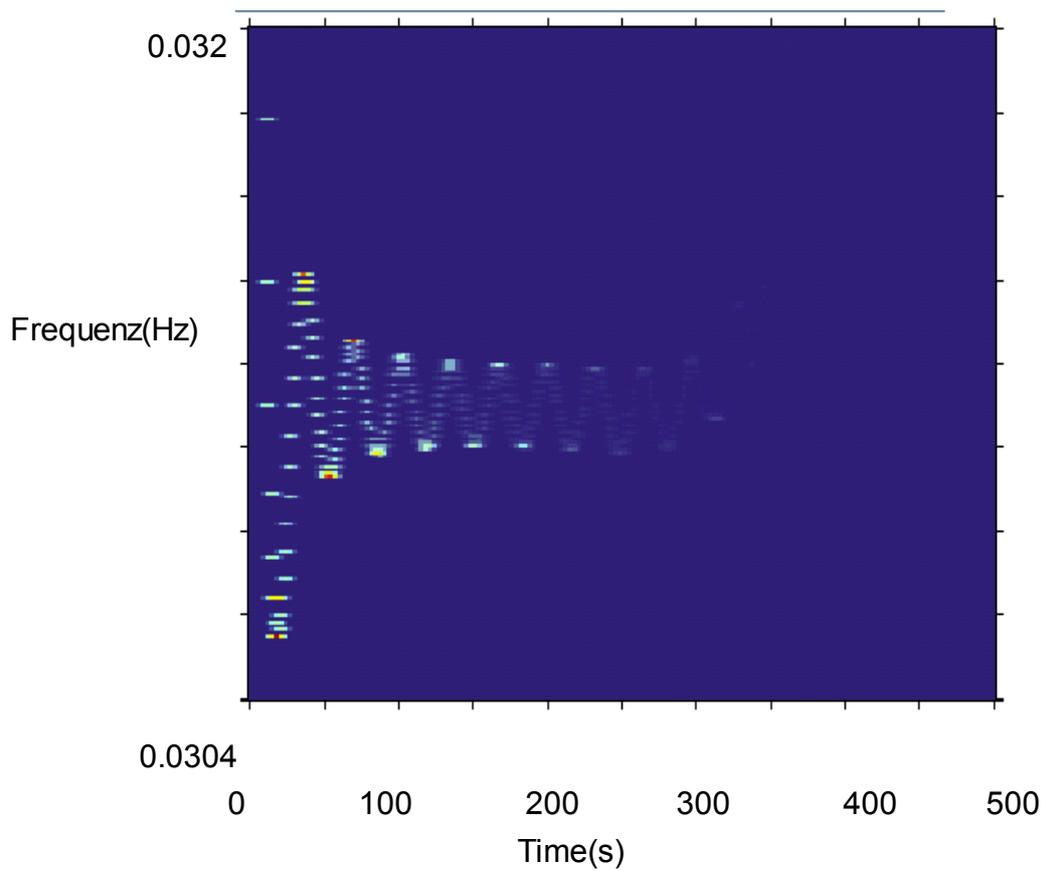


Abb 6 Hilbert-Spektrum des in Abb. 5 dargestellten Signals. Die *intra-wave modulation* wird hier mit 1.5% angegeben.

Zu erwarten wäre, dass das Hilbert-Spektrum nur eine Frequenz über die Zeit aufzeigt. Da aber der Modulator der Amplitudenmodulation eine abklingende Funktion ist und somit ein unendlich großes Spektrum aufweist, sind die spektralen Anteile des negativen Spiegelspektrums auch im positiven Spektralbereich anzutreffen. Wird die Trägerfrequenz analysiert, werden auch genau diese Anteile mit analysiert. Robert Höldrich meinte, dass der Begriff der *intrawave modulation* somit nicht stimmt, (siehe Seminar vom 26.01.2006)

2) Als zweites Beispiel wird ein Kosinus verwendet, der plötzlich seine Frequenz ändert.

