11 Digitale Audiotechnik

11.1 Analog - Digital

Ein analoges Signal ist *zeitkontinuierlich* und *wertekontinuierlich*. D.h. ein Analogsignal ist zu jedem Zeitpunkt – deren es unendlich viele gibt – definiert und das Signal kann jeden beliebigen Wert annehmen.
Ein digitales Signal ist dagegen *zeitdiskret* und *wertediskret*. D.h. ein Digitalsignal ist nur zu bestimmten Zeitpunkten definiert und kann auch nur bestimmte Werte annehmen.

11.1.1 Vorteile der Digitaltechnik

- Speicherplatzersparnis
- grafische Darstellung beim Editing
- non-destructive Editing
- prinzipiell Verlustfreies Kopieren / Übertragen
- kostengünstig

11.2 Digitale Audiokomponenten

11.2.1 Mischpult

Die Fader und Potis haben fast immer eine Mehrfachbelegung. Damit ist es möglich kleinere Pulte mit sehr vielen Funktionen zu bauen. Dadurch werden die Mischpulte jedoch auch sehr viel unübersichtlicher. Eine Total Recall Funktion, also die Möglichkeit alle Einstellungen zu speichern und wieder aufzurufen, ist auch fast immer zu finden. Genauso wie eine interne Effektsektion.

11.2.2 Effektgeräte

Hall wird heutzutage fast nur noch digital erzeugt. Eine Schwierigkeit stellen besonders kleine Räume dar, da dafür eine sehr hohe Rechenleistung benötigt wird. Weitere typische digitale Effekte sind z.B. Pitchshifting und Timestretching.

11.2.3 EQs, Filter

Nur digital lassen sich FIR-Filter realisieren welche – im Gegensatz zu IIR-Filtern – keine frequenzabhängigen Phasenverschiebungen erzeugen. Hohe Frequenzen sind besonderes bei günstigen EQs problematisch.

11.2.4 Dynamics

Oft ist eine grafische Darstellung der statischen Kennlinie integriert. Limiter mit 0ms Attacktime lassen sich nur digital realisieren.

11.2.5 Endstufen

Digitale Endstufen haben einen besonders hohen Wirkungsgrad und sind somit sehr klein und leicht.

11.2.6 Aufzeichnungsformate

Format	Besonderheiten
Festplatte	Sie sind schnell und günstig. Die Platten sollten jedoch regelmäßig defragmentiert werden.
R-DAT	Das rotary head-DAT ist ein bandgestütztes System mit 2 Kanälen, einer Samplerate von 48kHz und einer Wortbreite von 16 Bit. Da beim ersten Bespielen ein Timecode auf das Band geschrieben wird sollte man das Band mindestens einmal durchgängig bespielt werden.
ADAT	Dieses System benutzt handelsübliche SVHS-Kassetten bzw. SVHS-Recorder. Es können maximal 8 Spuren auf eine Kassette geschrieben werden.
DA-38	Hier werden Hi-8 Kassetten verwendet. Die Spurenzahl beträgt 8.
DASH	
Pro Digi	
(U-Matic)	Es werden Videokasetten verwendet. Es wird immer ein Fehlerprotokoll erstellt. Deswegen wird
(DLT)	DVD-Mastermedium
MOD	Diese “Disketten“ lassen sich bei Raumtemperatur nicht ummagnetisieren. Wird jedoch eine Stelle durch einen Laser erhitzt lässt sich dies schon durch ein schwaches Magnetfeld erreichen. Dadurch wird die MOD sehr viel dichter Beschreibbar und hat somit eine hohe Kapazität.
MD	Dies ist die Consumerversion der MOD. Die Daten werden verlustbehaftet gespeichert (→ ARTAC).
CD	Die Compact Disc ist ein optisches Medium, d.h. die Daten werden als Erhöhungen und Vertiefungen auf die CD gepresst und dann von einem Laser abgetastet. Die Samplerate beträgt 44,1 kHz, die Wortbreite 16Bit und es können 2 Kanäle gespeichert werden.
DVD	Die Digital Versatil Disc gibt es in verschiedenen Varianten (s.u.).

