

„Digital Basics“

Teil 3 des Skriptums für die VO „Studiogerätekunde“

Bachelorarbeit

durchgeführt von

Matti Kruse

Institut für Signalverarbeitung und Sprachkommunikation
der Technischen Universität Graz

Leiter: Univ.-Prof. Dipl.-Ing. Dr.techn. Gernot Kubin

Betreuer/in: Thorsten Rohde

Graz, im Jänner 2011

Zusammenfassung

Die Bachelor-Arbeit „Digital Basics - Teil 3 des Skriptums für die Vorlesung Studiogerätekunde“ wurde innerhalb eines Projektes erarbeitet, in dem ein neues Skriptum für die Vorlesung „Studiogerätekunde“ an der TU Graz erstellt wurde. Unter der Obhut der Projektleiterin Annika Neidhardt wurde das Skriptum in drei Abschnitte eingeteilt. Der hier vorliegende Teil beschäftigt sich mit den Grundlagen digitaler Audiotechnik. In der gesamten Arbeit werden die Inhalte, unterstützt von einfachen Rechenbeispielen und Abbildungen, praxisnah und verständlich dargestellt. Unter Berücksichtigung der Anforderungen, die an ein Skriptum für Studienanfänger gestellt werden, sowie der Inhalte der Vorlesung, haben sich vier Hauptabschnitte ergeben.

Der erste Teil beschäftigt sich mit dem Verhalten von Schwingungen auf Verbindungsleitungen in der analogen und digitalen Audiotechnik. Im zweiten Teil werden Grundlagen der Analog-Digital- und der Digital-Analog Umsetzung vorgestellt, wobei wichtige Begriffe erläutert werden. Teil drei enthält einen Überblick über gängige Schnittstellen in der digitalen Audiotechnik. Hier finden sich Informationen über die Arten der Datenübertragung und die Verwendung der richtigen Kabel und Stecker. Des Weiteren werden einige Gerätetypen aus der digitalen Studiotechnik vorgestellt. Der vierte Teil beschäftigt sich mit der Speicherung digitaler Audiodaten. Am Beispiel der Audio-CD wird eine Möglichkeit der Speicherung beschrieben und es werden verschiedene Speicherformate vorgestellt.

Abstract

The Bachelor-thesis "Digital Basics - Part 3 of the script for the lecture Studiogerätekunde" has been developed within a project, in which a new script for the lecture "Studiogerätekunde" at the Technical University of Graz has been developed. Under the supervision of project manager Annika Neidhardt, the script was divided into three sections. The present part is concerned with the basics of digital audio technology. Throughout the paper the content is presented in a practical and understandable manner, supported by simple arithmetic examples and illustrations. Taking into account the script demands for new students, as well as the contents of the lecture, four main sections emerged.

The first part deals with the behavior of oscillations of interconnections in analog and digital audio technology. In the second part the basics of analog-to-digital and digital-to-analog conversion are presented, and important terms are explained. Part three provides an overview of common interfaces in digital audio technology. This section contains information about the different types of data communication and about the use of suitable cables and connectors. Furthermore, several types of devices from digital studio technology are presented. The fourth part deals with the storage of digital audio data. With the example of the audio CD a way of storing it is described and various storage formats are introduced.

Inhaltsverzeichnis

Inhaltsverzeichnis	3
1 Einleitung	5
2 Verbindungsleitungen in der Analog- und Digitaltechnik	6
2.1 Ausbreitung von elektromagnetischen Wellen auf Leitungen	6
2.1.1 Die Schwingung	6
2.1.2 Die Welle.....	7
2.2 Wellenlängen auf Leitungen	8
2.3 Reflexionen in Leitungen	9
2.4 Folgerungen für die Digitaltechnik	14
2.5 Folgerungen für die Analogtechnik.....	15
3 AD- und DA-Umsetzung	17
3.1 Sampling	18
3.2 Quantisierung	22
3.3 Dithering	24
3.4 AD-Wandlerverfahren.....	25
3.5 DA-Wandlerverfahren.....	29
3.6 Zusammenfassung DA/AD-Wandlung	32
4 Geräte und Schnittstellen	34
4.1 Digitale Schnittstellen	34
4.1.1 AES3	35
4.1.2 AES3id und AES 42.....	36
4.1.3 S/PDIF	37
4.1.4 SDIF	38
4.1.5 MADI	38
4.1.6 ADAT	40
4.1.7 TDIF-1 und TDIF-2.....	40
4.1.8 Weitere digitale Schnittstellen.....	41
4.2 Audio über Bussysteme.....	43
4.2.1 Protokolle und Routing	43
4.2.2 Bussysteme	43

4.2.3 IEEE 1394 (Firewire)	44
4.2.4 USB	45
4.2.5 HDMI	46
4.3 Netzwerk Audio	46
4.3.1 Das ISO-OSI Schichtenmodell	47
4.3.2 Ethernet	48
4.3.3 AudioRail	49
4.3.4 CobraNet	49
4.3.5 EtherSound	50
4.4 Prozessoren	51
4.5 Digitale Pegel	53
4.6 Mobile Aufnahmegeräte und Peripheriegeräte mit Beispielen	55
4.6.1 Soundkarten	55
4.6.2 Mobile Aufnahmegeräte	57
4.6.3 Digitale Effektgeräte	61
5 Digitale Speicherung	63
5.1 Die Audio-CD	63
5.2 Digitale Speicherformate	65
Literaturverzeichnis	68

1 Einleitung

Die Digitaltechnik hat in der heutigen Tonstudioteknik, wie in fast allen Lebens- und Arbeitsbereichen, einen sehr großen Stellenwert. Oft wird das aufzunehmende akustische Signal direkt von einem analogen Signal in ein digitales überführt. Die weitere Bearbeitung findet dann auf digital verarbeiteten Komponenten, wie z. B. dem Computer, statt. Abgesehen davon werden viele Klänge, z. B. von Software-Synthesizern, digital erzeugt.

In diesem Kapitel wird beschrieben, wie sich das analoge Audiosignal in analogen Leitungen verhält, wie die Umsetzung in ein digitales Signal funktioniert und es werden gebräuchliche digitale Übertragungswege vorgestellt. Darüber hinaus werden Hinweise auf die Verwendung geeigneter Kabel und Stecker gegeben und einige Geräte aus dem Bereich der digitalen Studioteknik vorgestellt.

Es werden die Grundlagen der digitalen Audiotechnik vermittelt. Die Vorlesung „Digitale Audiotechnik 1“ von Ao. Univ. Prof. Dr. Gerhard Graber beschäftigt sich vertiefend mit der Thematik und greift weitere wichtige Punkte auf.

Empfehlenswerte Literatur für Interessierte ist das Buch „DAFX“ bzw. „Digitale Audiosignalverarbeitung“ von Udo Zölzer, sowie die Kapitel 14, 15, 17 und 18 aus dem Buch „Handbuch der Tonstudioteknik“, herausgegeben von Stefan Weinzierl und als eBook bei der Bibliothek der TU Graz einsehbar.

2 Verbindungsleitungen in der Analog- und Digitaltechnik

Der grundlegende Unterschied zwischen dem analogen und dem digitalen Signal besteht darin, dass das analoge Signal kontinuierlich verläuft und sich stetig ändert. Das digitale Signal besteht aus zeitdiskreten Abtastwerten, die binär dargestellt werden. Zunächst wird das Verhalten eines Signals in Leitungen besprochen.

2.1 Ausbreitung von elektromagnetischen Wellen auf Leitungen

Der durch den Schall entstehende Wechseldruck (vgl. Kapitel Mikrofone), der an der Membran eines Mikrofons herrscht, wird in elektrische Spannungsänderungen umgewandelt. Dieses elektrische Signal wird für die weitere Verwendung über Leitungen transportiert.

2.1.1 Die Schwingung

Wiederholt sich ein Vorgang innerhalb eines gewissen Zeitabschnittes, so wird dieser allgemein als Schwingung bezeichnet. In unserem Fall ist es eine sich ändernde elektrische Spannung.

Es wird zwischen periodischen und nichtperiodischen Schwingungen unterschieden. Periodische Schwingungen wiederholen sich in regelmäßigen Intervallen, nichtperiodische Schwingungen weisen unregelmäßige Zustandsänderungen auf.

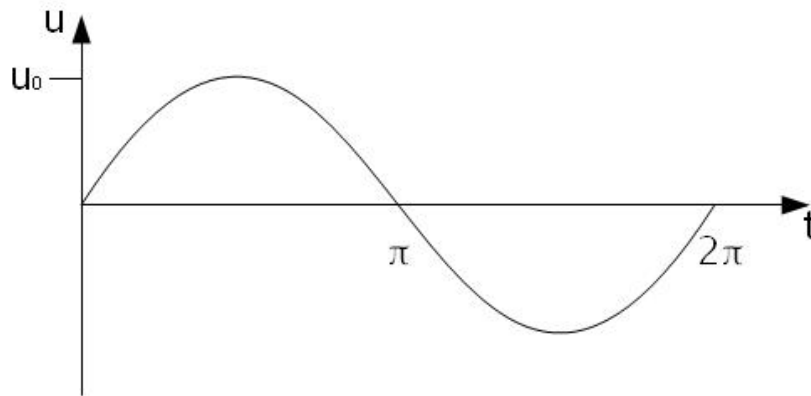


Abb 2-1: Schwingung

Eine Schwingung lässt sich beschreiben durch

$$u = u_0 \cos(\omega t + \varphi)$$

mit

u_0 ... Amplitude der Schwingung

ω ... Kreisfrequenz $\omega = 2\pi f$ mit der Frequenz $f = \frac{1}{T}$, wobei T die Dauer einer Periode darstellt.

φ ... Anfangsphase der Schwingung, bzw. die Verschiebung der Cosinusschwingung auf der t-Achse

Bei einer Schwingung werden die Auslenkung und die zeitliche Komponente eines Vorgangs betrachtet. Wird der örtliche Aspekt mit einbezogen, so kann ein Vorgang mittels einer Welle beschrieben werden.

2.1.2 Die Welle

Schwingt ein Molekül innerhalb eines Mediums (z. B. Luft, Wasser oder Metall), regt es auch seine Nachbarmoleküle zum Schwingen an. Auf diese Weise breitet sich die Schwingung innerhalb eines Mediums als Welle aus.

Wird eine Spannung auf eine Leitung induziert, fließt elektrischer Strom, wodurch sich um die Leitung ein magnetisches Feld bildet, welches sich in gleicher Weise ändert wie der Wechselstrom, bzw. die Wechselspannung. Bei dieser zeitlichen Änderung des magnetischen Feldes entsteht ein elektrisches Feld. Diese zeitlich schwankenden Felder haben immer elektrische und magnetische Komponenten. Sie bilden ein elektromagnetisches Feld. Das elektromagnetische Feld breitet sich in Form von elektromagnetischen Wellen im Raum aus.

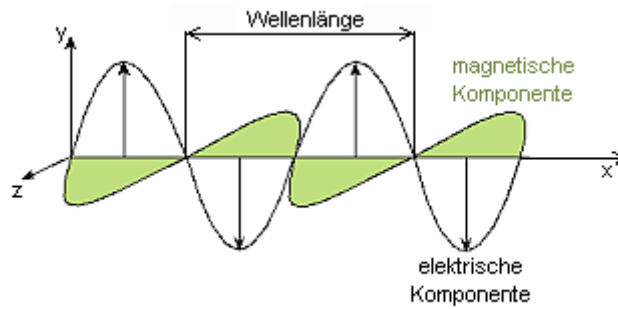


Abb 2-2: Ausbreitung elektromagnetischer Wellen [<http://bbs.saarburg.bildung-rp.de/emv/strahl~1.htm>]

Die allgemeine Beschreibung einer Welle findet sich in der Wellengleichung wieder. Ihre Lösung wird an anderer Stelle besprochen (z. B. in der Vorlesung „Raumakustik“). Für das grundlegende Verständnis wird die Betrachtung der Wellenfunktion aus der Quantenmechanik herangezogen.

$$u = u_0 \cos(\omega t + kx + \varphi)$$

Diese Gleichung ähnelt jener einer Schwingung sehr. x ist der Ortsvektor und die Kreiswellenzahl k hängt auf einfache Weise mit der Wellenlänge $\lambda = \frac{c}{f}$, der Kreisfrequenz ω und der Lichtgeschwindigkeit c zusammen:

$$k = \frac{2\pi}{\lambda} = \frac{\omega}{c}$$

2.2 Wellenlängen auf Leitungen

Elektrische Leitungen sind komplexe Widerstandsnetzwerke mit stark frequenzabhängigem Verhalten.

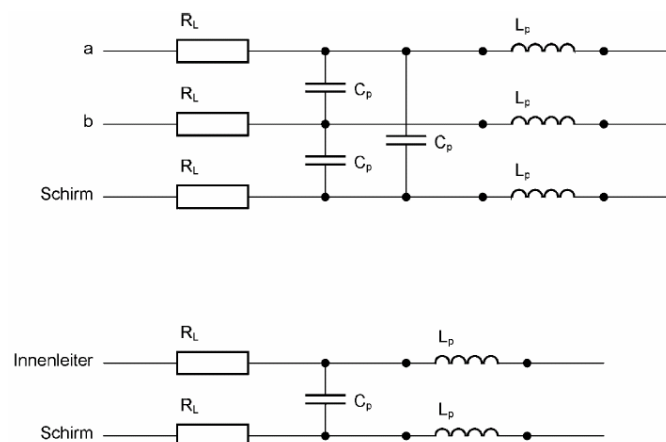


Abb 2-3: Ersatzschaltbild einer 2-poligen, abgeschirmten Leitung (oben) und einer koaxialen Leitung (unten) [WEINZIERL]

Einen deutlichen Unterschied zwischen der Übertragung von Audiosignalen in der Analog- und Digitaltechnik erkennt man, wenn man die Werte der zu übertragenden Frequenzen betrachtet. Analoge Audioleitungen transportieren Frequenzen bis etwa 20 kHz, digitale Audioleitungen noch weitaus höhere. So übertragen einige zweikanalige digitale Schnittstellen (Kapitel 4.1) Datenraten von ca. 6 MBit/s.

Für analoge Leitungen bedeutet das, dass bei einer Frequenz von 20 kHz Wellenlängen von 15 km auftreten.

$$\lambda = \frac{c}{f} = \frac{300 * 10^6 \frac{m}{s}}{20 * 10^3 \frac{1}{s}} = 15km$$

Wenn man die 6 MHz in einem digitalen Audio-Kabel exemplarisch betrachtet, treten Wellenlängen von nur 50 m auf.

$$\lambda = \frac{300 * 10^6 \frac{m}{s}}{6 * 10^6 \frac{1}{s}} = 50m$$

In diesen Fällen wären bei üblichen Leitungslängen von bis zu 100 m digital bereits zwei volle Wellenlängen sichtbar, während in analogen Leitungen nur eine Schwingung auftritt, das bedeutet, dass entlang der gesamten Leitung die gleiche Phase vorhanden ist. Die geringen Phasenänderungen über die - relativ zur Wellenlänge - kurze Strecke sind zu vernachlässigen.

2.3 Reflexionen in Leitungen

Die Ausbreitung elektromagnetischer Wellen in Leitungen erfolgt transversal, das heißt, dass die eine Schwingung senkrecht zur anderen erfolgt. Wobei der Strom (verantwortlich für das magnetische Feld) und die Spannung (verantwortlich für das elektrische Feld) zueinander in einem festen Verhältnis stehen. Dieses Verhältnis wird durch den Wellenwiderstand Z_0 beschrieben. Bei verlustlosen Leitungen gilt für die zur Senke hinlaufende Welle (Index h)

$$\frac{u_h}{i_h} = +Z_0$$

und für die rücklaufende Welle (Index r)

$$\frac{u_r}{i_r} = -Z_0.$$

Im Idealfall wird die gesamte von der Quelle gelieferte Energie vom Leitungsabschluss (der Senke) aufgenommen. Das negative Vorzeichen bei der Gleichung für die rücklaufende Welle indiziert jedoch, dass im Reflexionsfall Energie von der Senke zur Quelle zurückfließen kann. Solche Reflexionen in Leitungen entstehen dann, wenn die Impedanz des

Leistungsabschlusses (also der Senke) vom Wellenwiderstand der Leitung und/oder von der Impedanz der Quelle abweicht und ein Teil der Energie reflektiert wird. Auch Stoßstellen innerhalb von Leitungen, wie Steckverbinder, Kupplungen oder stark geknickte und deformierte Kabel, können dies verursachen. Die von der „falschen“ Abschlussimpedanz reflektierte Welle überlagert das Originalsignal und kann im Extremfall zur Verdopplung oder völligen Auslöschung des Signals führen. Der Reflexionsfaktor r ist das Amplitudenverhältnis zwischen der einfallenden und der reflektierten Welle und ergibt sich aus

$$\underline{r} = \frac{\underline{Z}_a - \underline{Z}_0}{\underline{Z}_a + \underline{Z}_0}, \quad |r| \leq 1,$$

mit

\underline{Z}_a ... Außenwiderstand des Verbrauchers

\underline{Z}_0 ... Wellenwiderstand der Leitung

Ist $r = 0$ kommt es zu keiner Reflexion. Bei $r = 1$ kommt es zur Vollständigen Reflexion der Welle.

Rechenbeispiel zum Reflexionsverhalten auf Leitungen in der Analogtechnik

Hierzu ein Rechenbeispiel, welches verdeutlicht, was an einer 12,5m bzw. 50m langen Leitung bei der Übertragung von 20 kHz passiert. Der Wellenwiderstand ist gleich dem Quellwiderstand des Senders (10 Ω). Der Lastwiderstand beträgt 47 kΩ. Berechnet wird der Reflexionsfaktor r , die Wellenzahl k und die Wellenlänge λ .

$$R_Q = 10\Omega, R_L = 47k\Omega, f = 20kHz$$

$$l_1 = 12,5m, l_2 = 50m$$

$$r = \frac{Z_L - Z_Q}{Z_L + Z_Q} = \frac{46990\Omega}{47010\Omega} = 0,99957 \approx 1$$

$$k = \frac{\omega}{c} = \frac{2\pi 20000 \frac{1}{s}}{300000000 \frac{m}{s}} = 0,0004 \frac{1}{m}$$

$$\lambda = \frac{c}{f} = \frac{300000000 \frac{m}{s}}{20000 \frac{1}{s}} = 15000m$$

Hier erkennt man, dass auf der gesamten Leitung nur eine Schwingung sichtbar ist ($\lambda \gg l$).

Nach unserer Gleichung für eine Welle wird zunächst für den Fall $l_1=12,5m$ die Spannung an der Quelle berechnet; zunächst für die hinlaufende Welle ($u_h, x=0$), dann für die rücklaufende Welle ($u_r, x = 2 \cdot 12,5 = 25$).

$$u_h = u_0 \cos(\omega t + kx), x = 0$$

$$u_h = u_0 \cos(\omega t)$$

$$u_r = u_0 \cos(\omega t + kx), x = 25$$

$$u_r = u_0 \cos(\omega t + 0,0004 \cdot 25) = u_0 \cos(\omega t + 0,01)$$

Zwischenrechnung Additionstheorem:

$$\cos(x + \varepsilon) = \cos x \cdot \cos \varepsilon - \sin x \cdot \sin \varepsilon$$

für $\varepsilon \ll 1$ gilt $\cos \varepsilon \cong 1$; $\sin \varepsilon \cong \varepsilon$

$$\cos(x + \varepsilon) = \cos x - \varepsilon \cdot \sin x \cong \cos x$$

$$u_r = u_0 \cos(\omega t)$$

$$u_h + u_r = u_0(\cos(\omega t) + \cos(\omega t)) = 2u_0 \cos(\omega t)$$

Bei längeren Leitungslängen verändert sich nur der Wert von kx . Dieser bleibt aber selbst bei einer Kabellänge von 50 m zu gering, um das Ergebnis signifikant zu verändern.

$$u_r = u_0 \cos(\omega t + 0,0004 \cdot 100) = u_0 \cos(\omega t + 0,04)$$

Bei derartig großen Wellenlängen in Bezug auf die Leitungslänge kommt es bei einer Reflexion also stets zu einer Verdopplung des Signals.

Rechenbeispiel zum Reflexionsverhalten auf Leitungen in der Digitaltechnik

Als nächstes werden die Werte für eine Frequenz, die bei der digitalen Signalübertragung vorkommt, von 6 MHz berechnet.

$$R_Q = 10\Omega, R_L = 47k\Omega, f = 6MHz$$

$$l_1 = 12,5m, l_2 = 50m$$

$$r = \frac{Z_L - Z_Q}{Z_L + Z_Q} = \frac{46990\Omega}{47010\Omega} = 0,99957 \approx 1$$

$$k = \frac{\omega}{c} = \frac{2\pi 6000000 \frac{1}{s}}{300000000 \frac{m}{s}} = 0,125 \frac{1}{m}$$

$$\lambda = \frac{c}{f} = \frac{300000000 \frac{m}{s}}{6000000 \frac{1}{s}} = 50m$$

$l_1=12,5$ m, Spannung für den Ort $x = 0$, also an der Quelle

hinlaufende Welle:

$$u_h = u_0 \cos(\omega t + kx), x = 0$$

$$u_h = u_0 \cos(\omega t)$$

rücklaufende Quelle:

$$u_r = u_0 \cos(\omega t + kx), x = 25$$

$$u_r = u_0 \cos(\omega t + 0,125 \cdot 25) = u_0 \cos(\omega t + 3,125) \approx -u_0 \cos(\omega t)$$

Überlagerung von hinlaufender und rücklaufender Welle:

$$u_h + u_r = u_0 (\cos(\omega t) - \cos(\omega t)) = 0$$

Das bedeutet, dass die reflektierte Welle am Senderausgang die umgekehrte Phase zur gesendeten Welle aufweist und sich diese beiden Wellen daher aufheben (destruktive Interferenz).

