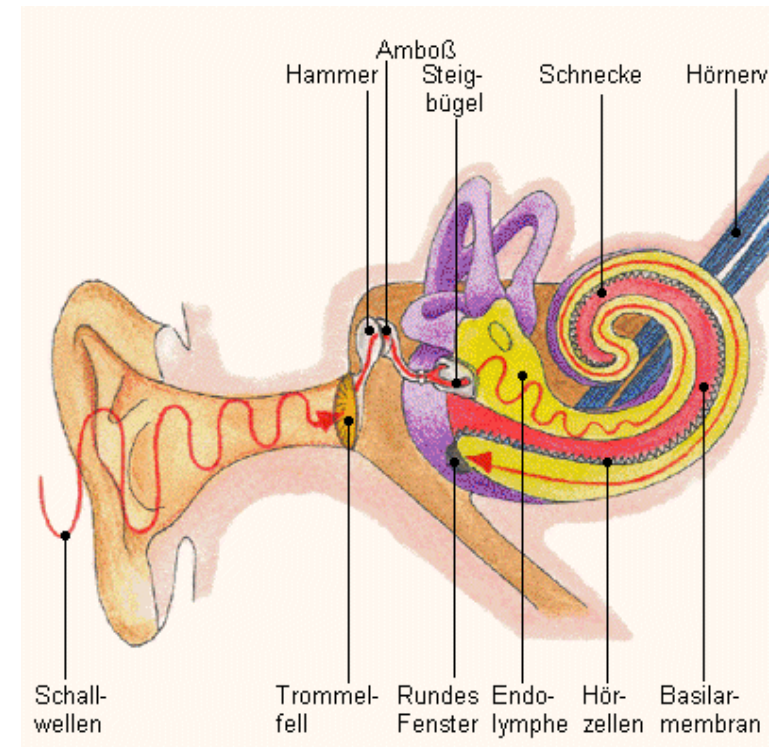


Audio: Agenda

- Schall und Schallausbreitung
- Aufbau des Gehörs
- Hörwahrnehmung
- Sprache, Vokalartikulation

- Signalverarbeitung
- Abtastung, Abtasttheorem
- A/D- und D/A-Wandlung
- Filter, elementare Algorithmen

- Kodierung von Audiodaten
- psychoakustische Verfahren, MP3



Audio: Literatur

K.v.d.Heide, Skripte zu "Signalverarbeitung" und " Nachrichtentechnik",

<http://tams-www.informatik.uni-hamburg.de/lehre/> (Interaktive Skripte, Matlab erforderlich)

Mathworks, Inc., Matlab 5.3 User and Toolbox Manuals, www.matlab.com

B. Kollmeier, Vorlesungsskripte Audiologie, www.physik.uni-oldenburg.de/Docs/medi/download/

M. Bosi, R.E. Goldberg, Introduction to digital audio coding and standards, Kluwer, 2004

U. Zölzer, Digitale Audiosignalverarbeitung, Teubner 1996, 2 Aufl. 1999

P.Gerdsen, P.Kröger, Digitale Signalverarbeitung in der Nachrichtenübertragung, Springer 1997

R.W.Hamming, Digital Filters, Prentice Hall, 1983

W.H.Press, B.P.Flannery, S.A.Teukolsky, W.T. Vetterling, Numerical Recipes, Cambridge Univ. Press

U. Tietze, Ch. Schenk, Halbleiter-Schaltungstechnik, Springer 1993 (analoge Schaltungen)

IEEE Journal Signal Processing

diverse Konferenzen zum Thema, Audio u.a. Audio Engineering Convention

DSP-Datenbücher (www.motorola.com, www.analog.com)

Motivation

wozu die Gehörwahrnehmung untersuchen?

- Forschung / Erkenntnisgewinn

möglichst optimale Verfahren zur:

- Aufzeichnung / Speicherung / Wiedergabe von Audio
- so gut wie nötig - aber nicht besser (!)

- psychoakustische Verfahren, etwa MP3-Kompression
- bessere Algorithmen für Musik und Musikproduktion
- Raumsimulation für Spiele oder virtuelle Realität

- möglichst effiziente (low-level) Sprachverarbeitung
- medizinische Anwendungen, insb. Hörgeräte

Schall

physikalischer Schallbegriff:

- alle elastischen Schwingungen eines Materials
- im allg. Fall longitudinal und transversal
- aber in Gasen (Luft) nur longitudinal (=Druckschwingungen)

- Schallgeschwindigkeit ca. 340 m/s in Luft
- Schallintensität lokale Druckänderung

Luftdruck auf Meereshöhe	1 bar	= 10.000 Pa
Schmerzschwelle	1 mbar	= 10 Pa
wahrnehmbare Druckänderung	10E-9 bar	= 0.00001 Pa

- Ohr hat optimale Empfindlichkeit:
- weitere Steigerung: Wahrnehmen der Molekularbewegung

Schall: Leistung, dB

- Kenngrößen des Schallfelds: Druckamplitude P bzw. Intensität I
- Gesamtleistung: Integral der Intensität über eine Oberfläche
- Lautstärke (Intensität): dB

$$L / \text{dB} \quad := 10 \log_{10}(I / I_{\text{ref}})$$
$$\quad := 20 \log_{10}(P / P_{\text{ref}})$$

$$P_{\text{ref}} \quad := 0.00002 \text{ Pa}$$

- alternativ: dB als Maß für das Verhältnis (p1/p2)

20 dB zehnfacher Schalldruck

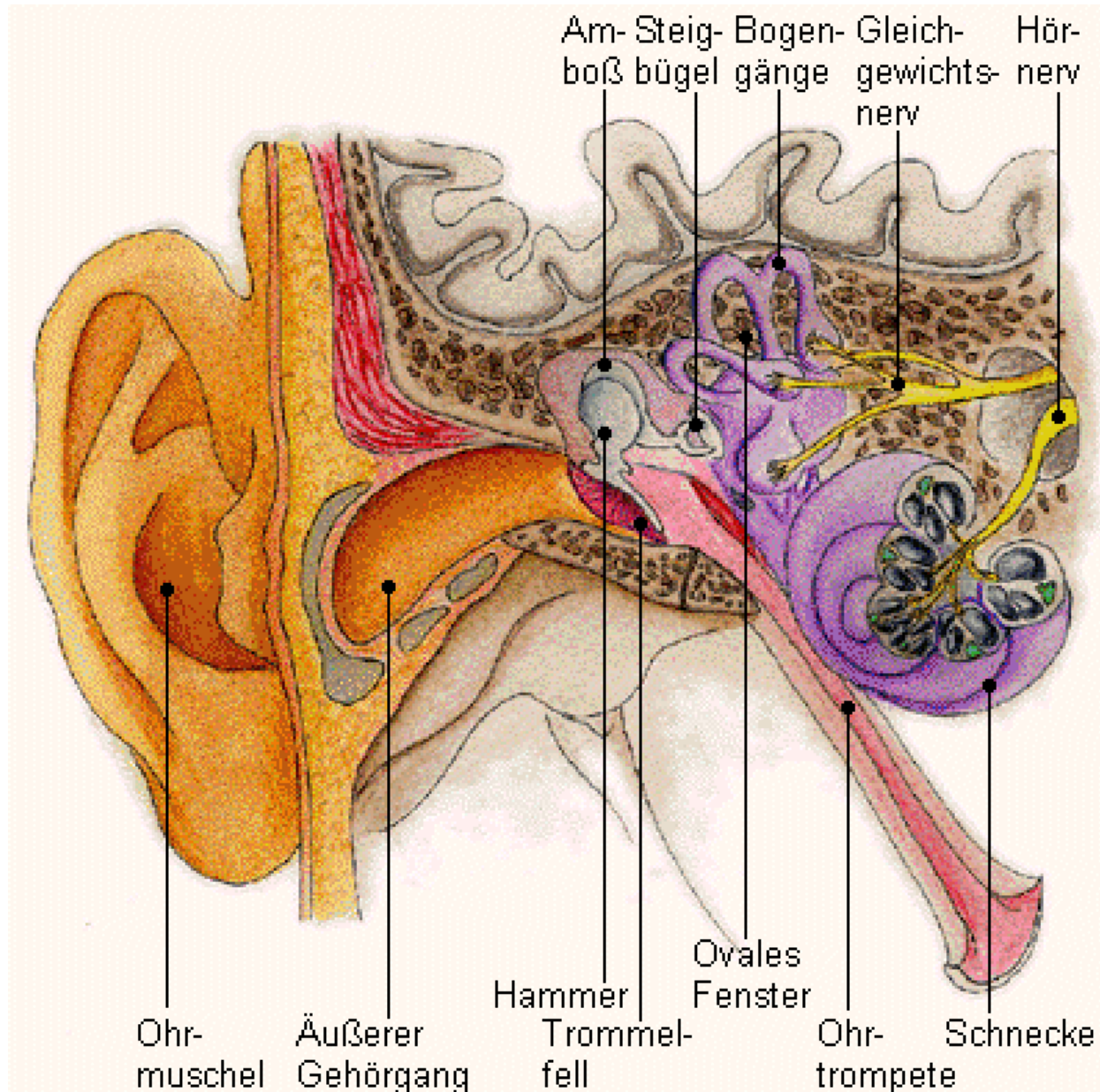
6 dB doppelter Schalldruck

-6 dB halber Schalldruck

$$20 * \log_{10}(2) = 6.0206 \text{ dB}$$

Quelle	
Sprache	10 μ W
Geige	1 mW
Lautsprecher	100 W

Ohr: Schema



drei Untereinheiten:

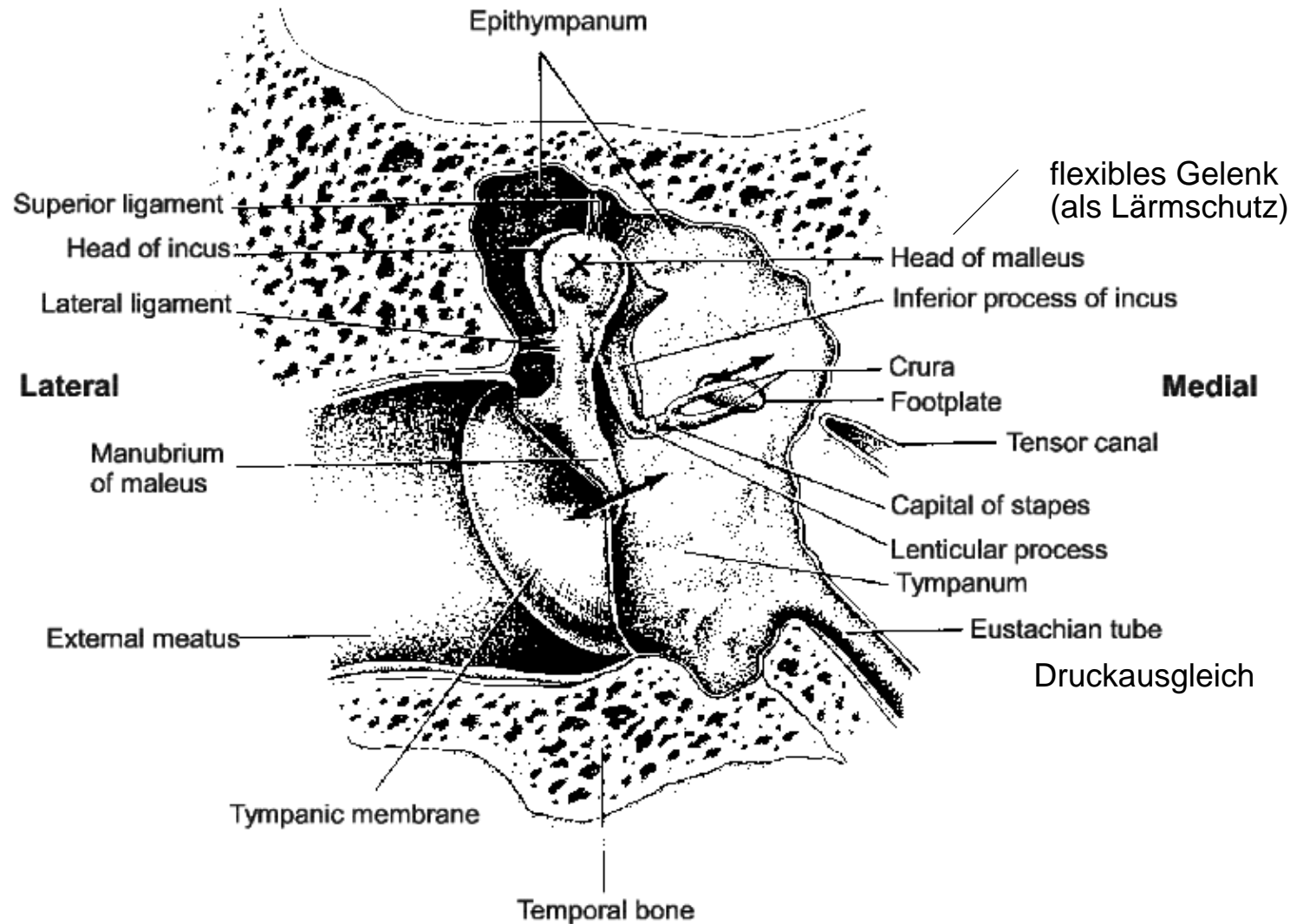
- Außenohr
- Mittelohr
- Innenohr

bzw.:

- Verstärkung, Delay
- Impedanzwandlung
- Frequenzanalyse

(glasklar Mensch 3D)

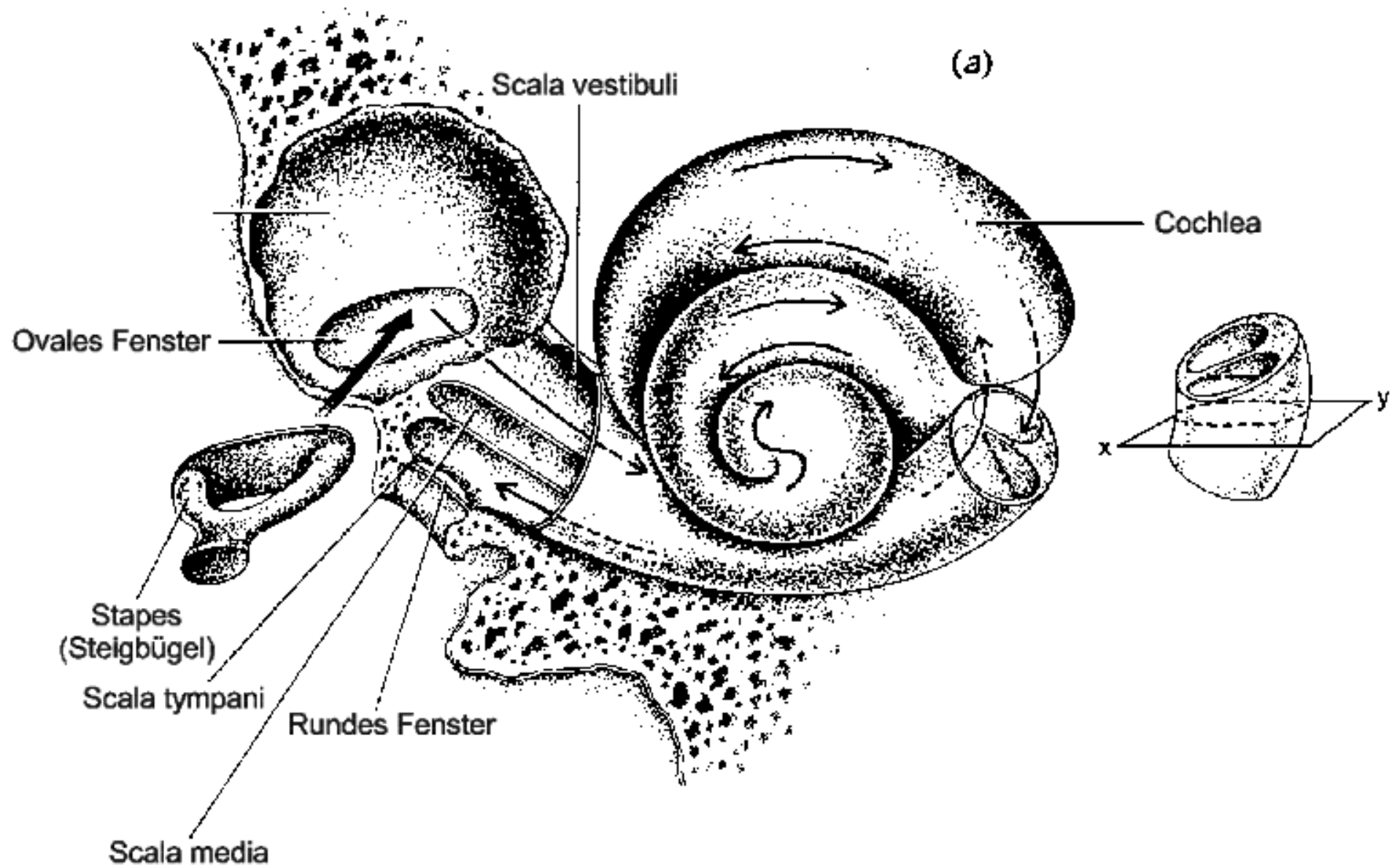
Mittelohr



- Verstärkung (Impedanzwandlung), Übergang Luft nach Wasser
- Trommelfeld ca. 20fache Fläche des ovalen Fensters

(Kollmeier)

Innenohr: Aufbau und Cochlea



(Kollmeier)

