

Modellbildung und Simulation

Sommersemester 2011

Klaus Kasper

Termine

Veranstaltung	Vorlesung Mo 18-21 (X) D14/104	Praktikum Mo 18-21 (Y) D15/107	Praktikum Fr 12-15 (X) D15/107
1. Termin	28.03.11	04.04.11	08.04.11
2. Termin	11.04.11	18.04.11	29.04.11
3. Termin	02.05.11	09.05.11	13.05.11
4. Termin	16.05.11	23.05.11	27.05.11
5. Termin	30.05.11	06.06.11	17.06.11
6. Termin	20.06.11	27.06.11	01.07.11

Organisation

- Moodle-Server: <http://lernen.h-da.de/>
- Kurs: Informatik/Informatik (M.Sc.)/Sommersemester 2011/Modellbildung und Simulation – Kasper - SS 2011
- Schlüssel: MuS_KAS_SS2011
- Vorlesungen, Aufgabenblätter und weitere Materialien werden auf dem Moodle-Server zur Verfügung gestellt.
- Fragen und Anmerkungen zur Veranstaltung bitte immer über das Diskussionsforum der Veranstaltung, so dass immer alle informiert sind.
- Wenn Sie interessante Informationen finden, stellen Sie diese bitte auch allen zur Verfügung.

Hinweise zum Praktikum

- freie Wahl der Programmiersprache
- Lösungen müssen dokumentiert werden
- Hinweise für die Durchführung der Kompilation
- Hinweise für die Ausführung der Lösungen

- 3 Aufgabenblätter
- Für eines der Aufgabenblätter bitte eine Lösung mit „ordentlichem“ GUI

Inhalt

- Einführung
- Beispielhafte Anwendungen
- Digitalisierung
- Parametrisierung
 - Fourier-Transformation

Einführung

Modellbildung

- Modelle sind Hilfsmittel für den Umgang mit Realität.
- Die Realität wird nicht direkt abgebildet, sondern abstrahiert.
 - Nicht zentrale Details werden ignoriert.
 - Komplexität des Modells wird minimiert.
- Die Leistungsfähigkeit aktueller Rechenmaschinen erweitert den Möglichkeitsraum der Modellbildung.

Prozess

- Die Entwicklung des Zustands eines Systems wird als Prozess bezeichnet.
- Exakt handelt es sich um dynamische Prozesse, da zeitliche Entwicklungen untersucht werden.
- Ziele der Modellierung von Prozessen:
 - Prognose zukünftiger Zustände des Systems.
 - Klassifikation von Muster, die vom untersuchten System erzeugt wurden.
 - Optimierung von Geschäftsabläufen.

Ziele der Modellierung

- Prognose
 - Wetter / Klima
 - Kursverläufe (Risikobewertung)
 - gesellschaftliche Dynamiken
- Klassifikation
 - Qualitätskontrolle
 - Bild- / Spracherkennung
- Optimierung bzw. Standardisierung von Prozessen
 - ITIL/eTOM
 - Mobilitätsmanagement

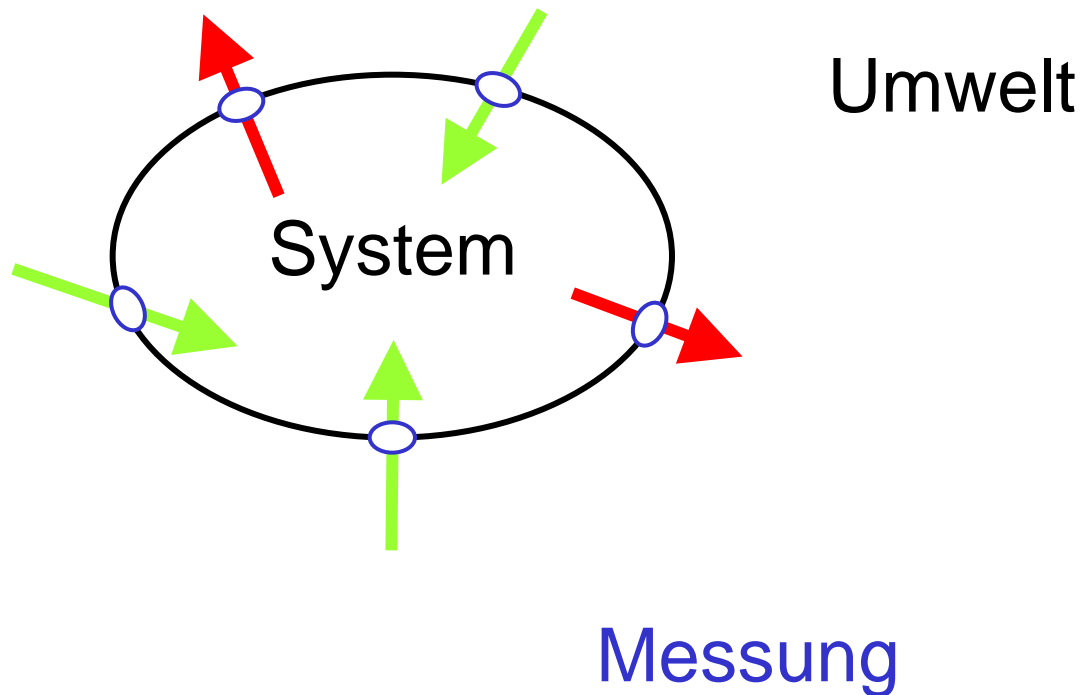
Modelltypen

- Alltagsmodelle (z.B. Bauernregeln)
- Analytische Modelle (z.B. Ökonomie (Marx), Physik (Newton))
- Datengetriebene Modelle (z.B. Klimamodelle)
- Akteurgetriebene Modelle (z.B. Modellierung gesellschaftlicher Dynamiken)
- Technische Modelle (z.B. Schaltungssimulation)

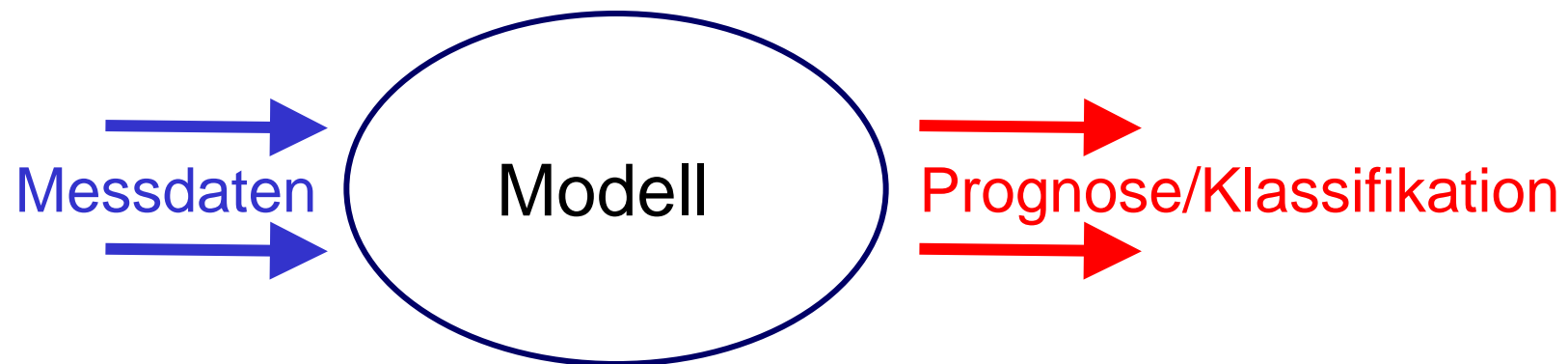
Datengetriebene Modellbildung und Simulation

- Definition System / Umwelt
- Austausch zwischen System und Umwelt
- Messung und Sammlung der Eingangsgrößen
- Größen für Beschreibung des Systemzustands (Ausgangsgrößen)
- Berechnung von unbekanntem Systemzuständen durch Simulation

Modellbildung



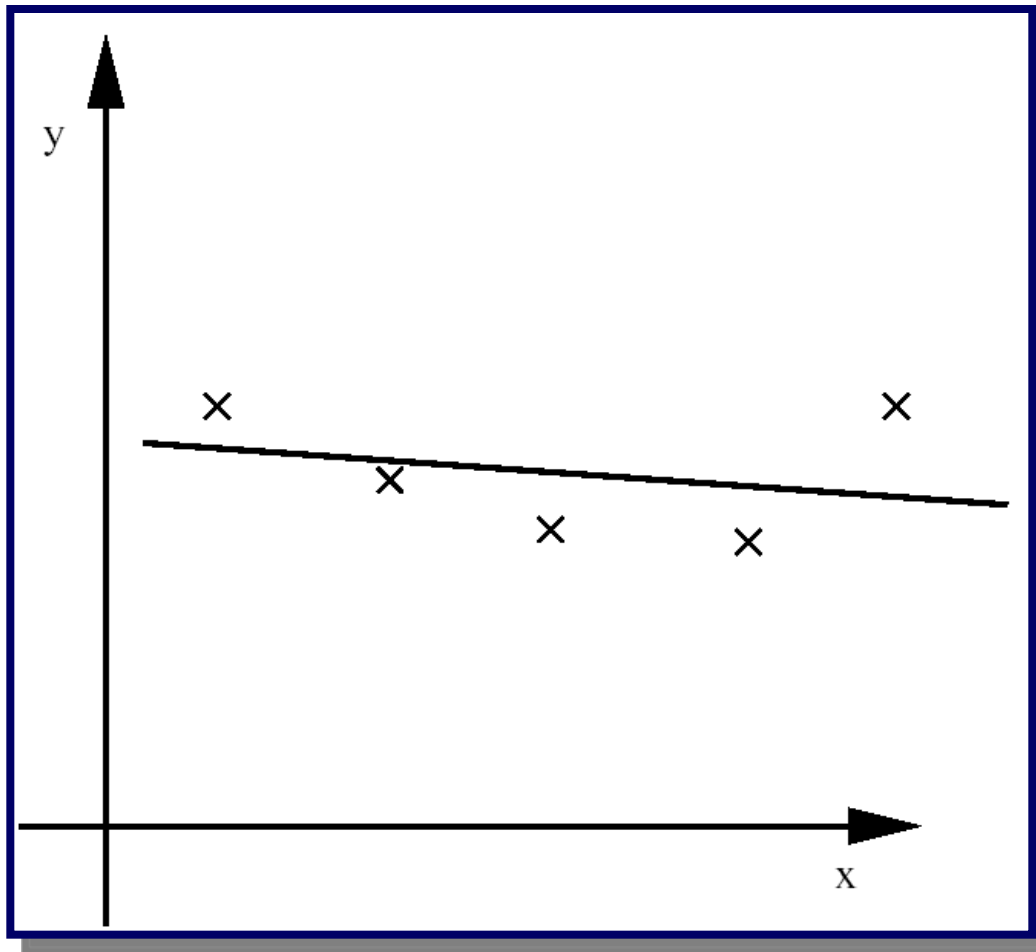
Simulation



Nichtlineare Systeme

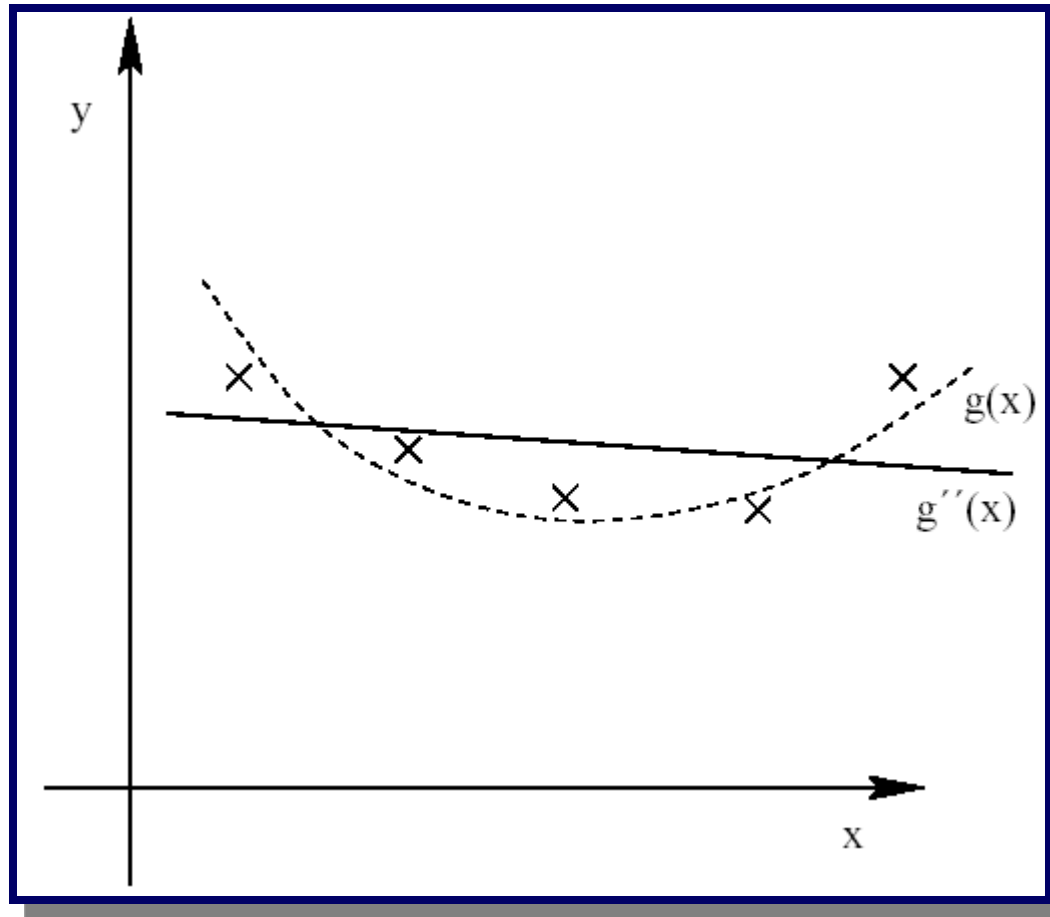
- **Nichtlineare Systeme** reagieren auf Störungen, im Gegensatz zu linearen Systemen, nicht proportional.
- Die meisten in der Natur vorkommenden Systeme sind nichtlinear.
- Beispiele sind die Dynamik der Weltwirtschaft, das ökologische System der Erde oder das Wetter.

