

Unterlagen zum Tutorium in den ersten Semestern des Tonmeisterstudiums

Alexander Lerch

Version 0.3

Inhaltsverzeichnis

1	Schwingungen und Wellen	1
1.1	Schwingungen	1
1.2	Frequenz und Periodendauer	2
1.3	Amplitude	4
1.4	Phasenverschiebung	4
1.5	Töne, Klänge und Klangfarben	6
1.6	Wellen	8
1.7	Ausbreitungsgeschwindigkeit	8
1.8	Wellenlänge	9
1.9	Ausbreitung von Wellen	10
1.10	Zusammenfassung	10
1.11	Aufgaben	11
2	Schall und Schallfeld	13
2.1	Schallschnelle und Schalldruck	13
2.2	Entfernungsabhängigkeit des Schalldrucks	15
2.3	Nahfeld und Fernfeld/ Entfernungsabhängigkeit der Schallschnelle	15
2.4	Schalleistung	17
2.5	Schallintensität	17
2.6	Gestörte Schallausbreitung	17
	2.6.1 Reflexion	18
	2.6.2 Beugung	19
2.7	Zusammenfassung	19
2.8	Aufgaben	20
3	Pegel	21
3.1	Das Weber-Fechnersche Gesetz	21
3.2	Schalldruckpegel & Co.	22
3.3	Pegel bei der Addition von Signalen	24
3.4	Zusammenfassung	28
3.5	Aufgaben	28

4	Raumakustik	31
4.1	Nachhallzeit	32
4.2	Absorption	33
4.3	Direktschallfeld und Diffusschallfeld	33
4.4	Hallradius	33
4.5	Frequenzabhängigkeit des Nachhalls	36
4.6	Flatterecho	36
4.7	Stehende Wellen	37
4.8	Zusammenfassung	38
4.9	Aufgaben	39
5	Psychoakustik	41
5.1	Physiologie des Ohres	41
5.1.1	Das Außenohr	41
5.1.2	Das Mittelohr	42
5.1.3	Das Innenohr	43
5.2	Schalldruckpegel, Lautstärke und Lautheit	45
5.3	Tonhöhenempfindung	46
5.4	Verdeckungseffekte	46
5.5	Richtungshören bei einer Schallquelle	48
5.5.1	Das Gesetz der ersten Wellenfront	49
5.6	Richtungshören im Stereodreieck	52
5.7	Zusammenfassung	52
5.8	Aufgaben	54
6	Signalverarbeitung	55
6.1	Signalbeschreibungen	55
6.1.1	Zeitverlauf	55
6.1.2	Spektrum	55
6.1.3	Amplitudendichteverteilung	57
6.2	Systembeschreibungen	60
6.2.1	Lineare Systeme	60
6.2.2	Nichtlineare Systeme	61
6.3	Grundlegende Signalformen bei der Signalsynthese	61
6.4	Geräte der Signalverarbeitung	63
6.4.1	Fader	63
6.4.2	Filter	63
6.4.3	Delay und Delay-basierte Effekte	66
6.4.4	Hallgerät	67
6.4.5	Dynamikbearbeitung	68
6.4.6	Verzerrer und Enhancer	69
6.4.7	andere Effekte	69
6.5	Qualitätsmerkmale	69

6.5.1	Übertragungsfrequenzgang	69
6.5.2	Dynamikumfang	70
6.5.3	Klirrfaktor	70
6.5.4	Total Harmonic Distortion and Noise	71
6.5.5	Modulationsverzerrungen	71
6.5.6	Übersprechdämpfung	72
6.5.7	Hörtests	72
6.6	Zusammenfassung	72
6.7	Aufgaben	73
7	Elektroakustische Wandler	75
7.1	Mikrophone	75
7.1.1	Charakteristische Kenngrößen	75
7.1.2	Mikrophontypen	79
7.1.3	Bauform	84
7.2	Lautsprecher	85
7.3	Zusammenfassung	85
7.4	Aufgaben	85
8	Digitaltechnik	87
8.1	Binäre Zahlendarstellung	87
8.2	Digitalisierung eines Analogsignals	89
8.2.1	Abtastung	89
8.2.2	Quantisierung	92
8.2.3	Wandler	100
8.3	Digitale Filter	100
8.3.1	Filterstrukturen	105
8.4	Speichermedien	105
8.4.1	Magnetbänder	105
8.4.2	Optische Medien	105
8.5	Übertragungstechnik/Fehlerbehandlung	105
8.5.1	Fehlerquellen bei der Übertragung	105
8.5.2	Jitter	105
8.5.3	Kanalcodes	105
8.6	Codierungsverfahren	105
8.6.1	Überblick über Codierungsverfahren	106
8.6.2	Wo können Bits gespart werden	106
8.6.3	Redundanzcodierung	106
8.6.4	Irrelevanzcodierung	107
8.7	Zusammenfassung	112
8.8	Aufgaben	112

Anhang	114
A Lösungen der Aufgaben	117
A.1 Schwingungen und Wellen	117
A.2 Schall und Schallfeld	119
A.3 Pegel	120
A.4 Raumakustik	121
A.5 Psychoakustik	122
A.6 Elektrotechnik	123
A.7 Signalverarbeitung	123
A.8 Mikrophone	124
A.9 Digitaltechnik	124
Verzeichnis der Abkürzungen und Formelzeichen	126
Abbildungsverzeichnis	129
Tabellenverzeichnis	132
Literaturverzeichnis	133
Index	137

Einleitung

Dieses Skript faßt einige Grundlagen zusammen, die in den ersten Semestern im Tutorium des Tonmeisterstudiums zu vermitteln sind. Es erhebt keinen Anspruch auf Vollständigkeit, sondern soll zur weiteren Beschäftigung mit der erhältlichen Fachliteratur anregen.

Da diese Unterlagen auch zum späteren Nachschlagen dienen können, ist eine Gliederung in Themenschwerpunkte sinnvoll. Die Abfolge einzelner Punkte im Tutorium kann also unter Umständen von der hier vorgenommenen Gliederung abweichen.

Kapitel 1

Schwingungen und Wellen

Ein Schallereignis entsteht, indem ein *Schallsender* die Luft (oder Luftteilchen) in Schwingungen versetzt. Ein Schallsender kann der Sprachtrakt eines Menschen mit seinen schwingenden Stimmbändern genauso sein wie die gezupfte, geschlagene oder gestrichene Saite eines Musikinstruments, die schwingende Luftsäule eines Musikinstruments oder die schwingende Membran eines Lautsprechers.

Die in Schwingung geratenen Luftteilchen regen wiederum ihre Nachbarteilchen zum Schwingen an; auf diese Art entsteht die *Schallwelle*, die das akustische Ereignis „transportiert“.

Ein Schallempfänger nimmt die Schwingungen der Luft auf und wandelt sie um, sei es im Mikrophon in elektrische Signale oder im Gehör über Trommelfell und andere Membrane in Nervenimpulse und anschließend in Informationen.

Im ersten Kapitel werden zunächst die typischen Merkmale akustischer Schwingungen eingeführt und behandelt. Auch die Begriffe Ton, Klang und Klangfarbe werden kurz erläutert. Anschließend geht es um die Eigenschaften von Wellen und ihre Ausbreitung.

1.1 Schwingungen

Eine Schwingung ist die zeitliche Zu- und Abnahme einer physikalischen Größe. Es wird zwischen *periodischen* und *nichtperiodischen* Schwingungsvorgängen unterschieden, je nachdem, ob sich die einzelnen Schwingungszustände in regelmäßigen Abständen wiederholen oder nicht. Das gängigste Beispiel für eine Schwingung ist eine Sinusschwingung:

Die in Abb. 1.1 dargestellte Schwingung ist in dieser Form erst einmal nicht-periodisch, weil sie zeitlich begrenzt ist. Stellt man sich diese Schwingung aber auf die gleiche Art und Weise rechts und links dieser Abbildung fortgesetzt vor, so erhält man eine periodische Schwingung.

Bemerkung

Sprache und Musik sind nichtperiodische Schwingungsvorgänge. Man kann solche

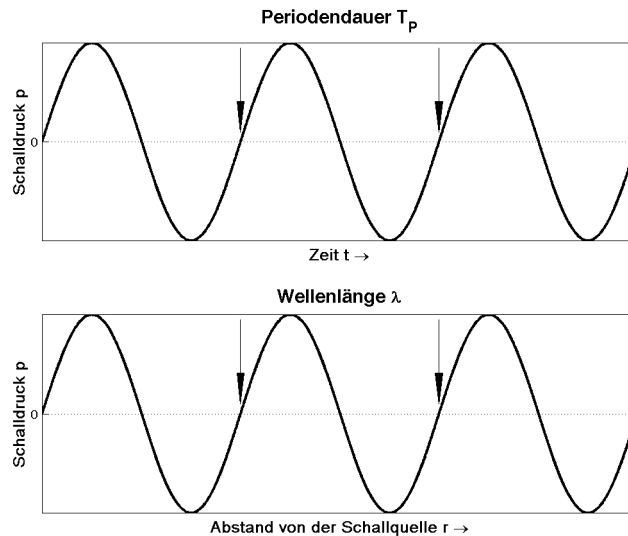


Abbildung 1.1: Periodendauer und Wellenlänge einer Schwingung

Signale aber oft in kleine Abschnitte aufteilen, in denen die Schwingung periodisch erscheint; dann spricht man von einer quasiperiodischen Schwingung.

Die nachfolgenden Größen beziehen sich auf periodische Schwingungen. Ihr Bezug zu realen Signalen wie Sprache und Musik beruht auf der Quasiperiodizität dieser Signale.

Die Punkte, an denen eine Sinusschwingung den Maximalausschlag, ihre Scheitelwerte, erreicht, nennt man auch **Schwingungsbauch**. Nimmt sie hingegen den Wert an, um den sie pendelt, so bezeichnet man diesen als **Schwingungsknoten** (in Abb. 1.1 der Nulldurchgang).

1.2 Frequenz und Periodendauer

Das Maß für die Häufigkeit, mit der sich eine periodische Schwingung wiederholt, ist die **Frequenz** f . Sie gibt die Anzahl der Schwingungsperioden pro Sekunde an.

$$f = \frac{\text{Anzahl der Schwingungen}}{s} \quad (1.1)$$

Die Einheit für die Frequenz ist somit die reziproke Sekunde $\left[\frac{1}{s}\right]$ bzw. Hertz $[Hz]$.

Der für Menschen hörbare Frequenzbereich liegt in etwa zwischen $20Hz$ und $20kHz$. Schwingungen mit niedriger Frequenz klingen tief, Schwingungen höherer Frequenz hoch.

Bemerkung

Das Tonhöhenempfinden ist *nicht* dem Betrag einer Frequenzänderung proportional, sondern dem Verhältnis der Änderung. Zum Beispiel hört man bei einem Sprung von 220Hz auf 440Hz eine Oktave, während man beim Sprung von 440Hz auf 660Hz eine Quinte hört, obwohl beide Male 220Hz zur unteren Frequenz addiert wurden. Ändert man aber die Frequenz von 440Hz auf 880Hz , so erklingt eine Oktave. Man hört also gleiche Intervalle, wenn das Verhältnis $\frac{\text{obere Frequenz}}{\text{untere Frequenz}}$ gleich ist.

Die Tabelle 1.2 gibt Anhaltspunkte über den Frequenzumfang einzelner Instrumente. Man erkennt, daß der hörbare Frequenzbereich um einiges über die von diesen Instrumenten erzeugten Grundschrwingungen hinausgeht. Wir würden nur ein sehr dumpfes, unvollständiges Klangbild erleben, wenn wir beispielsweise nur bis 5kHz hörten. Dies hängt mit den von Instrumenten erzeugten Obertönen zusammen, die ein wesentlich weiteres Frequenzspektrum umfassen können (s. Abschnitt 1.5).

Frequenz	Note	Instrument
16.5Hz	C_2	Taste C im 32' der Orgel
33Hz	C_1	C-Saite bei fünfseitigen Kontrabässen
66Hz	C	C-Saite der Violoncelli
131Hz	c	C-Saite der Bratschen
262Hz	c'	tiefstes c der Geigen
524Hz	c''	hohes c der Tenöre
1047Hz	c'''	hohes c der Soprane
2093Hz	c^4	höchstes c der Geigen
4185Hz	c^5	höchstes c der Piccolo-Flöten

Tabelle 1.1: Überblick über die Zuordnung von Frequenzwerten zu musikalischen Tonbezeichnungen (aus [Mey95])

Die Dauer einer Schwingungsperiode bezeichnet man als **Periodendauer** T_p . Sie ist der Kehrwert der Frequenz.

$$T_p = \frac{1}{f} \quad (1.2)$$

Die Einheit für die Periodendauer T_p ist damit die Sekunde [s].

Beispiel

Die Periodendauer der Schwingung einer Wasserwelle ist die Zeit, die ein Teilchen benötigt, das immer auf und ab schwingt, um von einem Wellental ins nächste zu gelangen.

Verwandte Themenbereiche

- Abschnitt 5.3

1.3 Amplitude

Die **Amplitude** einer Schwingung ist der „Scheitelwert“ einer Schwingung, also ihr maximaler Ausschlag. Dieser Begriff wird hauptsächlich für sinusförmige Schwingungen gebraucht. Im Zusammenhang mit andersförmigen Schwingungen wird allerdings oft für einen beliebigen Ausschlag der Begriff Amplitudenwert verwendet. Dies geschieht dann meistens zur Differenzierung z.B. zu Pegelwerten (s. Abschnitt 3).

Verwandte Themenbereiche

- Abschnitt 5.2
- Abschnitt 6.4.1

1.4 Phasenverschiebung

Verschiebt man eine periodische Schwingung zu sich selber und vergleicht dann die ursprüngliche mit der verschobenen Schwingung, so sind diese beiden Schwingungen bis auf die zeitliche Verschiebung Δt immer noch identisch. Für unterschiedliche Frequenzen liegen die beiden gleichen Schwingungen bei konstanter Verschiebung immer unterschiedlich zueinander. Aus diesem Grund gibt man die Verschiebung als Winkel an. Man setzt die Periodendauer gleich 360° und kann nun einen Winkel in Abhängigkeit von der zeitlichen Verschiebung und der Frequenz bestimmen:

$$\phi = \frac{\Delta t}{T_p} \cdot 360^\circ \quad (1.3)$$

Diesen Winkel nennt man dann **Phasenverschiebung** bzw. **Phasendifferenz** (s. Abb. 1.2), wenn man es ganz exakt ausdrücken will **Phasenwinkel**. Dieser Wert ist abhängig von dem Verhältnis von Verschiebung und Periodendauer und wird in der Praxis häufig verwendet.

Um von einer Phasenverschiebung zu sprechen, benötigt man immer einen Vergleichs- bzw. einen Bezugspunkt. Es ergibt keinen Sinn, eine Schwingung phasenverschoben zu nennen, wenn keine zweite Schwingung angegeben wird, zu der diese phasenverschoben ist. Es ist jedoch korrekt, von einer Phasendifferenz zwischen zwei Schwingungen zu sprechen, oder diese beiden Schwingungen zueinander (um einen Winkel) phasenverschoben zu nennen.

Bemerkung

Das Ohr ist nicht in der Lage, die absolute Phasenverschiebung einer Schwingung zu erkennen; es erkennt keinen Unterschied zwischen einer Schwingung und der zeitlich verschobenen Schwingung. Treten allerdings die ursprüngliche und die verschobene Schwingung gleichzeitig auf, so läßt sich durchaus ein deutlicher Unterschied wahrnehmen.

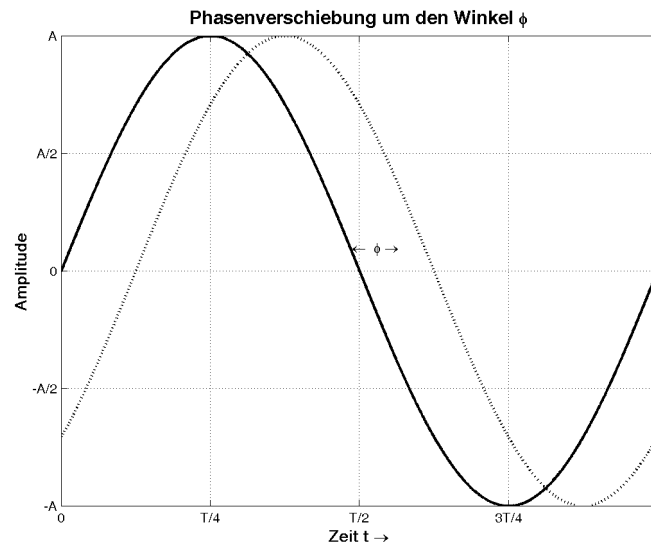


Abbildung 1.2: Phasenverschiebung um den Winkel ϕ bzw. um die Zeit Δt

Werden zwei oder mehrere Schallwellen gleicher Frequenz überlagert, so ergibt sich als resultierendes Signal je nach Phasenwinkel ein entweder verstärktes oder gedämpftes Signal gleicher Frequenz. Ein weiterer Spezialfall ist die Überlagerung von sinusförmigen Schwingungen leicht unterschiedlicher Frequenz. Diese führt zu einer sogenannten **Schwebung** (vgl. Abb. 1.3) mit einer leicht verschobenen hörbaren Frequenz und einer von der Differenzfrequenz abhängigen periodischen Amplitudenänderung. Eine solche Schwebung kann man leicht beim Stimmen eines Instrumentes hören; je näher die beiden Frequenzen sind, desto langsamer wird die hörbare Schwebung.

Herleitung

Die Addition zweier sinusförmiger Schwingungen $y(t) = x_1(t) + x_2(t)$ mit leicht unterschiedlicher Frequenz lässt sich beschreiben durch:

$$y(t) = \sin(2\pi f t) + \sin(2\pi (f - \Delta f) t) \quad (1.4)$$

Nach den Additionstheoremen für trigonometrische Funktionen ergibt sich nun

$$\begin{aligned} y(t) &= \sin\left(\frac{2\pi (f + f - \Delta f) t}{2}\right) \cdot \cos\left(\frac{2\pi [f - (f - \Delta f)] t}{2}\right) \\ &= \sin\left(2\pi \frac{f - \Delta f}{2} t\right) \cdot \cos\left(2\pi \frac{\Delta f}{2} t\right) \end{aligned} \quad (1.5)$$

Eine Überlagerung von Schwingungen wird **Interferenz** genannt.

Verwandte Themenbereiche

- Abschnitt [3.3](#)

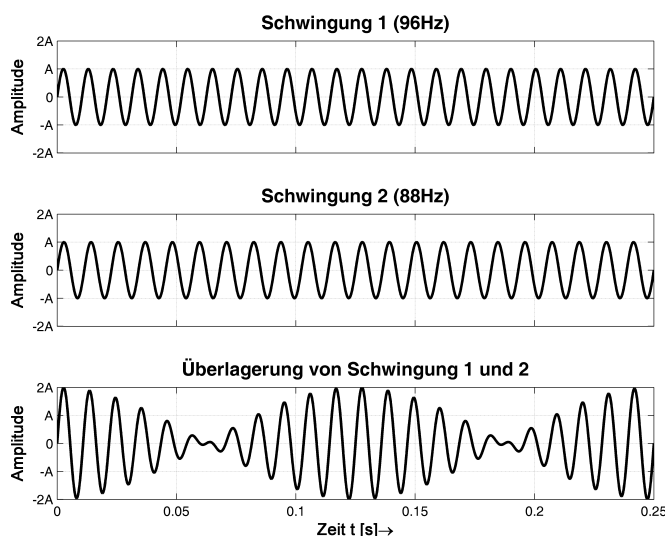


Abbildung 1.3: Schwebung aus der Addition (Überlagerung) zweier Schwingungen mit unterschiedlicher Frequenz

1.5 Töne, Klänge und Klangfarben

Tonmeister bzw. Musiker verwenden die Begriffe *Ton* und *Klang* anders als beispielsweise Akustiker. Um Verwirrungen vorzubeugen, soll hier kurz auf die beiden Begriffe eingegangen werden.

Im wissenschaftlichen Gebrauch¹ ist eine Sinusschwingung einer bestimmten Frequenz (im hörbaren Bereich) ein Ton. Ein einzelner Ton in der Musik ist (je nach Instrument) ein wesentlich komplexeres Gebilde, in jedem Fall aber ein Gemisch von mehreren oder vielen verschiedenen Frequenzen. Akustiker würden ein solches Frequenzgemisch wiederum Klang nennen.

Ein einzelner Ton eines beliebigen Musikinstrumentes zeichnet sich allgemein durch sogenannte harmonische **Obertöne** aus, deren Frequenz ein ganzzahliges Vielfaches der Grundfrequenz ist. Die Erhöhung der Frequenz jeweils um die Grundfrequenz führt dabei zu immer kleineren Intervallen. Aus der Obertonreihe lassen sich daher die Frequenzverhältnisse der Intervalle (genauer der „reinen“ Intervalle) herleiten. Tabelle 1.2 zeigt die Frequenzverhältnisse für die ersten Intervalle.

Bemerkung

Harmonische Obertöne spielen beim Hören eine sehr wichtige Rolle: aus der Struktur der harmonischen Obertöne kann der Hörer sogar selbst dann noch die Grundfrequenz hören, wenn diese physikalisch nur leise oder gar nicht im Schallereignis vorhanden ist.

¹nach DIN1320[DIN97]

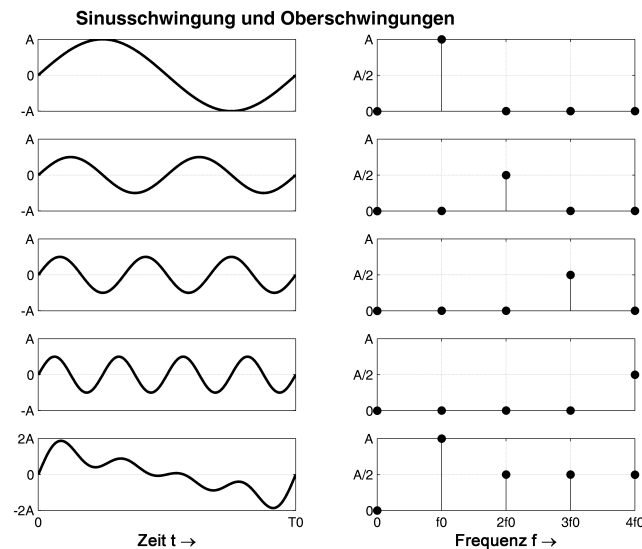


Abbildung 1.4: Sinusschwingung und harmonische Oberschwingungen im Zeitbereich (links) und im Frequenzbereich (rechts). Die unterste Darstellung zeigt die Addition der Einzelschwingungen ebenfalls im Zeit- und Frequenzbereich

Frequenzverhältnis f/f_0	reines Intervall
1/1	Prime
2/1	Oktave
3/2	Quinte
4/3	Quarte
5/4	gr. Terz
6/5	kl. Terz

Tabelle 1.2: Frequenzverhältnisse für reine Intervalle; die Frequenz einer reinen Quinte zur Grundfrequenz f_0 ist daher $\frac{3}{2}f_0$

Betrachtet man das Frequenzspektrum unterschiedlicher Instrumente, so erkennt man (unabhängig von dem gerade gespielten Ton) instrumentenspezifische Frequenzbereiche, die stärker angehoben sind als andere und erheblichen Einfluß auf die Klangfarbe eines Instrumentes haben. Solche Amplitudenmaxima innerhalb des Spektrums, deren Frequenz nicht von der gespielten Tonhöhe abhängt, nennt man **Formanten** oder **Formantbereiche**. Die Klangfarben unterschiedlicher Formanten lassen sich den Klangfarben von Vokalen zuordnen (s. Abb. 1.5). So klingen beispielsweise Instrumente mit einem starken Formanten u dunkel, während solche mit starkem Formanten i hell und spitz klingen.

Neben der Frequenzstruktur des Klanges darf aber auch die klangliche Wirkung des Zeitverlaufs nicht unterschätzt werden. Unterschiedliche Wirkungen

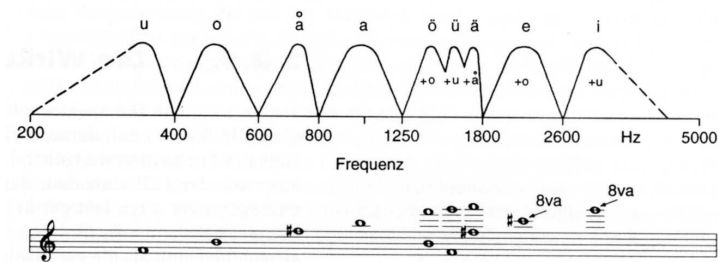


Abbildung 1.5: Frequenzlage der Formanten für die Vokale der deutschen Sprache (aus [Mey95])

verschiedener Musikinstrumente beruhen zu einem teilweise sehr hohen Maße auf der zeitlichen Struktur des Klanges. Einige wichtige Stichworte, die in späteren Abschnitten detaillierter erläutert werden sind z.B. die Hüllkurve und der Einschwingvorgang eines Klangereignisses.

Verwandte Themenbereiche

- Abschnitt [6.1.1](#), [6.1.2](#)
- Abschnitt [6.4.2](#)

1.6 Wellen

Nachdem weiter oben schon das Wort **Welle** gefallen ist, muß dieses noch richtig definiert werden. Innerhalb eines Mediums (Wasser, Luft, Metall etc.) lassen sich einzelne Moleküle zu Schwingungen anregen. Diese schwingenden Moleküle regen ihrerseits auch die benachbarten Moleküle zu Schwingungen an, die ihre Bewegung wiederum an die nächsten Teilchen weitergeben. Diese Art der Ausbreitung ist eine Welle.

1.7 Ausbreitungsgeschwindigkeit

Die Geschwindigkeit, mit der sich eine Welle in Ausbreitungsrichtung fortpflanzt, nennt sich **Ausbreitungsgeschwindigkeit** c . Wie jede Geschwindigkeit hat diese die Einheit $[\frac{m}{s}]$.

Beispiel

Die Ausbreitungsgeschwindigkeit einer Wasserwelle ist die Geschwindigkeit, mit der sich ein Wellenkamm in Ausbreitungsrichtung fortbewegt.

Die Ausbreitungsgeschwindigkeit ist abhängig von dem Medium, in dem sich die Welle fortpflanzt. Für Schallwellen wird sie als **Schallgeschwindigkeit** c

bezeichnet. In Luft ist die Schallgeschwindigkeit abhängig von Temperatur, Luftfeuchtigkeit, Luftdruck und Kohlendioxidgehalt. Sie ist *nicht* abhängig von der Frequenz. Während jedoch die letztgenannten Größen (Luftfeuchtigkeit, Luftdruck und Kohlendioxidgehalt) zu vernachlässigen sind, hat die Temperatur einen entscheidenden Einfluß (z.B. auf die Stimmung von Instrumenten mit schwingender Luftsäule wie Blasinstrumente oder Orgeln). Die Schallgeschwindigkeit läßt sich abhängig von der Temperatur mit folgendem Ausdruck berechnen:

$$c = 331.4 \frac{m}{s} \cdot \sqrt{\frac{\zeta + 273}{273}} \quad (1.6)$$

ζ ist die Temperatur in °C.

Für die meisten Überschlagsrechnungen wird ein Wert von $c = 340m/s$ angenommen.

Bemerkung

Für Zimmertemperatur ist der Wert $c = 340m/s$ etwas niedrig angesetzt. Nach Einsetzen in die Formel ergibt sich für eine Temperatur von 20°C $c = 343.3m/s$.

Bemerkung

Oft muß eingeschätzt werden, wie lange Schall braucht, um eine bestimmte Strecke zurückzulegen. Da man nicht immer einen Taschenrechner parat hat, ist es sinnvoll, sich zu merken, daß *1m ungefähr in 3ms zurückgelegt wird*.

1.8 Wellenlänge

Bei der Ausbreitung einer periodischen Schwingung in einem Medium treten in bestimmten Abständen in Ausbreitungsrichtung immer dieselben Schwingungszustände auf. Den Abstand benachbarter gleicher Schwingungszustände bezeichnet man als **Wellenlänge** λ mit der Einheit [m]. Die Wellenlänge muß somit von der Frequenz f und der Ausbreitungsgeschwindigkeit, mit der sich eine Welle in einem Medium fortpflanzt, abhängen. Man erhält

$$\lambda = \frac{c}{f} \quad (1.7)$$

Die Wellenlänge darf nicht mit der Periodendauer verwechselt werden; sie hängt zusätzlich noch von der Ausbreitungsgeschwindigkeit ab.

Beispiel

Die Ausbreitungsgeschwindigkeit einer Wasserwelle ist die Geschwindigkeit, mit der sich ein Wellenkamm in Ausbreitungsrichtung fortbewegt. Die Wellenlänge einer solchen Welle ist der Abstand zweier Wellenkämme in Ausbreitungsrichtung.

1.9 Ausbreitung von Wellen

Bei der Form der Ausbreitung einer Welle kann man zwischen zwei Grenzfällen unterscheiden, der **Kugelwelle** und der **Ebenen Welle**. Bei der Kugelwelle befinden sich alle Teilchen mit dem gleichen Abstand zur Schallquelle im gleichen Schwingungszustand, bei der Ebenen Welle befinden sich alle Teilchen auf einer Linie im gleichen Schwingungszustand. Von einer Ebenen Welle spricht man auch, wenn man sich weit von einer Kugelschallquelle entfernt befindet, so daß die Teilchen im gleichen Schwingungszustand sich näherungsweise auf einer Linie befinden (das gilt natürlich nur für einen begrenzten Ausschnitt aus der Kugel).

Beispiel

Wasserwellen am Strand sind näherungsweise Ebene Wellen, weil sich die gesamte Wellenfront in einer Linie auf den Strand zu bewegt. Wenn man in ruhiges Wasser einen Stein wirft, dann breiten sich kreisrunde Wellen aus. Hier könnte man von einer Kugelwelle sprechen.

Dieses Beispiel „hinkt“ insofern, als es sich auf zwei Dimensionen beschränkt. Übertragen auf den Schall im Medium Luft muß man das ganze auf drei Dimensionen abstrahieren: aus Kreisen werden also Kugeln, die sich ausdehnen.

Die Wasserwelle, bei der sich ein bestimmtes Teilchen immer auf und ab bewegt, ist sicherlich ein anschauliches Beispiel einer Welle. Man sollte dieses Beispiel aber nicht überstrapazieren, da einige Eigenschaften der Wasserwellen nicht mit denen von Schallwellen übereinstimmen. Vor allem schwingen die Teilchen in der Luft nicht auf und ab, sondern in Ausbreitungsrichtung des Schalls, d.h. wenn der Schall sich beispielsweise nach rechts ausbreitet, so schwingen die Teilchen auch von rechts nach links und nicht von oben nach unten. Eine solche Welle, bei der die Teilchen in Ausbreitungsrichtung schwingen, wird allgemein *longitudinale Welle* genannt.

1.10 Zusammenfassung

Eine Welle besteht aus Schwingungen einzelner Teilchen. Eine Schwingung wird gekennzeichnet durch die **Frequenz** f , welche die Anzahl an Schwingungen pro Sekunde angibt, bzw. die **Periodendauer** T_p , den Kehrwert von f und die Amplitude ihrer einzelnen Schwingungszustände. Bei gegeneinander verschobenen Schwingungen kann man die **Phasendifferenz** bzw. den **Phasenwinkel** angeben. Eine Welle kann prinzipiell beschrieben werden durch die **Wellenlänge** λ und die **Ausbreitungsgeschwindigkeit**. Die Ausbreitungsgeschwindigkeit hängt vom Medium und seinen Eigenschaften ab, bei Luft insbesondere auch von der Temperatur. Die Ausbreitungsgeschwindigkeit von Schall heißt **Schallgeschwindigkeit** c . Bei der Ausbreitungsart einer (Schall-)Welle wird zwischen den beiden Grenzfällen **Kugelwelle** und **Ebene Welle** unterschieden.

1.11 Aufgaben

1. In welcher Zeit legt Schall bei normaler Zimmertemperatur eine Strecke von $10m$ zurück?
2. Eine periodische Schwingung wird mit der gleichen um 180° verschobenen Schwingung addiert. Wie sieht das Ergebnis aus?
3. Was ist der Unterschied zwischen Periodendauer T_p und Wellenlänge λ ?
4. Berechne die Wellenlängen λ für folgende Frequenzen bei $c = 340m/s$.
 - $16Hz$
 - $100Hz$
 - $1000Hz$
 - $10000Hz$
 - $20000Hz$
5. Berechne die Frequenz des Tones a' für die angegebenen Temperaturen, wenn sie bei $14.35^\circ C$ $440Hz$ beträgt und die Wellenlänge konstant bleibt.
 - $16^\circ C$
 - $19^\circ C$
 - $22^\circ C$
6. Ist der oben berechnete Tonhöhenunterschied gravierend oder vernachlässigbar?

Kapitel 2

Schall und Schallfeld

Nachdem im vorigen Kapitel die Begriffe Schwingung und Welle erläutert wurden, wird dieses Kapitel die wichtigen Größen für Schall einführen, und auf die Besonderheiten der Schallausbreitung eingehen.

2.1 Schallschnelle und Schalldruck

Wie bereits in Abschnitt 1.6 beschrieben, breitet sich eine Schallwelle mittels Hin- und Herbewegung von Luftteilchen in Ausbreitungsrichtung aus. Die Geschwindigkeit, mit der die Teilchen sich bewegen, hat mit der Schallgeschwindigkeit c , mit der sich eine Welle fortpflanzt, *nichts* zu tun. Die Geschwindigkeit, mit der einzelne Teilchen um den Ort pendeln, an dem sie sich in Ruhe befänden, wird **Schallschnelle** v genannt.

Beispiel

Die Schallschnelle eines Teilchens ist *keine* konstante Größe. Sie ist an dem ursprünglichen Ruhepunkt des Teilchens am größten und bei der maximalen Auslenkung gleich Null. Zur Veranschaulichung kann man sich ein schwingendes Teilchen also wie ein schwingendes Pendel vorstellen. Das bedeutet auch, daß die Luftteilchen nicht mit der Welle mitwandern.

Die Luftteilchen schwingen in Ausbreitungsrichtung und regen ihrerseits Nachbarpartikel zum Schwingen an. Somit folgt eine Schwankung der Dichte und damit Änderungen des Drucks. Diese Druckänderungen bezeichnet man als **Schalldruck** p . Der Schalldruck wird vom Ohr wahrgenommen; das Ohr ist ein Schalldruckempfänger. Der Schalldruck hat einen Wert für einen bestimmten Punkt im Raum, die Schallschnelle hingegen hat zudem noch eine Richtung, weswegen man sie als Vektor schreibt.

Schallschnelle und Schalldruck sind Wechselgrößen, d.h. sie ändern sich ständig. Man möchte sie aber gerne durch jeweils einen einzelnen Wert ausdrücken. Hierfür hat sich der sogenannte **Effektivwert** (englisch: RMS für Root Mean Square) eingebürgert. Streng ist dieser definiert durch

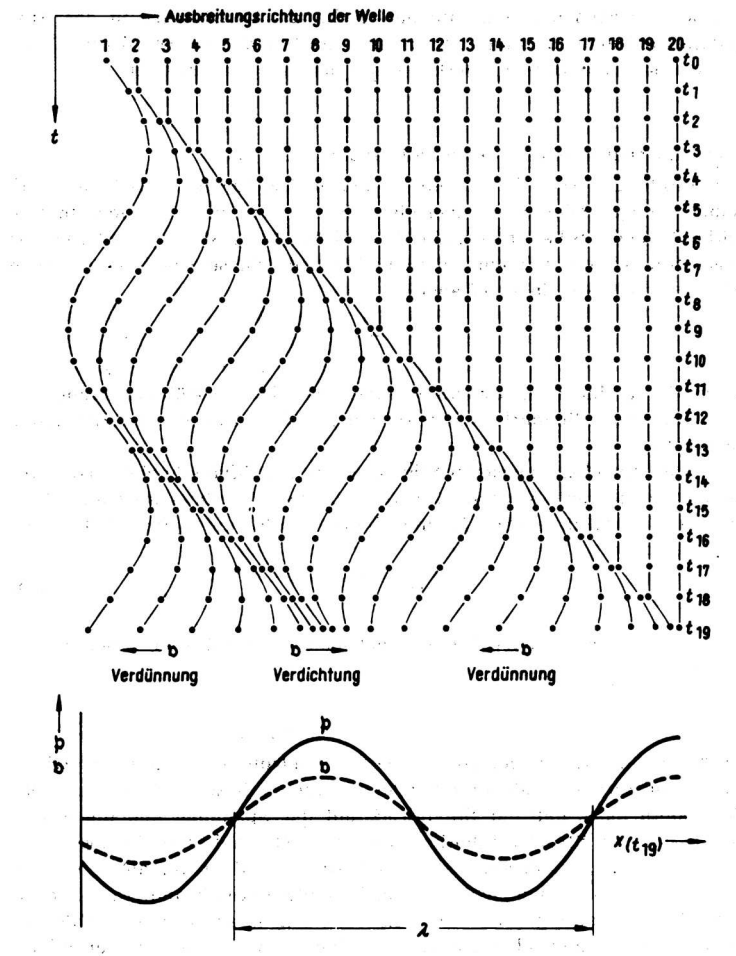


Abbildung 2.1: Ausbreitung einer Welle und Schalldruck (aus [Web94])

$$p_{eff} = \sqrt{\lim_{T \rightarrow \infty} \frac{1}{T} \int_T [p(t)]^2 dt} \quad [Pa], \quad (2.1)$$

für einfache periodische Signale (mit der Periodendauer T) kann man ihn aber auch durch

$$p_{eff} = \sqrt{\frac{1}{T} \int_T [p(t)]^2 dt} \quad [Pa] \quad (2.2)$$

berechnen. Ebenso läßt sich der Effektivwert der Schallschnelle ermitteln.

Bemerkung

Man meint, wenn von Schalldruck oder Schallschnelle als absoluten Größen die Rede ist, fast immer den Effektivwert (auch wenn das nicht explizit erwähnt wird), denn auf diese Weise lassen sich diese Wechselgrößen als einzelne Werte behandeln.

Bemerkung

Gegenüber dem statischen Luftdruck ist der Schalldruck außerordentlich klein. Der atmosphärische Druck beträgt $10^5 Pa$ ¹, während der Schalldruck bei mittlerer Lautstärke ca. $1/10 Pa$ beträgt. Der maximal erträgliche Schalldruck, die Schmerzschwelle, liegt bei ca. $100 Pa$.

2.2 Entfernungabhängigkeit des Schalldrucks

Der Schalldruck nimmt mit zunehmender Entfernung ab, eine weit entfernte Schallquelle wird wesentlich leiser wahrgenommen als eine nahe (wenn beide gleichstark abstrahlen). Es läßt sich zeigen, daß der Schalldruck umgekehrt proportional zur Entfernung r ist², d.h.

$$p \sim \frac{1}{r}. \quad (2.3)$$

Das bedeutet also, daß sich bei jeder Entfernungsverdopplung der Schalldruck halbiert. Diese Gesetzmäßigkeit nennt man oft **1/r-Gesetz**. Sie folgt direkt als Lösung der sogenannten Wellengleichung, die man für eine Kugelschallquelle aufstellen kann.

Bemerkung

Prinzipiell gilt dieses Gesetz genau wie die im folgenden Abschnitt 2.3 erläuterte Abhängigkeit der Schallschnelle von der Entfernung nur im sogenannten Freifeld (s. Abschnitt 4.3), d.h. nicht in normalen Räumen, sondern nur im Freien oder im reflexionsarmen Raum. Näherungsweise kann man aber auch in normalen Räumen von dieser Gesetzmäßigkeit ausgehen, allerdings nur in Nähe der Schallquelle (s. Abschnitt 4.3).

2.3 Nahfeld und Fernfeld/ Entfernungabhängigkeit der Schallschnelle

Als Lösung der Wellengleichung für die Schallschnelle ergibt sich eine etwas komplexere Beziehung als für den Schalldruck. Ist man weit von der Schallquelle entfernt, so nimmt der Betrag der Schallschnelle genauso wie beim Schalldruck mit $1/r$ ab. Nahe an der Schallquelle nimmt die Schnelle aber nicht nur mit $1/r$, sondern mit $1/r^2$ ab. Man unterscheidet je nach Verhalten der Schnelle zwischen **Nahfeld** ($1/r^2$ – Abnahme) und **Fernfeld** (Abnahme mit $1/r$).

Herleitung

Man erhält als Lösung der Wellengleichung für den Betrag der Schnelle

$$|\mathbf{v}| \sim \frac{1}{r} \sqrt{1 + \left(\frac{\lambda}{2\pi r} \right)^2} \quad (2.4)$$

¹ $1 Pa = 1 \frac{N}{m^2}$

²Dies gilt für Kugelschallquellen. Für andere Schallquellen nimmt der Schalldruck bis zu einer bestimmten Entfernung langsamer ab, und verhält sich dann wie im Kugelschallfeld.

r ist der Abstand zur Kugelschallquelle,
 λ ist die Wellenlänge des abgestrahlten Schalls

Der Betrag der Schallschnelle ist also sowohl von der Entfernung, als auch von der abgestrahlten Wellenlänge abhängig. Man erkennt, wenn man für r große Werte einsetzt, daß die Summe unter der Wurzel sehr nahe bei 1 liegt. Dies bedeutet, daß man für große Werte von r (d.h. große Abstände von der Schallquelle) die Beziehung

$$|\mathbf{v}| \sim \frac{1}{r} \sqrt{1} \quad (2.5)$$

erhält, also die gleiche Abnahme wie beim Schalldruck. Setzt man für r allerdings sehr kleine Werte ein, so kann man die 1 der Summe unter der Wurzel vernachlässigen und erhält

$$|\mathbf{v}| \sim \frac{1}{r^2}. \quad (2.6)$$

Den Übergang zwischen Nahfeld und Fernfeld findet man, wenn der zweite Summand $\frac{\lambda}{2\pi r}$ die gleiche Größenordnung hat wie der erste Summand, also ungefähr bei $\frac{\lambda}{2\pi r} = 1$. Der Punkt des Übergangs ist also frequenzabhängig, oder anders ausgedrückt: die Ausdehnung des Nahfeldes ist abhängig von der abgestrahlten Wellenlänge.