Nun die Berechnung für den Fall $l_2=50$ m,

hinlaufende Welle:

$$u_h = u_0 \cos(\omega t),$$

rücklaufende Quelle:

$$u_r = u_0 \cos(\omega t + kx), x = 100$$

$$u_r = u_0 \cos(\omega t + 0,125 \cdot 100) = u_0 \cos(\omega t + 12,5) \approx u_0 \cos(\omega t)$$

Überlagerung von hinlaufender und rücklaufender Welle:

$$u_h + u_r = u_0 (\cos(\omega t) + \cos(\omega t)) = 2u_0 \cos(\omega t)$$

Hier wird die Welle am Verbraucher gleichphasig zur gesendeten Welle reflektiert, das heißt, die Wellen addieren sich (konstruktive Interferenz).

In der Abbildung auf der nächsten Seite werden die letzten beiden Berechnungen noch einmal grafisch dargestellt. a.a) bildet die hinlaufende Welle in einem $\lambda/4$ langen Kabel ab; a.b) die rücklaufende Welle und a.c) die Addition der Wellen. Analog dazu bildet b.a) die hinlaufende Welle in einem λ langen Kabel ab; b.b) die rücklaufende Welle und b.c) die Addition beider Wellen.

Je nach Kabellänge kommt es also vor, dass die reflektierte Welle zur Auslöschung oder zur Verdopplung der gesendeten Welle führt. Dieser Effekt ist natürlich nicht erwünscht, weshalb bei digitalen Audioverbindungen eine Leistungsanpassung (Kapitel 2.4) zum Einsatz kommt.

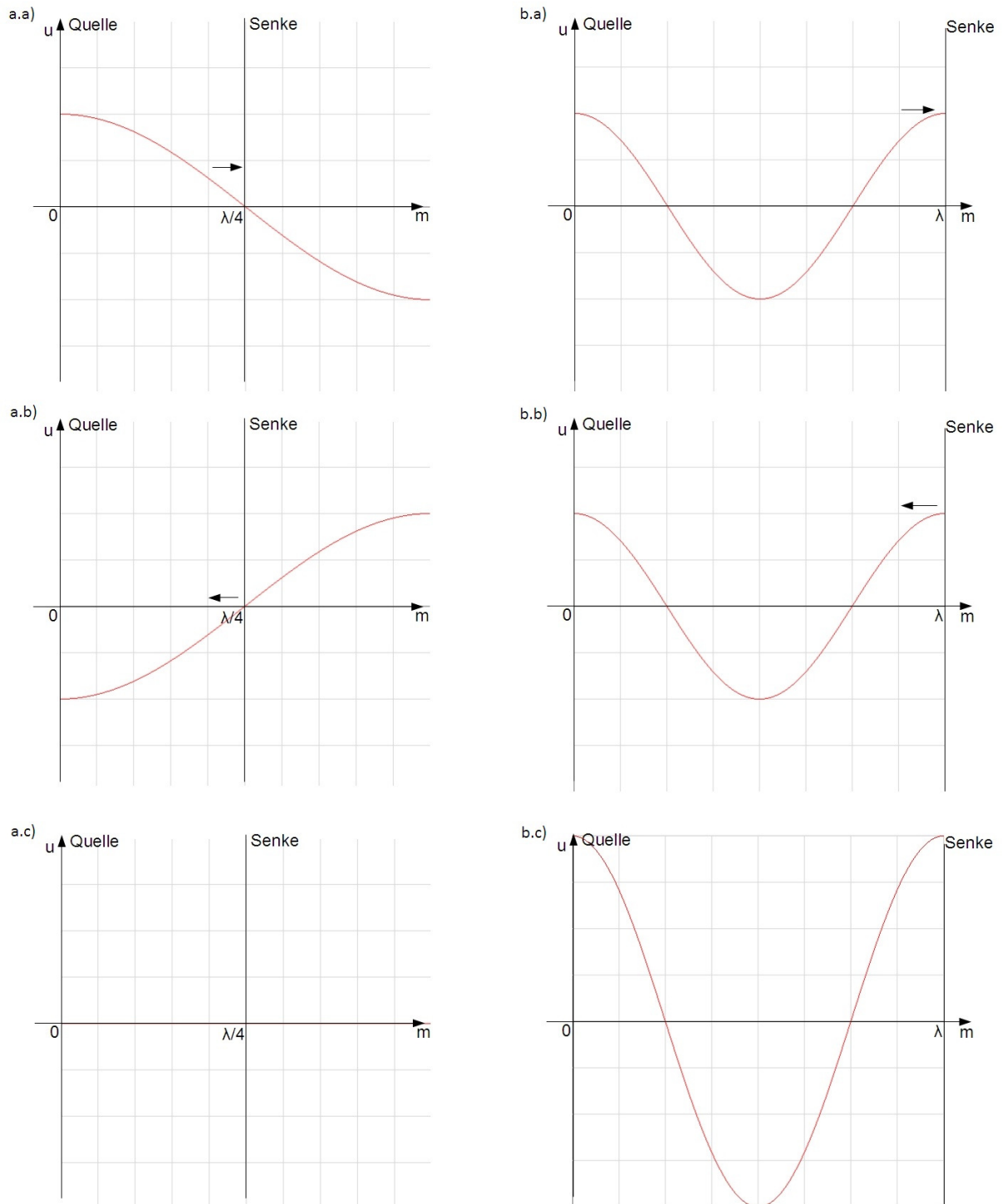


Abb 2-4: links: Kabellänge $\lambda/4$, a.a) hinlaufende Welle, a.b) an der Senke reflektierte rücklaufende Welle, a.c) addition der Wellen aus a.a) und a.b); rechts: Kabellänge λ , b.a) hinlaufende Welle, b.b) an der Senke reflektierte rücklaufende Welle, b.c) Addition der Wellen aus b.a) und b.b)

2.4 Folgerungen für die Digitaltechnik

Leistungsanpassung

Überall, wo Reflexionen auf Leitungen unbedingt vermieden werden müssen, kommt eine Leistungsanpassung zum Einsatz mit

$$R_i = R_a = Z_0,$$

da nur hier der Reflexionsfaktor $r=0$ ist und es zu keiner Reflexion kommt. Daher sind praktisch alle digitalen Audiogeräte mit Schnittstellen mit Leistungsanpassung ausgestattet. Gleiches gilt für die Antennenverbindungen von Funkmikrofonempfängern oder Sendern für In-Ear-Monitoring - also überall dort, wo hohe Frequenzen und damit kurze Wellenlängen, bzw. Wellenlängen im Bereich der Leitungslängen und kleiner auftreten.

Hier kann die maximale Leistung in den Verbraucher gebracht werden, wobei die eine Hälfte der Leistung im Quellwiderstand und die andere Hälfte der Leistung im Lastwiderstand verbraucht wird.

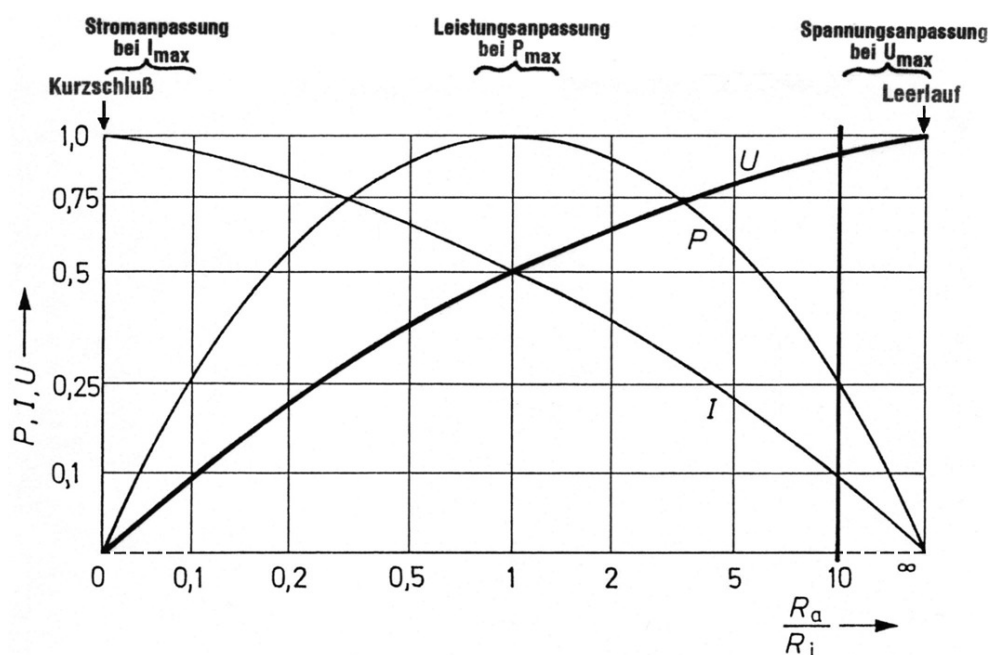


Abb 2-5: Strom I , Leistung P und Spannung U in Abhängigkeit vom Widerstandsverhältnis R_a/R_i [SENGPIEL1]

Es muss bei der Übertragung digital kodierter Audiosignale mit einer Leistungsanpassung gearbeitet werden. Daher ist es notwendig, einen Wellenwiderstand zu definieren: 110Ω für das AES3-Format (siehe Kapitel 4.1.1) für verdrehte Doppeladerkabel und 75Ω für Koaxialkabel. Der letztgenannte Wert ist aus der Videotechnik entlehnt, die häufig gemeinsam mit Audiotechnik verwendet wird. Dort war eine Definition eines Wellenwiderstandes von Anfang an notwendig, da die analoge Videotechnik bereits Frequenzen bis in den MHz-Bereich umfasst.

Des Weiteren folgt aus der Leistungsanpassung, dass immer nur ein Verbraucher an eine Quelle angeschlossen werden kann, da beim Anschluss mehrerer Geräte eine

Leistungsanpassung nicht mehr gegeben wäre. Für die Verteilung von digitalen Signalen benötigt man daher eigene Geräte (Splitter)

2.5 Folgerungen für die Analogtechnik

Spannungsanpassung

Wenn der Innenwiderstand R_i einer Quelle (z. B. eines Mikrofons) deutlich kleiner ist als der Abschlusswiderstand R_a der darauf folgenden Senke (z. B. Mischpulteingang), d. h. für

$$R_i \ll R_a$$

spricht man von einer Spannungsanpassung. Das Anpassungsverhältnis liegt in der analogen Audiotechnik üblicherweise zwischen 1:10 und 1:1000. Es bewirkt, dass die Quelle durch den Verbraucher kaum belastet wird, da wenig Strom fließt und die Ausgangsspannung vollständig am angeschlossenen Gerät abfällt. In der modernen analogen Audiotechnik kommt praktisch immer eine Spannungsanpassung zum Einsatz.

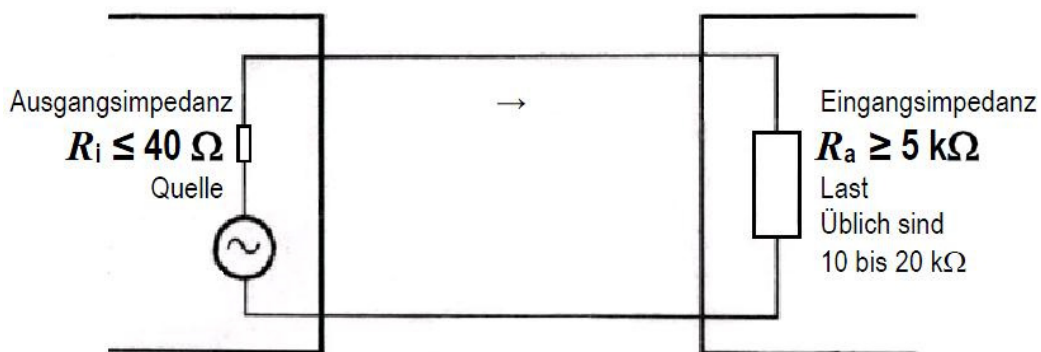


Abb 2-6: Geräteverbindung mit Ausgang und Eingang [SENGPIEL2]

Die unterschiedlichen Benennungsmöglichkeiten für die Widerstände können leicht Verwirrung stiften! Daher ist folgende Tabelle unbedingt zu beachten.

R_i	R_a
Innenwiderstand	Lastwiderstand
Quellwiderstand	Außenwiderstand
Austangswiderstand	Eingangswiderstand
	Abschlusswiderstand

Üblicherweise liegen die Werte für den Quellwiderstand zwischen 10Ω und 200Ω , der Lastwiderstand wird entsprechend groß gewählt ($1 \text{ k}\Omega$ bis $20 \text{ k}\Omega$). Auf der Leitung ist nur eine Schwingung zu sehen, da die Wellenlänge λ viel größer ist, als die Leitungslänge. Man kann hier den Vorteil nutzen, an eine Quelle mehrere Verbraucher parallel anzuschließen, indem ein einfaches Y-Kabel verwendet wird, oder ein Patchfeld ein Signalausgang auf mehrere Verbraucher geschaltet wird (z. B. Summenausgang auf Monitore und Aufnahmegerät).

3 AD- und DA-Umsetzung

Das analoge Signal muss für eine digitale Verarbeitung interpretiert werden. Da es in der digitalen Audiotechnik keinen Parameter gibt, der sich dem Schalldruckverlauf, bzw. dem Verlauf der Wechselspannung entsprechend ändert, wird das Signal in eine Abfolge binärer Informationen umgewandelt. Die Amplitude eines Signals wird in kurzen Zeitabständen gemessen und als absoluter Zahlenwert festgehalten. Die Grundlagen dieser Analog/Digital-Umsetzung (AD, ADC... Analog-Digital-Konverter) bzw. die Digital/Analog-Umsetzung (DA, DAC... Digital-Analog-Konverter) wird im folgenden Kapitel behandelt.

Die Wandlung eines analogen Signals in eine digitale Audioinformation spaltet sich im Wesentlichen in zwei Prozesse auf, dem Sampling und der Quantisierung.

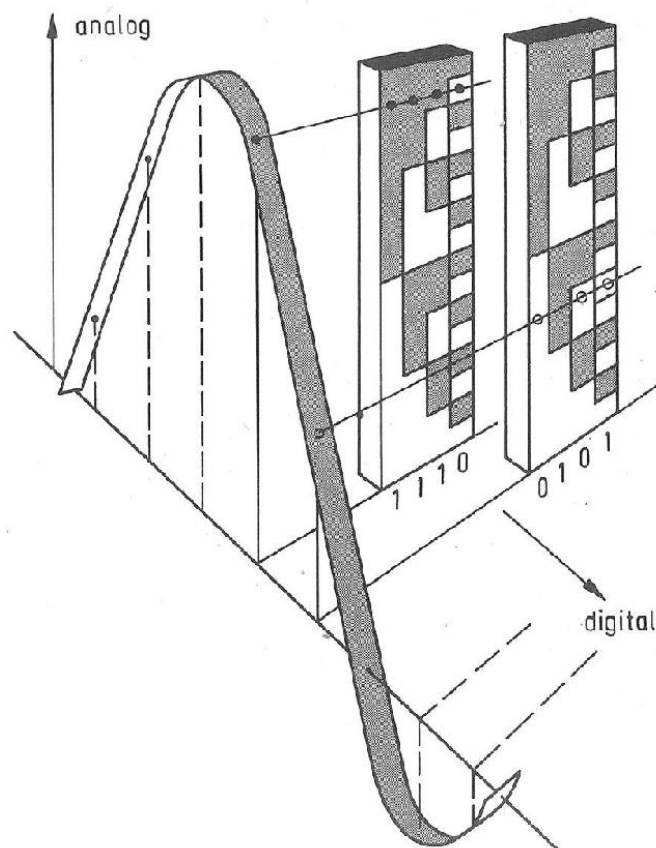


Abb 3-1: Prinzip der Umsetzung eines Analogsignals in eine Digitalinformation [WEBERS]

3.1 Sampling

Sample and Hold

Im Bereich der digitalen Audiotechnik bedeutet Sampling die regelmäßige Abtastung der Amplitude eines Signals. Ein Sample ist also ein Momentanwert eines Signals. Der Wert dieses Samples muss für eine gewisse Zeit lang festgehalten werden, nämlich bis die Amplitude exakt erfasst worden ist. Für diesen Vorgang wird die Bezeichnung *Sample & Hold* verwendet. Schaltungstechnisch wird dies mittels eines Schalters, der durch einen Sample-Impuls gesteuert wird und eines Haltekreises mit Kondensator realisiert.

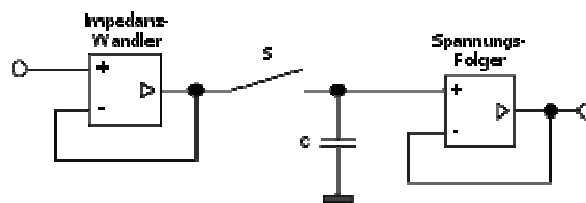


Abb 3-2: Schematische Anordnung einer Sample & Hold-Schaltung

Die Qualität der Abtastung hängt von der Gleichmäßigkeit des Taktgebers sowie den Kenngrößen der Sample & Hold-Schaltung ab:

Acquisition Time: Zeit (Einstellzeit) zwischen dem Sample-Befehl bzw. dem Hold-Befehl und dem Einschwingen innerhalb einer bestimmten Fehlergrenze.

Aperture Time: Zeit zwischen dem Öffnen bzw. Schließen des Halbleiterschalters und dem Einschwingen der Ausgangsspannung innerhalb einer bestimmten Fehlergrenze.

Effective Aperture Time: Effektive Öffnungszeit. Sie berücksichtigt zusätzlich die Laufzeiten innerhalb des Analog- und des Steuersignal-Weges.

Aperture Jitter: Diese statistische und signalabhängige Unsicherheit der Aperturzeit führt zu Amplitudenfehlern.

Slew Rate: Maximale Anstiegsgeschwindigkeit der Ausgangsspannung.

Offset Voltage: Differenz zwischen Ausgangsspannung und Eingangsspannung im Sample-Betrieb.

Charge Offset (Hold Step): Wenn sich das Kontroll-Signal („Hold“-Signal) ändert, erfolgt, durch diesen relativ großen Spannungssprung, über Streukapazitäten eine Einkoppelung in den Signalweg. Das „Hold“-Signal wirkt sich in Form eines Spannungsoffsets innerhalb des Signalwegs aus.

Droop Rate: Änderung der Ausgangsspannung im Hold-Betrieb, verursacht durch Kondensatorleckströme und Eingangsströme der Operationsverstärker.

Output Impedance: Der Ausgangswiderstand sollte möglichst niederohmig sein.

Bei der parallelen Abtastung verschiedener Signale (z. B. einer Mehrspuraufnahme) müssen alle Kanäle gleichzeitig abgetastet werden, um Phasenfehler zwischen den Kanälen zu vermeiden.

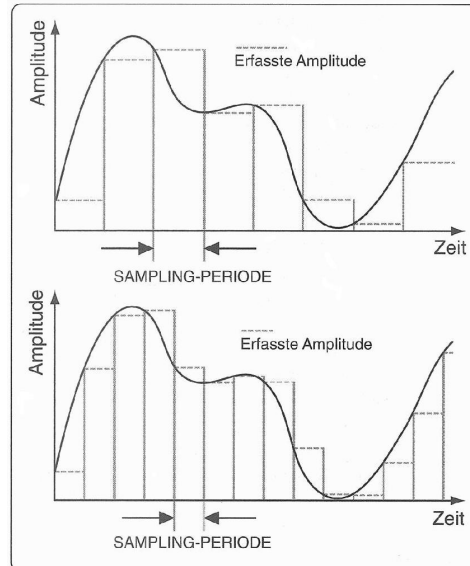


Abb 3-3: Sampling mit verschiedenen Abtastfrequenzen [HENLE]

In der Abbildung erkennt man, dass immer ein gewisser Anteil des analogen Signalverlaufs unberücksichtigt bleibt. Durch ausreichendes wiederholtes Abtasten kann das Signal nahezu verlustfrei reproduziert werden; eine sehr hohe Abtastung führt aber auch zu einer sehr hohen Informationsdichte auf Kosten von Rechenaufwand und Speicherplatz.

Jitter

Idealerweise erfolgt die Abtastung des Spannungsverlaufs in gleichen Zeitabständen. Tritt eine zeitliche Varianz der Flanken des Abtasttaktes gegenüber der Sollposition ein, so spricht man von einem Jitter-Fehler.

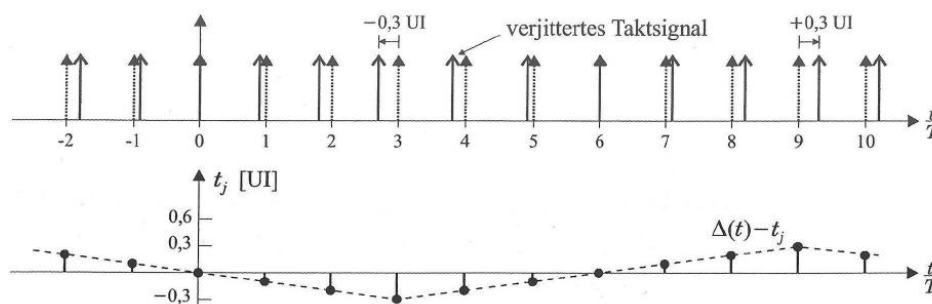


Abb 3-4: verjittertes Signal bei dreieckförmigem Jittersignal mit einer Jitterfrequenz von $f_c/12$ und einer Jitteramplitude von $3 UI$ (Unit Interval) [GRABER]

Abb 3.4 zeigt einen Jitter-Fehler mit regelmäßigen Abweichungen, welcher zu Störfrequenzen in Form von Verzerrungen führt. Bei zufälligen Abweichungen führen diese Jitter-Fehler zu einem Rauschen, dem sogenannten Phasenrauschen.