Lineare Regression



$$f(x) = ax + b$$

Linear versus nichtlinear

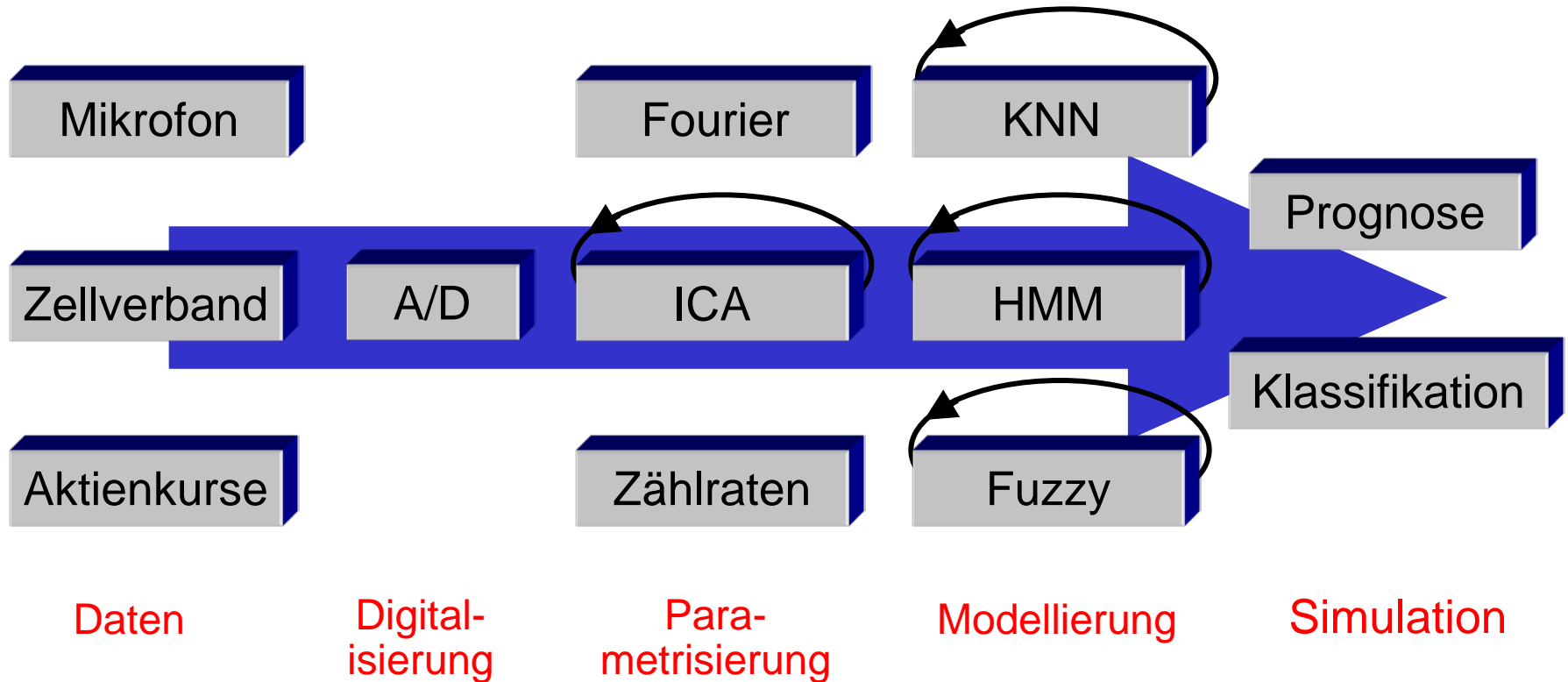


Berechnung linearer/nichtlinearer Zusammenhänge

- Lineare Regression
- Singulärwertzerlegung

- Künstliche Neuronale Netzwerke
- Stochastische Verfahren

Datengetriebene Modellierung

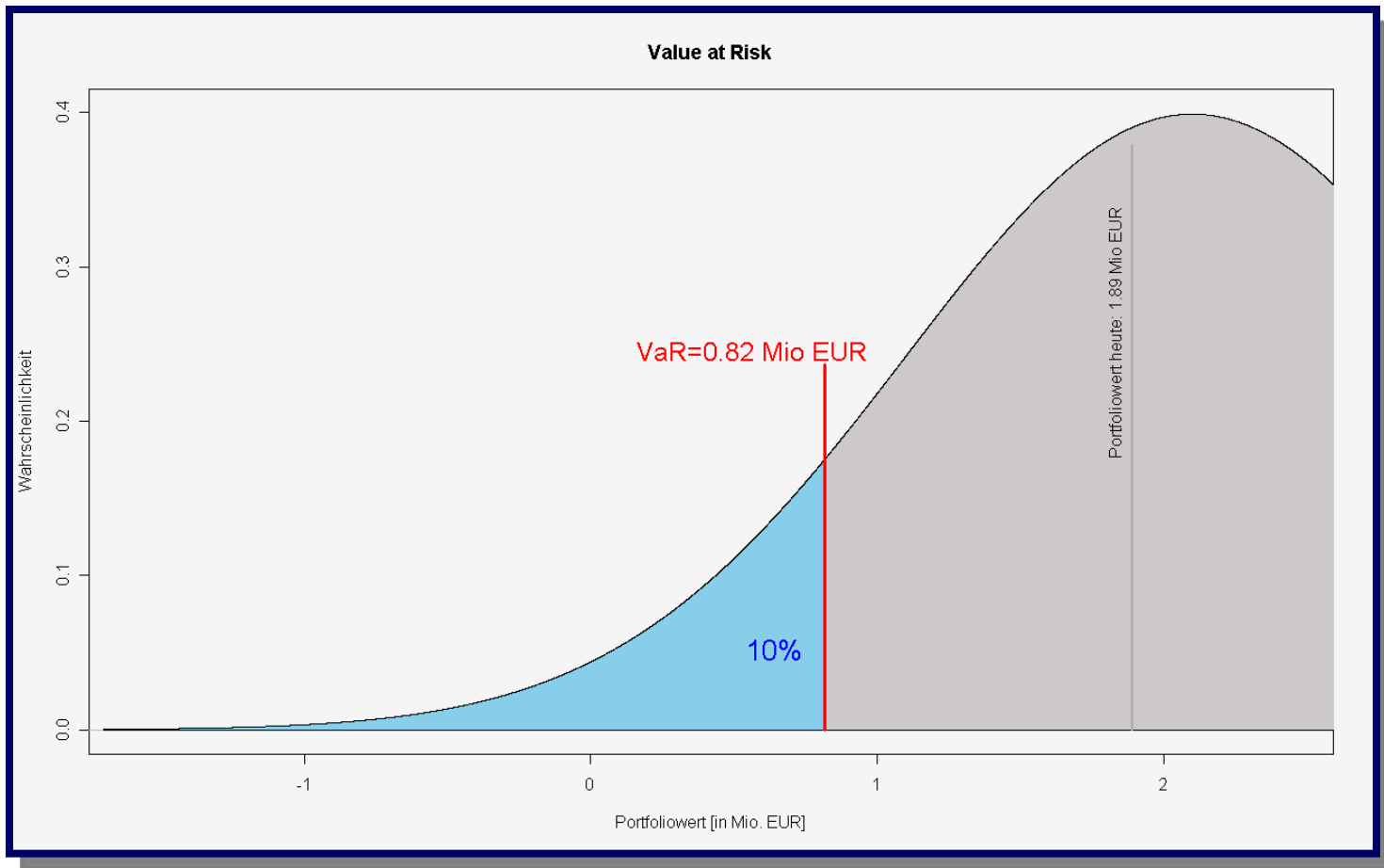


Anwendungen

Value at Risk

- **Monte Carlo Simulation**
- Sammlung und Analyse historischer Daten.
- Modellierung der Verteilung.
- Simulation von Kursentwicklungen gemäß der modellierten Verteilung.
- Große Zahl an Simulationen.
- Prognose von Kursentwicklungen.
- Berechnung des monetären Risikos für die Entwicklung des Portfolios.

Value at Risk

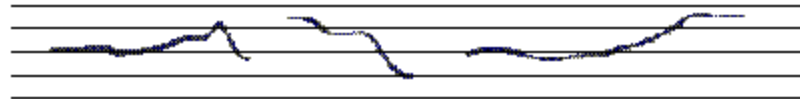


Schätzung der Prosodie für Text-to-Speech

- Rekurrente neuronale Netzwerke
- Sammlung und Analyse gesprochener Sprache eines Sprechers
- Training der Netzwerke
- Schätzung der Lautdauer
- Schätzung des Frequenzverlaufs für stimmhafte Laute

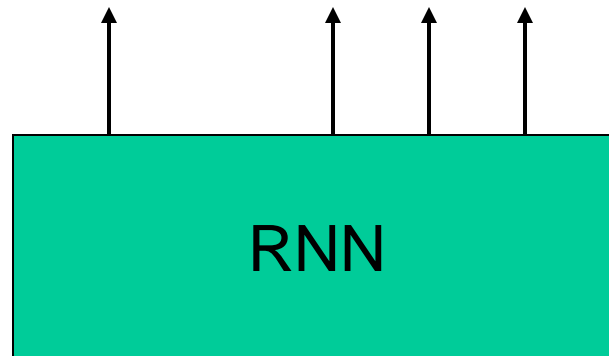
Prosodiegenerierung

Sprechmelodie und -rhythmus



Lautdauer

relative F0-Kontur



t I S **d E k** @
> 1 < **> 1 >** 2

- Wortposition relativ
- Wortart
- Phraseninformation

Deutsche Banken können mit dem Cash Management von Siemens 1 Milliarde Euro aus Kapitalbindung lösen

Wie viel Geld muss in einem Geldautomaten sein? Ein neues Prognose-Tool von Siemens auf der Grundlage jüngster Erkenntnisse der Neuroinformatik - das so genannte Intelligente Cash Management – ermöglicht eine exakte Planung der Befüllungsmengen und -zeitpunkte für Geldautomaten. Transport-, Versicherungs- und Kapitalbindungsaufwände können so auf ein Minimum reduziert werden.....

Dass Siemens bisher alle Vergleichsstudien im Auftrag verschiedener Banken gewonnen hat, führt Dr. Zimmermann auf die 15jährige Erfahrung des Teams von Corporate Technology mit neuronalen Netzen und der Prognose dynamischer Systeme zurück.

Quelle: <http://idw-online.de/pages/de/news75040>

Deutsche Banken können mit dem Cash Management von Siemens 1 Milliarde Euro aus Kapitalbindung lösen

Wie viel Geld muss in einem Geldautomaten sein? Ein neues Prognose-Tool von Siemens auf der Grundlage jüngster Erkenntnisse der Neuroinformatik - das so genannte Intelligente Cash Management – ermöglicht eine exakte Planung der Befüllungsmengen und -zeitpunkte für Geldautomaten. Transport-, Versicherungs- und Kapitalbindungsaufwände können so auf ein Minimum reduziert werden.....

Dass Siemens bisher alle Vergleichsstudien im Auftrag verschiedener Banken gewonnen hat, führt Dr. Zimmermann auf die 15jährige Erfahrung des Teams von Corporate Technology mit neuronalen Netzen und der Prognose dynamischer Systeme zurück.

Quelle: <http://idw-online.de/pages/de/news75040>

Digitale Hilfe bei Preisprognosen

Gute Prognosen der Marktentwicklung sind seit der Liberalisierung des Strommarktes für Stromeinkäufer notwendiger denn je. Werden die Preise im Laufe des Jahres steigen oder fallen und wann ist der günstigste Zeitpunkt, um größere Mengen einzukaufen? Keine einfache Entscheidung für Norbert Fuchs von Siemens Corporate Supply Chain and Procurement und Fritz Bullrich-Mörlbach von Siemens Real Estate Distrikt München, die das Jahreskontingent des Konzerns – etwa 2,3 TWh Strom für 580 Standorte – an börsentechnisch günstigen Tagen ordern müssen

Mit Hilfe eines patentierten mathematischen Modells auf Basis der "Software-Entwicklungsumgebung für neuronale Netze" (SENN) beobachten Zimmermann und sein Team den Strommarkt und errechnen die Preisentwicklung der kommenden zwölf Monate. Der Erfolg ist überzeugend: "Die Trefferquote für die kurzfristige Monatsbetrachtung liegt bei bis zu 80%

Quelle: [Link zur Siemens Website](#)

Energielastprognosen mit Neuronalen Netzen

Alle schienengebundenen Verkehrsmittel gelten als große Energieverbraucher. Der Energieverbrauch variiert allerdings durch bestimmte Rahmenparameter (Werk-, Sonn- und Feiertage; regelmäßige und unregelmäßige Stoßzeiten usw.).