DVD

Physikalische Formate

Name	Kapazität	Seiten	Layer
DVD-5	4,7GB	1	1

DVD-9	8,54GB	1	2
DVD-10	9,4GB	2	1
DVD-18	17,08GB	2	2
DVD-14	13,24	2	2/1

Logische Formate

Book A	DVD-ROM
Book B	DVD-Video
Book C	DVD-Audio
Book D	DVD-R (DVD-R / DVD+R)
Book E	DVD-RAM (DVD-RW / DVD+RW / DVD-RAM)

Audiospezifikationen

DVD-Video	DVD-Audio
LPCM AC-3 (Dolby Digital 5.1 / 7.1) DTS	LPCM (eventuell MLP-codiert) AC-3 (Dolby Digital 5.1 / 7.1) DTS
Für LPCM: • 48 / 96kHz • 16 / 20 / 24 Bit • max. 8 Kanäle bei 48 kHz / 16 Bit • max. 2 Kanäle bei 96 kHz / 24 Bit	Für LPCM: • 44,1 / 48 / 88,2 / 96 / 176,4 / 192 kHz • 16 / 20 / 24 Bit • max. 6 Kanäle bei 96 kHz / 24 Bit • max. 2 Kanäle bei 192 kHz / 24 Bit

11.3 AD-Wandlung

11.3.1 Abtastung

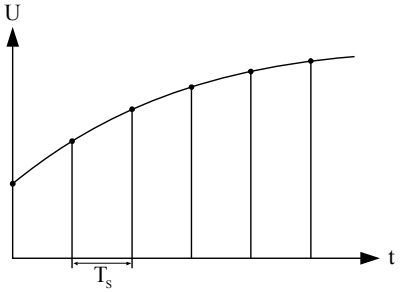


Abb. 1

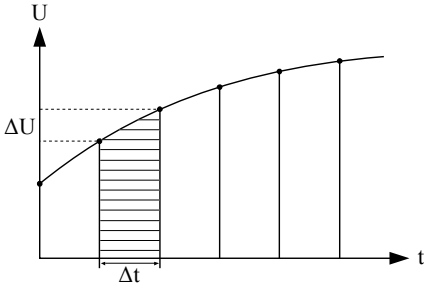


Abb. 2

Die Abtastung macht aus dem zeitkontinuierlichen Analogsignal ein zeitdiskretes Analogsignal. Wie oft das Signal pro Sekunde abgetastet wird bestimmt die *Samplefrequenz* f_s. Der Kehrwert, also die Zeit die zwischen zwei Abtastungen vergeht, wird *Sampletakt* T_s genannt. Der Abtaster selbst ist ein Schalter welcher sich im Takt einer Dirac-Impuls-Folge ständig öffnet und schließt. Ein Dirac-Impuls ist ein Impuls welcher so kurz wie möglich ist ($\Delta t \rightarrow 0$). Da das Eingangssignal von einer Dirac-Impuls-Folge in der Amplitude moduliert wird spricht man bei der Abtastung auch von einer *Puls-Amplituden-Modulation* (PAM) bzw. bei einem abgetasteten Signal von einem PAM-Signal. Nach dem Abtaster befindet sich ein Kondensator um den abgetasteten Wert über die Dauer des Sampletaktes beizubehalten. Die Verbindung aus Abtaster und Haltekondensator bezeichnet man als Sample-and-Hold-Schaltung.