Innenohr: cortisches Organ

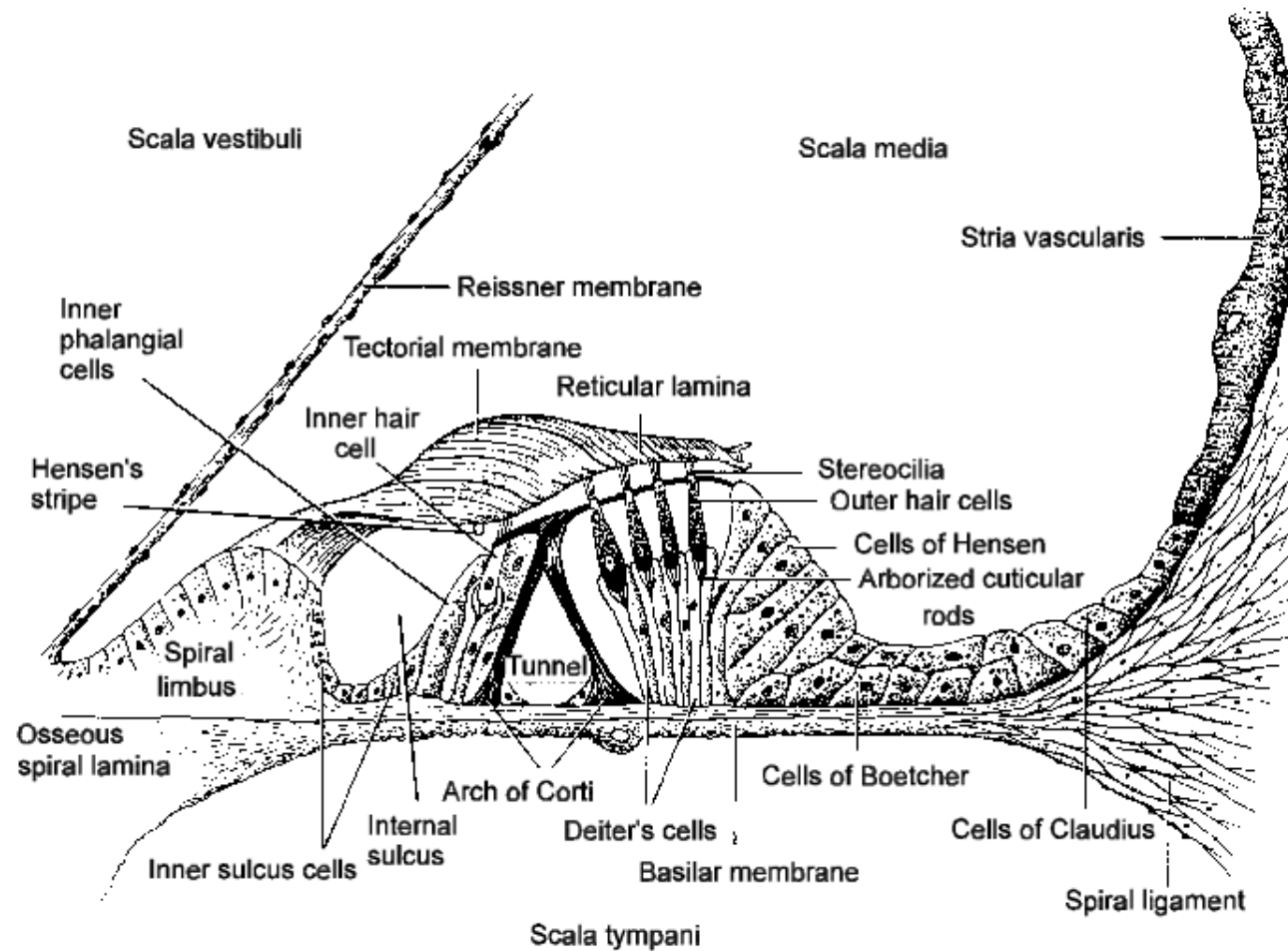
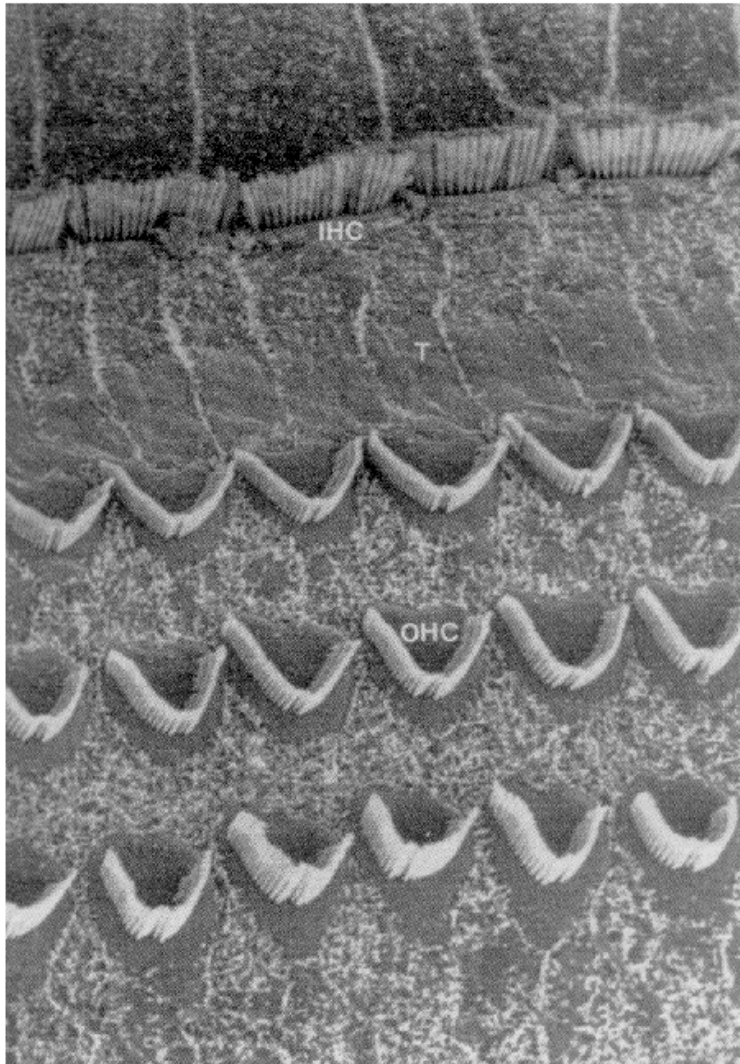


Abbildung 3.7: Schematische Darstellung des Cortischen Organs (aus Gulick, W. L., Gescheider, G. A., Frisina, R. D.: Hearing. Oxford University Press, Oxford 1989)

(Kollmeier)

Innenohr: Haarzellen



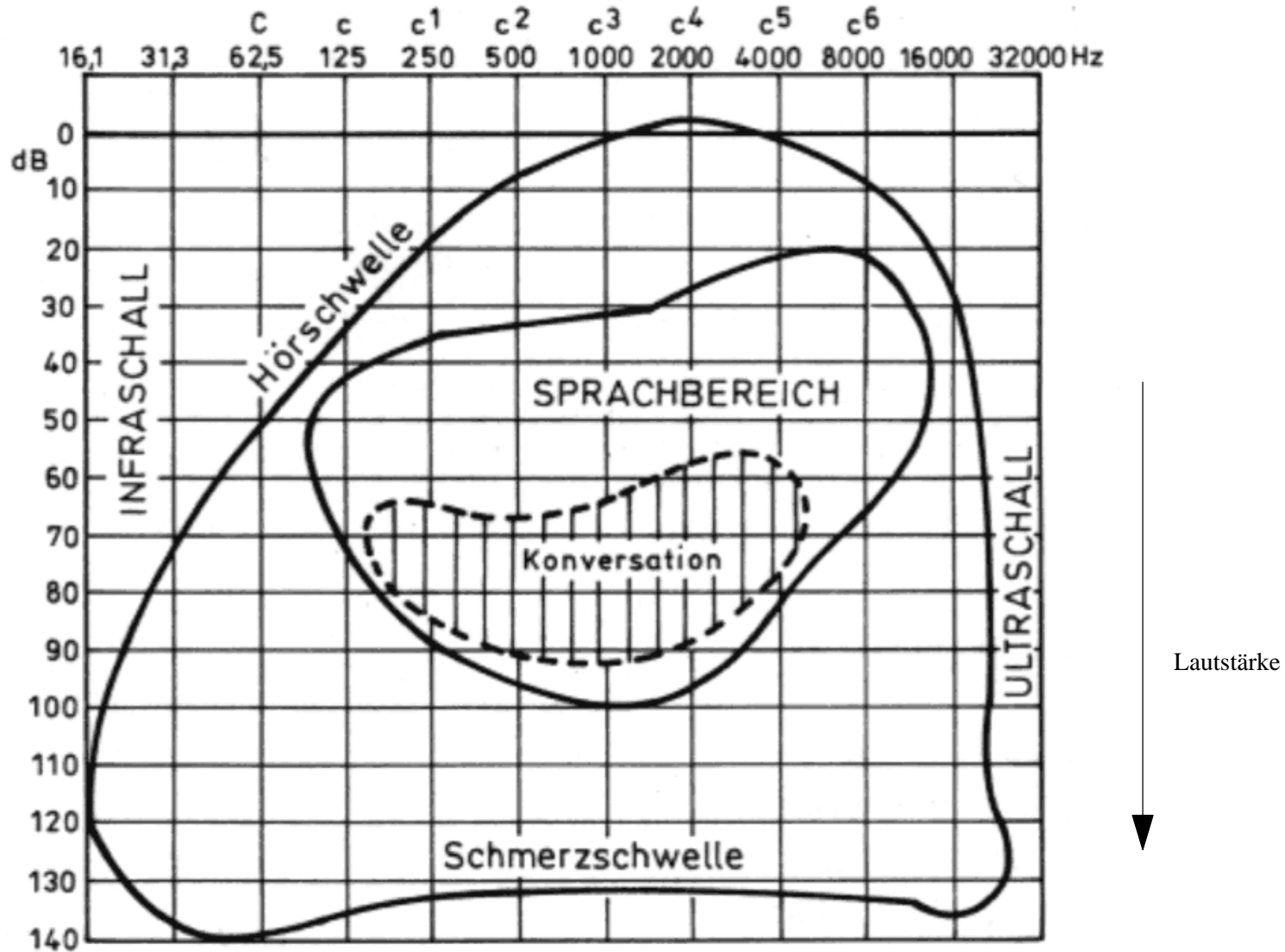
innere Haarzellen
"Sensoren"

äußere Haarzellen
"Aktoren"

Abbildung 3.8: Detailansicht der Innenohr-Anatomie, IHC = Innere Haarzellen, OHC = Äußere Haarzellen (aus Dallos, P., Popper, A. N., Fay, R. R. (Hrsg): The Cochlea. Springer Verlag, New York 1996)

(Kollmeier)

Hörbereich und Sprachbereich



gute Sprachverständlichkeit im "Hauptsprachbereich" (ca. 300-3500 Hz)

(c't 10/89)

absolute Hörschwelle

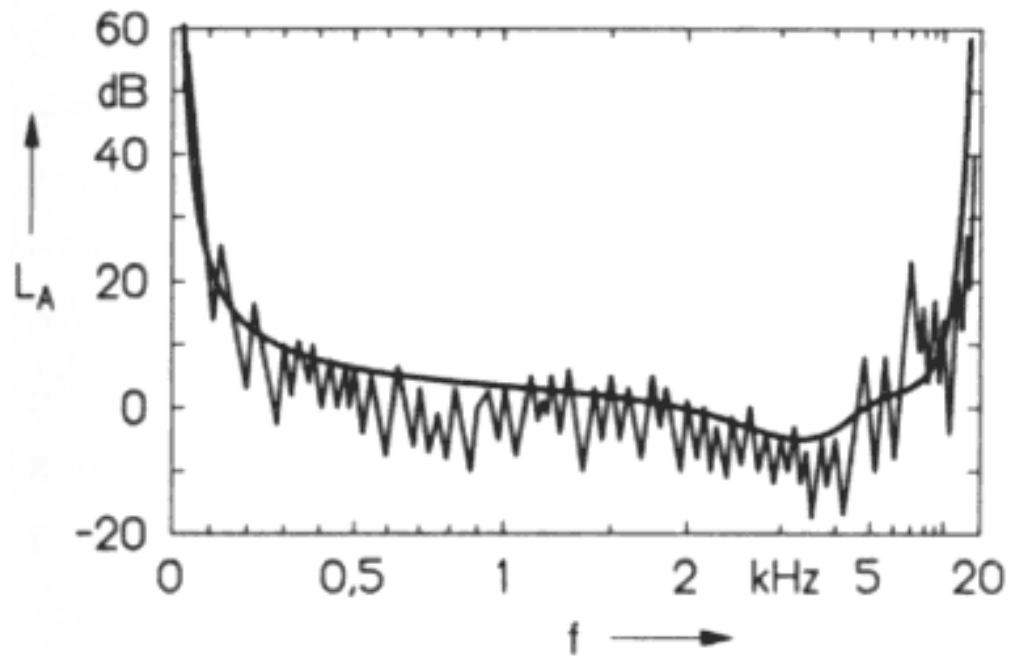


Abb. 9.11. Monaurale Absoluthörschwelle für Sinustöne. Glatte Kurve: Mittelwert des Schwellenpegels L_A von normalhörenden Personen. Gezackte Kurve: Mit der Methode des pendelnden Regels (Békésy tracking) gemessen (25jähriger Mann, rechtes Ohr). Abszisse: Tonfrequenz, SPINC-skaliert; Ordinate: Schallpegel

- geringe Empfindlichkeit bei tiefen und hohen Tönen
 - maximale Empfindlichkeit zwischen 2..5 KHz
- => Loudness-Kurven
- Demo

(Terhardt)

Lautheit

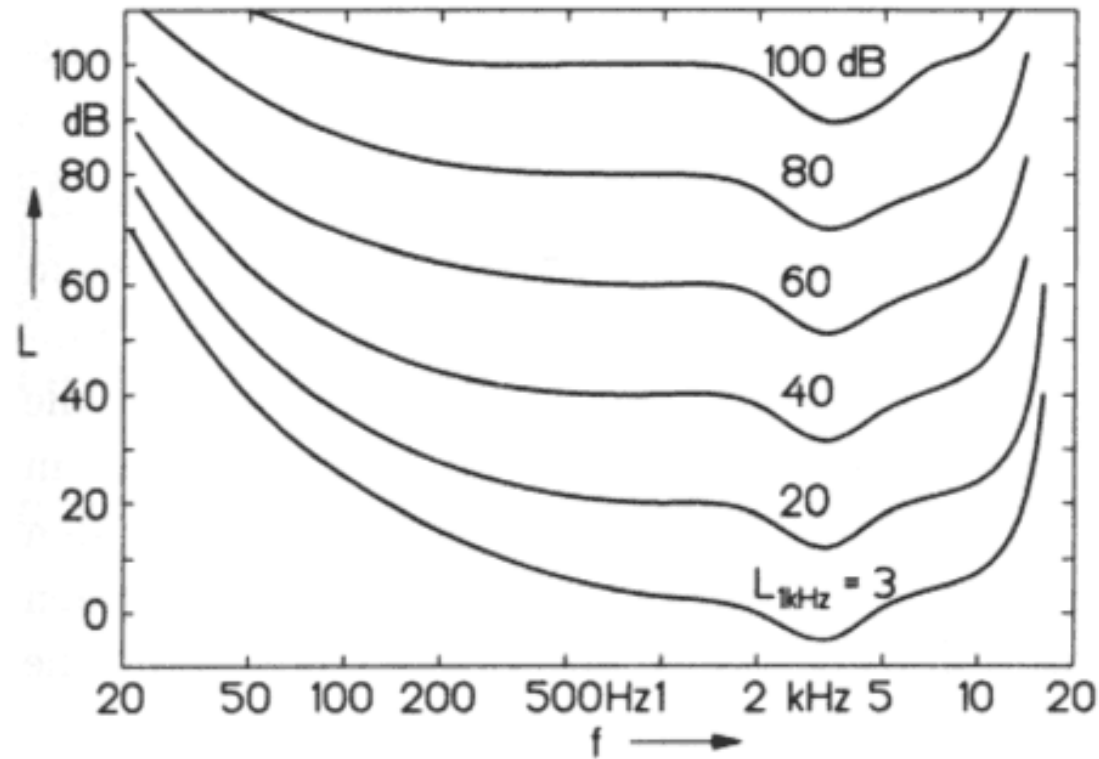


Abb. 10.8. Kurven gleicher Lautheit von Sinustönen im ebenen Schallfeld: Schallpegel L eines Sinustones (Ordinate), welcher die gleiche Lautheit hervorruft wie ein 1 kHz-Ton mit dem angegebenen Pegel, in Abhängigkeit von der Tonfrequenz f (Abszisse)

- bei leisen Tönen schwache Wahrnehmung von Höhen/Bässen
- Empfindlichkeitsmaximum bei ca. 3 KHz
- "loudness"-Regler am Hifi-Verstärker zur Kompensation

(Terhardt)

"Tuningkurven":

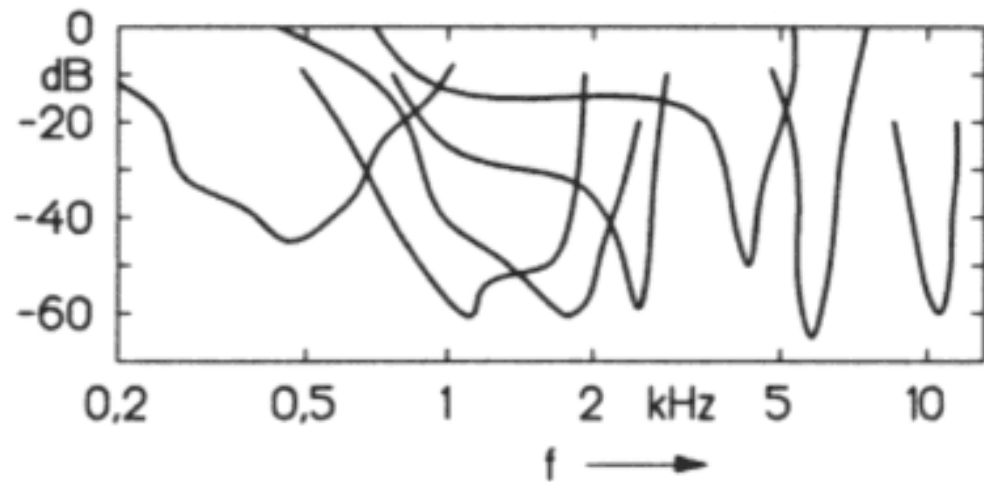


Abb. 9.10. Tuningkurven, gemessen an anästhesierten Katzen. Jede Kurve gehört zu einer anderen Faser des akustischen Nerven. Die Ordinate gibt denjenigen Schallpegel eines Sinustones an, welcher zur Erregung einer festen Anzahl von Aktionspotentialen pro Zeiteinheit erforderlich ist. Abszisse: Frequenz des Sinustones. Nach Kiang [501]

- hohe Güte der Frequenzselektion
- "aktive" Frequenzbereiche der Fasern überlappen
- menschliches Gehör entsprechend

(Terhardt)

Maskierung

Mithörschwelle eines Sinustones bei Verdeckung durch Schmalbandrauschen. In Abb.11.4 sind drei Mithörschwellen von Sinustönen verdeckt durch Schmalbandrauschen der angegebenen Mittenfrequenzen dargestellt [1125]. Die Fourier-Spektren der maskierenden Schmalbandgeräusche sind durch Verwendung von Bandpaßfiltern hoher Flankensteilheit auf Bandbreiten von ungefähr 100, 160 und 700 Hz beschränkt [CD 2].

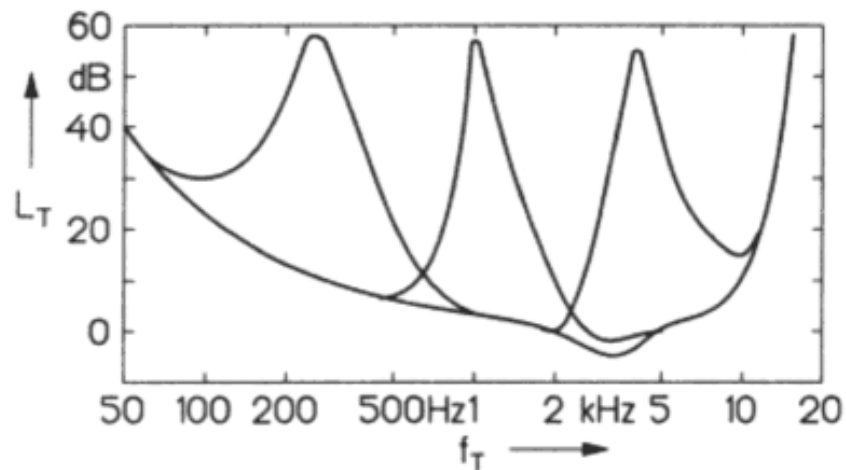
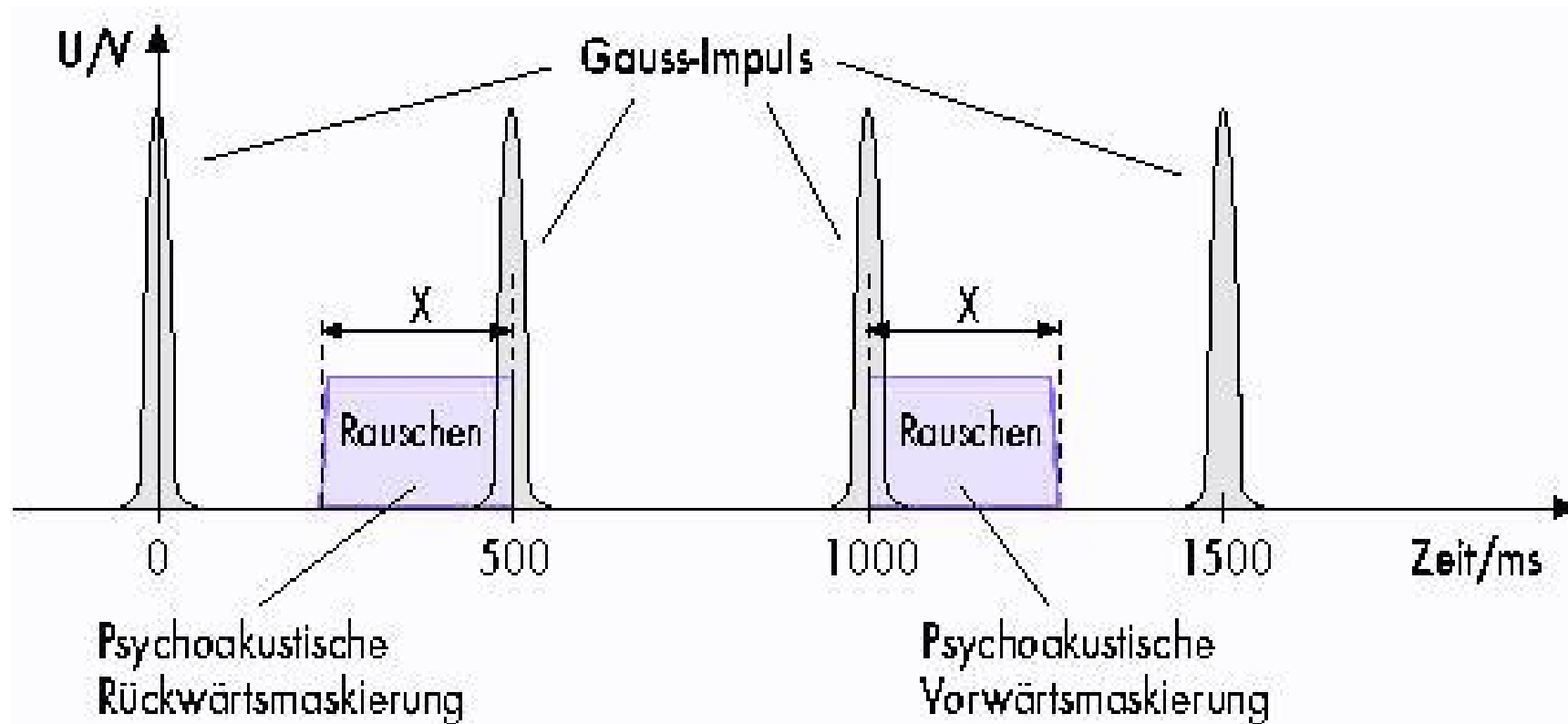