Je unpräziser die Vorhersage des jeweiligen Tagesverbrauches, desto höher sind die Einkaufskosten für elektrische Energie, da Mehrverbrauch unter Umständen zu teuren Konditionen nachberechnet wird bzw. den Betreibern bei Minderverbrauch mögliche Einsparungen entgehen....

Durch den Einsatz der von Siemens patentierten Lösung auf Basis Neuronaler Netze sind Prognosen von höchster Qualität möglich.

Diese Lastprognosen sind eine ideale Planungs- und Entscheidungsgrundlage für

- die strukturierte Strombeschaffung

- die Optimierung von Stromerzeugung

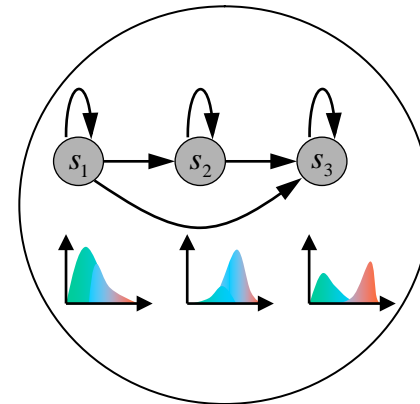
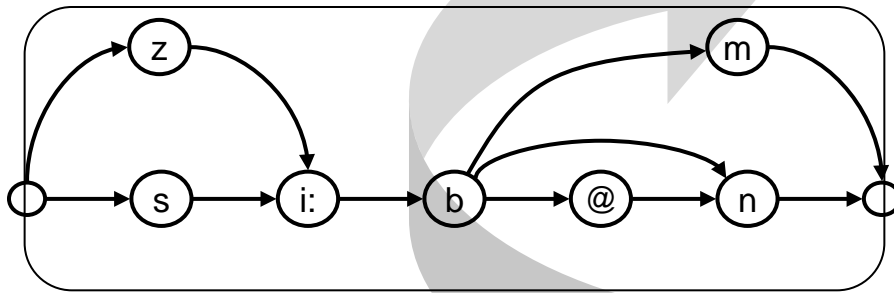
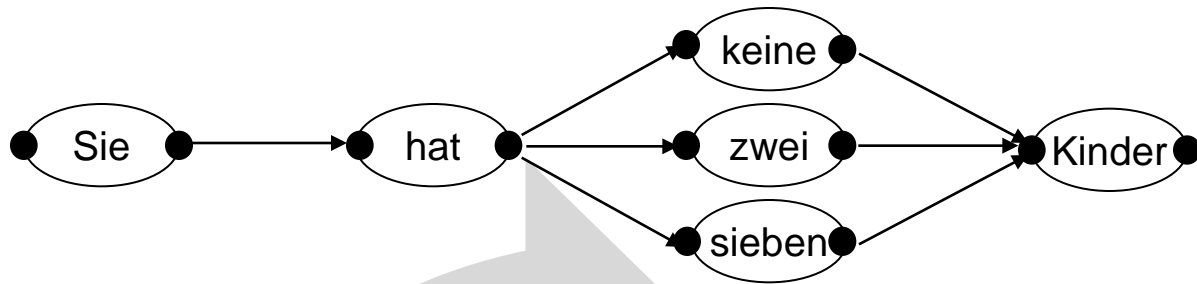
....

Quelle: [Link zur Siemens Website](#)

Spracherkennung

- Hidden Markov Modelle
- Sammlung und Analyse gesprochener Sprache vieler Sprecher
- Berechnung der Parameter des stochastischen Modells
- Schätzung von Emissionswahrscheinlichkeiten
- Berechnung von Hypothesen durch Decodierung von Markovketten 1. Ordnung und Verwendung des Viterbi-Algorithmus

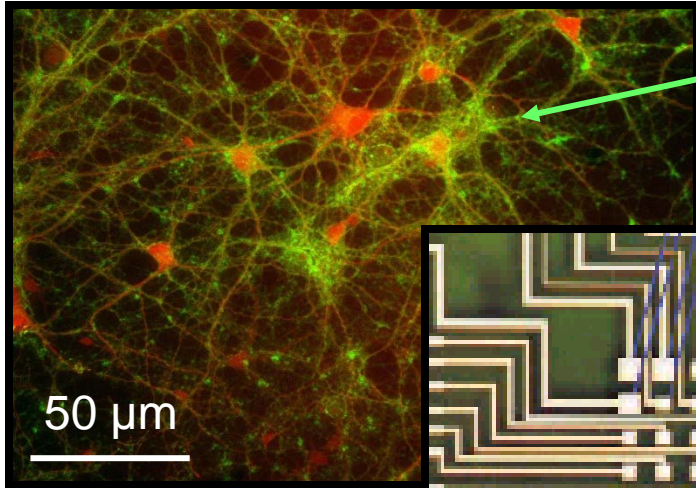
Modellierungsebenen



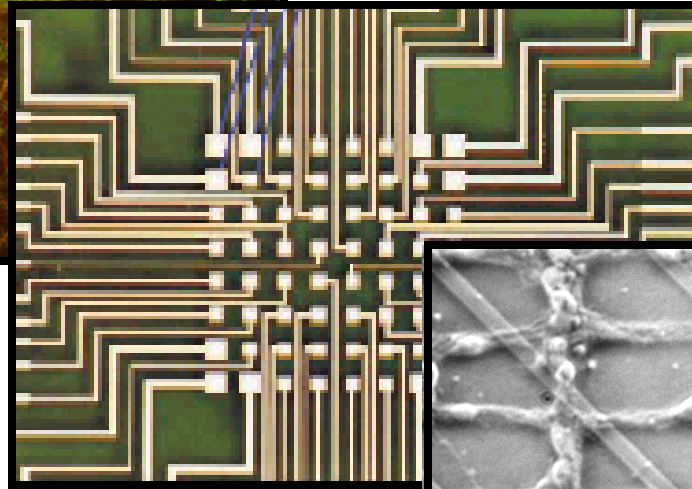
Biosensor

- Zellen (neuronale Zellen oder Herzmuskelzellen) wachsen auf einen Silizium-Chip auf
- Die Signale der Zellen werden abgeleitet und können analysiert werden
- die Zellen befinden sich in einer Nährlösung
- Reaktionsmuster des Zellverbandes auf Schadstoffeinträge oder Pharmazeutika wird modelliert

Recording of Neural Signals

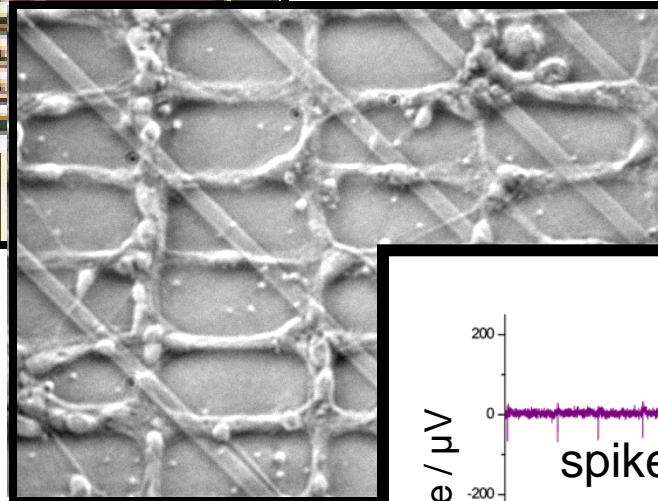


Formation of synapses are the prerequisite for neural communication

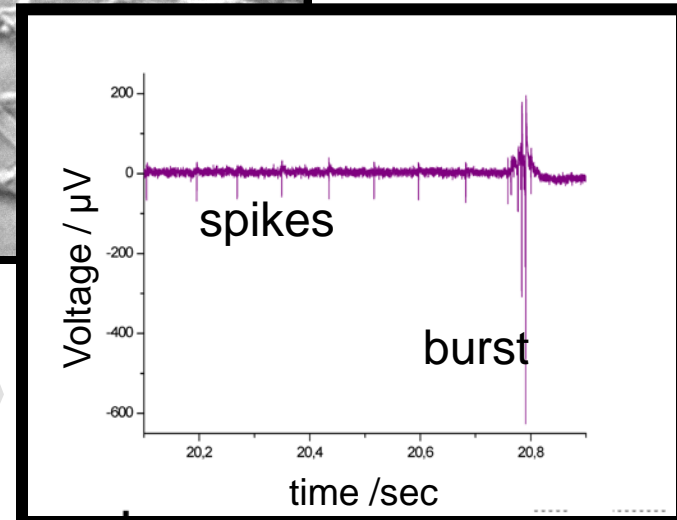


MicroElectrode-Array Chip

Mammalian Neural Net patterned onto MEA-chip

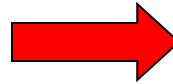
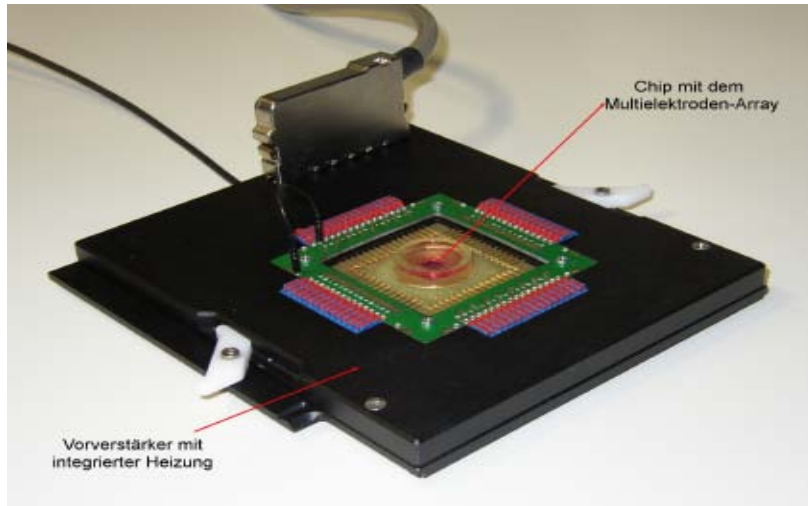