Abtastfrequenz

Die Anzahl der Samples, die pro Sekunde erfasst werden, gibt man als Abtastfrequenz (Sample Rate) an. Eine Abtastfrequenz von 48 kHz bedeutet, dass pro Sekunde 48.000 Samples erfasst werden. Bei Audio-CDs wird eine Abtastfrequenz von 44,1 kHz verwendet, bei DVD-Audio eine bis zu 196 kHz. Weitere typische Frequenzen sind 48 oder 96 kHz. Bei Verwendung einer Software lässt sich diese Frequenz oft regulieren.

Nyquist-Shannon-Abtasttheorem (Shannon'sches Abtasttheorem)

Nach dem Shannon'schen Abtasttheorem muss die Abtastfrequenz mindestens doppelt so hoch sein, wie die höchste zu erfassende Frequenz. Um Frequenzen, die den Hörbereich – ein Frequenzbereich von 20 Hz bis 20 kHz - betreffen, verlustfrei erfassen zu können ist eine Abtastfrequenz von mindestens 40 kHz erforderlich.

$$f_{\text{abtast}} > 2 \cdot f_{\text{max}}$$

Gleichzeitig folgt aus dem Shannon'schen Abtasttheorem, dass Signale mit einer Frequenz über der halben Abtastfrequenz (Nyquist-Frequenz) nicht mehr exakt wiederhergestellt werden können.

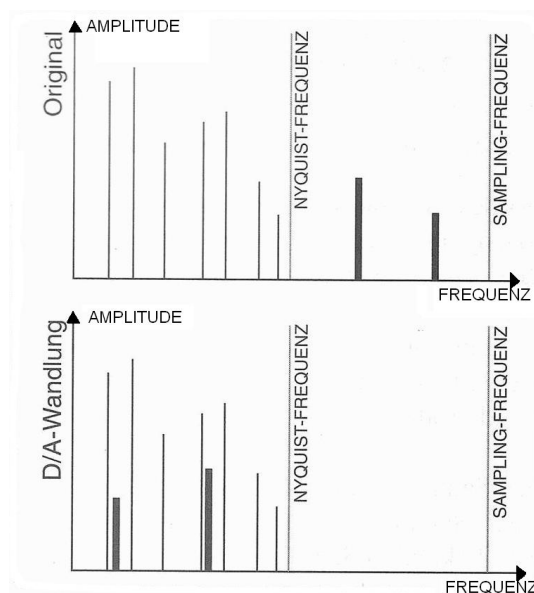


Abb 3-5: Wirkung von Aliasing-Fehlern [HENLE]

Bei der Rückwandlung des digitalen Signals in einen analogen Spannungsverlauf (DA-Wandlung) werden diese Frequenzen, die zuvor nicht hörbar waren, in den hörbaren Bereich „gespiegelt“. Dieses Phänomen wird als „Aliasing“ bezeichnet.

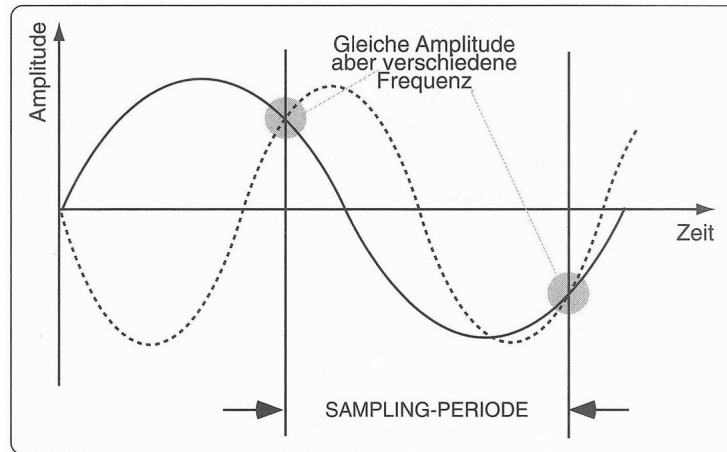


Abb 3-3-6: Zustandekommen des Aliasing-Fehlers [HENLE]

Das Abtasttheorem besagt, dass eine Frequenz mindestens zweimal innerhalb einer Wellenlänge abgetastet wird (durchgezogene Linie in Abb. 3-6). Die Samples eines Signals mit einer über der halben Sampling-Frequenz liegenden Frequenz (gestrichelte Linie) ergeben bei der DA-Umsetzung die gleiche Information.

Um Aliasing zu vermeiden, muss man entweder eine höhere Abtastfrequenz wählen oder das Signal entsprechend bandbegrenzen, d.h., vor der Wandlung muss ein steilflankiges Tiefpassfilter vorgeschaltet werden, das die sogenannten Alias-Frequenzen herausfiltert. Die Wahl einer höheren Abtastfrequenz hat den Nachteil eines erhöhten Rechenaufwands und Speicherplatzbedarfes. Die Tiefpassfilterung hat den Nachteil, dass ein Tiefpass in der Praxis nicht beliebig steil gewählt werden kann und es bei der Filterung zu Phasenveränderungen kommt, die wiederum Störungen im Bereich der Übertragungsfrequenz verursachen.

Oversampling

Eine Möglichkeit, Aliasing-Fehler zu umgehen und gleichzeitig das Filterdesign zu vereinfachen, ist das Oversampling. Bei der DA-Umsetzung wird das Eingangssignal zunächst mit einer höheren Sampling-Frequenz abgetastet. Dadurch muss das Anti-Aliasing-Filter weniger steil sein und kann auf eine weitaus höhere Eckfrequenz abgestimmt werden. Die eigentliche Anti-Aliasing-Filterung wird dann digital implementiert. Die zusätzlich abgetasteten Samples werden nicht übertragen oder aufgezeichnet, sondern dienen nur der Filterung.

Ein weiterer Vorteil dieses Verfahrens ist der Gewinn eines höheren Signal-Rausch-Abstandes. Durch eine schnellere Abtastung wird eine Verminderung des Quantisierungsrauschens erreicht (vgl. Kapitel 3.2)

Downsampling

Downsampling ist vom Prinzip her der umgekehrte Vorgang des Oversamplings. Hat man ein digitales Signal, dessen Frequenzband wesentlich kleiner ist, als die halbe Abtastfrequenz, so kann man die Taktfrequenz ohne Informationsverlust reduzieren. Grundsätzlich ist beim Downsampling darauf zu achten, dass die Bedingung des Shannon'schen Abtasttheorems

erfüllt bleibt, also vor dem Downsampling die Bandbegrenzung auf die Hälfte der neuen Abtastfrequenz erfolgt.

Downsampling wird auch bei der wahrnehmungsangepassten Datenreduktion verwendet, um bandbegrenzte Signale mit minimal möglicher Datenrate zu übertragen, zu speichern oder zu bearbeiten.

3.2 Quantisierung

Quantisierung ist die Übersetzung der durch die Sample & Hold-Schaltung ermittelten Spannungswerte in eine binäre Information. Die Genauigkeit der Quantisierung hängt davon ab, aus wie vielen Bits das digitale Wort zusammengesetzt wird. Wird eine Wortbreite von 16 Bit gewählt, erlaubt dies eine Unterscheidung in $2^{16} = 65.536$ Einzelwerte. Es können dementsprechend viele Spannungswerte bestimmt werden. Auch hier wird die Qualität des digitalen Signals mit einer höheren Wortbreite verbessert. Dadurch erhöht sich aber auch der Speicherplatzbedarf. Audio-Signale auf CDs werden mit 16 Bit kodiert. In der Audiotechnik wird aber oft mit einer höheren Bit-Quantisierung gearbeitet.

Digitale Verzerrungen

Bei der AD-Umsetzung treten bei der Überschreitung der Vollaussteuerung erhebliche Verzerrungen auf. Für den Umsetzer ergibt ein übersteuertes Signal eine Eingangsamplitude, die über der maximalen erlaubten Frequenz liegt. Dieser Bereich kann nicht mehr erfasst werden. Dies führt zu starken, in der Regel unerwünschten Verzerrungen.

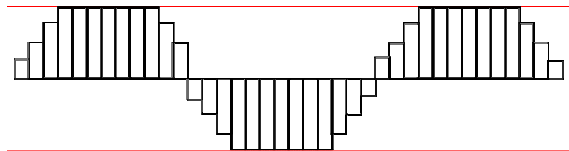


Abb 3-3-7: Übersteuerung eines digitalen Signals [Grafik: David Batley on Wikipedia]

Im Gegensatz dazu führen analoge Übersteuerungen - beispielsweise in Bandmaschinen oder Röhrenverstärkern - die im Sättigungsbereich betrieben werden, zu Effekten, die nicht unangenehm klingen müssen und unter Umständen sogar erwünscht sind.

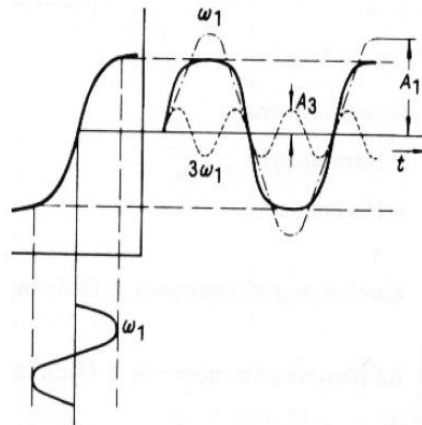


Abb 3-3-8: analoge Übersteuerung mit Erzeugung von ungeraden harmonischen Schwingungen [SENGPIEL3]

Am Beispiel der Abb. 3-8 werden durch die Übersteuerung harmonische Teiltöne erzeugt.

Quantisierungsrauschen

Wie in Abb. 3.2 ersichtlich ist, treten beim Quantisieren Abweichungen vom tatsächlichen Wert der gesampelten Eingangsspannung auf. Diese Abweichung führt zu einem Rauschen, dem sogenannten Quantisierungsrauschen (auch Granularrauschen).

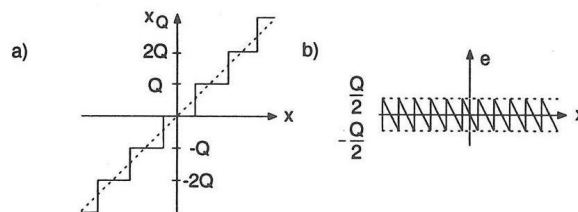


Abb 3-9: a) Nichtlineare Kennlinie des Quantisierers mit dem Eingangssignal x und dem quantisierten Ausgangssignal x_Q b) Quantisierungsfehler e [ZÖLZER]

Den Pegel dieses Rauschens bzw. die Differenz zwischen maximalem Pegel und Rauschen bei einer Quantisierung mit N Bits kann man mit einer einfachen Formel berechnen:

$$\text{Signal/Rausch-Abstand} = 6 N + 1,8 \text{ [dB]}$$

Bei einer 16-Bit-Quantisierung beträgt diese Differenz demzufolge 97,8 dB. Dieser komfortable Dynamikumfang wird jedoch in der Praxis reduziert, um einen Headroom für unvorhergesehene Pegelspitzen zu schaffen, damit Verzerrungen auch in solchen Fällen ausgeschlossen werden können.

Das Quantisierungsrauschen macht sich bei geringen Amplituden bemerkbar, die sich durch die gewählte Quantisierung nicht hinreichend präzise beschreiben lassen, sowie bei tiefen Frequenzen, bei denen lange schleifende Übergänge zwischen den Quantisierungsstufen auftreten. Es macht sich als schmalbandiges Störgeräusch im Nutzfrequenzbereich bemerkbar, welches als sehr störend wahrgenommen wird, da es in keinem harmonischen Zusammenhang mit dem Nutzsignal steht.

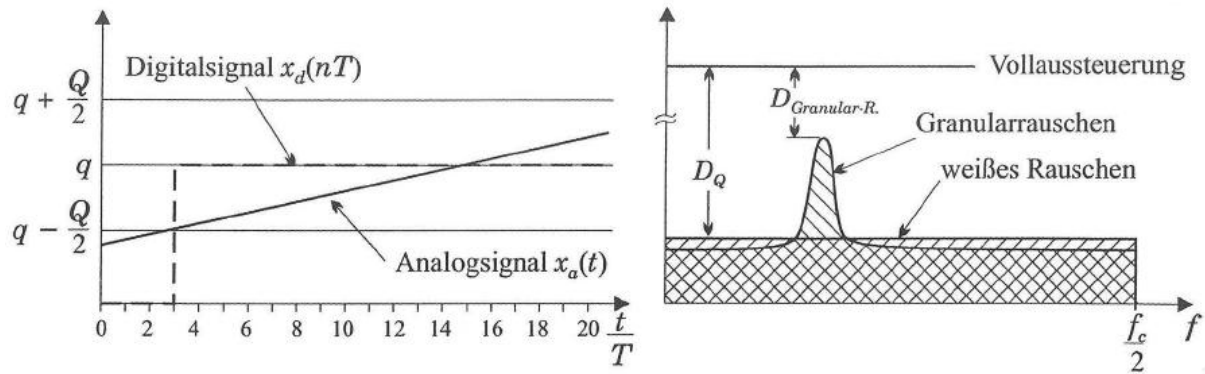


Abb 3-3-10: Granularrauschen als schmalbandiges Störgeräusch und Verringerung der verfügbaren Dynamik [GRABER]

3.3 Dithering

Es gibt zwei Anwendungen, für die Dithering genutzt wird: zum einen, um niedrige Signalpegel zu linearisieren, zum anderen bei der Re-Quantisierung, also der erneuten Quantisierung bereits quantisierter Signale. Letzteres tritt in der Praxis bei Wortlängenreduktion auf, also z. B. bei einer Überspielung von einem 24-Bit-Signal auf ein 16-Bit-Signal. Dies ist erforderlich, wenn ein Audiosignal, das rechnerintern mit 24 Bit verarbeitet wird, auf eine CD überspielt werden soll. Diese benötigt ein 16-Bit-Signal. Somit muss requantisiert werden.

Quantisierungsrauschen wird bei Signalen mit kleiner Leistung, also z. B. beim Ausklingen eines Musikstücks, hörbar. Um dies zu verhindern wird vor der eigentlichen Quantisierung dem Signal ein Rauschen, das Dither-Rauschen hinzugefügt. Das Dither-Rauschen entspricht einer digitalen Zufallsfolge und ähnelt in etwa weißem Rauschen.

Durch das Hinzufügen des Dither-Rauschens erhöht sich zwar die Gesamtrauschleistung, schmalbandiges und unangenehmes Granularrauschen wird jedoch vermieden.

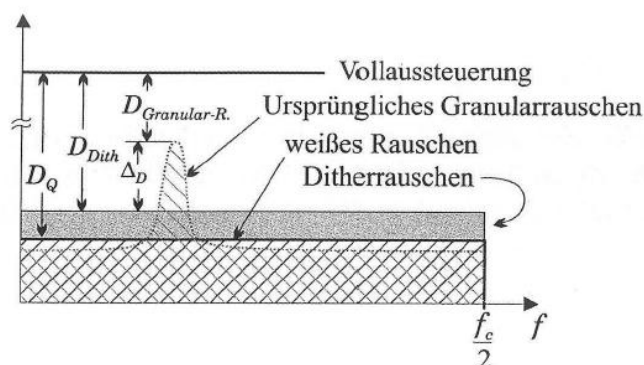


Abb 3-11: Verbesserung der verfügbaren Dynamik durch Dithering [GRABER]

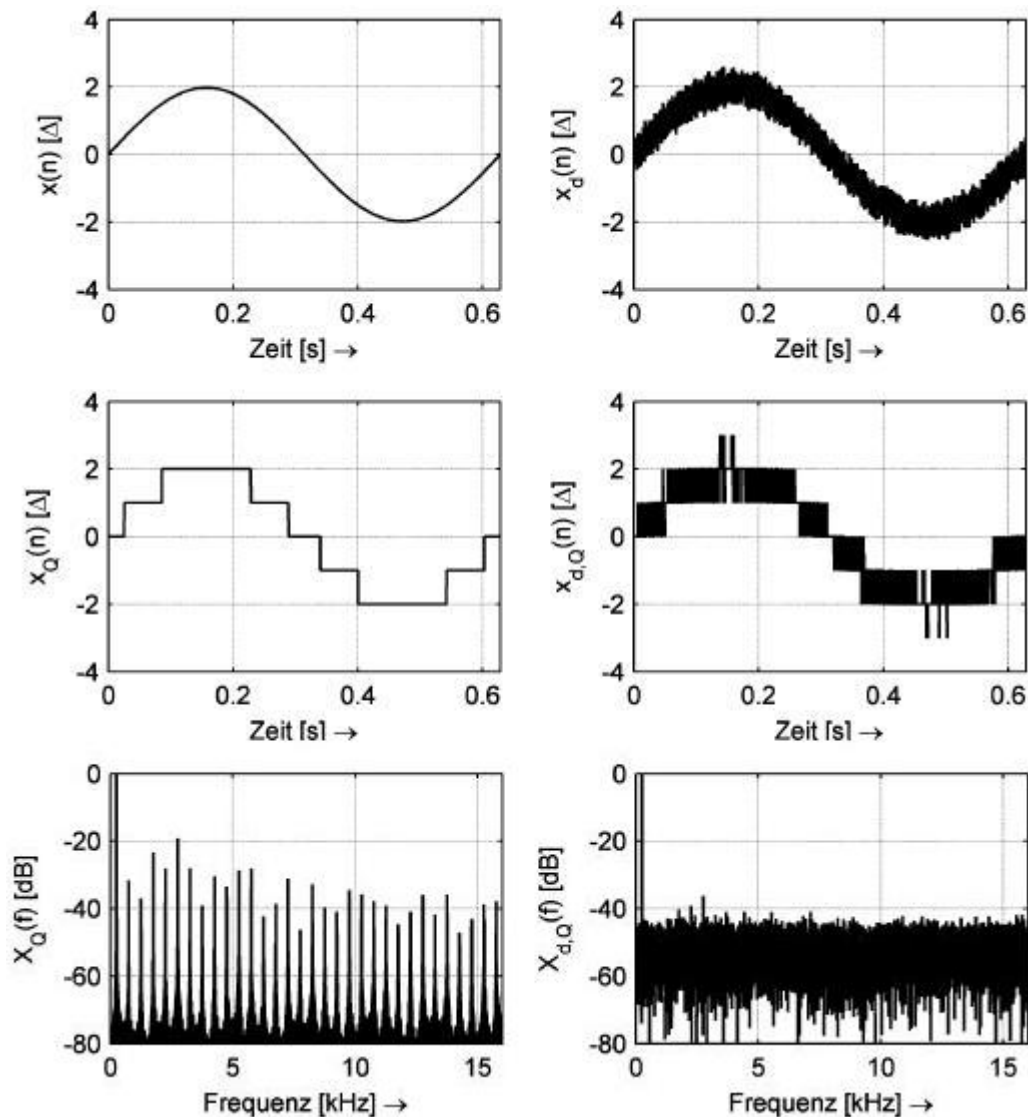


Abb 3-12: Von oben nach unten: Eingangssignal, mit 3 Bit Wortbreite quantisiertes Eingangssignal, Spektrum des quantisierten Signals. Links: ohne Dither. Rechts: mit Dither [WEINZIERL]

In Abbildung 3-12 ist die positive Wirkung des Ditherings deutlich erkennbar. Die störenden Frequenzen, wie sie im Bild unten links zu sehen sind, treten mit Dither nicht mehr auf.

3.4 AD-Wandlerverfahren

Pulse Code Modulation

Pulse Code Modulation oder kurz PCM beschreibt allgemein den Vorgang der AD-Wandlung. Für jedes Sample wird ein Datenwort erzeugt.

Es gibt mehrere Wege, um aus einer Spannung einen Zahlenwert abzuleiten. Wandler mit hoher Auflösung und Genauigkeit sind relativ langsam, Wandler mit relativ niedriger

Auflösung dagegen schneller. Im Folgenden werden unterschiedliche Wandlerverfahren kurz vorgestellt.

Zählwandler

Der DA-Wandler wird von einem Binärzähler angesteuert. Dieser erhält seine Zählimpulse aus einem Generator über ein UND-Gatter. Ein Komparator, der die Eingangsspannung mit der Ausgangsspannung des DA-Wandlers vergleicht, gibt das UND-Gatter frei.

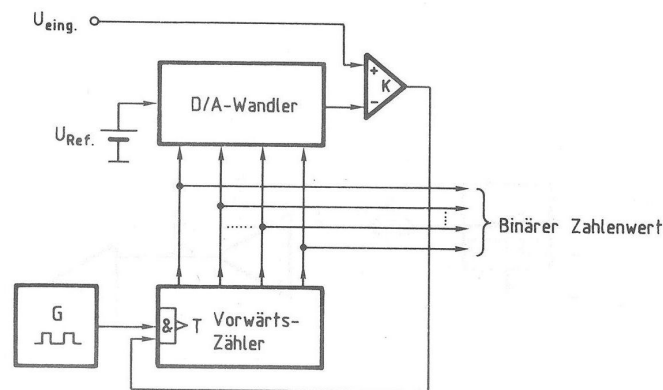


Abb 3-13: Zählwandler [DICKREITER]

Zu Beginn des Wandlungsvorgangs fängt der Zähler bei null an zu zählen. Die Ausgangsspannung des Wandlers erhöht sich stufenweise, bis sie den Wert der Eingangsspannung erreicht hat. Dann sperrt der Komparator das UND-Gatter und verbindet den Zähler mit dem Ausgang. Der Zählerstand entspricht dann dem Binärwert des Abtastwertes.

Der Vorteil dieses Wandlers liegt in seinem einfachen Aufbau, der Nachteil besteht darin, dass die Wandlungszeit von der Größe des Eingangssignals abhängt.

Dual-Slope-Wandler

Der Dual-Slope-Wandler funktioniert nach dem Prinzip eines Zählwandlers.