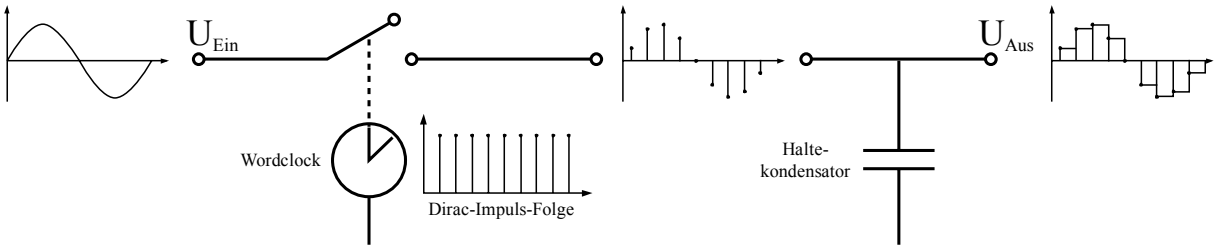
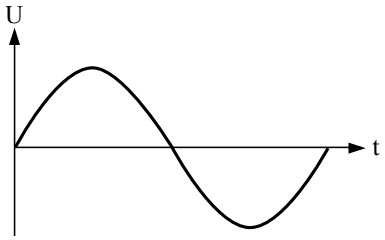
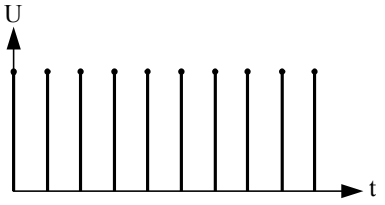
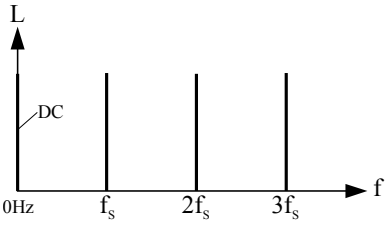
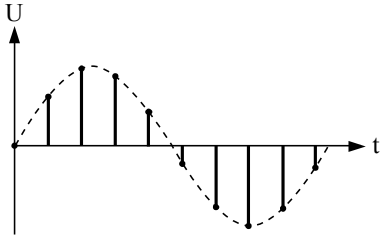
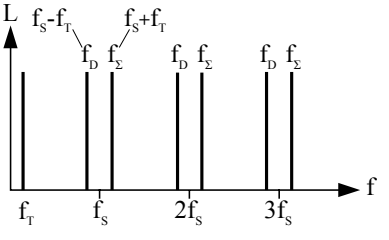
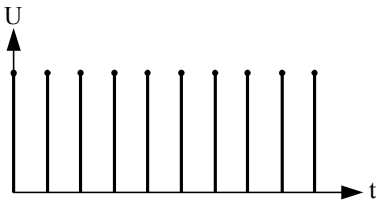
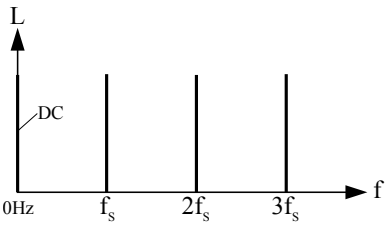
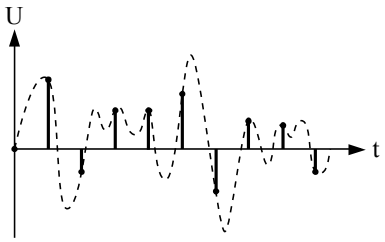
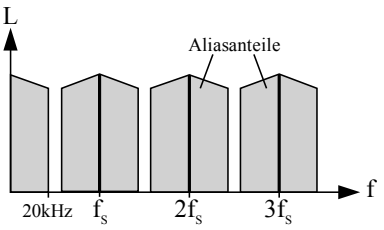


Abb. 3

Die nachfolgende Tabelle zeigt wie sich das Spektrum eines Signals ändert, wenn es abgetastet wird. Ober- und unterhalb der Samplefrequenz und deren Vielfachen tauchen im Abstand der abgetasteten Frequenz zusätzliche Frequenzen im Spektrum auf, so genannte *Aliasanteile*.