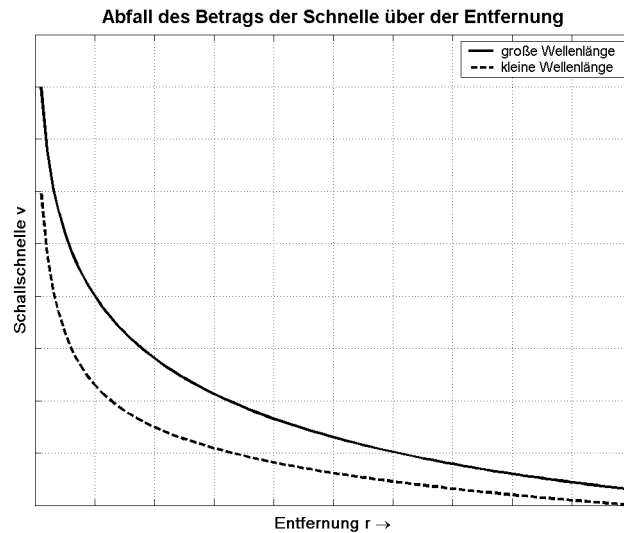


Abbildung 2.2: Verlauf der Schallschnelle über dem Abstand r von der Schallquelle für zwei verschiedene Frequenzen

Bei tiefen Frequenzen muß man wesentlich weiter von der Schallquelle entfernt sein, um sich im Fernfeld zu befinden, als für höhere. Den Übergang von Nah- zu Fernfeld findet man, wie in der Herleitung erwähnt, im Abstand

$$r = \frac{\lambda}{2\pi}. \quad (2.7)$$

Ein weiterer Unterschied zwischen Nah- und Fernfeld ist von Bedeutung. Im Fernfeld befinden sich Schalldruck und Schallschnelle in Phase, während sie im Nahfeld um bis zu 90° zueinander phasenverschoben sind.

Bemerkung

Für Mikrophone, die keine reinen Schalldruckempfänger sind, sondern auf die Schallschnelle ansprechen, ist das Wissen um Nah- und Fernfeld einer Schallquelle von Bedeutung. Schnelleempfänger liefern einerseits eine um 90° zu Druckempfängern verschobene Ausgangsspannung, andererseits erhält man bei solchen Mikrophenen im Nahfeld eine sehr starke Anhebung der tiefen Frequenzen. Dies wird *Nahbesprechungseffekt* genannt. Dazu in Abschnitt ?????? mehr.

2.4 Schalleistung

Die Stärke des Schalldrucks an einem bestimmten Punkt im Raum ist durch die „Stärke“ der Schallquelle bestimmt. Es ist von Interesse, die Stärke einer Schallquelle unabhängig vom Abstand des Hörers und den räumlichen Gegebenheiten angeben zu können. Die **Schalleistung** P_{ak} , also die abgestrahlte Energie pro Sekunde, ist eine Größe, die dies ermöglicht. Ihre Einheit ist Watt [W]. Sie ist

$$P = p_{eff}^2 = \lim_{T \rightarrow \infty} \frac{1}{T} \int_T [p(t)]^2 dt \text{ [W]}. \quad (2.8)$$

2.5 Schallintensität

Die **Schallintensität** ist die Schalleistung, die durch eine senkrecht zur Ausbreitungsrichtung stehende Fläche strömt.

$$\mathbf{J} = \frac{1}{T} \int_0^T p(t) \cdot \mathbf{v}(t) dt \left[\frac{W}{m^2} \right] \quad (2.9)$$

Für eine Kugelschallquelle gilt für die Intensität in Abhängigkeit des Abstandes r

$$J(r) = \frac{P_{ak}}{4\pi r^2} \left[\frac{W}{m^2} \right] \quad (2.10)$$

2.6 Gestörte Schallausbreitung

In den vorigen Abschnitten wurde davon ausgegangen, daß die Schallwelle sich ungestört ausbreiten kann, also auf keine Hindernisse stößt. In Räumen gibt es eine Vielzahl von Hindernissen, die den Schall an ungestörter Ausbreitung hindern, insbesondere natürlich die Wände. Je nach Art und Beschaffenheit des Hindernisses entstehen verschiedene Phänomene. Im folgenden soll kurz auf diese eingegangen werden.

2.6.1 Reflexion

Trifft eine Schallwelle auf eine ebene Fläche, die groß im Verhältnis zu ihrer Wellenlänge ist, so wird sie reflektiert. Hierfür gilt - genau wie in der Optik - das bekannte Gesetz Einfallswinkel = Ausfallswinkel, das besagt, daß die unter einem Winkel θ (die 0° -Achse steht senkrecht auf der Wand) auf die Wand treffende Welle mit einem Winkel $-\theta$ wieder reflektiert wird. Schall verhält sich hier also wie ein Strahl (s. Abb. 2.3).

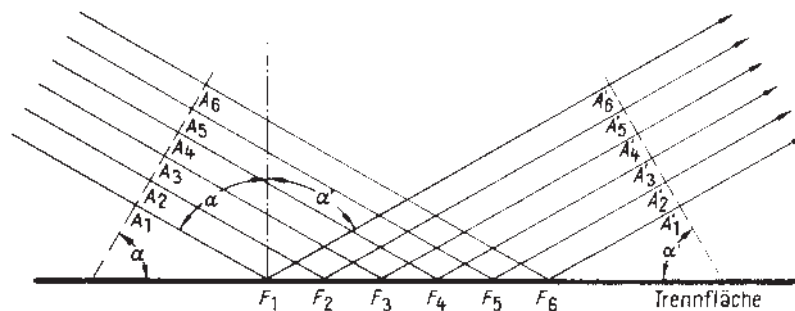


Abbildung 2.3: Reflexion einer Ebenen Welle mit Einfallswinkel α an einer Wand (aus [Web94])

Je nach Material der Wand sind die Reflexionen allerdings nicht mehr mit dem eintreffenden Schall identisch. Vielmehr reflektieren verschiedene Wandmaterialien verschiedene Frequenzen unterschiedlich gut. Beispielsweise reflektiert eine harte Steinwand den eintreffenden Schall fast verlustlos, während eine mit Stoff bespannte Wand hohe Frequenzen dämpfen wird.

Die Eigenschaft eines (Wand-)materials, nicht die gesamte Schallenergie zu reflektieren, sondern nur einen Teil, kann mit dem **Absorptionsgrad** α beschrieben werden. Der Absorptionsgrad eines Materials, das keinen Schall reflektiert, sondern alles absorbiert, ist $\alpha = 1$; für eine verlustfrei reflektierende Wand wäre $\alpha = 0$. Für reale Materialien kann der Absorptionsgrad also Werte zwischen 0 und 1 annehmen.

Bemerkung

Bei verlustbehafteter Reflexion (s.o.) spricht man von *Schalldämpfung*. Wenn man sich dagegen beispielsweise nicht im gleichen Raum wie das Schallereignis aufhält, sondern im Nebenraum, so erreicht den Hörer nur die Schallenergie, die weder reflektiert noch absorbiert, sondern durchgelassen wurde. Diese Eigenschaft einer Wand bezeichnet man als *Schalldämmung*.

Verwandte Themenbereiche

- Abschnitt 4

2.6.2 Beugung

Bei der Reflexion von Schall war die Voraussetzung, daß die reflektierende Fläche groß gegenüber der Wellenlänge ist. Ist die Wellenlänge hingegen in der gleichen Größenordnung wie die Fläche (resp. ihre kleinste Ausdehnung), so werden die Schallwellen nicht mehr reflektiert, sondern gebeugt. **Beugung** bedeutet, daß die Schallwellen sich um das Hindernis „herumbiegen“. Ist das Hindernis also klein gegenüber der Wellenlänge, so treten also dahinter kaum Abschattungswirkungen auf, sondern die Schallwelle pflanzt sich hinter dem Hindernis fast genau wie vor dem Hindernis fort (s. Abb. 2.4).



Abbildung 2.4: Beugung von Wellen hinter einem Hindernis: (a) Hindernis groß im Vergleich zu der Wellenlänge, (b) Hindernis gleich groß wie die Wellenlänge (aus [Web94])

2.7 Zusammenfassung

Schall breitet sich als *longitudinale Welle* aus. Die Luftteilchen schwingen um ihren Ruhepunkt mit der **Schallschnelle** v und erzeugen durch ihre Bewegung Druckschwankungen, den **Schalldruck** p . Beide Größen sind Wechselgrößen und werden zumeist nur mit ihrem *Effektivwert* angegeben. Seltener werden in der Tonstudientechnik für Berechnungen oder Angaben auch die **Schalleistung** P_{ak}

oder die **Schallintensität J** verwendet.

Der Schalldruck nimmt bei Entfernung von der Schallquellen mit $1/r$ ab, dies besagt das sogenannte **$1/r$ -Gesetz**. Im **Fernfeld** nimmt die Schallschnelle ebenfalls mit $1/r$ über der Entfernung ab. Im sogenannten **Nahfeld** nimmt sie hingegen mit $1/r^2$ ab und ist zum Schalldruck um maximal 90° in der Phase verschoben. Der Übergang von Nah- zu Fernfeld ist abhängig von der Frequenz bzw. der Wellenlänge. Das Nahfeld ist für tiefe Frequenzen ausgedehnter als für hohe.

Schallwellen werden an Gegenständen **reflektiert**, die groß im Vergleich zu ihrer Wellenlänge sind. Je nach Wandmaterial werden sie frequenzabhängig gedämpft. Ist das Hindernis dagegen in der gleichen Größenordnung wie die Wellenlänge, so wird die Welle um das Hindernis herum **gebeugt**.

2.8 Aufgaben

1. Berechne für folgende Frequenzen den Abstand, an welchem das Nahfeld in das Fernfeld übergeht
 - 50Hz
 - 200Hz
 - 500Hz
 - 1000Hz
 - 4000Hz
2. Um welchen Faktor verringert sich der Schalldruck, wenn man statt in 30cm Entfernung von der Schallquelle in 180cm Entfernung mißt?
3. Schallquelle und Zuhörer befinden sich im gleichen Abstand zu einer Wand, die 3m von beiden entfernt ist. Quelle und Hörer sind 5m voneinander entfernt. Wie viele Millisekunden nach dem Direktschall trifft die Reflexion von dieser Wand ein? Um welchen Faktor ist sie gegenüber dem Direktschall abgeschwächt, wenn die Wand verlustfrei reflektiert?
4. In kurzem Anstand vor dem Mikrophon befindet sich ein Pult der Abmessungen $30 \times 40\text{cm}$. Ab welcher Frequenz ist mit sehr deutlichem Pegelabfall zu rechnen?

Kapitel 3

Pegel

Bei der Arbeit im Tonstudio geht es kaum um die Größen Schalldruck oder gar Schallschnelle. Vielmehr taucht überall der Begriff Dezibel oder dB auf. Dieses Kapitel erklärt den Begriff Pegel in Dezibel und erläutert einige wichtige Grundsätze im Umgang mit Pegeln.

3.1 Das Weber-Fechnersche Gesetz

Inzwischen sind zwar die grundlegenden Kenngrößen eines Schallfeldes bekannt, aber bisher wurde kaum darauf eingegangen, wie diese Größen überhaupt vom Menschen wahrgenommen werden. Das sogenannte **Weber-Fechnersche Gesetz** stellt einen Zusammenhang zwischen der Stärke eines Reizes (in unserem Fall beispielsweise der Schalldruck) und der Stärke der zugehörigen Sinnesempfindung (entsprechend: empfundene Lautstärke) her. Weber und Fechner stellten ganz allgemein für die Reizwahrnehmung des Menschen fest, daß der gerade wahrnehmbare Zuwachs W von dem Verhältnis der Änderung eines Reizes R zur Stärke des schon vorhandenen Reizes abhängig ist:

$$dW = \text{const.} \cdot \frac{dR}{R} \quad (3.1)$$

Nach beidseitiger Integration dieser Differentialgleichung erhält man die Wahrnehmung als logarithmische Funktion irgendeines Reizes, d.h. - jetzt auf die Lautstärkeempfindung bezogen -, daß eine Verdoppelung des Schalldrucks **nicht** als eine Verdoppelung der Lautstärke wahrgenommen wird. Vielmehr müßte der Schalldruck exponentiell steigen, um eine lineare Erhöhung der empfundenen Lautstärke herbeizuführen.

Bemerkung

Leider stimmen die dem Gesetz zugrundegelegten Annahmen nicht ganz; in Wirklichkeit gilt dieses Gesetz nicht streng. Es stellt jedoch eine hinreichend gute Näherung an die Wirklichkeit dar.

3.2 Schalldruckpegel & Co.

In Anlehnung an das Weber-Fechnersche Gesetz wurde eine Pseudoeinheit namens **Dezibel**¹ [dB] eingeführt, die ein logarithmisches Maß darstellt. Werte mit der Einheit dB werden Pegel genannt. Durch den Übergang von der Einheit [Pa] zur Pseudoeinheit [dB] ist eine mehr oder weniger gute Anpassung an das menschliche Hörvermögen möglich.

Eine sehr häufig gebrauchte Größe ist in diesem Zusammenhang der **Schalldruckpegel**, der definiert ist mit

$$L_{SPL} = 10 \log \left(\frac{p_{eff}^2}{p_{0,eff}^2} \right). \quad (3.2)$$

Die geläufigste Schreibweise ist

$$L_{SPL} = 20 \log \left(\frac{p_{eff}}{p_{0,eff}} \right). \quad (3.3)$$

In dieser Gleichung taucht nicht nur der effektive Schalldruck p_{eff} ² auf, sondern auch der sogenannte *Bezugsschalldruck* $p_{0,eff}$. Dieser stellt eine Normierung dieser Funktion dar ($L_{SPL} = 0dB$ für $p = p_0$), da der Logarithmus von 1 *Null* ergibt. Für den absoluten Schalldruckpegel ist $p_{0,eff} = 20\mu Pa$ ³. Dies entspricht ungefähr der sogenannten Hörschwelle, also dem Wert des Schalldrucks, den man (bei mittleren Frequenzen) gerade wahrnehmen kann.

Umgekehrt berechnet sich der effektive Schalldruck aus dem Pegel mit

$$p_{eff} = p_{0,eff} \cdot 10^{\frac{L_{SPL}}{20}}, \quad (3.4)$$

was sich durch Umstellen der Gleichung (3.3) zeigen läßt.

Natürlich lassen sich auch andere Größen (mit anderen Bezugswerten) genau wie der Schalldruck als Pegel ausdrücken. Gebräuchlich sind auch noch der **absolute Leistungspegel**, der definiert ist mit

$$L_P = 10 \log \left(\frac{P}{P_0} \right) \quad (3.5)$$

und der **absolute Spannungspegel**. Dieser ist definiert mit

$$L_U = 10 \log \left(\frac{U_{eff}^2}{U_{0,eff}^2} \right) \quad (3.6)$$

¹dB ist keine richtige Einheit im physikalischen Sinn, da sich alle Einheiten bei der Rechnung herauskürzen

²Der Schalldruck ist eine Wechselgröße; da eine statische Größe zum Berechnen des Pegels eingesetzt werden muß, wird der quadrierte Effektivwert (= Leistung) gewählt. Zur Erinnerung: die Leistung $P = p_{eff}^2 = \lim_{T \rightarrow \infty} \frac{1}{T} \int_T [x(t)]^2 dt$

³Im folgenden wird, wie allgemein üblich, lediglich vom Schalldruckpegel gesprochen, wenn der absolute Schalldruckpegel gemeint ist.

Kurzzeichen	Bezeichnung	Bezugswert
dB, dBSPL	absoluter Schalldruckpegel	$20\mu Pa$
dB	alle relativen Pegel	kein Bezugswert
dB, dBSIL	absoluter Schallintensitätspegel	$10^{-12} \frac{W}{m^2}$
dB,dBSWL	absoluter Schalleistungspegel	$10^{-12} W$
dB,dBSVL	absoluter Schallschnellepegel	$5 \cdot 10^{-8} \frac{m}{s}$
dB(A),dB(B),dB(C)	absoluter Schalldruckpegel mit Filterkurven bewertet	$20\mu Pa$
dB	Funkhausnormpegel	$1.55V$
dBu	absoluter Spannungspegel	$0.775V$
dBv, dBV	absoluter Spannungspegel	$1V$
dBm	absoluter Leistungspegel	$1mW$
dbFs	Pegel bezogen auf Full Scale (digitale Geräte)	full scale

Tabelle 3.1: Übersicht über verschiedene Pegelwerte und deren Bezugswerte

beziehungsweise

$$L_U = 20 \log \left(\frac{U_{eff}}{U_{0,eff}} \right). \quad (3.7)$$

Bemerkung

Der Logarithmus in den oben angeführten Definitionen erscheint nach dem Weber-Fechner-Gesetz sinnvoll zu sein. Der Faktor 10 ist allerdings auf Anhieb nicht ganz zu verstehen.

Die Pseudo-Einheit *Bel* (nach dem Ingenieur Bell benannt) wurde festgelegt mit $\log \frac{x}{x_0}$. Als die Akustiker feststellten, daß das durchschnittliche subjektive Unterscheidungsvermögen bei ca. 1/10 dieses Wertes lag, wurde die heute gebräuchliche „Einheit“ Dezibel (*dB*) eingeführt, die einem Zehntel *Bel* entspricht.

Alle relativen Pegel (z.B. Differenzen zwischen zwei absoluten Pegeln) haben zwar keinen Bezugswert (wie z.B. $p_{0,eff}$), aber trotzdem die Einheit *dB*. Bei der Differenzbildung zweier Pegel (mit gleichen Bezugswerten) kürzt sich der Bezugswert weg.

Anhand der Tabelle 3.2, die einige gebräuchliche Pegel aufführt, wird deutlich, daß fast beliebige Größen durch Pegel ausgedrückt werden können.

Bemerkung

Immer, wenn mit der „Einheit“ *dB* gearbeitet und eine absolute Größe beschrieben wird, sollte man genau darauf achten, auf welchen Bezugswert diese sich bezieht. Für verschiedene Bezugswerte ergeben sich vollkommen unterschiedliche Pegel.

In Abb. 3.1 ist die typische **Hörfläche** abgebildet. Die Hörfläche ist die Fläche zwischen der *Ruhehörschwelle* und der *Schmerzgrenze*. Die Ruhehörschwelle ist die Grenze, ab der man Schallereignisse geringeren Schalldruckpegels nicht mehr wahrnehmen kann und die Schmerzgrenze diejenige, ab der man lautere Schallereignisse nur noch unter Schmerzen wahrnehmen kann. Interessant ist in dieser Abbildung auch die unterschiedliche Empfindlichkeit des Ohres für verschiedene Frequenzen (s. dazu Abschnitt 5.2).

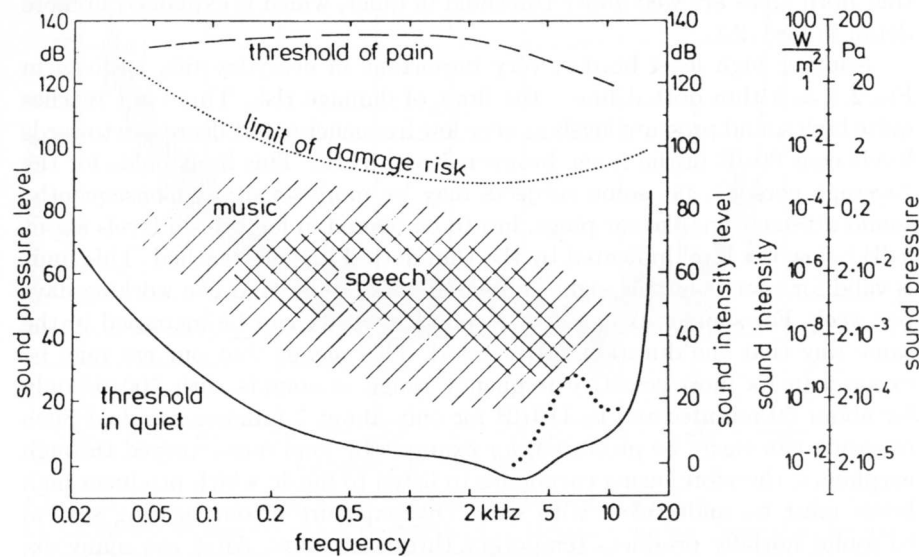


Abbildung 3.1: Hörfläche, d.h. Fläche zwischen Ruhehörschwelle und Schmerzgrenze. Zusätzlich ist zur Orientierung die typische Fläche für Sprache und Musik skizziert (aus [ZF99])

Bemerkung

Neben den oben genannten Größen ist in Abb. 3.1 auch die Risikogrenze für Gehörschäden angegeben. Diese Größe geht allerdings von einer dauerhaften Belastung (8 Stunden am Tag, 5 Tage die Woche) aus; der Schalldruckpegel kann etwas erhöht werden, wenn die Dauer der Belastung verkürzt wird. Wird das Ohr hohen Belastungen ausgesetzt, so kommt es zu einer vorübergehenden Verschiebung der Grenzen der Hörfläche zu höheren Pegeln. Geschieht dies aber öfters, so wird die temporäre Verschiebung zu einer permanenten, also einem Gehörschaden.

Verwandte Themenbereiche

- Abschnitt 1.3
- Abschnitt 5.2

3.3 Pegel bei der Addition von Signalen

Bis jetzt wurde noch nicht behandelt, wie hoch der Pegel eines Signals $u(t) = x(t) + y(t)$ ist, das die Summe zweier Signale darstellt. Wie in Abschnitt 3.2 ausgeführt, steht im Argument des Logarithmus ein Effektivwert. Leider lassen sich Effektivwerte resp. Leistungen nicht so einfach addieren, wie man das vielleicht erwarten könnte. Es zeigt sich nämlich, daß es bei der Addition eine Rolle spielt, ob sich die Signale, welche addiert werden, *ähnlich* sind oder nicht.

Zur Erläuterung müssen erst einige Begriffe eingeführt werden, die leider anschaulich schwer zu fassen sind.⁴ Zwei Signale heißen kohärent, wenn ein Signal nur gegen sich selber verschoben und skaliert wurde, beispielsweise wenn sie von der gleichen Quelle kommen. Laufzeit- bzw. Pegelunterschiede spielen also keine Rolle.

Zwei Signale heißen hingegen inkohärent, wenn sie einander zu keinem Zeitpunkt in irgendeiner Weise ähnlich sind.

Die Korrelation gibt die Verwandtschaft zweier Signale zu einem bestimmten Zeitpunkt pegelunabhängig an. Signale sind also korreliert, wenn sie bis auf einen Skalierungsfaktor übereinstimmen. Der sogenannte Korrelationsgrad zeigt die Ähnlichkeit auf einer Skala zwischen $[-1; 1]$ an; dabei steht „1“ für Gleichheit, „-1“ für Gegenphasigkeit und „0“ für Unkorreliertheit.

Dies bedeutet einerseits, daß inkohärente Signale immer unkorreliert sind, andererseits korrelierte Signale immer kohärent sind. Umgekehrt trifft dies nicht allgemein zu.

Beispiel

Rauschen ist zu allen Signalen (außer zu sich selbst) inkohärent und damit auch unkorreliert.

Eine Sinus- und eine Cosinusschwingung sind zwar unkorreliert zueinander, aber kohärent.

Die Ausgangssignale zweier Mikrophone sind (wenn die Mikrophone nicht zu nah beieinander stehen) näherungsweise unkorreliert.

Bemerkung

In der Technik und Mathematik wird oft der Begriff *orthogonal* statt *nicht korreliert* verwendet. Die Bedingung für Unkorreliertheit bzw. Orthogonalität zweier Signale $x(t), y(t)$ ist $\int x(t) \cdot y(t) dt = 0$

Man stelle sich nun die zu addierenden Signale als Vektoren vor (s. Abb. 3.2), deren Betrag ihr Effektivwert ist. Sind zwei Signale identisch, so haben die zugehörigen Vektoren den gleichen Betrag und die gleiche Orientierung. Sind sie hingegen unkorreliert, so stehen die Vektoren senkrecht aufeinander (90° Phasenverschiebung!). Das gewünschte Ergebnis, nämlich der Effektivwert des resultierenden Signals, ist der Betrag des aus der Addition resultierenden Vektors.

Für die Addition zweier gleicher Signale x erhält man also

$$u_{Ges,eff} = 2 \cdot x_{eff} \quad (3.8)$$

Der Pegel des resultierenden Signals nimmt also um $20 \log 2 = 6 \text{ [dB]}$ zu.

Bei der Addition zweier nicht korrelierter Signale x und y erhalten wir

$$u_{Ges,eff} = \sqrt{x_{eff}^2 + y_{eff}^2} \quad (3.9)$$

⁴Es kann sein, daß diese Begriffe an anderer Stelle anders eingeführt werden. Es wurde versucht, die Definition der beiden Begriffe Inkohärenz und Korrelation möglichst aus dem gängigen Gebrauch in der Aufnahmepraxis herzuleiten. Aus diesem Grund muß man sich bei der Lektüre verschiedener Literatur immer vergewissern, in welchem Sinn die Begriffe Kohärenz und Korrelation benutzt werden.

Addition gleicher Signale Addition unkorrelierter Signale



Abbildung 3.2: Veranschaulichung der Addition von Effektivwerten

Für $x_{eff} = y_{eff}$ erhält man hier eine Zunahme um $20 \log \sqrt{2} = 3 \text{ [dB]}$! Bei der Addition von unkorrelierten Signalen *gleichen* Pegels nimmt der Pegel also um 3 dB zu.

Beispiel

Das Rauschen zweier Tonbandspuren, die mit gleichem Pegel gemischt werden, ist um 3 dB lauter als das einer Spur.

Werden nun mehrere nicht korrelierte Quellen *gleichen* Pegels addiert, so erhält man analog dazu:

$$u_{Ges,eff} = \sqrt{u_{1,eff}^2 + u_{2,eff}^2 + \dots + u_{n,eff}^2} \quad (3.10)$$

und der Gesamtpegel nimmt dementsprechend um $20 \log \sqrt{n} = 10 \log n \text{ [dB]}$ zu. Haben die Quellen nicht denselben Pegel, so muß man in die obigen Gleichungen die entsprechenden Effektivwerte einsetzen und hat etwas mehr Arbeit. Die Rechnung ist aber im Prinzip identisch.

Bemerkung

Es gibt natürlich beliebige Zwischenstufen zwischen „unkorreliert“ und „identisch“ (s. Abb 3.3). In einem solchen Fall stehen die Vektoren dann nicht mehr senkrecht aufeinander und der resultierende Pegel wird bei der Addition von zwei Signalen zwischen 3 und 6 dB liegen.

Für die Leser, die es etwas genauer wissen wollen, folgt eine etwas detailliertere Herleitung.

Herleitung

Um die erhaltenen Resultate zu überprüfen, muß die Definition des Effektivwertes betrachtet werden, die ja lautete

$$u_{eff} = \sqrt{\lim_{T \rightarrow \infty} \frac{1}{T} \int_T [u(t)]^2 dt}. \quad (3.11)$$

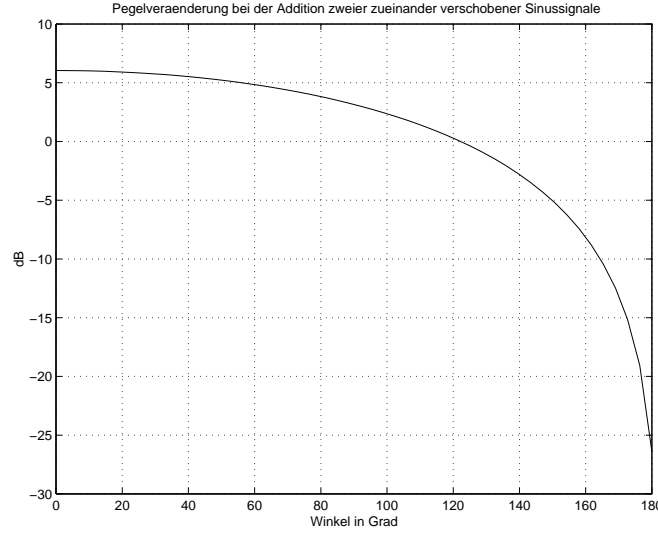


Abbildung 3.3: Pegeländerung bei der Addition zweier zueinander verschobener Sinussignale. Für einen Phasenwinkel von 0° ergeben sich 6dB Pegelzunahme, während der Pegel des resultierenden Signals für einen Winkel von 180° gegen $-\infty$ geht

Dementsprechend gilt also für den Effektivwert, wenn $u(t)$ die Summe zweier Signale $x(t)$ und $y(t)$ ist

$$\begin{aligned}
 u_{eff} &= \sqrt{\lim_{T \rightarrow \infty} \frac{1}{T} \int_T [x(t) + y(t)]^2 dt} \\
 &= \sqrt{\lim_{T \rightarrow \infty} \frac{1}{T} \int_T [x(t)^2 + y(t)^2 + 2x(t)y(t)] dt}. \quad (3.12)
 \end{aligned}$$

Für den Fall, daß $x(t) = y(t)$ ist, erhält man

$$\begin{aligned}
 u_{eff} &= \sqrt{\lim_{T \rightarrow \infty} \frac{1}{T} \int_T [2 \cdot x(t)]^2 dt} \\
 &= 2 \cdot x_{eff}, \quad (3.13)
 \end{aligned}$$

also das Ergebnis, das nach Gleichung 3.8 schon erwartet werden konnte.

Sind x und y nicht korreliert, so gilt für die Gleichung 3.12, daß das Integral über den gemischten Term gleich 0 ist (dies gilt aufgrund der Orthogonalitätseigenschaft, die als mathematische Definition der Unkorreliertheit gegeben wurde). Mit Hilfe der Orthogonalitätseigenschaft unkorrelierter Signale folgt also

$$u_{eff} = \sqrt{x_{eff}^2 + y_{eff}^2} \quad (3.14)$$

und damit Gleichung 3.9.

Bisher ging es um die *elektrische* Addition von Signalen. Sollen diese *akustisch* addiert werden, so ergeben sich andere Pegelunterschiede. Im allgemeinen sind

Schallwellen im Raum bei zwei Sendern nicht mehr korreliert, auch wenn die Sender selbst das gleiche Signal abstrahlen. Dementsprechend erhöht sich der Schalldruckpegel bei zwei Lautsprechern, welche das gleiche Signal abstrahlen, um $3dB$.

Beispiel

Zwei Lautsprecher, die mit dem gleichen Signal gespeist werden, erzeugen einen Schalldruckpegel, der um $3dB$ höher ist als der von einem einzigen Lautsprecher erzeugte Pegel.

Zwei gleiche und gleichlaute Musikinstrumente spielen unisono $3dB$ lauter als ein einzelnes.

3.4 Zusammenfassung

Es wurde eine (Pseudo-)Einheit **Dezibel** eingeführt. Mit dieser Einheit lassen sich viele Größen dem menschlichen Empfindungsvermögen aufgrund des Weber-Fechnersche Gesetzes anpassen. *Dezibel* ist keine richtige Einheit, sondern kann sowohl für Pegeldifferenzen (relative Pegel) als auch für absolute Größen (bezogen auf einen Bezugswert) stehen. Eine besonders wichtige Größe für Tonmeister ist der **Schalldruckpegel**. Dessen Bezugsgröße ist $2 \cdot 10^{-5} Pa$.

Jeder absolute Pegel ist auf eine bestimmte Bezugsgröße normiert. Nur bei Kenntnis dieser Bezugsgröße ist eine Angabe eines beliebigen Wertes in dB sinnvoll.

Der resultierende Pegel bei der Addition von Signalen ist *nicht* die Summe der Einzelpegel. Um den resultierenden Pegel berechnen zu können, muß zuerst der resultierende Effektivwert errechnet werden, daraufhin kann auch der Pegel bestimmt werden. Bei der elektrischen Addition zweier Signale gleichen Pegels ergibt sich $6dB$, falls die Signale korreliert sind, und $3dB$ im unkorrelierten Fall. Bei der akustischen Addition zweier Signale ergibt sich hingegen i.a. eine Pegelerhöhung von $3dB$.

3.5 Aufgaben

Vor den Aufgaben eine kurze Wiederholung der wichtigsten Eigenschaften der Logarithmusrechnung:

- $\log(a \cdot x) = \log a + \log x$
- $\log x^a = a \cdot \log x$
- $x = 10^{\log x}$
- $\log x = \frac{\ln x}{\ln 10}$
- $\log \frac{1}{x} = -\log x$
- Ableitung $\partial \frac{\log x}{\partial x} = \frac{1}{x}$

1. Leite aus Gleichung 3.1 das Weber-Fechner-Gesetz $W = \text{const.} \cdot \ln(R) + C$ her.

2. Wie groß ist die Abnahme des Schalldrucks und des Schalldruckpegels bei einer Verdoppelung der Entfernung, wenn der Schalldruck proportional $1/r$ abnimmt?
3. Berechne für folgende Spannungsverhältnisse die Pegeldifferenz in dB :
 - $1 : 1$
 - $1 : \sqrt{2}$
 - $1 : 2$
 - $1 : 10$
4. Berechne für folgende Pegeldifferenzen die Spannungsverhältnisse:
 - $3dB$
 - $6dB$
 - $10dB$
 - $20dB$
 - $100dB$
5. Berechne die Pegelzunahme bei der Addition von 4 bzw. 8 unkorrelierten Quellen gleichen Pegels.
6. Welcher Pegelunterschied ergibt sich zwischen den Ausgangsspannungen zweier Mikrophone (Schalldruckempfänger), die 60cm und 180cm von der Schallquelle entfernt sind?
7. Um wieviel Dezibel nimmt der Pegel zu, wenn statt 2 Lautsprechern 4 von dem gleichen Signal angesteuert werden?
8. Um wieviel Dezibel nimmt der Pegel zu, wenn statt 5 Geigen 10 spielen?
9. Berechne sowohl für den korrelierten („gleichen“) als auch für den unkorrelierten Fall die Pegelzunahme bei der Addition zweier Signale, deren Pegelunterschied
 - $6dB$
 - $10dB$ beträgt.

Kapitel 4

Raumakustik

Die Akustik eines Aufnahmerraums kann für die Klangqualität einer Aufnahme von entscheidender Bedeutung sein. Das folgende Kapitel gibt eine Übersicht über die wichtigsten Größen der Raumakustik und akustische Besonderheiten bestimmter Räume.

In einem Raum erreicht nicht nur der Schall das Ohr des Hörers, der von der Schallquelle den direkten Weg nimmt, sondern auch vielfältige Reflexionen von Wänden, Decke, Boden, etc. Diese Reflexionen erzeugen einen Raumeindruck, der sich von Raum zu Raum stark unterscheidet. Dieser Raumeindruck, also die *Akustik* des Raumes, wird dadurch festgelegt, wann, wie stark und wie häufig Reflexionen im Vergleich zum Direktschall beim Hörer eintreffen (s. Abb. 4.1).

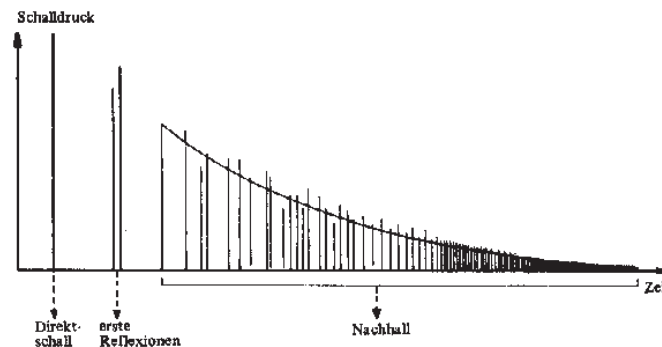


Abbildung 4.1: Zeitliche Folge der Reflexionen in einem Raum bei Impulsschall (aus [Dic87])

Im allgemeinen treffen die ersten Reflexionen nach dem Direktschall relativ vereinzelt ein und bestimmen durch ihre Verzögerung zum Direktschall und ihren Pegel stark den Klangeindruck des Raumes. Mit zunehmender Zeit liegen die einzelnen Reflexionen immer dichter beieinander und „verschwimmen“ zu einem abklingenden Nachhall.

Bemerkung

Tatsächlich ist die Verzögerung und Stärke der eintreffenden Reflexionen sowohl vom Standort des Hörers als auch von dem der Schallquelle abhängig. Das bedeutet nichts anderes, als daß verschiedene Punkte in einem Saal unterschiedliche Höreindrücke hervorrufen können. Es gibt aber doch einen mittleren Gesamteindruck des Raumes, den man meint, wenn man allgemein von der Akustik eines Raumes spricht.

Verwandte Themenbereiche

- Abschnitt [6.4.4](#)

4.1 Nachhallzeit

Die **Nachhallzeit** T ist eines der wichtigsten und bekanntesten Kriterien zur Beurteilung der Akustik eines Raumes. Sie ist die Zeit definiert, in der die Schallenergie eines (eingeschwungenen) Raumes nach „Abschalten“ der Schallquelle auf den millionsten Teil, d.h. um 60dB abnimmt. Diese Zeit entspricht nicht unbedingt der vom Zuhörer tatsächlich gehörten Zeit des Ausklangs, der **Nachhalldauer**, da diese abhängig von der Lautstärke der Schallquelle ist¹. Abhängig vom Volumen eines Raumes und den Absorptionseigenschaften des Wandmaterials ergeben sich für verschiedene Räume stark unterschiedliche Nachhallzeiten. Der Akustiker *Sabine* fand empirisch die *Sabinesche Nachhallformel* zur näherungsweisen rechnerischen Ermittlung der Nachhallzeit:

$$T \simeq 0.163 \text{sm}^{-1} \frac{V}{A} \quad (4.1)$$

V ist das Volumen des Raumes

A ist die sogenannte äquivalente Absorptionsfläche, d.h. die Summe aller Einzelflächen multipliziert mit ihrem jeweiligen Absorptionsgrad. Die Größe der Fläche A entspricht also der Größe einer theoretischen Fläche, welche keinen Schall reflektiert.

Räume für unterschiedliche Anwendungen sollten auch unterschiedliche Nachhallzeiten haben. So muß beispielsweise ein Theater wesentlich „trockener“ sein als ein Konzertsaal, um die Sprachverständlichkeit zu erhöhen. Typische Nachhallzeiten liegen in Konzertsälen oft um die 2s , in Theatern um 1.2s und in Kirchen sehr variabel von ca. 3s bis zu 7s .

Der Nachhall klingt näherungsweise exponentiell ab; wenn man also den Schalldruckpegel über der Zeit aufträgt, so erkennt man eine abfallende Gerade, die *Nachhallgerade* (s. Abb. [4.2](#)).

¹So ist die *Nachhalldauer* beispielsweise für ein Fingerschnippen sicherlich geringer als für ein Klatschen, die *Nachhallzeit* des Raumes bleibt aber dennoch immer die gleiche

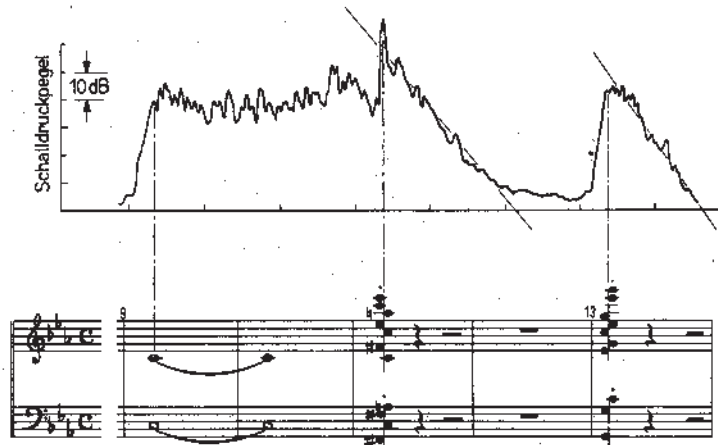


Abbildung 4.2: Nachhallaufzeichnung mit Pegelschreiber während eines Konzerts (Beethoven Op.62, Coriolan-Ouvertüre, Takt 9 ff) (aus [CM78])

4.2 Absorption

4.3 Direktschallfeld und Diffusschallfeld

Der **Direktschall** ist derjenige Schall, der den direkten Weg vom Sender zum Hörer nimmt, also als erste Wellenfront vom Sender eintrifft. Der **Diffusschall** oder Raumschall ist der Schall, der bei seinem Eintreffen beim Hörer bereits eine oder mehrere Reflexionen erfahren hat. Der Direktschall nimmt nach Gleichung (2.3) mit zunehmender Entfernung von der Schallquelle ab und bildet das sogenannte **Direktschallfeld** oder **Freifeld**. Hingegen spricht man von **diffusem Schallfeld** oder statistischem Schallfeld, wenn der Schalleinfall am Meßort aus allen Raumrichtungen gleich wahrscheinlich und gleich stark ist. Das bedeutet, daß im rein diffusen Schallfeld keine Lokalisation der Schallquelle mehr möglich ist. Der Pegel des Diffusschallfeldes ist bei längeren Nachhallzeiten im gesamten Raum gleich, hängt also nicht von der Entfernung zur Schallquelle ab (s. Abb. 4.3).

Bemerkung

Die Begriffe Direktschallfeld und Diffusschallfeld sind nicht zu verwechseln mit Nahfeld und Fernfeld, sondern stehen für etwas vollkommen anderes!

4.4 Hallradius

In einem Raum überlagern sich Direktschallfeld und Diffusschallfeld. Nahe an der Schallquelle ist das Direktschallfeld stark im Vergleich zum Diffusschallfeld.