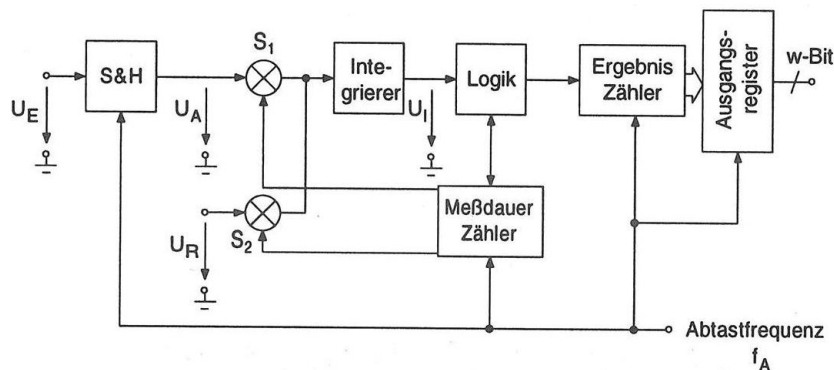


Abb 3-3-14: Dual-Slope-Wandler [ZÖLZER]

In der ersten Phase ist der Schalter S_1 für die Dauer, die vom Messdauerzähler vorgegeben wird geschlossen. Die Ausgangsspannung der Sample & Hold-Schaltung liegt auf dem

Integrierer. In der zweiten Phase öffnet sich der Schalter S_1 und der Schalter S_2 schließt sich. Die Referenzspannung, die eine umgekehrte Polarität aufweist liegt nun am Integrator, der sich dadurch entlädt. Der Ergebniszähler ermittelt durch eine Zählung von Taktimpulsen die Dauer des Entladungsvorgangs, der proportional zur Eingangsspannung verläuft.

Sigma-Delta-Wandler

Die größte Bedeutung unter den Wandlerprinzipien hat die Wandlung nach dem Sigma-Delta-Prinzip. Bei der Quantisierung wird nur die Differenz des aktuellen Samples zum vorherigen Sample erfasst.

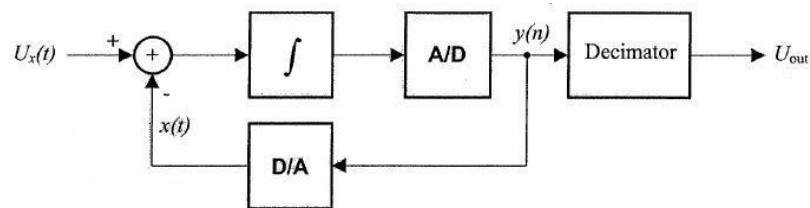


Abb 3-3-15: Ersatzschaltbild eines Sigma-Delta-Wandlers [BRASSEUR]

Im einfachsten Fall besteht der Sigma-Delta-Wandler aus einem ADC mit einem oder wenigen Bit Auflösung. Der Ausgang $y(n)$ des ADC liefert ein Digitalsignal mit hoher Taktrate, das zum einen ein digitales Filter höherer Ordnung speist, zum anderen mit einem Ein-Bit-D/A-Wandler in der Rückkopplungsschleife verbunden ist. Das Ausgangssignal $x(t)$ des DAC wird vom Eingangssignal abgezogen. So wird die Differenz zweier aufeinanderfolgender Samples erfasst.

Das mit hoher Taktrate gelieferte Ausgangssignal des AD-Wandlers $y(n)$ wird digital gefiltert, wodurch ein Wandlungsergebnis U_{out} mit relativ geringer Taktrate, aber großer Wortbreite entsteht.

Parallelwandler

Der Parallel- oder Flash-Wandler hat den Vorteil einer sehr geringen Umsetzungszeit, da der Wert der Eingangsspannung direkt ermittelt werden kann. Er besteht aus $2^w - 1$ Komparatoren für eine w -Bit-Auflösung. Aufgrund der hohen Anzahl an Komparatoren wird dieser Wandlertyp für eine Quantisierungsstufe von bis zu 10 Bit eingesetzt.

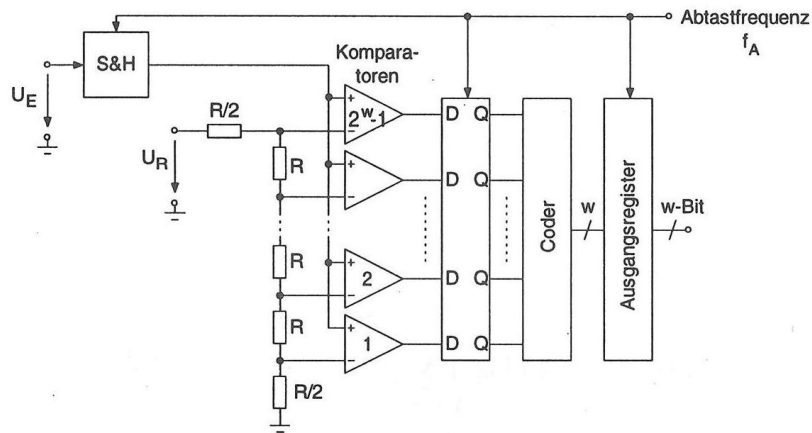


Abb 3-16: Parallelwandler [ZÖLZER]

Jeder Komparator vergleicht das Ausgangssignal der Sample & Hold-Schaltung mit einer Referenzspannung, die über einen Spannungsteiler aus der Referenzschaltung gewonnen wird. Alle Komparatoren, bei denen die Eingangsspannung größer ist als die Referenzspannung, werden am Ausgang einen High-Pegel abgeben, alle anderen behalten einen Low-Pegel. Damit erfolgt an dieser Stelle die Quantisierung. Die Ausgänge der Komparatoren werden in ein Register übernommen und über eine Kodierlogik in ein w-Bit-Datenwort umgesetzt.

Sukzessive Approximation

Der Wägewandler bzw. die Wandlung mit sukzessiver Approximation ist ein weiteres Wandlervorgehen mit relativ kurzer Umsetzzeit. Hier wird das Prinzip der schrittweisen Annäherung genutzt.

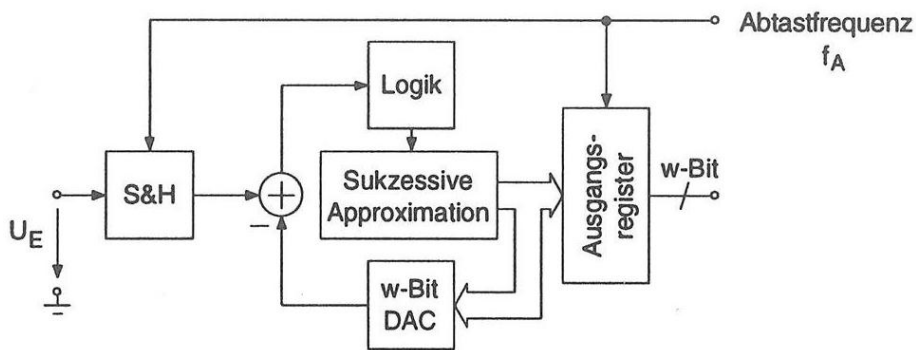


Abb 3-17: Wandlung mit sukzessiver Approximation [ZÖLZER]

Im ersten Schritt wird überprüft, ob eine positive oder negative Spannung am Komparator anliegt. Ist die Spannung positiv, wird im nächsten Schritt durch die Ausgabe von $+0,5 U_R$ über den DA-Wandler überprüft, ob die am Komparatorausgang anliegende Spannung größer oder kleiner als $+0,5 U_R$ ist. Danach erfolgt die Ausgabe von $(+0,5 \pm 0,25)U_R$ über den DA-Wandler, woraufhin der Komparatorausgang wieder ausgewertet wird. Dieser Vorgang wiederholt sich w mal und führt zu dem w-Bit Ausgangswort.

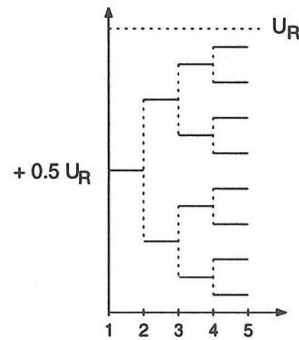


Abb 3-3-18: Sukzessive Approximation [ZÖLZER]

3.5 DA-Wandlerverfahren

Spätestens, wenn das digitale Audiosignal wiedergegeben werden soll, muss die digitale Bit-Folge wieder in Spannungswerte umgewandelt werden. Diese Spannungsverläufe verlaufen zunächst stufenartig. Eine Glättung kann mit einem digitalen Tiefpassfilter erreicht werden. Als eine weitere Möglichkeit zur Glättung kann bei der DA-Umsetzung auch Oversampling verwendet werden. Dazu werden zusätzliche Samples eingefügt, deren Wert mittels Interpolation ermittelt wird.

Auch hier gibt es verschiedene Modelle zur DA-Wandlung, wobei wiederum das Prinzip des Delta-Sigma-Wandler das am häufigsten verwendete und daher bedeutendste ist. Delta-Sigma-Wandler kommen beispielsweise in CD-Playern zum Einsatz.

Parallelverfahren

Mit dem Parallelverfahren lassen sich sehr hohe Abtastraten erzielen, jedoch nur eine begrenzte Auflösung. Daher findet es in direkter Form in der Audiotechnik keine Anwendung. Mit geeigneter Nachbearbeitung durch eine Delta-Sigma-Modulation lassen sich allerdings sowohl eine hohe Geschwindigkeit als auch eine hohe Genauigkeit erreichen.

Während bei der AD-Umsetzung beim Parallelverfahren die Eingangsspannung mit einer Referenzspannung verglichen wird, wird bei der DA-Umsetzung jedem möglichen Wert des Digitalsignals ein Schalter zugeordnet. Über einen Spannungsteiler werden alle durch den Wandler darstellbaren Spannungen bereitgestellt und mittels eines vom Digitalsignal angesteuerten Schalters auf den Analogausgang geschaltet.

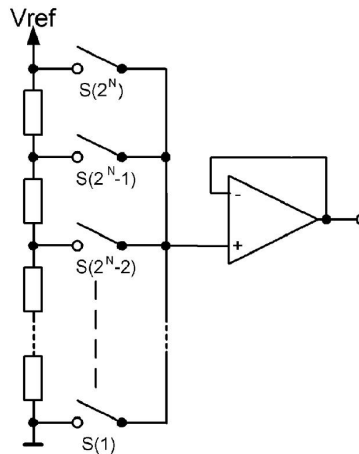


Abb 3-3-19: Parallel-DA-Wandler mit Spannungsteiler [WEINZIERL]

Serielle Verfahren

Bei einer Wandlung mit seriellem Verfahren wird nicht wie beim Parallelverfahren jedem Spannungswert ein Schalter zugeordnet, sondern jedem Bit. Mit Hilfe entsprechend gewichtete Widerstände und einer Spannungsreferenz kann dann die Ausgangsspannung aufsummiert werden.

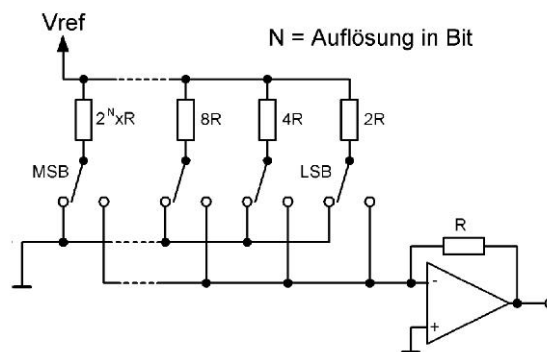


Abb 3-20: Wandler mit gewichteten Widerständen [WIEINZIERL]

Die Widerstandswerte wachsen dabei mit jeder weiteren Schalterposition auf das Doppelte. Da sich diese stark unterschiedlichen Werte allerdings kaum mit ausreichender Genauigkeit realisieren lassen, nutzt man bei einem R/2R-Wandler ein Leiternetzwerk, wie es hier abgebildet ist:

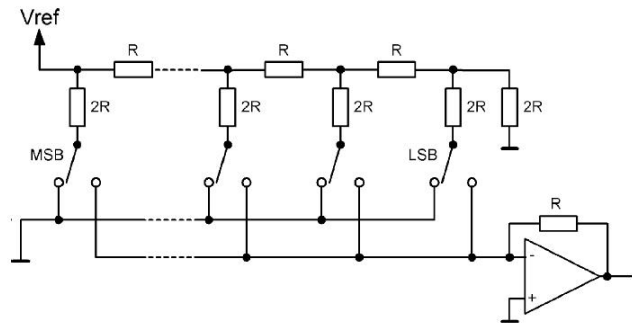


Abb 3-21: R/2R-Leiternetzwerk [WEINZIERL]

R/2R-Wandler erreichen hervorragende Signalrauschabstände und waren deshalb in der Anfangszeit der Digitalisierung die dominierende Technologie bei DA-Wandlern. Heute werden eher kostengünstige Technologien wie der Delta-Sigma-Wandler verwendet.

Glitches

Bei der Wandlung mittels Schaltnetzwerken besteht die Gefahr, dass nicht alle Schalter genau gleichzeitig schalten. Dadurch kann es zu Störimpulsen, den sogenannten Glitches kommen. Am Analogausgang kommt es zu kurzzeitigen steilflankigen Impulsen, bevor sich der richtige Wert einstellt. Dieser Effekt wird durch den Tiefpassfilter am Ausgang des DA-Wandlers geschwächt. Ferner helfen sogenannte Delischer. Hier wird die erzeugte Treppenfunktion von einer Sample & Hold-Schaltung mit einer Verzögerung von einem halben Abtastintervall noch einmal abgetastet.

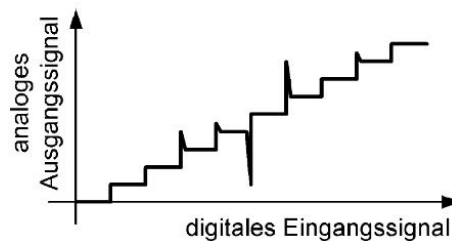


Abb 3-22: Glitches [WEINZIERL]

Delta-Sigma-Umsetzer

Der Delta-Sigma-DA-Umsetzer funktioniert umgekehrt wie der entsprechende AD-Umsetzer. Zunächst werden dem Umsetzer die Datenworte in ein Eingangsregister übergeben. Danach erfolgt eine Abtastratenerhöhung. Ein Datenwortsignal mit mehreren Bit wird zu einem 1-Bit-Signal quantisiert. Am Eingang liegt nun ein serielles 1-Bit-Signal an, dessen Mittelwert dem analogen Signal entspricht. Überwiegen im 1-Bit-Signal die Einsen gegenüber den Nullen, so stellt dies einen positiven Analogpegel dar. Überwiegen die Nullen, handelt es sich um einen negativen Analogpegel. Das Verhältnis von „1“ und „0“ bestimmt demzufolge den exakten Betrag.

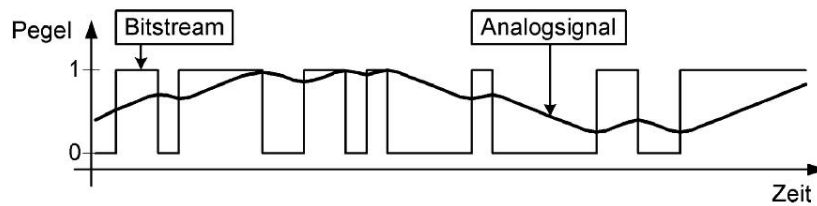


Abb 3-23: Bitstream und daraus gefiltertes Analogsignal [WEINZIERL]

Für ein genaues Ergebnis sind, anders als bei einer normalen Abtastrate, wesentlich schnellere Wechsel zwischen „1“ und „0“ erforderlich. Deshalb benötigen Delta-Sigma-Wandler ein sehr hohes Oversampling, um die hohe Taktrate für den Bitstream zu erreichen.

Es folgt hier noch ein analoges Tiefpassfilter zur Glättung des Spannungsverlaufs, mithilfe dessen hohe Frequenzanteile entfernt werden.

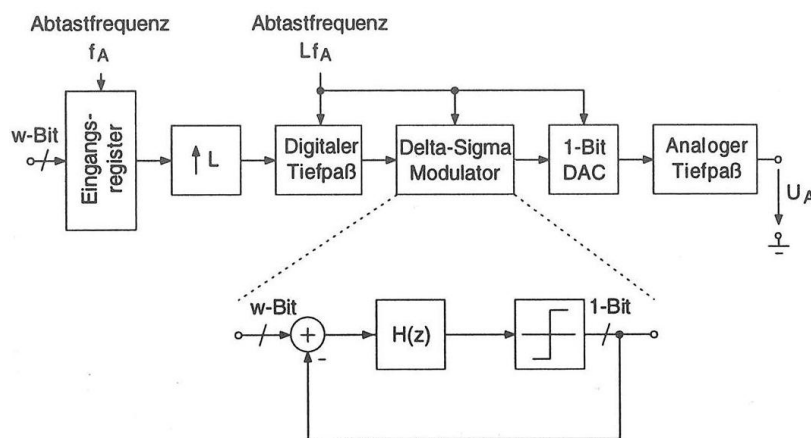


Abb 3-3-24: Delta-Sigma DA-Umsetzer [ZÖLZER]

3.6 Zusammenfassung DA/AD-Wandlung

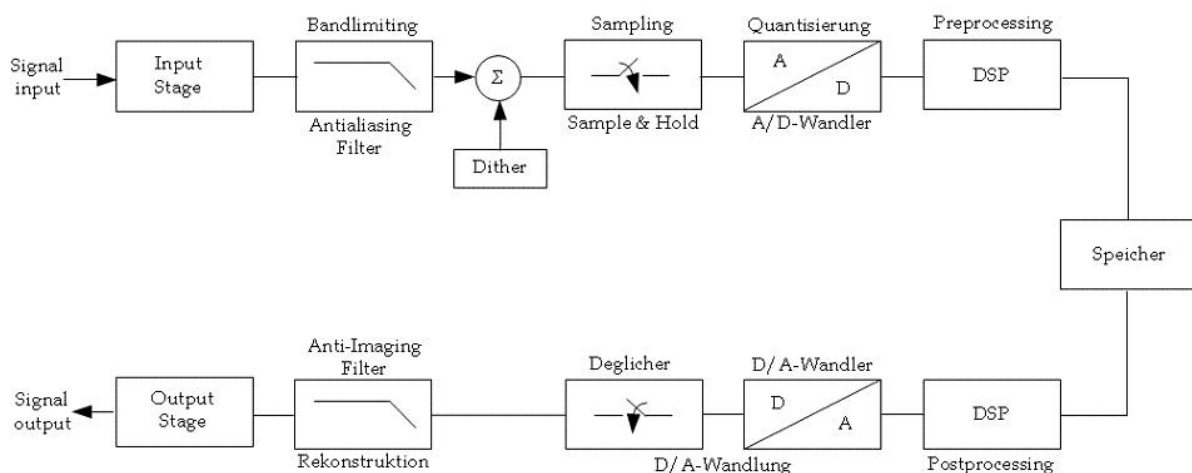


Abb 3-25: Blockschaltbild einer DA- und anschließender AD-Wandlung [KRUSE]

In obiger Abbildung ist noch einmal der Signalweg der AD- und DA-Wandlung dargestellt. Das Analogsignal wird gefiltert und ein Dithersignal wird aufsummiert. Eine Sample & Hold-Schaltung tastet die Spannungswerte in bestimmten Zeitintervallen ab. Im AD-Wandler wird das Signal quantisiert und steht zur digitalen Weiterverarbeitung oder Speicherung zur Verfügung (DSP = Digital Signal Processor). Im DA-Wandler werden aus dem quantisierten Signal wieder analoge Spannungsverläufe rekonstruiert. Der nachfolgende Deglitcher und der Tiefpassfilter schützen vor Störimpulsen und -geräuschen.

4 Geräte und Schnittstellen

4.1 Digitale Schnittstellen

In der digitalen Audiotechnik kommen unterschiedliche Datenströme vor, die Audioinformation beinhalten. Je nach Anwendung müssen diese Datenströme durch eine geeignete Kodierung aufbereitet werden. Ein direkter Austausch digitaler Audioinformationen durch das einfache Zusammenstecken zweier Kabel ist daher in aller Regel ausgeschlossen.

Um einen Datentransfer zwischen verschiedenen Geräten zu ermöglichen, beispielsweise für eine weitere Bearbeitung von Aufnahmen oder die Übertragung auf ein Medium zur kommerziellen Verbreitung (z. B. CD) gibt es verschiedene Schnittstellenformate. Damit können digitale Audiodaten verlustfrei von einem auf ein anderes Format kopiert werden. Diese Schnittstellen und Formate sind jedoch ihrerseits nicht miteinander kompatibel, was dann einen Formatkonverter erfordert, der die Daten von einem in das andere Format umwandelt.

Für eine digitale Verkopplung von Aufzeichnungsgeräten, Effektgeräten, etc. einerseits und digitalen Mischpulten andererseits, sind genormte Schnittstellen notwendig. Inzwischen gibt es eine Vielzahl verschiedener Schnittstellen. Im Folgenden werden die wichtigen beschrieben.

Wordclock

Zunächst soll jedoch auf eine wichtige Voraussetzung eingegangen werden, die von großer Bedeutung in der gesamten digitalen Audiotechnik ist. Wo immer digitale Audiodaten verarbeitet oder übertragen werden, steuert ein Taktsignal sämtliche Bearbeitungs- und Übertragungsvorgänge. Es stellt die eindeutige Zuordnung der einzelnen Datenbits auf der Zeitachse dar. Die Wordclock ist das häufigste verwendete Taktsignal und steht in den meisten Geräten zur Verfügung. Der Vollständigkeit halber sei erwähnt, dass die Wordclock jedoch nicht das einzige Taktsignal ist, welches Verwendung findet.