Abb. 11.4. Mithörschwellen von Sinustönen bei Verdeckung durch frequenzgruppenbreite Schmalbandrauschen mit den Mittenfrequenzen 250, 1000, 4000 Hz und dem Maskierpegel $L_M = 60$ dB. Nach [1125]

(Terhardt)

- Frequenzanalyse in der Cochlea nicht perfekt
- laute Töne verdecken benachbarte leisere
- Ausnutzung in psychoakustischen Modellen (MiniDisc, MP3, usw)

Maskierung: temporal



- laute Töne verdecken spätere leisere Töne
- und frühere leisere Töne (!!)
- Ausnutzung in psychoakustischen Modellen (MiniDisc, MP3, usw)

Frequenzgruppen

Tabelle 9.1: Frequenzgruppen nach Zwicker 1982

z /Bark	f_u /Hz	f_o /Hz	Δf_G /Hz	f_m /Hz
0	0	100	100	50
1	100	200	100	150
2	200	300	100	250
3	300	400	100	350
4	400	510	110	450
5	510	630	120	570
6	630	770	140	700
7	770	920	150	840
8	920	1080	160	1000
9	1080	1270	190	1170
10	1270	1480	210	1370
11	1480	1720	240	1600
12	1720	2000	280	1850
13	2000	2320	320	2150
14	2320	2700	380	2500
15	2700	3150	450	2900
16	3150	3700	550	3400
17	3700	4400	700	4000
18	4400	5300	900	4800
19	5300	6400	1100	5800
20	6400	7700	1300	7000
21	7700	9500	1800	8500
22	9500	12000	2500	10500
23	12000	15500	3500	13500
24	15500			

Tonhöhenwahrnehmung?

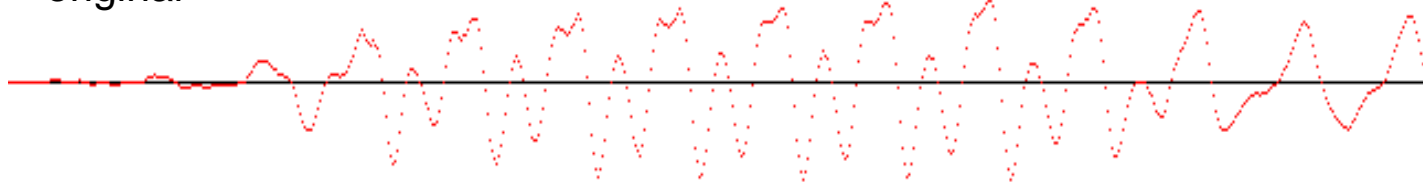
- Messungen
 - vgl. Tuningkurven

 - tiefe Frequenzen:
ca. 100 Hz Bandbreite
 - hohe Frequenzen:
log. Breite der Gruppen
- => "Bark" Frequenzskala
- Ausgangspunkt für
psychoakustische Modelle

(Tabelle aus Zoelzer)

Phasen: Hilbertfilter

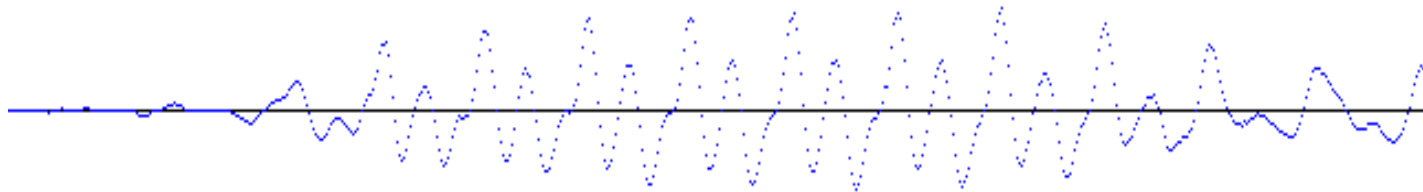
original



screenshot original / filtered signal

Sprachsample "Oberseminar"

phasengedreht



- Ohr unempfindlich gegen Phaseninformation
- verändertes Signal klingt einzeln (vollkommen) gleich
- erst bei Interferenzen mehrerer Signale hörbar
- Phasengang vieler Systeme (Telefon!) extrem schlecht
- Demo

(vdHeide)

Klangfarben

akustisches Signal mit vielen Teiltönen:

- keine Wahrnehmung aller Einzeltöne
- sondern Zuordnung zu Quellen/Instrumenten
- tiefste Frequenz bestimmt die wahrgenommene Tonhöhe
- Obertonspektrum identifiziert die Quelle
- mit Eigenschaft "Klangfarbe"

- Klangfarbe auch stark vom "attack" des Klangs abhängig
(=> vgl. Synthesizer mit kurzen Samples plus Schleifen)

- entsprechend automatische "Sprecheridentifikation"
- komplexe Zusammenhänge für Akkordwahrnehmung

Klangfarben: Spektrogramm

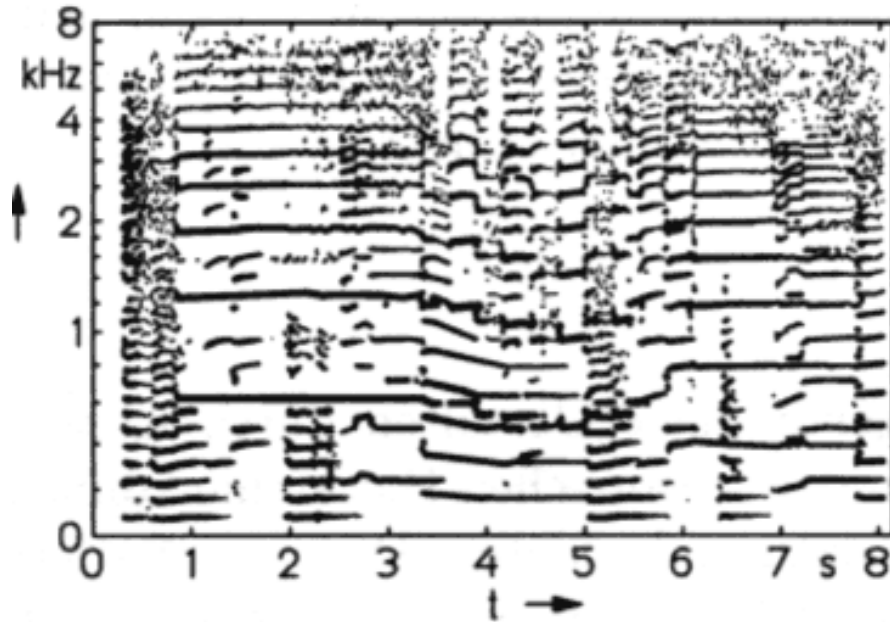


Abb. 1.9. Beispiel eines Teiltonzeitmusters. Ausschnitt einer mehrstimmigen Blasmusikdarbietung. Die hauptsächlich beteiligten Instrumente sind Trompete, Posaune, Klarinette und Tuba. Die instantane Teilton-Amplitude wird durch die Strichstärke angedeutet. Ordinate Bark-skaliert. Nach [969]



Abb. 1.10. Transkription der Trompeten- und der Tuba-Stimme des Musikbeispiels, dessen Teiltonzeitmuster in Abb. 1.9 dargestellt ist

(Terhardt)

Klangfarben: Beispiele

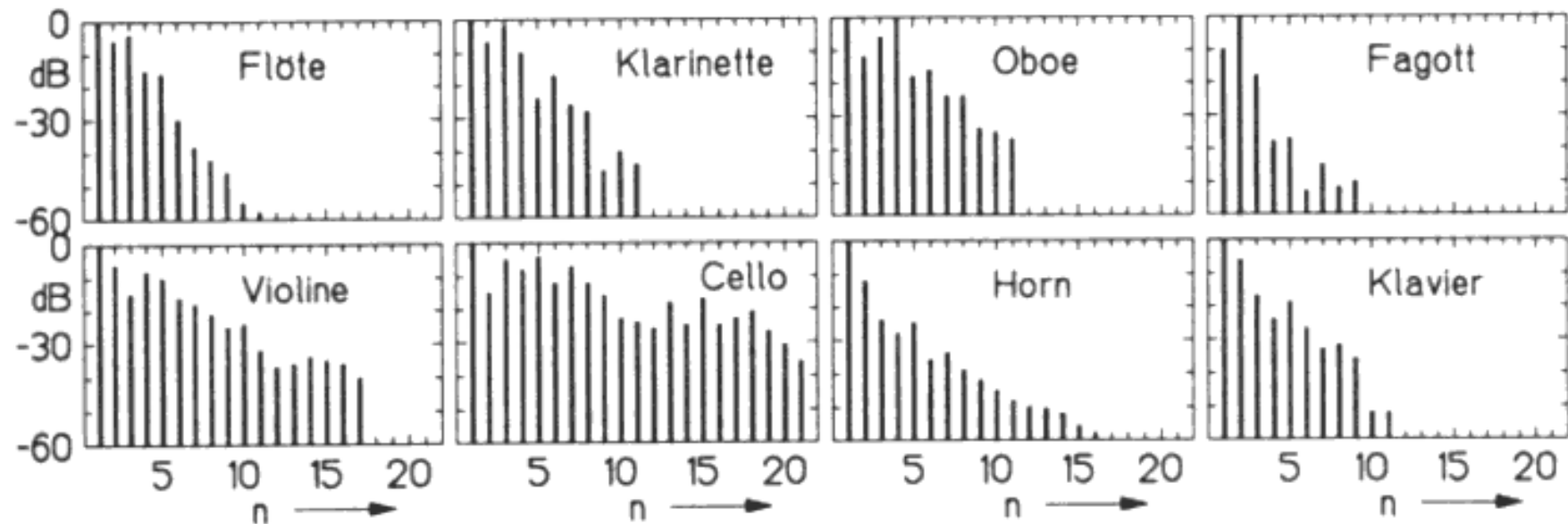


Abb. 8.12. Teiltonspektren verschiedener Musikinstrumente schematisch

(Terhardt)

- Streicher mit reichem Obertonspektrum
- Flöten mit sehr reinen Tönen (insb. Blockflöte)
- Klavier stark anschlag- und tonhöhenabhängig
- aber auch "Attack-Phase" maßgeblich für Klangwahrnehmung

Beispiel: "Akustischer Bass"

akustisches Signal mit vielen Teiltönen:

- keine Wahrnehmung aller Einzeltöne
- sondern Zuordnung zu Quellen/Instrumenten
- komplexe Zusammenhänge für Akkordwahrnehmung

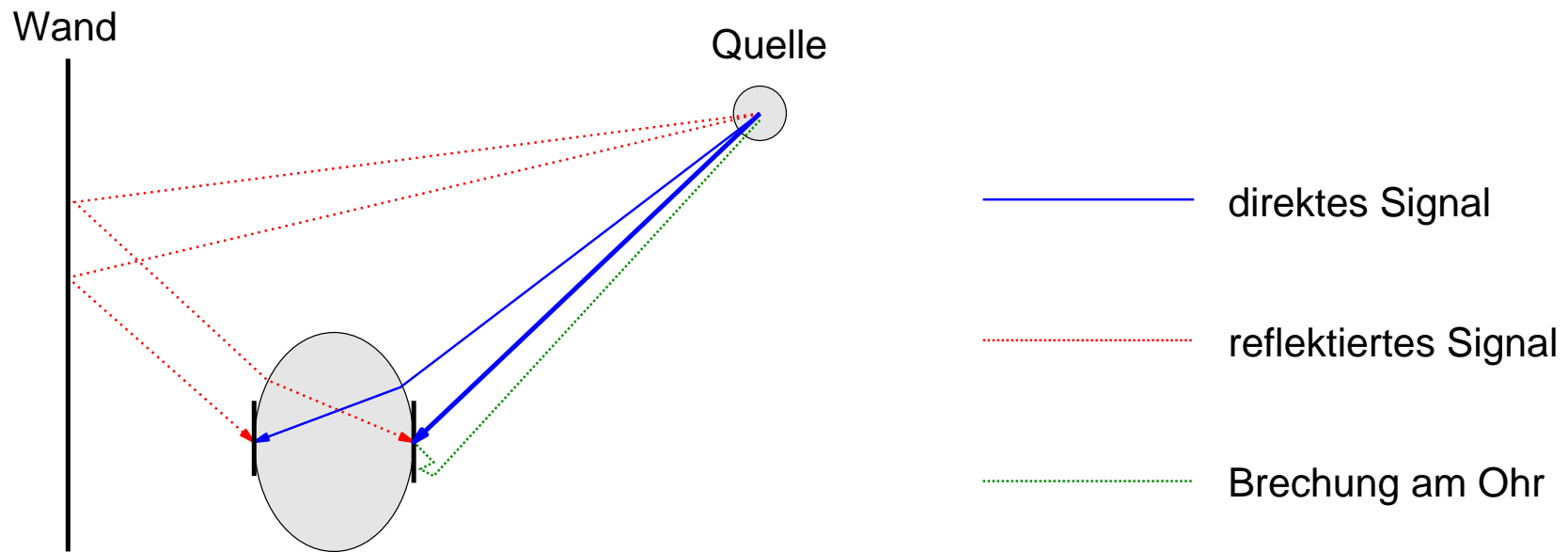
$$\sin(x) + \sin(y) = 2 * \sin[(x+y)/2] * \cos[(x-y)/2]$$

=> "akustischer Bass":

$$\sin(x) + \sin(3/2 x) = 2 * \sin[5/4 x] * \cos[x/4]$$

- Kirchenorgeln (seit ca. 1700): Grundton+Quinte = Suboktave
- ähnliche Tricks für "bass-booster" & Co.
- Demo

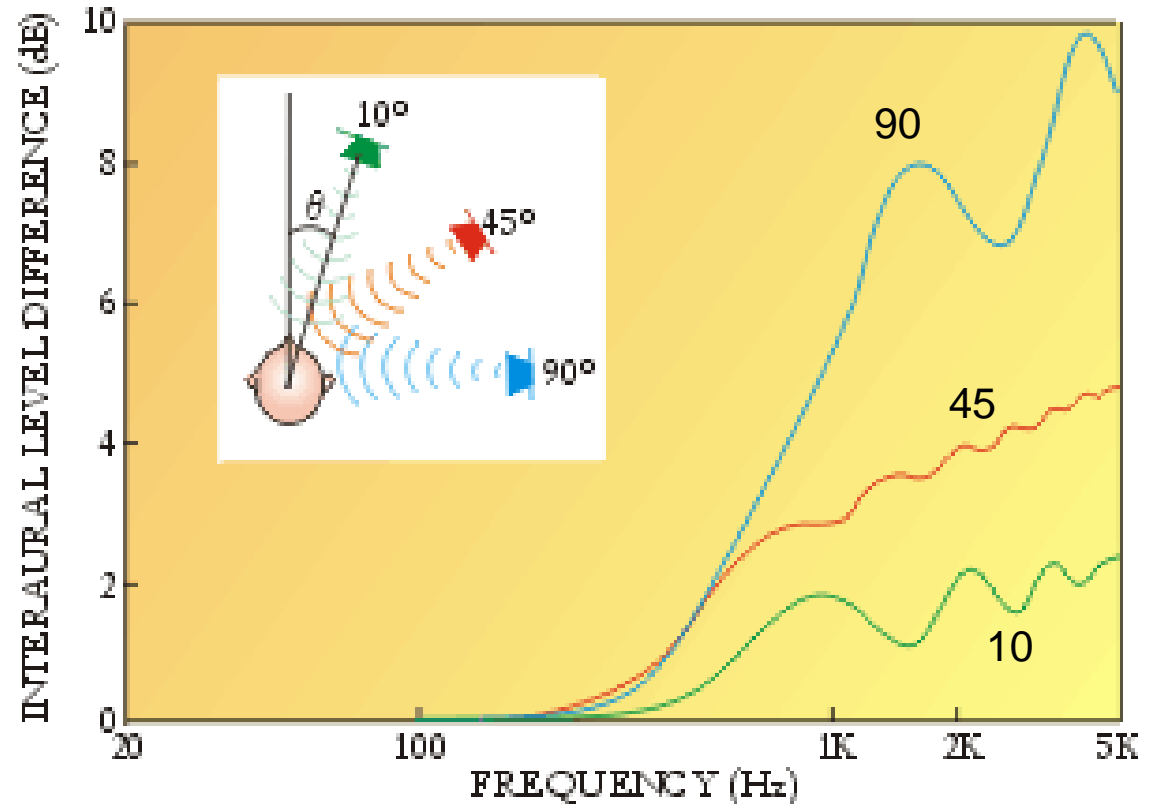
Akustische Raumwahrnehmung



- direktes Signal
- reflektierte Signale
- jeweils mit Beugung um den Kopf herum
- jeweils mit Brechung an Ohr, Schultern, ...
- zusätzlich "obstruction" und "occlusion" für verdeckte Quellen ...