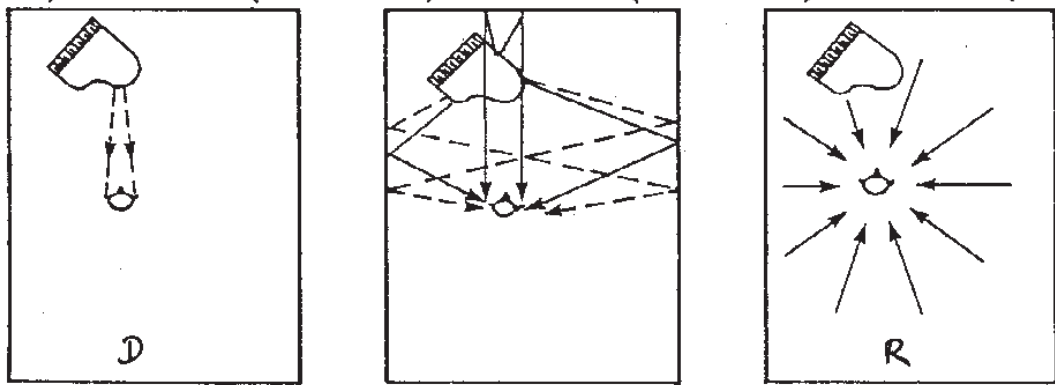


Abbildung 4.3: Direktschall, erste Reflexionen und Diffusschall am Ort des Hörers (aus [Sen98])

Je weiter man sich allerdings von der Quelle entfernt, desto schwächer wird der Pegel des Direktschalls und „verschwindet“ irgendwann im Diffusschall, der ja im gesamten Raum den gleichen Pegel hat (s. Abb. 4.4). Die Entfernung von der Schallquelle, bei welcher der Schallpegel des Direktschalles gleich dem des Raumschalles ist, wird **Hallradius** genannt:

$$r_H = 0.057 s^{1/2} m^{-1/2} \sqrt{\frac{V}{T}} \quad (4.2)$$

r_H ist der Hallradius

V ist das Volumen des Raumes

T ist die Nachhallzeit des Raumes

Bei einer Vergrößerung des Raumvolumens steigt der Hallradius also an, bei einer Verlängerung der Nachhallzeit verringert er sich (s. Abb. 4.5). Die sich aus dieser Gleichung ergebenden Hallradien sind erstaunlich gering; so beträgt er in einem $100m^3$ großen Zimmer mit der Nachhallzeit $0.5s$ lediglich $81cm$.

Wie schon der Name *Hallradius* impliziert, gilt diese Definition nur für eine vollkommen ungerichtete Schallabstrahlung (Kugelschallquelle). Da die meisten Schallquellen in der Regel aber nicht kugelförmig abstrahlen, erweitert man Gleichung (4.2) um den sogenannten *Bündelungsgrad*² γ und erhält in Hauptabstrahlrichtung der Schallquelle

$$r_{H,eff} = 0.057 s^{1/2} m^{-1/2} \sqrt{\gamma} \sqrt{\frac{V}{T}}. \quad (4.3)$$

²Der Bündelungsgrad ist für eine Kugelschallquelle 1 und wird größer, je stärker eine Quelle gerichtet abstrahlt.

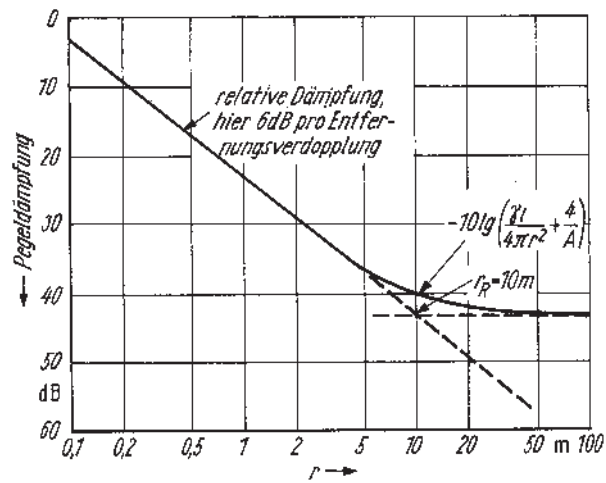


Abbildung 4.4: Pegeldämpfung in einem trockenen Raum über der Entfernung von der Schallquelle. Der Schallpegel verhält sich für kleine Abstände nach dem $1/r$ -Gesetz, für große Abstände ist er gleich dem Diffusschallpegel (aus [AS93])

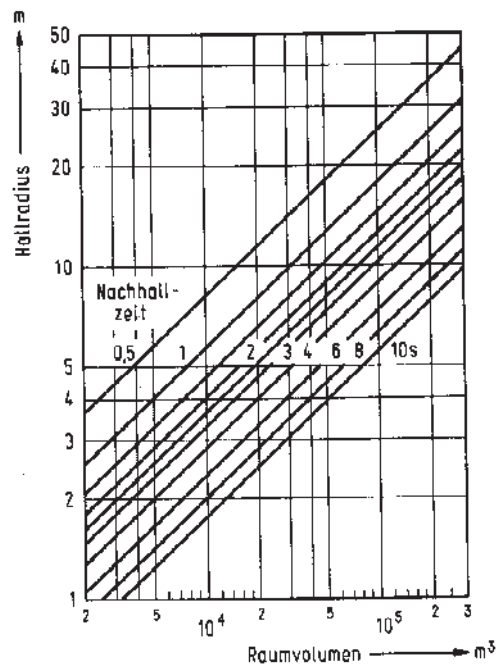


Abbildung 4.5: Hallradien in Abhängigkeit von Raumvolumen und Nachhallzeit (aus [Web94])

Dieser Wert wird als *effektiver Hallradius*, *Hallabstand* oder seltener als *Richtentfernung* bezeichnet.

Bemerkung

Für Aufnahmen ist die Kenntnis des Hallradius eines Raumes insofern wichtig, als man schon beim Aufstellen der Mikrophone ansatzweise einschätzen kann, ab welcher Entfernung von der Schallquelle der Direktschall leiser als der Diffusschall sein wird. Mit diesem Wissen lassen sich die Mikrophone schon vor dem ersten Testhören an einem Ort aufstellen, der der angestrebten Aufnahmeästhetik bezüglich des Verhältnisses von Direktschall und Raumanteil in etwa entsprechen könnte. Eine solche Überlegung kann aber nur erste Anhaltspunkte für den Mikrophonstandort liefern.

4.5 Frequenzabhängigkeit des Nachhalls

Die Nachhallzeit wurde oben als die Zeit definiert, in der die Energie eines eingeschwungenen Raumes um $60dB$ abnimmt (s. Abschnitt 4.1). Nun klingen allerdings auch Räume mit identischer Nachhallzeit oftmals sehr unterschiedlich. Die Nachhallzeit ist nämlich nicht für alle Frequenzen konstant, sondern unterscheidet sich zum Teil stark für verschiedene Frequenzen. In den meisten Räumen ist die Nachhallzeit bei tiefen Frequenzen länger als bei höheren Frequenzen. Eine Messung der frequenzabhängigen Nachhallzeit ist präziser und aussagekräftiger als die Messung der frequenzunabhängigen Nachhallzeit.

4.6 Flatterecho

Breitet ein Schallsignal sich derart aus, daß es über zwei (oder mehrere) Flächen wieder an den Ausgangspunkt zurückkehrt, so kann ein **Flatterecho** entstehen. Das bedeutet, daß der Schall immer wieder hin- und hergeworfen wird und somit - bei einem impulsartigem Schallereignis (Klatschen, etc.) - eine schnelle Abfolge leiser werdender Echos wahrzunehmen ist. Diese Echos werden i.a. als störend empfunden. Ein Flatterecho kann insbesondere dann entstehen, wenn zwei reflektierende Wände parallel zueinander stehen und die anderen Raumrichtungen stärker gedämpft sind.

Um ein Flatterecho zu beseitigen, sollte man entweder versuchen, eine der beiden gegenüberstehenden Wände z.B. durch einen Vorhang absorbierend zu machen, oder mit Hilfe von Diffusoren, großen Stellwänden etc. die Reflexionen in andere Richtungen abzulenken. Oft hilft auch einfach nur eine andere Platzierung der Musiker bzw. der Mikrophone.

4.7 Stehende Wellen

Trifft eine Schallwelle senkrecht auf eine Wand (Einfallrichtung 0°), so überlagern sich die einfallende und die reflektierte Welle.

Man stelle sich nun zwei parallel stehende reflektierende Wände vor; entspricht eine halbe Wellenlänge oder ein Vielfaches davon genau dem Abstand der beiden Wände, so werden sich Direktschall und ein- bzw. mehrfach reflektierter Schall immer phasengleich überlagern. Lediglich die Ausbreitungsrichtung kann unterschiedlich sein. Es entsteht ein stationäres Schallfeld zwischen den beiden Wänden mit Schwingungsknoten, in denen der Schalldruck fast 0 ist und Schwingungsbäuchen, in denen er maximal ist; diese sind ortsfest und verändern sich nicht über der Zeit. Dies sind sogenannte **stehende Wellen** (s. Abb. 4.6); sie werden teilweise auch als *Raumresonanzen* oder *Raummoden* bezeichnet.

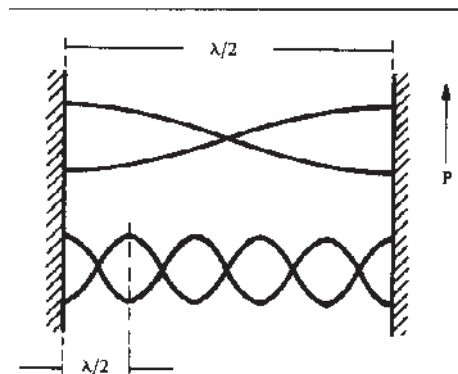


Abbildung 4.6: Druckverteilung in stehenden Wellen zwischen zwei parallelen Wänden für zwei verschiedene Frequenzen (aus [Dic87])

Bemerkung

Stehende Wellen können nicht nur zwischen zwei parallel stehenden Wänden auftreten, sondern auch durch mehrfache Reflexionen an verschiedenen Wänden, wenn die Bedingung erfüllt ist, daß der zurückgelegte Weg ein Vielfaches der halben Wellenlänge ist. Solche stehenden Wellen werden dann Raummoden höherer Ordnung genannt.

Abbildung 4.7 zeigt den Zusammenhang zwischen Druck und Schnelle bei einer stehenden Welle.

Stehende Wellen fallen vor allem in kleineren Räumen bei tiefen Frequenzen auf. Bei hohen Frequenzen liegen die Schwingungsknoten und -bäuche so dicht nebeneinander, daß man sie nicht mehr wahrnehmen kann. Je kleiner ein Raum ist, desto weiter liegen die Frequenzen auseinander, bei denen sich stehende Wellen ausbilden können. Deshalb kann man stehende Wellen in kleinen Räumen deutlicher wahrnehmen, da sie sich nur bei sehr wenigen Frequenzen ausbilden,

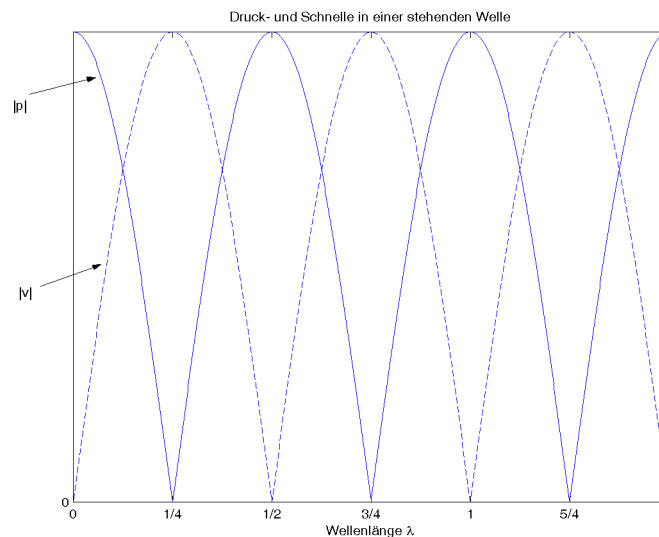


Abbildung 4.7: Druck- und Schnelleverlauf in einer stehenden Welle

die sich vom Hörer deutlich unterscheiden lassen. Aus diesem Grund wird bei der akustischen Konzeption eines Raumes darauf geachtet, daß die sogenannte *Eigenfrequenzdichte* möglichst hoch ist, d.h. daß die Frequenzen der einzelnen Moden möglichst so dicht nebeneinander liegen, daß die Wahrnehmung einer einzelnen Mode nicht mehr wahrnehmbar ist.

Bemerkung

Theoretisch tritt das Problem von stehenden Wellen erst auf, wenn der Raum mit einem reinen Sinuston dauerhaft angeregt wird, und nicht bei zeitlich veränderlichen (Musik-)Signalen. Tatsächlich kann es bei einer Aufnahme jedoch passieren, daß eine gewisse Baßlastigkeit oder Baßarmut auftritt, die dann durch leichtes Verrücken der Mikrophone verschwinden kann.

4.8 Zusammenfassung

Die sogenannte Akustik eines Raumes wird gekennzeichnet durch den Pegel, die Zeit und die Anzahl, mit der Reflexionen im Vergleich zum Direktschall beim Zuhörer eintreffen. Ein sehr wichtiges Kriterium für den Größeneindruck des Raumes ist die **Nachhallzeit** T . Sie ist abhängig von dem Raumvolumen und der Menge und Art von absorbierendem Material im Raum und entspricht im allgemeinen nicht der tatsächlich gehörten **Nachhalldauer**. Die Nachhallzeit ist **frequenzabhängig**, d.h. verschiedene Frequenzen klingen unterschiedlich schnell nach. Im allgemeinen ist die Nachhallzeit bei tiefen Frequenzen länger als bei hohen.

Durch die vielfältigen Reflexionen an den Begrenzungsflächen entsteht zusätzlich zu dem von der Schallquelle ausgesendeten **Direktschallfeld** ein näherungsweise **diffuses Schallfeld**. Hier trifft der Schall von allen Seiten gleich stark und mit gleicher Wahrscheinlichkeit ein. Der Radius der Kugel um die Schallquelle, an dem der Pegel des Direkt- und des Diffusschallfeldes gleich sind, wird **Hallradius** genannt. Für gerichtete Schallquellen muß zusätzlich noch der Bündelungsgrad der Schallquelle berücksichtigt werden. Man spricht dann vom **effektiven Hallradius**. Insbesondere in kleineren Räumen können zwischen parallel verlaufenden Wänden **stehende Wellen** entstehen. Diese bewirken ortsfeste Druckmaxima bzw. -minima, die vor allem bei tiefen Frequenzen wahrnehmbar werden.

4.9 Aufgaben

1. Berechne die Nachhallzeit für einen Raum mit dem Volumen $V = 21000m^3$ und einer äquivalenten Absorptionsfläche von $1711.5m^2$.
2. Berechne den Hallradius des o.g. Raums.
3. Berechne den effektiven Hallradius dieses Raums, wenn die Schallquelle eine Trompete mit einem ungefähren Bündelungsgrad von XXXX ist.
4. Berechne die Frequenzen möglicher stehender Wellen (1. Ordnung) in einem Raum mit der Grundfläche $5m \times 8m$.
5. Berechne die Verzögerung der ersten Wandreflexionen zum Direktschall in einem quadratischen Raum von $2500m^2$ Grundfläche, wenn Quelle und Hörer sich auf der Mittelachse gegenüberstehen, aber $30m$ Entfernung zueinander haben.

Kapitel 5

Psychoakustik

Das Verständnis von physikalischen Größen ist von großer Bedeutung, allerdings ist es ebenfalls von größter Bedeutung, sich darüber im klaren zu sein, wie starken Einfluß Änderungen von physikalischen Größen auf die Hörwahrnehmung haben. Aus diesem Grund ist es wichtig, die Funktionsweise des Ohres und die Besonderheiten der menschlichen Schallwahrnehmung zu kennen. In diesem Kapitel geht es um die Physiologie des Ohres sowie um psychoakustische Größen wie Lautstärke, Tonhöhe, Verdeckungseffekte und Richtungshören.

5.1 Physiologie des Ohres

Die Physiologie des Ohres gehört eigentlich nur bedingt zur Psychoakustik. Da das Verständnis der prinzipiellen Funktionsweise des Ohres aber zumindest nicht schaden kann, erscheint dies noch als der geeignetste Ort für dieses Thema. Das Ohr (s. Abb. 5.1) hat die Aufgabe, mechanische Energie aufzunehmen und so aufzubereiten, daß das Gehirn den eintreffenden Schall in einen Höreindruck umwandeln kann. Das Gehör kann anatomisch in drei Teile aufgeteilt werden:

- das Außenohr,
- das Mittelohr und
- das Innenohr.

5.1.1 Das Außenohr

Das **Außenohr** ist das, was umgangssprachlich oft einfach nur als Ohr bezeichnet wird, nämlich die nach außen „sichtbaren“ Teile unseres Gehörorgans. Es besteht aus der *Ohrmuschel*, dem *äußeren Gehörgang* und dem *Trommelfell*. Die Ohrmuschel stellt eine Art Schalltrichter da, die den Schall auffängt und deren Übertragungsfunktion abhängig von der Einfallsrichtung des Schalls ist. Der Schall

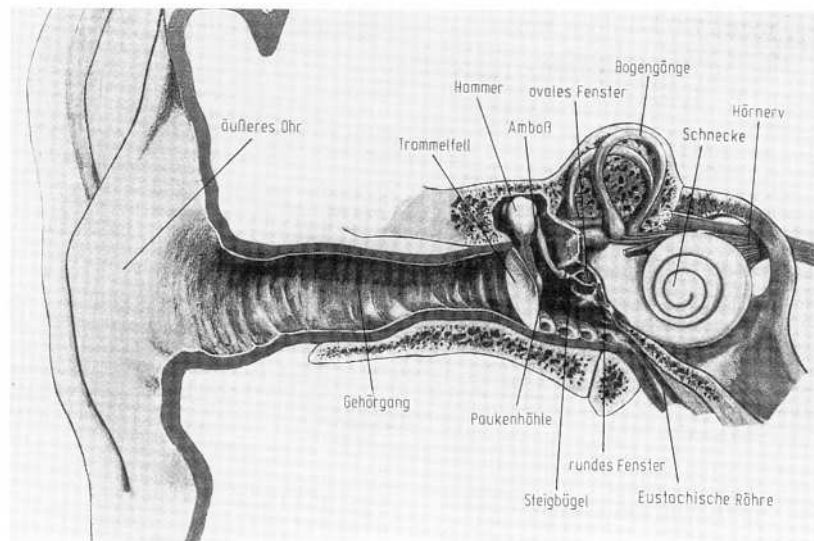


Abbildung 5.1: Schnitt durch das Gehör (aus [HM94])

gelangt dann über den 3 – 4cm langen Gehörgang zum Trommelfell, einer Membran, die durch den Schall zum Schwingen angeregt wird. Das Trommelfell stellt die Grenze zum Mittelohr dar.

5.1.2 Das Mittelohr

Im **Mittelohr** befinden sich die Gehörknöchelchen *Hammer*, *Amboß* und *Steigbügel*, mit deren Hilfe die Schwingungen des Trommelfells auf eine weitere Membran, das sogenannte *ovale Fenster*, weitergeleitet werden. Der Hammer ist mit dem Trommelfell verwachsen und ist durch ein Gelenk mit dem Amboß verbunden, der wiederum über ein Gelenk mit dem Steigbügel verbunden ist.

Bemerkung

Zwei kleine Muskeln setzen an Hammer bzw. Steigbügel an. Auf sehr laute Geräusche reagieren diese mit einer Anspannung und dämpfen auf diese Weise allzu starke Bewegungen; sie üben also eine gewisse Schutzfunktion aus.

Die Gehörknöchelchen wandeln Schwingungen großer Amplitude und kleiner Kraft in solche kleiner Amplitude und großer Kraft um. Der Steigbügel überträgt die gewandelten Schwingungen auf das *ovale Fenster*, die Grenze zum Innenohr. Damit das Mittelohr seine Funktion gut erfüllen kann, muß der (statische) Druck innerhalb des Mittelohres dem Druck der Umwelt entsprechen. Der Druckausgleich erfolgt über die *Eustachische Röhre*, auch *Ohrtrumpete* genannt.

Bemerkung

Die Eustachische Röhre verbindet das Mittelohr mit dem Nasen-Rachen-Raum; aus

diesem Grund läßt sich das unangenehme Druckgefühl bei plötzlicher Änderung des Außendrucks, beispielsweise beim Starten eines Flugzeuges, durch Kauen oder Gähnen beseitigen.

5.1.3 Das Innenohr

Das **Innenohr** enthält das Gleichgewichtsorgan und das eigentliche Hörorgan, das die Form einer Schnecke hat und dementsprechend auch *Schnecke* oder *Cochlea* heißt. Die Schnecke ist im Gegensatz zum Mittelohr mit Flüssigkeit gefüllt. Die Schnecke läßt sich in drei Kanäle aufteilen, welche *Scala vestibuli* (Vorhoftrappe), *Scala media* und *Scala tympani* (Paukentreppe) genannt werden.

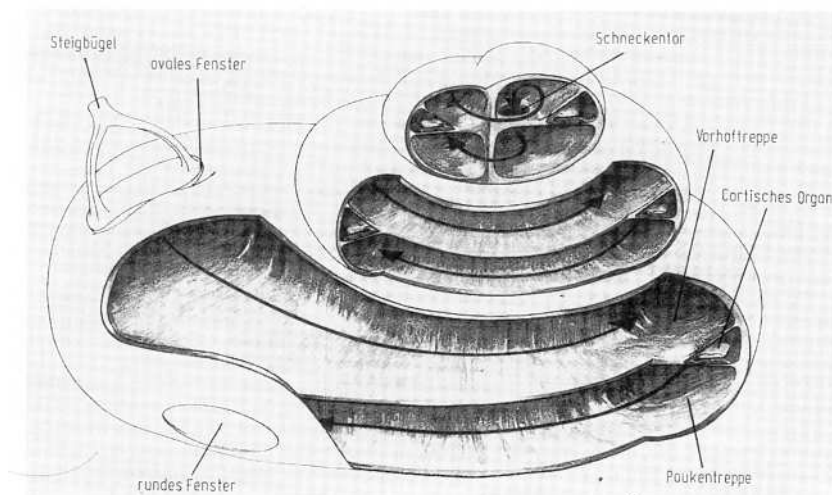


Abbildung 5.2: Schnitt durch die Schnecke (aus [HM94])

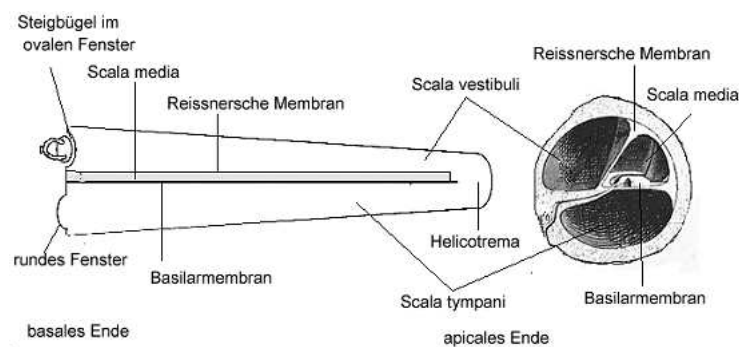


Abbildung 5.3: Darstellung der Schnecke in abgerolltem Zustand

Die Schwingungen des Steigbügels werden über das ovale Fenster auf die Flüssigkeit übertragen. Aufgrund der Druckänderung in den Scalen wird die Basilarmembran, die sich zwischen Scala media und Scala tympani befindet zum Schwingen angeregt. Auf dieser Membran bilden sich Wanderwellen aus (s. Abb. 5.4). Dabei ist die Amplitude dieser Wanderwellen (bzw. die Einhüllende der

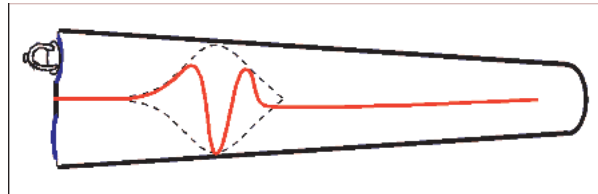


Abbildung 5.4: Wanderwelle auf der Basilarmembran

Schwingung) abhängig von der anregenden Schwingung und von der Position der Welle auf der Membran. Bei einer sinusförmigen Anregung steigt die Amplitude zuerst langsam an, bis sie an einer bestimmten frequenzabhängigen Stelle ihr Maximum erreicht und danach relativ schnell auf Null abklingt. Das bedeutet, daß sich auf der Basilarmembran bestimmte Stellen auch bestimmten Frequenzen zuweisen lassen. Töne hoher Frequenz werden in der Nähe des ovalen Fensters abgebildet, Töne niedriger Frequenz in der Nähe des *Schneckentors*.

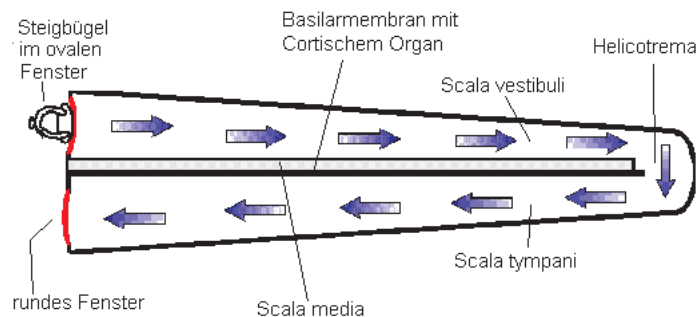


Abbildung 5.5: Flüssigkeitsströmung in der Schnecke

Auf der Basilarmembran befindet sich das *Cortische Organ*, welches die Aufgabe hat, die Schwingungen in elektrische Impulsmuster umzusetzen. Das Cortische Organ besitzt eine Vielzahl von Sinneszellen¹, auch Haarzellen genannt, welche die Schwingungsform der Basilarmembran abtasten. Die Haarzellen in der Umgebung der Maximalamplitude der Wanderwelle werden am stärksten angeregt; da jede Haarzelle von Einzelfasern des Hörnervs kontaktiert wird, kann das Gehör eine

¹Die Anzahl ist ca. 14000.

Tonhöhe zuordnen. Die Codierung der Reizintensität geschieht durch die Frequenz, mit der die Neuronen feuern. Mit zunehmendem Schalldruck nimmt also auch die Impulsfrequenz der Nervenzellen im Innenohr zu.

5.2 Schalldruckpegel, Lautstärke und Lautheit

Die empfundene Lautstärke eines Schallereignisses ist nicht nur abhängig von dem Schalldruckpegel, sondern auch von dessen Frequenzzusammensetzung. Beispielsweise rufen zwei Sinustöne unterschiedlicher Frequenz aber gleichen Pegels im allgemeinen eine unterschiedliche Lautstärkeempfindung hervor. Dieser Sachverhalt wird verdeutlicht durch die sogenannten **Kurven gleicher Lautstärkepegel** (s. Abb. 5.6). Hier erkennt man, daß der Schalldruckpegel für unterschiedliche Fre-

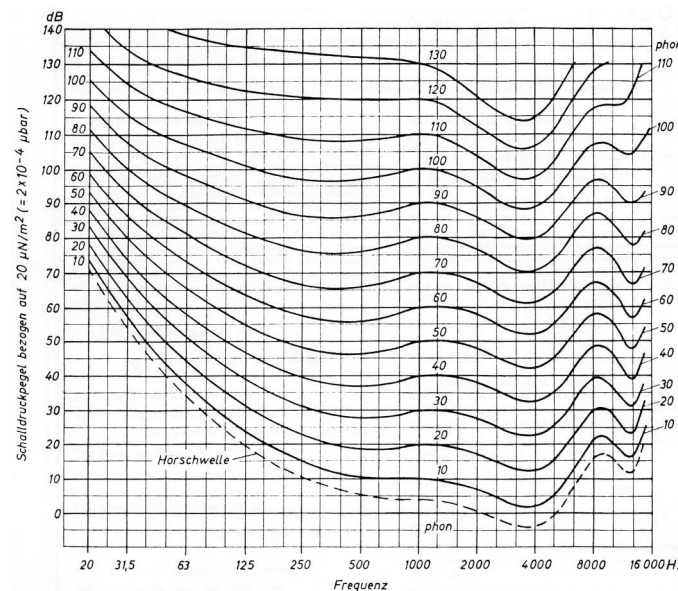


Abbildung 5.6: Normalkurven gleicher Lautstärkepegel (aus DIN 45 630)

quenzen sich teilweise stark verändern muß, um das gleiche Lautstärkeempfinden hervorzurufen. Die Einheit *phon* versucht, das Lautstärkeempfinden nachzubilden. Eine einzelne Kurve der Kurven gleicher Lautstärkepegel hat einen festen *phon*-Wert. So wird beispielsweise ein Ton beliebiger Frequenz mit der Lautstärke *80phon* als gleichlaut mit einem Sinuston der Frequenz *1kHz* mit dem Schalldruckpegel *80dB* empfunden. Besonders im Bereich zwischen 2 und *5kHz* ist das Ohr sehr empfindlich.

Die Kurven gleicher Lautstärkepegel verlaufen nicht parallel zueinander, sondern haben für unterschiedliche Lautstärken auch einen leicht unterschiedlichen Verlauf. Insbesondere wird ein leiser wiedergegebenes Schallereignis baß- und auch leicht

höhenärmer erscheinen als ein lautes.

Bemerkung

Für Aufnahme und Mischung kann dieses Wissen von Bedeutung sein. Ein wichtiger Anhaltspunkt ist, wie laut die Mischung später erwartungsgemäß abgehört wird, ob die Abhörlautstärke bei der Mischung darüber oder darunter liegt und wie sich das auf die wahrgenommene Frequenzverteilung auswirkt. Ein Probehören der Mischung in anderen Lautstärken kann hilfreich sein.

Die **Lautheit** wird in der Einheit *sone* angegeben. Als Bezugswert wurde dem Wert 1 sone ein Sinuston der Frequenz 1 kHz und dem Schalldruckpegel 40 dB zugeordnet. Ein doppelt so laut empfundener Schallvorgang hat die Lautheit 2 sone (was $+10\text{ phon}$ entspricht), ein vierfach so laut empfundener 4 sone .

Bemerkung

Im Tonstudioalltag wird man sich mit der Einheit *sone* i.a. nicht beschäftigen müssen. In der akustischen Meßtechnik spielt sie allerdings eine Rolle; z.B. wird in Qualitätstests von Computerfestplatten oft der Störschall in *sone* angegeben.

Die Kurven gleicher Lautstärke betrachten nur den pegelabhängigen Frequenzgang des Ohres. Auch andere Parameter können aber das Lautstärkeempfinden beeinflussen. So wirkt beispielsweise eine sehr kurzes Schallsignal leiser als das gleiche Signal gleichen Pegels länger ausgehalten.

Verwandte Themenbereiche

- Abschnitt [1.3](#)
- Abschnitt [3.2](#)

5.3 Tonhöhenempfindung

Die Tonhöhenempfindung ist nicht ausschließlich abhängig von der Frequenz (wie in [1.2](#) beschrieben), sondern auch von der Lautstärke eines Tones. Bei Frequenzen über zweitausend Hertz steigt die empfundene Tonhöhe mit zunehmendem Pegel leicht an während sie bei Frequenzen unter 1 kHz mit zunehmenden Pegel leicht abfällt. Zur weitergehenden Vertiefung sei auf [ZF99] und [VM75] verwiesen.

Verwandte Themenbereiche

- Abschnitt [1.2](#)
- Abschnitt [3.1](#)

5.4 Verdeckungseffekte

Es lassen sich nicht immer alle Informationen wahrnehmen, die in einem Musiksignal enthalten sind. Vielmehr treten in bestimmten Fällen sogenannte **Verdeckungseffekte** auf. So verdeckt beispielsweise ein einzelner Sinuston dicht neben dieser Frequenz liegende Signalanteile (s. Abb. [5.7](#)); diese sind unhörbar.

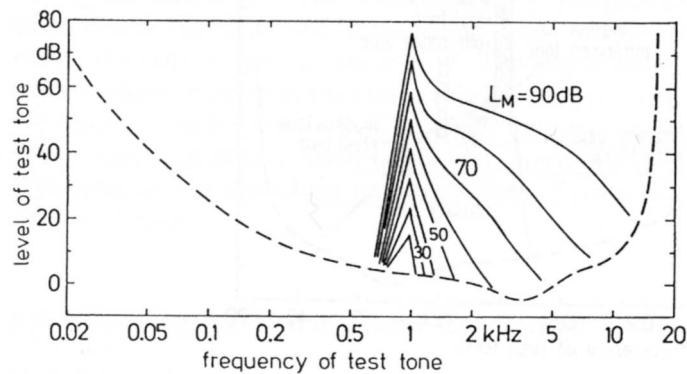


Abbildung 5.7: Pegel eines Sinustons, der von einem 1 kHz -Sinuston unterschiedlichen Pegels maskiert wird, als Funktion der Frequenz des Testtones (aus [ZF99])

Neben den Verdeckungseffekten im Frequenzbereich gibt es auch noch zeitliche Verdeckungseffekte, nämlich die **Nachverdeckung** (auch *Forward Masking* oder *Postmasking*) und die **Vorverdeckung** (auch *Backward Masking* oder *Premasking*). In Abb. 5.8 sind die Auswirkungen dieser Effekte dargestellt. Es gibt also kurz vor und kurz nach einem Signal eine Zeit, in welcher wir unter einer gewissen Schwelle liegenden Schallereignisse nicht mehr wahrnehmen können.

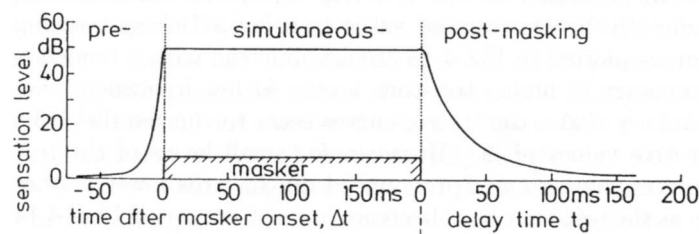


Abbildung 5.8: Bereiche, in denen Pre- und Postmasking auftritt (aus [ZF99])

Bemerkung

Bei der verlustbehafteten Codierung von Audiosignalen (z.B. die MPEG-Formate, Dolby AC-3, Sony ATRAC, etc.) wird versucht, diese Eigenschaften des Gehörs bzw. der Wahrnehmung auszunutzen und solche Signalanteile, die i.a. nicht mehr wahrgenommen werden können, gleich „wegzulassen“.

Verwandte Themenbereiche

- Abschnitt 8.6

5.5 Richtungshören bei einer Schallquelle

Viele Untersuchungen haben sich mit der Richtungslokalisierung von Schallquellen beschäftigt. Die wahrzunehmende Auflösung in Blickrichtung (**Lokalisations-schärfe**) ist abhängig von Versuchsaufbau und verwendetem Tonsignal. Um diese Abhängigkeit zu veranschaulichen, sind in der Tabelle 5.1 die Ergebnisse einiger Testreihen mit unterschiedlichen Schallquellen aufgeführt.

Die Lokalisation einer Schallquelle wird vor allem durch das beidohrige (*binaurale*)

Veröffentlichung	Art des Signals	Lokalisationsschärfe
Klemm (1920)	Impulses (clicks)	0.75°-2°
King and Laird (1930)	Impulse (click) train	1.6°
Stevens and Newman (1936)	Sinusoids	4.4°
Schmidt et al. (1953)	Sinusoids	>1°
Sandel et al. (1955)	Sinusoids	1.1°-4.0°
Mills (1958)	Sinusoids	1.0°-3.1°
Stiller (1960)	Narrow-band noise, \cos^2 tone bursts	1.4°-2.8°
Boerger (1965a)	Gaussian tone bursts	0.8°-3.3°
Gardner (1968a)	Speech	0.9°
Perrott (1969)	Tone bursts	1.8°-11.8°
Blauert (1970b)	Speech	1.5°
Haustein and Schirmer (1970)	Broadband noise	3.2°

Tabelle 5.1: unterschiedliche Messungen zur Lokalisationsschärfe bei horizontaler Verschiebung der Schallquelle aus der Blickrichtung (aus [Bla97])

Hören ermöglicht. Vom Gehirn werden die (*interauralen*) Unterschiede zwischen den Schallsignalen am linken und am rechten Ohr ausgewertet und die Richtung der Schallquelle bestimmt. Maßgebliche Parameter der Richtungslokalisierung sind *Laufzeitunterschiede* und *Pegelunterschiede* zwischen den Ohren. Wie Abb. 5.9 schematisch veranschaulicht, gelangt ein Schallsignal im allgemeinen nicht gleichzeitig an beide Ohren, sondern wird an dem der Schallquelle zugewandten Ohr früher eintreffen. Auch gelangt durch die Abschattung des Kopfes das Schallsignal frequenzabhängig mit geringerem Pegel an das der Schallquelle abgewandte Ohr.

Bemerkung

Diese Dämpfung ist frequenzabhängig; hohe Frequenzen werden stärker gedämpft als tiefe (s. Abschnitt 2.6.2).

Zusätzlich zu der Auswertung der Signalunterschiede zwischen beiden Ohren gibt es auch noch die Übertragungsfunktion unserer Außenohren. Durch diese wird es ermöglicht, zum Beispiel zwischen vorne und hinten zu unterscheiden, was ja anhand von Signalunterschieden nicht möglich wäre². Der Einfluß der Außenohren

²Allein aufgrund der Übertragungsfunktion des Außenohres ist das noch nicht ausreichend gut möglich; zusätzlich zu den Außenohrübertragungsfunktionen werden vom Hörer auch sehr geringe Kopfbewegungen ausgeführt, um durch die dadurch auftretenden Laufzeit- und Pegeldifferenzen zwischen den möglichen Einfallsrichtungen zu unterscheiden

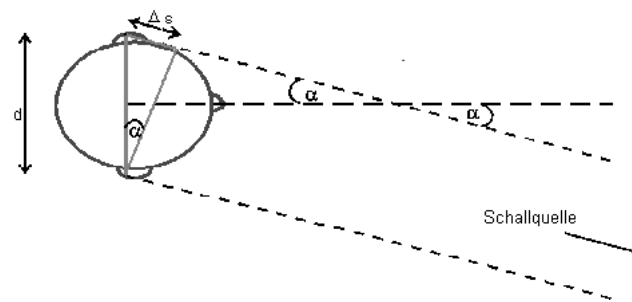


Abbildung 5.9: Laufzeitunterschiede zwischen beiden Ohren für eine Schallquelle

auf den „gehörten“ Frequenzgang des eintreffenden Schalls ist abhängig von der Einfallsrichtung und ist bei jedem Menschen leicht unterschiedlich. Es lassen sich aber einige typische Frequenzen angeben, die für einen bestimmten Richtungseindruck stehen. Abb. 5.10 zeigt Untersuchungsergebnisse, welche verdeutlichen, daß der Richtungseindruck stark von der Frequenz eines Signals abhängt. Hier wurde zur Messung Terzbandrauschen mit unterschiedlicher Länge verwendet. Man erkennt beispielsweise, daß bei einer Mittenfrequenz von 1 kHz ca. 90 Prozent der Testpersonen das Signal hinter sich orten, während das Schallereignis bei einer Mittenfrequenz von 8 kHz oben lokalisiert wird. Sind Frequenzen um 4 kHz in einem Signal enthalten, so erhöht sich der „Vorne-Eindruck“, die Präsenz eines Signals. Diese Frequenzbereiche werden **richtungsbestimmende Bänder** genannt. Das Wissen um die Wirkung dieser verschiedenen Frequenzbänder kann beim Filtern durchaus hilfreich sein, da man schon einen Anhaltspunkt hat, in welchem Frequenzbereich man zunächst suchen sollte, um den gewünschten Effekt zu erzielen.