Bei Zusammenschaltung verschiedener digitaler Audiokomponenten müssen diese von derselben Wordclock getaktet werden. Bei einigen Schnittstellenformaten muss die Wordclock über eine eigene Verbindung übertragen werden, bei anderen Formaten kann das Aufnahmegerät die Wordclock aus dem Takt der eingehenden Daten generieren. Es lassen sich auch verschiedene digitale Komponenten extern takten, so dass auch die Möglichkeit besteht, beide Geräte bei einer Überspielung extern von derselben Wordclock zu steuern.

In einer Studioumgebung müssen alle Komponenten von einer gemeinsamen Wordclock getaktet werden. Hierzu wird im Allgemeinen ein spezieller Taktgenerator verwendet, der sämtliche Komponenten mit einem gemeinsamen Takt versorgt.

Die meisten Geräte verfügen über einen eigenen Anschluss, der meist mit WCLK (Wordclock) beschriftet und als BNC-Buchse ausgeführt ist.

4.1.1 AES3

AES3-Interface ist eine zweikanalige Schnittstelle. Aufgrund der engen Verwandtschaft zwischen AES3 und dem EBU Tech. 3250-E-Interface wird sie häufig auch AES/EBU genannt. Es wurde 1985 von der Audio Engineering Society (AES) als das erste normierte digitale Audio Interface definiert. Die meisten verfügbaren Zwei- oder Mehrkanal-Schnittstellen basieren weitgehend auf diesem Standard, der seinerseits auf den Grundüberlegungen zur SDIF-Schnittstelle der Firma Sony beruht. Zweikanalige digitale Audiointerfaces mit differentieller Übertragung wie AES3 oder EBU Tech. 3250-E sind untereinander weitgehend kompatibel. Sie verwenden differentielle, also symmetrische Signalübertragung mit einem Spannungspegel zwischen 2 und 7 Volt, 2-polig geschirmte Leitungen mit einem Wellenwiderstand von 110Ω und XLR-Steckverbindungen. Es müssen bei einer Samplingfrequenz von 48 kHz Entfernungen von mindestens 100 m fehlerfrei überbrückbar sein.

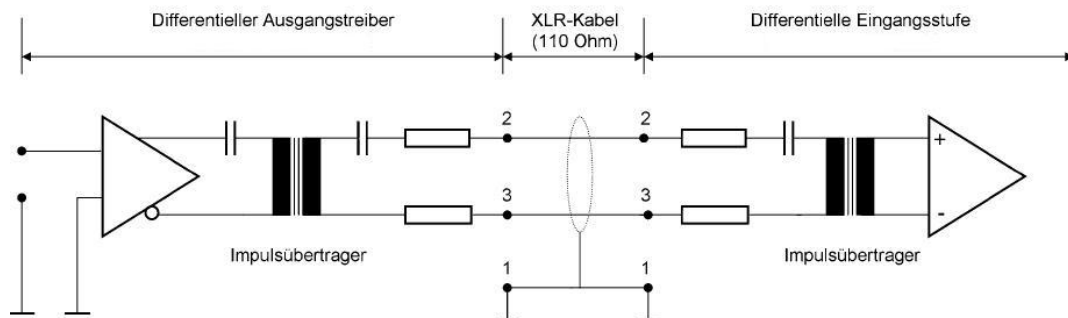


Abb 4-4-1: Differentielles Interface mit galvanischer Trennung nach EBU Tech. 3250-E, kompatibel zu AES3 [WEINZIERL]

Datenformat

Jeder AES3-Datenblock besteht aus 192 so genannten Frames, die jeweils in zwei Subframes mit den Daten für jeweils einen Kanal (L/R) aufgeteilt sind. Eine Wordclock muss nicht separat übertragen werden, da der Empfänger diese aus den eingehenden Daten selbst ableiten kann. Ein Subframe ist 32 Bit lang, von denen 20 Bit für die digitalen Audioinformationen vorgesehen sind, 4 Sync-Bits für die Präambel und 4 Auxiliary-Bits die wahlweise für Zusatzinformationen oder als Erweiterung der Audiodaten von 20 Bit auf 24 Bit zur Verfügung stehen. Nach den eigentlichen Audiodaten befinden sich am Ende des Subframes die Bits „Validity“, „User“, „Channel Status“ und „Parity“.

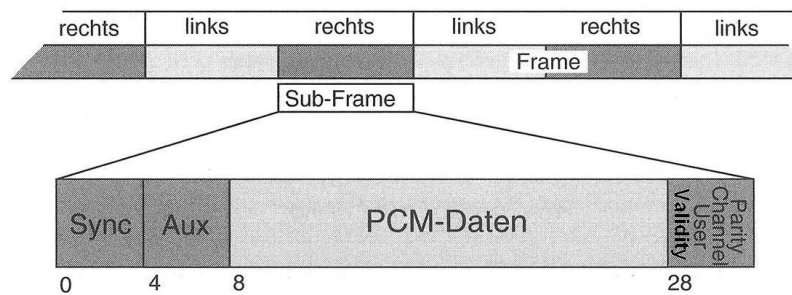


Abb 4-2: Datenformat der AES3-Schnittstelle: Block, Frame und Subframe [HENLE]

Das *Validity-Bit* (auch V-Bit oder Gültigkeitsbit) signalisiert, ob die Audio-Informationen zur Wiedergabe geeignet sind, oder nicht. Der Wert Logisch 1 signalisiert Störungen in der Signalquelle oder allgemein „ungültiges Audio“. Der Wert Logisch 0 steht für „gültiges Audio“, wobei die Entscheidung, wann ein gültiges oder ungültiges Audio vorliegt, weitgehend herstellerabhängig ist.

Das *User-Bit* (auch U- oder Anwender-Bit, User Channel) kann vielfältige Informationen übertragen. Bei der Anwendung im Heimbereich kann es z. B. zur Übertragung von Tracknummer, Titelname und Interpret zwischen CD-Player und z. B. MiniDisc-Recorder eingesetzt werden. Die Informationen werden in diesem Fall über jeweils einen Block gesammelt.

Die *Channel-Status-Bit* (auch C-Bit oder Kanalstatus-Bit) jedes Subframes werden über die Dauer eines Blocks gesammelt. Die so entstehenden 192 Bit pro Kanal ergeben 24 Bytes zu je acht Bit. Bei 48 kHz Samplingfrequenz wird so alle 4 ms ein vollständiger Informationsblock aus 24 Bytes erreicht. Die Informationen, die über diese Bits übertragen werden können, sind sehr vielfältig und daher ein häufiger Grund für Inkompatibilitäten zwischen zwei Schnittstellen.

Das *Parity-Bit* (auch P-Bit oder Prüf-Bit) wird in Abhängigkeit des Bitinhalts der Bits 4 bis 31 so gesetzt, dass sich immer eine gradzahlige Prüfsumme über alle Bits der Wertigkeit Logisch 1 ergibt. Auf diese Weise kann eine ungerade Anzahl an Fehlern pro Subframe erkannt werden.

4.1.2 AES3id und AES 42

Das unsymmetrische Interface AES3id kommt vor allem im Broadcastbereich und bei videonahen Audiogeräten wie Dolby E und Dolby Digital Encodern und professionellen Videomaschinen zum Einsatz. Es arbeitet mit einer videotypischen Signalspannung von maximal 1 V, verwendet Videokabel mit 75 Ω Wellenwiderstand und BNC-Steckverbinder. Das Protokoll entspricht dem des AES3-Interfaces.



Abb 4-3: Interface nach AES3-id mit Impedanzanpassung an AES3 [WEINZIERL]

AES42 definiert ein Zweikanal-Interface mit integrierter Spannungsversorgung und Synchronisation. Die Spannungsversorgung erfolgt mit 10 V (statt 48 V Phantomspeisung). Als Strom sind 250 mA spezifiziert. Die Fernsteuerung des Mikrofons erfolgt über eine Pulsmodulation der Versorgungsspannung mit $2\text{ V} \pm 0,2\text{ V}$ und einer Datenrate von 750 Bit/s (bei 48 kHz Samplingfrequenz).

4.1.3 S/PDIF

S/PDIF (Sony/Philips Digital Interface) weist gegenüber AES3 Unterschiede bei den elektrischen Parametern und bei der Verwendung der VUCP-Informationen auf. Es wurde aus Gründen des Urheberrechtsschutzes mit einem Kopierschutz ausgestattet. Das Interface verwendet eine unsymmetrische Schnittstelle mit einer Impedanz von $75\ \Omega$. Als Steckverbindung kommt eine Cinch-Buchse zum Einsatz. Die Channel-Status-Daten enthalten oft nicht nur Informationen über einen Kopierschutz, sondern auch über die Art des sendenden Gerätes (Gerätetyp und Herkunftsland).

Da jedoch nur wenige Hersteller alle Informationen senden bzw. auswerten, funktioniert die Verbindung der ungleichen Schnittstellen meist recht gut. So darf etwa ein digitales Heimgerät nach S/PDIF keine höheren Abtastraten als 48 kHz über das Interface ausgeben, um das Kopieren hochwertiger Audioinhalte unmöglich zu machen. In der Praxis hält sich jedoch kaum ein Hersteller an diese Empfehlung, so dass DVD-Video- und DVD-Audio-Player mit bis zu 192 kHz Samplingfrequenz über die S/PDIF-Schnittstelle senden, um den Anschluss an hochwertige Heimkino-Anlagen mit digitalen Decodern und Verstärkern zu ermöglichen.

Das TOSLink-Interface, welches von der Firma Toshiba entwickelt wurde und auch als S/PDIF Optical bekannt ist, verwendet als Übertragungsmedium Lichtwellenleiter aus Kunststoff, die mit Licht von 660 nm Wellenlänge arbeiten. Die maximale Datenrate der Interfaces beträgt etwa 15 MBit/s, was für Audio mit bis zu 96 kHz Abtastfrequenz ausreichend ist.



Abb 4-4-4: TOSLink Stecker [Cyvh on en.wikipedia]

4.1.4 SDIF

Das bereits 1982 eingeführte SDIF-2 (Sony Digital Interface) verwendet getrennte Leitungen für den linken und rechten Kanal und eine zusätzliche Verbindung für das Clocksignal. Zur Übertragung werden Koaxialleitungen mit 75Ω Wellenwiderstand und BNC-Steckern verwendet. Der Spannungspegel beträgt 5 V. Bei einigen digitalen Mehrspurrekordern kommt eine Variante mit differentieller Übertragung mit 50-poligen Sub-D-Verbindungen zum Einsatz. Die Wordclock wird auch hier über eine separate Koaxial-Leitung übertragen. Zusätzlich zu den ursprünglichen 20 Bit breiten Audiodaten werden Informationen über die verwendete Emphase sowie über den Kopierschutz mitgesendet. Pro Abtastwort werden 32 Bit übertragen, von denen 20 für Audio und 9 für User- und Kontrollbits reserviert sind. 3 Bits werden zur Synchronisation eingesetzt. Die Datenrate pro Kanal beträgt 1,53 MBit/s bei 48 kHz und 1,21 Mit/s bei 44,1 MHz.

SDIF-3 wurde für ein bitserielles Abtastverfahren entwickelt, das bei der Aufnahme und beim Mastering von Material für Super Audio CDs (SACD) zum Einsatz kommt.

4.1.5 MADI

Die bisher besprochenen Interfaces übertragen nur Stereo-Signale. In der Praxis werden z. B. auch 3 AES3-Leitungen für die Übertragung von 5.1 Sound verwendet. Wenn man jedoch mehr Kanäle übertragen möchte, wird eine solche Lösung eher unübersichtlich. Daher werden für die Anwendungsfälle Multichannel Interfaces verwendet.

Das Multichannel Audio Digital Interface (MADI) basiert auf einer Initiative der Firmen AMS-Neve, Mitsubishi, Sony und SSL aus dem Jahr 1988 zur Vereinfachung der Verbindung zwischen Mehrspurrekordern und Mischpulten. Heute ist MADI durch AES10-1991 und ANSI S4.43-1991 international standardisiert. Je nach System können 56 oder 64 Audiokanäle über eine Leitung übertragen werden. Die maximalen Abtastraten liegen derzeit bei 96 kHz, die Auflösung beträgt wahlweise 16, 20 oder 24 Bit. MADI kann über elektrische und optische Verbindungen übertragen werden.

Die elektrische Variante verwendet Koaxialkabel mit 75Ω Wellenwiderstand. 56 (bzw. 64) Subframes werden pro MADIFrame übertragen, wobei die einzelnen Subframes den AES3 Frames fast gleichen. Der einzige Unterschied ist eine MADISpezifische Synchronisierung in den Bits 0 – 3.

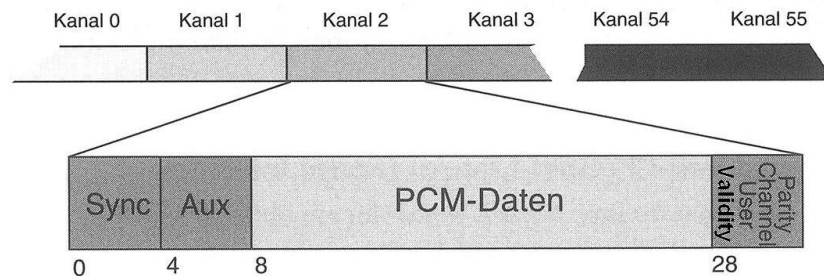


Abb 4-4-5: MADIFormat [HENLE]

Die Übertragungsrate der MADISchnittstelle ist unabhängig von der Samplingfrequenz oder der Anzahl der Audiokanäle. Die Datentransferrate liegt bei ca. 100 MBit/s, die Gesamtdatenrate bei 125 MBit/s. Dieser Unterschied ergibt sich aus der Kodierung, bei der aus Gründen der Synchronisierung zweier Einheiten, in einem Schritt ein 4-Bit-Wort in ein 5-Bit-Wort kodiert wird.

Aufgrund strenger Spezifikationen v. a. für Jitter und die Qualität des Kabelmaterials, kann eine Kabellänge von etwa 50 m erreicht werden.

MADI Optical überträgt das gleiche Datenformat über Lichtwellenleiter aus echter Glasfaser. Zur Umsetzung zwischen Optical und Electrical MADI gibt es bidirektionale Medienkonverter.



Abb 4-6: Bidirektionaler Medienkonverter für MADI Coax (BNC) und MADI Optical (Duplex Fibre Connector) [Soundscape Sydec Audio Engineering]

4.1.6 ADAT

Ende der 1980er-Jahre entwickelte die Firma Alesis achtspurige digitale Aufnahmemaschinen. Zur verlustlosen digitalen Verbindung wurden die Geräte mit optischen Schnittstellen namens ADAT Lightpipe ausgestattet. Das optische Interface ist eine serielle, selbsttaktende Schnittstelle zur Übertragung von acht Audiokanälen mit bis zu 24 Bit Auflösung. Die Abtastrate beträgt üblicherweise 48 kHz. Wenn Audiosignale mit 96 kHz Samplingfrequenz übertragen werden sollen, kommt „Double Speed“ zum Einsatz, wodurch die Anzahl der Kanäle pro Leitung von acht auf vier reduziert wird.

Das Datenformat kombiniert die Subframes aller acht Audiokanäle zu einem Frame, das zusätzlich 11 Sync-Bits und vier User-Bits enthält. Als Steckverbindung wird das von S/PDIF bekannte TosLink verwendet.

4.1.7 TDIF-1 und TDIF-2

Das Tascam Digital Interface wurde entwickelt, um acht Audiokanäle bidirektional zwischen gleichartigen Achtspur-Recordern des Tascam Digital Tape Recording Systems überspielen zu können, entwickelte sich jedoch bald zu einem universellen Interface für Produkte unterschiedlicher Hersteller.

Für die Schnittstelle werden 25-polige Sub-D-Steckverbinder verwendet, über die sowohl Eingänge als auch Ausgänge geführt werden. Die Daten werden ähnlich wie bei AES3 in vier Paaren übertragen. Die Leitungen sind jedoch unsymmetrisch.

Neben den Audiodaten werden Statusbits und ein Taktsignal übertragen, gleichzeitig wird der rechte und linke Kanal eines Kanalpaars identifiziert. Ein Datenwort ist hier 32 Bit lang, wobei die Audioinformation 24 Bit lang sein kann. Die Abtastfrequenz ist auf 48 kHz begrenzt. Bei Verdopplung reduziert sich die Kanalzahl auf die Hälfte. Die Spannungen liegen zwischen 2 und 3 Volt. Das Interface ist nur für kurze Kabellängen bis 5 m geeignet.

TDIF-2 wurde ein wenig stabiler entworfen und erlaubt, mit einem richtigen Kabel, Verbindungen bis etwa 10 m.



Abb 4-7: 25-Poliger Sub-D Stecker [music123.com]

4.1.8 Weitere digitale Schnittstellen

Natürlich gibt es noch weitere digitale Schnittstellen, von denen einige hier noch Erwähnung finden.

Y2 ist eine Schnittstelle der Firma Yamaha, die ursprünglich zur Kaskadierung von Mischpulten und Effektgeräten entwickelt wurde. Sie erlaubt in zwei Varianten die Übertragung von zweikanaligem (*Y2*) als auch mehrkanaligem (*MEL2*) Audio mit 24 Bit Auflösung.

Das *ProDigi*-Format ist eine Entwicklung von Mitsubishi. Jeweils 16 Spuren werden über eine 50-Polige Sub-D-Steckverbindung übertragen. Ein Datenwort ist hier 32 Bit lang, wobei nur 16 Bit für Audiodaten genutzt werden.

Das Interface der Firma Roland nennt sich *R-Bus*. Es werden acht Audiokanäle sowie einige Zusatzdaten parallel und bidirektional übertragen. Es sind Abtastraten bis zu 96 Hz möglich und eine Auflösung bis zu 24 Bit. Als Steckverbindung kommen 25-polige Sub-D-Verbindungen zum Einsatz

Das serielle digitale Interface (*SDI*) wird in der professionellen Videotechnik verwendet. Es erlaubt die Übertragung von linearen, nicht datenreduzierten Videosignalen, einer Vielzahl von Kontroll- und Steuerinformationen und insgesamt 16 Audiokanälen über ein koaxiales Kabel.

Interface	AES3	AES3-id	AES42	ADAT Lightpipe	MADI Electrical	MADI Optical	SDIF-2	TDIF
Anzahl der Kanäle	2	2	2	8	56 oder 64	56 oder 64	2 oder 24	8
Übliche Datenrate	6,144 Mbit/s	6,144 Mbit/s	6,144 Mbit/s	9,936 Mbit/s	98,304 Mbit/s	98,304 Mbit/s	1,536 Mbit/s	12,288 Mbit/s
Carrier-Signal	2-7 V	1 V	2-7 V	optisch	1 V	optisch	5 V	2-3 V
Impedanz	110 Ω	75 Ω	110 Ω	optisch	75 Ω	optisch	75 Ω	Quelle ~ 56 Ω Senke 140 – 1500 Ω
Steckverbindung	XLR	BNC	XLR	TOSLink	BNC	Dupelx Fibre Connector SC oder ST	BNC oder Sub-D 50	Sub-D 25
Leitungslänge	100 m	1000 m	100 m	5 m	50 m	500 – 2000 m	10 m	5 m
Anmerkung			Für digitale Mikrofone	Je nach Ausführung ADAT-Sync zusätzlich erforderlich			Getrennte Leitungen für Left, Right und Clock	Bidirektional

Tabelle 3-1: Verbreitete digitale Mehrkanal-Schnittstellen

4.2 Audio über Bussysteme

4.2.1 Protokolle und Routing

Übertragungsverfahren für Bus- und Netzwerkverbindungen verwenden meist paketorientierte Übertragungsprotokolle. Ein Datenpaket besteht, ähnlich wie die Audioframes der AES3-Schnittstelle, aus Kontrollinformationen und der eigentlichen Nutzlast. Typische Datenpakete enthalten einen Paketkopf mit Informationen zur Empfangsadresse, Sendeadresse, Paketlänge, Paketnummer und Kontrollbits. Die genaue Zusammensetzung des Paketkopfes kann ebenso wie die Paketlänge von Netz zu Netz variieren.

Die meisten Protokolle orientieren sich an den Regeln des ISO-OSI Schichtenmodells (Kapitel 4.3.1). Audionetze wie AudioRail, CobraNet und EtherSound verwenden entweder eigene oder adaptierte Übertragungsprotokolle.

4.2.2 Bussysteme

Ein Bus ist ein System zur Datenübertragung zwischen mehreren Teilnehmern über einen gemeinsamen Übertragungsweg, bei dem einzelne Teilnehmer nicht an der Datenübertragung zwischen anderer beteiligt sind.

Viele Studiokomponenten wie z. B. externe Soundkarten verwenden die Bus-Interfaces Firewire oder USB zur Signalübertragung. Sie erlauben die Verbindung mit Audio-Workstations auf Basis der Betriebssysteme Mac OS und Windows, wo sie in den allermeisten Fällen durch einfaches Anstecken (Plug & Play) betriebsbereit werden. Bei anderen Betriebssystemen wie Linux kann die Konfiguration von Firewire und USB aufwändiger sein. Auch bei MacOS und Windows tritt, ohne eine Installation von speziellen Treibern, vom externen Gerät über den Bus in den Rechner und wieder zurück, eine erhebliche Verzögerung (Latenz) auf. Hier müssen Low-Latency-Treiber wie ASIO 2, WDM Kernel-Streaming (PC) oder Core Audio (Mac) installiert werden, um die betriebssystemeigenen Treiber zu umgehen. Latenzzeiten unter 0,5 ms stellen wahrnehmungspsychologisch (z. B. für ein Playback über Kopfhörer) in der Praxis noch kein Problem dar. Wird jedoch das verzögerte Signal mit einem unverzögerten elektrisch oder akustisch gemischt wird, kommt es unweigerlich zu störenden Kammfiltereffekten (Phasing). Im Vergleich zu den bereits besprochenen Interfaces haben die Bus-Systeme wie Firewire und USB den Vorteil, dass Peripheriegeräte während des Computerbetriebs ein- und ausgesteckt werden können. Ein weiterer Vorteil ist, dass die Versorgungsspannung auf dem Bus liegt und viele externe Geräte ohne weitere Kabel auskommen.