ILD: Lautstärkedifferenz

interaural level difference:



(Physics Today Nov. 99)

- Welle kaum gestört, wenn Wellenlänge > Kopf
- daher kaum Effekt bei niedrigen Frequenzen
- Bässe schlecht ortbar

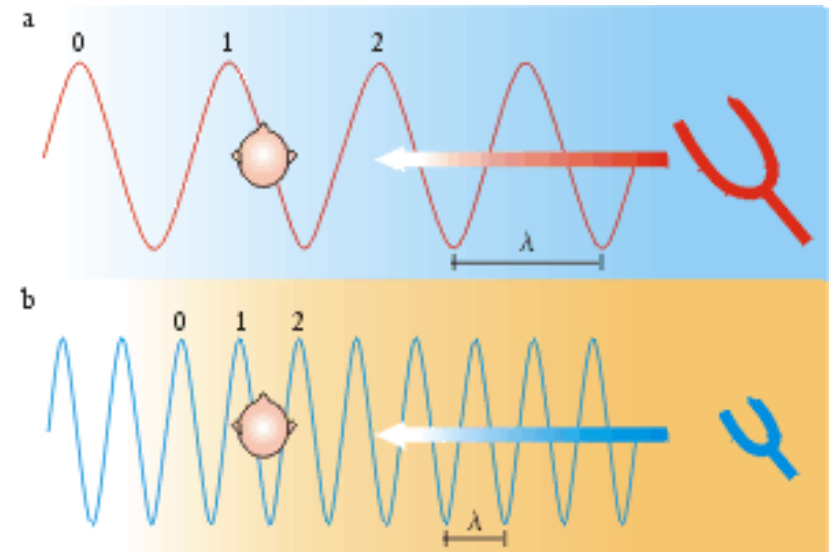
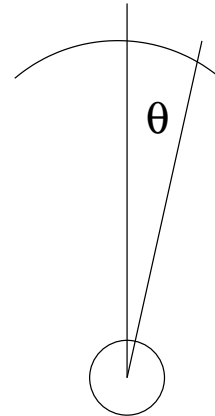
ITD: Zeit- und Phasendifferenz

interaural time difference:

$$dt = \frac{3a}{c} \sin \theta$$

a = Kopfradius ~ 8.75 cm

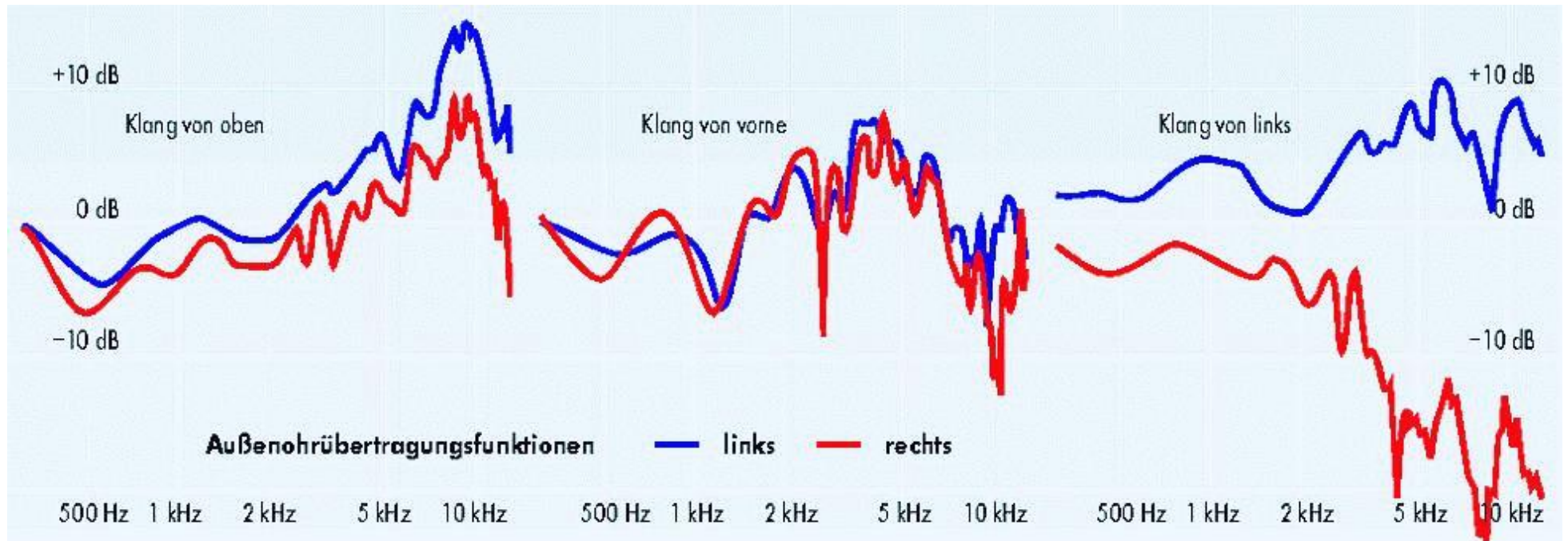
c = Schallgeschw ~ 340 m/s



(Physics Today Nov. 99)

- 500 Hz Sinuston von vorne ortbar mit ca. 1 Grad Auflösung
=> ca. 13 ms Zeitauflösung
- mehrdeutig, wenn $n \cdot \text{Wellenlänge} \approx \text{Kopfdurchmesser}$
- Ohr unempfindlich für Phaseninformation oberhalb ca. 1 KHz
- vgl. Hilbertfilter-Experiment

HRTF



(ct 04/97 160)

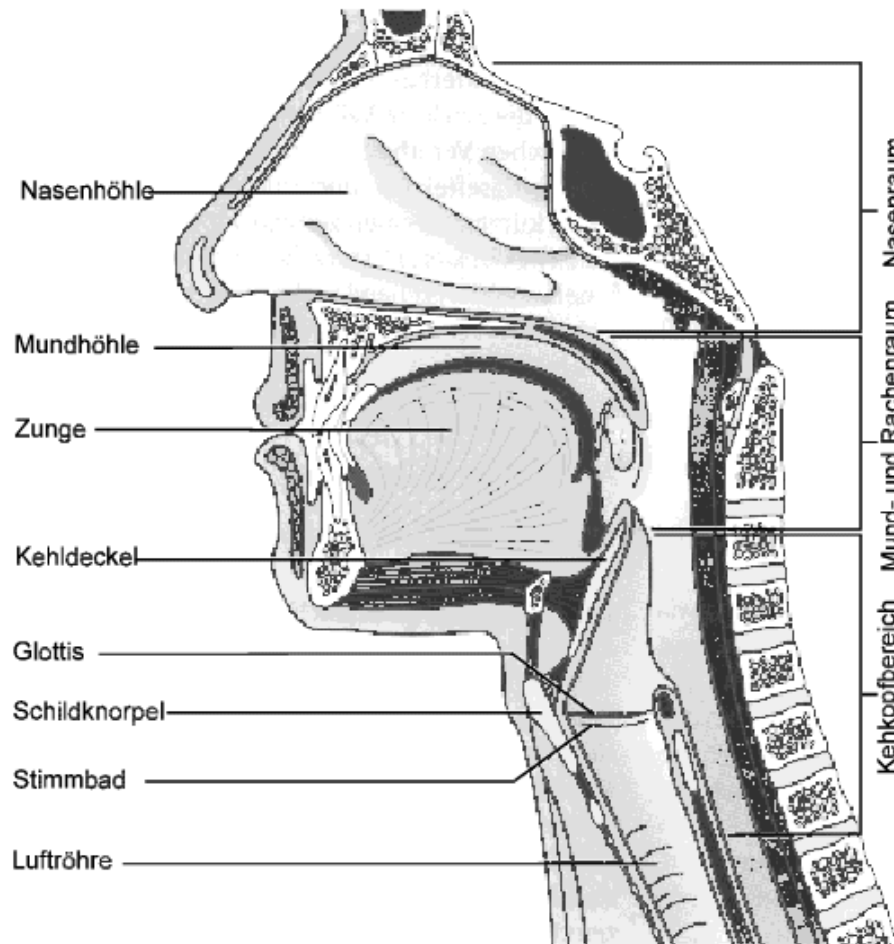
"head related transfer function": Außenohr/Kopf-Übertragungsfunktion

- Einfluß von Kopf, Schultern, Außenohr auf Schallsignale
- sehr komplexe Winkel- und Frequenzabhängigkeit
- Verbesserung der Ortung durch (kleine) Kopfbewegungen
- Demos später (Direct X)

Audio-Demos:

Denon, "Audio Technical CD" (1984)	Tracks:
• Channel-Check, Phase-Check	1,2
• Hörbeispiele	3, ..., 9
• 1 kHz Sinus, Lautstärke -15, -16, -18, -20 dB	13,14,15,16
• Sinustöne: 40 Hz .. 20 KHz	46, ..., 55
• Frequenz-Sweep, 5 Hz .. 22.5 kHz	65
• weißes und rosa Rauschen	79, 90
• Quantisierung: Lautstärke 0, -20, -40, -60 dB	96, 95, 94, 93
Terhardt, "Akustische Kommunikation"	
• Mithörschwelle	1,2
• Frequenzgruppen, Tonhöhen	3,4
• Maskierung, Residualtöne	22
• Akustischer Bass	27

Sprache: Vokaltrakt



- Glottis, Stimmband
- Klang zusätzlich durch Zungen-/Lippenstellung
- typische Spektren:
Mann / Frau / Kind
normal / erregt

Abbildung 6.1: Vokaltrakt im Querschnitt (schematisch) (aus Kießling, J., Kollmeier, B., Diller, G.: Versorgung und Rehabilitation mit Hörgeräten. Thieme Verlag, Stuttgart 1997)

(Kollmeier)

Sprache: Spektrogramm

Formanten, d. h. die Frequenzbereiche, bei denen der Vokaltrakt eine besonders hohe Verstärkung des akustischen Signals bewirkt. Die Formanten stellen sich damit als Spitzen im Spektrum dar und sind charakteristisch für den jeweils artikulierten Vokal (s. unten).

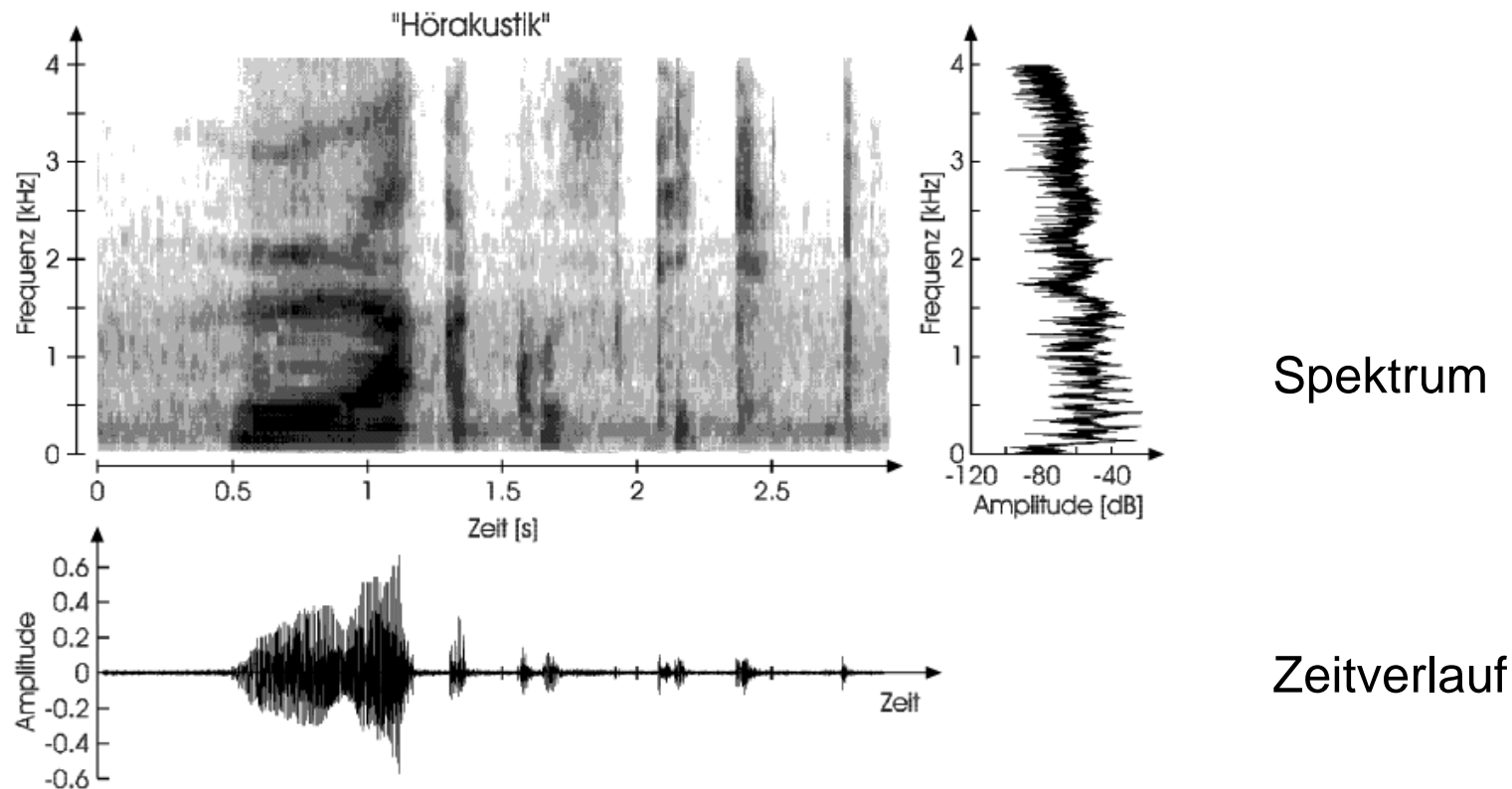


Abbildung 6.2: Zeit-Signal (unten), Spektrum (rechts) und Spektrogramm (links oben) des Wortes „Hörakustik“

(Kollmeier)

Sprache: Artikulation

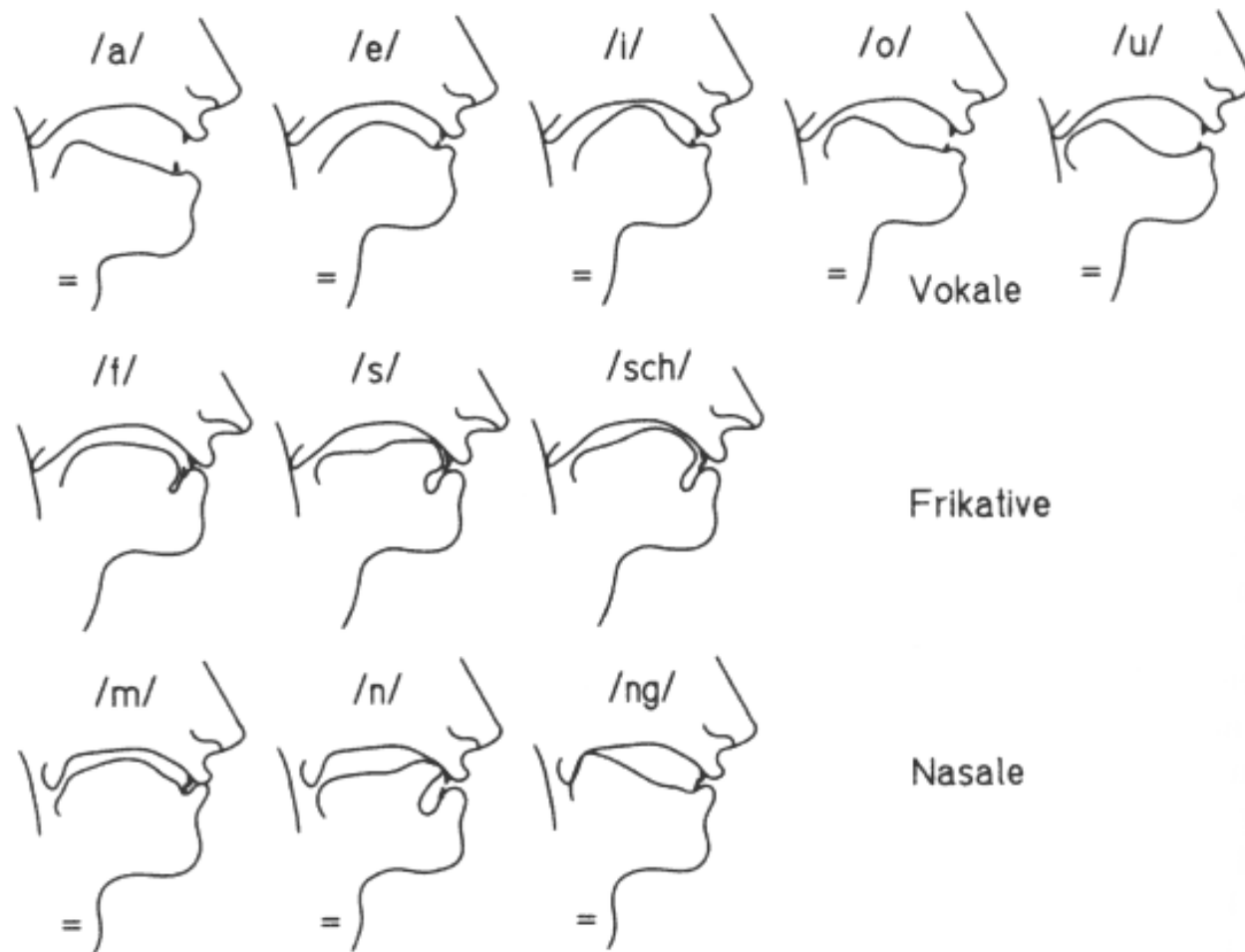
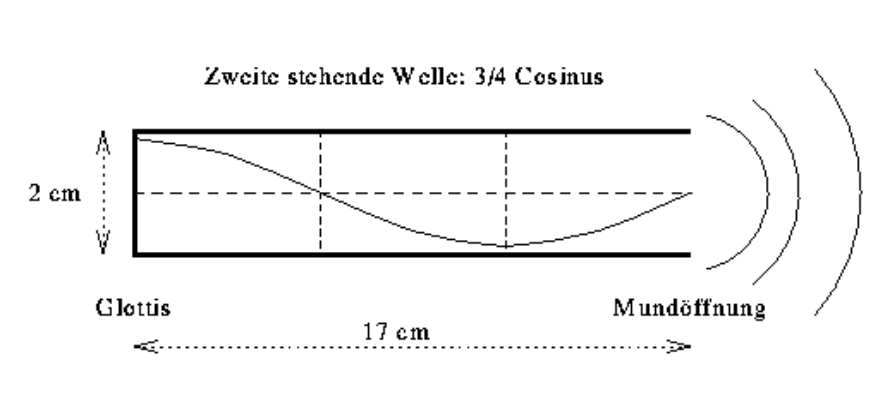
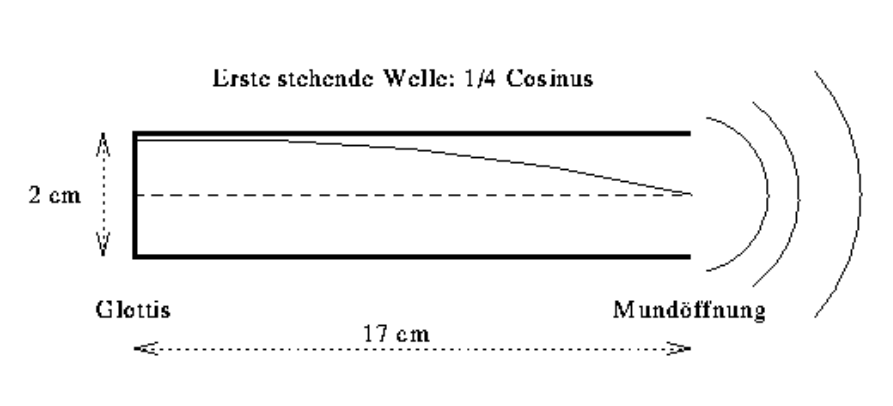
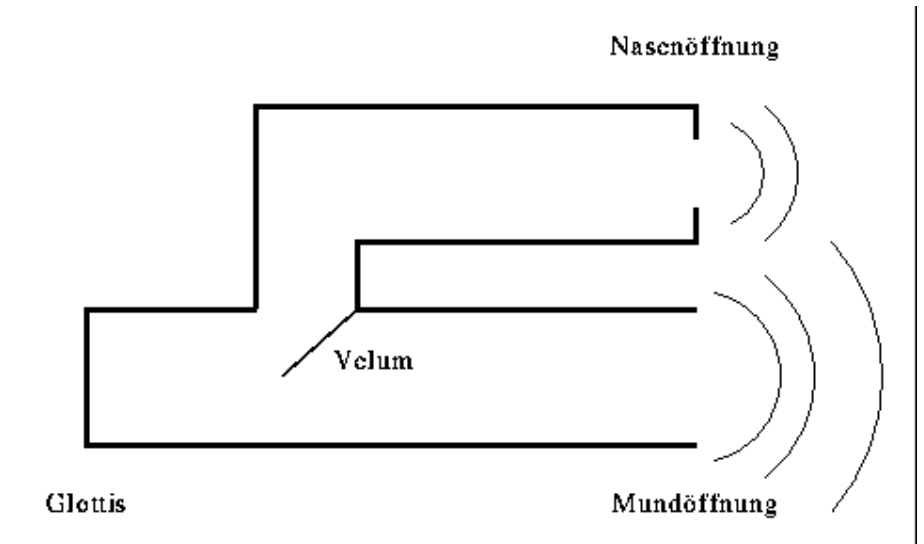
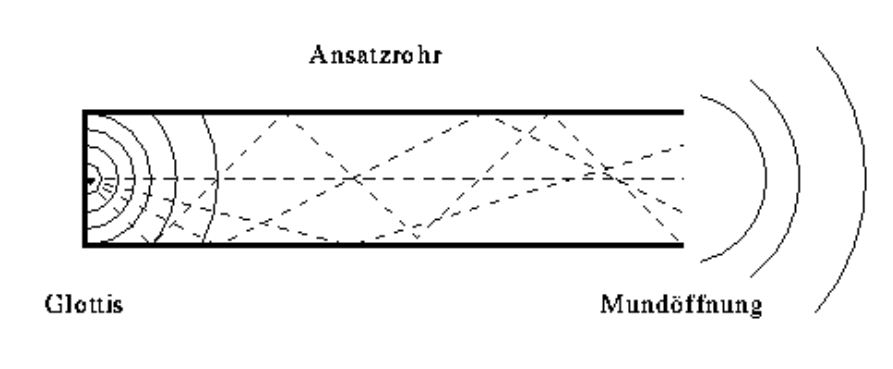