Verwandte Themenbereiche

- Abschnitt [6.4.2](#)

5.5.1 Das Gesetz der ersten Wellenfront

Zwei ähnliche Signale, die aus unterschiedlichen Richtungen kommen (z.B. Direktschall und ein Rückwurf), werden aus der Einfallsrichtung lokalisiert, aus welcher die erste Wellenfront eintrifft. Dieser Sachverhalt wird **Gesetz der ersten Wellenfront** bzw. **Prezedenzeffekt** genannt. Die Verzögerung des zweiten Signals darf allerdings eine gewisse Schwelle (*Echoschwelle*) nicht überschreiten, da der Hörer in einem solchen Fall zwei einzelne Signale (z.B. Direktschall und Echo) wahrnimmt. Die *Echoschwelle* ist abhängig von der Verzögerungszeit und dem Pegel des zweiten Signals (s. Abb. 5.12). Ist die Verzögerung des zweiten Signals sehr klein, so entsteht ein anderes Phänomen, die Summenlokalisierung (s. Abschnitt 5.6. Versucht man bei zwei identischen Signalen, die durch allmähliche

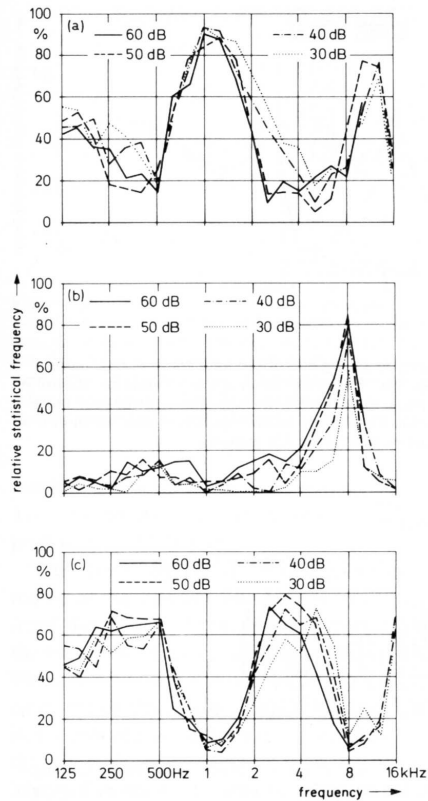


Abbildung 5.10: relative Häufigkeit der Antworten *hinten*, *oben*, *vorne* über der Mittenfrequenz eines Terzbandrauschens (einmal von vorne und einmal von hinten eingespielt) aufgetragen (aus [Bla97])

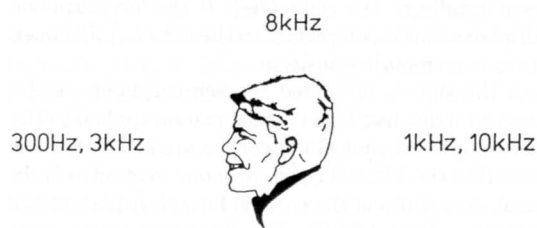


Abbildung 5.11: Veranschaulichung der richtungsbestimmenden Bänder in der Medianebene, unabhängig von der Position der Quelle (aus [ZF99])

Verzögerung des zweiten Signals verschobene Richtungswahrnehmung wiederum durch Pegelerhöhung des zweiten Signals auszugleichen, was *Trading* genannt wird,

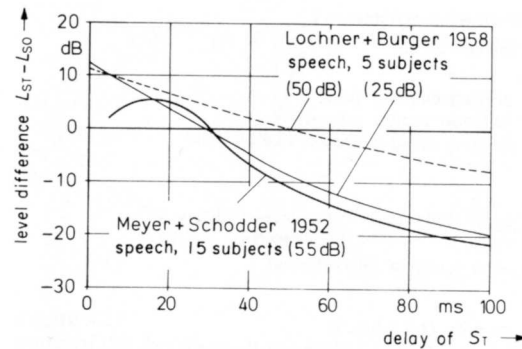


Abbildung 5.12: Verschiedene Messungen zur Echoschwelle bei Sprache (aus [Bla97])

so erhält man die in Abb. 5.13 dargestellte Funktion. Man könnte also ein um 15ms

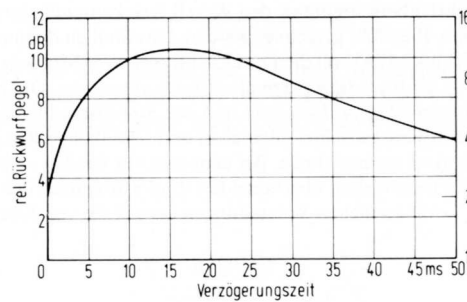


Abbildung 5.13: Pegel eines Rückwurfes, bei dem die Richtungslokalisierung des Direktschalls zerstört wird (aus [HM94])

verzögertes Signal um bis zu 10dB anheben, ohne daß dieses Signal als einzelnes wahrzunehmen wäre. Diese Messung erfolgte allerdings mit Sprachsignalen, läßt sich also demnach nur näherungsweise auf Musik übertragen.

Bemerkung

Bei Beschallungen mit mehreren Lautsprechersystemen wird oft mit entsprechenden Verzögerungen gearbeitet, um den Pegel des zu verstärkenden Signals erhöhen zu können, ohne die Ortbarkeit des Direktsignals zu verlieren.

Verwandte Themenbereiche

- Abschnitt 6.4.3

5.6 Richtungshören im Stereodreieck

Gibt man im Stereodreieck (Abb. 5.14) auf die Lautsprecher korrelierte Signale gleichen Pegels, so ortet man die Schallquelle auf der Lautsprecherbasis, also in der Mitte. Eine solche virtuelle Schallquelle heißt **Phantomschallquelle**³.

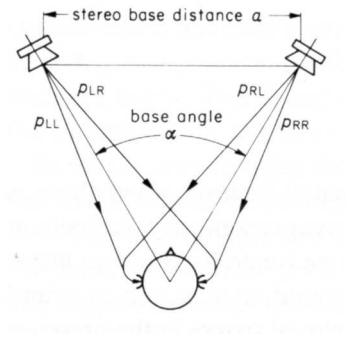


Abbildung 5.14: Übliche Lautsprecher-Stereoaufstellung, der Winkel α zwischen den beiden Lautsprechern beträgt normalerweise ca. 60° (aus [Bla97])

Um die Phantomschallquelle aus der Mitte heraus zu bewegen, kann man Pegel- und/oder Laufzeitdifferenzen benutzen. Bei welchen Pegel- bzw. Laufzeitunterschieden zwischen den Signalen *links* und *rechts* verschiebt sich nun die Ortung einer Phantomschallquelle im klassischen Stereo-Panorama (s. Abb. 5.14)? Abb. 5.15 und Abb. 5.16 veranschaulichen, wie weit sich die Ortung der Phantomschallquelle bei Änderung des Pegelunterschiedes oder der Verzögerung zwischen *links* und *rechts* aus der Mitte verschiebt. Auch bei den Abb. 5.15 und 5.16 wird wiederum deutlich, wie stark die Messungen von dem Testsignal abhängen. Eine generelle Aussage, z.B. ab welchem Laufzeitunterschied oder welcher Pegeldifferenz ein Signal in einem Lautsprecher geortet wird, ist nur sehr grob möglich.

5.7 Zusammenfassung

Das Ohr läßt sich in drei Teile untergliedern: **Außenohr**, **Mittelohr** und **Innenohr**. Der Schall gelangt über den Gehörgang zum *Trommelfell* und versetzt dieses in Schwingung. Die *Gehörknöchelchen* im Mittelohr übertragen die Schwingungen des Trommelfells auf das *ovale Fenster*, die Grenze zum Innenohr. Im Innenohr befindet sich die Schnecke, das eigentliche Hörorgan. Hier werden die Schwingungen auf die *Basilarmembran* übertragen, deren Schwingungen dann die *Haarzellen*, Nerven auf der Basilarmembran, zum Feuern anregen.

³siehe in der einschlägigen Literatur unter *Summenlokalisation* und *Assoziationsmodell von Theile*.

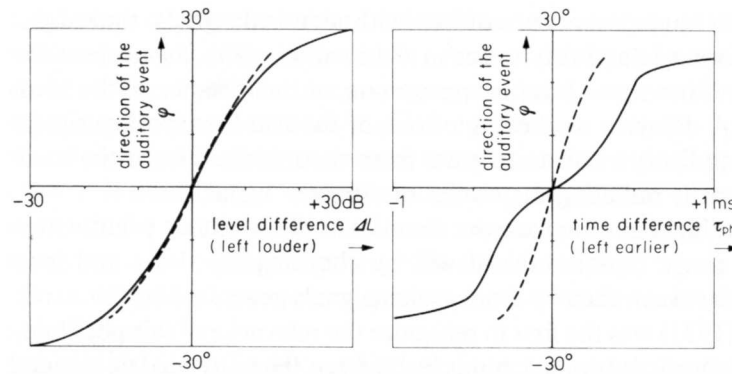


Abbildung 5.15: Lokalisation einer Phantomschallquelle im Stereodreieck (nach Abb. 5.14); links Pegelunterschiede, rechts Laufzeitunterschiede; Signal Sprache, gestrichelte Linien für bewegbaren Kopf, durchgezogene Kurven für feststehenden Kopf (aus [Bla97])

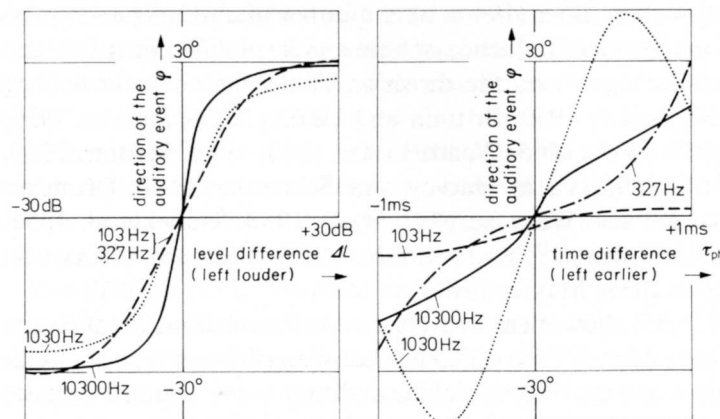


Abbildung 5.16: Lokalisation einer Phantomschallquelle im Stereodreieck nach Abb. 5.14; links Pegelunterschiede, rechts Laufzeitunterschiede; Signal terzbreite Sinusbursts verschiedener Frequenzen, gestrichelte Linien für bewegbaren Kopf, durchgezogene Kurven für feststehenden Kopf (aus [Bla97])

Die **Lautstärke**, mit der wir ein Schallereignis empfinden, läßt sich nur näherungsweise mit dem absoluten Schalldruckpegel beschreiben. Tatsächlich ist das Lautstärkeempfinden sehr stark abhängig von der Art des Schallereignisses, sowohl von dessen zeitlicher als auch frequenzmäßiger Zusammensetzung. Die **Kurven gleicher Lautstärkepegel** zeigen das frequenzabhängige Lautstärkeempfinden. Einheiten, die dieses Empfinden nachvollziehen, sind beispielsweise *phon* und *sone*. Nicht alle in einem Audiosignal enthaltenen Informationen müssen wirklich hör-

bar sein. Es gibt **Verdeckungseffekte**, bei denen ein bestimmtes Schallereignis andere Ereignisse unhörbar macht. Diese Effekte existieren sowohl im Zeit- als auch im Frequenzbereich.

Der Mensch kann eine Schallquelle v.a. aufgrund der *Laufzeit-* und *Pegelunterschiede* zwischen beiden Ohren im Raum lokalisieren. Eine etwas untergeordnete Rolle spielen für die Ortung auch die *Außenohrübertragungsfunktionen*. Alle drei Parameter der Lokalisation lassen sich auch zusammenfassen in den sogenannten kopfbezogenen Übertragungsfunktionen (HRTF: *Head Related Transfer Functions*).

Der **Prezedenzeffekt** bzw. das **Gesetz der ersten Wellenfront** besagt, daß Schall aus der Richtung lokalisiert wird, aus der die erste Wellenfront beim Hörer eintrifft.

Im Stereodreieck hat man zwei Schallquellen (die Lautsprecher). Liegen auf beiden Lautsprechern korrelierte Signale, so bildet sich eine **Phantomschallquelle** zwischen den beiden Lautsprechern. Durch Laufzeit- und/oder Pegeldifferenzen zwischen den beiden Lautsprechersignalen läßt sich die Phantomschallquelle zwischen den beiden Lautsprechern bewegen. Abhängig vom Signal sind Pegeldifferenzen zwischen 15 und 20dB bzw. Laufzeitunterschiede zwischen 0.7 und 1ms, u.U. bis zu 2ms notwendig, um die Phantomschallquelle ganz in einen Lautsprecher zu verschieben. Diese Werte sind nicht zu verwechseln mit den *interauralen* Laufzeit- und Pegeldifferenzen, die wesentlich kleiner sind.

5.8 Aufgaben

1. Zwei parallel ausgerichtete Mikrophone stehen im Abstand von 50cm. Im Abstand von 5m zu der gedachten Verbindungslinie zwischen den Mikrophen befindet sich eine Schallquelle, die aber um 25cm aus der Mitte verschoben ist, so daß Mikrophone und Quelle ein rechtwinkliges Dreieck bilden. Wie groß sind die Pegel- und Laufzeitunterschiede zwischen den Mikrophen?
2. Wie weit müßte man das weiter entfernte Mikrofon verschieben, um eine Laufzeitdifferenz von 1ms zu erreichen?
3. Wie groß kann der Laufzeitunterschied zwischen den beiden Ohren maximal sein, wenn das Signal komplett um den Kopf herum gebeugt wird? Der Kopf soll dabei durch eine Kugel mit Radius 10cm angenähert werden, die Schallgeschwindigkeit sei $340 \frac{m}{s}$.

Kapitel 6

Signalverarbeitung

Dieses Kapitel stellt eine Einführung in die wichtigsten Begriffe der Signalverarbeitung dar. Zunächst geht es relativ theoretisch um Systembeschreibungen, anschließend geht es um einen kurzen Überblick über verschiedene Effekte und Qualitätsmerkmale.

6.1 Signalbeschreibungen

Nachdem in vorherigen Abschnitten wiederholt sowohl Darstellungen des Zeitverlaufs wie auch Frequenzdarstellungen von Signalen auftauchten, soll dieses Kapitel diese Darstellungen nun detaillierter erklären.

6.1.1 Zeitverlauf

Der **Zeitverlauf** eines Signals ist die geläufigste und am einfachsten verständliche Darstellung eines Audiosignals. Nimmt man z.B. den Schalldruck eines akustischen Ereignisses an einem Ort auf, und zeichnet die Druckschwankungen über der Zeit auf, so betrachtet man den Zeitverlauf des Drucks. Dabei nimmt die Zeit nach rechts hin zu, und die Amplitude des Drucks wird nach oben aufgetragen. Nichts anderes ist auch die Wellenformdarstellung unterschiedlichster Audibearbeitungsprogramme, wo die Änderungen der Amplitude des Signals über der Zeit dargestellt werden. Abb. 6.1 zeigt beispielhaft die Zeitdarstellung eines Audiosignals.

6.1.2 Spektrum

Ebenso wie mit der Zeitdarstellung eines Signals, wurde auch schon mit seiner **Frequenzdarstellung**, dem Spektrum eines Signals in vorhergehenden Kapiteln gearbeitet. Beispielsweise zeigt die Abb. 1.4 links den Zeitverlauf und rechts das Frequenzspektrum von Sinusschwingungen. Je höher die Frequenz einer Schwingung ist, desto mehr rutscht der Peak dieser Schwingung auf der Frequenzachse

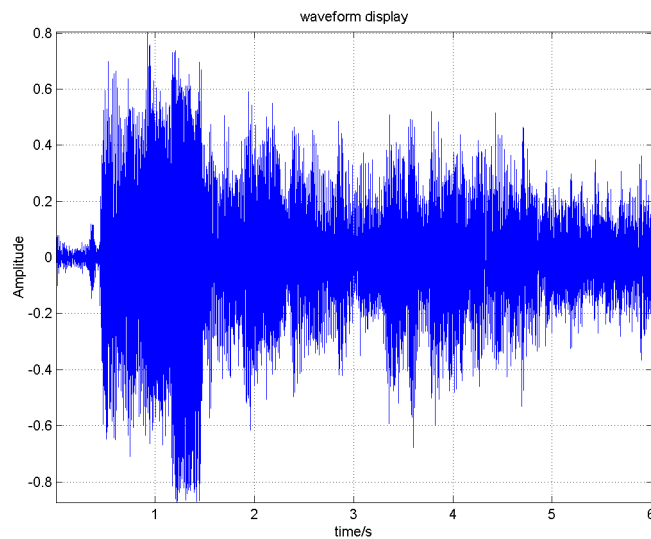


Abbildung 6.1: Zeitdarstellung eines Audiosignals; diese Beispiel ist ein kurzer Ausschnitt aus einem Stück mit der Besetzung Gesang und Klavier

nach rechts. Sinusförmige Schwingungen besitzen nur einen einzigen Frequenzanteil; in realen Signalen können dagegen sehr viele Frequenzen zur gleichen Zeit auftreten. Die Frequenzdarstellung wird auch Fouriertransformation des Signals genannt. Da Musiksignale sich auch über die Zeit ständig ändern, kann man allerdings aus Langzeitspektren kaum interessante Informationen sehen. Aus diesem Grund werden selten Frequenzanalysen eines kompletten längeren Signalabschnitts gemacht. Eine wichtige Eigenschaft der Fouriertransformation ist, daß die Transformation in beide Richtungen möglich ist, d.h. jedes Spektrum läßt sich genauso leicht in den Zeitverlauf umrechnen wie andersherum.

Führt man eine Spektralanalyse (eines digitalen Signals) über N Samples durch, so erhält man genau $N/2$ komplexe Spektralwerte, die gleichförmig über das Spektrum verteilt sind. Nimmt man beispielsweise 1024 Signalwerte eines Signals der Abtastfrequenz 48kHz , so hat man eine zeitliche Auflösung von 21.3ms und eine Frequenzauflösung von 46.9Hz . Je gröber also die zeitliche Auflösung wird, desto feiner wird die Frequenzauflösung und umgekehrt. Dies ist wichtig, denn das Signal sollte in sich innerhalb eines Analysefensters nicht zu sehr ändern. Abb. 6.2 zeigt die verschiedene Darstellungsmöglichkeiten der komplexen Werte des schon in Abb. 6.1 verwendeten Signals. Dieses Signal ist eigentlich schon viel zu lang, um eine Frequenzanalyse davon durchzuführen, aus diesem Grund ist das Ergebnis auch nicht sonderlich interessant zu betrachten. Abb. 6.3 ist die geläufige Darstellungsweise eines Signalspektrum; hier wird der Betragsfrequenzgang in Dezibel gezeigt. Da der Signalausschnitt jedoch für eine FFT-Analyse eigentlich zu lang ist, gibt es noch weitere Darstellungsmöglichkeiten, in der Zeit-

und Frequenzverlauf in gewisser Hinsicht vereint werden, z.B. das sogenannte Spektrogramm, das in Abb. 6.4 dargestellt ist, oder eine Pseudo-3D-Grafik wie in Abb.6.5.

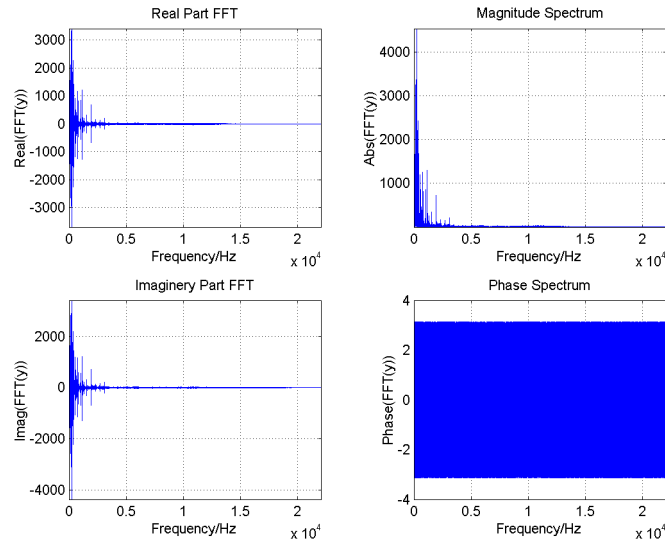


Abbildung 6.2: Frequenzspektrum eines Audiosignals, links als Real- und Imaginärteil, rechts mit Betrag- und Phasenspektrum

6.1.3 Amplitudendichteverteilung

Die Amplitudendichteverteilung (ADV oder englisch: PDF für Probability Density Function) gibt an, wie wahrscheinlich bzw. wie häufig bestimmte Amplitudenwerte in einem Signal vorkommen. Dabei werden alle möglichen Amplituden auf der Abszisse (*x-Achse*) aufgetragen, und die Häufigkeit jedes einzelnen Wertes auf der Ordinate (*y-Achse*). Da ein Sprach- oder Musiksignal meistens ebenso viel positive wie negative Signalwerte besitzt, ist die ADV solcher Signale symmetrisch zur Ordinate. Bestünde ein Signal nur aus positiven Werten, so wäre auf der negativen Abszisse nur der Wert 0 aufgetragen. Sind in einem Signal die kleineren Amplitudenwerte häufiger als große, so wird die ADV nach außen hin abflachen. Dies ist praktisch bei allen Sprach- und Musiksignalen der Fall. Damit die ADV sich für identische Signale, die aber unterschiedliche Länge haben, nicht unterscheidet, wird sie zumeist auf die Anzahl der betrachteten Signalwerte normiert. Dies führt dann dazu, daß die Summe aller normierten ADV-Werte 1 ergibt.

Bemerkung

Die beschriebene Funktion ist eigentlich nur eine Näherung der ADV, nämlich eine *Häufigkeitsverteilung*(HV). Lediglich die Verallgemeinerung dieser HV-Messung für

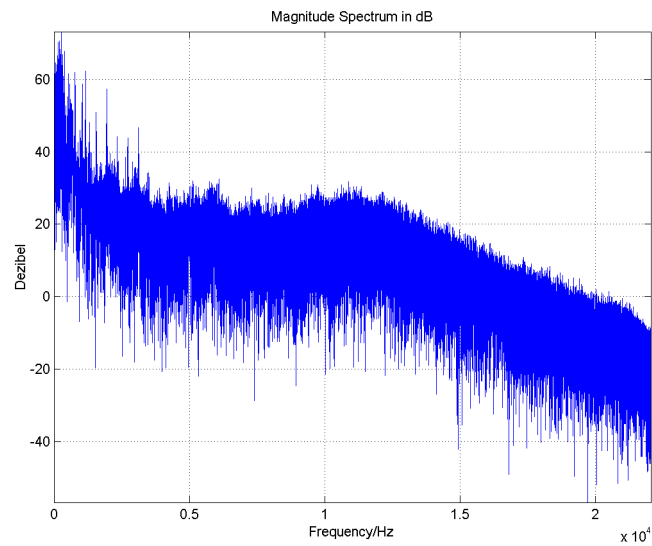


Abbildung 6.3: Betragsfrequenzgang eines Audiosignals in dB; gebräuchlichste Frequenzdarstellung

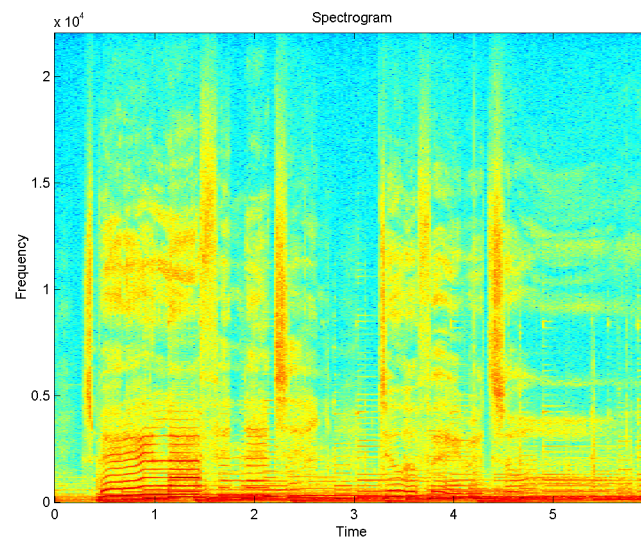


Abbildung 6.4: Spektrogrammdarstellung eines Audiosignals; die roten Bereiche besitzen die höchste Amplitude

kontinuierliche Amplitudenwerte aller Signalausdehnungen liefert dann die ADV bzw. die tatsächliche Wahrscheinlichkeitsdichtefunktion. Dennoch wird eine solche Häufigkeitsverteilung im folgenden unter der Bezeichnung ADV verwendet.

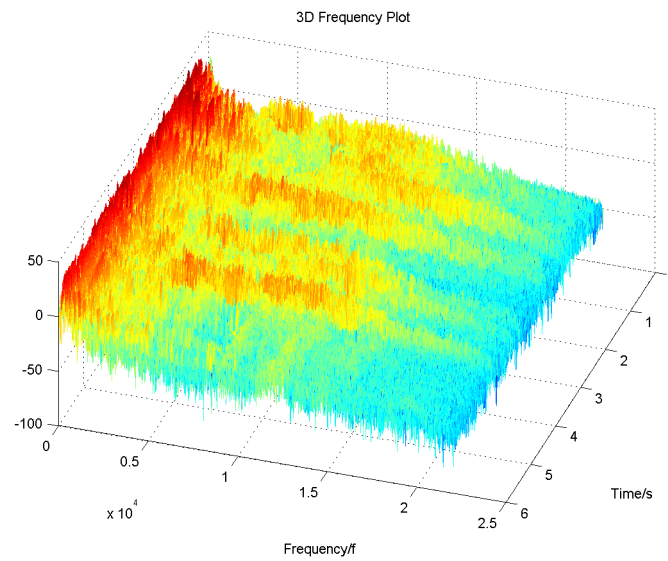


Abbildung 6.5: Pseudo-3D-Darstellung eines Frequenzverlaufs über der Zeit

Bemerkung

Da auf der Ordinate die Häufigkeit von Amplitudenwerten aufgetragen ist, kann kein Wert der ADV jemals negativ werden.

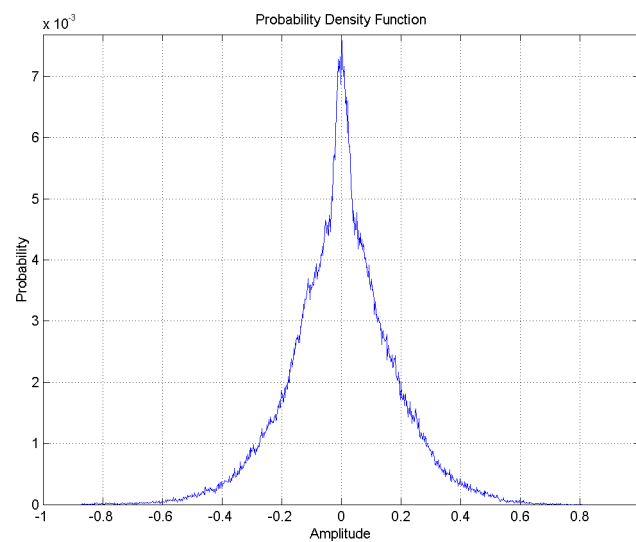


Abbildung 6.6: Amplitudendichteverteilung eines Musiksignals

6.2 Systembeschreibungen

Der Begriff System kann in der Signaltheorie für praktisch jedes signalverarbeitende Gerät verwendet werden. Beispiele aus der Tontechnik sind Effekte wie Delay und Hallgerät ebenso wie ein normaler Verstärker oder ein Kompressor (vgl. Abschnitt 6.4). Ebenso kann aber auch jeder Raum als System verstanden werden, der jedes akustische Signal verhallt. Der Begriff System steht also erstmal für eine black box mit einem vom Eingangsaudiosignal abhängigen Ausgang.

Der Begriff System kann also recht abstrakt für ziemlich viele Dinge stehen. Parameter, die ein System in seiner Grundfunktionalität beschreiben, sind vor allem die sog. *Linearität* und die *Zeitvarianz*. Ein System wird linear genannt, wenn es bestehende Signalanteile (z.B. Frequenzbänder) des Eingangssignals verstärkt oder dämpft, während ein sog. nichtlineares System dem Signal zuvor nicht enthaltene Signalanteile hinzufügt. Beispielsweise sind Verstärker oder Filter lineare Systeme, während ein Verzerrer oder ein Exciter nichtlineare Systeme sind.

Bemerkung

In der Realität besitzt jedes analoge System wie Verstärker und Filter auch einen nichtlinearen Anteil. Dieser Anteil soll dann möglichst klein sein; ein Beispiel für die Messung der Auswirkung nichtlinearer Systemanteile (Verzerrungen) ist die Klirrfaktormessung (vgl. Abschnitt chap:klirrfaktor).

Ein weiteres wichtiges Merkmal von Systemen ist ihre Zeitabhängigkeit. Ändern sich die Eigenschaften des Systems mit der Zeit, so nennt man das System zeitvariant, andernfalls zeitinvariant. Ein Filter ist ein zeitinvariantes System (wenn man nicht gerade die Filterparameter verändert), während ein Kompressor ein zeitvariantes System ist.

6.2.1 Lineare Systeme

Lineare Zeitinvariante System lassen sich in zwei Darstellungen beschreiben, dem *Übertragungsfrequenzgang* (auch Transferfunktion) $H(f)$ und der *Impulsantwort* $h(t)$. Dabei gilt, daß die Übertragungsfunktion die Fouriertransformation, das Spektrum, der Impulsantwort ist. Die Impulsantwort eines Systems ist - wie der Name schon sagt - die Antwort des Systems auf einen sehr kurzen Impuls, z.B. ist die Impulsantwort eines Raumes nach einem kurzen lauten Händeklatschen näherungsweise wahrzunehmen. Die Impulsantwort eines idealen Systems ist ein unendlich kurzer Impuls; der Übertragungsfrequenzgang eines idealen Systems ist ein horizontal verlaufendes Spektrum, d.h. daß jede Frequenz das System unverändert passieren kann.

6.2.1.1 Faltung

Mit Hilfe der Transferfunktion $H(f)$ läßt sich das Spektrum des Ausgangs $Y(f)$ aus dem Spektrum des Eingangs $X(f)$ berechnen:

$$Y(f) = X(f) \cdot H(f) \quad (6.1)$$

Da sich Spektrum und Zeitsignal beliebig ineinander überführen lassen, ist der Ausgang eines linearen Systems aufgrund seiner Übertragungsfunktion leicht zu bestimmen. Am Beispiel eines Filters, der alle höheren Frequenzen aus dem Signal herausfiltert (Tiefpaßfilter), läßt sich das veranschaulichen: alle Frequenzen über der Grenzfrequenz f_G werden abgeschnitten, während die darunterliegenden Frequenzanteile das System unbeeinflußt passieren. Damit enthält das Ausgangssignal keine hohen Frequenzen mehr.

Wie im Frequenzbereich über Multiplikation läßt sich im Zeitbereich der Ausgang eines Systems $y(t)$ mit der sog. *Faltung* (auch: Convolution) bestimmen. Bei der Faltungsoperation (eines Signals $x(t)$ mit einer Impulsantwort $h(t)$) wird das Signal an jedem Zeitpunkt mit der Impulsantwort multipliziert und das Ergebnis über alle Zeitpunkte aufsummiert. Die Faltung wird in der Literatur i.a. mit einem Stern gekennzeichnet:

$$y(t) = x(t) * h(t) \quad (6.2)$$

6.2.2 Nichtlineare Systeme

Nichtlineare System können dem Signal bisher noch nicht vorhandene Anteile hinzufügen. Das bedeutet, nichtlineare System führen zu *Verzerrungen*. Zu nichtlinearen Systemen gibt es keine so einfach zu verallgemeinernde Beschreibung wie für lineare Systeme; einen nichtlinearen Effekt kann man auf vielerlei Weise erreichen. Ein typisches Beispiel sind Übersteuerungen; aber Nichtlinearitäten werden auch bewußt eingesetzt, z.B. bei Distortioneffekten, beim Kompressor oder beim Exciter.

6.3 Grundlegende Signalformen bei der Signalsynthese

Um ein besseres Gefühl für Frequenzdarstellungen eines Signals zu bekommen, ist in Abb. 6.7 die Zeit- und Frequenzdarstellung von in der Klangsynthese häufig verwendeten tonalen Signalgrundformen dargestellt. Dabei handelt es sich um ein Sinussignal, ein Rechtecksignal, ein Sägezahnsignal und ein Dreiecksignal. Es ist deutlich zu sehen, daß die unterschiedlichen Signalverläufe in der Zeitdarstellung sich auch deutlich auf den Spektralverlauf „auswirken“.

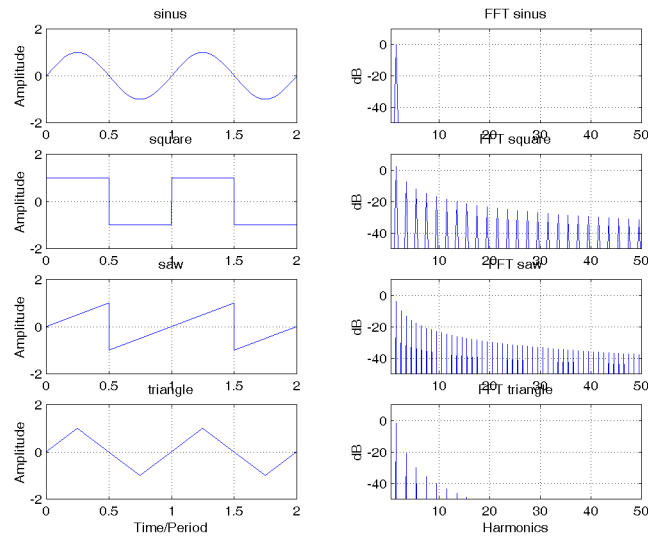


Abbildung 6.7: Sinus-, Rechteck-, Sägezahn-, und Dreieckssignal, links als Zeitsignal, rechts das Spektrum mit Achsenskalierung in Vielfachen der Grundfrequenz

Bemerkung

Als Daumenregel kann man davon ausgehen, daß der Anteil an hohen Frequenzen immer mehr zunimmt, je plötzlichere oder größere Sprünge ein Signalverlauf vorweist.

Eine weitere sehr wichtige, aber nicht so einfach definierbare Signalform ist das Rauschen. Die technische Definition des Rauschen besagt, daß es zu allen anderen Signalen inkohärent ist. Dies bedeutet z.B., daß theoretisch die Addition von zwei gleichlauten Rauschen immer nur eine Pegelerhöhung von $3dB$ ergeben kann.

Im Zusammenhang mit Rauschen werden oft Farben als nähere Bezeichnung für Rauschen verwendet, wie z.B. *weißes Rauschen* oder *rosa Rauschen*. Diese Farben beschreiben den Frequenzverlauf des Rauschen und sind damit auch so eine Art Klangfarbenbeschreibung des Rauschens. Abb. 6.8 zeigt verschiedenfarbige Rauschen als Zeitsignal und in Frequenzdarstellung. Ihre Energie nimmt hin zu hohen Frequenzen folgendermaßen ab:

- weißes Rauschen: $0dB/Octave$
- rosa Rauschen: $3dB/Octave$
- braunes Rauschen: $6dB/Octave$

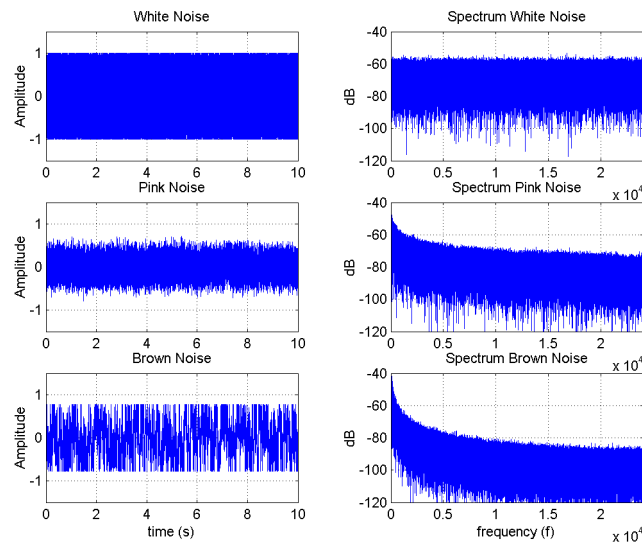


Abbildung 6.8: weißes, rosa, und braunes Rauschen, links als Zeitsignal, rechts das Spektrum

6.4 Geräte der Signalverarbeitung

In diesem Abschnitt sollen die typischen Tonstudiogeräte in Ihrer Funktionsweise beschrieben werden.

6.4.1 Fader

Ein **Fader** ist einfach nur ein Stellregler, mit dem die Amplitude eines Signals verändert werden kann. Fader findet man z.B. in Mischpulten. Sie verwenden praktisch immer eine dB-Skala, die im Bereich um 0dB besonders fein aufgelöst ist. Ein Fader ist ein lineares System und - wenn der Fader nicht bewegt wird - zeitinvariant.

Verwandte Themenbereiche

- Abschnitt [1.3](#)
- Abschnitt [3](#)

6.4.2 Filter

Filter (auch *Entzerrer*) sind die mit am häufigsten im Tonstudio eingesetzten Werkzeuge zur Klangbearbeitung und -verarbeitung.

Einfache Filter zur Klangbearbeitung kennt man zum Beispiel von der Stereoanlage oder einfacher Musiksoftware, wo sich Bässe, Mitten und Höhen anheben

bzw. absenken lassen. Genau das ist die Aufgabe eines Filters bzw. Entzerrers: die Anhebung oder Dämpfung bestimmter Frequenzbereiche. Für verschiedene Einsatzbereiche müssen Filter verschiedene Übertragungsfunktionen haben; der Verlauf dieser Übertragungsfunktionen bestimmt den *Filtertyp*.

Bevor die einzelnen Filtertypen kurz vorgestellt werden, müssen kurz einige Begriffe eingeführt werden:

- *Grenzfrequenz f_G* : Bei einseitigen Filtern die Frequenz (in *Hz*), ab der die Übertragungsfunktion eines Filter um mehr als $\pm 3dB$ von $0dB$ abweicht.
- *Mittenfrequenz f_C* : Bei symmetrischen Filtern die Frequenz (in *Hz*), um die die Übertragungsfunktion des Filters symmetrisch ist.
- *Bandbreite eines Filters B_G* : Bei symmetrischen Filtern die Breite des Frequenzbandes (in *Hz*) in dem die Übertragungsfunktion um weniger als $\pm 3dB$ von $0dB$ abweicht
- *Güte Q* : Bei symmetrischen Filtern ist die Güte ein aussagekräftigere Darstellung der Bandbreite. Die Güte wird berechnet durch das Verhältnis von

$$\frac{f_C}{B_G} \quad (6.3)$$

. Ist also die Bandbreite des Filters gleich der Mittenfrequenz, so nimmt die Güte den Wert 1 an. Ist die Bandbreite schmaler als die Mittenfrequenz, so wird die Güte größer als 1 und andersherum. Die Güte ist dimensionslos.

- *Flankensteilheit F* : Gibt in der Pseudo-Einheit $\left[\frac{dB}{Octave}\right]$ an, wie steil die abfallende Flanke der Übertragungsfunktion fällt. Typische Werte sind ganzzahlige Vielfache von 6 wie 6, 12, 18 und $24\frac{dB}{Octave}$.

Filter teilt man je nach Verlauf des Betragsfrequenzganges in verschiedene Kategorien ein. Die wichtigsten sind:

1. Tiefpaß-Filter:

Tiefe Frequenzen bis zur Grenzfrequenz f_G passieren das Filter (nahezu) unverändert, höhere Frequenzen werden zunehmend gedämpft. Die Stärke der Dämpfung nimmt mit der Flankensteilheit zu. Ein typischer Wert für die Flankensteilheit der meisten Tiefpaß-Filter liegt bei $12\frac{dB}{Octave}$. Die einstellbaren Parameter eines Tiefpasses können sein: *Grenzfrequenz*, *Flankensteilheit*

2. Hochpaß-Filter:

Alle Frequenzen über der Grenzfrequenz f_G passieren das Filter (nahezu) unverändert, Frequenzen darunter werden gedämpft. Je tiefer die Frequenz ist, desto stärker ist die Dämpfung, die mit der Flankensteilheit zunimmt. Das Hochpaß-Filter ist das Gegenstück zum Tiefpaßfilter, aus diesem Grund

können die einstellbaren Parameter wie beim Tiefpaß *Grenzfrequenz* und *Flankensteilheit* sein.

Sonderfälle des Hochpaß-Filters sind der *Trittschallfilter*, dessen Grenzfrequenz meist zwischen 50 und 180 Hz liegt und der zur Unterdrückung von Rumpelgeräuschen dient, und der *DC-Block-Filter* mit einer Grenzfrequenz um die 20 Hz zur Unterdrückung des Gleichanteils in einem Signal.

3. **Shelving-Filter** (manchmal auch Kuhschwanz-Filter):

Das Shelving-Filter scheint sich für den Anwender zunächst ähnlich zu verhalten wie Tiefpaß- oder Hochpaßfilter. Ab einer bestimmten Grenzfrequenz können Frequenzen angehoben oder abgesenkt werden. Der Unterschied zu den beiden oben genannten Filtertypen ist jedoch, daß diese Anhebung/Verstärkung oder Absenkung/Dämpfung für alle Frequenzen ab der Grenzfrequenz gleichstark ist und nicht mit zu- bzw. abnehmender Frequenz zunimmt. Je nachdem, ob die Anhebung oder Absenkung ober- oder unterhalb der Grenzfrequenz stattfindet, spricht man von *Höhenshelving* oder *Tiefenshelving*. Die einstellbaren Parameter eines Shelvingfilters sind *Grenzfrequenz* und *Anhebung/Dämpfung* in dB.

4. **Peak-Filter** (manchmal auch Präsenz-Filter):

Die Aufgabe des Peakfilters ist die Anhebung oder Absenkung eines bestimmten Frequenzbandes. Dazu kann i.a. die *Mittenfrequenz* des zu bearbeitenden Frequenzbandes und die *Güte* (und dadurch die Bandbreite) des Filters eingestellt werden.

5. **Bandpaß-Filter**: Das Bandpaß-Filter kann als Reihenschaltung von Tiefpaß mit hoher Grenzfrequenz und Hochpaß niedriger Grenzfrequenz interpretiert werden. Es läßt nur Frequenzen innerhalb eines bestimmten Frequenzbandes passieren. Einstellbare Parameter sind die *Mittenfrequenz*, die *Bandbreite* und die *Flankensteilheit* des Bandpasses.

6. **Bandsperre-Filter**:

Die Bandsperre ist das genaue Gegenteil des Bandpasses; alle Frequenzen, die nicht innerhalb eines bestimmten Frequenzbandes liegen, können das Filter (nahezu) ungehindert passieren. Frequenzen innerhalb des Frequenzbandes werden unterdrückt. Die einstellbaren Parameter sind wie beim Bandpaß die *Mittenfrequenz*, die *Bandbreite* und die *Flankensteilheit*.

7. **Notch-Filter**:

Das Notchfilter kann aus Anwendersicht als Spezialfall der Bandsperre mit einem sehr schmalen Frequenzband interpretiert werden. Alle Frequenzen bis auf eine ganz spezielle können das Filter (nahezu) ungehindert passieren, diese Frequenz wird komplett herausgefiltert. Als Parameter läßt sich zumeist nur die *Mittenfrequenz* einstellen, manchmal auch die *Güte* des Filters.