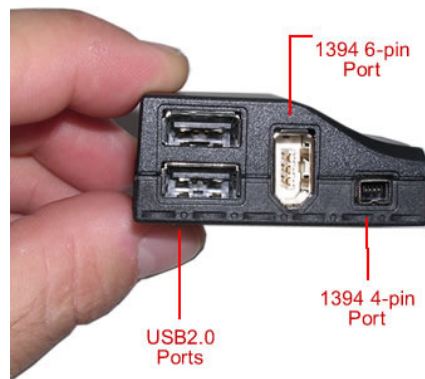


Abb 4-8: Frontansicht einer PCMCIA Karte mit Firewire- und USB 2.0-Anschlüssen [usbgear.com]

4.2.3 IEEE 1394 (Firewire)

Der Überbegriff IEEE 1394 oder Firewire bezeichnet mehrere herstellerspezifische, jedoch kompatible Busschnittstellen. Firewire stammt ursprünglich aus dem Hause Apple, i.Link ist die videoorientierte Version der Firma Sony und mLAN die audiospezifische Variante der Firma Yamaha. Am Rechner oder am externen Gerät erkennt man den Anschluss in der Regel am Schriftzug „1394“ oder an folgenden Symbolen:



Abb 4-9: Symbole für Firewire und i.Link

i.Link und Firewire unterscheiden sich im Wesentlichen durch die Steckerverbindung. Es existieren 4- und 6-polige Stecker, wobei die 6-poligen vornehmlich für Firewire und die 4-poligen für i.Link. Eine externe Stromversorgung ist nur mit den 6-poligen Steckern möglich. Firewire-Geräte mit 6-poligem Anschluss können an einem Sony i.Link-Interface benutzt werden, sofern sie über eine externe Stromversorgung verfügen.



Abb 4-10: 4- und 6-polige Steckervariante für eine IEEE 1394-Verbindung [kbsinc.com]

Mit Firewire können Datenraten von 400 MBit/s (IEEE 1394a oder Firewire 400) oder 800 MBit/s (IEEE 1394b oder Firewire 800) erreicht werden. Der neueste Standard IEEE 1394-2008 oder Firewire S3200 aus dem Jahr 2008 ermöglicht eine Datenübertragungsrate von 3,2

Gbit/s. Es können bis zu 63 Geräte an einen Bus angeschlossen werden, wobei die Kabellänge zwischen zwei Geräten im Allgemeinen auf 4,5 m begrenzt ist. .

Die Daten werden paketorientiert innerhalb eines Zyklus von 125 μ s übertragen und dafür in sogenannte Quadlets von 32 Bit unterteilt.

Firewire ist „audiofreundlich“. Das auf Yamahas mLAN-Standard basierende Audio and Music Data Transmission Protocol unterstützt sogar die Übertragung von datenreduziertem Mehrkanalton (Dolby Digital, DTS), Metadaten sowie die transparente Übertragung von Audiodaten nach IEC 60958 (S/PDIF).

Um mLAN-fähige Audiogeräte mit einem PC oder Mac zu verbinden, benötigt man einen OHCI-fähigen Firewire-Port am Rechner. OHCI steht für Open Host Controller Interface und beschreibt eine standardisierte Schnittstelle zur Kommunikation zwischen Firewire-Geräten.

Ausführliche Informationen und Spezifikationen über Firewire werden auf der Internetseite der 1394 Trade Association (<http://www.1394ta.org>, zuletzt aufgerufen am 23.11.2011) angegeben.

4.2.4 USB

Obwohl USB und Firewire nicht direkt miteinander verwandt sind, gibt es eine Reihe von Gemeinsamkeiten. Ebenso wie Firewire unterstützt USB Hot-Plug (Anschließen während des Betriebs) und verschiedene Datenraten. Der Datenaustausch zu Festplatten, Druckern oder Mischpulten ist ebenso möglich wie die Übertragung von Audio- und Videodaten (Streaming Media). Die Kommunikation über USB findet ähnlich wie bei dem OSI-Standard (Kapitel 4.3.1) auf mehreren Ebenen statt.

Von ursprünglich 1,5 MBit/s und 12 MBit/s unter USB 1.1 wurde die Datenrate mit USB 2.0 auf 480 MBit/s erhöht. USB 3.0 verfügt über eine Datenrate von bis zu 4000 MBit/s. Die Signalübertragung erfolgt über 4-polige Steckverbindungen, die sowohl Versorgungsspannung (5 V, max. 500 mA) als auch eine differentielle Datenleitung bereitstellen. Im Gegensatz zu Firewire gibt es keine eigene Clockleitung, der Takt ist im Datenbus integriert. Die einzelnen Datenpakete sind 1 ms lang, seit USB 2.0 gibt es aber auch Microframes mit 125 μ s. An der Spitze jedes USB-Verbundes steht der Hostadapter. Bis zu 127 Geräte können an einem USB-Host angeschlossen werden, die maximale Leitungslänge beträgt im Allgemeinen 5 m.

Kontaktbelegung

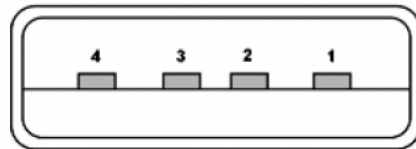
Pin 1 ... +5 Volt Versorgung

Pin 2 ... Daten -

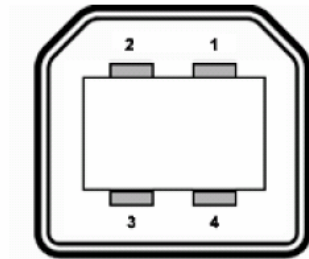
Pin 3 ... Daten +

Pin 4 ... Masse (GND)

Gehäuse ... Schirm



USB-Steckverbindung „A-Type“



USB-Steckverbindung „B-Type“

Abb 4-11: Pinbelegung und Ansicht der USB-Schnittstelle [WEINZIERL]

4.2.5 HDMI

Das High-Definition Multimedia Interface (HDMI) ist eine digitale Schnittstelle für multimediale Anwendungen im Heimbereich. Im Gegensatz zu Firewire oder USB benötigt sie keinen Hostrechner, sondern verfügt über eine eigene Intelligenz. Mit einer Bandbreite von bis zu 5 GBit/s lassen sich alle heute bekannten Video- und Audioformate digital und bei Bedarf ohne Datenreduktion über ein einziges Kabel transportieren. Unterstützt werden Standard- und High Definition Video, Steuersignale und Mehrkanalton bis zu acht Audiokanälen mit je 24 Bit und Abtastraten zwischen 32 und 192 kHz.



Abb 4-12: HDMI-Stecker [D-Kuru on Wikipedia]

4.3 Netzwerk Audio

Um preisgünstige, standardisierte Netzwerkkomponenten wie Router, Switches, Hubs und Interfaces verwenden zu können, entsprechen die grundlegenden Eigenschaften der allermeisten Audio-Netzwerke dem Ethernet-Standard IEEE 802.3. Dazu ist die Einhaltung der Bedingungen von ISO-Layer 1 (Physical) und ISO-Layer 2 (Data Link) aus dem ISO-OSI-Schichtenmodell erforderlich. Die Verwendung angepasster Protokolle in Audionetzwerken führt zu geringen Latenzzeiten von 5 μ s bis 1,6 ms.

4.3.1 Das ISO-OSI Schichtenmodell

Das Open Systems Interconnection Reference Model (OSI) ist ein offenes Schichtenmodell für die Kommunikation informationsverarbeitender Systeme. Es wurde 1979 vom internationalem Normungsgremium ISO (International Standardization Organisation) definiert und seither kontinuierlich weiterentwickelt. Das ISO-OSI-Modell definiert und vereinheitlicht Verfahren und Regeln für den Austausch digitaler Daten.

Das OSI-Modell definiert sieben Schichten oder Ebenen, so genannte ISO-Layer. Jede Schicht hat bestimmte Aufgaben, für die sie die darunter liegende Schicht als Transportmedium oder Hilfsmittel benutzt. Dadurch wird ein komplexer Kommunikationsprozess in kleinere, leichter überschaubare Teilprozesse zerlegt.

Die Schichten 1 bis 4 sind transportorientiert, die Schichten 5 bis 7 anwendungsorientiert.

Schicht 1: Bitübertragung (Physical Layer)

In dieser Schicht sind die physikalischen Parameter der Verbindung und deren fortgesetzte Betriebsbereitschaft festgelegt. Alle elektrischen Parameter wie Impedanz, Signalpegel und Datenrate, aber auch die Art der Steckverbinder sowie funktionale und prozedurale Parameter sind hier definiert.

Schicht 2: Sicherung (Data Link Layer)

Diese Ebene dient der Verbindungsherstellung zwischen Quelle und Senke bzw. Quelle und Netzwerk. Ihre Aufgaben liegen in der Synchronisierung einerseits, aber auch im Erkennen, Behandeln und Korrigieren etwaiger hardwarebedingter Fehler. Dazu werden sog. Übertragungsprotokolle eingesetzt.

Schicht 3: Vermittlung (Network Layer)

Diese Ebene ist unter anderem für Vermittlung, Adressierung und transparenten Datentransport verantwortlich. Bei leistungsorientierten Diensten sorgt die Vermittlungsschicht für das Schalten von Verbindungen, bei paketorientierten Diensten für die Weitervermittlung von Datenpaketen über mehrere Zwischenstationen. Auch Netzadressen gehören zu dieser Schicht.

Schicht 4: Transport (Transport Layer)

Die vierte Ebene garantiert zwei miteinander kommunizierenden Anwendungen eine transparente lückenlose und gesicherte Ende-zu-Ende-Kommunikation. Sie bietet den anwendungsorientierten Schichten 5 bis 7 einen einheitlichen Zugriff, so dass diese die Eigenschaften des Kommunikationsnetzes nicht beachten müssen. Die meisten Audionetzwerke verwenden hier spezielle, an Audioerfordernisse angepasste Protokolle.

Schicht 5: Sitzungsschicht (Session Layer)

Um kurzzeitige Unterbrechungen und ähnliche Probleme zu beheben, stellt diese Schicht Dienste für einen organisierten Datenaustausch zur Verfügung. Zu diesem Zweck werden Wiederaufsetzpunkte (Check Points) gesetzt, an denen die Sitzung nach einem Ausfall der Transportverbindung synchronisiert werden kann, ohne dass die Übertragung wieder von vorne beginnen muss. Dezentrierte Audionetzwerke wie CobraNet oder EtherSound umgehen die Schichten 5 bis 7, da sie stark latenzbehaftet sind.

Schicht 6: Darstellung (Presentation Layer)

Die Darstellungsschicht setzt die an sich systemabhängige Darstellung der Daten in eine unabhängige Form um und ermöglicht somit den sinnvollen Informationsaustausch zwischen unterschiedlichen Systemen. Hier wird gewährleistet, dass Daten, die von der Anwendungsschicht eines Systems gesendet werden, von der Anwendungsschicht eines anderen Systems gelesen werden können. Falls erforderlich, agiert die Darstellungsschicht als Übersetzer zwischen verschiedenen Dateiformaten, indem sie ein für beide Systeme verständliches Datenformat verwendet.

Schicht 7: Anwendung (Application Layer)

Hier werden für unterschiedliche Anwendungen eine Vielzahl an Funktionalitäten zur Verfügung gestellt. Bei Audionetzwerken wie CobraNet oder EtherSound kommen spezielle Application Layer zum Einsatz.

4.3.2 Ethernet

Ethernet definiert bestimmte Kabeltypen (Cat5, Cat6), eine bestimmte Signalisierung für die Bitübertragung sowie Paketformate und Protokolle für die Steuerung des Medienzugriffs und die Sicherungsschicht des OSI-Schichtenmodells. Ethernet ist seit Beginn der 1990er-Jahre die meistverwendete Netzwerktechnologie und nach IEEE 802.3 genormt. Es kann die Basis für verschiedene Netzwerkprotokolle wie etwa TCP/IP oder spezielle Audionetzwerkprotokolle sein. Um Kollisionen bei der Datenübertragung zu vermeiden, kommt ein Verfahren namens Carrier Sense Multiple Access With Collision Detection (CSMA/CD) zum Einsatz. Die Übertragungsrate beträgt je nach Ausführung 10 MBit/s (10Base-T), 100 MBit/s (100Base-T) und 1000 MBit/s (1000Base-T, Gigabit-Ethernet)

Als elektronisches Signal kommt bei Ethernet mit 100 MBit/s Multilevel Transmission Encoding mit drei Pegeln (MLT3) mit -1V, 0V und +1V zum Einsatz. Der Wellenwiderstand des vierpaarigen Kabels beträgt 100 Ω . Als maximale Entfernung werden meist 100m zwischen einzelnen Stationen angegeben.



Abb 4-13: Cat5 Stecker

4.3.3 AudioRail

AudioRail ist eine Audio-Netzwerklösung mit Latenzzeiten in der Größenordnung von etwa $4,5 \mu\text{s}$. Pro Teilnehmer wird diese Zeit um etwa $0,25 \mu\text{s}$ verlängert. Diese äußerst geringe Latenz wird durch eine einfache Datenstruktur ermöglicht. Die physikalische Übertragung erfolgt mit einem 4 Bit breiten, 25 MHz schnellen Datenstrom im Zeitmultiplex-Verfahren.

Insgesamt können über AudioRail 32 Audiokanäle mit jeweils 24 Bit und 48 kHz Abtastfrequenz bidirektional übertragen werden, bei 96 kHz halbiert sich die Zahl auf 16. Im Gegensatz zu den meisten anderen Lösungen kann jedes Audiosignal mit seiner nativen Taktfrequenz übertragen werden. Das erlaubt es, unterschiedlich getaktete Signale ohne gemeinsame Zeitreferenz über das gleiche Netz zu senden. Somit entspricht AudioRail einem Bündel getrennter Audioleitungen, die über ein gemeinsames Netzkabel transportiert werden. Wenn jedoch alle Quellsignale auf den gleichen Takt bezogen sind, wie z. B. die digitalen Ausgänge eines Mehrfach-Mikrofonvorverstärkers oder einer Audioworkstation, erfolgt die Übertragung starr verkoppelt. AudioRail kann derzeit keine Steuer- und MIDI-Daten übertragen, wie sie etwa zur Fernbedienung abgesetzter Mikrofon-Vorverstärker benötigt werden.



Abb 4-14: AudioRail: 32 Audiokanäle über ein Netzkabel nach Cat.5 Spezifikation (Foto: Audiorail)

4.3.4 CobraNet

CobraNet wurde von Peak Audio, einer Abteilung des Chip-Herstellers Cirrus Logic entwickelt. Es basiert auf Standard-Ethernet nach IEEE 802.3, verwendet das Ethernet-Protokoll MAC (Media Access Control) sowie eine eigenes Anwendungsprotokoll auf ISO-Layer 7 (Application Layer). IP (Internet Protokoll) und UDP (User Datagram Protokoll)

kommen bei der Audioübertragung nicht zum Einsatz, stehen jedoch für andere Zwecke zur Verfügung.

Mit CobraNet ist es möglich, einen Datenverkehr mit einer Bandbreite von 100 MBit/s zu erzielen. Die besondere Struktur des CobraNet-Protokolls nutzt über 90 % der möglichen Bandbreite für den Datentransfer, wohingegen in normalen Netzwerken nur 30 % bis 40 % für die eigentliche Nutzlast zur Verfügung stehen. Bis zu 64 Audiokanäle mit je 20 Bit und 48 kHz (56 Kanäle bei 24 Bit) können über ein Cat.5-Kabel übertragen werden. Auch Samplingfrequenzen von 96 kHz sind möglich, 44,1 kHz wird jedoch nicht unterstützt.

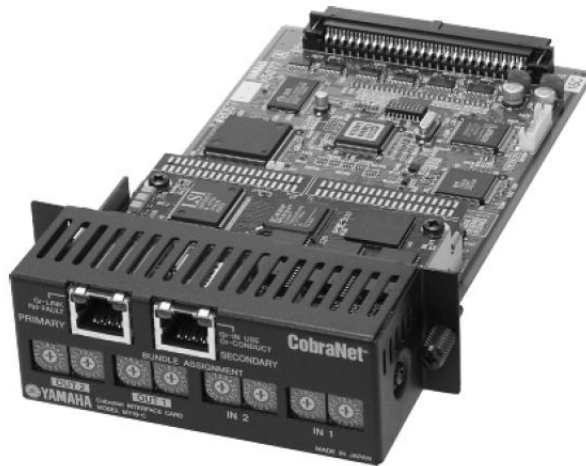


Abb 4-15 CobraNet-Interface für digitale Audiomischpulte (Foto: Yamaha Corporation)

Die Latenzzeit von einem Audioeingang über das Netzwerk zu einem anderen Audioausgang beträgt vier Zyklen zu je 1,33 ms, in Summe 5,33 ms.

Neben reinen Audiosignalen lassen sich hier auch MIDI-Daten und Steuersignale zur Fernbedienung diverser Audiogeräte übertragen.

4.3.5 EtherSound

Auch das EtherSound-Protokoll der Firma Digigram ist auf physikalischer Ebene Ethernet-kompatibel und verwendet auch das MAC-Protokoll. Der Vorzug von EtherSound ist eine kurze Latenzzeit von 125µs bei 48 kHz Abtastrate. Das Einfügen von zusätzlichen Teilnehmern im Signalweg verlängert die Latenzzeit um jeweils 1,5µs. Das Datenformat ist paketorientiert, wobei zwei Pakete pro Frame zum Einsatz kommen. Paket 1 enthält Steuerinformationen, vergleichbar zu den Channel Status Informationen nach AES3. Das zweite Paket enthält die Daten von bis zu 64 Audio-Kanälen. Jeder Kanal kann Wortbreiten von bis zu 24 Bit und Abtastraten zwischen 44,1 oder 48 kHz transportieren. Abtastraten von 88,2 kHz, 96 kHz, 176,4 kHz und 192 kHz werden ebenfalls unterstützt, jedoch reduziert sich dadurch die Anzahl der maximalen Audiokanäle. Ein EtherSound-Netzwerk mit 100 MBit/s kann eine Vielzahl unterschiedlicher Audiostreams übertragen: 64 Kanäle mit 48 kHz oder 62 Kanäle mit 48 kHz plus einen Audiostream mit 96 kHz bis hin zu 16 Kanälen mit 192 kHz.

4.4 Prozessoren

Zur Verarbeitung digitaler Audiosignale können sowohl Mikroprozessoren zum Einsatz kommen, wie sie als zentrale Recheneinheit (Central Processing Unit, CPU) in handelsüblichen PCs verwendet werden, als auch Signalprozessoren (Digital Signal Processors, DSPs). Die Unterschiede zwischen beiden Prozessortypen liegen in ihrem prinzipiellen Aufbau (Architektur), in der Organisation des internen Datenflusses, in der Art der Ein- und Ausgabeschchnittstellen und im Umfang des für die Programmierung zur Verfügung stehenden Befehlssatzes.

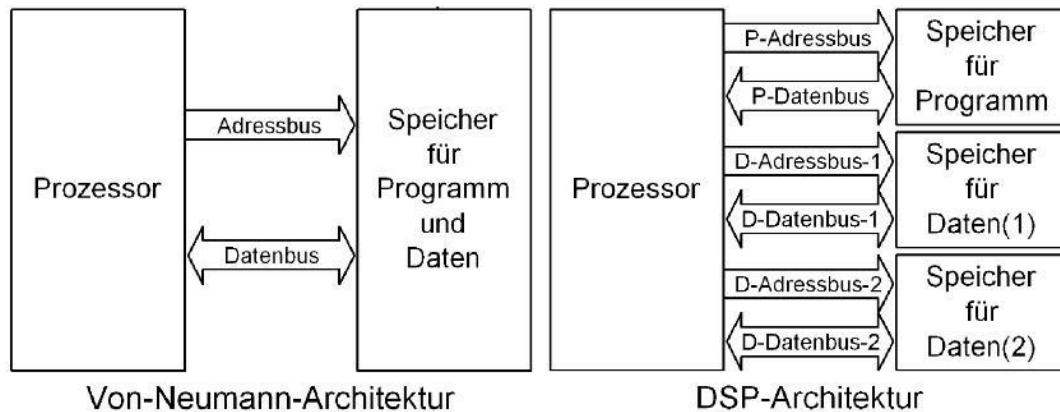


Abb 4-16: Von-Neumann- und DSP-Architektur von Prozessoren [WEINZIERL]

Mikroprozessoren sind in der Regel nach einer sogenannten Von-Neumann-Architektur aufgebaut, die einen gemeinsamen Bus und einen gemeinsamen Speicher für Programmcode und Daten vorsieht. Signalprozessoren dagegen verfügen über getrennte Busse und getrennte Speicher für Programme und Daten. Dieses Grundprinzip findet sich auch bei modernen Prozessoren mit zunehmend komplexer Architektur wieder.