Abb. 7.1. Schematische Darstellung der Artikulation einiger Sprachlaute

- Zungen- und Lippenstellung maßgeblich

(Terhardt)

Sprache: Vokaltrakt



- Stimmband + Resonanzrohr
- Lage der Formanten:
Wellenlänge: $L/4 = 17\text{cm}$

$$F1 = c/L = 340\text{m/s} / 68\text{cm} = 500 \text{ Hz}$$

$$F2 = 1500 \text{ Hz}$$

$$F3 = 2500 \text{ Hz}$$

$$F3 = 3500 \text{ Hz}$$

Sprache: Vokaldreieck

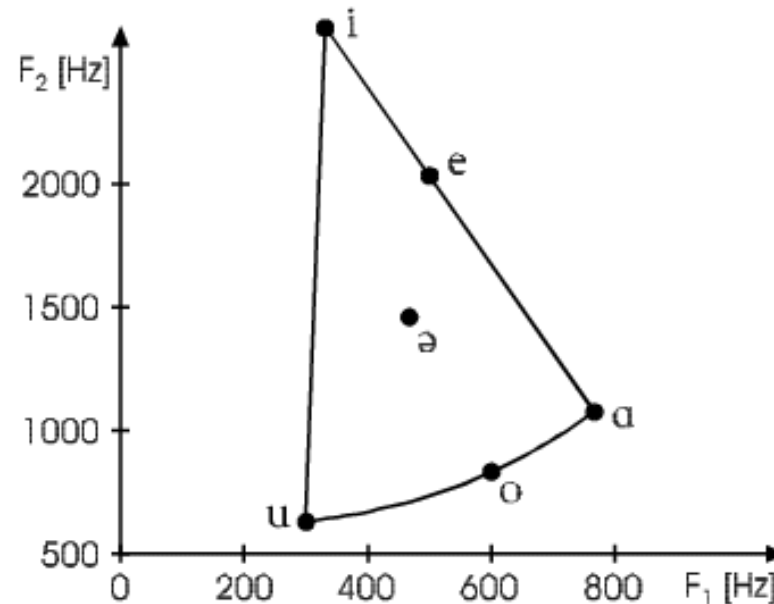


Abbildung 6.1.4: Vokaldreieck. Aufgetragen ist die Frequenz des ersten Formanten F_1 (Abzisse) und des zweiten Formanten F_2 (Ordinate) für unterschiedliche Vokale (schematisch)

(Kollmeier)

- typische Frequenzen / Formanten der Vokale
- verhältnismäßig einfache Erkennung von stimmhaften Lauten

Gehör: Zusammenfassung

Gehör außerordentlich leistungsfähig:

- überdeckt 13 Größenordnungen Schalldruck
- fast optimale Empfindlichkeit
- Hörbereich etwa 20Hz .. 20 KHz
- Frequenzanalyse "mechanisch" in der Cochlea
- komplexe Verarbeitung/Erkennung im Gehirn
- Wahrnehmung Klang vs. Geräusch

insbesondere zur Raumwahrnehmung:

- Zeit- und Intensitätsdifferenzen der Signale
- Kopfübertragungsfunktion
- Auswertung von Echos

=> entsprechende Modelle für effiziente Signalverarbeitung

Digitale Signalverarbeitung

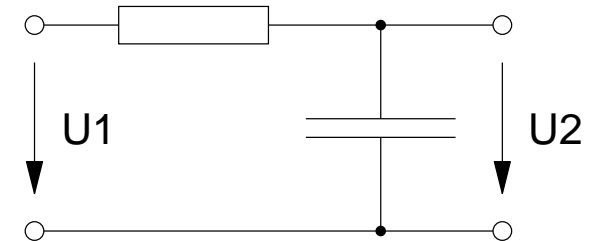
- Motivation
- analoge vs. digitale Signalverarbeitung
- Zahlenfolgen, LTI-Systeme
- Abtasttheorem
- Quantisierung, AD/DA Konverter, Dithering

- Beispiele für Audio-Algorithmen
- digitale Filter

DSP: Motivation

analoge Signalverarbeitung:

- Modellierung mit Differentialgleichungen
- direkte Umsetzung mit diskreten Bauteilen
- hauptsächlich mit "analoger" Elektronik
- Widerstände, Kondensatoren, Spulen, Transistoren



aber:

- entweder geringe Genauigkeit (z.B. 10%) oder sehr teuer
 - Alterung, Exemplarstreuungen, aufwendiger Abgleich
 - wirtschaftlich nur für geringe Anzahl der Bauelemente
 - Datenspeicherung (analog) problematisch
- => komplexe Algorithmen nicht umsetzbar
=> digitale Verarbeitung verspricht Abhilfe

DSP: Definition

"digital signal processing" (digitale Signalverarbeitung):

- | | |
|-----------------------------------------------------------|------------------------|
| Verbeitung von | für: |
| • zeitdiskreten | Nachrichtenübertragung |
| • wertdiskreten | Bildverarbeitung |
| • Zahlenfolgen | Audioverarbeitung |
| auf Digitalrechnern | usw. |
| • als Ersatz analoger Verfahren | |
| • Differenzen- statt Differentialgleichungen | |
| • oft mit Echtzeitanforderungen | |
| • im Prinzip beliebige Genauigkeit | |
| • oft mit Spezialhardware, etwa Signalprozessoren / ASICs | |

DSP: Eigenschaften

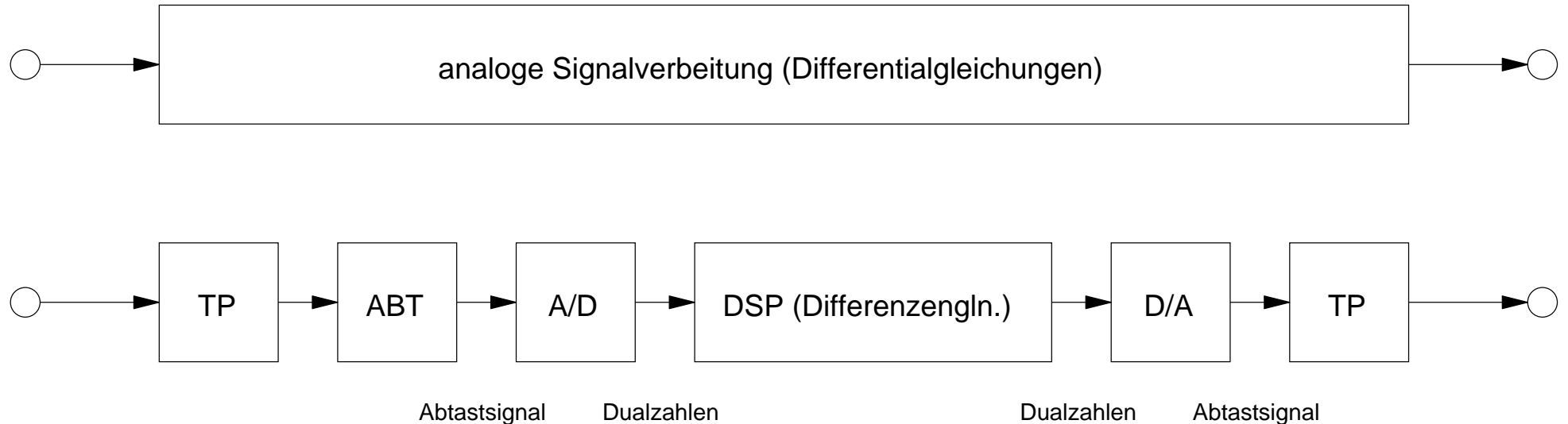
Verarbeitung zeit- und wertdiskreter Zahlenfolgen:

- als Ersatz für zeit/wert-kontinuierliche Verarbeitung
- wenn Abtasttheorem erfüllt (genügend hohe Abtastrate)
- beliebige Genauigkeit (Wortbreite anpassen)
- optimale Stabilität (z.B. keine Temperaturabhängigkeit)
- kein Abgleich notwendig (z.B. Nullpunktseinstellung)
- keine Alterung, keine Toleranzen

- auch für sehr komplexe Algorithmen
- die analog nicht (wirtschaftlich) realisierbar sind
- insbesondere: Datenspeicherung

- VLSI-Technologie erlaubt sehr hohe Performance

DSP: vs. analoge Verarbeitung



"analoge"

Signalverarbeitung

"digitale"

- zeitkontinuierlich
- wertkontinuierlich
- Differentialgleichungen
- Drift, Rauschen, ...

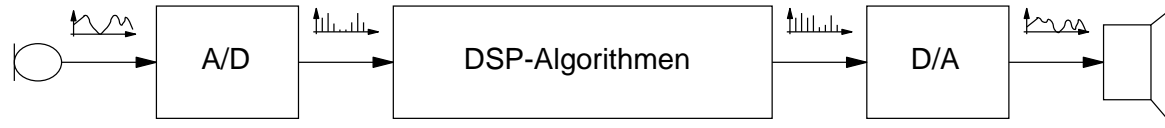
- zeitdiskret (Abtastung)
- wertdiskret (Quantisierung)
- Differenzengleichungen
- Überlauf, Rundungsfehler

DSP: Audio-Anwendungen

Signalverarbeitung im Audio-Bereich:

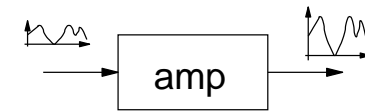
- Datenspeicherung (CD, DVD, ...)
 - Datenkompression (MP3, AC3, MiniDisc, ...)
 - Datenübertragung (DSR, GSM, . . .)
- jeweils für Musik und Sprache
- Musikproduktion (digitales Tonstudio)
 - Musiksynthese (z.B. virtuell analog)
 - Audioeffekte (z.B. Hall)
 - Surround-Verfahren (z.B. DirectX mit HRTF)
- Samplefrequenz typ. 44.1 / 48 KHz
 - hohe Genauigkeit erforderlich: 16 .. 24 bit

DSP: Verstärker, Mixer



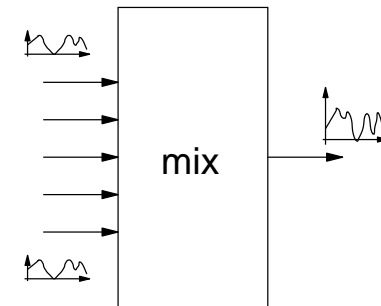
- digitaler Verstärker

```
output[t] = input[t] * gain;
```



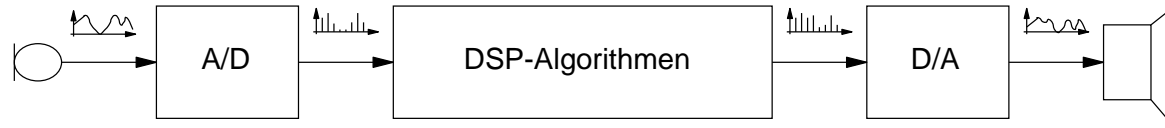
- digitaler Mixer:

```
output[t] = 0;
for( int i=0; i < n_inputs; i++) {
    output[t] += input[i][t] * gain[i];
}
```



- viele MAC-Operationen (multiply-accumulate)
- Problem Aussteuerung vs. Überlauf => saturation arithmetic

DSP: *Echo, Hall, und mehr*



- Verzögerung: auf alte Abtastwerte zurückgreifen

```
delay[t] = input[t]*gain + input[t-t2]*gain2;
```

- Rekursion möglich: Echo, Hall, Raumsimulation

```
hall[t] = input[t]*gain + hall[t-t2]*gain2;
```

- Algorithmen für viele Anwendungen:
Verzögerung, Echo, Hall, Verzerrung, "Overdrive", ...
Filter, Formatfilter, Tonhöhenänderung, Tempoänderung, ...

DSP: Sampling und Synthese

- Sampling: Abtastwerte abspeichern

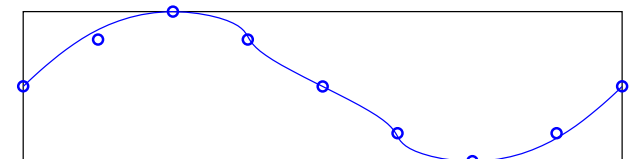
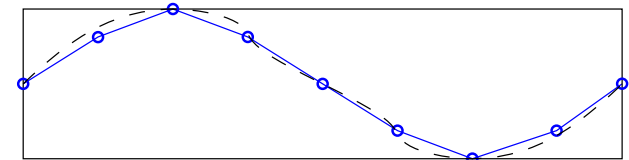
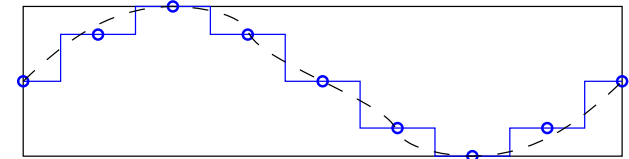
```
sample[t] = input[t];
```

- Samples direkt abspielbar (CD, Spiele)

- "Wavetable"-Synthesizer:

```
output[t] = sample[t*pitch]
            interpolate( sample[] )
            filter( interpolate( sample[] ) )
            effects( filter( interpolate( sample[] ) ) )
```

- Interpolation: nearest / linear / splines / ...



keine/lineare/Spline-Interpolation

Audiosysteme: Anforderungen

Grundanforderungen:

- CD/DAT Qualität 16 bit linear, 44.1 / 48 KHz
- höhere Genauigkeit 24 bit linear PCM
- oder Gleitkomma 32 bit IEEE

Beispiele:

	# MAC/s	(48 KHz)
• 10 Kanal Mixer	480.000	
• 100-tap FIR-Filter	4.800.000	
• 1000-tap FIR-Filter	48.000.000	
• 10 Kanal Mixer, je 3 100-tap FIR	144.480.000	
• . . .		
• Pentium-200 (10 Takte/Mult.)	20.000.000	
• Pentium-III, 1 GHz, ISSE (max:)	4.000.000.000	

=> Spezialhardware oder aktuelle Prozessoren

Zahlenfolgen, Elementarfolgen

Wertemenge $\{Y(k)\}$

Zuordnung: Index $k \rightarrow Y(k)$

wichtige "Elementarfolgen":

Impulsfolge

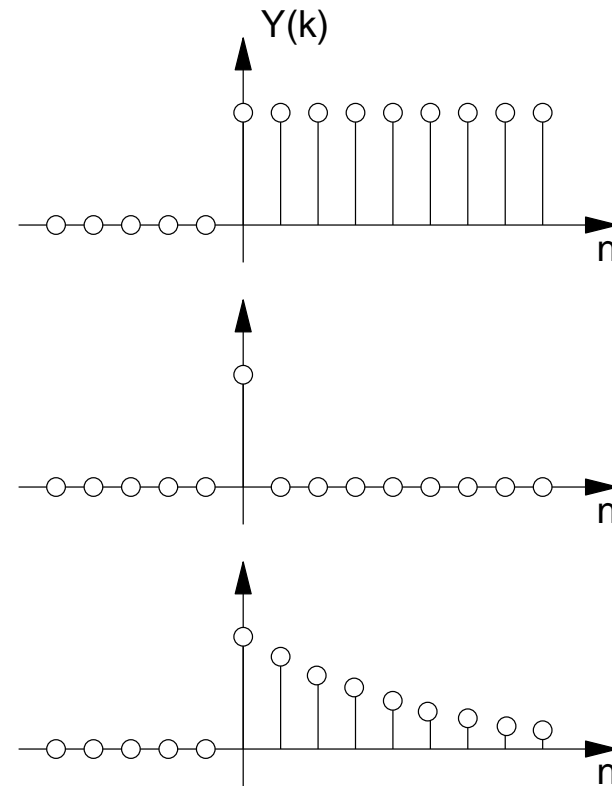
Sprungfolge

Exponentialfolge

Sinusfolge

lineare Systeme:

- Signale aus Elementarfolgen zusammensetzen
- Systemverhalten entsprechend berechenbar



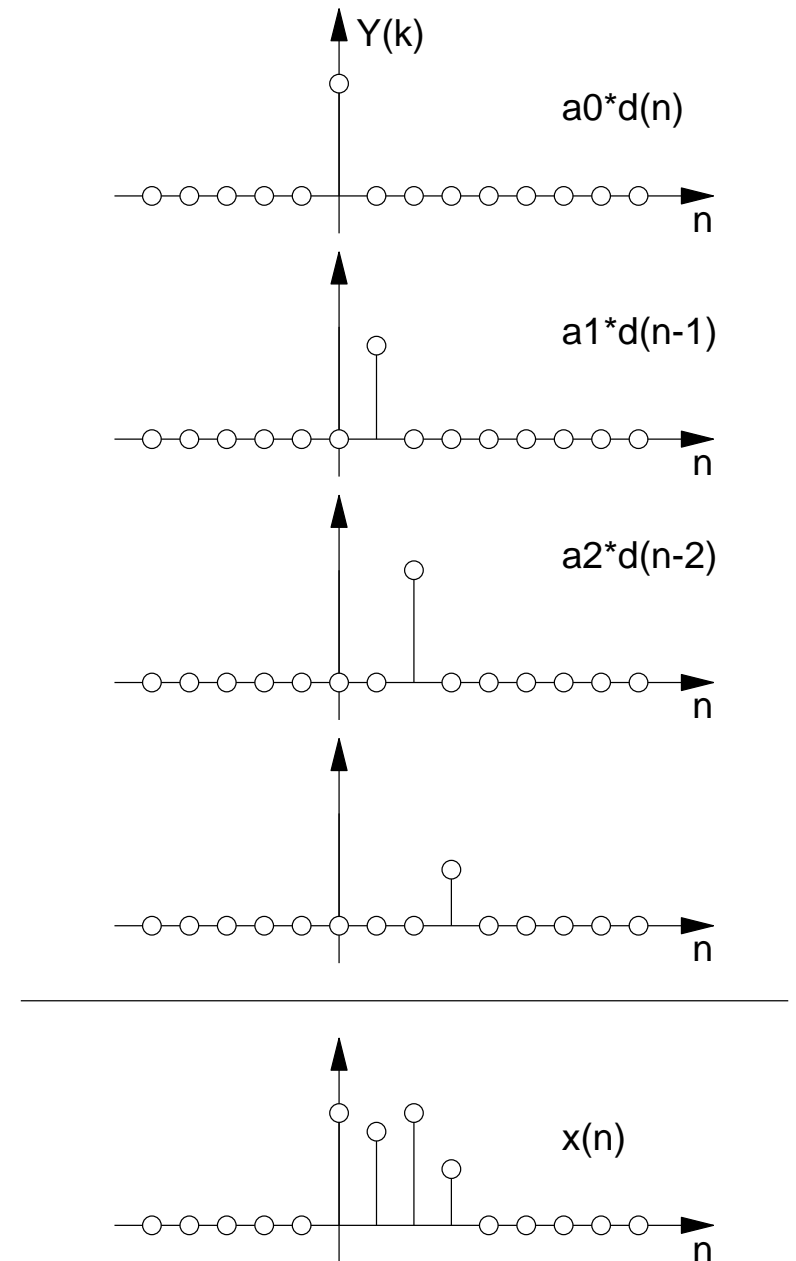
Impulsfolge:

Konstruktion beliebiger Zahlenfolgen:

- Summe
- skaliertes
- zeitverschobenes
Impulsfolgen

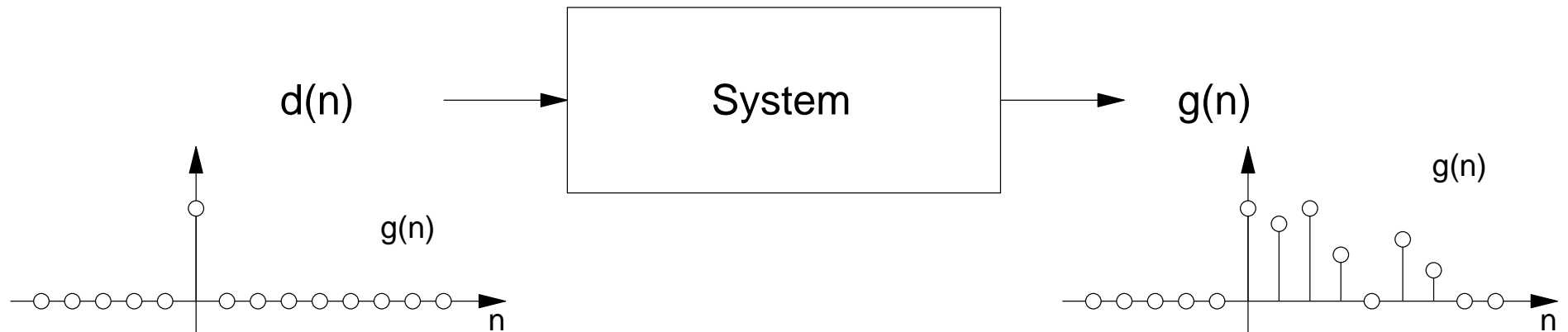
$$x(n) = a_0 * d(n) + a_1 * d(n-1) + \dots$$

- gleichwertig: Konstruktion mit
- skalierten, zeitverschobenen
- Sprungfolgen



Impulsantwort

"Impulsantwort" := Ausgangsfolge als Reaktion auf die Impulsfolge



- => liefert vollständige Beschreibung von LTI-Systemen
"linear, time-invariant"
z.B. Verstärker, Filter, usw.
- => auch als lineare Approximation nichtlinearer Systeme

Differenzengleichungen

Berechnung des Ausgangswerts $Y(n)$

- aus gegenwärtigem Eingangswert $X(n)$
- aus früheren Eingangswerten $X(n-k)$
- aus früheren Ausgangswerten $Y(n-k)$ (Rekursion)

- besonders wichtig: linear, konstante Koeffizienten, 2. Ordnung

$$y(n) = A_0 x(n) + A_1 x(n-1) + A_2 x(n-2) - B_1 y(n-1) - B_2 y(n-2)$$

Ordnung N : $B_1 \dots B_N$ kommen vor, $A_1 \dots A_M$, $M < N$

- Übertragungsfunktion $H(z)$ durch "Z-Transformation":

$$Y(z) = A_0 X(z) + A_1 X(z)/z + A_2 X(z)/z^2 - B_1 Y(z)/z + B_2 Y(z) / z^2$$

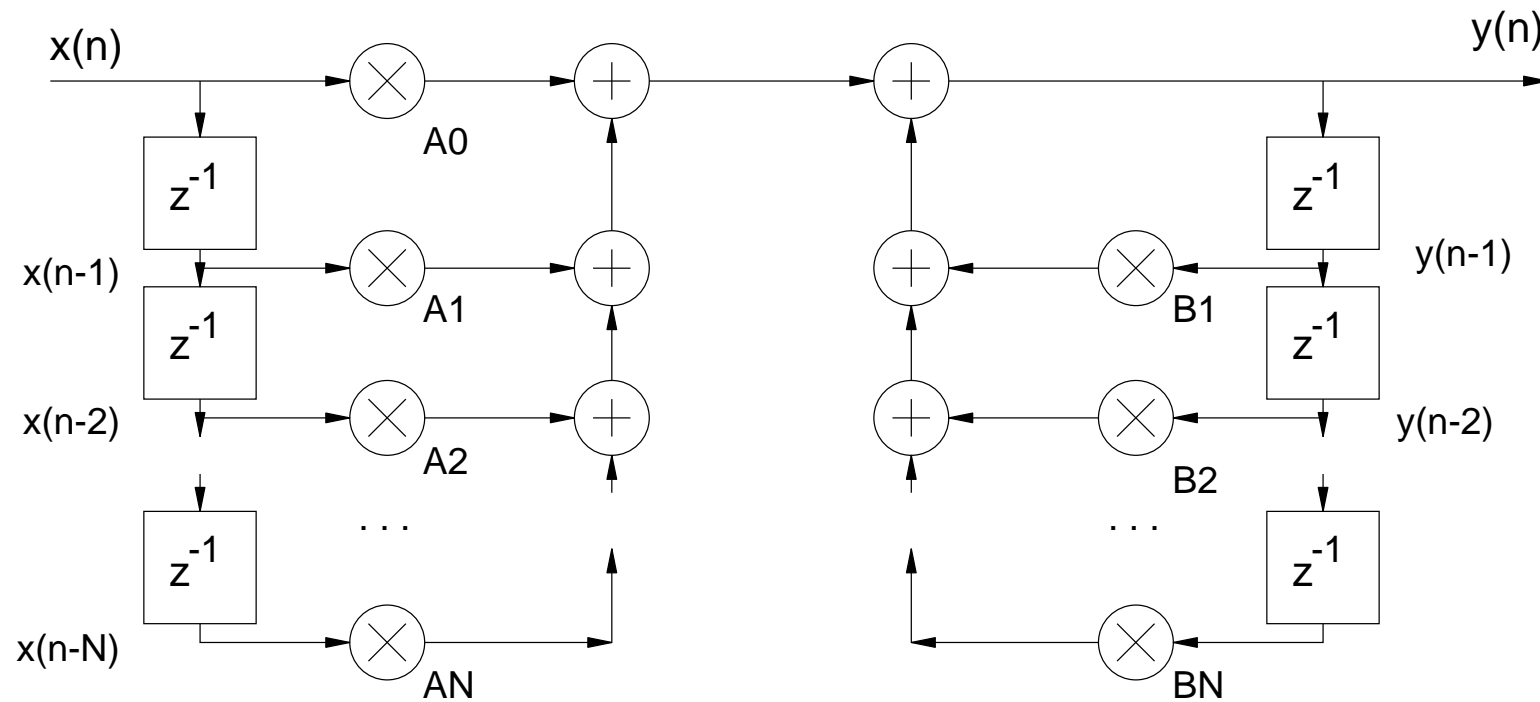
DGL: Direktform

Differenzengleichungen, linear, konstante Koeffizienten

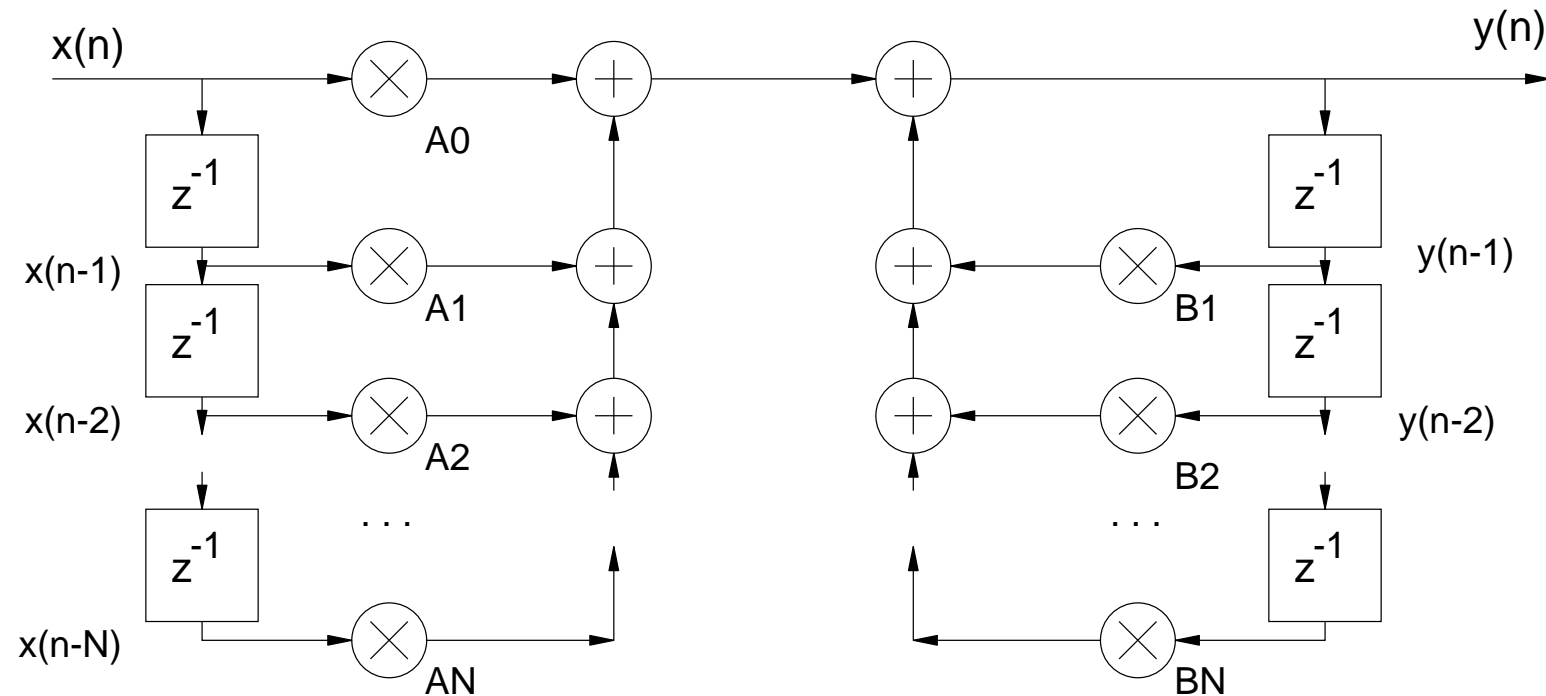
$$y(n] = A_0 x(n) + A_1 x(n-1) + A_2 x(n-2) - B_1 y(n-1) - B_2 y(n-2)$$

Umsetzung in "Direktform":

- Addition, Multiplikation, Verzögerung (=Register)



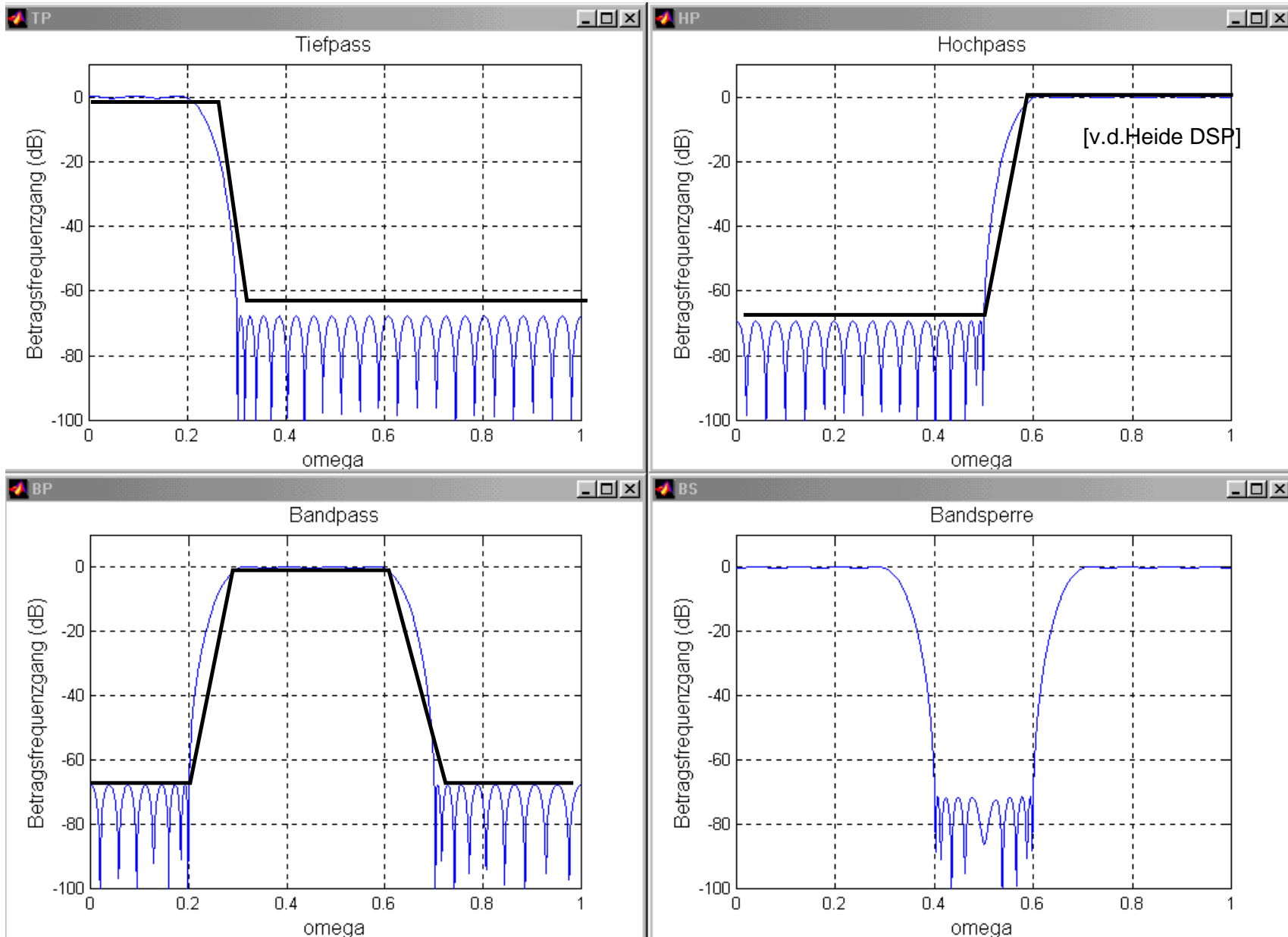
DGLs: Blockschaltbilder



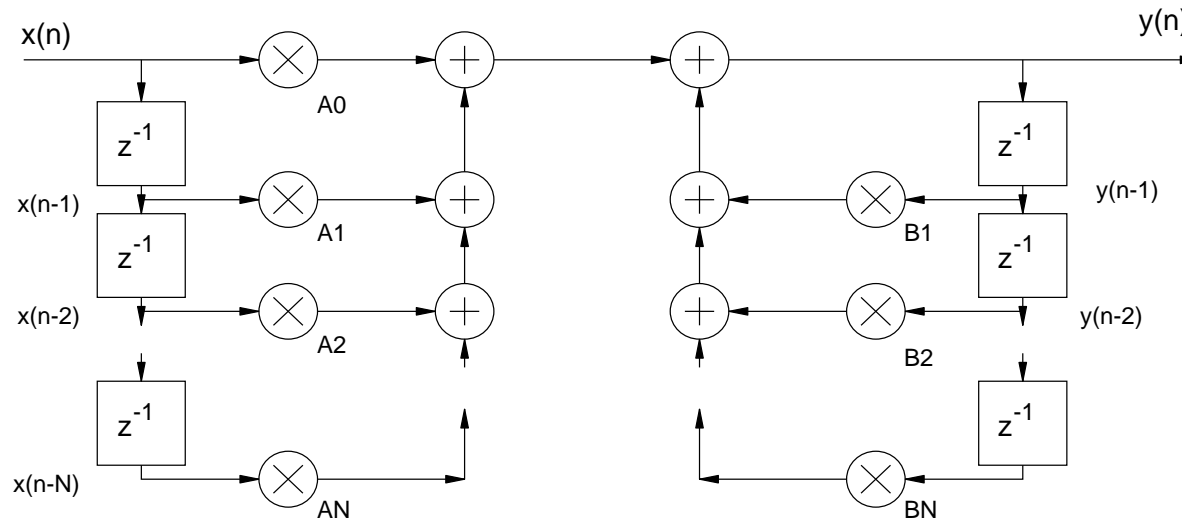
- Direktform erlaubt sofortige Realisierung von DSP-Algorithmen
- Varianten möglich (Vertauschen von Registern/Operatoren)
- bei Bedarf zusätzlich weitere (z.B. nicht-lineare) Operatoren

=> Darstellung mit Blockschaltbildern

Filter: Tiefpaß, Hochpaß, etc.



Filter: Frequenzgang



- normierte Frequenz $F = f/FA$
- direkte Berechnung des Frequenzgangs via $H(z)$:

$$|\underline{A}(j\omega)| = \sqrt{\frac{\left[\sum_{k=0}^n \alpha_k \cos 2\pi k F \right]^2 + \left[\sum_{k=0}^n \alpha_k \sin 2\pi k F \right]^2}{\left[\sum_{k=0}^n \beta_k \cos 2\pi k F \right]^2 + \left[\sum_{k=0}^n \beta_k \sin 2\pi k F \right]^2}}$$

Abtasttheorem

macht die Verarbeitung von Zahlenfolgen überhaupt Sinn ?

Abtasttheorem:

*ein auf die halbe Abtastfrequenz (=Nyquistfrequenz)
bandbegrenzte Signal kann nach der Abtastung
perfekt rekonstruiert werden*

- Grundlage aller digitalen Signalverarbeitung
- Beweis mit Fouriertransformation

- siehe [Shannon 48]
- siehe Matlab-Skript Signalverarbeitung

Abtasttheorem

7.4 Abtasttheorem für Zeitfunktionen

Ein reelles bandbegrenzttes Signal $u(t)$ hat eine Signalgrenzfrequenz f_{gs} . Für die Fourier-Transformierte $\underline{U}(f)$ dieses Signals gilt:

$$\underline{U}(f) = 0 \quad \text{für } |f| > f_{gs} \quad . \quad (7.5)$$

Durch periodische Wiederholung von $\underline{U}(f)$ mit der Periode $2 f_{gs}$ entsteht eine periodische Frequenzfunktion, die in eine Fourier-Reihe

$$\underline{U}_{\text{per}}(f) = \sum_{n = -\infty}^{\infty} c_n \exp(-j 2\pi n \frac{f}{2 f_{gs}}) \quad (7.6)$$

entwickelt werden kann. Die Zeitfunktion

$$u(t) = \int_{-\infty}^{\infty} \underline{U}(f) \exp(j 2\pi ft) df \quad (7.7)$$

ergibt sich als inverse Fourier-Transformierte der Frequenzfunktion $\underline{U}(f)$. Da diese

(Gerdson u. Kröger)

Abtasttheorem

nur im Bereich

$$-f_{gs} < f < +f_{gs} \quad (7.8)$$

von null verschieden ist und innerhalb dieses Bereiches mit $U_{per}(f)$ übereinstimmt, kann für die Zeitfunktion auch

$$u(t) = \int_{-f_{gs}}^{f_{gs}} \left\{ \sum_{n=-\infty}^{\infty} c_n \exp(-j 2\pi n \frac{f}{2f_{gs}}) \right\} \exp(j 2\pi f t) df \quad (7.9)$$

geschrieben werden. Man faßt die e-Funktionen zusammen und bildet die Summe der Integrale. So erhält man:

$$u(t) = \sum_{n=-\infty}^{\infty} c_n \int_{-f_{gs}}^{f_{gs}} \exp(-j 2\pi f [t - \frac{n}{2f_{gs}}]) df \quad (7.10)$$

Die Lösung des Integrals in dieser Summe ergibt:

$$\int_{-f_{gs}}^{f_{gs}} \exp(-j 2\pi f [t - \frac{n}{2f_{gs}}]) df = 2 f_{gs} \frac{\sin(2\pi f_{gs} [t - \frac{n}{2f_{gs}}])}{2\pi f_{gs} [t - \frac{n}{2f_{gs}}]} \quad (7.11)$$

Zur Bestimmung der Fourier-Koeffizienten c_n wird

$$t = k \frac{1}{2 f_{gs}} \quad k \text{ ganze Zahl} \quad (7.12)$$

gesetzt. Dann sind in der Summe (7.10) alle Terme mit $n \neq k$ null. Damit gilt für

Abtasttheorem

die Koeffizienten:

$$c_n = \frac{1}{f_{gs}} u\left(n \frac{1}{2 f_{gs}}\right) \quad . \quad (7.13)$$

So entsteht aus (7.10) die Interpolationsformel

$$u(t) = \sum_{n=-\infty}^{\infty} u\left(n \frac{1}{2 f_{gs}}\right) \frac{\sin\left(2\pi f_{gs}\left[t - \frac{n}{2 f_{gs}}\right]\right)}{2\pi f_{gs}\left[t - \frac{n}{2 f_{gs}}\right]} \quad , \quad (7.14)$$

die aussagt, daß ein bandbegrenzttes Signal $u(t)$ mit der Signalgrenzfrequenz f_{gs} vollständig beschrieben ist durch seine Abtastwerte zu den Zeiten

$$t_k = k \frac{1}{2 f_{gs}} \quad k \text{ ganze Zahl} \quad . \quad (7.15)$$

Dieses Abtasttheorem für Zeitfunktionen ist von grundlegender Bedeutung für die digitale Signalverarbeitung.