Der Einsatzzweck eines Notch-Filters ist zum Beispiel das Herausfiltern von störendem Netzbrummen von einer Aufnahme. Dabei werden meistens mehrere Notchfilter eingesetzt, deren Mittenfrequenzen ganzzahlige Vielfache der Grundfrequenz sind, da Netzbrummen normalerweise nicht nur aus einem sinusförmigen Signal besteht.

Die Phasenverschiebungen um die Mittenfrequenz sind bei einem Notch-Filter sehr stark, und das gefilterte Signal sollte nach der Bearbeitung daraufhin überprüft werden.

Verschiedene dieser Filter (v.a. Tiefpaß und Bandpaß) gibt es auch als sogenannte *Resonanz-Filter*. Diese zeichnen sich dadurch aus, daß sie an der Grenzfrequenz bzw. der Mittenfrequenz noch eine starke (regelbare) Anhebung haben. Dadurch wird klanglich diese Frequenz betont, was insbesondere bei Filterungen interessant ist, wo zeitliche Veränderungen der Grenzfrequenz/Mittenfrequenz als musikalisches Stilmittel eingesetzt werden.

Filter unterschiedlicher Hersteller klingen im allgemeinen unterschiedlich. Diese Klangunterschiede sind das Resultat unterschiedlicher Herangehensweise beim Filterentwurf. Bei analogen Filtern spielen auch die verwendeten Bauteile eine Rolle. Diskussionsstoff liefert oft auch der Phasenfrequenzgang eines Filters. Hier läßt sich nicht abschließend sagen, welchen Einfluß dieser auf die empfundene Qualität eines Filters hat. Filter mit linearen Phasengang sind in der Analogwelt nicht realisierbar. Ein Filter ist ein lineares System und - wenn die Parameter nicht geändert werden - zeitinvariant.

Mit dem Einsatz von Filtern lassen sich auch andere Effekte erreichen, wie z.B. der **Phaser**-Effekt. Dieser Effekt lässt sich erreichen durch eine Reihenschaltung mehrerer Notch-Filter mit unterschiedlichen Mittenfrequenzen. Dieser Mittenfrequenzen werden durch einen Oszillator leicht moduliert. Die starken Phasenverschiebungen um die Mittenfrequenz des Notchfilters führen bei der Addition mit dem Originalsignal zu Interferenzen (vgl. Abschnitt 1.4, die sich zeitlich mit der Frequenz des Oszillators verändern).

Verwandte Themenbereiche

- Abschnitt 5.5
- Abschnitt 1.5
- Abschnitt 1.4

6.4.3 Delay und Delay-basierte Effekte

Das **Delay** ist ein Effekt, der schon seit langer Zeit insbesondere im Pop-/Rockbereich eingesetzt wird. Im einfachsten Fall wird dieser Effekt erzeugt, indem das Eingangssignal verzögert wird und gedämpft dem Originalsignal hinzugemischt wird. Je nach Intensität des verzögerten Signals und der Verzögerungszeit entsteht so oft der Eindruck eines Echos (vgl. Abschnitt 5.5.1 zur Echoschwelle). Auch

wenn man diesen einfachen Fall schon als Delayeffekt bezeichnen kann, so sind die i.a. eingesetzten Delayeffekte noch etwas ausgefeilter. Der typische Delayeffekt ist rückgekoppelt, d.h. der Ausgang des Effekts wird (etwas gedämpft) wiederum auf ein Verzögerungsglied gegeben. Auf diese Art und Weise erreicht man ein ständig leiser werdendes Ausklingen. Abhängig von den Einstellungen (Zeit der Verzögerung, Dämpfung des verzögerten Signals) läßt sich die Länge und die Dichte des Effekts von einem beatweise wiederholen bis zum hallähnlichen Nachklang variieren.

Ein Delay ist i.a. ein lineares System und meistens zeitinvariant.

Andere Effekte bauen auf der zeitlichen Veränderung der Verzögerungszeit auf. Der **Flanger** arbeitet zumeist mit einer Zeit $< 15ms$, die kontinuierlich - beispielsweise sinusförmig - moduliert werden. Die Modulationsgeschwindigkeit liegt in der Größenordnung von $1Hz$.

Der **Chorus** ähnelt dem Funktionsprinzip des Flangers sehr. Hier wird aber nicht nur ein verzögertes Signal auf das Originalsignal addiert, sondern mehrere mit unterschiedlichen Delayzeiten (meistens zwischen 10 und $25ms$). Die Verzögerungszeiten werden dabei ständig leicht verändert. Im Gegensatz zum Flanger sind diese kleinen Änderungen zufällig, und werden nicht vom einem deterministischen Signal wie einem Sinus hervorgerufen.

Flanger und Chorus sind lineare zeitinvariante Systeme. Klanglich ist der Phaser ein verwandter Effekt.

Verwandte Themenbereiche

- Abschnitt [5.5.1](#)

6.4.4 Hallgerät

Aufgabe eines **Hallgerätes** ist es, das Eingangssignal in einer Weise zu verarbeiten, daß es mit einem räumlichen Eindruck versehen wird, wie eine Schallquelle in einem Raum. Technisch gesprochen versucht ein Hallgerät die Raumimpulsantwort eines Raumes nachzubilden, und das Eingangssignal mit dieser zu falten. Es existieren auch Hallgeräte, die auf real gemessenen Rauminpulsantworten beruhen, und die das Eingangssignal mit dieser Impulsantwort falten. Zur Optimierung des Halls für das spezielle Eingangssignal und die Anwendung werden meistens die unterschiedlichsten Räume als Presets angeboten, die dann mit verschiedenen Parametern angepaßt werden können. Die wichtigsten Parameter sind hierbei die Nachhallzeit, verschiedene Filterparameter insbesondere zur Tiefenanhebung/-absenkung oder Hochpaßfilterung und das sogenannte Pre-Delay, mit dem sich eine Verzögerung des verhallten Signals erzielen läßt.

Ein Hallgerät ist ein lineares System und - wenn die Parameter nicht geändert werden - zeitinvariant.

Verwandte Themenbereiche

- Abschnitt 4
- Abschnitt 6.2.1.1

6.4.5 Dynamikbearbeitung

Geräte zur Dynamikbearbeitung regeln im allgemeinen die Verstärkung/Dämpfung für ein bestimmtes Eingangssignal automatisch in Abhängigkeit des Eingangspegels. Dazu kann man bei fast allen Geräten die sogenannte *Attack Time* und *Release Time* einstellen. Die *Attack Time* ist eine Zeitkonstante zur Einstellung der Reaktionszeit des Systems auf plötzliche Peaks oder laute Passagen, während über die *Release Time* die Reaktionszeit beim Wechsel von lauten hin zu leisen Passagen einstellbar ist. Bekannte Geräte zur Dynamikbearbeitung sind:

1. Noise Gate:

Die Aufgabe eines Noise Gates ist, bei Signalen mit einem geringeren Pegel als der eingestellten *Threshold* kein Signal auszugeben, bei lauterem Signalen das Signal allerdings unverändert zu lassen. Dies geschieht über einen pegelabhängigen *Gain Factor* der sich allerdings nicht unmittelbar ändert. Oft lassen sich auch über die *Attack Time* und die *Release Time* die Reaktionszeiten des Systems einstellen. Zu kurze Zeiten führen hierbei oft zu einem zerstückelten Klang, während zu lange Reaktionszeiten entweder dazu führen, daß der Anfang relevanter, laute, Passagen fehlt oder das Gate bei nichtgewollten Signalen mit kleinem Pegel zur spät „zumacht“.

2. Limiter:

Ein Limiter soll starke Peaks eines Signals kontrollieren (insbesondere um Übersteuerungen zu vermeiden), das Eingangssignal und seinen Dynamikbereich aber möglichst unverändert lassen. Hierzu muß die Reaktionszeit des Systems insbesondere für die Detektion von lauten Signalanteilen möglichst kurz sein. Wird der Eingangspegel des Signals bei gleichen Limitereinstellungen angehoben, so resultiert dies sowohl in einer Lautstärkeerhöhung des Signals als auch in einigen zusätzlichen hohen Frequenzenanteilen.

3. Compressor und Expander:

Compressoren sind die am häufigsten eingesetzten Geräte zur Dynamikbearbeitung. Sie werden zur Eingrenzung des Dynamikumfangs benutzt. Signalanteile mit kleinem Pegel passieren das System unbeeinflusst, während hohe Pegel oder laute Teile entsprechend der Kompressorkennlinie gedämpft werden. Durch Anhebung des Eingangspegels kann so ein lauterer Signal mit geringem Dynamikumfang erreicht werden.

Expander tun genau das Gegenteil von Kompressoren; sie senken leise Signalanteile im Pegel ab, während sie laute Signalanteile anheben. Das Resultat ist eine „lebendigere“ Klangcharakteristik.

Es gibt eine ganze Reihe Effektgeräte oder PlugIns, die prinzipiell zur Familie der Dynamikprozessoren gehören, aber aufgrund von Modifikationen im Algorithmus und anderen Gründen anders genannt werden. Mögliche Namen sind z.B. Finalizer, Maximizer, etc.

Die Geräte zur Dynamikbearbeitung sind i.a. nichtlineare und zeitvariante Systeme.

Verwandte Themenbereiche

- Abschnitt [3.2](#)
- Abschnitt [5.2](#)

6.4.6 Verzerrer und Enhancer

6.4.7 andere Effekte

Ein sog. **De-Esser** wird zur Verarbeitung von Sprache und Gesang verwendet. Bei diesen Signalen werden die Zisch- und S-laute oft als zu dominant und störend empfunden. Die Aufgabe eines De-Essers ist es, diese Signalanteile, die meistens im Bereich zwischen 2 und 6kHz liegen, zu dämpfen. Hierzu wird meistens ein Filter verwendet, dessen Dämpfung abhängig von der Signalamplitude in diesem Frequenzbereich ist.

Röhrensimulation, Tape Saturation

6.5 Qualitätsmerkmale

Die Einschätzung der Qualität von verwendetem Equipment kann sehr wichtig sein. Wenn man selbst keine Messungen durchführen kann oder will, ist man zumeist auf Tests in Fachzeitschriften und die Herstellerangaben angewiesen. Die dabei häufig angegebenen Daten sollen in diesem Abschnitt kurz erläutert werden.

6.5.1 Übertragungsfrequenzgang

Der **Übertragungsfrequenzgang** eines Gerätes ist v.a. bei Verstärkern oder (elektroakustischen) Wandlern von Interesse, da diese - im Gegensatz zu Effekten - das Signal möglichst wenig beeinflussen sollen. Angestrebt wird daher ein möglichst linearer Frequenzgang über einen möglichst großen Frequenzbereich. Die Linearität wird dabei meist durch die maximale Abweichung (wie z.B. $\pm 0.5dB$) vom idealen (horizontal) verlaufenden Frequenzgang angegeben. Dazu wird der Frequenzbereich (z.B. 15Hz - 22kHz) angegeben, für den diese maximale Abweichung gilt. Diese Angaben lassen allerdings keine Rückschlüsse auf den Phasengang des Systems zu.

Die (Hersteller-)Angabe eines Frequenzganges ist umso aussagekräftiger, je höher die Auflösung der dB-Skala ist. Manchmal wird der Frequenzgang auf

einer Skala mit dermaßen hoher Auflösung gezeigt, daß eine Aussage über seine Linearität gar nicht mehr möglich ist. Auch durch Verdickung der aufgetragenen Linie wirkt der Frequenzgang für das Auge des flüchtigen Betrachters zunächst linear, obwohl er das nicht sein muß.

6.5.2 Dynamikumfang

Der **Dynamikumfang** eines Systems ist leicht vereinfacht gesagt der Pegelwert, um den ein vollausgesteuertes Signal gedämpft werden kann, bis es im Grundrauschen des Systems verschwindet. Dieser Wert berücksichtigt keine nicht-linearen Verzerrungen und wird in dB angegeben. Dieser Wert ist vergleichbar zum Signal-Rauschabstand (SNR) eines Systems.

6.5.2.1 Bewertungsfilter

Das Ohr hat, wie in Abschnitt 5.2 beschrieben, keinen linearen Frequenzgang. Die spektrale Verteilung des Systemrauschens bei der Messung des Dynamikumfangs wird bei einer Messung wie oben beschrieben allerdings nicht berücksichtigt. Wenn das Rauschen beispielsweise sehr starke Anteile oberhalb von 12kHz hätte, so würden diese genauso stark gewichtet wie Anteile um 3kHz , was nicht dem subjektiven menschlichen Lautstärkeempfinden entspräche.

Aus diesem Grund gibt es die sogenannten **Bewertungsfilter** oder **Gewichtungsfilter**, welche grob den Verlauf der Kurven gleicher Lautstärkepegel approximieren. Die in der Tonmeßtechnik am häufigsten verwendete Bewertungskurve ist die sogenannte *A-Bewertung*, da diese der Empfindlichkeit des Ohres für geringe Pegel nachempfunden ist. Die B- und C-Kurven sind für Messungen bei mittlerem bzw. hohem Pegel gedacht. Werte, die mit A-Bewertung gemessen wurden, werden durch ein angehängtes A kenntlich gemacht, z.B. $95\text{dB}(A)$ oder $110\text{dB}(C)$. Zu beachten ist, daß A-bewertete Messungen praktisch immer „bessere“ Ergebnisse liefern als unbewertete.

Als alternative Bewertungskurve zu der A-Bewertung existiert noch die CCIR-Bewertungskurve.

Verwandte Themenbereiche

- Abschnitt 3.2

6.5.3 Klirrfaktor

Der **Klirrfaktor** (englisch: *THD* für *Total Harmonic Distortion*) eines Systems ist ein Maß für seine nichtlinearen Verzerrungen. Ein nichtlineares System erzeugt Verzerrungen bei Frequenzen, die ganzzahligen Vielfachen der Eingangsfrequenz entsprechen. Zur Messung des Klirrfaktors wird an den Eingang des Systems ein

Sinussignal (meistens der Frequenz 1000Hz gelegt. Der Klirrfaktor k berechnet sich dann mit

$$k = \sqrt{k_2^2 + k_3^2}, \quad (6.4)$$

wobei die Summanden das Verhältnis des Spannungswertes bei der n -fachen Grundfrequenz und der Gesamtspannung des Systems U_{ges} sind:

$$k_n = \frac{U_n}{U_{Ges}} \quad (6.5)$$

Der Pegel des Sinustons ist von immenser Bedeutung für die Aussagekraft der Messung. Im allgemeinen wächst der Klirrfaktor mit zunehmendem Eingangspegel, so daß die Messung möglichst bei Volllaussteuerung durchgeführt werden sollte.

Manchmal wird statt des Klirrfaktors auch das sogenannte *Klirrdämpfungsmaß* a_k angegeben:

$$a_k = 20 \log \left(\frac{1}{k} \right) \quad (6.6)$$

6.5.4 Total Harmonic Distortion and Noise

Die Messung des **THD+N**-Wertes ist wie der Klirrfaktor ein Maß für die nichtlinearen Verzerrungen, berücksichtigt aber auch das Rauschen des Testsystems. Im Falle eines idealen linearen Systems (ohne nichtlineare Verzerrungen) entsprechen die THD+N-Werte dem Dynamikumfang des Systems.

6.5.5 Modulationsverzerrungen

Durch eine nichtlineare Verzerrung entstehen nicht nur harmonische Vielfache der Grundfrequenz, sondern auch sogenannte Summen- und Differenztöne, die bei der Klirrfaktormessung nicht berücksichtigt werden. Die Stärke dieser Verzerrungen versucht man mit dem kaum mehr gebräuchlichen **Intermodulationsfaktor**, für dessen Messung die Summe der Spannungen an den Frequenzen der Summen- und Differenztöne gebildet wird, oder mit dem sog. **Differenztonfaktor** zu messen. Der Differenztonfaktor gibt das Verhältnis von der Spannung bei den Differenztonfrequenzen zu der Gesamtspannung an. Es können verschiedene Ordnungen des Differenztonfaktors angegeben werden:

$$d_2 = \frac{U_{f_2-f_1}}{\sqrt{2} \cdot U_{Ges}} \quad (6.7)$$

$$d_3 = \frac{U_{2f_2-f_1} + U_{2f_1-f_2}}{\sqrt{2} \cdot U_{Ges}} \quad (6.8)$$

Die Differenz von f_2 und f_1 sollte dabei 70Hz betragen.

6.5.6 Übersprechdämpfung

Die **Übersprechdämpfung** eines mehrkanaligen Systems gibt an, mit welcher Dämpfung in dB das Audiosignal eines Kanals auf dem anderen zu hören ist. Übersprechen sollte in digitalen Geräten nicht auftreten.

6.5.7 Hörtests

Mit den beschriebenen Merkmalen lassen sich zeitvariante Systeme kaum in Ihrer Qualität beurteilen. Desweiteren sind auch Systeme vorstellbar, die technisch schlechtere Meßwerte aufweisen, aber dennoch subjektiv besser klingen. Um solche Qualitätskriterien besser in den Griff zu bekommen, kann man Hörtests durchführen, bei denen die Testhörer die Qualität (u.U im Vergleich zu einem Referenzsignal) aufgrund ihres subjektiven Eindrucks beurteilen. Insbesondere für Systeme wie AD/DA-Wandler (vgl. Abschnitt 8.2.3) oder Codierungsverfahren (vgl. Abschnitt 8.6) sind solche Hörtests von Bedeutung.

Die Durchführung von Hörtests ist aber zumeist sehr aufwendig, wenn die Ergebnisse aussagekräftig sein sollen. In der Vorbereitung muß man sich genaue Gedanken machen über die Auswahl des Testmaterials, die Auswahl und Menge der Testhörer, die Testmethode (Bewertungsskala, Referenzsignal, etc.), die Testbedingungen wie z.B. die verwendeten Kopfhörer oder Lautsprecher und die Akustik des Raums und letztendlich auch die Methode zur statistischen Auswertung dieser Tests. Aufgrund dieser Komplexität bei Planung und Durchführung eines Hörtests sowie um Ergebnisse unterschiedlicher Hörtests miteinander vergleichbar zu machen, gibt es internationale Standards, die sich mit genau dieser Thematik auseinandersetzen (s. z.B. [IR94]).

Verwandte Themenbereiche

- Abschnitt 8.2.3
- Abschnitt 8.6

6.6 Zusammenfassung

Es existieren verschiedene Möglichkeiten, ein Audiosignal und seine Eigenschaften darzustellen. Die wichtigsten sind der **Zeitverlauf**, bei dem die Amplitudenwerte über der Zeit aufgetragen werden, verschiedene Formen des **Spektrums**, wo meistens die Amplitude einer bestimmten Frequenz für alle im Signal enthaltenen Frequenzen angezeigt wird, und die Darstellung in Form der **Amplitudendichteverteilung**, wo die Häufigkeit bestimmter Amplitudenwerte über das Signal bestimmt wurden.

Jedes Gerät in der Tontechnik, sei es ein Effektgerät, ein Mischpult, ein Mikrofon oder ein Lautsprecher, selbst ein Raum kann technisch als **System**

interpretiert werden. Für Systeme sind die übergeordnetsten Kategorien **Linearität**/bldNichtlinearität und **Zeit(in-)varianz**. Nichtlineare Systeme fügen im Gegensatz von linearen Systemen neue Frequenzanteile in ein Signal ein. Der Ausgang von linearen und zeitinvarianten Systemen läßt sich über die sogenannte **Faltung** berechnen.

Es existieren Studiogeräte für die unterschiedlichsten Einsatzbereiche. Diese können sowohl linear als auch nichtlinear, sowohl zeitinvariant als auch zeitvariant sein.

Bei der Messung von Audioqualität ist es von großer Wichtigkeit, die von Hersteller angegebenen Werte richtig zu interpretieren. Die wichtigsten Punkte sind hier der lineare **Übertragungsfrequenzgang**, der **Dynamikumfang** und geringe **nichtlineare Verzerrungen**.

6.7 Aufgaben

1. Warum ist ein Fader ein lineares System, obwohl seine Skala doch nichtlinear ist?
2. Berechne die Güte Q eines Filters mit der Mittenfrequenz $3.2kHz$ mit folgenden Bandbreiten:
 - $1600Hz$,
 - $3200Hz$,
 - $4000Hz$
 - $6400Hz$
3. Ab welcher Frequenz ist bei einem Tiefpaß mit der Grenzfrequenz $100Hz$ und der Flankensteilheit $12\frac{dB}{Octave}$ die Dämpfung größer als $15dB$?
4. Auf welche Zeitspanne muß die Verzögerungszeit eines Delayeffekts gestellt werden, wenn der Effekt „im Takt“ klingen soll und das Musikstück ein Tempo von 90 Schlägen pro Minute (BPM = Beats per minute) besitzt?
5. Auf wieviel Prozent muß der Anteil des Rückkopplungsziweiges eines Delayeffekts gesetzt werden, wenn der Pegel jeder Rückkopplung um $6dB$ abnehmen soll?
6. Mit dem Pre-Delay verzögert man das verhallte Signal gegenüber dem trockenen, unverhallten Signal. Wird bei Erhöhung des Pre-Delays der Gesamteindruck eher nach hinten, in die Tiefe des Raums rutschen oder im Gegenteil präsenter bzw. näher klingen?
7. Compressor-Einstellung

Kapitel 7

Elektroakustische Wandler

Der Sinn eines elektroakustischen Wandlers ist die Umwandlung einer physikalischen Größe des Schallfeldes wie z.B. den Schalldruck oder die Schallschnelle in eine elektrische Größe wie z.B. eine Spannung umzuwandeln, oder umgekehrt (Schallsender). Alle in der Tonstudioteknik gebräuchlichen Wandler besitzen eine Membran, welche durch das Schallfeld in Schwingungen versetzt wird, und einen elektrischen Teil, der diese mechanischen Schwingungen in elektrische Schwingungen wandelt.

7.1 Mikrophone

Mikrophone als Schallempfänger wandeln also akustische Energie in elektrische Energie.

Einfluß auf den Prozeß dieser Wandlung haben sowohl der *mechanische Aufbau* als auch der *elektrische Aufbau* des Mikrophons. Diesen beiden Teile werden, nach einer kurzen Erläuterung einiger grundlegender Wandlerkenngrößen, umfassend behandelt. Anschließend wird der Einfluß der Bauform auf das Mikrophons erläutert.

7.1.1 Charakteristische Kenngrößen

7.1.1.1 Pegel-Kenngrößen

Die in Mikrophondatenblättern angegebenen Größen sind oft einzahlige Pegelwerte, deren Aussage im folgenden kurz beschrieben werden soll.

- **Feldübertragungsfaktor**

Der Feldübertragungsfaktor gibt den Effektivwert der Ausgangsspannung an, wenn das Mikrophon einem Schalldruck von $1Pa$ bei einer Frequenz von $1kHz$ ausgesetzt wird (in [V] oder [dBV]).

- **Grenzschalldruckpegel**

Der Grenzschalldruckpegel eines Mikrophons ist der Schalldruckpegel, ab dem der Klirrfaktor größer als 0.5% wird (in [dBSPL]).

Verwandte Themenbereiche

- Abschnitt [6.5.3](#)

- **Maximaler Ausgangspegel**

Der maximale Ausgangspegel gibt den Effektivwert der Spannung am Ausgang des Mikrophons beim Grenzschalldruck an (in [V] oder [dBu]).

- **Eigenstörspannung**

Die Eigenstörspannung ist die effektive Ausgangsspannung des Mikrophons, wenn es keinem Schall ausgesetzt ist (in [V] oder [dBu]).

- **Ersatzgeräuschpegel**

- **Geräuschpegelabstand**

- **Dynamikumfang**

Der Dynamikumfang ergibt sich aus der Differenz zwischen Grenzschalldruckpegel und Ersatzgeräuschpegel.

7.1.1.2 Übertragungsfrequenzgang

Ein wichtiges Merkmal zur Charakterisierung eines Mikrophons ist natürlich sein Übertragungsfrequenzgang (vgl. [6.5.1](#)). Soll das Mikrophon das Schallereignis neutral übertragen, so ist natürlich ein linearer Frequenzgang mit gleichbleibendem Feldübertragungsmaß für alle Frequenzen wünschenswert. Es werden allerdings in vielen Fällen bewußt auch Mikrophone mit Abweichungen von diesem Frequenzgang produziert und eingesetzt, da sich auf diese Weise ein anderer, gewünschter Klangeindruck erzielen läßt.

Bei der Angabe des Übertragungsfrequenzganges eines Mikrophons wird zu- meist nicht nur eine einzelne Kurve angegeben, sondern mehrere für den Schallein- fall aus verschiedenen Richtungen. Diese sind sogenannte **Freifeldfrequenzgänge**.

Freifeld/Diffusfeld

Der Übertragungsfrequenzgang eines Mikrophons setzt sich hauptsächlich aus zwei Komponenten zusammen: die Übertragungsfunktion des *mechanischen Aufbaus* und die des *elektrischen Aufbaus*.

Verwandte Themenbereiche

- Abschnitt [6.5.1](#)

7.1.1.3 Richtcharakteristik

Es ist anschaulich leicht vorzustellen, daß ein Mikrophon nicht unbedingt in alle Richtungen gleich empfindlich ist, d.h. bei Besprechen eines Mikrophons von hinten oder von der Seite ist das Mikrophon unter Umständen wesentlich unempfindlicher. Für verschiedene Einsatzzwecke gibt es verschiedene sogenannte **Richtcharakteristiken**, denn manchmal soll die Empfindlichkeit eines Mikrophons besonders in bestimmte Richtungen hoch sein, manchmal sollen aber alle Richtungen gleich behandelt werden. Genau dieses Verhalten wird mit der Richtcharakteristik (s. Abb. 7.1 für die geläufigsten Richtcharakteristiken) angegeben, wo eine Linie die Empfindlichkeit für verschiedene Schalleinfallrichtungen zeigt. Bei der Darstellung der Richtcharakteristiken befindet sich in der Mitte das gedachte Mikrophon, das in die 0° -Richtung zeigt. Dann wird für jeden Winkel die Empfindlichkeit des Mikrophons aufgetragen und man erhält eine sich schließende Linie - die Richtcharakteristik. Technisch ausgedrückt wird der Feldübertragungsfaktor für jeden Einfallswinkel aufgetragen. Setzt man an einer bestimmten Einfallrichtung den Feldübertragungsfaktor ins Verhältnis zum Feldübertragungsfaktor der 0° -Bezugsrichtung, so erhält man den sogenannten **Richtungsfaktor** Γ des Mikrophons für diesen Einfallswinkel θ . Der 20-fache Logarithmus des Richtungsfaktors ist das **Richtungsmaß**. Somit ist das Richtungsmaß für die Bezugsrichtung immer 0dB .

Die dargestellten Richtcharakteristiken sind idealisiert. In den Datenblättern realer Mikrophone kann man deutliche Abweichungen zu den idealen Kurven sehen. Insbesondere spielt hier die Frequenzabhängigkeit der Richtcharakteristik eine große Rolle. So führen Reflexions- und Beugungseffekte am Mikrophonkorpus i.a. dazu, daß die Empfindlichkeit des Mikrophones sich verändert. Eigentlich kugelförmige Richtcharakteristiken können so beispielsweise bei höheren Frequenzen deutlich gerichtet werden. Aus diesem Grund werden von Mikrophonherstellern i.a. verschiedene Graphen für verschiedene Frequenzen angegeben.

Die Richtwirkung des Mikrophones wird teilweise auch durch einzelne Werte beschrieben, v.a. dem **Bündelungsgrad** und dem **Bündelungsmaß**. Der Bündelungsgrad vergleicht die Richtcharakteristik des Mikrophons mit einer Kugelcharakteristik. Je geringer die Leistung des Raumschalls im Verhältnis zu der Raumschalleistung eines Mikrophons mit Kugelcharakteristik und dem gleichen Feldübertragungsmaß ist, desto höher ist der Bündelungsgrad.

Bemerkung

Mathematisch berechnet sich der Bündelungsgrad γ aus dem Richtungsfaktor Γ in Abhängigkeit vom Einfallswinkel θ :

$$\gamma = \frac{2}{\int_0^\pi \Gamma^2(\theta) \sin(\theta) d\theta} \quad (7.1)$$

Das Bündelungsmaß ist der 10fache Logarithmus des Bündelungsgrades und ist

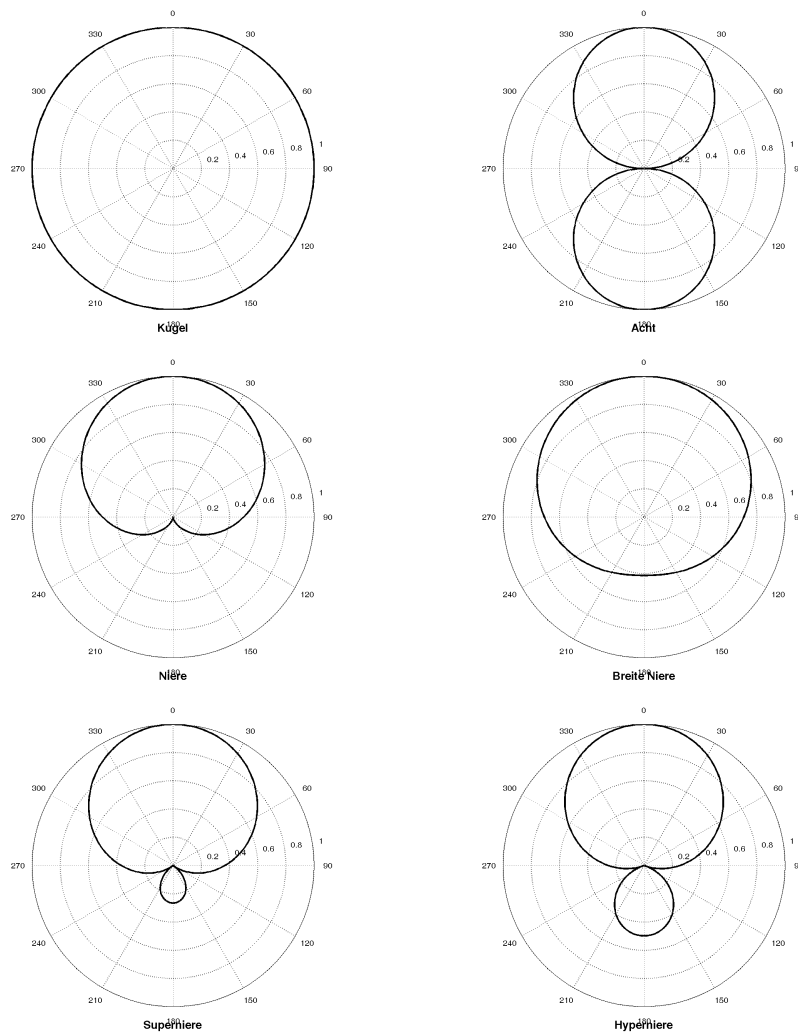


Abbildung 7.1: Richtcharakteristiken von Mikrofonen; oben Kugel, Niere und Acht, unten Breite Niere, Superniere und Hypernieren

dementsprechend 0 für Mikrophone mit Kugelcharakteristik. Auch einige andere, mit dem Bündelungsgrad verwandte Größen werden manchmal angegeben:

- **Vergrößerungsfaktor** (auch *(relativer) Abstandsfaktor* oder *Distance Factor* (DSF)): dieser Wert gibt an, wie weit der Abstand des Mikrophons zur Schallquelle im Vergleich zur Kugelcharakteristik entfernt werden kann, so daß das Verhältnis von Direkt- und Diffusschall gleich bleibt. Der Vergrößerungsfaktor berechnet sich über die Quadratwurzel aus dem Bündelungsgrad.
- **Random Energy Efficiency** (REE): aufgenommener Störschall im Verhältnis zur Kugelcharakteristik. Dieser Wert ist reziprok zum Bündelungsgrad.

Abb. 7.2 zeigt die verschiedenen Maße Bündelungsgrad, DSF und REE in der Übersicht.

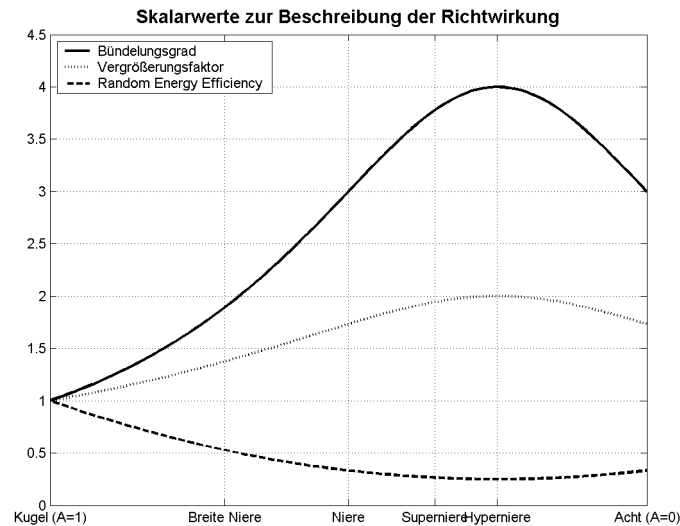


Abbildung 7.2: Bündelungsgrad, DSF und REE für verschiedene Richtcharakteristiken

Verwandte Themenbereiche

- Abschnitt [4.4](#)

7.1.1.4 Impulsverhalten

Auch wenn zwei unterschiedliche Mikrophone vergleichbare Kenndaten haben, können sie dennoch unterschiedlich klingen. Dies liegt dann meistens am unterschiedlichen **Impulsverhalten** der Mikrophone. Dabei wird nichts anderes als die Impulsantwort (vgl. Abschnitt [6.2.1](#)) des Mikrophons gemessen. Je nach Mikrophon ist die Impulsantwort unterschiedlich lang und teilweise klanglich verfärbt. Dies kann sich deutlich auf das Klangbild auswirken, muß aber nicht unerwünscht sein.

7.1.2 Mikrophontypen

Praktisch alle gebräuchlichen Studiomikrophone verfügen über eine Membran, die von Bewegungen der Luftteilchen in Schwingung versetzt wird wie auch das menschliche Trommelfell. Abhängig davon, wie die Schwingungen in elektrische Signale umgewandelt werden spricht man von einem unterschiedlichen **Arbeitsprinzip** des Mikrophons. Hier trennt man zwischen den **Elongationsempfängern**,

bei denen die erzeugte elektrische Spannung proportional zur Auslenkung der Membran ist, und den **Geschwindigkeitsempfängern**, bei denen diese Spannung proportional zur Geschwindigkeit der Membran ist. Im folgenden sollen die beiden wichtigsten Vertreter dieser beiden Gattungen eingeführt werden: das Kondensatormikrophon als bekanntester Elongationsempfänger und das dynamische Mikrophon als dem verbreitetsten Geschwindigkeitsempfänger.

7.1.2.1 Kondensatormikrophon

Mechanischer Aufbau

Die Größen, die im mechanischen Aufbau des Kondensatormikrophons die Schwingungseigenschaften und damit den Übertragungsfrequenzgang beeinflussen können, sind in Abb. 7.3 dargestellt. Die wichtigsten sind:

- die Masse der Membran,
- die Steife des Luftpolsters hinter der Membran, und
- die Stärke der Reibung an der Membranaufhängung.

In der Mechanik spricht man in diesem Fall von einem *einfachen Schwinger*.

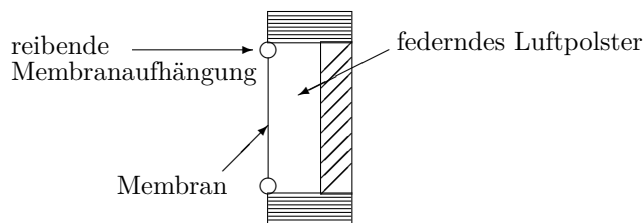


Abbildung 7.3: mechanischer Aufbau eines Mikrophons

Die Eigenschaften bei der Bewegung der Mikrophonmembran werden daher maßgeblich von folgenden Kräften bestimmt:

- F_m : die Kraft, um die Membranmasse zu beschleunigen/in Bewegung zu setzen
- F_r : die Dämpfungs- oder Reibungskraft
- F_s : die Kraft, mit der die Steife des Luftpolsters die Schwingung beeinflusst

Die Kraft, die das System von außen in Bewegung setzt, muß gleich der Summe dieser drei Kräfte sein:

$$F = F_m + F_r + F_s \quad (7.2)$$

Aus dieser Gleichung lässt sich die in Abb. 7.4 dargestellte frequenzabhängige Membranauslenkung in Abhängigkeit der Membranmasse m , der Dämpfungskonstante r und der Federkonstante s herleiten. Das Kondensatormikrophon als Elongationsempfänger ist direkt von der Auslenkung der Membran abhängig, daher kann man bei der dargestellten Kurve von der Übertragungsfunktion des mechanischen Teils des Kondensatormikrophons sprechen.

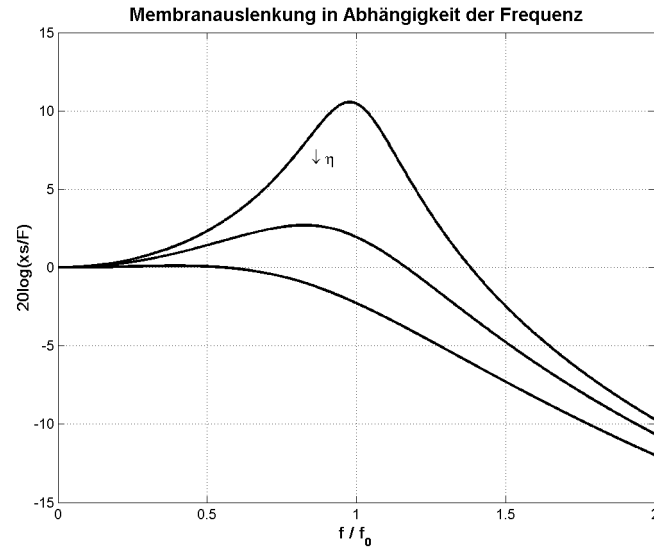


Abbildung 7.4: Übertragungsfrequenzgang eines Elongationsempfängers in Abhängigkeit verschiedener Dämpfungswerte

Herleitung

In der Mechanik ergibt sich für Glg. (7.2)

$$F = m\ddot{x} + r\dot{x} + sx \quad (7.3)$$

Jeder Punkt steht hierbei für eine Ableitung. Nimmt man eine cosinusförmige Membranauslenkung an und schreibt diese als komplexe Zahl, so erhält man (mit $x = e^{j2\pi ft}$):

$$\begin{aligned} F &= -m(2\pi f)^2 e^{j2\pi ft} + jr2\pi f e^{j2\pi ft} + se^{j2\pi ft} \\ &= (-m(2\pi f)^2 + jr2\pi f + s) x \end{aligned} \quad (7.4)$$

Die gesuchte Übertragungsfunktion ist die Auslenkung x durch die Kraft F :

$$\begin{aligned}
 x/F &= \frac{1}{-m(2\pi f)^2 + jr2\pi f + s} \\
 &= \frac{1/s}{-\underbrace{\frac{m}{s}}_{:=(2\pi f_0)^2} (2\pi f)^2 + j\underbrace{\frac{r}{s}}_{:=\eta/(2\pi f_0)} 2\pi f + 1} \\
 &= \frac{1/s}{1 - \frac{f^2}{f_0^2} + j\eta \frac{f}{f_0}} \tag{7.5}
 \end{aligned}$$

Aus diesem Ergebnis für die Auslenkung der Membran lassen sich folgende Schlüsse ziehen:

- ist die Anregungsfrequenz f sehr klein ist ist die Membranauslenkung ausschließlich von der Federsteife s abhängig und verläuft näherungsweise konstant
- Es gibt eine Resonanzüberhöhung an der Stelle $2\pi f_0 = \sqrt{s/m}$.
- Die Stärke der Resonanzüberhöhung hängt von $\eta = r2\pi f_0/s$, also insbesondere von der Dämpfung r ab.
- Für Frequenzen oberhalb von f_0 spielt praktisch nur noch der quadratische Anteil des Nenners eine Rolle. Bei diesen Frequenzen nimmt der Frequenzgang also mit $12dB/Oct$ ab.

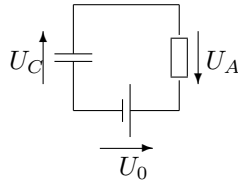
Insgesamt hat der mechanische Aufbau eines Mikrophons also eine Tiefpascharakteristik. Die verschiedenen Parameter werden zumeist wie folgt abgestimmt: die Masse der Membran m möglichst klein, um die Resonanzfrequenz f_0 steigen zu lassen und damit den linearen Übertragungsbereich zu vergrößern, die Federsteife s aus dem gleichen Grund möglichst groß, andererseits aber auch nicht zu groß, um den Übertragungsfaktor nicht zu sehr zu verkleinern, und die Dämpfung r eher groß, um die Resonanzüberhöhung breit und niedrig zu gestalten.

Elektrischer Aufbau

Die Membran eines Kondensatormikrophones besteht aus Metall oder metallisiertem Kunststoff und stellt eine Elektrode eines Kondensators dar. Sie ist in geringem Abstand vor der elektrisch leitenden Gegenelektrode angeordnet. Jede Schwingung der Membran verändert den Abstand zwischen beiden Elektroden und damit die Kapazität dieses Kondensators, was eine Änderung der Kondensatorspannung U_C und eine Änderung der Ausgangsspannung U_A zur Folge hat. Das prinzipielle Schaltbild des elektrischen Teils des Kondensatormikrophons ist in Abb. 7.5 dargestellt.

Kondensatormikrophone benötigen eine Spannungsquelle U_0 zu Vorspannung. Diese wird oft *Phantomspannung* genannt.

Der elektrische Aufbau des Kondensatormikrophons hat eine hochpaßähnliche Übertragungsfunktion.