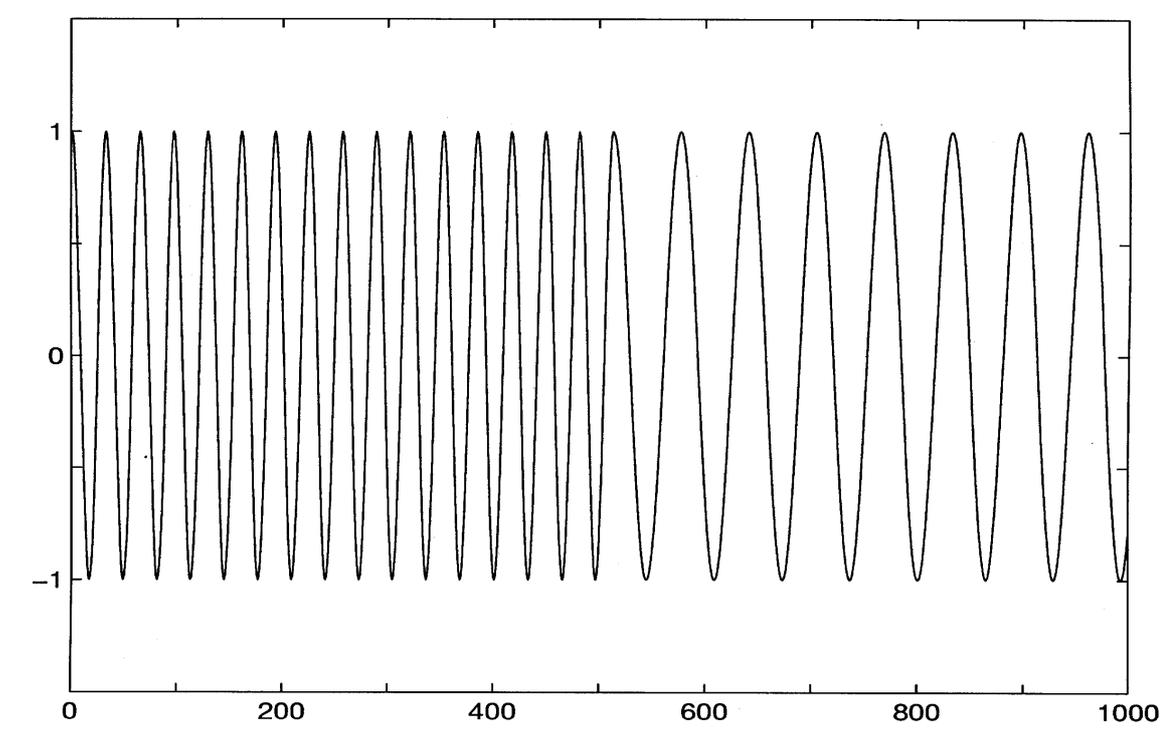


Abb. 7 Schwingungsform eines Kosinus im Zeitbereich, der seine Frequenz schlagartig ändert.

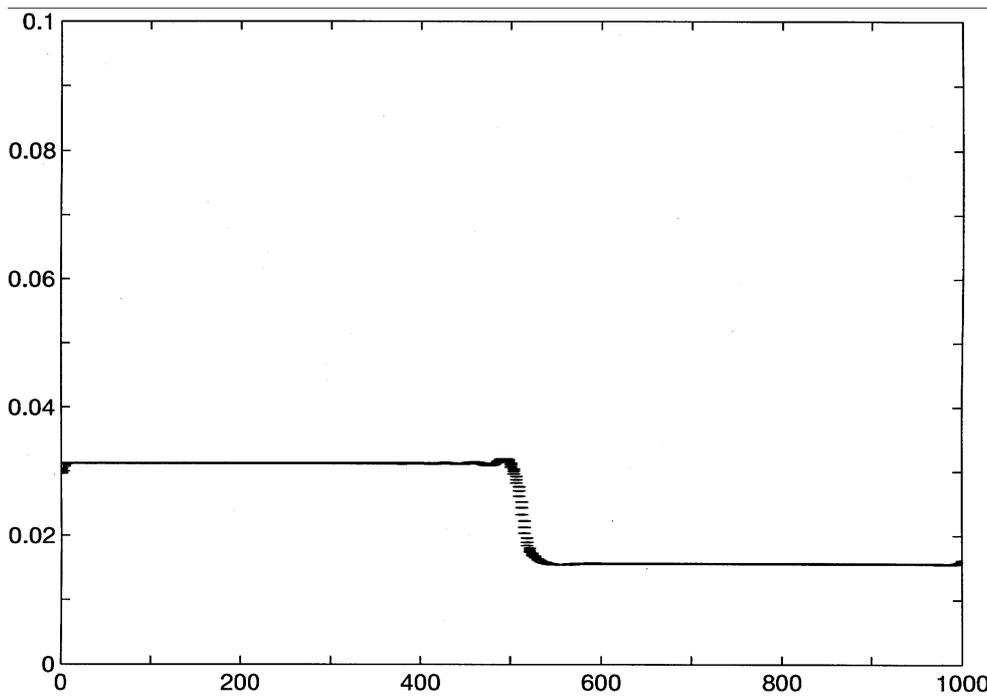


Abb. 8 Hilbert-Spektrum, mit genauer Frequenz- und Zeitlokalisation, die leichten Frequenzoszillationen werden auf das Gibbs-Phänomen zurückgeführt.