Abtasttheorem: Schema

Eingangsspektrum

Eingangstiefpaß

bandbegrenztes Signal

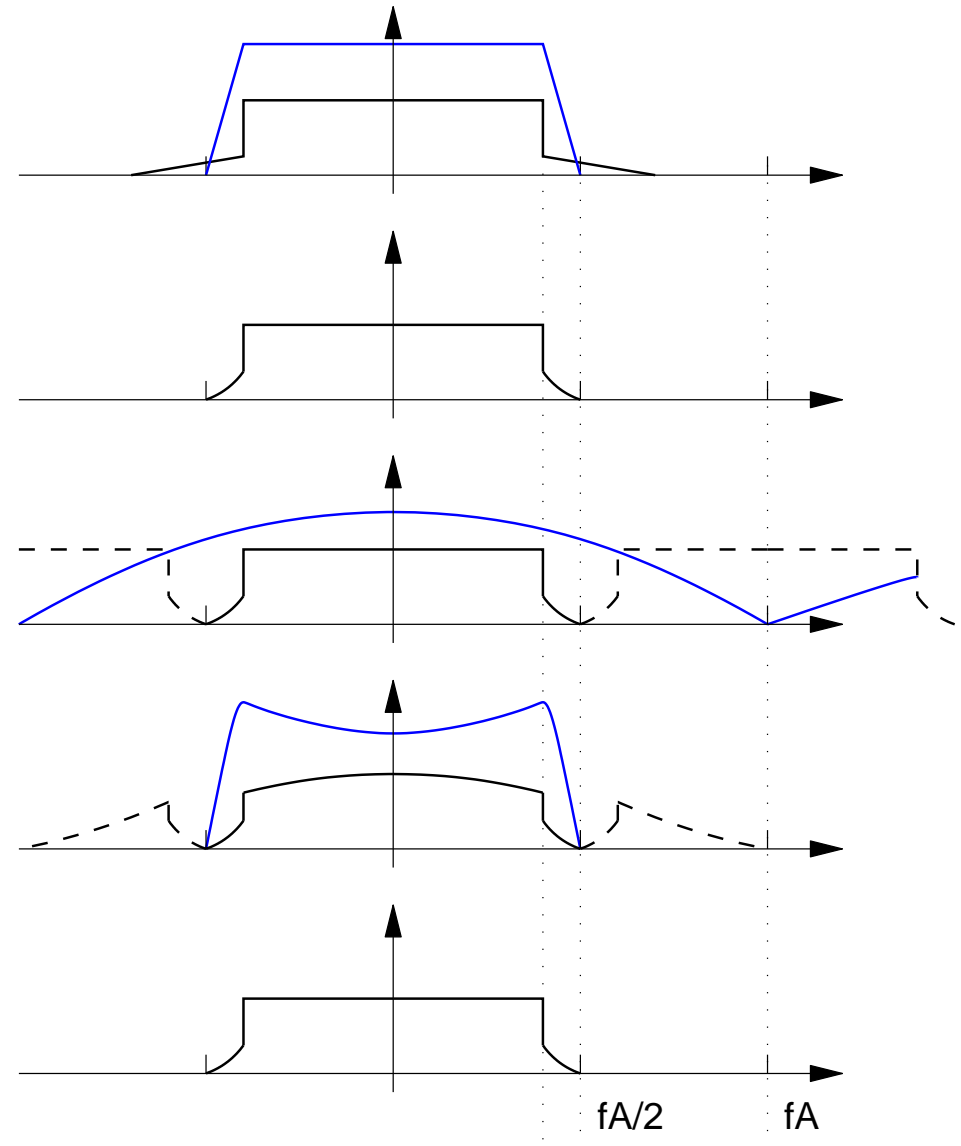
Abtastung

Abtastung $\sin(x)/x$

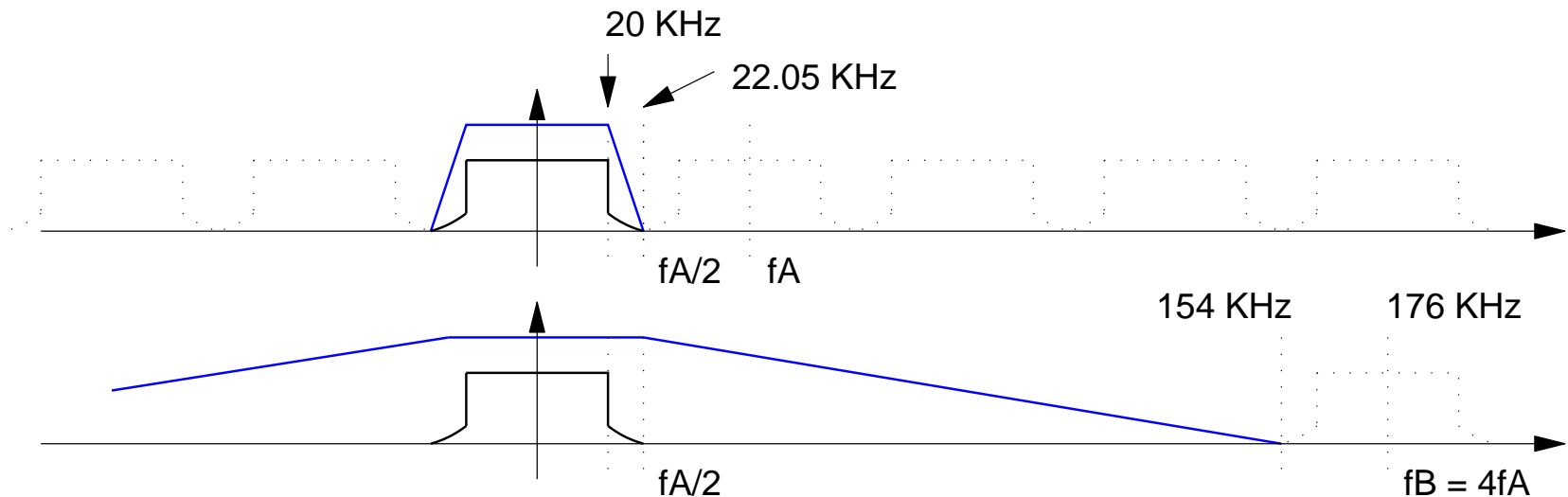
Rekonstruktion

Ausgangstiefpaß

Ausgangsspektrum



Abtasttheorem: Überabtastung



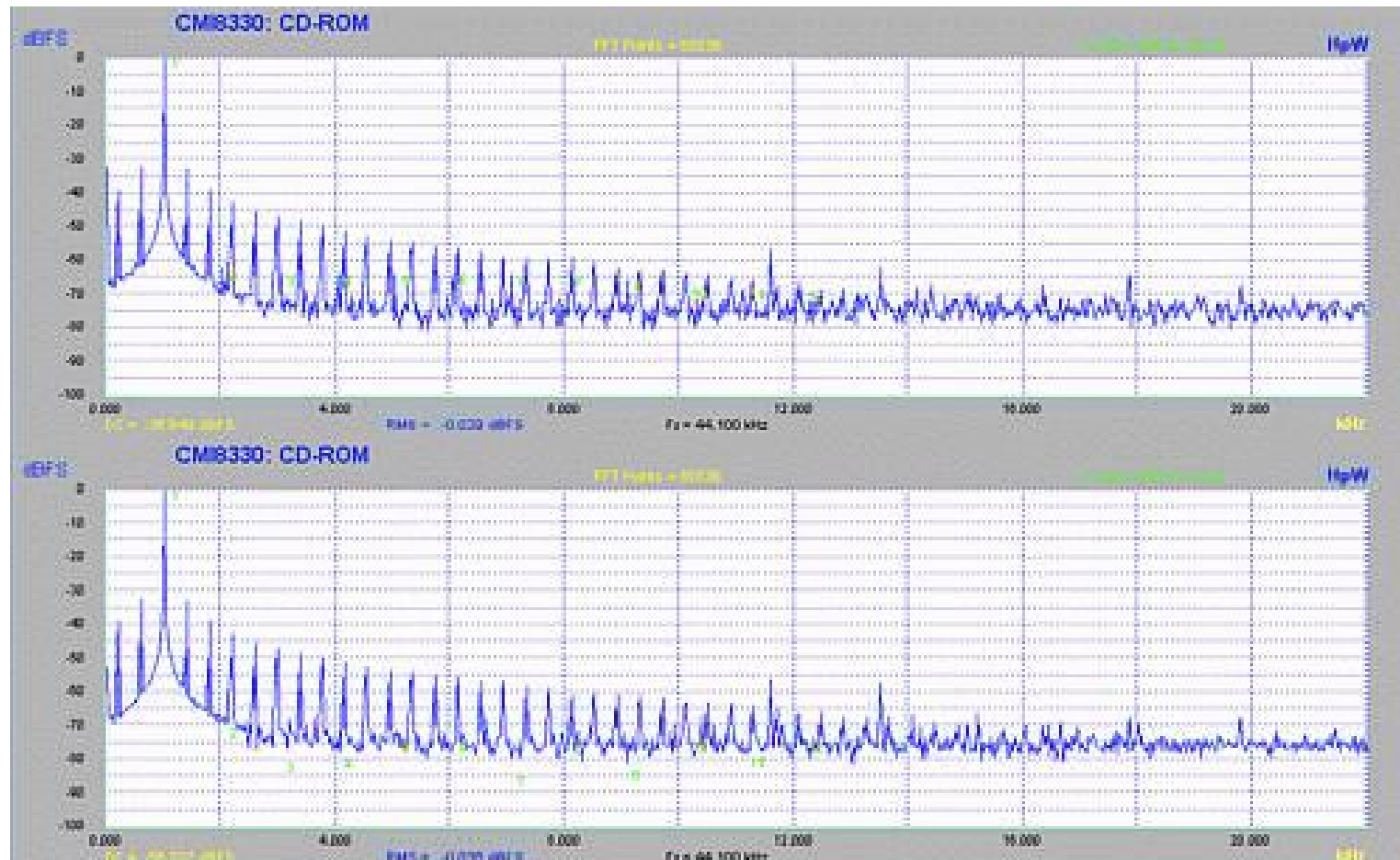
Abtasttheorem:

- vollständige Rekonstruktion des Signals, aber:
- erfordert analogen Tiefpaß sehr hoher Qualität
- Beispiel CD: Dämpfung 96 dB von 20 KHz bis 22.05 KHz
- typische analoge Filter: 6 .. 24 dB / Oktave

Oversampling:

- digitale Neu-Abtastung mit höherer Rate (z.B. 4x, 8x)
- digitale Tiefpaßfilterung beseitigt die Spiegelfrequenzen
- einfacher analoger Tiefpaß genügt

Abtasttheorem: Aliasing-Fehler . . .



(c't 10/98)

- Meßwerte der CMI 8330 Soundkarte
- Sinussignal und massive Anti-Aliasing Fehler (Spiegelfrequenzen)
- massiver Qualitätsverlust durch Stümperei

A/D-Wandlung

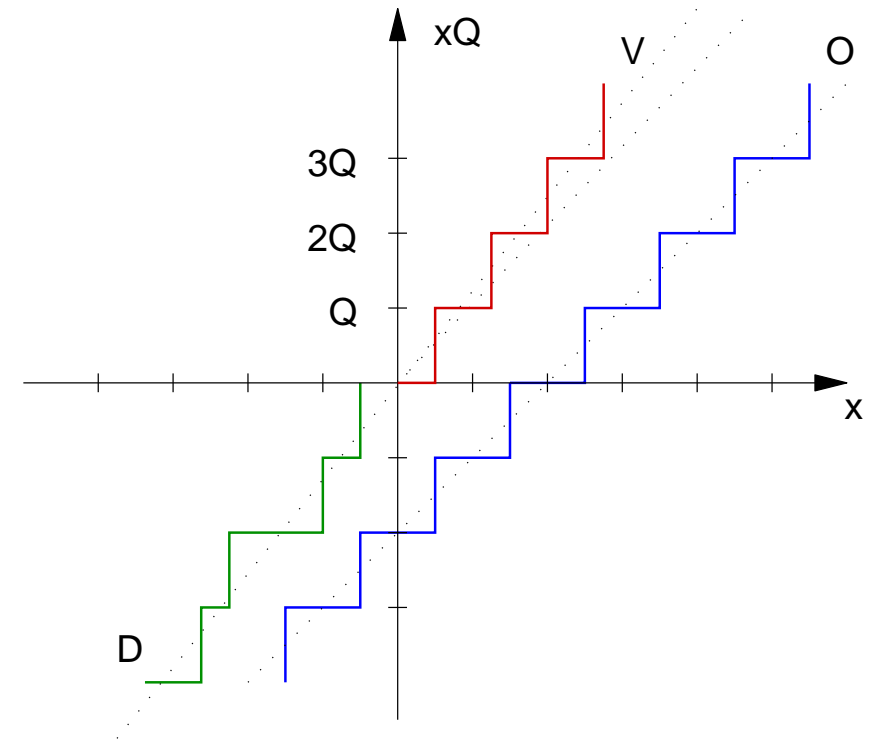
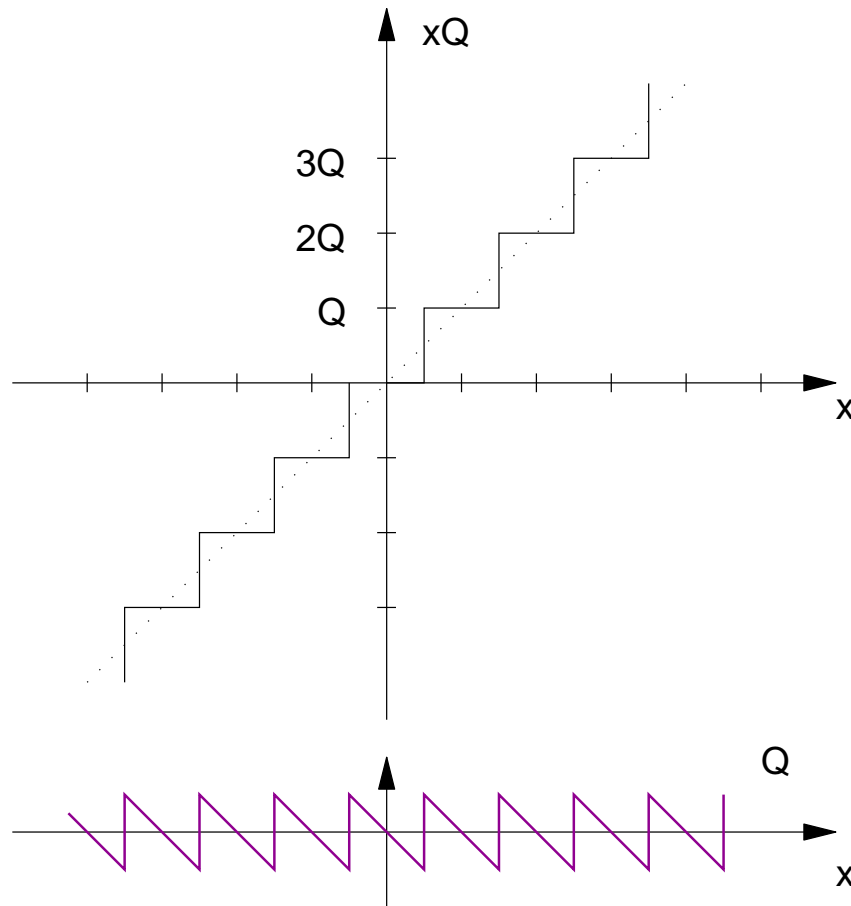
Digitalisierung eines Eingangssignals:

- Zeitquantisierung (Abtaster, Sample&Hold)
- Amplitudenquantisierung (A/D-Konverter)
- Codierung (meistens 2-Komplement)

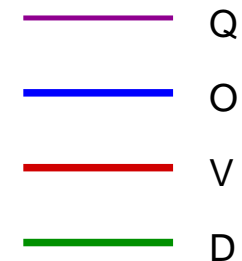
- nur für bandbegrenzte Eingangssignale
- analoger Tiefpaß vor dem Eingang erforderlich

- diverse Verfahren
- Trade-off zwischen Performance / Genauigkeit / Kosten
- für Audio: 24-bit Wandler bereits erhältlich

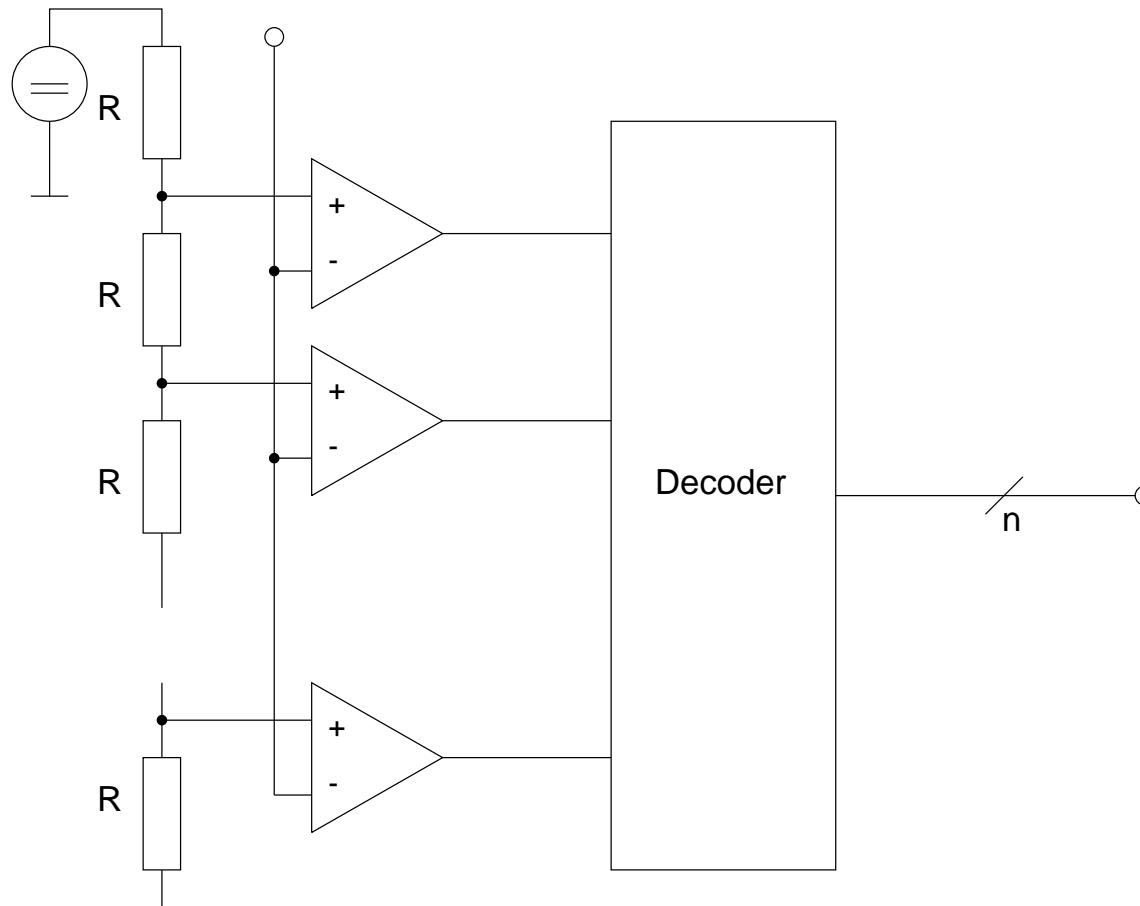
A/D: Parameter



- Quantisierungsfehler
- Offset / Nullpunktsfehler
- Verstärkungsfehler (falsche Steigung)
- differentielle / integrale Nichtlinearität