**Abbildung 7.5:** elektrisches Prinzip des elektrostatischen Wandlers**Herleitung**

Die Kapazität C des Kondensators ist abhängig vom Abstand der beiden Elektroden im Ruhezustand d und der Membranauslenkung x , sowie von der Ruhekapazität C_0 . Es gilt:

$$C = C_0 \frac{1}{1 - \frac{x}{d}} \quad (7.6)$$

Die Ladung des Kondensators Q verändert sich ebenfalls mit Bewegung der Membran. Die Ladung entspricht der Summe von Ruheladung Q_0 und Wechselladung q . Es ergibt sich

$$Q = Q_0 + q \quad (7.7)$$

Die Spannung über dem Kondensator U_C ergibt sich aus dem Quotienten von Ladung und Kapazität. Mit den Gln. (7.6) und (7.7) ergibt sich näherungsweise:

$$\begin{aligned} U_C &= \frac{Q}{C} \\ &= \frac{Q_0 + q}{C_0} \left(1 - \frac{x}{d}\right) \\ &\approx U_0 \left(1 - \frac{x}{d}\right) + \frac{q}{C_0} \end{aligned} \quad (7.8)$$

Die Approximation ist gültig, wenn die Wechselladung $q \ll Q_0$ und die Membranauslenkung $x \ll d$.

7.1.2.2 Dynamisches Mikrophon**Mechanischer Aufbau**

Der mechanische Aufbau eines dynamischen Mikrophones entspricht dem des Kondensatormikrophons. Der entscheidende Unterschied ist jedoch, daß die gesuchte Ausgangsgröße nicht wie bei dem Kondensatormikrofon die Membranauslenkung, sondern die Membrangeschwindigkeit v ist. Daher ergibt sich für den mechanischen Aufbau des dynamischen Mikrophons ein anderer prinzipieller Frequenzgang, der Abb. 7.6 dargestellt ist.

Herleitung

Da sich die Geschwindigkeit eines Körpers aus der zeitlichen Ableitung des Ortes (in diesem Fall also der Membranauslenkung) ergibt, muß die erhaltene Gleichung (7.5) zeitlich abgeleitet werden. Wie schon bei Glg. (7.3) resultiert eine Ableitung in einer Multiplikation mit $j2\pi f$. Der resultierende Frequenzgang ist also

$$v/F = \frac{j2\pi f/s}{1 - \frac{f^2}{f_0^2} + j\eta \frac{f}{f_0}} \quad (7.9)$$

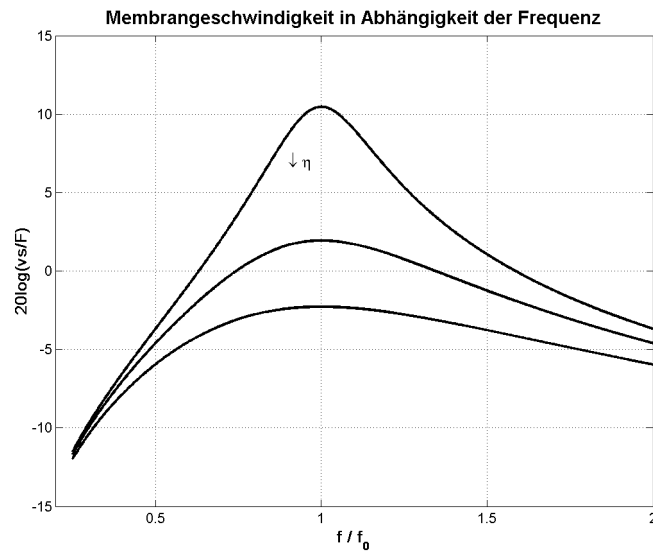


Abbildung 7.6: Übertragungsfrequenzgang eines Geschwindigkeitsempfängers in Abhängigkeit verschiedener Dämpfungswerte

Der für Frequenzen kleiner der Resonanzfrequenz steigt die Membrangeschwindigkeit mit $6dB/Oct$ an, für größere fällt sie ebenfalls mit $6dB/Oct$. Da der in Abb. 7.6 dargestellte Übertragungsfrequenzgang keinen linearen Teil mehr aufweisen kann, liegt es nicht mehr im Interesse des Entwicklers, die Resonanz hoch abzustimmen. Vielmehr wird die Resonanz nun in die Mitte des Frequenzbereichs gelegt und möglichst stark gedämpft, um den Frequenzgang zu linearisieren.

Elektrischer Aufbau

7.1.2.3 Andere Mikrofontypen

7.1.3 Bauform

7.1.3.1 Druckempfänger

7.1.3.2 Druckgradientenempfänger

7.1.3.3 Mischformen

Herleitung

Die Funktion des Richtungsfaktors bei für gemischte Charakteristiken lautet $\Gamma(\theta) = A + B \cos(\theta)$ mit A: Anteil der Kugelcharakteristik und B: Anteil der Achter-Charakteristik, wobei $A + B = 1$. Eingesetzt in Gleichung (7.1) ergibt sich somit:

$$\gamma = \frac{2}{\int_0^\pi (A + B \cos(\theta))^2 \sin(\theta) d\theta} \quad (7.10)$$

Für das Integral im Nenner ergibt sich

$$\begin{aligned}
 \int_0^\pi A^2 \sin(\theta) d\theta + \int_0^\pi B^2 \underbrace{\cos^2(\theta)}_{=\frac{1}{2}(1+\cos(2\theta))} \sin(\theta) d\theta + \int_0^\pi 2AB \cos(\theta) \sin(\theta) d\theta &= \\
 [-A^2 \cos(\theta)]_0^\pi + \frac{B^2}{2} \left(\int_0^\pi \sin(\theta) d\theta + \int_0^\pi \underbrace{\cos(2\theta) \sin(\theta)}_{\frac{1}{2}(\sin(-\theta)+\sin(3\theta))} d\theta \right) + [AB \sin^2(\theta)]_0^\pi &= \\
 [-A^2 \cos(\theta)]_0^\pi + \frac{B^2}{2} \left[-\cos(\theta) - \frac{1}{2} \cos(-\theta) - \frac{1}{6} \cos(3\theta) \right]_0^\pi + [AB \sin^2(\theta)]_0^\pi &= \\
 \left[-\left(A^2 + \frac{3B^2}{4} \right) \cos(\theta) - \frac{B^2}{12} \cos(3\theta) \right]_0^\pi &= \\
 2A^2 + \frac{2B^2}{3} &
 \end{aligned}$$

Daraus folgt für den Bündelungsgrad:

$$\gamma = \frac{1}{A^2 + \frac{B^2}{3}} \quad (7.11)$$

7.1.3.4 Andere Bauformen

7.2 Lautsprecher

akustischer Kurzschluß, Lautsprecherzeile

7.3 Zusammenfassung

7.4 Aufgaben

1. warum 12dB/oct bei Auslenkungsübertragungsfrequenzgang?

Kapitel 8

Digitaltechnik

Die Digitaltechnik und die digitale Signalverarbeitung stellen in den meisten Studios die wichtigsten Komponenten des Tonstudioalltags. Dieses Kapitel gibt eine Einführung in die Digitaltechnik und eine Übersicht über die verschiedensten mit Tontechnik oder Audiosignalverarbeitung verknüpften Applikationen.

8.1 Binäre Zahlendarstellung

Digitale Systeme arbeiten nicht mit der uns geläufigen dezimalen Zahlendarstellung, sondern mit der sogenannten **binären Zahlendarstellung**. Hervorstechendstes Merkmal dieses Systems ist, daß es ausschließlich auf zwei Ziffern basiert: 0 und 1.

Die sogenannte **Wortbreite** W eines digitalen Systems besagt, wie viele solcher einzelnen Ziffern, die im folgenden *Bits* genannt werden, zu einem Wort und damit zu einer Zahl zusammengefügt werden. Bei einer Wortbreite von 16bit , wie sie auf einer CD zu finden ist, bilden somit immer 16bit eine Zahl bzw. einen Amplitudenwert. Jedes einzelne Bit kann zwei verschiedene Zustände annehmen, d.h. insgesamt können $2 \cdot 2 \cdot 2 \cdot \dots \cdot 2 \cdot 2 \cdot 2 = 2^{16} = 65536$ verschiedene Zahlen dargestellt werden.

Die dargestellte Zahl berechnet sich mit Hilfe von Zweierpotenzen: beispielsweise steht die binäre Zahl 10011011 mit der Wortbreite 8bit für die dezimale Zahl $1 \cdot 2^7 + 0 \cdot 2^6 + 0 \cdot 2^5 + 1 \cdot 2^4 + 1 \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0 = 155$. Eine Änderung an der Stelle mit der höchsten Potenz (links) führt die größten Veränderungen herbei, daher wird dieses Bit das *Most Significant Bit (MSB)* genannt. Analog dazu führen Änderungen im Bit rechts die kleinsten Änderungen herbei; es wird *Least Significant Bit (LSB)* genannt.

Wie man sieht, lassen sich hier positive Zahlen von 0 bis 255 darstellen; im Falle von realen Audiosignalen bewegt sich die Amplitude des Signals um den Nullpunkt herum, d.h. es treten positive wie negative Amplitudenwerte auf. Dem wird dadurch Rechnung getragen, daß das MSB als eine Art Vorzeichen-Bit

interpretiert wird; ist es 0, so ist die Zahl positiv, im anderen Falle negativ. Das führt bei 16bit Wortbreite zu dem Zahlenbereich -32768 bis 32767 . Die Aufteilung des Zahlenbereichs geschieht so, daß die positiven Zahlen genauso behandelt werden wie im vorzeichenlosen Fall, die negativen aber von der negativen Maximalamplitude hochgezählt werden.

Beispiel

Es werden einige Beispiele in vorzeichenbehafteter Binärdarstellung mit 4 Bit Wortbreite gegeben:

Binär	Dezimal
0001	1
0101	5
0111	7
1000	-8
1001	-7
1111	-1

Tabelle 8.1: Beispiel für vorzeichenbehaftete Binärwerte der Wortbreite 4 Bit

Bemerkung

Der Binärwert eines vorzeichenbehafteten Signals nicht immer so einfach zu berechnen, wie man auf den ersten Blick meinen würde. Falls das Vorzeichenbit nicht gesetzt ist, kann die Zahl wie oben gezeigt berechnet werden; im Fall einer negativen Zahl müssen folgende Schritte durchgeführt werden:

1. die Bitfolge in ihre *Komplementärdarstellung* zu *invertieren*, d.h. aus jeder 0 eine 1 zu machen und umgekehrt (Beispiel: 1001 0111 1100 0001 \rightarrow 0110 1000 0011 1110,
2. die resultierende positive Zahl wie oben erläutert zu bestimmen (Beispiel: 0110 1000 0011 1110 entspricht der dezimalen Zahl 26686),
3. die Zahl mit einem negativen Vorzeichen zu versehen (Beispiel: 26686 \rightarrow -26686),
4. 1 von der Zahl zu subtrahieren (Beispiel: Endergebnis ist -26687)

Bemerkung

Das erläuterte Zahlenformat kann nur ganzzahlige Werte darstellen, es wird **Festkommaformat** (auch Fixed-Point-Format) genannt. Gerade im Audiobereich wird aber zum Teil auch mit einem anderen Format gearbeitet, dem sogenannten **Fließkommaformat** (auch Floating-Point-Format oder Gleitkommaformat). Hier werden nur Nachkommastellen und ein Exponent abgespeichert. Der Vorteil dieses Formats für Audibearbeitung ist, daß sehr kleine Zahlen sehr genau dargestellt werden können, sehr große Zahlen hingegen ungenauer. Dies entspricht dem menschlichen Hörvermögen recht gut.

Das Floating-Point-Format wird im allgemeinen bei der Audibearbeitung auf dem Computer und in einigen Hardwaresystemen verwendet. Zur Übertragung und Speicherung findet es normalerweise keinen Einsatz.

8.2 Digitalisierung eines Analoogsignals

8.2.1 Abtastung

Wenn wir den Verlauf einer Spannung oder des Schalldrucks betrachten, sehen wir *kontinuierliche* Veränderungen des Signals über der Zeit. Solche kontinuierlichen Veränderungen kennzeichnen *analoge* Signale. Wir können uns einen kurzen Zeitabschnitt betrachten und immer weiter hineinzoomen, ohne daß der Verlauf des Signals jemals unterbrochen wird. Ein solcher unterbrechungsfreier Zeitverlauf ist in der Digitaltechnik nicht darstellbar, da hierzu unendlich viele Punkte dargestellt werden müßten. Stattdessen wird das analoge Signal in bestimmten Zeitabständen abgetastet und es werden nur die einzelnen Amplitudenwerte zum Abtastzeitpunkt gespeichert. Die Frequenz dieser Abtastung wird **Abtastrate** (englisch: **Sample Rate**) genannt, ihre Einheit ist natürlich $[Hz]$. Abbildung 8.1 zeigt einen kleinen Ausschnitt eines kontinuierlichen (analogen) Signals und die resultierende abgetastete Folge.

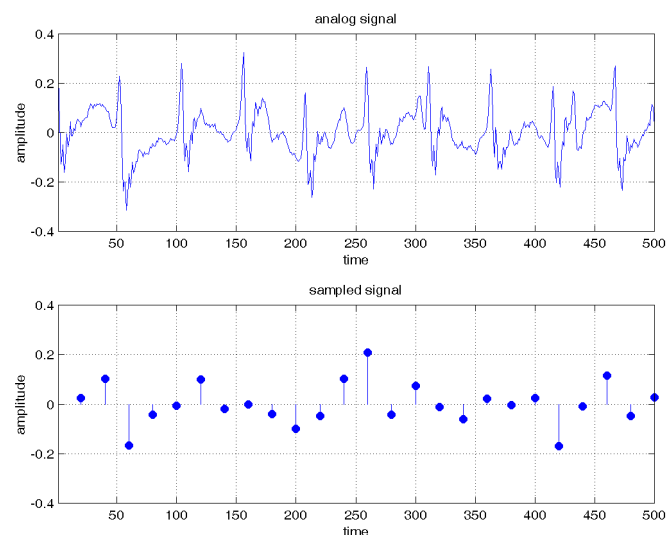


Abbildung 8.1: Vergleichende Darstellung eines kontinuierlichen und eines abgetasteten Signalverlaufs, oben: kontinuierlicher Zeitverlauf, unten: entsprechender abgetasteter Zeitverlauf

8.2.1.1 periodisches Spektrum

Eine wichtige Eigenschaft der Abtastung läßt sich an einem Beispiel veranschaulichen; dazu betrachten wir mehrere Sinusschwingungen, die mit einer Frequenz von $6kHz$ abgetastet werden. Die Originalsignale und die abgetasteten Signale für

die Frequenzen 1kHz , 7kHz und 13kHz sind in Abb. 8.2 dargestellt, während die Abb. 8.3 die Originalsignale und abgetasteten Signale für die Frequenzen 1kHz , 5kHz , 11kHz und 17kHz zeigt.

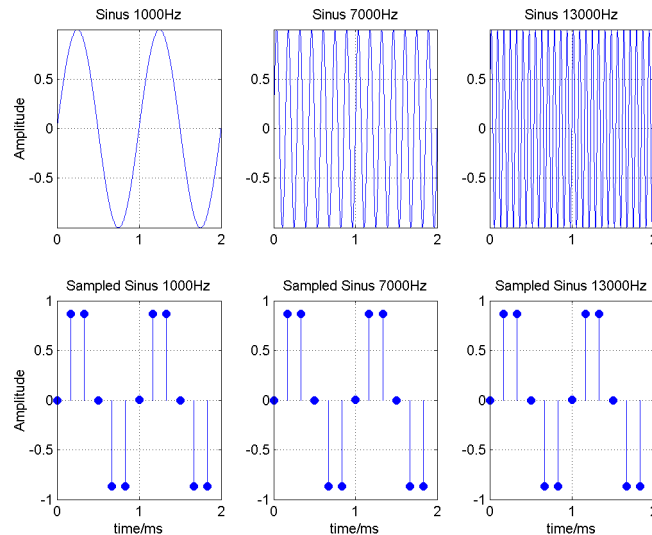


Abbildung 8.2: Darstellung von analogem und abgetastetem Zeitverlauf von Sinusschwingungen der Frequenzen 1kHz , 7kHz und 13kHz , die Abtastfrequenz ist 6kHz ; oben: kontinuierlicher Zeitverlauf, unten entsprechender abgetasteter Zeitverlauf

Abb. 8.2 zeigt deutlich, daß die abgetasteten Signale bei diesen bestimmten Frequenzen identisch sind. Die abgetasteten Signale aus Abb. 8.3 sind zwar nicht identisch, haben aber die gleiche Frequenz und Amplitude, lediglich eine andere Phasenlage. Mit Hilfe dieser Beispiele ist man einer der wichtigsten Eigenschaften abgetasteter Signale auf der Spur, die lautet: **die Spektren abgetasteter Signale sind periodisch mit der Abtastfrequenz.**

Herleitung

Ohne zu sehr ins mathematische Detail gehen zu wollen, läßt sich der Weg der Herleitung wie folgt beschreiben: die Abtastung eines Signals läßt sich darstellen als die Multiplikation des analogen Signals mit einem Signal, das genau an den Abtastwerten den Wert 1 hat und an allen anderen Stellen den Wert 0. Der Verlauf dieses Abtastsignals erinnert an einen Kamm, die Abstände der einzelnen Zähne („1“) des Kamms entsprechen dem Abtastintervall. Es läßt sich nun zeigen, daß die Frequenztransformation eines solchen Signals ebenfalls wie solcher Kamm aussieht, wobei der Abstand der Zähne dieses Kamms nun der Abtastfrequenz entspricht. In Abschnitt 6.2.1.1 wurde darauf hingewiesen, daß eine Multiplikation zweier Signale im Zeitbereich einer Faltung der Frequenzdarstellungen dieser Signale entspricht. Es wird nun also das Spektrum des analogen Signals gefaltet mit dem kammförmigen Signal. Diese Faltung führt dazu, daß das Spektrum des analogen Signals periodisch mit der Abtastfrequenz wieder auftaucht.

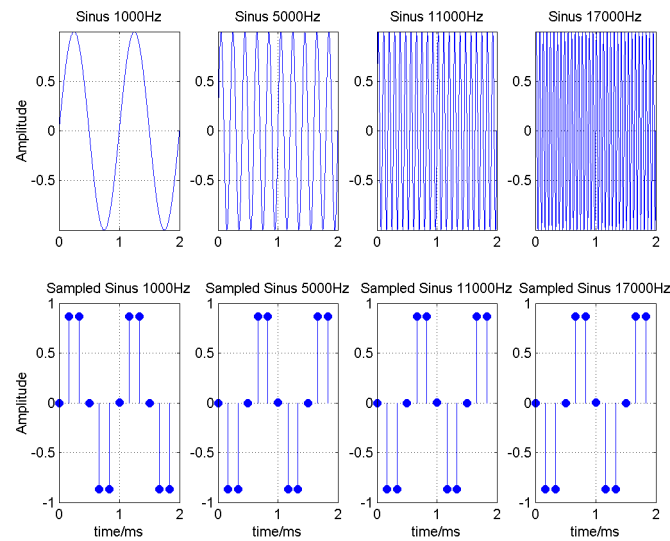


Abbildung 8.3: Darstellung von analogem und abgetastetem Zeitverlauf von Sinusschwingungen der Frequenzen 1kHz , 5kHz , 11kHz und 17kHz , die Abtastfrequenz ist 6kHz ; oben: kontinuierlicher Zeitverlauf, unten: entsprechender abgetasteter Zeitverlauf

8.2.1.2 Abtasttheorem und Aliasing

Aus der Periodizität des Spektrums des abgetasteten Signals folgt eine wichtige Eigenschaft; die (hohen) Frequenzen der periodischen Spektren können sich überlappen. Tun sie dies, so läßt sich nicht mehr unterscheiden, ob es sich um Frequenzen des Originalspektrums handelt oder um hineingespiegelte Frequenzen. Es zeigt sich, daß sich die periodischen Spektren nicht überlappen, wenn die Bandbreite des Signals kleiner als die halbe Abtastfrequenz ist. Dies ist das sogenannte **Abtasttheorem**. Wird es nicht eingehalten (die sog. *Unterabtastung*, so entstehen „Spiegelfrequenzen“, die sehr unangenehm klingen können. Dieser Effekt wird **Aliasing** genannt. Aus diesem Grund befindet sich vor jedem AD-Wandler ein Tiefpaßfilter, der (möglichst) alle Frequenzanteile oberhalb der halben Abtastfrequenz herausfiltert und somit die Einhaltung des Abtasttheorems gewährleistet. Dieses Filter wird *Antialiasingfilter* genannt. Die Güte dieses Filters kann großen Einfluß auf die Wandlertüte haben (siehe Abschnitt 8.2.1.3).

Bemerkung

Ein anschauliches Beispiel einer Unterabtastung im visuellen findet man in vielen Westernfilmen. Die Speichenräder einer Kutsche drehen sich mit der erwarteten Geschwindigkeit und Richtung, solange die Kutsche langsam fährt. Übersteigt die Speichengeschwindigkeit allerdings die halbe Abtastfrequenz der Kamera, so nimmt die wahrgenommene Geschwindigkeit des Rades wieder ab.

8.2.1.3 Rekonstruktion des Signals

Zur Rekonstruktion des analogen Signals aus dem abgetasteten Signal muß das abgetastete Signal tiefpaßgefiltert werden, um die periodischen Fortsetzungen des Spektrums „herauszufiltern“. Das Tiefpaßfilter hat die Aufgabe, nur die Signalfrequenzen kleiner als die halbe Abtastfrequenz passieren zu lassen. Es wird *Rekonstruktionsfilter* genannt. Theoretisch ist ein analoges Signal, welches unter Berücksichtigung des Abtasttheorems gesampelt wurde, ohne Verlust wieder rekonstruierbar. Leider gilt dies aber nur für den Fall, daß sowohl Antialiasingfilter als auch Rekonstruktionsfilter ideal sind, d.h. alle Frequenzen bis zur halben Abtastfrequenz ungehindert passieren lassen, aber keine Frequenzanteile oberhalb der halben Abtastfrequenz. Ein solches Filter ist praktisch nicht zu realisieren, sondern wird bei vorhandenen Wandlern nur mehr oder weniger gut approximiert.

Bemerkung

Als Entdecker des Abtasttheorem werden oft Nyquist oder Shannon genannt. Nyquist war wohl der erste, der das Abtasttheorem ca. 1930 veröffentlichte, während Shannon nach dem 2. Weltkrieg eine Generalisierung des Abtasttheorems publizierte. Daher werden als gleichbedeutend mit Begriff Abtasttheorem manchmal auch die Begriffe Nyquisttheorem und Shannonththeorem verwendet.

8.2.2 Quantisierung

Ebenso wie ein digitales Signal keinen kontinuierlichen Zeitverlauf haben kann (s. Abschnitt 8.2.1), kann es auch keinen kontinuierlichen Amplitudenverlauf besitzen, da nur einzelne Werte abgespeichert werden können. Dies bedeutet, daß die Amplitude - ebenso wie die Zeit bei der Abtastung - in bestimmte Abschnitte oder besser Stufen eingeteilt werden muß. Dieser Vorgang wird **Quantisierung** genannt. Eine kontinuierlich steigende Amplitude wird durch den Quantisierungsvorgang auf eine Treppenfunktion abgebildet. Diese Treppenfunktion wird auch als Quantisierungskennlinie bezeichnet und ist in Abb. 8.4 rechts dargestellt. Wie an dieser Abbildung veranschaulicht, ist das Prinzip der Quantisierung, bestimmte Amplitudenbereiche auf einzelne *quantisierte* Werte abzubilden. Die Quantisierungskennlinie ist damit eine treppenförmige Funktion. Je mehr einzelne Stufen erlaubt werden, desto kleiner werden die Stufen und desto ähnlicher wird das quantisierte Signal dem originalen kontinuierlichen Signal. Die Anzahl der Stufen bestimmt die Zahl der zur Speicherung eines Wertes benötigten Bits. 16 Stufen lassen sich mit 4 Bit darstellen; im Fall der Abb. 8.4 handelt es sich also um einen 4-Bit-Quantisierer. Wortbreite eines digitalen Signals und Anzahl der Quantisierungsstufen sind also direkt ineinander überzuführen: um ein digitales Signal der Wortbreite 16 Bit zu erhalten, muß die Anzahl der Stufen 65536 betragen. Soll die Wortbreite voll ausgenutzt werden, so müssen Minimal- und Maximalamplitude (oder -spannung) des Signals und Eingangsminimal- und Maximalspannung des Quantisierers übereinstimmen. Je geringer die Aussteuerung des Eingangssignals ist, desto größer wird der Quantisierungsfehler verglichen zum Eingangssignal. Andererseits muß

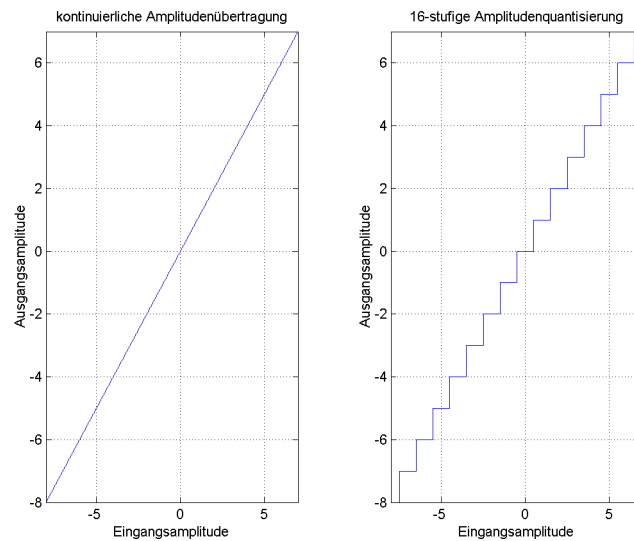


Abbildung 8.4: links: Amplitudenkennlinie bei amplitudenkontinuierlicher Übertragung, jeder Eingangsamplitudenwert wird auf den exakt gleichen Ausgangsamplitudenwert abgebildet; rechts: Amplitudenkennlinie bei Quantisierung mit 16 Amplitudenstufen, bestimmte Amplitudenbereiche der Eingangsamplitude werden auf einzelne feste Werte abgebildet

eine Übersteuerung des Quantisierers unter allen Umständen vermieden werden, da die Übersteuerungsartefakte, zu sehr unangenehmen Verzerrungen führen können.

8.2.2.1 Quantisierungsfehler

Die Abtastung eines Signals ist theoretisch ohne Verluste rückgängig zu machen, bei der Quantisierung allerdings ist dies auch in der Theorie nicht möglich. Bei jeder Quantisierung wird unvermeidlich ein Fehler gemacht, der sog. **Quantisierungsfehler** q (auch Quantisierungsrauschen). Dieser berechnet sich aus der Differenz zwischen dem Originalsignal x und dem quantisierten Signal x_q für einen beliebigen Zeitwert t .

$$q(t) = x(t) - x_q(t) \quad (8.1)$$

Es wird i.a. davon gesprochen, daß dieser Fehler dem Signal hinzugefügt wird; tatsächlich ergibt ja die Subtraktion von $x(t) - q(t)$ das quantisierte Signal $x_q(t)$. Abb. 8.5 zeigt den Quantisierungsfehler eines mit 4 Bit quantisierten optimal ausgesteuerten Sinussignals. Es ist anschaulich klar, daß der Quantisierungsfehler kleiner wird, wenn die Zahl der Stufen bei gleichbleibender Maximalamplitude erhöht wird, denn der Betrag der Amplitude des Quantisierungsfehlers kann maximal eine halbe Stufenhöhe Δ betragen. Mit diesem Wissen läßt sich der (theoretische) Rauschabstand eines Quantisierers berechnen. Da die Stufenhöhe Δ

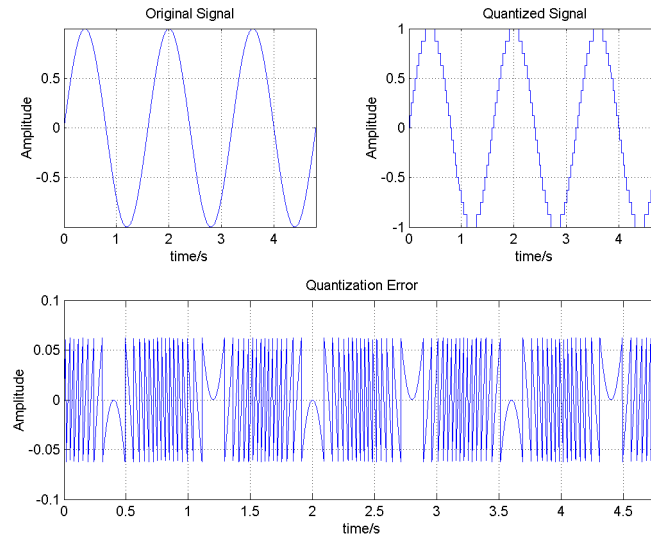


Abbildung 8.5: links oben: das kontinuierliche Originalsignal; rechts oben: das mit einer Auflösung von 4 Bit quantisierte Signal; unten: der dabei gemachte Quantisierungsfehler

durch die Wortbreite W des resultierenden digitalen Signals bestimmt wird, kann man den Rauschabstand SNR direkt auf die Wortbreite des Signals zurückführen und gelangt zu der Beziehung:

$$SNR = 6.02 \cdot W + const. \text{ [dB]} \quad (8.2)$$

Die Konstante $const.$ ist abhängig von Eingangssignal; sie ist für ein Sinussignal 1.76dB und für ein Sprach- bzw. Musiksignal im Bereich -7 bis -8dB . Als Daumenregel klingt die Gleichung (8.2) so: *Wird die Wortbreite eines Quantisierers um 1 Bit erhöht, so gewinnt man 6dB Rauschspannungsabstand.*

Herleitung

Die Maximalamplitude des Quantisierungsfehlers ist $\frac{\Delta}{2}$. Geht man davon aus, daß alle Amplitudenwerte zwischen den Maximalausschlägen des Quantisierers gleichhäufig vorkommen (das ist als Näherung erlaubt, wenn die Stufenanzahl groß ist), so kann die Leistung des Quantisierungsfehlers berechnet werden mit

$$\sigma_q^2 = \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} \frac{1}{\Delta} x^2 dx = \frac{\Delta^2}{12} \quad (8.3)$$

Der SNR berechnet sich aus dem Quotienten von Signalleistung zu Fehlerleistung. Als Beispiel nehmen wir nun eine optimal ausgesteuerte Sinusschwingung. Die Amplitude entspricht daher der Hälfte der Zahl der Stufen 2^W multipliziert mit der Stufenhöhe Δ . Daraus folgt:

$$\sigma_{\text{Sinus}}^2 = \int_0^{2\pi} (\text{Amplitude} \cdot \sin(x))^2 dx \frac{(\text{Amplitude})^2}{2} = \frac{(2^W \cdot \Delta)^2}{8} \quad (8.4)$$

Somit ergibt sich für den Signalrauschabstand des Quantisierers:

$$\begin{aligned}
 SNR &= 10 \log \left(\frac{\sigma_{\text{Sinus}}^2}{\sigma_q^2} \right) \\
 &= 10 \log \left(\frac{(2^W * \Delta)^2 \cdot 12}{8 \cdot \Delta^2} \right) \\
 &= 10 \log(2) \cdot 2 \cdot W + 10 \log \left(\frac{3}{2} \right) \\
 &= 6.02 \cdot W + 1.76 \text{ [dB]}
 \end{aligned}$$

8.2.2.2 Oversampling

Um die Qualität einer Digitalisierung zu verbessern, wird oftmals mit sog. **Oversampling** gearbeitet. Oversampling bedeutet, daß das Audiosignal mit einer höheren Frequenz abgetastet wird als benötigt wird bzw. als am Ausgang des Wandlers auftritt. Es gibt v.a. zwei Gründe für diese Überabtastung. Der erste Grund ist die effiziente technische Realisierung: um maximale Audiobandbreite bis nah an die halbe Abtastfrequenz ohne aufwendiges Antialiasingfilter realisieren zu können, wird die Abtastrate so hochgesetzt, daß ein einfaches Antialiasingfilter genügt um das Abtasttheorem zu erfüllen. Anschließend wird das Signal im digitalen Bereich tiefpaßgefiltert, so daß es die Anforderungen des Abtasttheorems für die eigentlich gewollte Abtastfrequenz erfüllt.

Dieses Vorgehen hat einen erwünschten Nebeneffekt, welcher der zweite Grund für die temporäre Erhöhung der Abtastfrequenz ist: der Signal-Rauschabstand (SNR) kann verbessert werden. Das ist zunächst überraschend, da die Abtastrate zunächst nur die Bandbreite des digitalisierten Signals beeinflusst, nicht den SNR. Zwei wichtige Eigenschaften bei der Digitalisierung führen jedoch zu einer Erklärung:

- Die Gesamtleistung des Quantisierungsrauschens ist unabhängig von der Abtastfrequenz.
- Das Quantisierungsrauschen ist näherungsweise weißes Rauschen, das über die gesamte Bandbreite des Signals gleichmäßig verteilt wird.

Wenn also die Gesamtleistung des Quantisierungsfehlers gleich bleibt, obwohl die Abtastfrequenz erhöht wird, dann wird bei Erhöhung der Abtastfrequenz die durchschnittliche Leistung in einem bestimmten Frequenzbereich sinken, da es sich über einen größeren Frequenzbereich erstrecken kann. Wendet man anschließend den oben angesprochenen digitalen Antialiasingfilter an, so wird der Anteil des Quantisierungsrauschens über der endgültigen halben Abtastfrequenz „weggefiltert“, und der SNR steigt. Man gewinnt mit solchen Oversamplingverfahren pro Frequenzverdopplung ca. 3dB Signal-Rauschabstand.

8.2.2.3 Dither

Das Verhältnis bzw. der Signal-Rauschabstand des quantisierten Signals wird schlechter je geringer das Signal angesteuert wird. Doch eine schlechte Aussteuerung führt nicht nur zu höherem Rauschpegel, sondern kann einen weiteren Effekt haben - das Quantisierungsrauschen ist kein weißes Rauschen mehr wie bei hohen Amplituden, sondern wird abhängig vom Eingangssignal. Abbildung 8.6 zeigt ein mit drei Stufen quantisiertes Signal und dessen Quantisierungsfehler. Der hier eingeführte Quantisierungsfehler hört sich an wie sehr störende Verzerrungen und ist besonders deutlich bei niedriger Aussteuerung und tiefen Signalfrequenzen. Diese sind zwar bei hochauflösenden Wandlern mit großem Dynamikumfang kaum mehr hörbar, sollen aber bei niedriger auflösenden Wandlern oder bei Wortbreitenkonvertierungen wie z.B. 24bit nach 16bit oder niedriger möglichst vermieden werden. Eine naheliegende Idee ist nun, das Signal etwas „analoger“ zu machen,

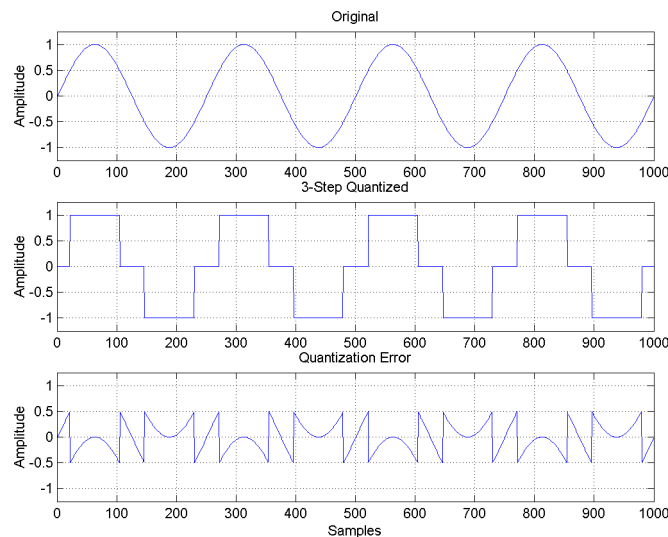


Abbildung 8.6: oben: Originalsignal, mitte: 3-stufig quantisiertes Signal, unten: Quantisierungsfehler

indem vor dem Quantisierungsprozeß einfach ein Rauschen addiert wird. Dieses Rauschen wird **Dither** genannt.

Zunächst naheliegend scheint die Annahme, dieses Rauschen müßte so stark sein, daß es die o.g. Verzerrungen akustisch verdeckt; das muß aber *nicht* der Fall sein. Vielmehr genügt ein schwaches Rauschen, das zu einer nicht mehr deterministischen, gleichbleibenden Abfolge der ausgewählten Quantisierungsstufen führt, sondern zu einer eher zufälligen. Wenn beispielsweise am Eingang eine Gleichspannung von 1.3mV anläge, der Quantisierer das Signal allerdings in 1mV -Schritten quantisiert, dann wird das quantisierte Signal bei einem ungeditherten Eingang

konstant bei 1mV liegen. Wird das Eingangssignal hingegen gedithert, so wird es manchmal bei 2mV , häufiger bei 1mV und sehr selten bei anderen Quantisierungswerten liegen. Tatsächlich wird dann aber beim quantisierten ein Mittelwert von 1.3mV feststellbar sein, d.h. im Mittel ist die geditherte Quantisierung genauer, da beliebige Quantisierungswerte möglich gemacht werden. Abbildung 8.7 zeigt das obige Beispiel mit einem hinzugefügten Dithersignal. Abbildung 8.8 zeigt die Spektren des quantisierten Signals und des gedithert quantisierten Signals.

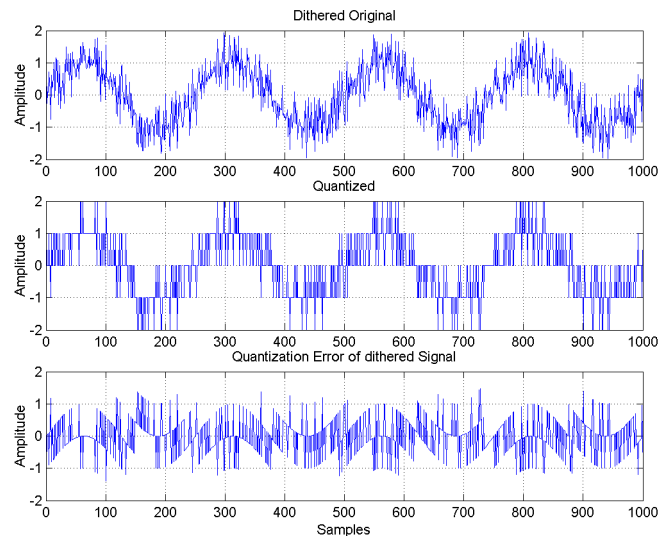


Abbildung 8.7: oben: gedithertes Originalsignal, mitte: grob-stufig quantisiertes Signal, unten: Quantisierungsfehler bei gedithertem Eingang

Beispiel

Die Wirkung des Dithering läßt sich leicht anhand eines Beispiels veranschaulichen. Hält man sich eine Hand mit leicht geöffneten Fingern vor die Augen, so wird ein Großteil des Gesichtsfeldes von den Fingern abgedeckt, und nur durch die Zwischenräume läßt sich etwas erkennen. Bewegt man diese Hand allerdings sehr schnell, so lassen sich - wenn auch etwas undeutlich - auch die Bereiche erkennen, die zuvor von den Fingern verdeckt waren.

Die Erfahrung hat gezeigt, daß Ditheramplituden im Bereich der halben bis zur 1.5-fachen Stufenbreite zu besten Ergebnissen führen.

Das Ditherspektrum ist zumeist entweder weiß oder leicht hochpaßgefiltert. Eine Hochpaßfilterung führt in den meisten Fällen zu einer subjektiven Qualitätsverbesserung, da die Rauschleistung etwas aus dem Hörbereich herausgeschoben wird. Ein wichtiger Aspekt des Ditherrauschens ist seine Amplitudendichteverteilung (ADV), d.h. die Wahrscheinlichkeit unterschiedlicher Signalwerte. Das „reale“ zur Amplitudendichteverteilung ist die Häufigkeitsverteilung, die eine Näherung

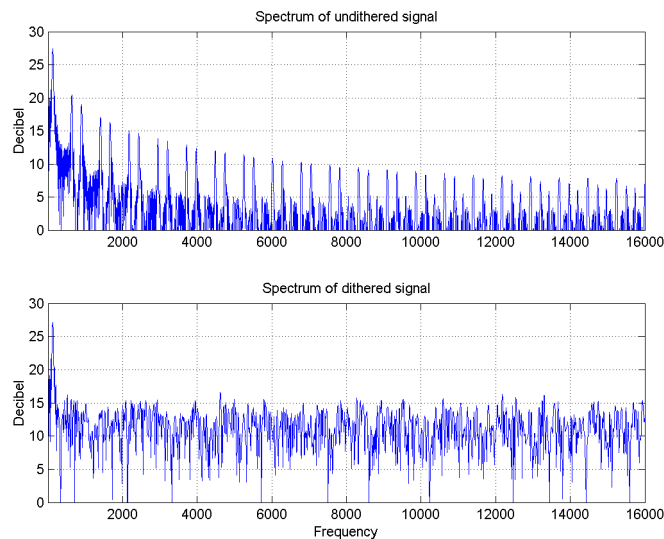


Abbildung 8.8: Spektrum eines mit und ohne Dither quantisierten Signals; oben: ungedithert quantisiertes Signal, unten: gleichförmig gedithert quantisiertes Signal

der ADV darstellt (vgl. Abschnitt 6.1.3). Wenn alle Signalwerte gleichhäufig vorkommen, spricht man von gleichverteiltem Rauschen und die Amplitudendichteverteilung ist rechteckförmig. Man ist aber i.a. der Meinung, daß ein gleichverteiltes Rauschen nicht zu optimalen Dithering-Ergebnissen führt, während ein dreieckförmig verteiltes Rauschen zu einem besseren Höreindruck führt. Abbildung 8.9 zeigt die Zeitverläufe und die Amplitudendichteverteilungen von gleich- und dreieckförmig verteiltem Rauschen. Teilweise kann man auch anders geformte Dither verwenden; z.B. gibt es häufig gaußförmige Rauschen. Eine gaußförmige Verteilung sieht glockenförmig aus.