Spontaneous neural signals recorded by the MEA-chip

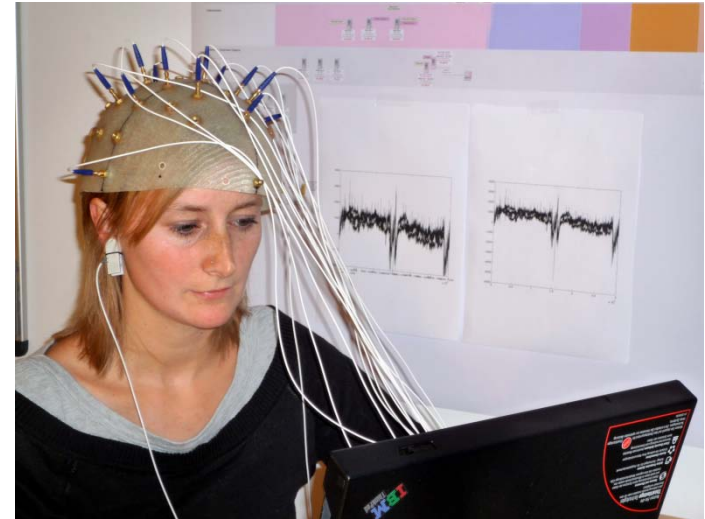


Anwendung 5: Brain-Computer Interface

HMM-basierte Biosensoren

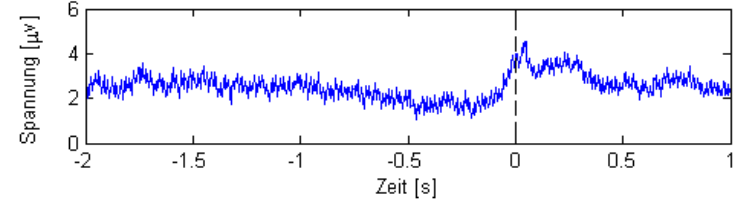
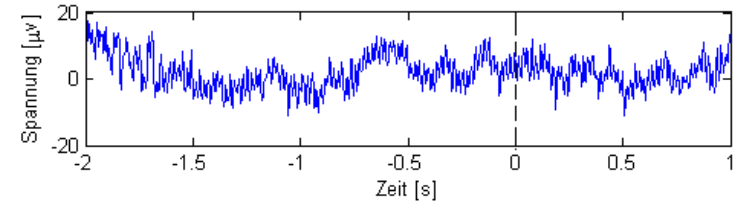
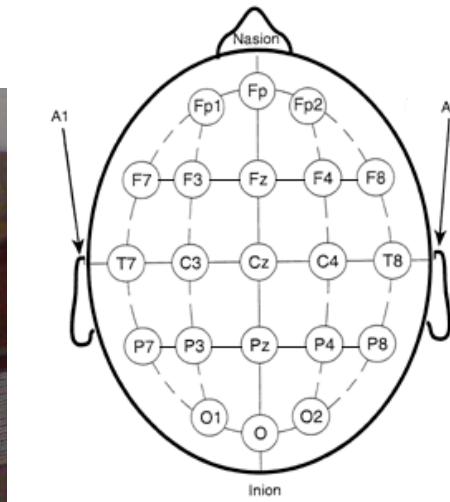
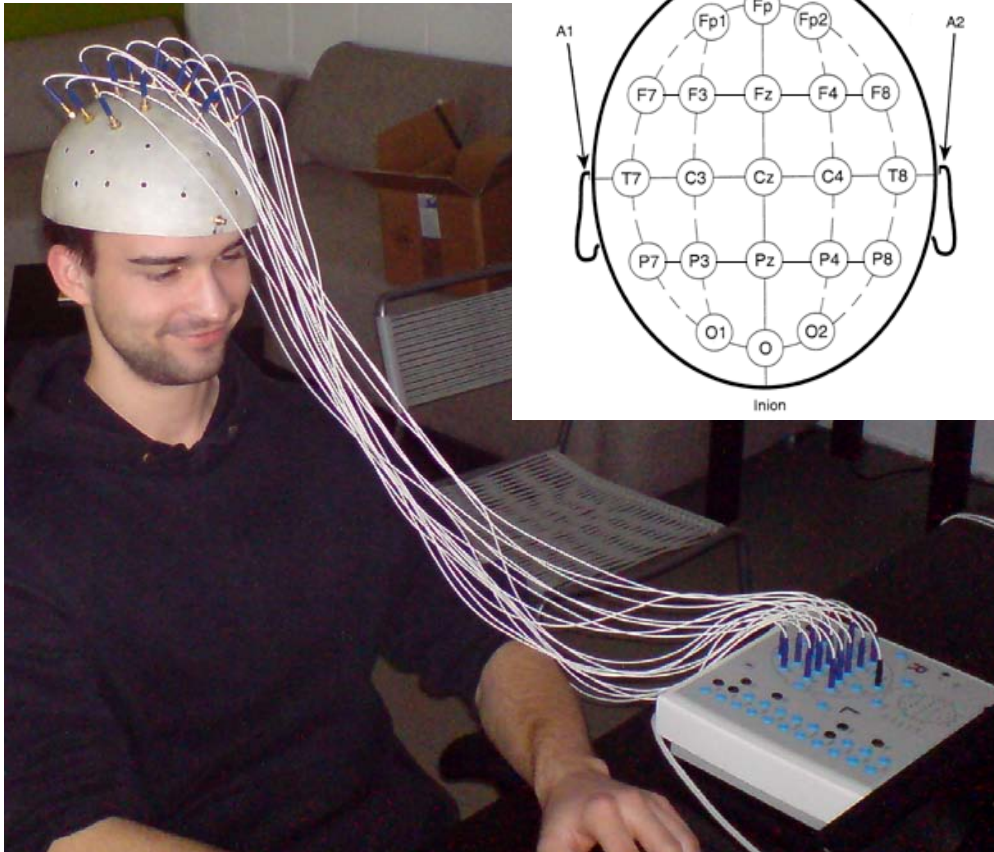


Brain-Computer-Interface



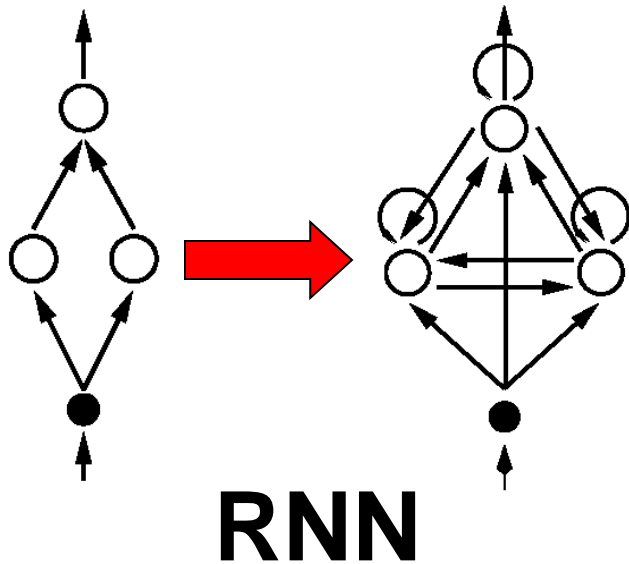
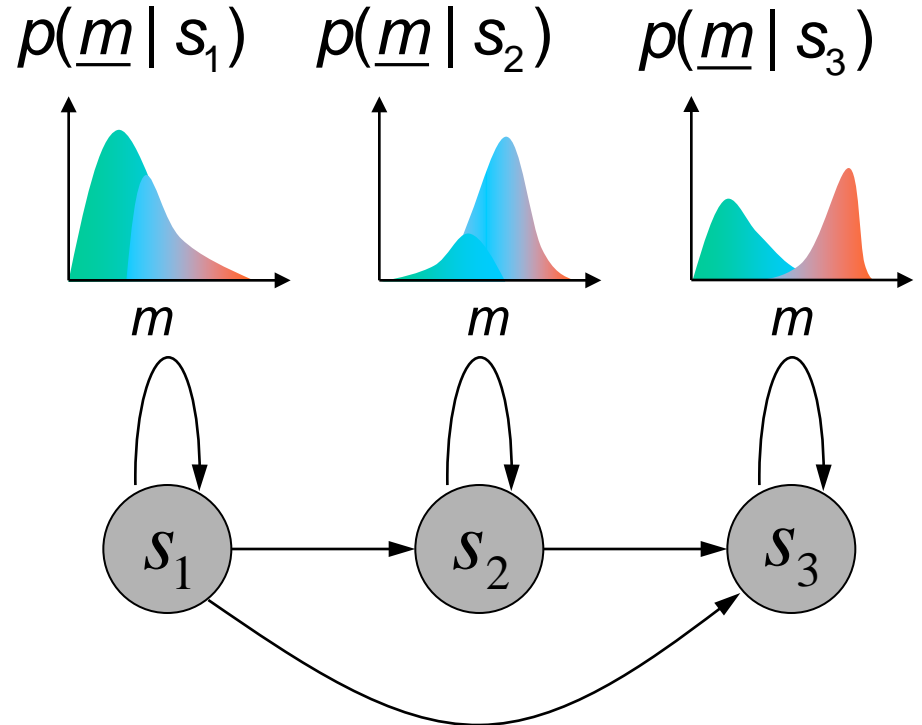
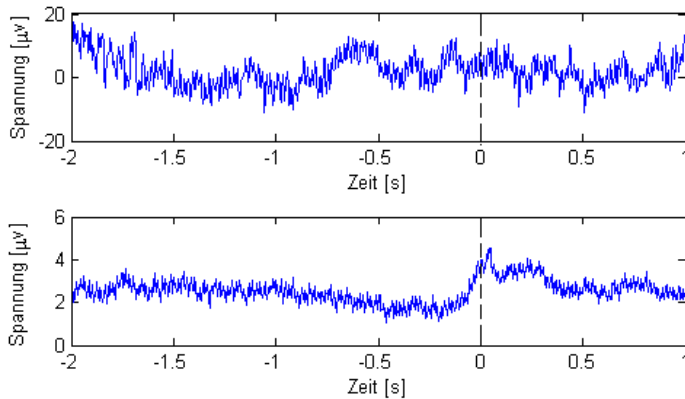
Brain-Computer Interface: Eine Anwendung von Biosensoren

Versuchsaufbau



Aufforderung zur Muskelbewegung bzw. zur gedanklichen Aktivität wird eingeblendet. EEG wird zur Aufzeichnung der neuronalen Aktivitäten genutzt. Das mehrkanalige Signal wird für eine spätere Analyse aufgezeichnet.

Klassifikation von Ereignissequenzen

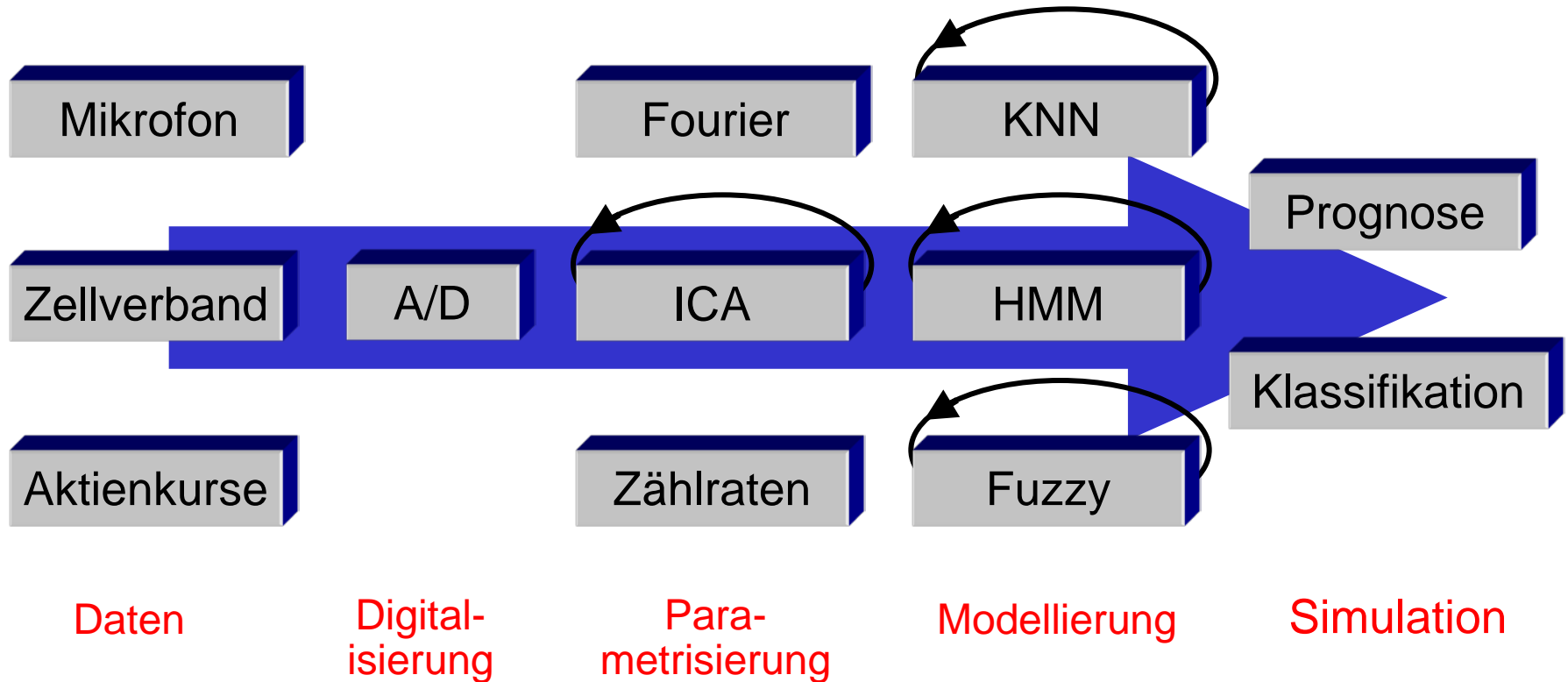


HMM

Konzepttransfer

- Modellbildung ist keiner Disziplin explizit zugeordnet.
- In unterschiedlichen Anwendungsbereichen sind die Techniken zu Modellbildung sehr weit entwickelt.
- Geringe Verbreitung des Wissens.
- Wegen hoher Abstraktion gute Chancen für Konzepttransfer.

Datengetriebene Modellierung



Digitalisierung

Glossar

Digitalisierung

Konvertierung analoger Signale in eine digitale Darstellung.
Beispiele: Tonsignale, Bildsignale, auch: Text

Quantisierung

Darstellung einer Größe in einem System, das nur diskrete Werte darstellen kann (vgl. Quantenphysik).

Abtastung

Aufzeichnung von Messwerten zu diskreten, meist äquidistanten, Zeitpunkten.

Dimensionen eines analogen Signals

- Räumlich (Bild)
- Zeitlich (Sprache)
- Kombination (Film)

Werte eines analogen Signals

- Wertebereich und Genauigkeit hängen von der Sensorik ab.
- Genauigkeit ist auch durch die Möglichkeiten des Aus- bzw. Ablesens der Werte beschränkt.
- Werte können prinzipiell in beliebiger Auflösung erfasst werden.

Wie wird digitalisiert?

- räumliche und/oder zeitliche Abtastung des analogen Signals
- Quantisierung der Messwerte

Abtastung

- Das analoge Signal wird i.a. in äquidistanten Abständen abgetastet.
- Die auf diese Weise gefundene Repräsentation soll eine möglichst vollständige Darstellung des analogen Signals liefern.
- Bei einer Rekonstruktion des analogen Signals soll dies als transparent wahrgenommen werden.

Quantisierung

- Der Wertebereich des Signals wird in Intervalle oder Zellen eingeteilt, die durch einen Wert oder Vektor repräsentiert werden.
- Das Eingangssignal wird durch den Wert oder Vektor quantisiert, der das Intervall bzw. die Zelle repräsentiert, der das Eingangssignal zugeordnet werden kann.

Beispiel: Audiosignale

Glossar

Amplitude

Bezeichnung für die maximale Auslenkung einer Schwingung bzw. einer Welle aus der Mittellage.