A/D: Parallel-Umsetzer

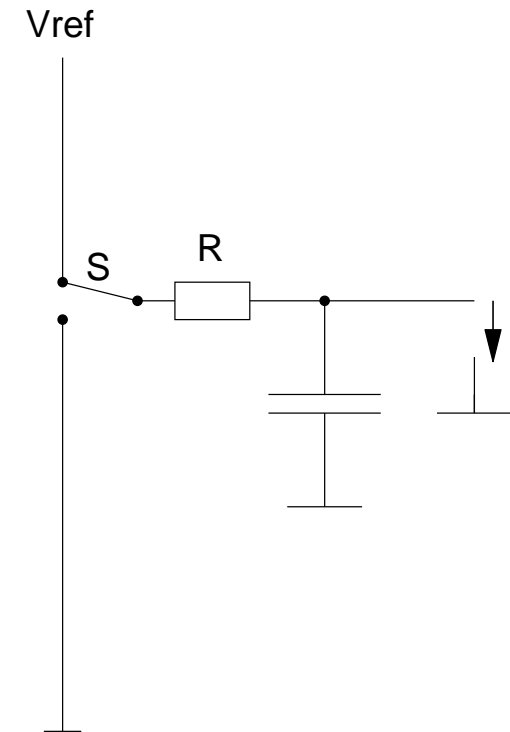
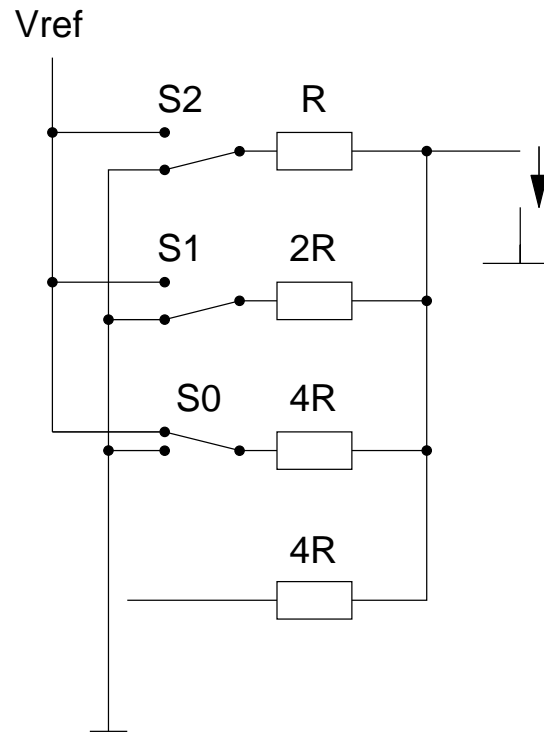
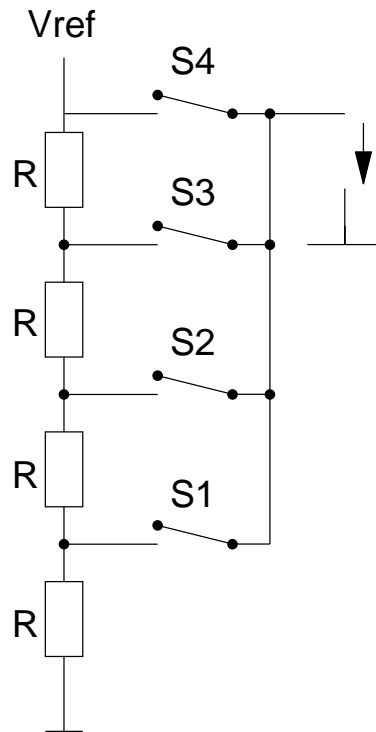


- erfordert $2^{*}n$ Komparatoren
- sehr schnell
- aber nur für geringe Auflösungen (z.B. 8-bit)

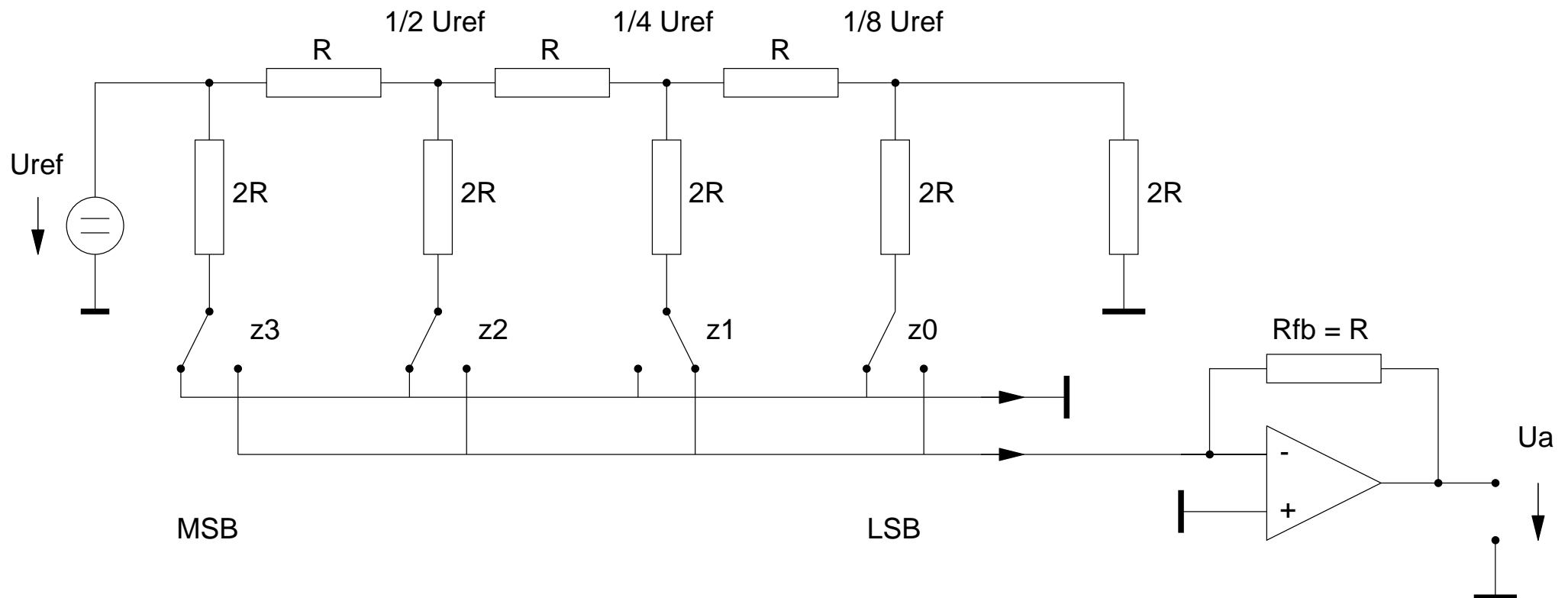
D/A-Wandlung

D/A: Umwandlung numerischer Werte in analoge Werte (Spannungen)

- Parallelverfahren
- Wägeverfahren
- Zählverfahren



D/A: R-2R Netzwerk

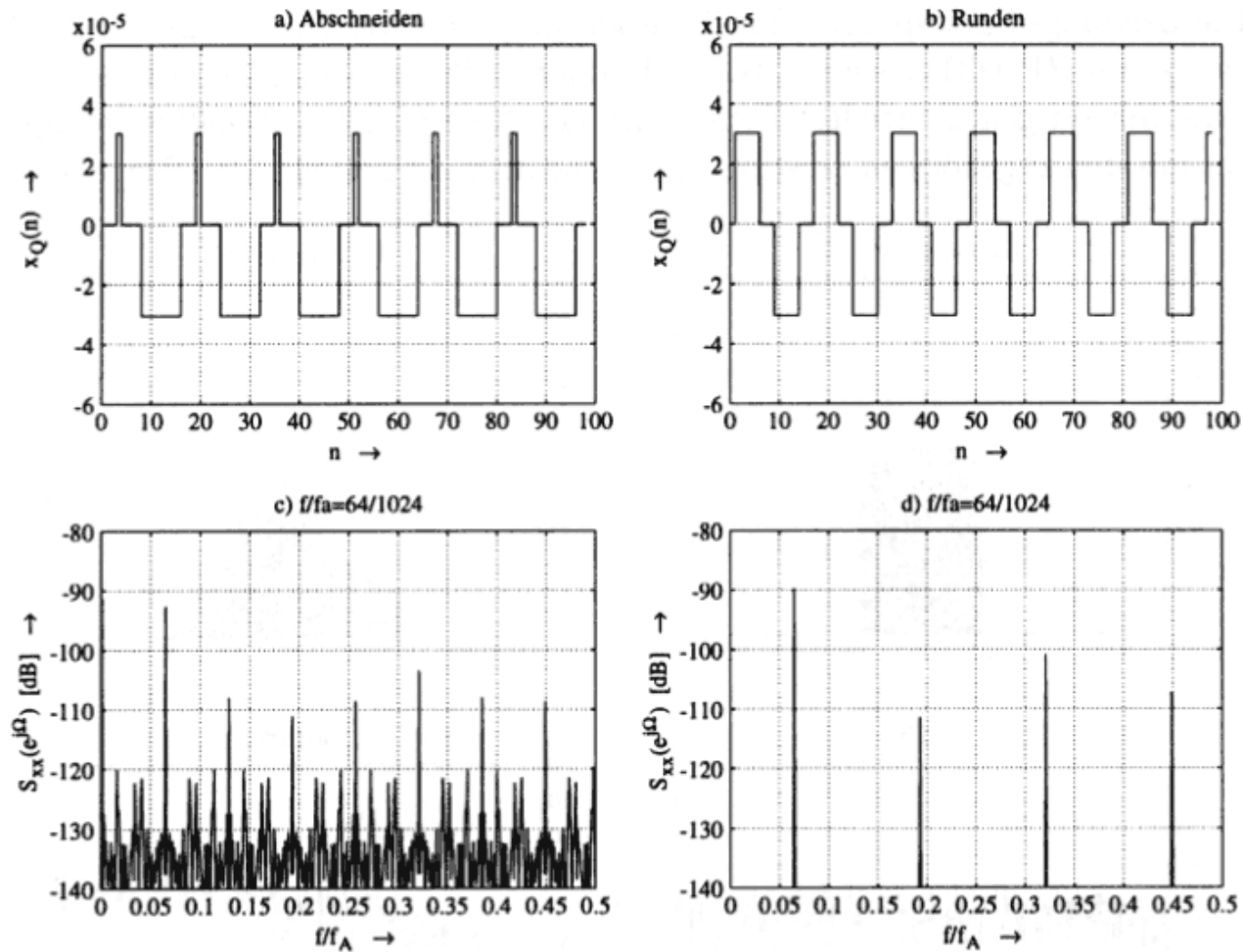


- VLSI: Herstellung unterschiedlicher Widerstände schwierig
 - aber gleiche Widerstände mit hoher Genauigkeit
- => fortgesetzte Spannungsteilung mit R-2R Netzwerk

$$U_a = -U_{ref} * (R_{fb}/16R) * (8z_3 + 4z_2 + 2z_1 + z_0)$$

- in CMOS-Technologie oft verwendet

Quantisierungsprobleme

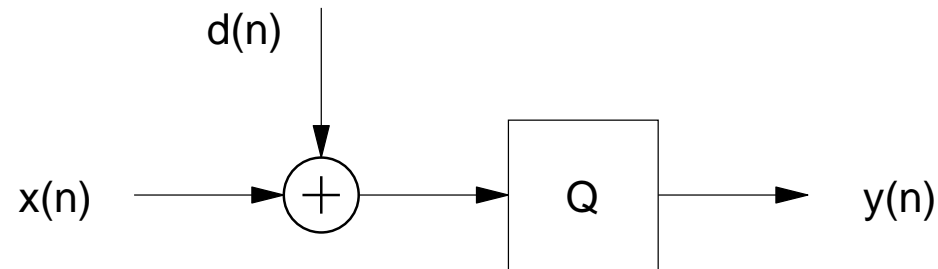


(Zölzer)

Bild 2.19: 1 Bit Aussteuerung des Quantisierers bei Abschneiden a/c und Runden b/d

- Quantisierung erzeugt störende "Nebensignale"
- Verbesserung durch Hinzufügen von Rauschen: "Dithering"

Dither: Prinzip



		bits
$x(n)$	Eingangssignal	$w+r$
$d(n)$	Zufallssignal (dither)	$s < r$
$y(n)$	Ausgangssignal	w

- Q: Re-Quantisierung des Signals, oder D/A-Wandler
- Amplitude des Dithers kleiner als LSB des Eingangssignals
- Verteilung (Statistik) des Dithers frei wählbar: noise shaping
- Mathematik siehe Zölzer S.38 ff

Dither: bei 1 bit Aussteuerung

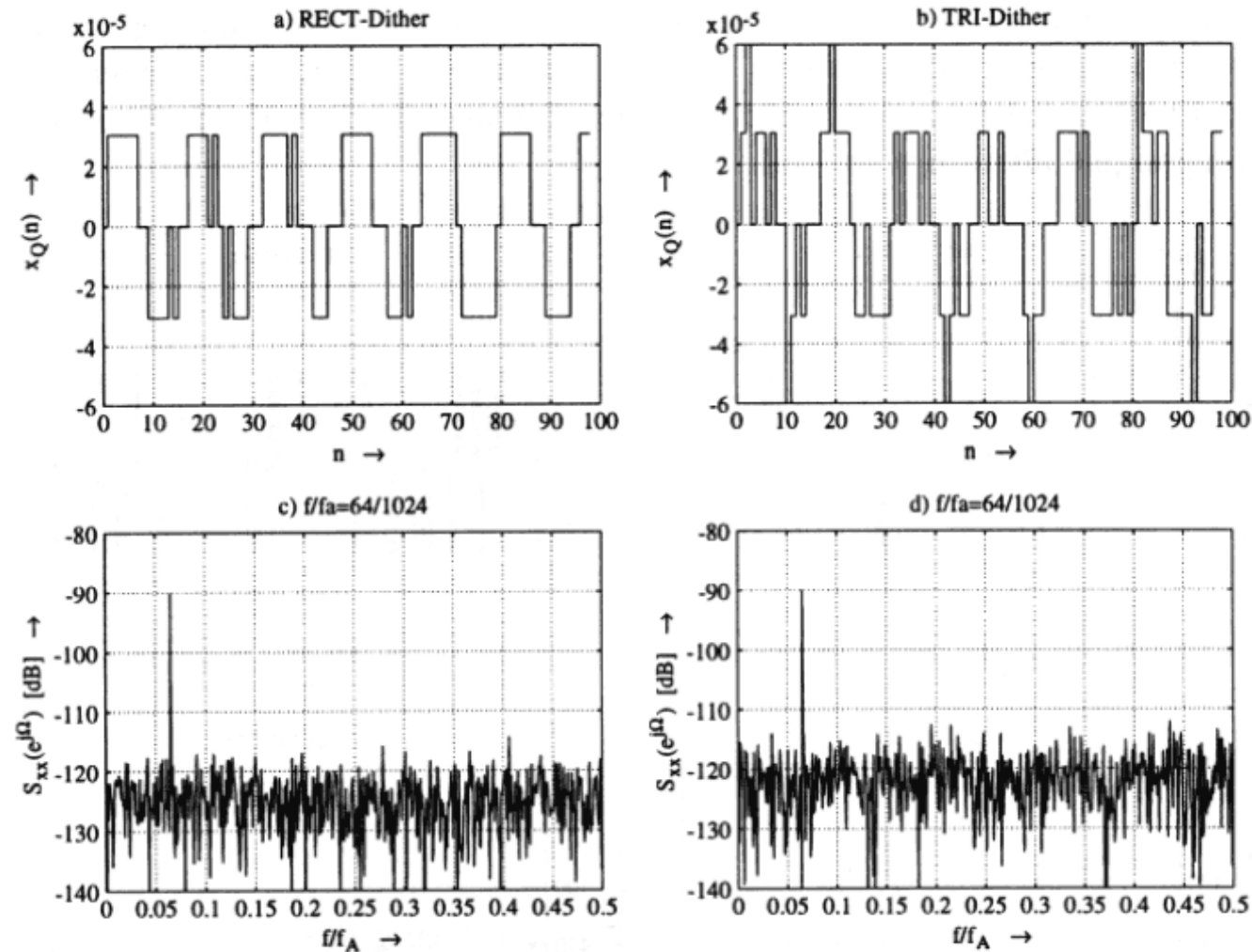


Bild 2.20: 1 Bit Aussteuerung - Rundungskennlinie mit RECT-Dither a/c und TRI-Dither b/d

(Zölzer)

Dither: bei 0.25 bit Aussteuerung

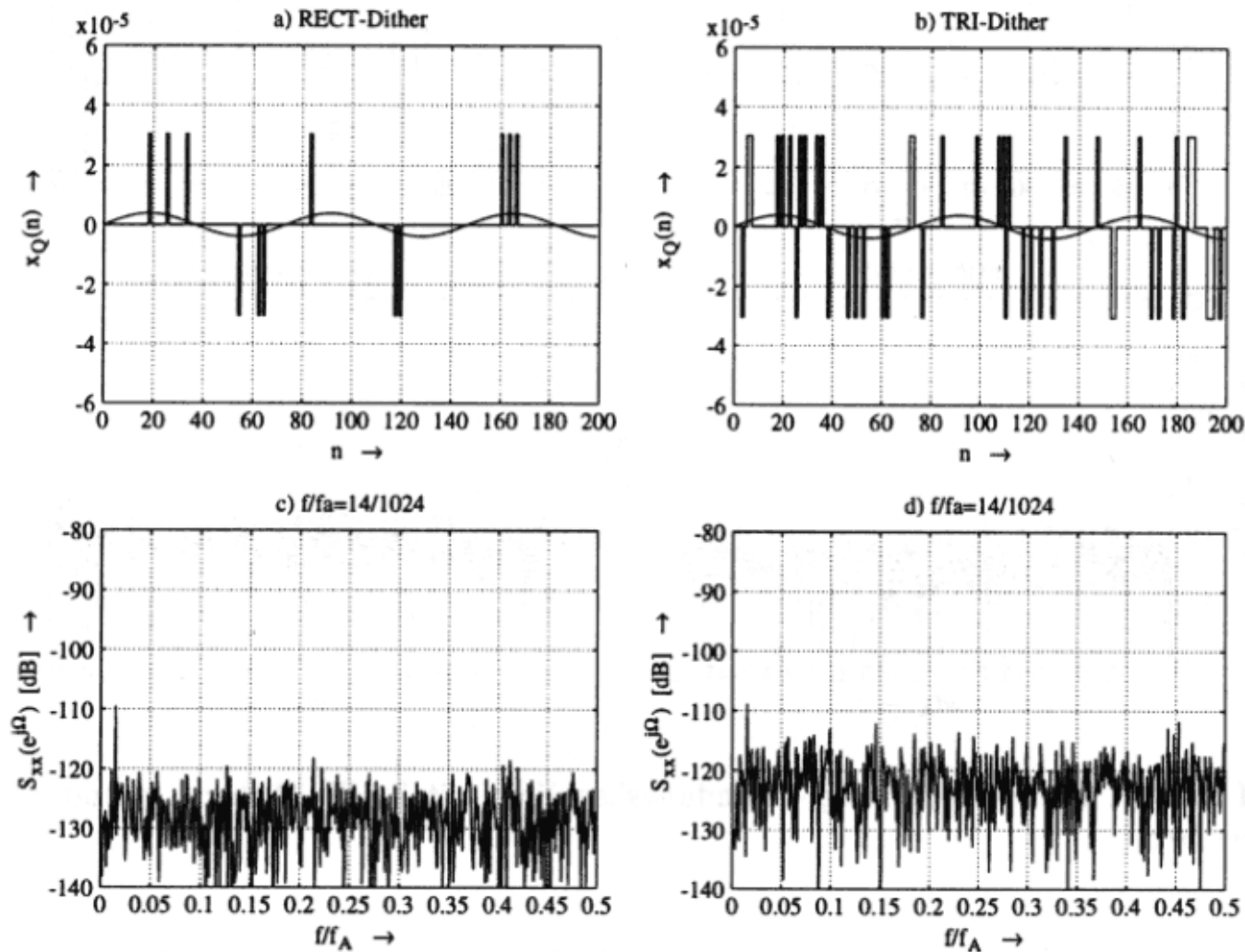


Bild 2.21: 0.25 Bit Aussteuerung - Rundungskennlinie mit RECT-Dither a/c und TRI-Dither b/d

(Zölzer)