Die Verwendung unterschiedlicher Ditherformen führt auch zu unterschiedlichem Rauschpegel. Der eingefügte Rauschpegel bewegt sich bei den o.g. Ditheramplituden zwischen 3 und 6dB.

Verwandte Themenbereiche

- Abschnitt 6.1.3

8.2.2.4 Noise-Shaping

Noise-Shaping ist eine Methode, die Qualität eines Wandlers oder einer Wortbreitenkonvertierung zu erhöhen. Der Quantisierungsfehler, der bei normaler Quantisierung näherungsweise ein weißes Spektrum hat, wird spektral geformt. Idealerweise wird dabei die Rauschleistung von Frequenzbereichen hoher Gehörempfindlichkeit (wie z.B. 2 – 4kHz, vgl. Abschnitt 5.2) in Bereiche geringerer

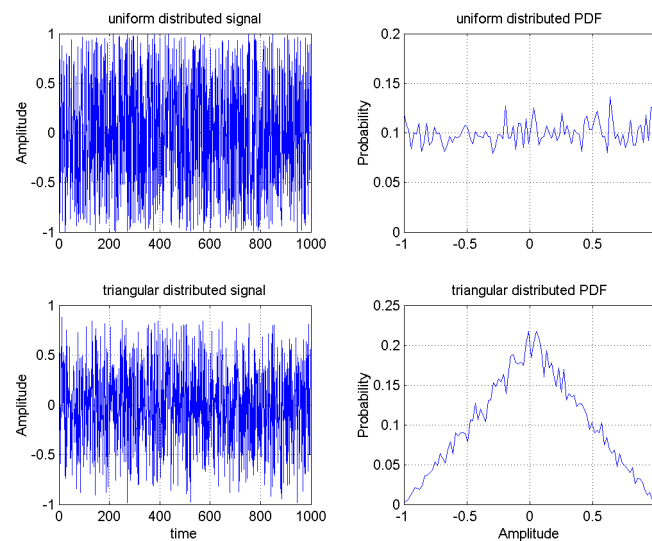


Abbildung 8.9: Zeitverläufe (links) und Amplitudendichteverteilungen (rechts) von Rauschen; oben: gleichverteilt, unten: dreieckförmig verteilt

Empfindlichkeit verschoben (zumeist hohe Frequenzbereiche). Diese Frequenzverschiebung wird durch eine Rückkopplung des Quantisierungsfehlers erreicht, je nachdem wieviele Koeffizienten diese Rückkopplung hat, spricht man von Noise-Shaping verschiedener Ordnungen. Im Fall von Noise-Shaping erster Ordnung wird der Quantisierungsfehler festgestellt und vom darauffolgenden Sample subtrahiert. Dadurch entsteht eine Verschiebung des Quantisierungsfehlers hin zu höheren Frequenzen. Komplexere Verschiebungen, die verschiedene Frequenzbereiche unterschiedlich gewichten, lassen sich nur mit Noise-Shaping höherer Ordnung bewerkstelligen. Es ist natürlich ein naheliegender Gedanke, den Spektralverlauf des Quantisierungsrauschens so zu gewichten, daß er optimal der Ruhehörschwelle angepaßt ist, ein es existieren in der Tat einige solche System auf dem Markt. Welche Systeme nun besser sind und welche schlechter, sollte aber mit kritischem Material (z.B. Ausklingvorgänge) selbst getestet werden.

Noise-Shaping wird meistens in Zusammenhang mit Dither verwendet, um unerwünschte Effekte bei der Rückkopplung des Quantisierungsfehlers zu vermeiden.

Wortbreitenkonvertierungen und damit auch Noise-Shaping sollten möglichst nur einmal am Ende der Verarbeitungskette eingesetzt werden.

Verwandte Themenbereiche

- Abschnitt [6.4.2](#)

8.2.3 Wandler

8.2.3.1 Analog-Digital-Wandler

8.2.3.2 Digital-Analog-Wandler

8.3 Digitale Filter

Digitale Filter lassen sich ausschließlich durch „Verzögerungsglieder“ und gewichtete Addition realisieren. Das bedeutet, daß jeder einzelne Abtastwert um bestimmte Zeiten verzögert wird, mit unterschiedlichen Faktoren multipliziert wird und mit anderen, ebenfalls verzögerten und gewichteten Abtastwerten aufaddiert wird. Warum sich Filter auf diese Art realisieren lassen, verdeutlicht man sich am besten anhand von Beispielen. Zuvor sollen die verwendeten Bestandteile der Filtergraphen kurz erläutert werden. Ein Pfeil zeigt immer die Richtung des Signalfusses an. Ein Kreis symbolisiert eine Rechenoperation: befindet sich ein $+$ im Kreis, so werden zwei Signale aufaddiert, befindet sich ein x im Kreis, so wird das Signal mit der nebenstehenden Konstante (*Filterkoeffizient*) multipliziert. Ein Kästchen mit dem Inhalt z^{-n} bedeutet, daß der eingehende Abtastwert um n Abtasttakte verzögert wird.

Beispiel 1

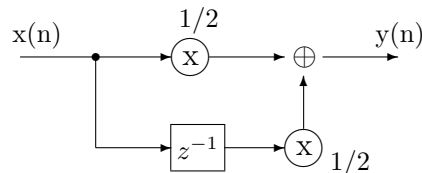


Abbildung 8.10: Filter Beispiel 1

Hier wird jeder einzelne Abtastwert mit 0.5 multipliziert und mit dem vorhergehenden, ebenfalls mit 0.5 multiplizierten Abtastwert addiert. Eine hypothetische Samplefolge $x(n) = [1 \ 110 \ 51 \ 20 \ 30 \ 0]$ führt also zu einem Ausgang $y(n) = [0.5 \ 55.5 \ 80.5 \ 35.5 \ 25 \ 15]$. Wie lassen sich nun die Eigenschaften eines solchen Filters bestimmen? Die Impulsantwort läßt sich einfach auf die gleiche Art bestimmen, wie das schon mit der Samplefolge geschehen ist. Als Eingangssignal dient dazu eine einzelne 1 (ein *Impuls*). Man erhält als Impulsantwort also $[0.5 \ 0.5]$. Der Frequenzgang läßt sich natürlich berechnen, für so ein einfaches Beispiel allerdings auch relativ einfach abschätzen. Dafür stellen wir uns am Eingang des Filters ein Sinussignal variabler Frequenz vor. Bei sehr tiefen Frequenzen wissen wir, daß sich aufeinanderfolgende Abtastwerte stark ähneln. Da am Ausgang des Filters der Mittelwert des aktuellen und des vorhergehenden Abtastwertes erscheint (aus diesem Grund heißt ein solches Filter auch Moving-Average- (MA-)Filter), hat

dieses Filter für tiefe Frequenzen kaum Auswirkungen. Für höhere Eingangsfrequenzen unterscheiden sich aufeinanderfolgende Werte zunehmend, so daß das Filter dämpfend auf das Signal wirkt. Ist das Eingangssignal des Filters im Grenzfall ein Sinussignal mit der halben Abtastfrequenz, so ist der Mittelwert zwischen zwei aufeinanderfolgenden Abtastwerten und damit auch der Ausgang des Filters immer 0. Somit handelt es sich in diesem Beispiel um ein *Tiefpaßfilter*.

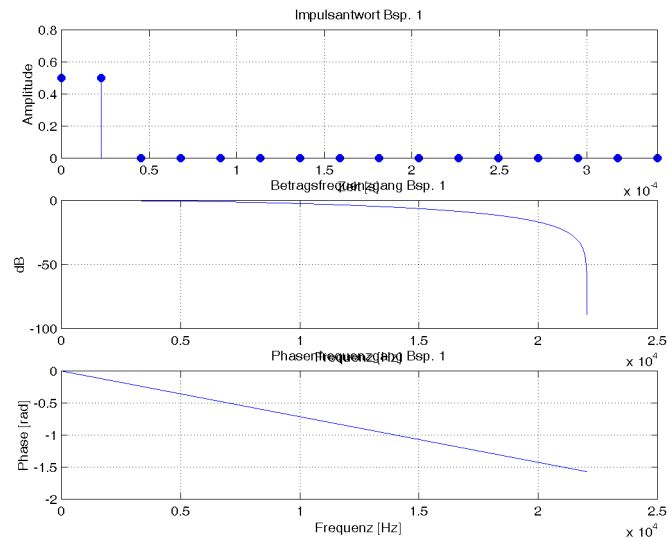


Abbildung 8.11: Impulsantwort, Betragsfrequenzgang und Phasenfrequenzgang von Beispielfilter Nr.1

Abbildung 8.11 zeigt die Impulsantwort, den Betragsfrequenzgang und den Phasenfrequenzgang dieses Filters. Deutlich zu erkennen ist die Tiefpaßcharakteristik, die allerdings erst bei sehr hohen Frequenzen deutlich zum Tragen kommt. Der Phasenfrequenzgang ist linear.

Beispiel 2

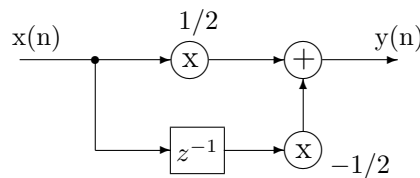


Abbildung 8.12: Filter Beispiel 2

Dieses Filter gleicht dem aus Beispiel 1 bis auf eine kleine Änderung genau. Lediglich der Faktor im Verzögerungsteil ist nun -0.5 statt 0.5 . Die Impulsantwort lautet

dementsprechend $[0.5 \text{ } -0.5]$. Wie schon im Beispiel 1 läßt sich der Frequenzgang des Filters schätzen. Für sehr tiefe Frequenzen mit ähnlichen aufeinanderfolgenden Abtastwerten ist der Ausgang des Filters stark gedämpft. Sind die aufeinanderfolgenden Abtastwerte gleich, so ist der Ausgang des Filters 0. Mit zunehmender Eingangsfrequenz nimmt auch die Amplitude des Ausgangs zu, bis er bei der halben Abtastfrequenz maximal wird. Es handelt sich also um ein *Hochpaßfilter*.

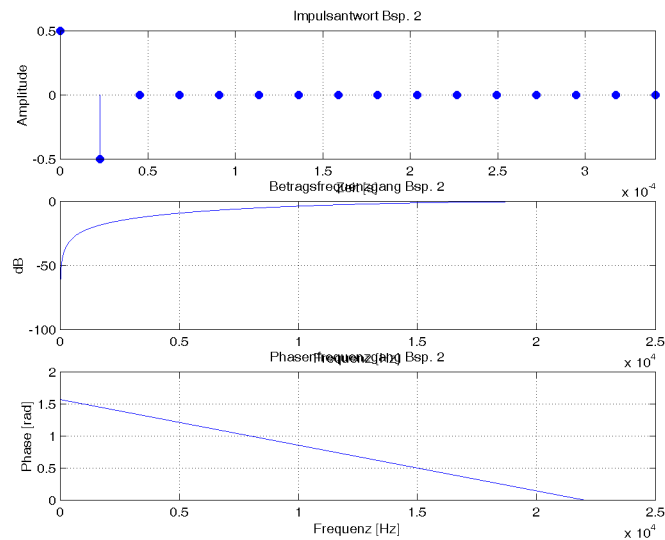


Abbildung 8.13: Impulsantwort, Betragsfrequenzgang und Phasenfrequenzgang von Beispielfilter Nr.2

Abbildung 8.13, in der Impulsantwort, Betragsfrequenzgang und Phasenfrequenzgang des Filters gezeigt werden, bestätigt die Herleitung der Impulsantwort und des Betragsfrequenzgangs. Der Phasenfrequenzgang ist wie in Beispiel 1 linear.

Beispiel 3

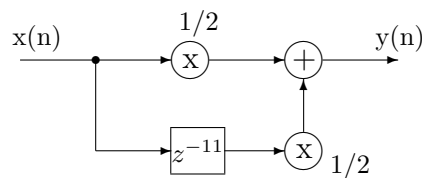


Abbildung 8.14: Filter Beispiel 3

Auch dieses Beispiel weist große Ähnlichkeit mit Beispiel 1 auf. Die Filterkoeffizienten sind identisch, das Verzögerungsglied verzögert hingegen nicht mehr nur

um einen Abtasttakt, sondern um 11. Wenn wir uns den Betragsfrequenzgang wie bei den obigen Beispielen herleiten, kommen wir zu dem Schluß, daß immer eine ganz bestimmte Frequenz und deren ganzzahlige Vielfache unterdrückt werden. Es handelt sich also um ein *Kammfilter*.

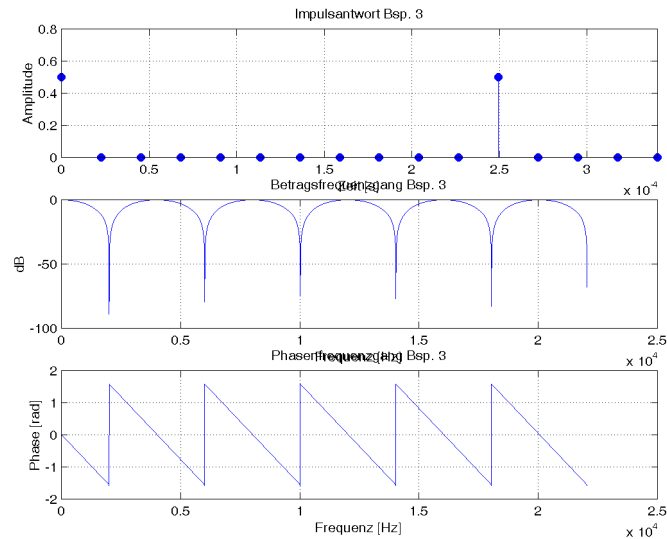


Abbildung 8.15: Impulsantwort, Betragsfrequenzgang und Phasenfrequenzgang von Beispielfilter Nr.3

Abbildung 8.13 zeigt die Impulsantwort, den Betragsfrequenzgang und den Phasenfrequenzgang des Filters. Auch wenn es nicht so aussieht, ist auch in diesem Beispiel die Phase linear, lediglich die Darstellung kann Vielfache von $\frac{\pi}{2}$ nicht von $\frac{\pi}{2}$ selbst unterscheiden.

Beispiel 4

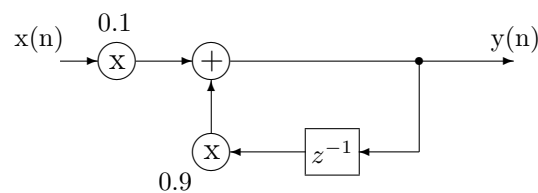


Abbildung 8.16: Filter Beispiel 4

In dem vierten und letzten Beispiel liegt eine andere Filterstruktur vor: der Verzögerungspfad ist nun rückgekoppelt, d.h. daß jeder Ausgangswert wieder gewichtet mit einem Eingangswert addiert wird. In den bisherigen Beispielen

hatte im Gegensatz dazu der Ausgangswert keinen Einfluß auf den folgenden Ausgangswert. Die Impulsantwort wird aufgrund der Rückkopplung niemals ganz abklingen (jedenfalls bei unendlicher Rechengenauigkeit), also unendlich ausgedehnt sein. Wäre der Faktor im Rückkopplungsweig größer als 1, so wird das Filter instabil und liefert immer größere Werte.

Der Amplitudenfrequenzgang dieses Filters läßt sich anschaulich nicht mehr so leicht herleiten wie in den obigen Beispielen. Wir sehen allerdings, daß der Rückkopplungsweig stärker den Ausgangswert bestimmt als der Eingang des Filters. Verändert sich also der Eingang des Filters nicht oder nur sehr langsam, so wird der Ausgang des Filters dem Eingang ähnlich sein. Ändert sich der Eingang hingegen schnell, so werden sich diese Änderungen nur wenig Auswirkungen auf den Ausgang haben. Also liegt die Vermutung nahe, daß es sich in diesem Beispiel wiederum um ein *Tiefpaßfilter* handelt. Abbildung 8.17 bestätigt diese Annahme. Der Phasengang ist nicht linear, der Betragsfrequenzgang dieses Tiefpaßfilters unterscheidet sich deutlich von dem aus Beispiel 1.

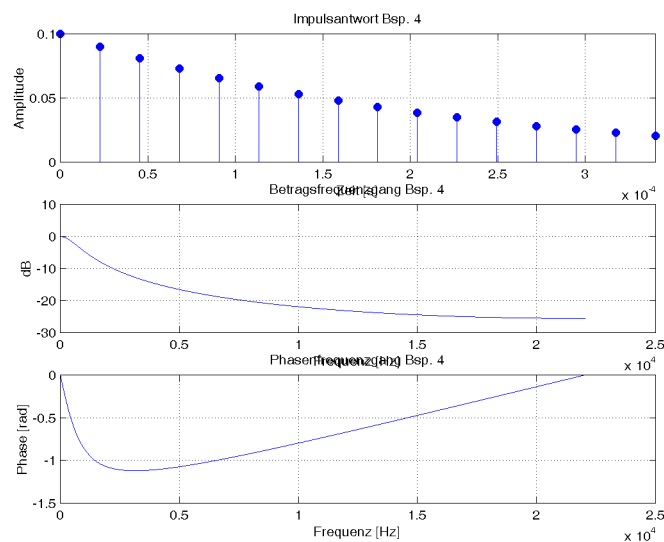


Abbildung 8.17: Impulsantwort, Betragsfrequenzgang und Phasenfrequenzgang von Beispielfilter Nr.4

Die ersten drei Beispiele sind Filter mit endlich langer Impulsantwort, aus diesem Grund werden sie als **FIR**-Filter (FIR steht für *Finite Impulse Response*) bezeichnet. Da FIR-Filter keine Rückkopplungsweig besitzen, stellen die FIR-Filterkoeffizienten auch die Impulsantwort des Filters dar. Beispiel 4 hingegen ist ein Filter mit (zumindest theoretisch) unendlich langer Impulsantwort, Filter dieser Art werden **IIR**-Filter (IIR für *Infinite Impulse Response*) genannt. Die Information, wie ein Filter denn nun tatsächlich klingt, verbirgt sich ausschließlich in den Koeffizienten bzw. der Berechnung der Koeffizienten (jedenfalls wenn man nichts falsch macht). Hier kann eine kleine Änderung schon deutliche klangliche Konsequenzen nach sich ziehen.

Verwandte Themenbereiche

- Abschnitt [6.4.2](#)

8.3.1 Filterstrukturen

Die Elemente eines IIR-Filters lassen sich auf unterschiedliche Art und Weise anordnen. Unterschiedliche Strukturen können unterschiedliche Vor- und Nachteile zur Folge haben, beeinflusst werden v.a.

- Rechenaufwand,
- Rechengenauigkeit und die
- Größe des benötigten Speichers.

Eine Änderung der Filterstruktur sollte, wenn korrekt durchgeführt, keine Auswirkungen auf die Audioqualität des Filters haben.

8.4 Speichermedien**8.4.1 Magnetbänder****8.4.2 Optische Medien****8.5 Übertragungstechnik/Fehlerbehandlung****8.5.1 Fehlerquellen bei der Übertragung****8.5.2 Jitter****8.5.3 Kanalcodes****8.5.3.1 Fehlerrobustheit****8.5.3.2 Fehlererkennung****8.5.3.3 Fehlerbehandlung****8.6 Codierungsverfahren**

Die für hochqualitative Audioübertragung oder -speicherung benötigte Bandbreite bzw. der Speicherplatz liegen sehr hoch. Aus diesem Grund ist man bemüht, die Menge der Daten zu reduzieren, ohne die Qualität (stark) zu beeinflussen. Dies ist Aufgabe von **Audiocodierungsverfahren** (auch *Audiokompressionsverfahren*¹).

¹Nicht zu verwechseln mit Compressoren (s. Abschnitt [6.4.5](#)), die nichts mit Datencodierung zu tun haben)

Diese spielen in vielen Bereichen in zunehmenden Maße eine bedeutende Rolle, auch wenn man dies nicht immer wahrnimmt (was ja auch der Sinn dieser Verfahren ist). So sind sie nicht nur im Internet mit ihrem sehr prominenten Vertreter *MP3* (MPEG-1 layer 3) vertreten, sondern werden auch in der Telefonie, beim Rundfunk und im Fernsehen, im Kino, auf DVDs und auch in alltäglichen Geräten wie z.B. dem MiniDisc-Recorder eingesetzt. Aus diesem Grund ist es auch für Toningenieure und Tonmeister sinnvoll, sich mit den Stärken und Schwächen dieser Verfahren auseinanderzusetzen.

Wie schon gesagt ist es Sinn und Zweck von Codierungsverfahren, die Datenmenge zur Übertragung oder Speicherung möglichst ohne hörbare Fehler (sogenannte **Artefakte**) für den Anwender zu gestalten. Damit dies gut funktioniert, sind teilweise recht komplexe Algorithmen erforderlich, die im folgenden ansatzweise beschrieben werden sollen.

8.6.1 Überblick über Codierungsverfahren

Eigenschaften von Codierungsverfahren (Bitrate, Kompressionsrate, CBR/VBR

8.6.2 Wo können Bits gespart werden

Es stellt sich natürlich die Frage, wieso Codierungsverfahren, die Neun Zehntel oder mehr der Daten einsparen, überhaupt so gut funktionieren können. Wandelt man z.B. ein *16bit*-Signal in ein *8bit*-Signal um, was lediglich einer Kompressionsrate von 1:2 entspricht, so hört man schon deutliche Artefakte. Der Kompressionsgewinn der Codierungsverfahren beruht einerseits auf Eigenschaften des Audiosignals selbst, andererseits auf Eigenschaften des menschlichen Gehörs.

Das Audiosignal besitzt Eigenschaften, die bekannt sind und die keine neue Information liefern. So ist z.B. fast allen Signalen zueigen, daß niedrige Amplitudenwerte wesentlich häufiger auftreten als hohe. Auch ist es z.B. im Falle eines positiven Amplitudenwertes sehr viel wahrscheinlicher, daß der darauffolgende Wert ebenfalls positiv ist als daß er negativ ist. Solche Eigenschaften nutzt die **Redundanzcodierung** aus, indem nur nicht-redundante Information übertragen oder gespeichert wird. Dabei gehen keine Daten/Informationen verloren.

8.6.3 Redundanzcodierung

Da bei der Redundanzcodierung keine Daten oder Informationen verloren gehen, handelt es sich um ein *verlustloses* (lossless) Codierungsverfahren. Da eine bitgenaue Rekonstruktion des Signals möglich ist, kann auch mehrfache Codierung und Decodierung die Signalqualität nicht verschlechtern. Ein etabliertes verlustloses Codierungsverfahren auf dem PC ist beispielsweise das ZIP-Verfahren, das Packprogramm genannt wird.

Verlustlose Verfahren haben neben dem Vorteil der Verlustlosigkeit auch zwei wichtige Nachteile:

- die Kompressionsrate ist mit Faktor 1.5 – 4 relativ gering
- die Ausgangsbitrate ist materialabhängig und kann nicht konstant gehalten werden

Redundanzcodierungsverfahren arbeiten heutzutage zumeist nach dem Prädiktionsprinzip; aufgrund der letzten Abtastwerte wird versucht, die kommenden Abtastwerte vorherzusagen. Der dabei gemachte Fehler ist zum ursprünglichen Audiosignal meistens sehr klein. Nun muß nur noch der Prädiktionsfehler und einige Information zur Signalvorhersage beim Empfänger übertragen werden, was wesentlich effizienter ist.

8.6.4 Irrelevanzcodierung

Die Verfahren der **Irrelevanzcodierung** versuchen, für das menschliche Gehör wichtige Signalanteile von unwichtigen Signalanteilen zu trennen, und die unwichtigen Anteile gar nicht oder sehr verrauscht zu übertragen. Dabei sind je nach Verfahren Kompressionsraten von 1 : 4 – 1 : 30 erzielbar. Zudem ist es mit verlustbehafteten Verfahren meistens auch möglich, eine feststehende Bitrate zu erzielen, was für die Signalübertragung große Bedeutung haben kann. Im allgemeinen hat der Decoder keinen Einfluß auf die Qualität des encodierten und wieder decodierten Signals, diese wird meist ausschließlich von Encoder bestimmt.

8.6.4.1 Grundsätzlicher Aufbau verlustbehafteter Verfahren

Abb. 8.18 zeigt den prinzipiellen Aufbau der meisten wahrnehmungsangepaßten Codierungsverfahren.

Da das Codierungsverfahren versucht, wichtige (relevante) Signalanteile von unwichtigen zu trennen, ist eine umfassende Analyse des Eingangssignals nötig. Diese geschieht im sogenannten **psychoakustischen Modell**. Diese Analyse sowie die spätere Codierung werden im Frequenzbereich durchgeführt, wobei die Transformation mittels einer **Filterbank** durchgeführt wird. Dabei werden wichtige Eigenschaften des Gehörs wie die Verdeckungseffekte (vgl. Abschnitt 5.4) und die Frequenzauflösung des Gehörs modelliert. Das Ergebnis des psychoakustischen Modells sagt dann den anderen Codierungstools, welche Frequenzbänder besonders wichtig sind, und welche vernachlässigbar sind. Vor der eigentlichen Quantisierung des Signals kommen - abhängig vom jeweils betrachteten Codierungsverfahren - noch einige Tools, welche die Codierungseffizienz noch weiter steigern. Beispiele sind die Ausnutzung von Redundanzen zwischen zwei Stereokanälen, die Prediktion von Spektralwerten sowie die Veränderung der zeitlichen Struktur des Quantisierungsrauschens.

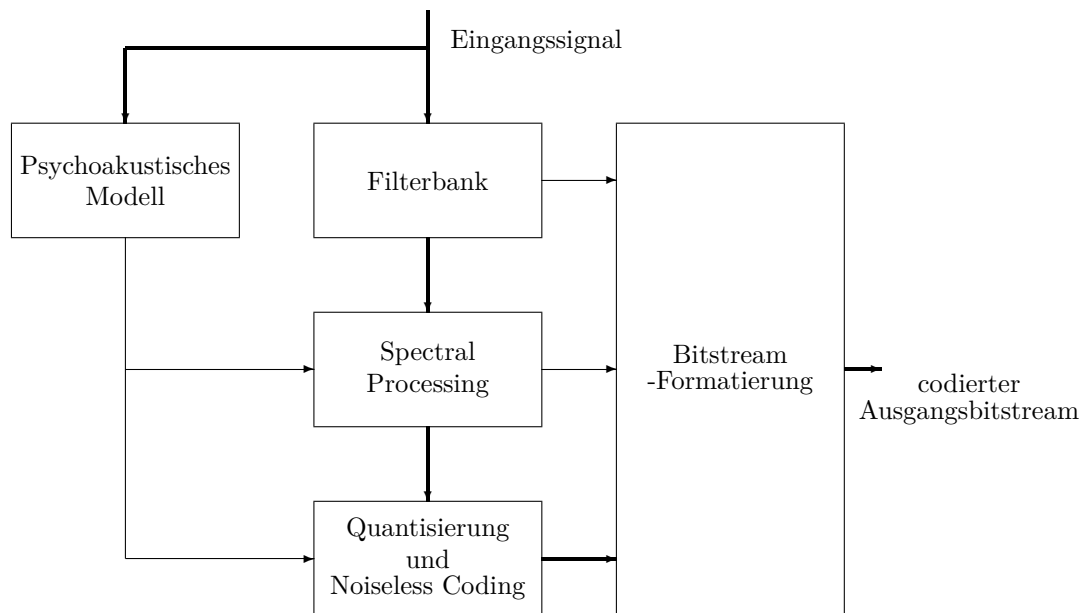


Abbildung 8.18: typischer Ablauf eines wahrnehmungsangepaßten Codierungsverfahrens, die dicken Pfeile markieren den Fluß der Audioinformationen, die dünnen den Fluß der Kontrolldaten

Der letzte und wichtigste Schritt des Codierungsverfahrens ist die Quantisierung. Basierend auf der Analyse des psychoakustischen Modells versucht der **Quantisierer**, wichtige Spektralanteile hochauflösend zu quantisieren und unwichtigere sehr grob zu quantisieren. Die Quantisierung im Zusammenhang mit der nachgeschalteten Redundanzcodierung der quantisierten Werte resultiert dann in dem Codierungsgewinn.

8.6.4.2 Typische Artefakte

- Bandbegrenzung
- Pre-Echo

- Quantisierungsrauschen
- Zwitschern
- Verzerrungen
- Aliasing
- Schwankungen/Verzerrungen des Stereobildes

8.6.4.3 Qualitätsbeurteilung

Die Qualität eines Codierungsverfahrens ist neben dem Algorithmus selbst abhängig von

- dem verwendeten Eingangssignal,
- dem verwendeten Encoder (es können mehrere verschiedene Encoder existieren, die auf dem gleichen Algorithmus beruhen),
- und den verwendeten Encodieroptionen.

Es gibt für jeden Encoder kritische und unkritische Testsignale. Bei unkritischen Testsignalen kann die Qualität selbst bei niedrigen Ausgangsbitraten sehr gut sein. Zur Qualitätsbeurteilung eines Verfahrens sollten möglichst kritische Testsequenzen ausgesucht werden, damit die Stärken und Schwächen (vielleicht auch im Vergleich mit anderen Verfahren) deutlich hervortreten.

Auch wenn verschiedene Encoder auf dem gleichen Prinzip beruhen, bedeutet das nicht automatisch, daß die resultierende Qualität gleich ist. Hier sollte man gegebenenfalls verschiedene Encoder vergleichen.

Mit ein wenig Tuning an den Encodieroptionen läßt sich die Encodierungsqualität oftmals signifikant im Hinblick auf das verwendete Eingangssignal und die angestrebte Ausgangsbitrate optimieren.

Die etablierten Verfahren zur Qualitätsmessung versagen im Zusammenhang mit Codierungsverfahren. Dies hat v.a. drei Gründe:

- die hohe Zeitinvarianz der Codierungsverfahren, die ca. alle 10 – 20ms ihr Übertragungsverfahren ändern können
- die Ausgangsqualität hängt stark vom Eingangssignal ab,
- durch die intensive Ausnutzung von psychoakustischen Erkenntnissen wird bewußt Rauschen insbesondere in verdeckten Frequenzbereichen eingeführt; wird der Pegel dieses Rauschens mit einfachen Mitteln wie dem SNR gemessen, so wird die „Unhörbarkeit“ dieses Rauschens nicht berücksichtigt.

Es existieren zwar Systeme, die versuchen, die Qualität von Codierungsverfahren objektiv zu messen, diese können allerdings bisher nur mit gewissen Einschränkungen verwendet werden. Somit bleibt als einzige und letzte Alternative zur Qualitätsbeurteilung von Codierungsverfahren nur die subjektive Qualitätsbeurteilung. Will man die Ergebnisse dieser subjektiven Beurteilung etwas objektivieren, so bleibt nur der aufwendige Hörtest (s. Abschnitt 6.5.7).

Verwandte Themenbereiche

- Abschnitt 6.5.7

8.6.4.4 Encodieroptionen

Dieser Abschnitt soll einen kurzen Überblick über die geläufigsten Encodieroptionen bieten und kurz ihre Auswirkungen skizzieren.

Die naheliegendsten und am häufigsten benutzten Encodieroptionen sind *Bitrate* und/oder *Qualitätsstufe*. Je höher die Bitrate, desto besser klingt i.a. das encodierte Signal. Daher beeinflussen sich diese beiden Parameter oft gegenseitig. Viele Encoder haben einen sog. *VBR*-Modus, für den lediglich noch die gewünschte Qualität selektiert wird und kein Einfluß mehr auf die Ausgangsbitrate genommen werden kann.

Über die einstellbare Grenzfrequenz des Tiefpaßfilters läßt sich eine Tiefpaßfilterung vor dem eigentlichen Encodiervorgang durchführen. Ein Tiefpaßfilter hilft dem Encoder unter Umständen, Zwischterartefakte zu vermeiden.

Unterschreitet ein Encoder seinen optimalen Bitratenbereich, kann sich die Qualität rapide mit sinkender Bitrate verschlechtern. Durch ein Downsampling des Eingangsbereichs läßt sich die Kompressionsrate wieder etwas verringern, so daß die empfundene Qualität steigt.

Bei sehr niedrigen Bitraten sinkt die Qualität oft so rasch, daß der Verzicht auf die Stereo-/Multichannelinformation sinnvoller ist als die deutlich hörbaren Codierungsartefakte in Kauf zu nehmen. Bei einem Downmix von Stereo nach Mono halbiert sich die Kompressionsrate, so daß der Encoder wieder Spielraum zur Qualitätsoptimierung hat.

8.6.4.5 Auswahlkriterien von Codierungsverfahren

Es existiert kein Audiocodierungsverfahren, das in jedem Einsatzbereich uneingeschränkt eingesetzt werden kann. Abhängig von Einsatzbereich lassen sich unterschiedliche Kriterien benennen, die im folgenden stichpunktartig dargestellt und erläutert werden sollen.

- **Audioqualität:** Die Qualität des codierten und wieder decodierten Signals ist sicherlich i.a. das wichtigste Kriterium bei der Auswahl des Codierungsverfahren und hängt mehr oder weniger direkt mit vielen der nachfolgenden Punkte zusammen. Ein sehr wichtiger Punkt im Zusammenhang mit der

Qualität wurde schon weiter oben angeführt: die Qualität ist bei zeitgemäßen Codierungsverfahren immer abhängig vom Eingangssignal, somit kann auch die Wahl des Codierungsverfahren vom zu codierenden Signal abhängen. Ist die Audioqualität das einzige Kriterium, so ist einem verlustloses Verfahren der Vorzug zu geben.

- **Bitrate:** Bitrate und Qualität haben direkt aufeinander Einfluß. Als Daumenregel kann man festhalten, daß die Qualität mit höherer Bitrate steigt. Verschiedene Verfahren sind immer auf bestimmte Bitraten ausgelegt und erzielen bei diesen Bitraten die besten Ergebnisse; bei abweichenden Bitraten können sie deutlich schlechter klingen als andere Verfahren. Es spielt ebenfalls eine Rolle, ob ein Verfahren eine *konstante* oder *variable* Bitrate erlaubt. Beispielsweise sind Verfahren mit variabler Bitrate für Streaminglösungen eher ungeeignet.
- **Komplexität:** Die Komplexität eines Verfahrens zeigt sich in der erforderlichen Rechenleistung für eine Codierung/Decodierung. Je komplexer ein Verfahren ist, desto mehr steigt die Auslastung des Rechners/Chips. Als Daumenregel gilt hier: je komplexer ein Verfahren ist, desto besser ist die Qualität.

Bemerkung

Im allgemeinen sind die Decoder wesentlich weniger aufwendig als die Encoder. Aus diesem Grund sind Decoder billiger und einfacher zu realisieren (z.B. portabel), während Encoder u.U. hohen Entwicklungsaufwand kosten

- **Delay:** In Einzelfällen, insbesondere im Falle zweiseitiger Kommunikation wie z.B. mit dem Telefon ist auch das Encodierungs-/Decodierungsdelay ein wichtiges Auswahlkriterium. Wenn dieses Delay groß ist wie z.B. bei den meisten MPEG-Verfahren, leidet der Gesprächsfluß unter dieser Einschränkung. Mit steigendem Delay steigt meistens auch die Qualität eines Verfahrens leicht an.
- **Verbreitung:** Je verbreiteter ein Verfahren ist, desto mehr Menschen können codierte Dateien ohne große Probleme abspielen. Will man also z.B. Demodateien für möglichst viele Hörer zugänglich machen, so ist die Wahl eines (im angestrebten Zielmarkt) sehr verbreiteten Verfahrens sinnvoll.
- **Kosten:** Die Kosten für die Benutzung eines Verfahrens schwanken. Teilweise dürfen Verfahren kostenlos benutzt werden, teilweise muß man vor der Benutzung eine Lizenz erwerben (meistens im Kaufpreis enthalten). Es kann sogar vorkommen, daß pro encodiertem Material Lizenzgebühren an den Rechteinhaber fällig werden.
- **Zukunftssicherheit:** Gerade bei der Anwendung von Codierungsverfahren für Archivierungen spielt die Frage der Zukunftssicherheit eines Verfahrens

eine wichtige Rolle, denn funktionsfähige Decoder müssen auch noch in mehreren/vielen Jahren zur Verfügung stehen. Anhaltspunkte dafür sind zum Beispiel, ob das Verfahren international standardisiert ist, ob es sich um einen internationalen de facto-Standard handelt und ob im Internet Quelltexte zu dem Verfahren zur Verfügung stehen.

- **Rechtssicherheit:** Auf den meisten Verfahren liegen mehrere Patente. Es existieren teilweise Codierungsverfahren, die diese Patente verletzen, den Anwender aber nicht darauf aufmerksam machen. Auch wenn die Patente im Moment nicht verfolgt werden, muß das nicht heißen, daß diese Verfahren auch in Zukunft bedenkenlos benutzt werden können.

8.7 Zusammenfassung

Die Zahlendarstellung in digitalen Systemen ist **binär**, sie besteht ausschließlich aus den Symbolen 0 und 1. Die **Wortbreite** eines Signals oder eines Systems beschreibt, wieviele Bits benötigt werden, um einen Wert darzustellen.

Bei der Konvertierung eines analogen Signal in ein digitales Signal muß das analoge Signal abgetastet und quantisiert werden. Die Abtastung findet in äquidistanten Abständen, die durch die **Abtastrate** festgelegt werden, statt. Es ist wichtig, daß das **Abtasttheorem** berücksichtigt wird, das als maximal erlaubte Frequenz die halbe Abtastfrequenz erlaubt. Die **Quantisierung** bildet die kontinuierliche Amplitude des Eingangssignals auf Quantisierungsstufen ab, deren Zahl von der Wortbreite abhängt. Der erzielte SNR hängt direkt mit der Wortbreite zusammen; als Daumenwert gilt ein Gewinn von $6dB$ SNR pro zur Wortbreite hinzugefügtem Bit.

Zur Erhöhung Qualität der Quantisierung, sowohl bei der A/D-Wandlung wie auch bei späterer Wortbreitenreduktion, können Methoden wie **Dithering** (Hinzufügen von Rauschen minimaler Leistung) und/oder **Noise-Shaping** (spektrale Formung des Quantisierungsfehlers) verwendet werden.

8.8 Aufgaben

1. Warum reicht der Zahlenbereich bei $16bit$ von -32768 bis 32767 und nicht bis 32768 ?
2. Berechne die Amplitudenwerte für folgende Binärzahlen (vorzeichenbehaftet):
 - 0000 0000 1001 1000
 - 0111 0001 0010 1111
 - 1000 1110 1101 0000

- 1111 1111 0110 0111
3. Wieviele unterschiedliche Amplitudenwerte können mit *24bit* Wortbreite dargestellt werden?
 4. Wie groß ist der Dynamikumfang eines Systems mit *20bit* bzw. *24bit* Wortbreite theoretisch?
 5. Berechne die Datenrate eines
 - *16bit*-Stereosignals der Abtastrate *44.1kHz*
 - *20bit*-Stereosignals der Abtastrate *48kHz*
 - *24bit*-Stereosignals den Abtastraten *48kHz*, *96kHz*, *192kHz*
 6. Mit welcher Frequenz muß ein Signal mindestens abgetastet werden, damit alle Frequenzen zwischen *0Hz* und *24kHz* theoretisch fehlerfrei rekonstruiert werden können?
 7. Warum ist die Zahl der Quantisierungsstufen immer eine Zweierpotenz (2,...,256,65536)?
 8. Warum kann die Verdopplung der Abtastfrequenz bei Oversampling zu einem Gewinn von *3dB* beim SNR führen?
 9. Wieviele dB SNR-Gewinn lassen sich theoretisch durch 128-fach Oversampling erzielen?
 10. CD Speicherdichte - Computer vs. Audio
 11. Kompressionsrate vs. Bitrate vs. Abtastrate
 12. Gibt es einen Zusammenhang zwischen Audiokompressionsverfahren und Compressoren? Wenn ja, wo liegt dieser?

Anhang

Anhang A

Lösungen der Aufgaben

A.1 Schwingungen und Wellen

1. Es gilt „Länge gleich Zeit mal Geschwindigkeit“:

$$\begin{aligned}\Delta l &= t \cdot c \\ t &= \frac{\Delta l}{c}\end{aligned}$$

Mit der Schallgeschwindigkeit $340 \frac{m}{s}$ ergibt sich

$$t = \frac{10m}{340 \frac{m}{s}} = 29.41ms$$

Nach der Daumenregel ergibt sich ein Wert von $10m \cdot 3 \frac{ms}{m} = 30ms$

2. Die Vorzeichen der beiden Schwingungen sind genau entgegengesetzt, d.h. sie löschen sich gegenseitig aus. Mathematisch lässt sich das für sinusförmige Schwingungen wie folgt interpretieren. Die erste Schwingung definiert man durch $x_1(t) = \sin\left(2 \cdot \pi \frac{t}{T_p}\right)$. Eine Verschiebung dieser Schwingung um 180° führt zu der Schwingung $x_2(t) = \sin\left(2 \cdot \pi \frac{t}{T_p} + \pi\right)$. Die Schwingung $x_2(t)$ kann nun mathematisch ausgedrückt werden durch $x_2(t) = -\sin\left(2 \cdot \pi \frac{t}{T_p}\right)$, wodurch man für das resultierende Signal $y(t)$ erhält:

$$y(t) = \sin\left(2 \cdot \pi \frac{t}{T_p}\right) - \sin\left(2 \cdot \pi \frac{t}{T_p}\right) = 0$$

3. Die Periodendauer ist die Zeit, in der an einem *festen Ort* der gleiche Zustand (Amplitudenwert) wiederkehrt. Die Wellenlänge hingegen gibt die minimale Entfernung zweier gleicher Schwingungszustände zu einem *festen Zeitpunkt* an.