Digitale Filteroperationen als typische Anwendung der digitalen Signalverarbeitung bestehen im innersten Kern aus der Multiplikation eines Signalabtwertes mit einem Koeffizienten, der Aufsummierung dieser Produkte sowie der Beschaffung der nächsten Signalwerte und Koeffizienten. Während diese Operationen auf Mikroprozessoren sequenziell ausgeführt werden müssen, können sie auf Signalprozessoren parallel und meist in einem Taktzyklus ausgeführt werden. Damit sind Signalprozessoren insbesondere für die Echtzeitverarbeitung großer Datenmengen prädestiniert, wie sie in zeitkritischen Anwendungsfeldern wie der Audio- und Bildverarbeitung sowie in der Regelungs-, Übertragungs- und Medizintechnik anfallen. Zu der Zeit, als Mikroprozessoren mit nur wenigen MHz getaktet waren, verfügten nur die um 1980 eingeführten Signalprozessoren über eine ausreichende Rechenleistung zur Echtzeitverarbeitung digitaler Audiosignale. Heutige Mikroprozessoren haben allerdings mit Taktfrequenzen von über 3 GHz Signalprozessoren mit Taktfrequenzen von unter 1 GHz überholt. Somit können auch auf den Prozessoren moderner PCs heute große Datenraten in Echtzeit verarbeitet werden, während Signalprozessoren überwiegend in spezialisierter Hardware wie Mischpulten, Effektgeräten, DSP-basierten Audioworkstations oder Digitalcontrollern für Lautsprecher eingesetzt werden.

Neben der taktfrequenzbezogenen Rechenleistung ist auch die Transferleistung in und aus einem Prozessor von Bedeutung, damit dessen Hochgeschwindigkeits-Parallelrechenwerke

vollständig genutzt werden können. Signalprozessoren besitzen hier meistens spezielle Hochgeschwindigkeitsschnittstellen, die Mikroprozessoren in der Regel fehlen.

Große Signalverarbeitungssysteme wie Mischpulte oder Workstations kommen kaum mit einem einzigen Prozessor aus. Ein Multiprozessorsystem verhält sich wie ein vielfach verästeltes Förderbandsystem, über das Elemente (Audiosamples oder –blöcke) von Verarbeitungstation zu Verarbeitungstation weitergereicht werden. Hier darf sich weder etwas stauen noch verloren gehen. Die Kommunikation dieser vielen Prozessoren stellt einen nicht unerheblichen Teil eines Gesamtsystems dar – so können im Signalverarbeitungskern eines modernen Digital-Mischpultes weit über 100 Signalprozessoren zusammenarbeiten.

Digitale Signalprozessoren werden natürlich nicht nur für Audioanwendungen verwendet, wie folgende Grafik verdeutlicht.

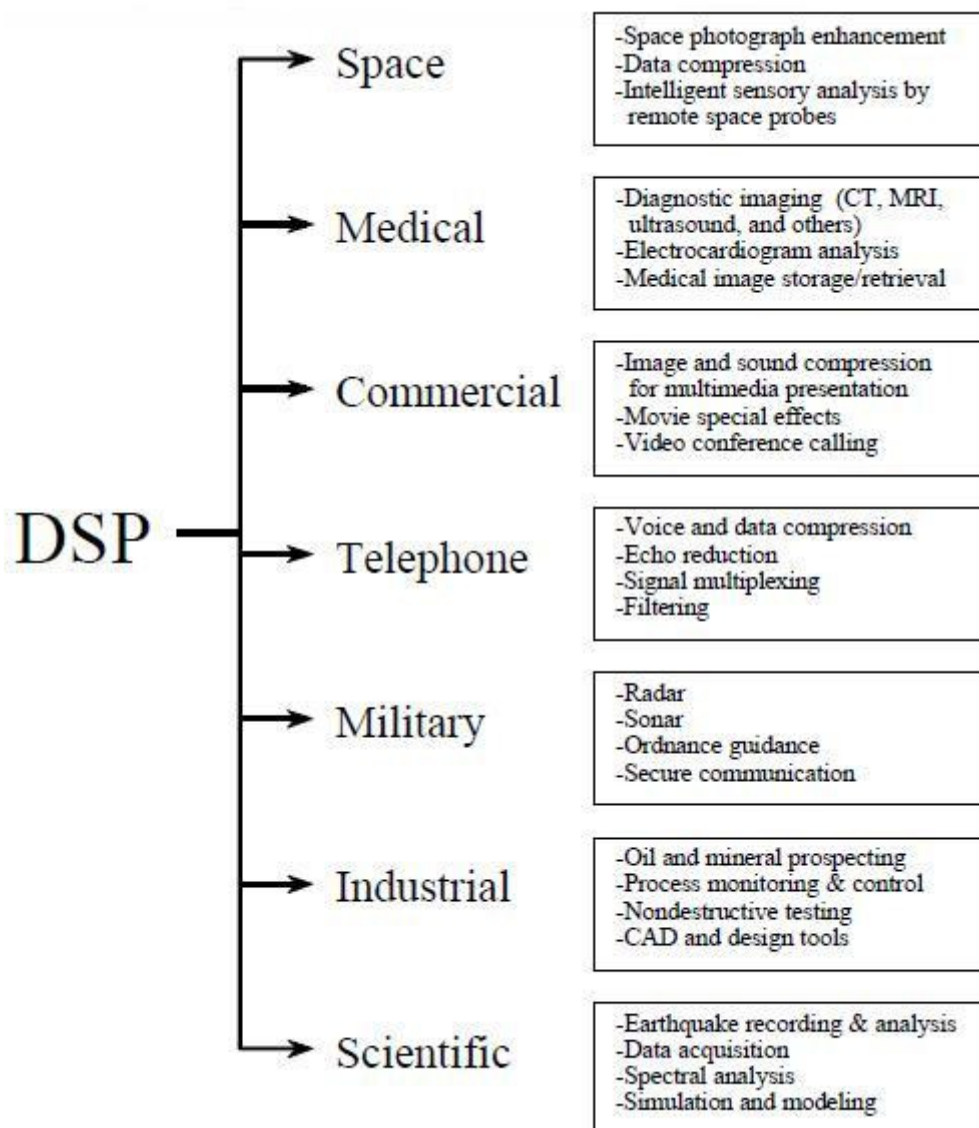


Abb 4-17: Einsatzbereiche digitaler Signalprozessoren [SMITH]

4.5 Digitale Pegel

Digitale Pegel werden in dBFS angegeben. dBFS steht für Dezibel Full Scale. 0 dBFS stellt den höchstmöglichen Pegelwert steht für einen digitalen Wert in dem alle Bits den Wert 1 besitzen. Der niedrigste Wert ist, z. B. für 16-Bit Audiosignal *0000 0000 0000 0001*, was einem Wert von -96 dB entspricht.

$$20 \cdot \log\left(\frac{1}{2^{16}}\right) = -96,33 \text{dBFS}$$

Der Dynamikbereich für ein 16Bit-Signal beträgt somit 96dB, für ein 20 Bit-Signal 120dB und bei 24 Bit 144 dB.

An diesem Beispiel erkennt man, dass sich digitale Pegel nicht in analoge Pegel umrechnen lassen. Es wurde oft versucht einen bestimmten dBu-Wert dem Wert 0 dBFS zuzuordnen, aber die Standards unterscheiden sich sogar innerhalb Europas, wobei dort meist die EBU Empfehlung R58-2000 verwendet wird, die einem Wert von +18 dBu den Wert 0 dBFS zuordnet.

Beim deutschen Funkhausnormalpegel dBr gilt hingegen:

$$+6\text{dBu} = 0\text{dBr} = -9\text{dBFS}$$

-9dBFS ist hier eine virtuelle, vereinbarte Vollaussteuerungsgrenze, die einen *Headroom* von 9dBFS zulässt, um Clipping bei unerwarteten Signalspitzen auszuschließen. Diese Empfehlung wird aber oft ignoriert, vor allen Dingen dann, wenn sich Limiter und Gates im Signalweg befinden, die verhindern, dass solche unerwarteten Signalspitzen auftreten. Dann würde das Audiosignal unnötig leiser sein.

Ein Headroom schützt jedoch nicht nur vor unerwarteten Pegelspitzen, sondern auch durch ungünstige interne Berechnungen in der Audio-Software können Fehler entstehen. So ist es beispielsweise möglich, dass es bei der Normalisierung eines Signals zu digitalen Übersteuerungen kommt, wenn die digitalen Audiosamples nicht an den Pegelspitzen liegen, und das Audiosignal dann auf 0 dBFS normalisiert wird. Die Abbildungen 4-18 bis 4-20 verdeutlichen diesen Effekt. Dasselbe Problem ergibt sich bei der Anzeige von digitalen Audiowerten, wenn dort Maximalwerte digital nicht angezeigt werden. Somit kann ein Sicherheitsabstand zur digitalen Vollaussteuerung sinnvoll sein. Diese ist jedoch frei wählbar. Oft wird ein Sicherheitsabstand von 1 bis 3 dBFS empfohlen.

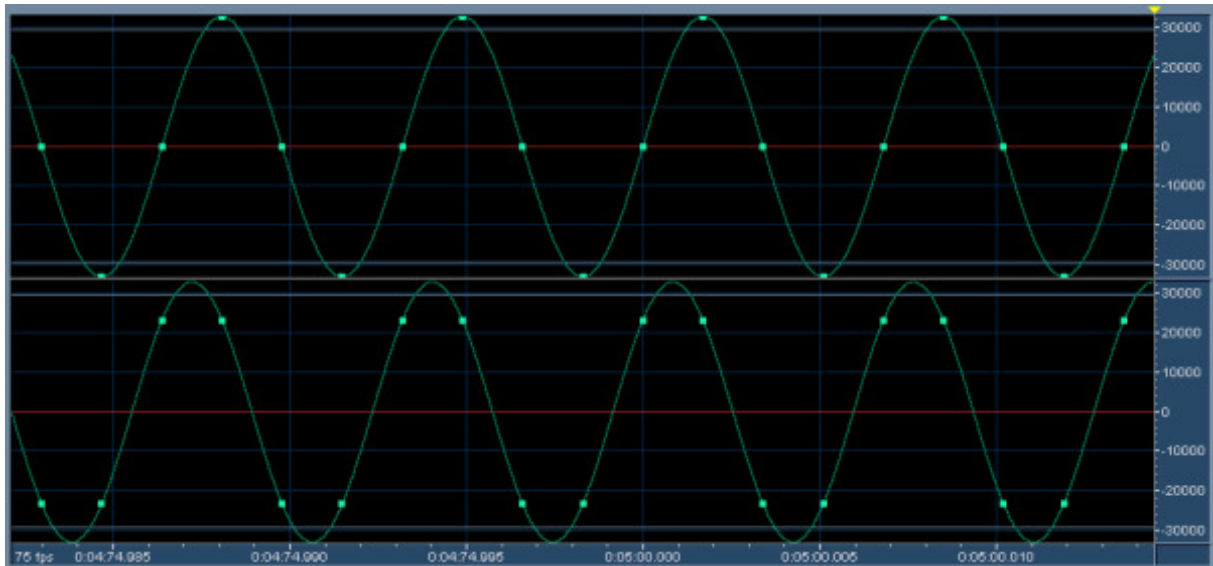


Abb 4-18: Die grünen Punkte stellen die Abtastsamples eines Sinus-Signals dar. Der untere Sinus-Ton ist zum oberen Signal um 45° verschoben [SCHUBERT]



Abb 4-19: Pegelanzeige eines DAT-Recorders beim Abspielen des Signals aus Abb 4-11. [SCHUBERT]

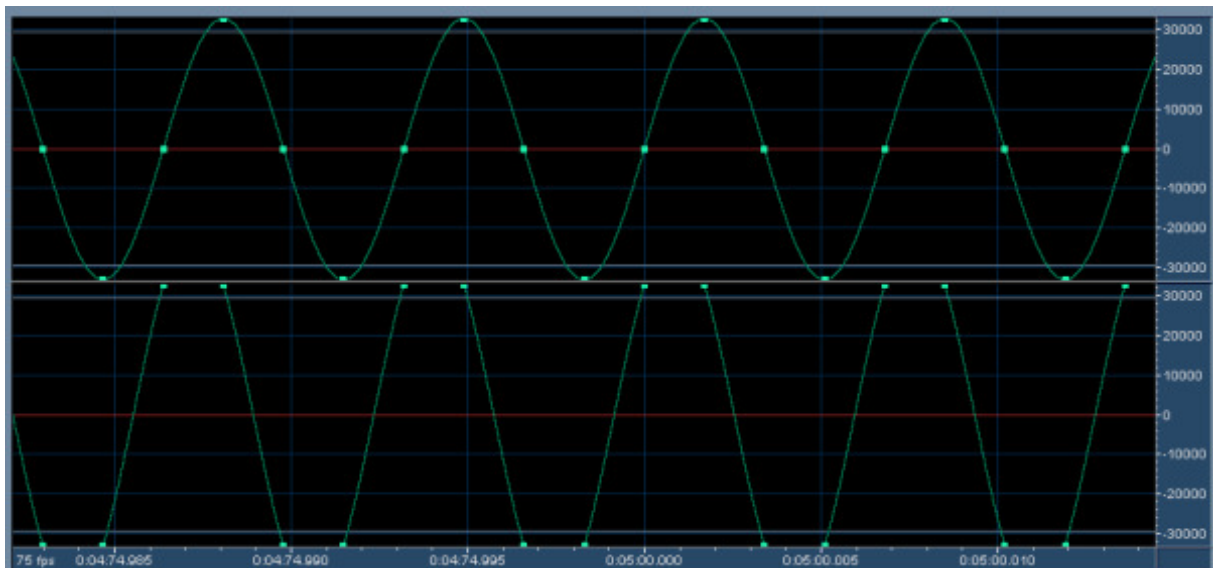


Abb 4-20: Signal, nachdem beide Kanäle getrennt auf den Maximal-Pegel normalisiert worden. [SCHUBERT]

4.6 Mobile Aufnahmegeräte und Peripheriegeräte mit Beispielen

Im Bereich der digitalen Studiotechnik gibt es eine unübersichtliche Vielzahl von digital verarbeiteten Audiogeräten. Die folgenden Kapitel enthalten eine grobe Marktübersicht über derzeit gängige Soundkarten, mobile Aufnahmegeräte und digitale Effektgeräte. Es werden jeweils drei Geräte unterschiedlicher Preis-, Qualitäts- oder Verwendungsklassen vorgestellt, die jedoch keinesfalls als Empfehlung missverstanden werden dürfen. Sie sollen nur den derzeitigen Stand der Technik repräsentieren.

Im Internet kann man sich auf den Seiten der Hersteller und Online-Musikshops einen guten und aktuellen Überblick verschaffen und die persönlichen Präferenzen berücksichtigen.

4.6.1 Soundkarten

Digitale Audiointerfaces unterscheiden sich zunächst mal in der Art der Schnittstelle. So gibt es PCI-Karten oder PCMCIA-Karten zum Einbau in den Computer, sowie externe Geräte mit z. B. USB- oder FIREWIRE-Schnittstelle. Des Weiteren gibt es zahllose Interfaces mit anwendungsspezifischen Funktionen (z. B. für DJs oder ein im Mischpult integriertes Interface).

RME HDSP 9632

Das HDSP 9632-PCI-Interface von RME ist eine vielseitige Soundkarte zum Einbau in den PC. Alle maximal Ein- und Ausgänge sind gleichzeitig nutzbar. Mit einem Preis von ca. 330 Euro liegt die Soundkarte im höheren Preissegment, wenn man bedenkt, dass sehr einfache PCI-Soundkarten schon ab ca. 10 Euro erhältlich sind. Es gibt auf dem Markt aber auch Produkte dieser Kategorie, die weit über 1000 Euro kosten.

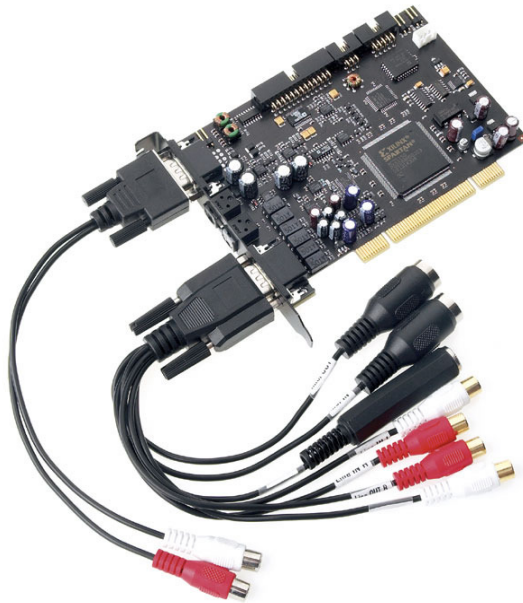


Abb 4-21: RME HDSP 9632 [music-service.de]

- Stereo Analog Ein- und Ausgang (192 kHz/24 Bit)
- 1 x ADAT Digital In/Out, 96-kHz-fähig
- 1 x SPDIF Digital In/Out, 192-kHz-fähig
- 1 x Stereo Kopfhörerausgang
- 1 x MIDI I/O für 16 Kanäle
- Mitgelieferte Software: DIGICheck (Meter- und Analysetool), HDSP Meter Bridge (Levelmeter mit Peak- und RMS Berechnung), TotalMix (512-Kanal Mischer)

M-AUDIO FAST TRACK PRO

Die FAST TRACK PRO ist eine einfache externe USB-Soundkarte der Firma M-Audio. Mit einem Preis von ca. 150 Euro ist sie eher im unteren Preissegment anzuordnen, in dem es jedoch recht viele Angebote gibt. Natürlich kann man auch hier weit über 1000 Euro investieren.



Abb 4-22: M-AUDIO FAST TRACK PRO [kwick.de]

- Stereobetrieb bis 96 kHz / 24 Bit
- 2 Mic/Instrumenten-Vorverstärker
- 2 x Insert
- 2 x Line Out plus 4 zusätzliche Outs über Cinch
- S/PDIF coaxial In/Output. S/PDIF-Ausgang auch für AC3/DTS Surround.
- 6,3mm Kopfhörerausgang mit Lautstärkeregler und A/B-Quellenschalter
- MIDI In/Out

MOTU 8 PRE

Das Firewire-Interface 8 Pre der Firma MOTU erlaubt eine gleichzeitige Aufnahme von bis zu 16 Kanälen (8 Kanäle über interne Wandler, 8 Kanäle über ADAT). Der Preis beträgt derzeit ca. 500 Euro.



Abb 4-23: MOTU 8PRE (Vorder- und Rückansicht) [av-versand.de]

- 8 Mikrofonvorverstärker (XLR/Klinke)
- ADAT In/Out (96 kHz fähig)
- MIDI In/Out
- 2x Analog Out (6,3 mm Klinke)
- Kopfhörer Out
- Bis zu 4 Geräte kaskadierbar
- Format: 19“, 1 HE
- Software: CueMix™ (Mixer für alle 16 Kanäle)

4.6.2 Mobile Aufnahmegeräte

Im Bereich der mobilen Aufnahme hat sich in den letzten Jahren sehr viel verändert. So sind einfach zu handhabende Geräte auf dem Markt, die in einem Preissegment liegen, das es

jedem ermöglicht, Aufnahmen in guter Qualität zu erstellen. Diese Geräte zeichnen sich auch dadurch aus, dass sie mit Batterie betrieben werden können und somit überall einsetzbar sind.

Zoom H2

Mobile digitale Aufnahmegeräte sind im Handel bereit ab ca. 30 Euro erhältlich. Der Zoom H2 ist ein weit verbreitetes Gerät und befindet sich mit einem Preis von ca. 150 Euro im unteren Preissegment. Übliche Einsatzbereiche sind Aufnahme von Interviews, Live-Mitschnitte, Proberaumaufnahmen, etc.



Abb 4-24: Zoom H2 [audiotranskription.de]

- gleichzeitige Aufnahme von 2 oder 4 Kanälen
- 2 eingebaute Mikrofone
- Aufnahme im mp3-Format (bis zu 329 kbps) oder wav (24 Bit/96 kHz)
- Ein zusätzlicher Mikrofon/Line-Eingang
- Kopfhörer-Anschluss/Line-Out-Ausgang
- USB-Schnittstelle
- Speicherung auf SD-Karten (bis zu 32 GB)
- Benötigte Batterien (2 mal AA)
- Bis zu 4 Stunden Aufnahme im Batteriebetrieb
- Maße (LxBxH): 63,5 x 110 x 32 mm
- Gewicht: 110 g ohne Batterien
- Zusätzliche Funktionen: Normalisierungsfunktion, Tiefpassfilter, Kompressor, Limiter, Chromatic Tuner und Metronom

Edirol R-44

Der Edirol R-44 Feldrecorder ist mit ca. 825 Euro erheblich teurer, leistet aber auch deutlich mehr. Dieses Gerät wird im professionellen Bereichen eingesetzt. Ein Exemplar befindet sich auch im Tonbüro der Studienvertretung Elektrotechnik-Toningenieur an der TU Graz!



Abb 4-25: Edirol R-44 [thomann.de]

- Parallele Aufnahme von 4 Spuren
- Eingebautes Stereomikrofon und Lautsprecher
- Unkomprimierte Aufnahme (bis zu 24 Bit / 192 kHz)
- 4 analoge XLR/Klinke-Eingänge
- 4 analoge Cinch-Ausgänge
- Digitaler Ein- und Ausgang (Cinch)
- USB-Schnittstelle
- Aufnahme auf SD-Karte (bis zu 32 GB)
- Benötigte Batterien (4 mal AA)
- Bis zu 4 Stunden Aufnahmezeit bei Aufnahme mit 16 Bit/44,1 kHz
- Maße (LxBxH): 157 x 183 x 61 mm
- Gewicht: 1,3 kg (inkl. Batterien)
- Zusätzliche Funktionen: zwei Geräte kaskadierbar zur parallelen Aufnahme von 8 Spuren, Limiter, Tiefpassfilter, div. Effekte

TASCAM HS-P82

Der Tascam HS-P82 Fieldrecorder wurde für Audioaufnahmen in Film und Fernsehen entwickelt, die im Freien gemacht werden müssen. Hier handelt es sich um professionelles Equipment auf hohem Preisniveau (ca. 4900 Euro).