Auffallend ist die Schärfe, mit der die Hilbert-Transformierte den Frequenzsprung auflöst.

3) Als letztes Beispiel wird die lineare Summierung von zwei Kosinus angeführt.

$$x(t) = \cos(2/30 \cdot \pi \cdot t) + \cos(2/34 \cdot \pi \cdot t) \quad (15)$$

Der Siebungsprozess würde bei solchen Daten keine IMFs ausgeben, da die vorliegenden Daten schon eine IMF für die Approximation erster Ordnung repräsentieren. Um die EMD durchzuführen muss das Abbruchkriterium einen äußerst kleinen Wert annehmen. Nach 3000 Siebungen gibt die EMD acht IMF Komponenten aus. Die Ersten vier IMFs sind in Abb. 10, S. 16 dargestellt. Von diesen IMFs wird das Hilbert-Spektrum berechnet (Abb. 11, S. 17 ). Diese Abbildung zeigt eindrücklich, dass weder die Morlet-, die Fourier- noch die Hilbert-Transformation imstande sind die beiden dicht beieinander liegenden Frequenzen abzubilden. Dieses Beispiel soll die Grenze der Auflösung dieses Verfahrens zeigen.

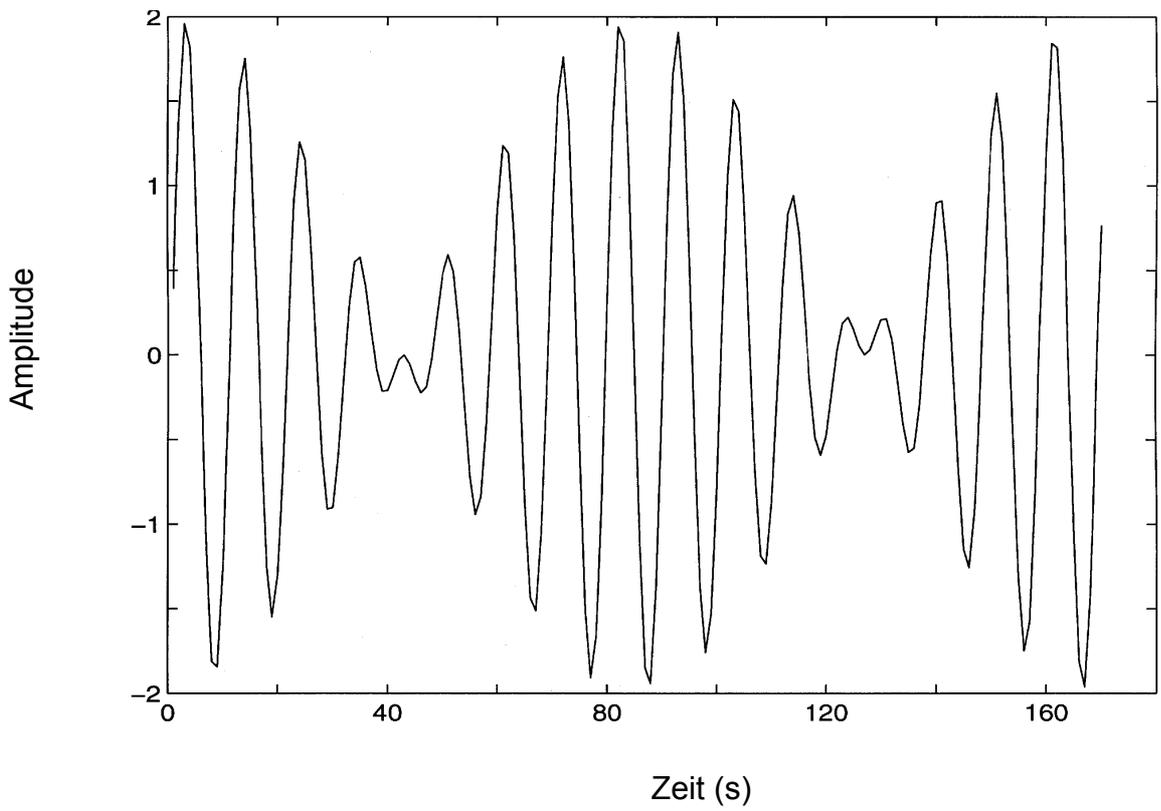


Abb. 9 Zeitlicher Verlauf von  $x(t)$  Gleichung (15)