Frequenz

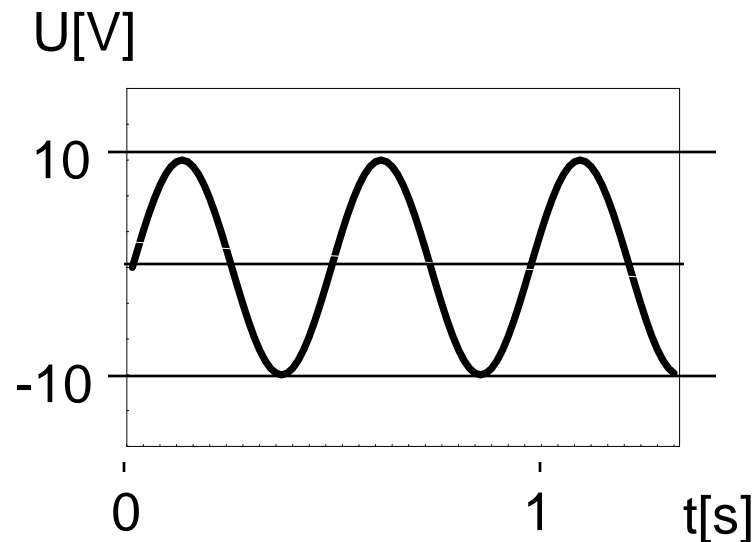
Bezeichnung für die Anzahl von Ereignissen innerhalb eines bestimmten Zeitraums (Hertz (Hz): Ereignisse/Sekunde).

Tiefpass

Filterung eines Signals, so dass Signalanteile mit Frequenzen unterhalb der Grenzfrequenz nicht verändert werden und Signalanteile mit Frequenzen oberhalb der Grenzfrequenz stark abgeschwächt werden.

Frequenz, Phase, Amplitude

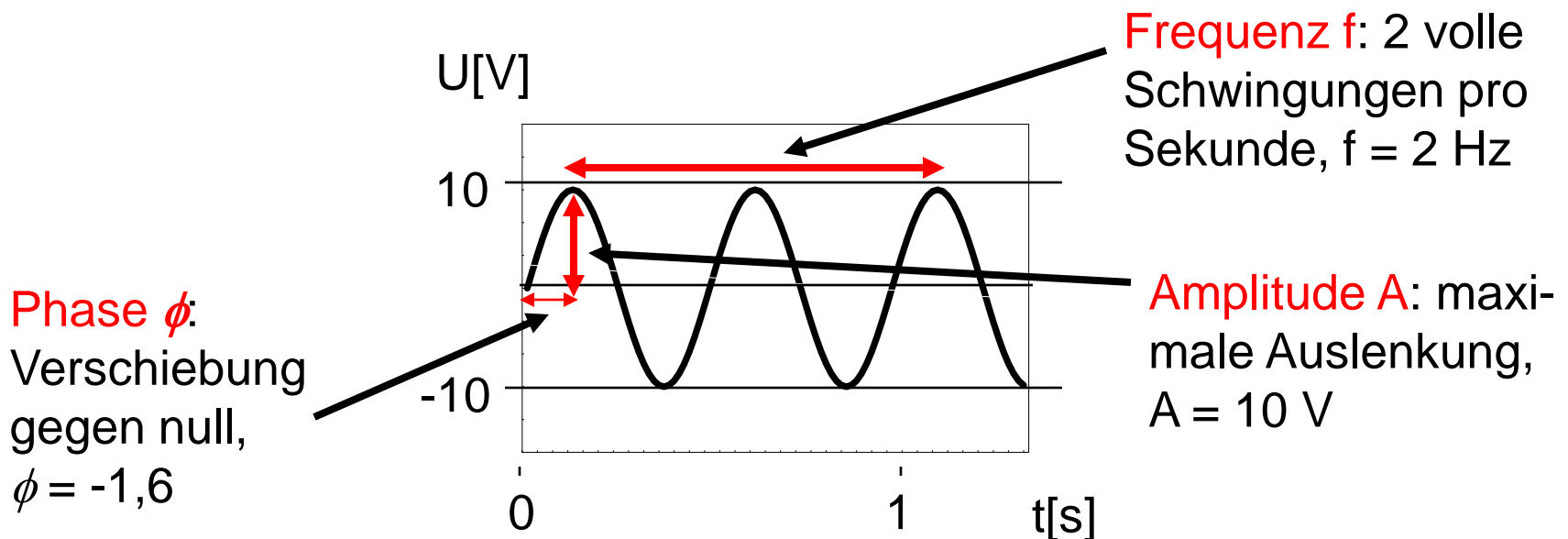
Eine cos-Schwingung (oder sin-Schwingung) $x(t)$ lässt sich durch drei Parameter vollständig beschreiben: Frequenz, Phase und Amplitude



$$x(t) = A \cos(2\pi ft + \phi)$$

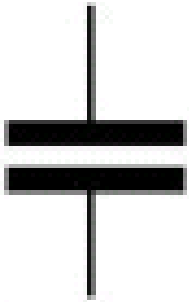
Frequenz, Phase, Amplitude

Eine cos-Schwingung (oder sin-Schwingung) $x(t)$ lässt sich durch drei Parameter vollständig beschreiben: Frequenz, Phase und Amplitude



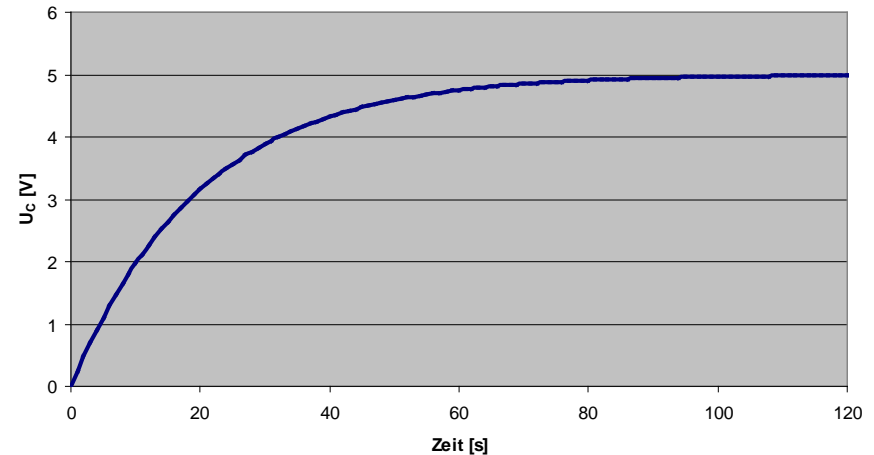
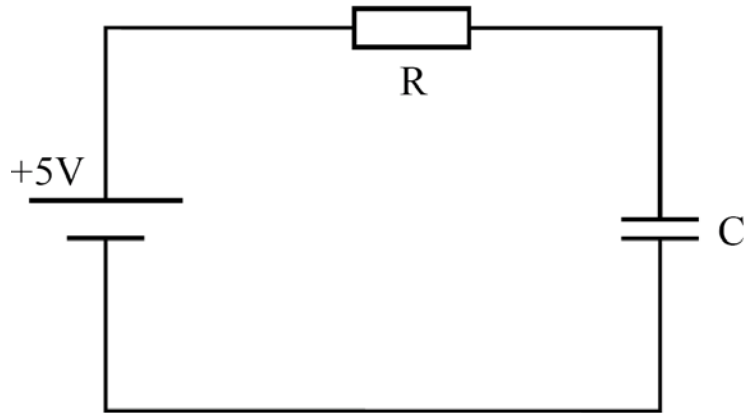
$$x(t) = A \cos(2\pi ft + \phi)$$

Kondensator



- Ladungsspeicherndes Element
- Besteht aus zwei Elektroden, die durch ein nichtleitendes Dielektrikum getrennt sind.
- Die Kapazität C eines Kondensators ist der Quotient aus Ladung Q und Spannung U .
- $C = Q/U$
- Die Einheit der Kapazität C ist das Farad [F].

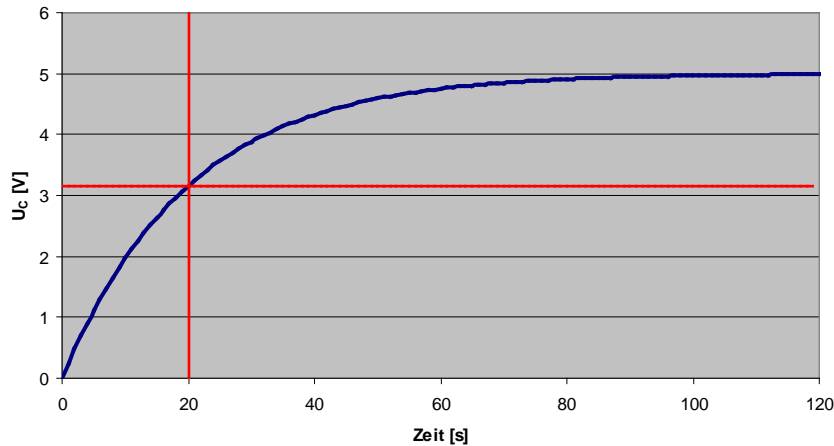
Aufladung eines Kondensators



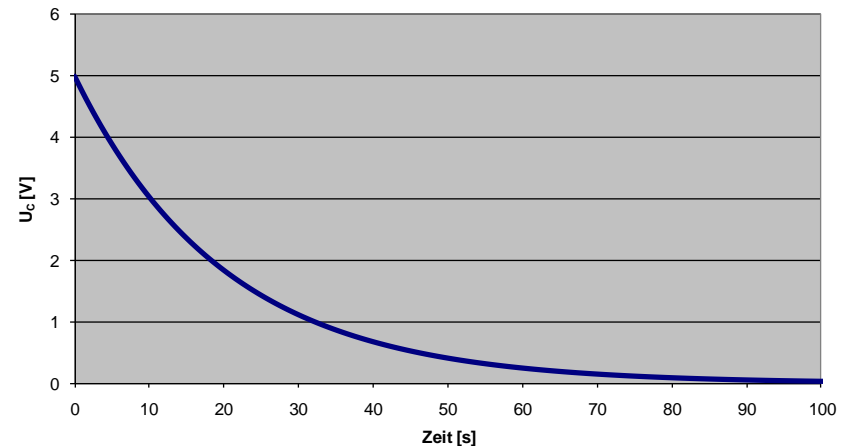
Zeitlicher Verlauf der Aufladung mit $R = 200\text{k}\Omega$ und $C = 100\mu\text{F}$.

$$U_C(t) = U_B (1 - e^{-t/RC})$$

Auf- und Entladung eines Kondensators



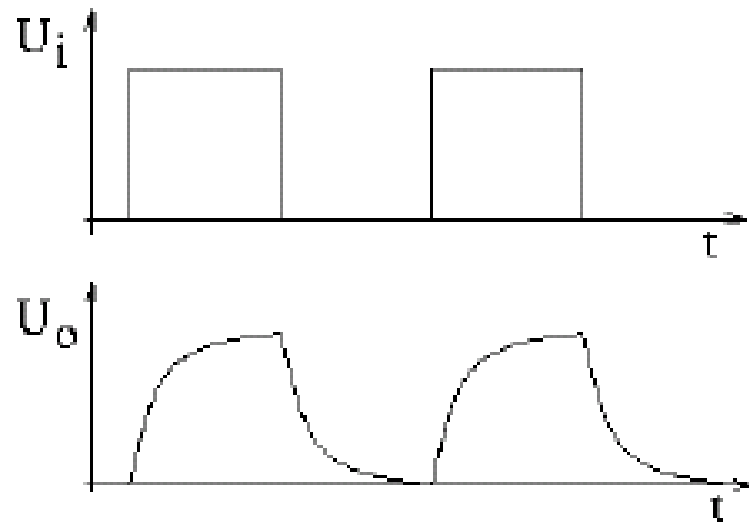
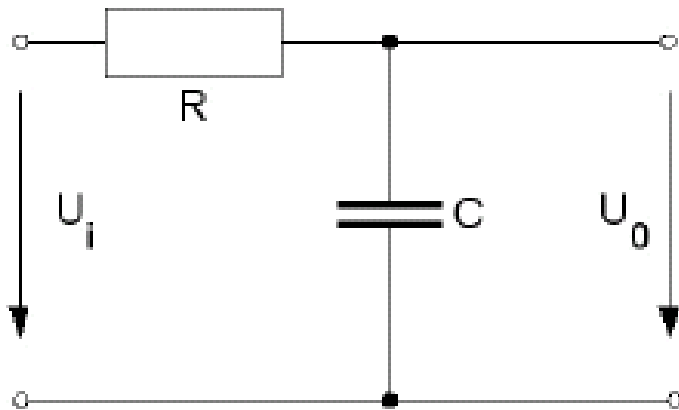
$$U_C(t) = U_B(1 - e^{-t/RC})$$