Abb 4-26: Tascam HS-P82 [dv247.com]

- 8 Aufnahmespuren und 2 Masterspuren
- Aufnahme mit bis zu 24 Bit/192 kHz (bei paralleler Aufnahme von 4 Spuren)
- 8 XLR Eingänge
- 8 AES42 Ein- und Ausgänge (DB-25-Steckverbinder)
- 2 XLR Ausgänge
- Stereo-Digitalausgang (BNC)
- Kopfhöerausgang (6,3 mm Stereoklinke)
- PS2-Anschluss für Computertastatur
- USB-Schnittstelle
- Speichermedium: Compact-Flash-Karten
- Benötigte Batterien (10 mal AA)
- 2 bis 5 Stunden Aufnahmezeit
- Maße (LxBxH): 270 x 100 x 260 mm
- Gewicht: 3,65 kg (ohne Batterien)
- Zusätzliche Funktionen: herausklappbarer TFT-Touchscreen

4.6.3 Digitale Effektgeräte

Auf dem Markt existieren sehr viele externe digitale Effektgeräte. Im Studio werden diese oft durch Software ersetzt, die sich direkt auf dem Rechner befindet.

Alesis Nanoverb 2

Der Nanoverb 2 von der Firma Alesis ist ein kleiner einfacher Effektprozessor im niedrigen Preissegment (ca. 100 Euro) und wurde sowohl für den Studio als auch für den Liveeinsatz entwickelt.



Abb 4-27: Alesis Nanoverb 2 (Vorder- und Rückansicht) [rekkerd.org, wikizic.org]

- 256 Presets unterteilt in 16 Effektprogramme je 16 Variationen
- Effekte: Tremolo, Chorus, Flange, Delay, Plate, Tape, Spring, Halls und Mischprogramme
- Pegel für Ein-/Ausgang regelbar
- 6,3 mm Klinkenanschlüsse

TC Electronic Fireworx

Dieser Multieffekt-Prozessor befindet sich mit einem Preis von ca. 770 Euro im mittlerem Preissegment und stellt sehr viele Möglichkeiten und Anschlüsse bereit.



Abb 4-28: TC Electronic Fireworx [audiofanzine.com]

- 24 Bit AD/DA-Wandler
- AES/EBU, S/PDIF und ADAT Digital Ein- und Ausgänge
- 2 Analoge XLR-Inputs

- Formate: AES/EBU (24 Bit), S/PDIF (20 Bit), Toslink, ADAT Lite pipe, IEC 958, EIAJ CP-340
- Sample Raten, 44,1 kHz; 48 kHz
- Sehr viele Effekte: u. a. Expander/Gate, Kompressoren, Filter, Phaser, Distortion, Vocoder, Synth, Ring Modulator, Chorus, Flanger, Delay, Reverb, Tremolo, Panner, Enhancer, Equalizer

Lexicon PCM96

Der Lexicon PCM96 ist ein professioneller Multieffektprozessor für Studio- und Liveanwendung mit hohem Preisniveau (ca. 3200 Euro).



Abb 4-29: Lexicon PCM96 [djwarehouse.com.au]

- 2 x AES/EBU In/Out
- 2 x Analog XLR In /Out
- MIDI In/Out/Thru
- Dual-Ethernet- und Firewire-Anschlüsse
- 24 Bit DA-Wandler
- Sample Raten: 44,1, 48, 88,2 und 96 kHz
- Reverb, Modulations- und Delay-Effekte, Stereo-Effekte,

5 Digitale Speicherung

Die digitale Speicherung von Daten stellt immer ein Problem dar, weil es eine Vielzahl von Möglichkeiten gibt, wie und auf welchem Medium das Signal gespeichert werden kann. Das führt dazu, dass Medien und Technologien veralten und Dateien nicht mehr lesbar werden. So ist z. B. die 3,5“ Diskette, die in den 1990er-Jahren sehr populär war für die meisten Anwender nicht mehr lesbar, weil die entsprechenden Geräte aus den Haushalten verschwunden sind. Die CD, die sich immerhin gut 30 Jahre auf dem Markt gehalten hat, wird in aktueller Literatur mitunter schon als „historisches Speichermedium“ bezeichnet. Im professionellen Audiobereich wurde viel mit ADAT gearbeitet, im Consumer-Bereich hatte die MiniDisc über einen gewissen Zeitraum hinweg einen großen Erfolg.

Abgesehen von der technischen Entwicklung gibt es auch wirtschaftliche Hindernisse. Die großen Hersteller von Speichermedien haben oft Probleme sich auf einheitliche Formate zu einigen. So gibt es immer wieder unterschiedliche Medien und Formate von unterschiedlichen Herstellern mit dazugehörigen Abspiel- und Aufnahmegeräten, die untereinander oft nicht kompatibel sind. Es gibt sehr viele moderne Formate die heute einen gewissen Bekanntheitsgrad haben. CD, DVD, Blu-ray Disc, HD-Festplatten, SSD, SD-Karten, Compact-Flash, etc.

Die Verantwortung für eine sichere und nachhaltige Speicherung liegt somit beim Anwender. Um digitale Daten über lange Zeiträume zu sichern, müssen diese immer wieder auf ein aktuelles Medium überspielt werden, da sonst die Wahrscheinlichkeit eines Datenverlustes sehr hoch ist.

5.1 Die Audio-CD

Die Audio CD wurde 1982 von Phillips und Sony auf den Markt gebracht und hat sich als digitales Musikspeicherformat im Konsumentenmarkt über 30 Jahre hinweg durchgesetzt. Die Audiodaten werden mit einer Sampling-Frequenz von 44,1 kHz und einer Quantisierung von 16 Bit Stereo gespeichert. Damit wird ein Signal-Rauschabstand von 97,8 dB erreicht.

Auf der CD befinden sich die digitalen Informationen auf einer von innen nach außen verlaufenden, spiralförmigen Spur als Folge von Vertiefungen, den sogenannten Pits. Die Flächen zwischen den Pits werden Lands genannt. Die Pits sind 0,6 μm breit und 0,12 μm tief. Der Abstand zwischen den Spuren beträgt 1,6 μm . Somit kommen auf einen Millimeter Plattenradius rund 600 Spuren, insgesamt ist die Spur ca. 5,7 km lang. Die Umdrehungsgeschwindigkeit ist nicht konstant. Sie wird so gesteuert, dass eine konstante

Geschwindigkeit der Spuren relativ zum Abtaster erreicht wird und beträgt bei der Abtastung zwischen 500 und 200 Umdrehungen pro Minute. Der Datenfluss ist damit konstant und beträgt 4,3 MBit/s

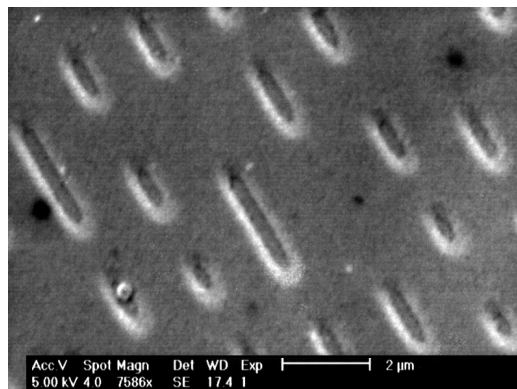


Abb 5-1: Aufnahme einer CD mit einem Raster-Elektronen-Mikroskop [bigdaddy1234 on Wikipedia]

Abgetastet wird eine CD mit einem Laserstrahl, der von der Plattenoberfläche unterschiedlich reflektiert wird, je nachdem, ob eine Vertiefung vorhanden ist oder nicht. Diese Unterschiede werden ausgewertet und zum Digital/Analog-Wandler weitergegeben. Da sich die Spuren mit konstanter Geschwindigkeit am Abtaster vorbeibewegen, werden sie mit einer ebenfalls konstanten Frequenz abgetastet. Auszuwerten ist dabei, ob die Flanke einer Vertiefung vorhanden ist oder nicht. Jede solche Flanke steht für ein gesetztes Bit, also den Binärwert 1. Ist keine Flanke vorhanden, ist das Bit nicht gesetzt, also 0.

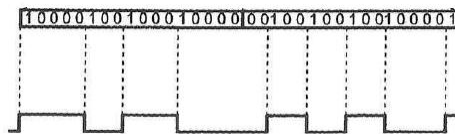


Abb 5-2: Pits und Lands mit interpretiertem Code [WEBERS]

Der Datenstrom auf einer CD wird seriell ausgelesen. Ein Datenrahmen (Frame) auf einer CD enthält 6 Samples zu je 2×16 Bit also $192 \text{ Bit} = 24 \text{ Byte}$ Audioinformationen. Zusätzlich enthält ein solcher Frame noch 8 Byte zur Fehlererkennung und Korrigierung (Cross Interleaved Read Salomon Code oder kurz CIRC), ein Byte für einen Control & Display Code (C&D-Code), in dem Informationen wie Musik an/aus oder Tracknummer gespeichert sind. Jedes der nun 33 Byte wird mit 17 Kanalbits zur Lesesicherheit (eight-to-fourteen-modulation) kodiert. Letztlich wird dem Frame noch ein 27 Bit großes Synchronisations-Muster (sync pattern) hinzugefügt. Ein Frame besteht demnach aus 588 Bit. Pro Sekunde werden 7350 Frames ausgelesen.

$$H = \frac{f_a \cdot n}{T_a} = \frac{44100 \cdot 588}{6} = 4,3218 \text{ Mbit/s}$$

Anzahl Frames/Sekunde: $44100/6 = 7350 \text{ Frames/s}$

5.2 Digitale Speicherformate

Ein Dateiformat definiert die Anordnung und die Beschreibung von Werten innerhalb einer Datei. Es gibt eine Vielzahl digitaler Speicherformate für Audio. Einige konnten sich durchsetzen, andere eher weniger. Da die CD als Speichermedium zunehmend an Bedeutung verliert, sei hier noch ein Überblick über einige andere Dateiformate gegeben.

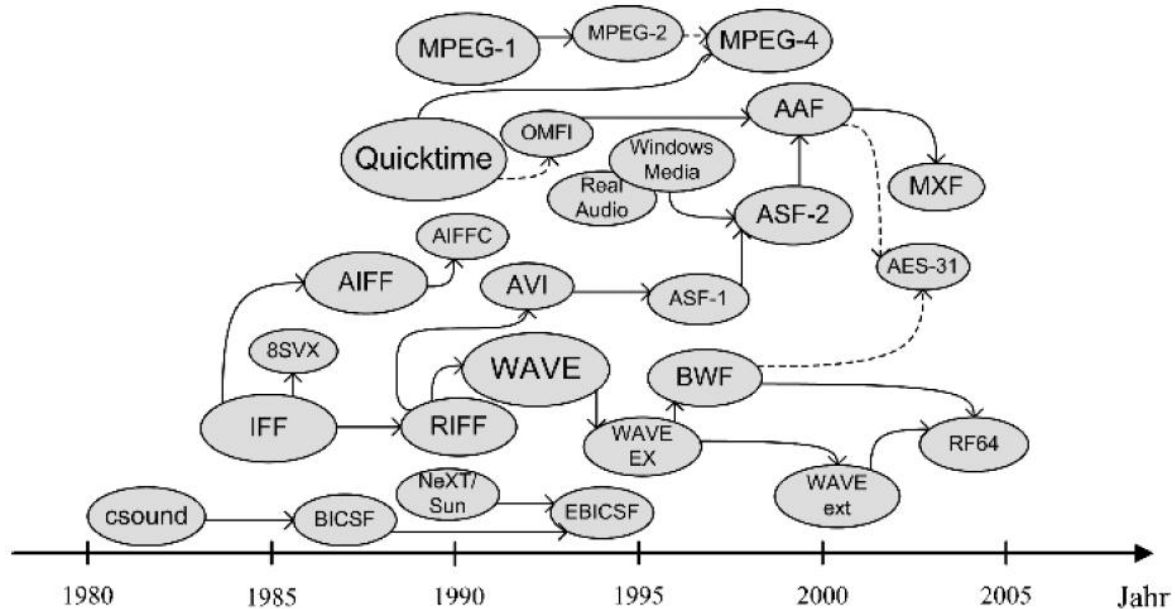


Abb 5-3: Evolution der Dateiformate für Audio. Die Jahresleiste gibt den Zeitraum der Einführung an, Symbol- und Schriftgrößen sind ein Indikator für den Beitrag des Formats zur technologischen Weiterentwicklung. [WEINZIERL]

WAVE (*.wav)

Das WAVE-Format war eines der ersten vollständig definierten Audio-Dateiformate. Es basiert auf dem *Resource Interchange File Format* (RIFF), das von Microsoft für das Betriebssystem Windows definiert wurde. RIFF ist ein so genanntes Containerformat zur Speicherung von Multimediadaten.

In der Regel sind in einer WAVE-Datei Audiodaten in einer unkomprimierten, „rohen“ Form vorhanden, die aus PCM-Daten gewonnen werden. Die Samples können in bis zu 24 Bit quantisiert sein. Durch eine Begrenzung von 4 GB Audi-Datenblock-Größe ist eine einzige WAVE-Datei oft nicht zur Speicherung von Mehrkanal-Audio geeignet.

Das *Audio Interchange File Format* (AIFF) ist das Apple-Pendant zu WAVE.

Andere auf RIFF basierte Formate sind z. B. *BMF*, welches für den Austausch von Audiomaterial zwischen Rundfunkanstalten und Studios erforderlich ist, da es ausführlichere Beschreibungsdaten und Kennzeichnungen enthält. *RF64* erweitert das WAVE-Format, in dem die 4 GB-Beschränkung aufgehoben wird. *AVI* wird in der Regel kombiniert mit Video-Daten verwendet.

MPEG-1

Die Moving Picture Coding Experts Group (MPEG), eine Arbeitsgruppe des technischen Komitees der ISO/IEC, hatte 1988 begonnen, neue Standards zur Audio/Video-Übertragung und Speicherung auszuarbeiten. Wesentlich war dabei die Blickrichtung auf verlustbehaftete Datenreduktion, um eine effiziente Distribution und Mediennutzung zu ermöglichen. 1992 wurde der Audiostandard MPEG-1 mit drei verschiedenen Layern der Audiokodierung fertig gestellt. Layer I wurde wenig genutzt, Layer II ist die Basis für Audiospeicherung im Rundfunkbereich, für Übertragung im digitalen Rundfunk und das meistgenutzte Audioformat für Digital Video Broadcasting. Layer III wurde durch die gute Klangqualität bei geringen Bitraten für Popmusik sehr populär und wird allgemein als *mp3* bezeichnet.

MPEG-1 Layer 3 (*.mp3)

Dieses Format dient zur Speicherung verlustbehafteter komprimierter Audiodaten. Es werden psychoakustische Phänomene ausgenutzt, um nur die Audiosignale zu speichern, die für den Menschen hörbar sind. Dadurch wird eine Reduzierung der Datenmenge möglich, die die wahrgenommene Audioqualität nicht, oder nur kaum verringert.

Weitere MPEG Formate

MPEG-2 bietet erhebliche Verbesserungen in der Videokodierung. *MPEG-2-AAC* (Advanced Audio Coding) ist ein Format, welches u. a. von Apple für den Online-Musikvertrieb verwendet wurde. *MPEG-4* zeichnet sich durch eine stärkere Video-Kompression aus. *MPEG-7* bietet eine umfassende Beschreibung von multimedialen Inhalten. *MPEG-21* versucht die Voraussetzungen für ein einheitliches Multimedia-Netzwerk zu schaffen. *MPEG-7* und *MPEG-21* zielen nicht auf Datenkompression ab.

Ogg Vorbis (*.ogg)

Ogg Vorbis ist ein freies Audioformat für verlustbehaftete komprimierte Audiodateien und definiert sich vor allem über seine explizite Patent- und Lizenzfreiheit als Open Source-Code-Alternative zu anderen Containerformaten. Dieses Format wird von der Stiftung Xiph.org gepflegt.

Quicktime

Quicktime beschreibt die Formt-Familie für Multimedialinhalte der Firma Apple, bestehend aus Dateiformat, Wiedergabe-Applikation, Encoding- und Authoring-Paket, sowie Streaming-Server. Teile vom Quicktime-Format wurden in den MPEG-4-ISO-Standard übernommen, umgekehrt wurde der ISO-Standard in Quicktime integriert.

Dolby AC3

Dolby AC3 oder Dolby Digital ist ein Mehrkanal-Tonsystem der Firma Dolby, das in der Filmtechnik zum Einsatz kommt. Dolby Digital unterstützt bis zu sechs diskrete Kanäle und verwendet ein psychoakustisches, verlustbehaftetes Verfahren zur Datenkompression. Weitere Mehrkanal-Tonsysteme sind z. B. *DTS* und *SDDS*.

Verlustfreie komprimierte Datenformate

FLAC (Free Lossless Audio Codec) ist ein Codec zur verlustfreien Audiodatenkompression, der im Rahmen der Xiph.org Foundation entwickelt wird. Ebenso wie Ogg Vorbis ist es frei verfügbar und die Nutzung ist nicht durch Softwarepatente beschränkt.

ALAC (Apple Lossless Audio Codec) ist das Pendant der Firma Apple.

Darüber hinaus gibt es noch eine große, unübersichtliche Anzahl an verschiedenen Kodierformaten und -verfahren.

Literaturverzeichnis

[AES] Audio Engineering Society, Inc. – *AES Recommended Practice for Digital Audio Engineering – Serial Multichannel Audio Digital Interface (MADI)*, Audio Engineering Society, 2003

(<http://www.iis.ee.ethz.ch/~felber/DataSheets/AES-EBU/aes10-2003.pdf>, zuletzt aufgerufen am 23.11.2011)

[BRASSEUR] Univ.-Prof. Dipl. Ing. Dr. Techn. BRASSEUR, Georg: *Skriptum zur Vorlesung Elektrische Messtechnik 1*, Institut für Elektrische Meßtechnik, TU Graz, 2005

[CONRAD] CONRAD Jan-Friedrich: *Taschenlexikon Musikelektronik*. 3. Auflage, Presse Projekt Verlags GmbH, Bergkirchen, 2000

[DICKREITER2] DICKREITER Michael: *Handbuch der Tonstudioteknik – Band 2*. 5., völlig Neubearb. u. erg. Aufl., K.G. Saur Verlag GmbH & Co. KG, München, 1990

[EMMET] EMMET, John: *Engineering Guidelines - The EBU/AES Digital Audio Interface*, European Broadcasting Union, Grand-Saconnex, 1995
(<http://tech.ebu.ch/docs/other/aes-ebu-eg.pdf>, zuletzt aufgerufen am 23.11.2011)

[FEY] FEY, Fritz: „Jenseits von 0dBFS – Wie man zuverlässig D/A-Wandler übersteuert“, Studio Presse Verlag, Oberhausen, 2003

(http://www.ansatheus.de/_at_dokserver/docs/audio/Jenseits_von_0_dBFS.pdf, zuletzt aufgerufen am 23.11.2011)

[GRABER] Ao. Univ. Prof. Dr. GRABER Gerhard: *Digitale Audiotechnik 1 – Skriptum zur Vorlesung*, Version 8.1, Institut für Breitbandkommunikation, TU Graz, 2006

[HENLE] HENLE Hubert: *Das Tonstudio Handbuch*. 5., komplett überarbeitete Auflage, GC Carstensen Verlag, München, 2001

[IEEE] IEEE Computer Society, *IEEE Standard for a High-Performance Serial Bus*, New York, 2008

[LERCH] LERCH, Reinhard: „Elektrische Messtechnik – analoge, digitale und computergestützte Verfahren“, Springer Verlag, Berlin, 1996

[SCHUBERT] SCHUBERT, Christian: „Das Thema Headroom bzw. Aussteuerungsreserve bei dBFS – ein häufiges Missverständnis“

(<http://www.sengpielaudio.com/HeadroomUndAussteuerung.htm>, zuletzt aufgerufen am 23.11.2011)

[SENGPIEL1] Dipl.-Ing SENGPIEL, Eberhard, Berlin

(<http://www.sengpielaudio.com/AnpassungInDerTonstudioteknik.pdf>, zuletzt aufgerufen am 23.11.2011)

- [SENGPIEL2] Dipl.-Ing SENGPIEL, Eberhard, Berlin
(<http://www.sengpielaudio.com/GroesseDerImpedanzen.pdf>, zuletzt aufgerufen am 23.11.2011)
- [SENGPIEL3] Dipl.-Ing SENGPIEL, Eberhard, Berlin
(<http://www.sengpielaudio.com/Harmonische-Partialtoene-Obertoene.pdf>, zuletzt aufgerufen am 23.11.2011)
- [SENGPIEL4] Dipl.-Ing SENGPIEL, Eberhard, Berlin
(<http://www.sengpielaudio.com/Rechner-db-volt.htm>, zuletzt aufgerufen am 23.11.2011)
- [SMITH] SMITH, Steven W.: „The Scientist & Engineer’s Guide to Digital Signal Processing
(http://analogical/en/processors-dsp/learning-and-development/content/scientist_engineers_guide/fca.html, zuletzt aufgerufen am 23.11.2011)
- [UNGER] UNGER Hans-Georg: *Elektromagnetische Wellen auf Leitungen*. 4. Auflage, Hüthig, Heidelberg, 1996
- [WATKINSON] WATKINSON John: *The Art of Digital Audio*, Revised Reprint, Focal Press, Oxford, 1991
- [WEBERS] WEBERS Johannes: *Tonstudioteknik*. 5., neu bearbeitete Auflage, Fränzis, München, 1989
- [WEINZIERL] Professor Dr. WEINZIERL, Stefan [Hrsg.]: *Handbuch der Audiotechnik*, Springer-Verlag, Berlin, 2008
- [ZELLMER] DR. ZELLMER GMBH, Sankt Augustin, (<http://www.firewire-infos.de/>, zuletzt aufgerufen am 23.11.2011)
- [ZÖLZER] Dr.-Ing. Habil. ZÖLZER Udo: *Digitale Audiosignalverarbeitung*. B.G Teubner, Stuttgart, 1996