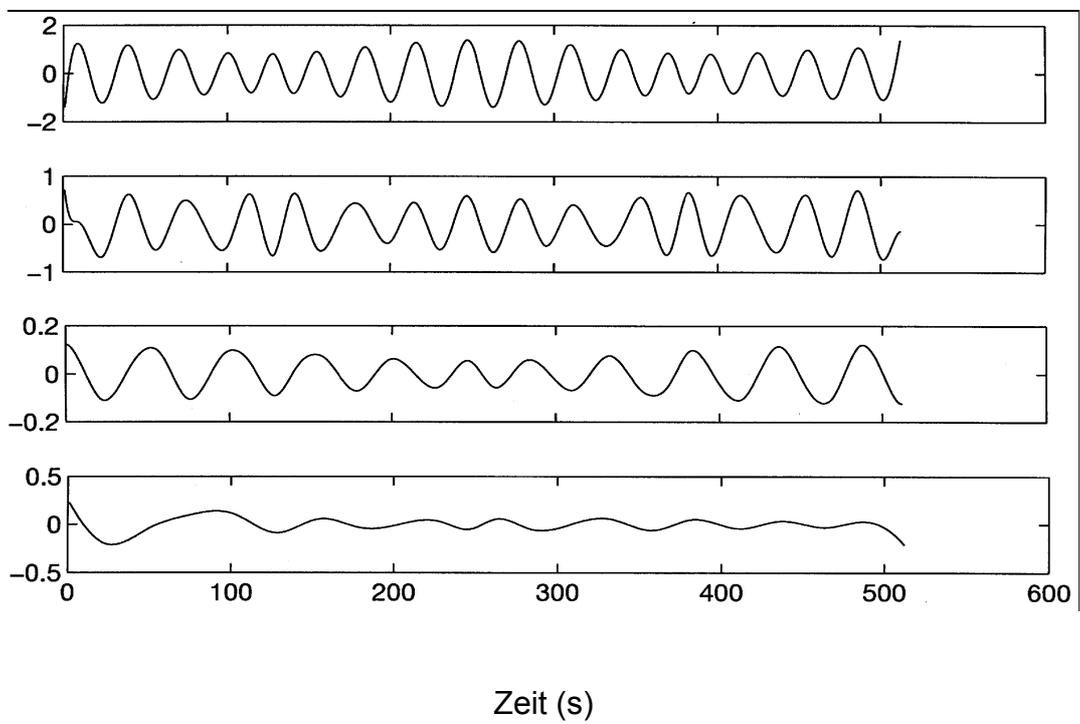


Abb. 10 Die ersten vier IMFs

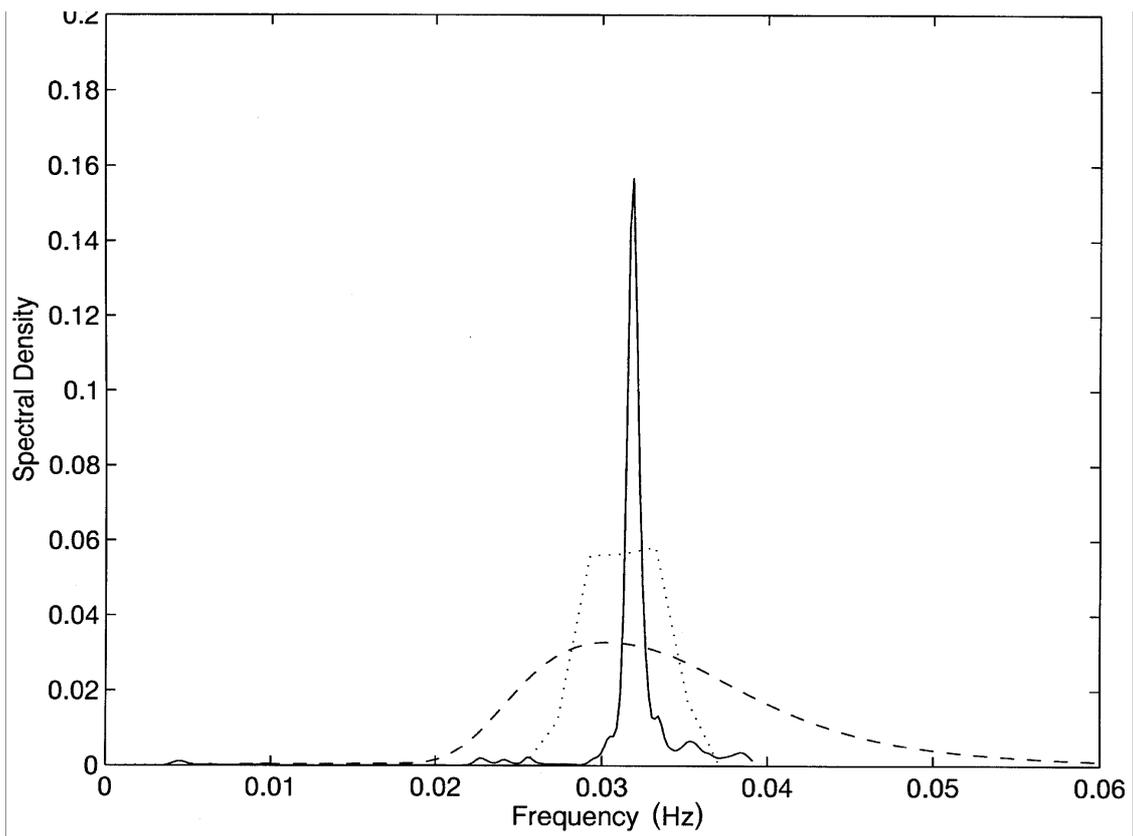


Abb 11 Vergleich von Fourier (gepunktete Linie), Hilbert (durchgezogene Linie) und Wavelet Spektrum (gestrichelte Linie). Keines der drei Verfahren trennt die zwei Eingangsfrequenzen.

## VI. Conclusio

Die EMD Methode ermöglicht es, nicht stationäre und nicht lineare Daten zu analysieren. Es ist die erste Methode, die lokal und adaptiv in der Frequenz-Zeit-Domäne analysiert. Der adaptive Siebungsprozess zerlegt das Signal in IMFs, die abhängig vom Eingangssignal, sich wie Oktavfilterbanken verhalten[7]. Ein großer Vorteil ist, dass die EMD in Kombination mit der Hilbert-Transformation keine ganze Schwingungsperiode braucht, um die lokale Frequenz zu definieren. Die Hilbert-Transformation gibt immer den am besten passenden Sinus oder Kosinus für die lokalen Daten aus. Die Momentanfrequenz wird dann entweder mit der stationären Phasenmethode oder mit der Ableitung der Phase bestimmt. Ein weiterer Vorteil von transienten Vorgängen, die keinen Nulldurchgang oder keinen Mittelwert aufweisen, ist, dass der Gleichanteil automatisch eliminiert wird. Das Hilbert-Spektrum ermöglicht es, auch die Stationarität von Daten zu messen.

Ein gewissen Nachteil stellt die lange Rechenzeit dieser Methode dar. Das Interpolationsverfahren zwischen den einzelnen Extremwerten mit den sogenannten „endeffects“ ist noch verbesserungswürdig.