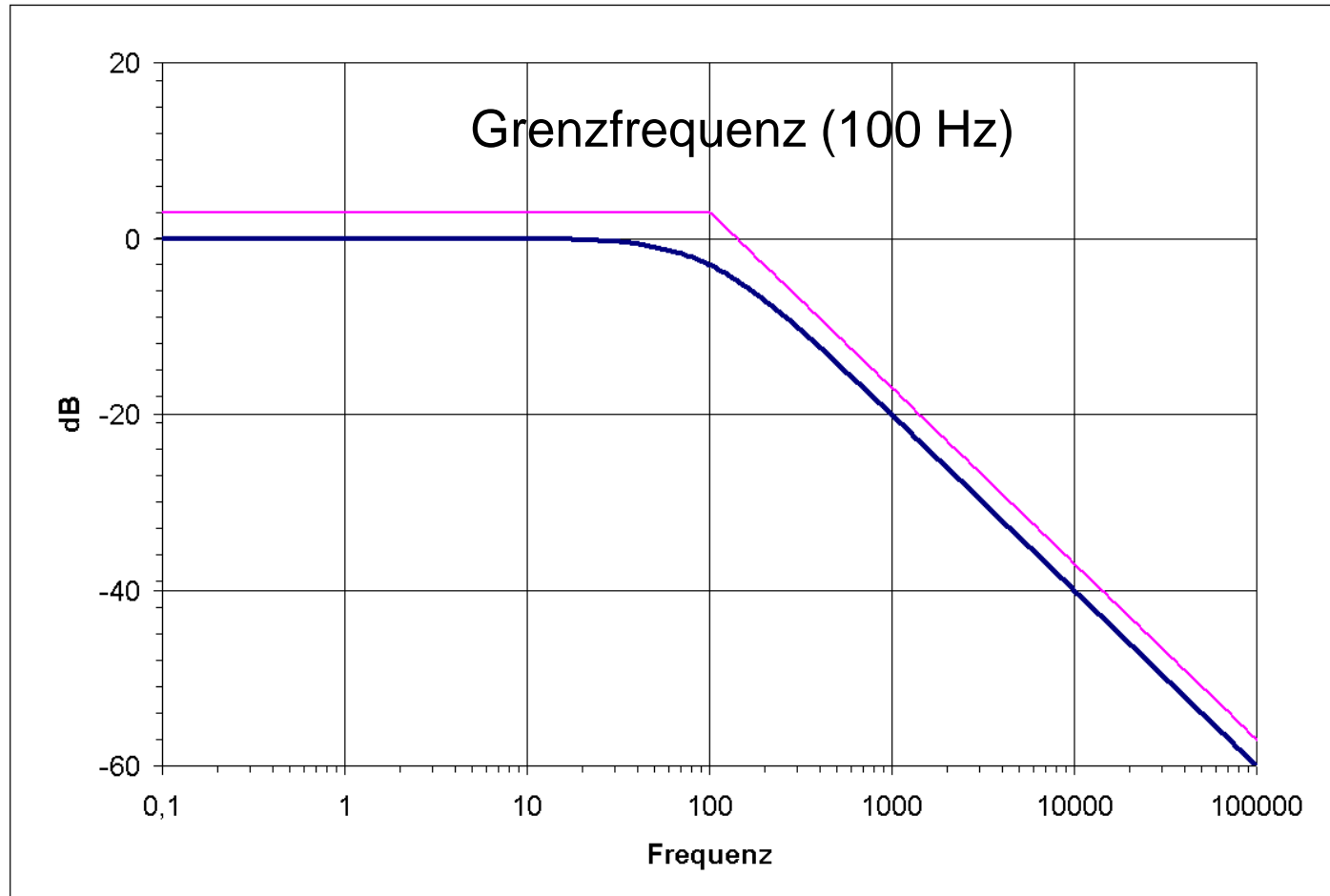
$$U_C(t) = U_B(e^{-t/RC})$$

Die Zeit, die RC entspricht, wird als Zeitkonstante τ der Schaltung bezeichnet.

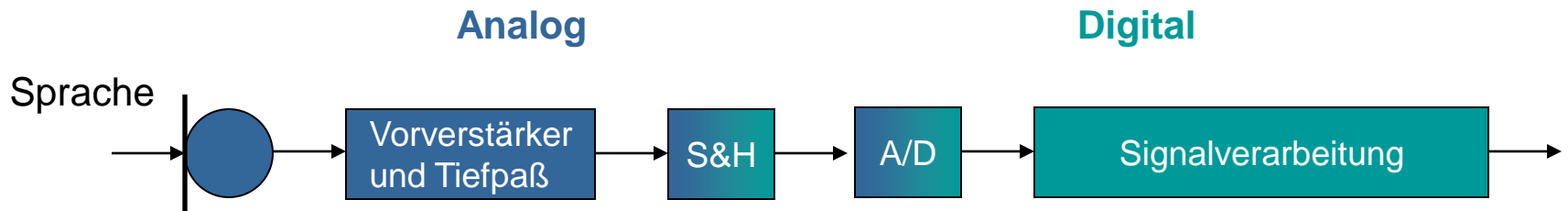
Tiefpass



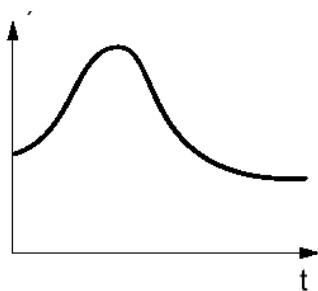
Tiefpass



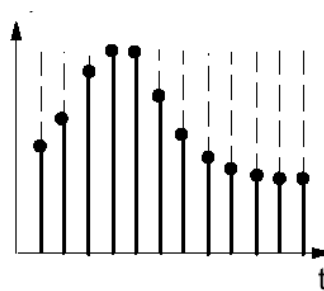
Digitalisierung eines analogen Signals



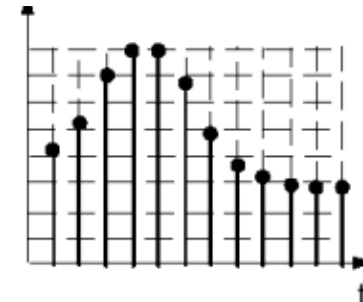
zeitkontinuierlich



zeitdiskret



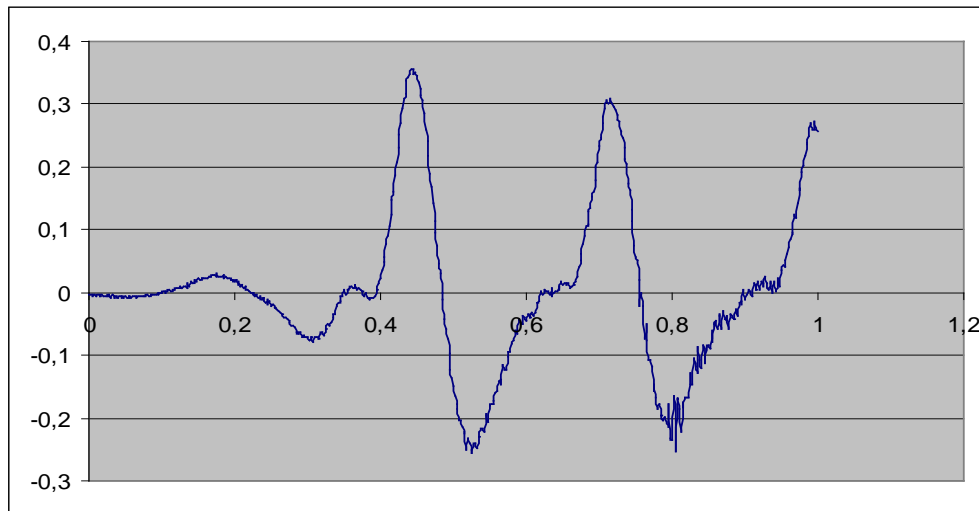
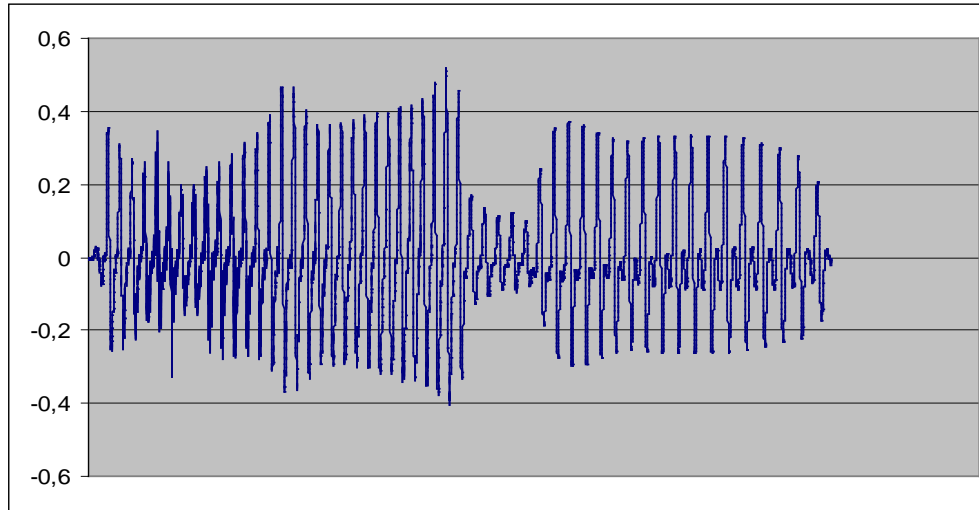
amplitudenquantisiert



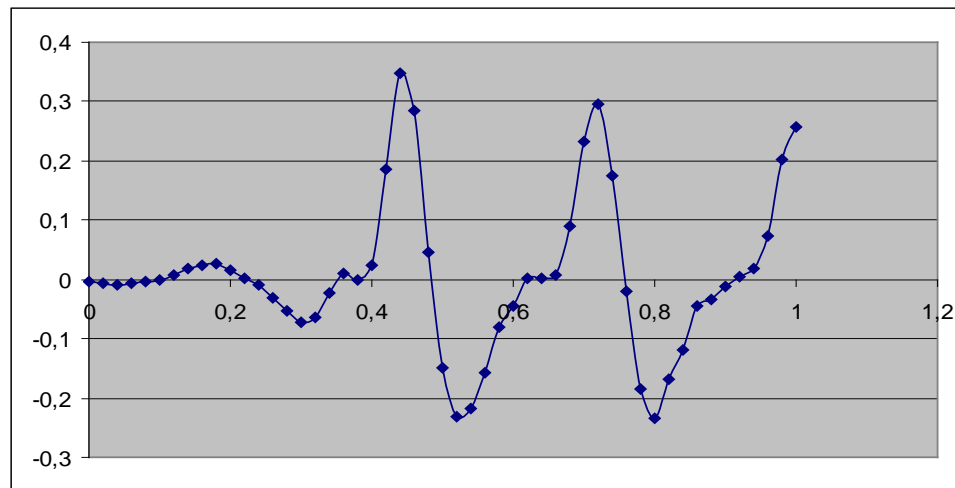
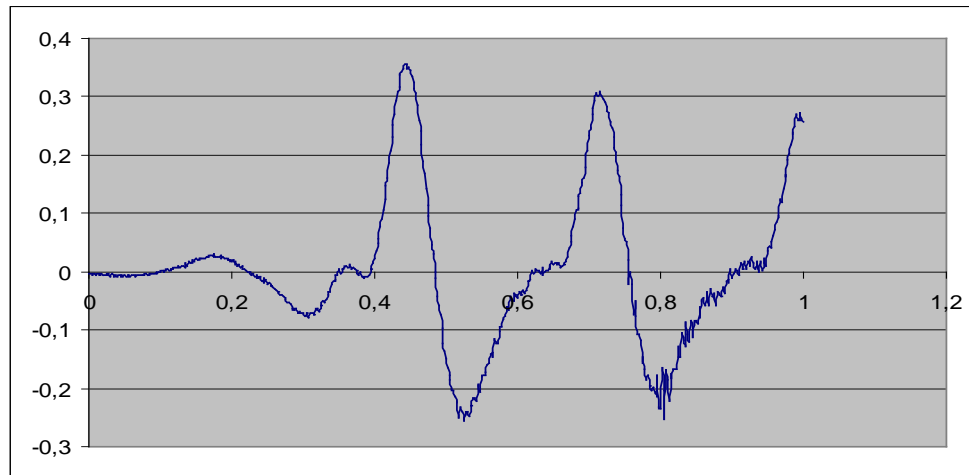
Analog-Digital (AD) - Wandlung

- analoges Mikrofonsignal
 - Verstärker
 - Tiefpassfilter
 - Abtastung (Zeitdiskretisierung)
 - Quantisierung der Amplituden
-
- Speicherung der digitalen Daten in einer Datei