4. Es gilt nach Gleichung (1.7) $\lambda = \frac{c}{f}$. Somit berechnen sich folgende Wellenlängen:

$$\begin{aligned}\lambda_{16Hz} &= 21.25m \\ \lambda_{100Hz} &= 3.4m \\ \lambda_{1kHz} &= 0.34m \\ \lambda_{10kHz} &= 0.034m = 3.4cm \\ \lambda_{20kHz} &= 0.017m = 1.7cm\end{aligned}$$

5. Zunächst können die verschiedenen Ausbreitungsgeschwindigkeiten mit Gleichung (1.6) berechnet werden. Es ergibt sich:

$$\begin{aligned}c_{14.35^\circ} &= 331.4 \frac{m}{s} \cdot \sqrt{\frac{287.35}{273}} = 340.00 \frac{m}{s} \\ c_{16^\circ} &= 331.4 \frac{m}{s} \cdot \sqrt{\frac{289}{273}} = 340.97 \frac{m}{s} \\ c_{19^\circ} &= 331.4 \frac{m}{s} \cdot \sqrt{\frac{292}{273}} = 342.74 \frac{m}{s} \\ c_{22^\circ} &= 331.4 \frac{m}{s} \cdot \sqrt{\frac{295}{273}} = 344.49 \frac{m}{s}\end{aligned}$$

Im zweiten Schritt wird die Wellenlänge nach Gleichung (1.7) berechnet:

$$\begin{aligned}\lambda_{const} &= \frac{c}{f} \\ &= \frac{340 \frac{m}{s}}{440 \frac{1}{s}} \\ &= 0.77m\end{aligned}$$

Über die Gleichung (1.7) können nun die Frequenzen für die Temperaturen berechnet werden:

$$\begin{aligned}f_{16^\circ} &= \frac{340.97 \frac{m}{s}}{0.77m} = 442.82Hz \\ f_{19^\circ} &= \frac{342.74 \frac{m}{s}}{0.77m} = 445.12Hz \\ f_{22^\circ} &= \frac{344.49 \frac{m}{s}}{0.77m} = 447.39Hz\end{aligned}$$

- 6.

A.2 Schall und Schallfeld

1. Das Nahfeld geht nach Gleichung (2.7) abhängig von der Wellenlänge ins Fernfeld über mit:

$$r = \frac{\lambda}{2\pi} = \frac{c}{2\pi \cdot f}$$

Somit berechnet sich der Abstand r für unterschiedliche Frequenzen zu:

$$\begin{aligned} r_{50} &= 1.08m \\ r_{200} &= 0.27m \\ r_{500} &= 0.11m \\ r_{1000} &= 0.05m \\ r_{4000} &= 0.01m \end{aligned}$$

2. Aus Gleichung (2.3) ist bekannt, daß der Schalldruck umgekehrt proportional zur Entfernung ist. Der zu berechnende Faktor F ist das Verhältnis des Schalldrucks in $30cm$ Entfernung zu dem in $180cm$ Entfernung. Es gilt also:

$$F = \frac{\frac{1}{0.3m}}{\frac{1}{1.8m}} = \frac{180}{30} = 6$$

3. Zunächst muß der Weg berechnet werden, den der reflektierte Schall zusätzlich zurücklegen muß. Dieser läßt sich über den bekannten *Satz des Pythagoras* bestimmen, der besagt, daß in einem rechtwinkligen Dreieck das Quadrat Länge der Hypotenuse c_H (längster Dreieckschenkel, gegenüber dem rechten Winkel) gleich der Summe der Quadrate der Längen der beiden Katheten a_K, b_K sind:

$$c_H^2 = a_K^2 + b_K^2$$

Nach einigem Überlegen ergibt sich für den Gesamtweg der Reflexion l_R

$$l_R = 2 \cdot \sqrt{[3m]^2 + \left[\frac{5}{2}m\right]^2} = 7.82m$$

und es folgt daraus der zusätzlich zurückgelegte Weg Δl

$$\Delta l = l_R - 5m = 2.82m$$

Mit der Beziehung „Länge gleich Zeit mal Geschwindigkeit“ erhält man für die Zeitverzögerung der Reflexion:

$$\Delta t = \frac{\Delta l}{c} = 8ms$$

Der Dämpfungsfaktor berechnet sich wie in Aufgabe 2 über das Verhältnis der beiden Strecken:

$$F = \frac{\frac{1}{5m}}{\frac{1}{5m + \Delta l}} = \frac{7.82}{5} = 1.56$$

4. Als Daumenregel gilt, daß ein Hindernis den Schall reflektiert, dessen Wellenlänge kleiner ist als die Dimension des Hindernisses. Hierfür gilt die kleinere Abmessung des Hindernisses als Anhaltspunkt, in diesem Fall also 30cm . Eine Wellenlänge von 30cm entspricht nach Gleichung (1.7) einer Frequenz von ca. 1100Hz . Höhere Frequenzen werden stark gedämpft.

A.3 Pegel

1. Gleichung (3.1) lautet:

$$dW = \text{const.} \cdot \frac{dR}{R}$$

Zur Lösung dieser Differentialgleichung müssen beide Seiten integriert werden, so daß man erhält:

$$\begin{aligned} \int 1dW &= \int \text{const.} \cdot \frac{1}{R} dR \\ \Leftrightarrow \\ W + C_1 &= \text{const.} \cdot \ln R + \text{const.} \cdot C_2 \\ \Leftrightarrow \\ W &= \text{const.} \cdot \ln R + \underbrace{\text{const.} \cdot C_2 - C_1}_C \end{aligned}$$

2. In den Aufgaben zu Abschnitt 2 wurde schon deutlich, daß sich der Faktor zur Verringerung des Schalldrucks aus dem Verhältnis der Entfernungen berechnet. Um den Pegelwert zu berechnen, muß der Logarithmus dieses Verhältnisses berechnet werden:

$$20 \log \frac{1/r}{1/2r} = 20 \log 2 = 6.02 \simeq 6 \text{ [dB]}$$

3. Die Pegeldifferenz berechnet sich analog zu Gleichung (3.7):

$$\begin{aligned} 20 \log \frac{1}{1} &= 0 \text{ [dB]} \\ 20 \log \frac{1}{\sqrt{2}} &= -3.01 \text{ [dB]} \\ 20 \log \frac{1}{2} &= -6.02 \text{ [dB]} \\ 20 \log \frac{1}{10} &= -20 \text{ [dB]} \end{aligned}$$

4. Das Spannungsverhältnis läßt sich aus Gleichung (3.4) herleiten:

$$\begin{aligned} 10^{3/20} &= 1.41 : 1 \\ 10^{6/20} &= 2.00 : 1 \\ 10^{10/20} &= \sqrt{10} : 1 = 3.16 : 1 \\ 10^{20/20} &= 10 : 1 \\ 10^{100/20} &= 100000 : 1 \end{aligned}$$

5. Nach Gleichung (3.10) ergibt sich für 4 Quellen:

$$10 \cdot \log \frac{4u_{1,eff}}{u_{1,eff}} = 6.02 [dB]$$

Für 8 Quellen ergibt sich analog:

$$10 \cdot \log \frac{8u_{1,eff}}{u_{1,eff}} = 9.03 [dB]$$

6. Wie in Aufgabe 2 berechnet sich die Pegeldifferenz über das Verhältnis:

$$20 \log \frac{1/60}{1/180} = 20 \log 3 = 9.03 \simeq 9 [dB]$$

7. Es handelt sich bei den Lautsprechern um eine akustische Addition, so daß gilt:

$$10 \cdot \log 4 - 10 \cdot \log 2 = 10 \cdot \log 2 = 3.01 [dB]$$

8. Diese Aufgabe ist in der Art ihrer Berechnung und in ihrer Lösung identisch zu Aufgabe 7, da es sich in beiden Fällen um eine Verdoppelung der Schallquellen und um eine akustische Addition handelt.

- 9.

A.4 Raumakustik

1. Die Nachhallzeit T berechnet sich nach Gleichung (4.1) mit

$$T \simeq 0.163 sm^{-1} \frac{21000m^3}{1711.5m^2} = 2s$$

Die Maße entsprechen grob der besetzten Berliner Philharmonie.

2. Der Hallradius r_H berechnet sich nach Gleichung (4.2) für die o.g. Werte mit

$$r_H = 0.057 s^{1/2} m^{-1/2} \cdot \sqrt{\frac{21000 m^3}{2s}} = 5.84 m$$

3.

4. Stehende Wellen können entstehen, wenn die Raumabmaße der halben Wellenlänge oder einem Vielfachen k der halben Wellenlänge entsprechen. Es gilt also für die tiefste gesuchte Wellenlänge

$$\lambda_5 = 2 \cdot 5m = 10m$$

$$\lambda_8 = 2 \cdot 8m = 16m$$

Die tiefsten Frequenzen sind dann (bei einer Schallgeschwindigkeit von $340 \frac{m}{s}$)

$$f_5 = \frac{340 m/s}{10m} = 34 Hz$$

$$f_8 = \frac{340 m/s}{16m} = 21.25 Hz$$

Da auch ganzzahlige Vielfache dieser Grundwellen als stehende Wellen auftreten, sind die auftretenden stehenden Wellen also bei den Frequenzen $k \cdot 34 Hz$ und $k \cdot 21.25 Hz$ (k ist eine ganze Zahl größer 0)

5. Die Seitenlängen des quadratischen Raums sind $\sqrt{2500} = 50m$. Der von der Reflexion zusätzliche zurückgelegte Weg berechnet sich über den *Satz des Pythagoras* mit

$$\Delta l = 2 \cdot \sqrt{[25m]^2 + [15m]^2} - 30m = 28.31m$$

Der Laufzeitunterschied ist somit

$$\Delta t = \frac{\Delta l}{c} = 83.26 ms$$

A.5 Psychoakustik

1. Der Abstand Δl , den das zweite Mikrophon weiter entfernt ist, kann wiederum über den *Satz des Pythagoras* berechnet werden. Es ergibt sich:

$$\Delta l = \sqrt{[5m]^2 + [0.5m]^2} - 5m = 2.49 cm$$

Daraus berechnet sich der Laufzeitunterschied bei einer Schallgeschwindigkeit von $340 \frac{m}{s}$ mit

$$\Delta t = \frac{\Delta l}{c} = 73.35 \mu s$$

und der Pegelunterschied zu

$$20 \log \frac{1/5}{1/5.0249} = 0.43 \text{ [dB]}$$

Sowohl Laufzeitunterschied als auch Pegelunterschied sind vernachlässigbar klein.

2. Zunächst muß die einem Laufzeitunterschied von $1ms$ entsprechende Längendifferenz Δl bestimmt werden:

$$\Delta l = 340 \frac{m}{s} \cdot 0.001s = 0.34m$$

Die gesuchte Mikrofonbasis B berechnet sich nun wie folgt:

$$\begin{aligned} \sqrt{[5m]^2 + B^2} - 5m &= 0.34m \\ \Leftrightarrow \sqrt{[5m]^2 + B^2} &= 0.34m + 5m \\ \Leftrightarrow B &= \sqrt{[5.34m]^2 - [5m]^2} \\ \Leftrightarrow B &= 1.87m \end{aligned}$$

3. Das Signal muß infolge der Beugung einen zusätzlichen weg entsprechend des halben Kopfumfangs zurücklegen. Der Kopfumfang K kann mithilfe des Radius r berechnet werden mit

$$K = 2 \cdot \pi \cdot r$$

Somit muß der Schall einen zusätzlichen Weg $\Delta l = K/2 = \pi \cdot 0.1m = 0.31m$ zurücklegen. Bei einer Schallgeschwindigkeit c von $340 \frac{m}{s}$ ergibt sich somit ein Laufzeitunterschied

$$\Delta t = \frac{\Delta l}{c} = 0.9ms$$

A.6 Elektrotechnik

A.7 Signalverarbeitung

1. Es ist gleichgültig, ob die Regelung der Dämpfung oder Verstärkung nichtlinear gemacht wird; ausschlaggebend ist, daß ausschließlich eine Skalierung des Signals stattfindet, die somit linear ist.

2. Nach Gleichung (6.3) berechnet sich die Güte eines Filters durch das Verhältnis von Mittenfrequenz zu Bandbreite. Daher lautet das Ergebnis:
 - $Q_{1600} = 2$
 - $Q_{3200} = 1$
 - $Q_{4000} = 0.8$
 - $Q_{6400} = \frac{1}{2}$
3. Da die Flankensteilheit des Filters $12 \frac{dB}{Octave}$ beträgt, ist die Frequenz $200Hz$ um $12dB$ gedämpft. Die Flankensteilheit könnte genauso mit $15 \frac{dB}{\text{kleine Dezime}}$ angegeben werden. Daraus folgt, daß die gesuchte Frequenz eine kleine Dezime über $100Hz$ liegt. Da eine kleine Dezime zur Grundfrequenz das Verhältnis $\frac{12}{5}$ aufweist, ist die gesuchte Frequenz $240Hz$.
Alternativ könnte man auch in der temperierten Stimmung rechnen: eine kleine Dezime entspricht 15 Halbtönen; da sich die Frequenz eines Halbtones zu der Frequenz f mit $f \cdot \sqrt[12]{2}$ berechnen läßt, liegt demnach die gesuchte Frequenz bei $100Hz \cdot \left(\sqrt[12]{2}\right)^{15} = 200Hz \cdot 2^{\frac{1}{4}} = 237.84Hz$
4. Die gesuchte Zeit entspricht genau der Dauer eines Schlages. Das Tempo $90BPM$ entspricht $\frac{90}{60} = 1.5 \text{ Beats per second}$. Daraus folgt, daß die Dauer eines Schlages $\frac{2}{3}s = 0.667s$ beträgt.
5. Eine Dämpfung der Amplitude um 50% entspricht einer Dämpfung um $6dB$, vgl. Abschnitt 3
6. Mit Erhöhung des Pre-Delays wird das Signal präsenter bzw. näher klingen, da durch die Verzögerung des (sehr weit im Raum klingenden) verhallten Signals zunächst der trockene, unverhallte Anteil des Signals wahrzunehmen ist.

A.8 Mikrophone

A.9 Digitaltechnik

1. Insgesamt stehen einem $16bit$ -Signal 65536 verschiedene Amplitudenwerte zur Verfügung. Da ein möglicher Amplitudenwert die 0 ist, steht keine gerade Anzahl an restlichen Werten zur Verfügung, die man auf der positiven wie negativen Seite verteilen könnte. Durch die Wahl des ersten Bits als Vorzeichenbit ergibt es sich nun, daß die positive Seite über einen Amplitudenwert weniger verfügt.
2. • $0000\ 0000\ 1001\ 1000 \hat{=} 2^7 + 2^4 + 2^3 = 152$

- $0111\ 0001\ 0010\ 1111 \hat{=} 2^{14} + 2^{13} + 2^{12} + 2^8 + 2^5 + 2^3 + 2^2 + 2^1 + 2^0 = 28975$
 - $1000\ 1110\ 1101\ 0000 \hat{=} -(2^{14} + 2^{13} + 2^{12} + 2^8 + 2^5 + 2^3 + 2^2 + 2^1 + 2^0) - 1 = -28976$
 - $1111\ 1111\ 0110\ 0111 \hat{=} -(2^7 + 2^4 + 2^3) - 1 = -153$
3. Die Anzahl der möglichen Amplitudenwerte für ein 24bit -Signal ist $2^{24} = 16777216$.
 4. Der Dynamikumfang des Systems läßt sich berechnen mit 6.02dB mehr Dynamik pro Bit. Daher ist der Dynamikumfang eines 20bit -Systems theoretisch 120.4dB und der eines 24bit -Systems 144.5dB . Die Werte realer Wandler liegen deutlich darunter.
 5. Die Datenrate eines Mono-Audiosignals berechnet sich mit $\text{Bits_per_Sample} \cdot \text{Samples_per_second}$. Somit erhält man

$$BR_{16,44.1} = 16 \frac{\text{bit}}{\text{Sample}} \cdot 44100 \frac{\text{Sample}}{\text{s}} = 705.6\text{ kbit/s}$$

$$BR_{20,48} = 20 \frac{\text{bit}}{\text{Sample}} \cdot 48000 \frac{\text{Sample}}{\text{s}} = 960\text{ kbit/s}$$

$$BR_{24,48} = 24 \frac{\text{bit}}{\text{Sample}} \cdot 48000 \frac{\text{Sample}}{\text{s}} = 1.152\text{ Mbit/s}$$

$$BR_{24,96} = 24 \frac{\text{bit}}{\text{Sample}} \cdot 96000 \frac{\text{Sample}}{\text{s}} = 2.304\text{ Mbit/s}$$

$$BR_{24,192} = 24 \frac{\text{bit}}{\text{Sample}} \cdot 192000 \frac{\text{Sample}}{\text{s}} = 4.608\text{ Mbit/s}$$

6. Nach dem Abtasttheorem muß die Abtastrate mehr als doppelt so hoch sein wie die höchste darzustellende Frequenz. Daher gilt hier

$$f_s > 48\text{ kHz} = 2 \cdot 24\text{ kHz}$$

7. Die Zahl der maximal darstellbaren Amplitudenstufen in Bitdarstellung ist immer eine Zweierpotenz, da es sich um das Binäresystem handelt: 2bit können $2^2 = 4$ Werte darstellen, 3bit $2^3 = 8$ Werte, 4bit $2^4 = 16$ Werte etc. Soll der Wertebereich des Wortes voll ausgeschöpft werden - was sinnvoll ist -, so muß die Anzahl der Wandlerstufen einer Zweierpotenz entsprechen.
8. Das Spektrum des Quantisierungsfehlers ist näherungsweise weiß, d.h. jede Frequenz ist gleichstark vertreten. Somit bildet das Spektrum näherungsweise eine horizontale (über die Frequenzen) Linie. Über die Integration über das Quadrat dieser Linie errechnet sich die Leistung des Quantisierungsrauschens. Wird die Abtastfrequenz verdoppelt, so muß das Ergebnis

des Integrals gleich bleiben, da die Leistung des Rauschens nicht von der Abtastfrequenz abhängig ist; das Rauschen erstreckt sich allerdings über einen doppelt so großen Frequenzbereich, so daß die Amplitude pro Frequenz sich halbiert haben muß. Nach Tiefpaßfilterung verschwindet die obere Hälfte des Quantisierungsrauschens, so daß die Gesamtleistung des Rauschens sich halbiert hat. Aus der Halbierung der Leistung des Quantisierungsfehlers folgt eine Erhöhung des Signal-Rauschabstands um $3dB$.

9. Jede Verdopplung der Abtastfrequenz führt zu einer Erhöhung des SNR um $3dB$. Wievielen Verdopplungen entspricht eine 128-fache Abtastfrequenz? Dies läßt sich über die Gleichung $2^x = 128$ bestimmen. Es ergibt sich:

$$x = \text{ld}(128) = 7;$$

7 Verdopplungen führen zu einer Erhöhung des Signal-Rauschabstands um $21dB$.

10.

Verzeichnis der Abkürzungen und Formelzeichen

Abkürzung	Erläuterung
ADV	Amplitudendichteverteilung (englisch: PDF für probability density function)
CBR	Constant Bit Rate
DRF	Directivity Factor (Bündelungsgrad)
DSF	Distance Factor (Vergrößerungsfaktor oder Abstandsfaktor)
FIR	Finite Impulse Response
IIR	Infinite Impulse Response
LSB	Least Significant Bit
MSB	Most Significant Bit
PDF	probability density function (s. auch ADV)
PZM	pressure zone microphone (s. auch Grenzflächenmikrophon)
REE	Random Energy Efficiency
RMS	Root Mean Square (Effektivwert)
SNR	Signal-to-Noise-Ratio (Signal-Rauschabstand) in dB
THD	Total Harmonic Distortion (Klirrfaktor)
THD+N	Total Harmonic Distortion and Noise
VBR	Variable Bit Rate

Einheiten	Erläuterung
$[^{\circ}C]$	Grad Celsius
$[Hz]$	Hertz

$[m]$	Meter
$[Pa]$	Pascal ($1Pa = 10\mu bar$)
$[s]$	Sekunde
$[V]$	Volt
$[W]$	Watt

$[W]$	Watt
-------	------

Formelzeichen**Erläuterung**

α	Absorptionsgrad
A	äquivalente Absorptionsfläche [m^2]
B_G	Bandbreite eines Filters [Hz]
c	Schallgeschwindigkeit [$\frac{m}{s}$]
Δ	Stufenhöhe eines Quantisierers
F	Flankensteilheit eines Filters [$\frac{dB}{Octave}$]
f	Frequenz [Hz]
f_0	Resonanzfrequenz [Hz]
f_S	Abtastrate [Hz]
f_C	Mittenfrequenz eines Filters/ Frequenzbandes [Hz]
f_G	Grenzfrequenz eines Filters [Hz]
γ	Bündelungsgrad
$\Gamma(\theta)$	Richtungsfaktor
$h(t)$	Impulsantwort eines Systems
$H(f)$	Transferfunktion oder Übertragungsfrequenzgang eines Systems
k	Klirrfaktor
J	Schallintensität [$\frac{W}{m^2}$]
λ	Wellenlänge [m]
L_{SPL}	absoluter Schalldruckpegel
L_P	absoluter Schalleistungspegel
L_U	absoluter Spannungspegel
$\ln(\cdot)$	Logarithmus Naturalis
$\log(\cdot)$	Logarithmus zur Basis 10
$\lg(\cdot)$	Logarithmus zur Basis 2
ϕ	Phasenwinkel
p	Schalldruck [Pa]

P	Schalleistung $[W]$
p_{eff}	effektiver Schalldruck $[Pa]$
$q(t)$	Quantisierungsfehler
Q	Güte eines Filters
r	Entfernung $[m]$
r_H	Hallradius $[m]$
$r_{H,eff}$	effektiver Hallradius $[m]$
θ	Einfallswinkel
t	Zeit $[s]$
T	Nachhallzeit $[s]$
T_p	Periodendauer $[s]$
v	Schallschnelle $[\frac{m}{s}]$
V	Volumen $[m^3]$
W	digitale Wortbreite $[Bit]$
ζ	Temperatur

Abbildungsverzeichnis

1.1	Periodendauer und Wellenlänge einer Schwingung	2
1.2	Phasenverschiebung um den Winkel ϕ	5
1.3	Schwebung	6
1.4	Sinusschwingung mit harmonischen Oberschwingungen	7
1.5	Frequenzlage der Formanten für die Vokale der deutschen Sprache	8
2.1	Ausbreitung einer Welle und Schalldruck	14
2.2	Verlauf der Schallschnelle über dem Abstand r	16
2.3	Reflexion einer Ebenen Welle	18
2.4	Beugung von Wellen hinter einem Hindernis	19
3.1	Darstellung der Hörfläche	24
3.2	Veranschaulichung der Addition von Effektivwerten	26
3.3	Addition zueinander verschobener Sinusignale	27
4.1	Reflexionsverhalten eines Raumes	31
4.2	Nachhall während eines Konzerts	33
4.3	Direktschall, erste Reflexionen und Diffusschall am Ort des Hörers	34
4.4	Pegeldämpfung über der Entfernung von der Schallquelle	35
4.5	Hallradien in Abhängigkeit von Raumvolumen und Nachhallzeit	35
4.6	Druckverteilung in stehenden Wellen zwischen zwei Wänden	37
4.7	Druck- und Schnelleverlauf in einer stehenden Welle	38
5.1	Schnitt durch das Gehör	42
5.2	Schnitt durch die Schnecke	43
5.3	Darstellung der Schnecke in abgerolltem Zustand	43
5.4	Wanderwelle auf der Basilarmembran	44
5.5	Flüssigkeitsströmung in der Schnecke	44
5.6	Normalkurven gleicher Lautstärkepegel	45
5.7	Mithörschwelle eines Sinustones	47
5.8	Pre- und Postmasking	47
5.9	Laufzeitunterschiede zwischen beiden Ohren für eine Schallquelle	49
5.10	Richtungseindruck aufgrund der Signalfrequenz	50
5.11	richtungsbestimmende Bänder in der Medianebene	50

5.12	Echoschwelle bei Sprache	51
5.13	Tradingeffekt	51
5.14	Übliche Lautsprecher-Stereoaufstellung	52
5.15	Lokalisation einer Phantomschallquelle (Sprache)	53
5.16	Lokalisation einer Phantomschallquelle (Sinusburst)	53
6.1	Zeitdarstellung eines Audiosignals	56
6.2	FFT eines Audiosignals	57
6.3	Betragsfrequenzgang	58
6.4	Spektrogramm eines Audiosignals	58
6.5	3D-Frequenz/Zeitverlaufsdarstellung	59
6.6	ADV eines Musiksignals	59
6.7	Grundsignale	62
6.8	verschiedene Rauschen	63
7.1	Mikrophon-Richtcharakteristiken	78
7.2	Maße zur Beschreibung der Richtwirkung	79
7.3	mechanischer Mikrophon-Aufbau	80
7.4	Mikrophonmembranauslenkung	81
7.5	elektrischer Kondensatormikrophon-Aufbau	83
7.6	Mikrophonmembrangeschwindigkeit	84
8.1	Abtastung eines analogen Signals	89
8.2	Abtastung von Sinusschwingungen unterschiedlicher Frequenz (1)	90
8.3	Abtastung von Sinusschwingungen unterschiedlicher Frequenz (2)	91
8.4	Quantisierungskennlinie	93
8.5	Quantisierungsfehler eines optimal ausgesteuerten Sinussignals	94
8.6	Quantisierungsfehler bei einer 3-stufigen Quantisierung	96
8.7	Quantisierungsfehler bei einer leicht geditherten Quantisierung	97
8.8	Spektrum eines geditherten Signals	98
8.9	Ditherformen	99
8.10	Filter Beispiel 1	100
8.11	Impulsantwort und Übertragungsfunktion von Filter Bsp.1	101
8.12	Filter Beispiel 2	101
8.13	Impulsantwort und Übertragungsfunktion von Filter Bsp.2	102
8.14	Filter Beispiel 3	102
8.15	Impulsantwort und Übertragungsfunktion von Filter Bsp.3	103
8.16	Filter Beispiel 4	103
8.17	Impulsantwort und Übertragungsfunktion von Filter Bsp.4	104
8.18	Irrelevanzcodierung	108

Tabellenverzeichnis

1.1	Zuordnung von Frequenzwerten zu musikalischen Tonbezeichnungen	3
1.2	Frequenzverhältnisse reiner Intervalle	7
3.1	Beispiele für Pegel und Bezugswerte	23
5.1	Lokalisationsschärfe bei horizontaler Verschiebung der Schallquelle	48
8.1	Binärwerte mit Vorzeichen	88

Literaturverzeichnis

- [AS93] AHNERT, Wolfgang ; STEFFEN, Frank: *Beschallungstechnik*. S. Hirzel Verlag Stuttgart, 1993
- [Bla97] BLAUERT, Jens: *Spatial Hearing*. 3. MIT Press Cambridge, Massachusetts, 1997
- [BP99] BORÉ, Gerhardt ; PEUS, Stefan. *Mikrophone*. 1999
- [CH90] CREMER, Lothar ; HUBERT, Matthias: *Vorlesungen über Technische Akustik*. 4. Springer-Verlag Berlin, 1990
- [CM78] CREMER, Lothar ; MÜLLER, Helmut A.: *Die wissenschaftlichen Grundlagen der Raumakustik*. 2. S. Hirzel Verlag Stuttgart, 1978
- [Dic87] DICKREITER, Michael: *Handbuch der Tonstudioteknik*. 5. K.G.Saur München, 1987
- [DIN97] DIN (Hrsg.): *Akustik Grundbegriffe (DIN 1320)*. 1997
- [Ear96] EARGLE, John M.: *Handbook of Recording Engineering*. 3. Chapman & Hall, 1996
- [Gör94] GÖRNE, Tom: *Mikrophone in Theorie und Praxis*. Elektor, 1994
- [HM94] HECKL, Manfred (Hrsg.) ; MÜLLER, Helmut A. (Hrsg.): *Handbuch der Technischen Akustik*. 2. Springer-Verlag Berlin, 1994
- [IR94] ITU-R (Hrsg.): *Methods for the subjective Assessment of small Impairments in Audio Systems including Multichannel Sound Systems (BS.1116)*. 1994
- [Mey95] MEYER, Jürgen: *Akustik und musikalische Aufführungspraxis*. 3. Verlag Erwin Bochinsky Frankfurt a. M., 1995
- [Moo97] MOORE, Brian C.: *An Introduction to the Psychology of Hearing*. 4. Academic Press, 1997

- [Mös91] MÖSER, Michael: *Unterlagen zu der Vorlesung Technische Akustik*. TU Berlin, 1991
- [Nol96] NOLL, Peter: *Unterlagen zu den Vorlesungen Nachrichtenübertragung 1 & 2*. TU Berlin, 1996
- [Nol97] NOLL, Peter: *Unterlagen zu der Vorlesung Statistische Nachrichtentheorie*. TU Berlin, 1997
- [Poh00] POHLMANN, Ken C.: *Principles of Digital Audio*. 4. McGraw-Hill, 2000
- [San85] SANK, Jon R.: Microphones. In: *Journal of the Audio Engineering Society* Bd. 33. 1985
- [Sen98] SENGPIEL, Eberhard: *Unterlagen zur Vorlesung Musikübertragung*. HdK Berlin, 1998
- [Ste55] STEVENS, S. S.: The Measurement of Loudness. In: *Journal of the Acoustical Society of America* Bd. 27. 1955
- [VM75] VERSCHUURE, J. ; MEETEREN, A. A.: The Effect of Intensity on Pitch. In: *Acustica* 32 (1975)
- [Wat94] WATKINSON, John: *An Introduction to Digital Audio*. Butterworth-Heinemann, 1994
- [Web94] WEBERS, Johannes: *Handbuch der Tonstudioteknik*. 6. Franzis-Verlag GmbH, 1994
- [Wut98] WUTTKE, Jörg. Mikrofonaufsätze. 1998
- [ZF99] ZWICKER, Eberhard ; FASTL, Hugo: *Psychoacoustics. Facts and Models*. 2. Springer Verlag Berlin, 1999
- [Zöl02] ZÖLZER, Udo (Hrsg.): *DAFX - Digital Audio Effects*. Wiley, 2002

Index

- Übersprechdämpfung, 72
Übertragungsfrequenzgang, 60, 61, 69, 76, 128
- Absorption, 33
Absorptionsfläche, 128
Absorptionsgrad, 18, 128
Abstandsfaktor, *siehe* Vergrößerungsfaktor
Abtastfrequenz, *siehe* Abtastrate
Abtastrate, 89, 112, 128
Abtasttheorem, 91, 112
Abtastung, 89
ADV, *siehe* Amplitudendichteverteilung
Aliasing, 91
Amplitude, 4
Amplitudendichteverteilung, 57, 97
Antialiasingfilter, 91
Arbeitsprinzip, 79
Artefakte, 106
Attack Time, 68
Außenohr, 41
Audiocodierungsverfahren, 105
Audiokompressionsverfahren, *siehe* Audiocodierungsverfahren
Ausbreitungsgeschwindigkeit, 8
Ausgangspegel
 maximaler, 76
- Bündelungsgrad, 34, 77, 128
Bündelungsmaß, 77
Backward Masking, *siehe* Vorverdeckung
Bandbreite eines Filters, 64
Bandpaß-Filter, 65
Bandsperre-Filter, 65
Beugung, 19
- Bewertungsfilter, 70
Bezugsschalldruck, 22
binaural, 48
Bitrate, 106, 110
 Constant, *siehe* Constant Bitrate
 Variable, *siehe* Variable Bitrate
Braunes Rauschen, 62
- CBR, *siehe* Constant Bitrate
Chorus, 67
Codierung
 Irrelevanz-, *siehe* Irrelevanzcodierung
 Redundanz-, *siehe* Redundanzcodierung
Codierungsverfahren, 105
Compressor, 68
Constant Bitrate, 106
Convolution, *siehe* Faltung
- dB, *siehe* Dezibel
DC-Block-Filter, 65
De-Esser, 69
Delay, 66
Dezibel, 22, 23
 dB(A), 23, 70
 dB(B), 23
 dB(C), 23, 70
 dBFs, 23
 dBm, 23
 dBSIL, 23
 dBSPL, 23
 dBSVL, 23
 dBSWL, 23
 dBu, 23
 dBV, 23
Differenztonfaktor, 71

- Diffusschall, [33](#)
- Diffusschallfeld, [33](#)
- Direktschall, [33](#)
- Direktschallfeld, [33](#)
- Distance Factor, *siehe* Vergrößerungsfaktor
- Dither, [96](#), [99](#)
- Dithering, [112](#)
- DSF, *siehe* Vergrößerungsfaktor
- Dynamikbearbeitung, [68](#)
- Dynamikumfang, [70](#), [76](#)
- Dynamisches Mikrophon, [83](#)

- Echo, [66](#)
- Echschwelle, [49](#)
- Effektivwert, [13](#)
- Eigenstörspannung, [76](#)
- Einschwingvorgang, [8](#)
- Elongationsempfänger, [79](#), [81](#)
- Entzerrer, [63](#)
- Ersatzgeräuschpegel, [76](#)
- Expander, [68](#)

- Faltung, [61](#)
- Fehler
 - Quantisierungs-, *siehe* Quantisierungsfehler
- Feldübertragungsfaktor, [75](#)
- Fernfeld, [15](#)
- Festkommaformat, [88](#)
- Filter, [63](#)
 - Bandbreite, [128](#)
 - Bandpaß-, [65](#)
 - Bandsperre, [65](#)
 - Bewertungs-, [70](#)
 - DC-Block-, [65](#)
 - digitale, [100](#)
 - FIR, *siehe* FIR-Filter
 - Flankensteilheit, [64](#), [128](#)
 - Güte, [64](#), [129](#)
 - Gewichtungs-, [70](#)
 - Grenzfrequenz, [64](#), [128](#)
 - Hochpaß-, [64](#)
 - IIR, *siehe* IIR-Filter
 - Mittenfrequenz, [64](#), [128](#)
 - Notch-, [65](#)
 - Peak-, [65](#)
 - Präsenz-, [65](#)
 - Resonanz, [66](#)
 - Shelving-, [65](#)
 - Struktur, [105](#)
 - Tiefpaß-, [64](#)
 - Trittschall-, [65](#)
- Filterbank, [107](#)
- Filtergraph, [100](#)
- Filterkoeffizient, [100](#)
- Filterstruktur, [105](#)
- Finite Impulse Response, [104](#)
- FIR, *siehe* Finite Impulse Response
- FIR-Filter, [104](#)
- Fixed-Point-Format, *siehe* Festkommaformat
- Flanger, [67](#)
- Flankensteilheit, [64](#)
- Flatterecho, [36](#)
- Fließkommaformat, [88](#)
- Floating-Point-Format, *siehe* Fließkommaformat
- Formant, [7](#)
 - bereich, [7](#)
- Forward Masking, *siehe* Nachverdeckung
- Fouriertransformation, [55](#)
- Freifeld, [33](#)
- Freifeldfrequenzgang, [76](#)
- Frequenz, [2](#), [128](#)
- Frequenzdarstellung, [55](#)
- Full Scale, [23](#)
- Funkhausnormpegel, [23](#)

- Güte, [64](#)
- Geräuschpegelabstand, [76](#)
- Geschwindigkeitsempfänger, [80](#), [84](#)
- Gesetz
 - 1/r-, [15](#)
 - der ersten Wellenfront, [49](#)
 - Weber-Fechnersches, [21](#)

- GewichtungsfILTER, 70
Gleitkommaformat, *siehe* Fließkommaformat
Grenzfrequenz, 64
Grenzschalldruckpegel, 76

Häufigkeitsverteilung, 57
Hörfläche, 23
Hüllkurve, 8
Hall, 67
Hallabstand, 36
Hallgerät, 67
Hallradius, 34, 129
 effektiver, 36
Hochpaß-Filter, 64

IIR, *siehe* Infinite Impulse Response
IIR-Filter, 104
Impulsantwort, 60, 61, 128
Impulsverhalten, 79
Infinite Impulse Response, 104
Inkohärenz, 25
Innenohr, 43
Interferenz, 5
Intermodulationsfaktor, 71
Irrelevanzcodierung, 107

Klang, 6
Klirrdämpfungsmaß, 71
Klirrfaktor, 70, 128
Kohärenz, 25
Kompressionsrate, 106
Kondensatormikrophon, 80
Korrelation, 25
Korrelationsgrad, 25
Kurven gleicher Lautstärkepegel, 45

Lautheit, 46
Least Significant Bit, 87
Leistungspegel, 22, 23
Limiter, 68
Linearität, 60
Lokalisationsschärfe, 48
LSB, *siehe* Least Significant Bit

Maskierung
 Nach-, *siehe* Nachverdeckung
 Simultan-, *siehe* Simultanmaskierung
 Vor-, *siehe* Vorverdeckung
Mikrophon
 Dynamisches, *siehe* Dynamisches Mikrophon
 Kondensator-, *siehe* Kondensatormikrophon
Mittelohr, 42
Mittenfrequenz, 64
Modell
 psychoakustisches, 107
Most Significant Bit, 87
MSB, *siehe* Most Significant Bit

Nachhalldauer, 32
Nachhallzeit, 32, 36, 67, 129
Nachverdeckung, 47
Nahbesprechungseffekt, 17
Nahfeld, 15
Noise Gate, 68
Noise-Shaping, 98, 112
Notch-Filter, 65

Oberschwingung, 7
Oberton, 6
Oversampling, 95

PDF, *siehe* Amplitudendichteverteilung
Peak-Filter, 65
Pegel, 21
Periodendauer, 3, 129
Phantomschallquelle, 52
Phantomspeisung, 82
Phasendifferenz, 4
Phasenverschiebung, 4
Phasenwinkel, 4, 128
Phaser, 66, 67
phon, 45
Postmasking, *siehe* Nachverdeckung
Präsenz-Filter, 65
Pre-Delay, 67
Premasking, *siehe* Vorverdeckung

- Pressure Zone Mikrophone, *siehe* Grenzflächenmikrophon
- Prezedenzeffekt, [49](#)
- Probability Density Function, *siehe* Amplitudendichteverteilung
- Pythagoras, Satz des, [119](#)
- PZM, *siehe* Grenzflächenmikrophon
- Quantisierer, [108](#)
- Quantisierung, [92](#), [112](#)
- Quantisierungsfehler, [93](#), [129](#)
- Quantisierungsrauschen, [93](#)
- Random Energy Efficiency, [78](#)
- Raummode, [37](#)
- Rauschen
 - Braunes, *siehe* Braunes Rauschen
 - dreieckförmig verteiltes, [98](#)
 - gleichverteiltes, [98](#)
 - Quantisierungs-, *siehe* Quantisierungsrauschen
 - Rosa , *siehe* Rosa Rauschen
 - Weißes, *siehe* Weißes Rauschen
- Redundanzcodierung, [106](#)
- REE, *siehe* Random Energy Efficiency
- Rekonstruktionsfilter, [92](#)
- Release Time, [68](#)
- Resonanz-Filter, [66](#)
- Richtcharakteristik, [77](#)
- Richtentfernung, [36](#)
- richtungsbestimmende Bänder, [49](#)
- Richtungsfaktor, [77](#), [128](#)
- Richtungsmaß, [77](#)
- RMS, *siehe* Effektivwert
- Root Mean Square, *siehe* Effektivwert
- Rosa Rauschen, [62](#)
- Ruhehörschwelle, [23](#)
- Sample Rate, *siehe* Abtastrate
- Sampling, *siehe* Abtastung
- Schalldämmung, [18](#)
- Schalldämpfung, [18](#)
- Schalldruck, [13](#), [128](#), [129](#)
- Schalldruckpegel, [22](#), [23](#), [128](#)
- bewerteter, [23](#)
- Schalleistung, [17](#), [129](#)
- Schalleistungspegel, [22](#), [23](#), [128](#)
- Schallgeschwindigkeit, [8](#), [128](#)
- Schallintensität, [17](#), [128](#)
- Schallintensitätspegel, [23](#)
- Schallschnelle, [13](#), [129](#)
- Schallschnellepegel, [23](#)
- Schmerzgrenze, [23](#)
- Schwebung, [5](#)
- Schwingungen, [1](#)
- Schwingungsbauch, [2](#)
- Schwingungsknoten, [2](#)
- Shelving-Filter, [65](#)
- Signal-Rauschabstand, [70](#), [95](#)
- Signal-to-Noise Ratio, *siehe* Signal-Rauschabstand
- SNR, *siehe* Signal-Rauschabstand
- sone, [46](#)
- Spannungspegel, [23](#), [128](#)
- Spektrogramm, [57](#)
- Spektrum, [55](#)
- System, [60](#)
 - lineares, [60](#)
 - nichtlineares, [61](#)
- THD, *siehe* Klirrfaktor
- THD+N, *siehe* Total Harmonic Distortion and Noise
- Tiefpaß-Filter, [64](#)
- Ton, [6](#)
- Total Harmonic Distortion (THD), *siehe* Klirrfaktor
- Total Harmonic Distortion and Noise, [71](#)
- Trading, [50](#)
- Transferfunktion, [60](#), [61](#), [128](#)
- Trittschallfilter, [65](#)
- Unterabtastung, [91](#)
- Variable Bitrate, [106](#), [110](#)
- VBR, *siehe* Variable Bitrate
- Verdeckungseffekte, [46](#)
- Vergrößerungsfaktor, [78](#)

- Verzerrung, [60](#)
- Vorverdeckung, [47](#)
- weißes Rauschen, [62](#)
- Welle, [8](#)
 - ebene, [10](#)
 - Kugel-, [10](#)
 - stehende, [37](#)
- Wellenlänge, [9](#), [128](#)
- Wortbreite, [87](#), [92](#), [112](#)
- Zahlendarstellung
 - binäre, [87](#)
- Zeitdarstellung, [55](#)
- Zeitvarianz, [60](#)