Diese neue Methode gibt weiters Einblicke in andere nicht lineare und nicht stationäre Prozesse.

Abschließend ein Zitat von Norden E. Huang: *„We have only just begun to explore the full physical interpretations of the Hilbert spectra for complicated data“.*

## VII. Quellennachweis

- [1] Huang, N.E., Shen, Z. and Long, S.R., et al., „The empirical mode decomposition and the Hilbert spectrum for non-linear and non-stationary time series analysis“, Proc. R. Soc. London A, pp 903-995, 1998.
- [2] Cohen, L., „Time -frequency analysis“, Engelwood Cliffs, NJ: Prentice-Hall, 1995.
- [3] Schwartz, M., Bennett, W.R. & Stein, S., „Communications systems and techniques,“ New York: McGraw-Hill, 1996.
- [4] Titchmarsh, E. C., „Introduction to the theory of Fourier integrals.“ Oxford: University Press, 1948.
- [5] Drazin, P. G, „Nonlinear systems“. Cambridge University Press. 1992.
- [6] Heydarian, P., Reiss, J.,D., „Extraction of long-term structures in musical signals using the empirical mode decomposition“, Proc. of the 8<sup>th</sup> Int. Conference on digital audio effects (DAFX-05), Madrid, Spain, September 20-22, pp 258-261, 2005.
- [7] Wu, Z., Hunag, N. E. , „A study of the characteristics of white noise using the empirical mode decomposition method“, London: The Royal Society, 2004.
- [8] Tung-Shin, H., „An Introduction to Hilbert/Huang Transformation Analysis and its applications“, Ess 265, Spring, 2005.