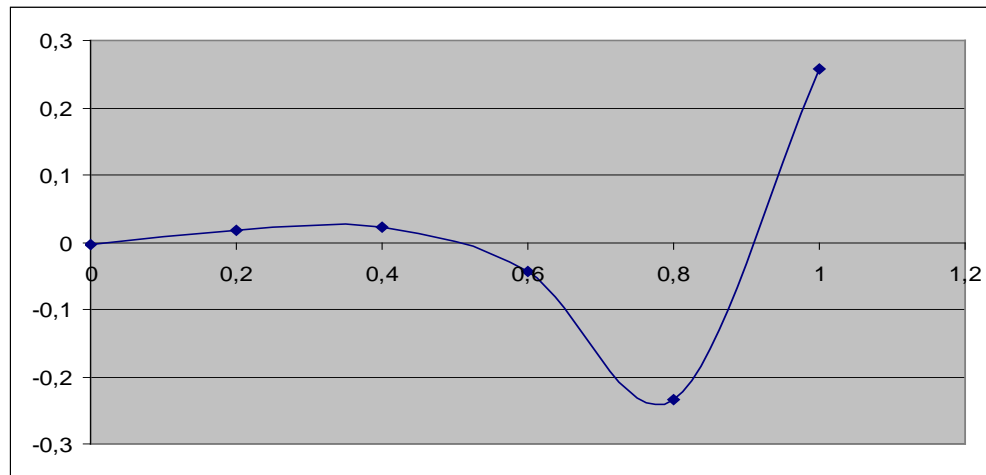
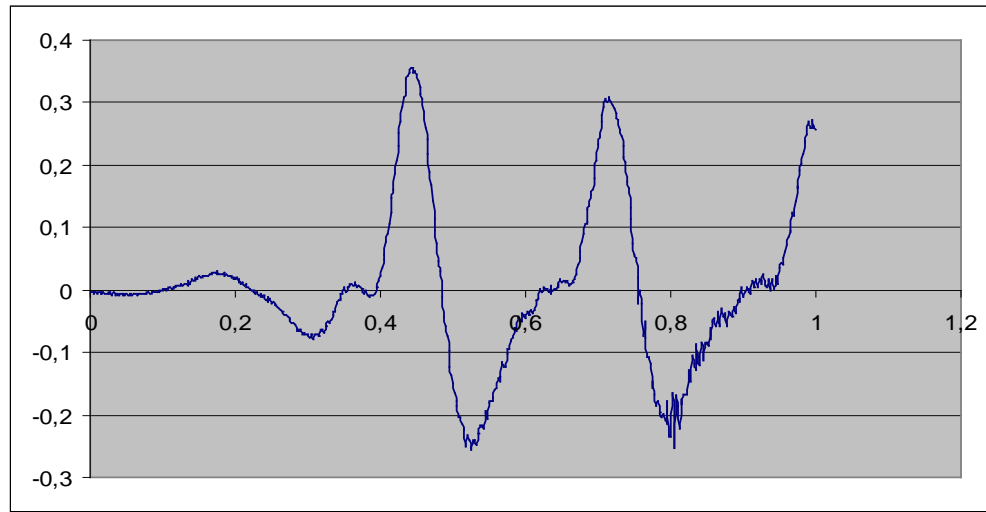
Analoges Signal



Abtastung des Signals



Abtastung des Signals



Abtasttheorem

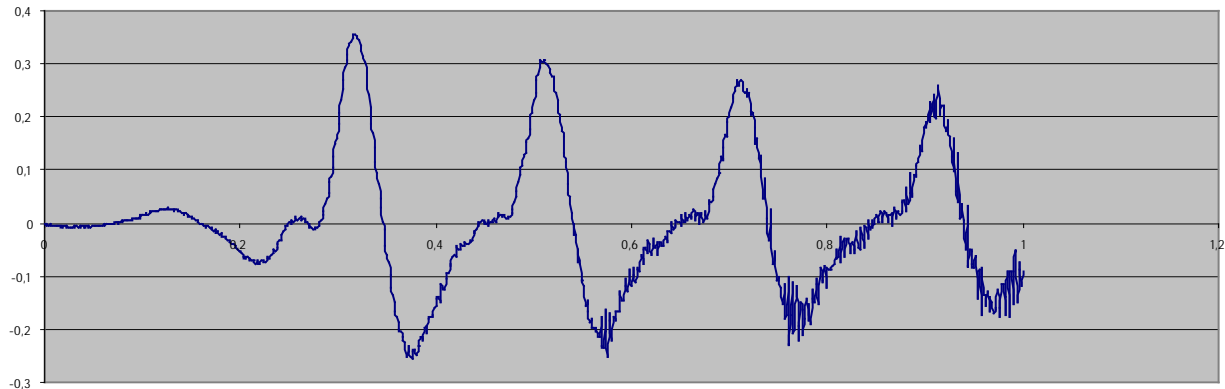
- Die Abtastfrequenz muss mindestens doppelt so hoch sein wie die maximale Frequenz, die im abgetasteten Signal enthalten ist (Shannon).
- Zur Gewährleistung dieser Bedingung muss das Eingangssignal entsprechend gefiltert werden (Tiefpass).

Typische Abtastraten

- DAT 48 kHz
- CD 44.1 kHz
- Breitbandsprache 16 kHz
- Diktiersysteme 11.025 kHz
- ISDN 8 kHz

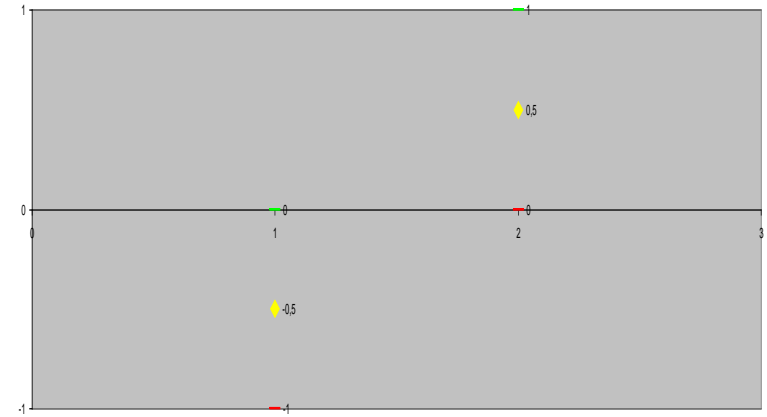
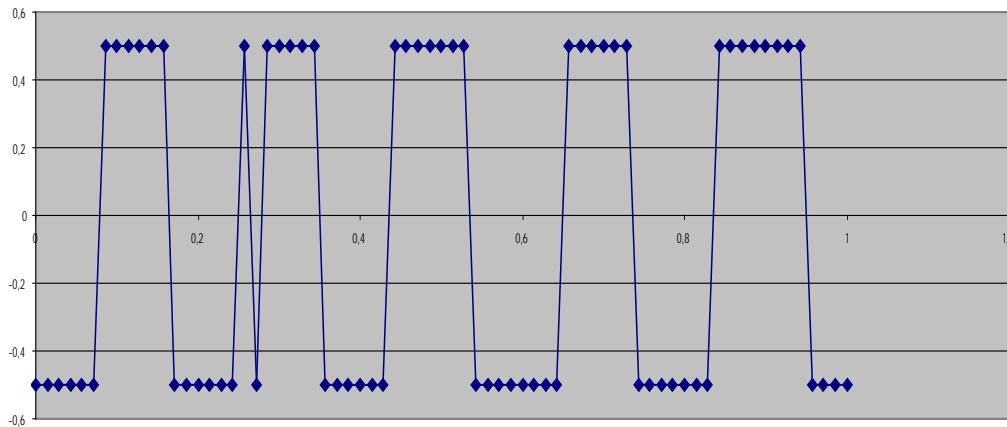
Amplitudenquantisierung

Original



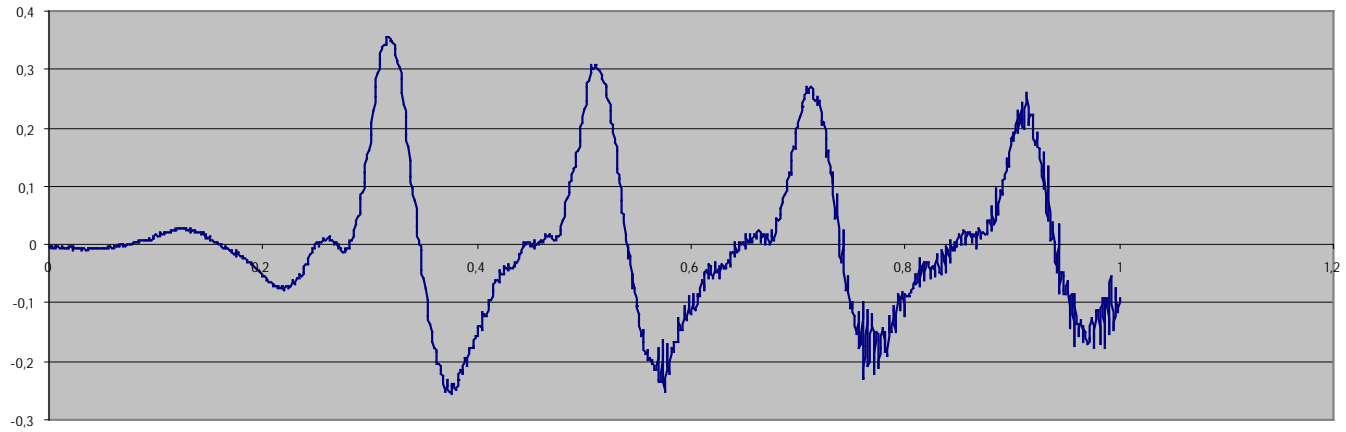
Quantisiert (1 bit)

Intervalle

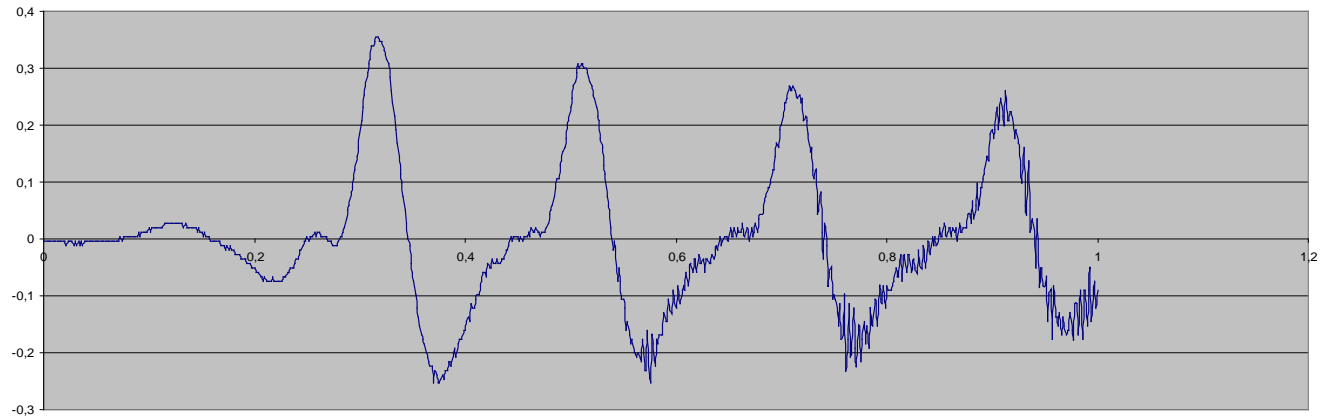


- Untere Grenze - Obere Grenze ♦ Repräsentant

Original



Quantisiert (8 bit)



Optimale Quantisierung

- Die Intervalle bzw. Zellen werden so gewählt, dass bei gegebener Anzahl der Quantisierungsstufen bzgl. einer gegebenen Datenmenge der Quantisierungsfehler minimal wird.
- In diesen Prozess geht insbesondere die Verteilungsdichte der gegebenen Datenmenge ein.

Ziel

- Das Ziel ist die Realisierung einer digitalen Repräsentation eines analogen Signals, die für den Betrachter oder bzgl. des Analysesystems denselben Informationsgehalt hat wie das analoge Signal.
- Beispiel: MPEG

Optimierte Quantisierung für Sprache

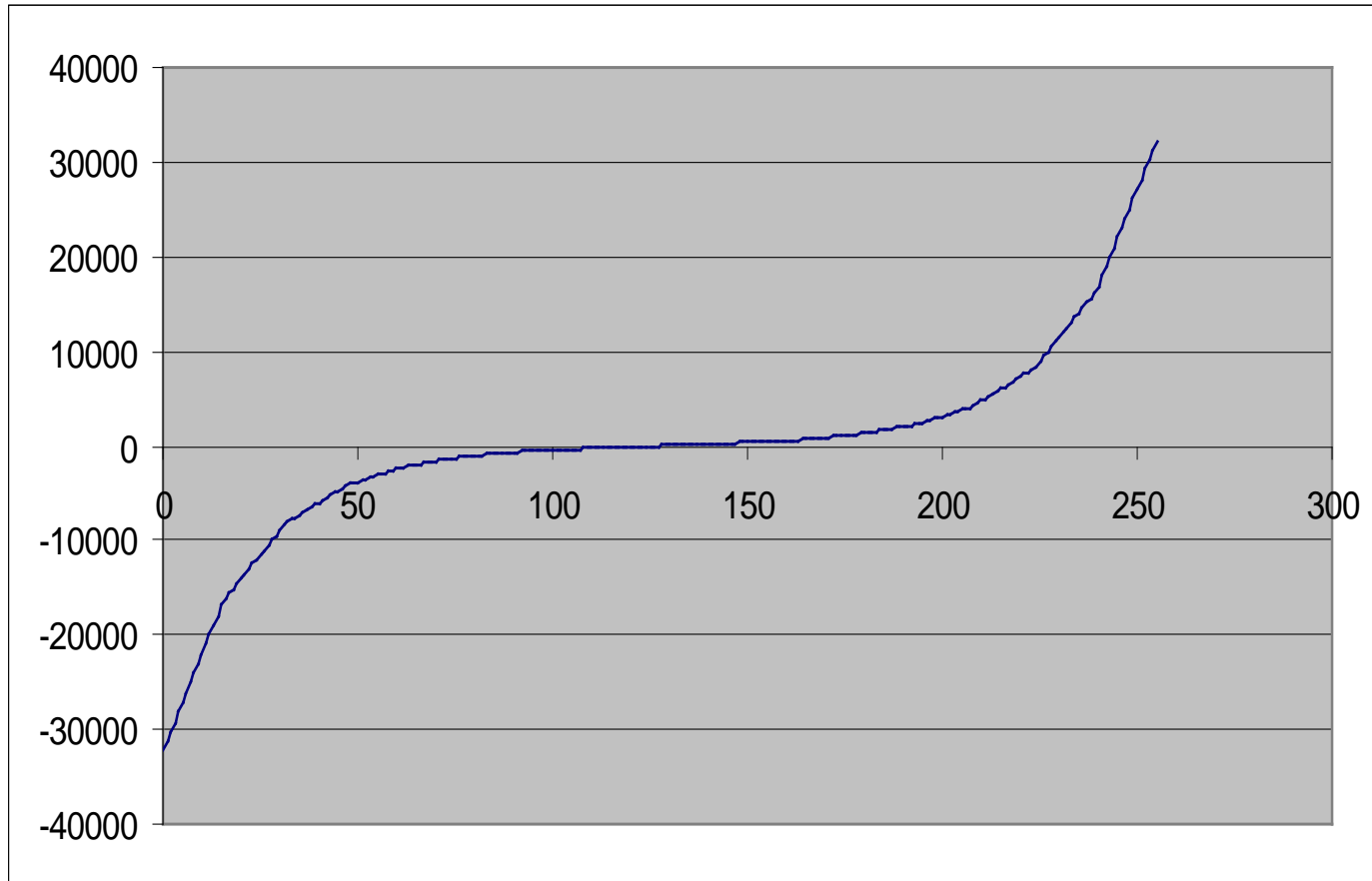
- Das menschliche Gehör nimmt Unterschiede bei leisen Signalen stärker wahr als bei lauten.
- Bei gleichförmiger Quantisierung (konstante Größe der Intervalle) werden Quantisierungsfehler daher bei leisen Signalen deutlicher wahrgenommen als bei lauten.
- Für eine optimierte Quantisierung werden die Intervallgrößen mit Hilfe von Segmenten unterschiedlicher Steigung gewählt.
- Mit 8 Bit gelingt eine Quantisierung, die der Qualität einer gleichförmigen Quantisierung mit 12 Bit entspricht.
- Amerika, Japan: μ -law (15 Segmente); Europa: a-law (13 Segmente)

Tabelle a-law

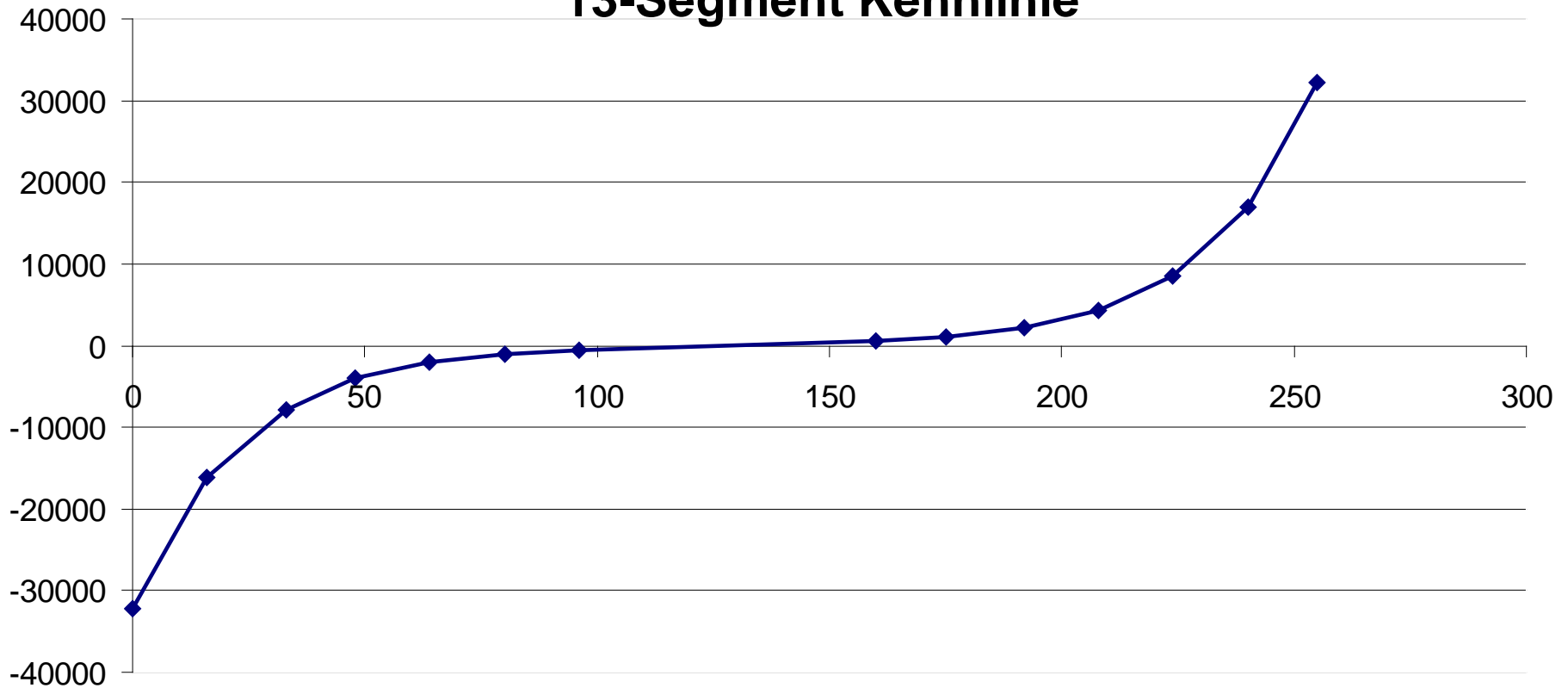
a-law	short int
0	-32256
1	-31232
2	-30208
...	...
125	-40
126	-24
127	-8

a-law	short int
128	8
129	24
130	40
...	...
253	30208
254	31232
255	32256

a-law Codierung (8 bit)



13-Segment Kennlinie



Beschränkungen der Digitalisierung

- **Abtastung**

bedingt eine Beschränkung der maximalen Frequenz des analogen Signals

- **Quantisierung**

Fehler durch die Quantisierung der Amplitude kann durch den Signal-Rausch-Abstand (SNR) bewertet werden:

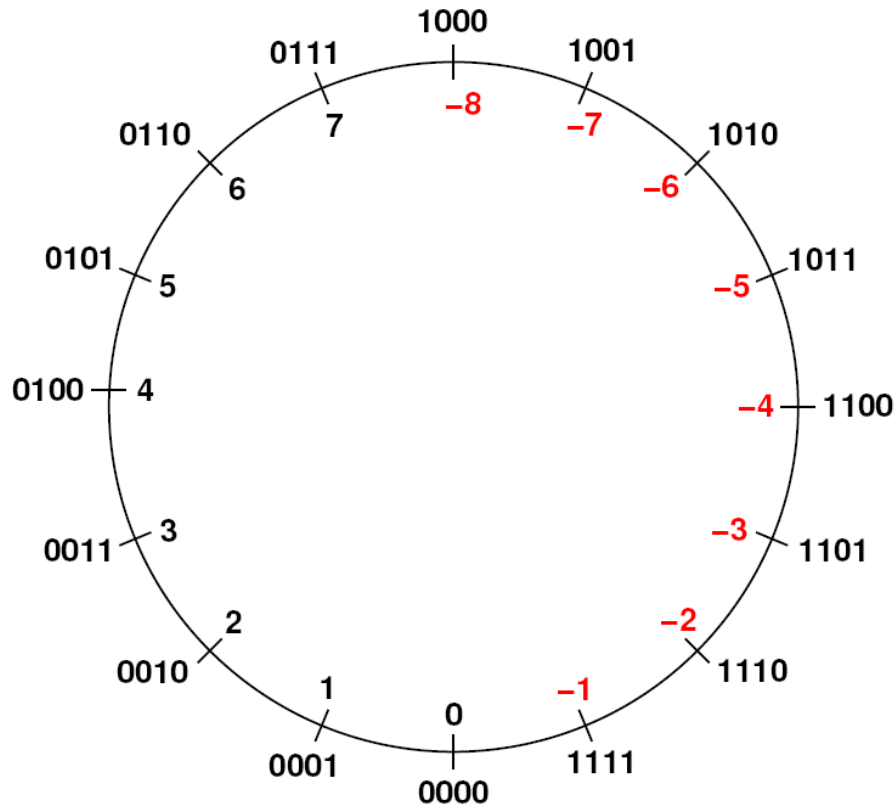
$$SNR = 10 * \log \frac{\sum x[n]^2}{\sum (\tilde{x}[n] - x[n])^2}$$

Dateiformate

- pcm, raw: ohne Header
- wav: Microsoft
- au: Sun Microsystems
- mp3: MPEG Layer 3

- beachte: Byte-Order

Integer-Zahlen (2er-Komplement)



Most Significant Bit (MSB) definiert das Vorzeichen.
positive Zahl: MSB = 0,
3 niederwertigen Bits werden wie gewohnt ausgewertet

negative Zahl: MSB = 1,
3 niederwertigen Bits werden invertiert, 1 addiert und anschließend ausgewertet

Wertebereich:


$$-2^{n-1} \leq Z \leq 2^{n-1} - 1$$

Beispiel 2er-Komplement

Zu berechnen: $(3)_{10} - (6)_{10} = (-3)_{10}$

$(3)_{10}$: $(0011)_2$

-6: Betrag 110, invertiert 001, 1 addiert 010

 $(-6)_{10}$: $(1010)_{2-2k}$

	0	0	1	1	
+	1	0	1	0	101, invertiert 010, 1 addiert 011
	0	0	1	0	
<hr/>					
	1	1	0	1	Ergebnis: $(1101)_{2-2k} = (-3)_{10}$

Little Endian – Big Endian

Bei der Interpretation, z.B. einer Integer-Zahl, ist zu unterscheiden, ob das least significant Byte an der höchsten Adresse („big endian“) oder der niedrigsten Adresse („little endian“) gespeichert wird.

Beispielsweise Speicherung eines 4 Byte Integer:

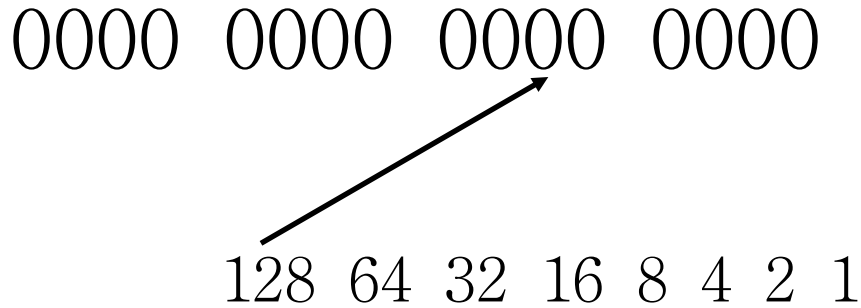
MSB → Byte3 Byte2 Byte1 Byte0 ← **LSB**

<u>Speicheradresse</u>	<u>Little Endian</u>	<u>Big Endian</u>
BasisAdresse+0	Byte0	Byte3
BasisAdresse+1	Byte1	Byte2
BasisAdresse+2	Byte2	Byte1
BasisAdresse+3	Byte3	Byte0

Little Endian: PC (Intel, AMD), Apple (Intel), ...

Big Endian: Apple (Motorola), Sun, ...

Beispiel: -96 (signed short integer)



000 0000 0110 0000

111 1111 1001 1111

111 1111 1010 0000

1111 1111 1010 0000

FF A0  A0 FF

Betrag

1er-Komplement

2er-Komplement

Vorzeichen

Byte-Order

Hex-Editor auf einem PC (Intel)

Hex-Darstellung
in einem PC

Interpretation als
signed 2 Byte Integer

4F 00

79

FF FF

-1

A1 FF

-95

Header einer .wav-Datei (Struktur)

Length	description	String
4 Byte	< Magic Number RIFF >	“RIFF”
4 Byte	< size of file >	
4 Byte	< Magic Number WAVE >	“WAVE”
 chunks		
4 Byte	< ID of chunk >	f.e. “fmt “
4 Byte	< size of chunk >	
..... data		
4 Byte	< ID of chunk >	f.e. “fact”
4 Byte	< size of chunk >	
..... data		
... more chunks		

Header einer .wav-Datei (Details)

Length	description	name
chunk: header information		
4 Byte	< ID of header >	“fmt “
4 Byte	< size of header segment (Byte) >	
2 Byte	< format tag >	
2 Byte	< number of channels >	
4 Byte	< number of samples per second >	
4 Byte	< number of bytes per second >	
.....	more information	
chunk: audio data		
4 Byte	< ID of data >	“data”
4 Byte	< size of data segment (Byte) >	
.....	audio data (f.e. 2 Byte/sample)	

Fourier-Transformation

Fourier-Transformation

- Jean Baptiste Joseph Fourier (1768-1830)
- Französischer Mathematiker und Physiker
- Fourier (1807): „Ein beliebiges kontinuierliches und periodisches Signal kann durch adäquat gewählte Sinus- und Cosinus-Funktionen vollständig repräsentiert werden.“

Fourier-Transformation

$$F(\omega) = \frac{1}{\sqrt{2\pi}} \int e^{-i\omega t} f(t) dt$$

$$f(t) = \frac{1}{\sqrt{2\pi}} \int e^{i\omega t} F(\omega) d\omega$$

Fourier-Transformation