Testsignal	Spektrum Testsignal
	
Dirac-Impuls-Folge	Spektrum Dirac-Impuls-Folge
	
abgetastetes Signal	Spektrum abgetastetes Signal
	

Wird ein beliebiges auf 20kHz bandbegrenztes Audiosignal abgetastet entstehen zusätzliche 20kHz breite Frequenzbänder ober- und unterhalb der Samplefrequenz und deren Vielfachen.

beliebiges Audiosignal	Spektrum beliebiges Audiosignal
Dirac-Impuls-Folge	Spektrum Dirac-Impuls-Folge
	
beliebiges abgetastetes Audiosignal	Spektrum abgetastetes Audiosignal
	

Shannon-Nyquist-Theorem

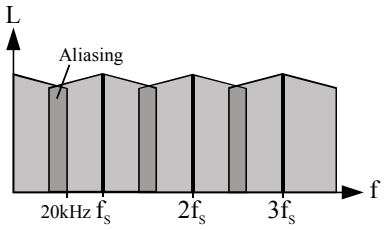


Abb. 16

Ist die Samplefrequenz f_S zu niedrig, überlappen sich die Aliasanteile und es kommt zu dem so genannten *Aliasing*. Das Shannon-Nyquist-Theorem besagt, dass die Samplefrequenz immer mindestens doppelt so hoch sein muss wie die höchste abzutastende Audiofrequenz (*Nyquist-Frequenz*), um Aliasing zu verhindern.

$$f_N = \frac{f_S}{2}$$

f_S: Samplefrequenz
f_N: Nyquist-Frequenz

Um zu verhindern, dass auch Signalanteile oberhalb der Nyquist-Frequenz an den Abtaster gelangen, wird vor dem Abtaster ein entsprechendes Tiefpassfilter angebracht. Abgeleitet vom Einsatzzweck wird dieses Filter auch *Anti-Alias-Filter* genannt.

Samplefrequenzen

22,05 kHz	Multimedia
32 kHz	Digital Radio (DAB: Digital Audio Broadcasting)
44,1 kHz	CD
48 kHz	DAT
96 kHz	DVD
192 kHz	DVD-A

Vorteile von hohen Samplingfrequenzen

- EQ arbeiten besser
- bessere Lokalisation bei Laufzeitstereofonie
- Differenztöne bei Naturinstrumenten durch sehr hohe Obertöne außerhalb des Hörbereichs

11.3.2 AD/DA-Wandlerkette

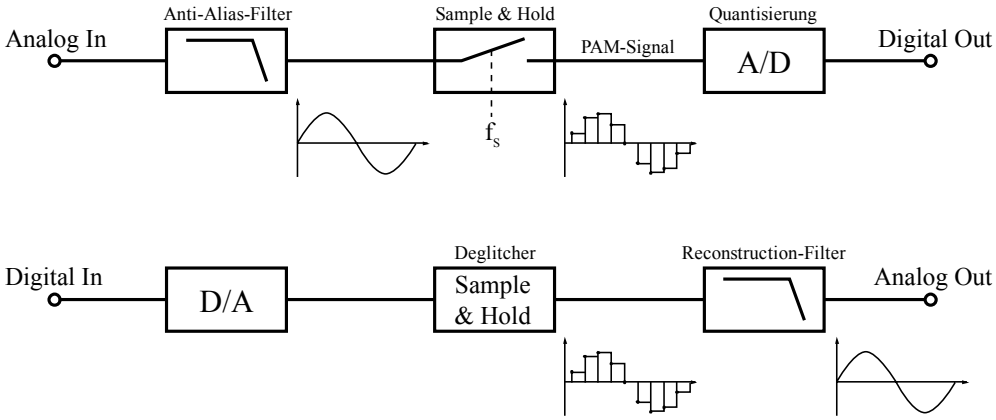


Abb. 17

Der D/A-Wandler benötigt immer eine gewisse Zeit bis er sich auf den vorgegebenen Wert eingeschwungen hat. Dadurch entstehen Glitches (siehe Abb. 18). Um diese zu entfernen wird hinter den D/A-Wandler eine Sample & Hold Schaltung angebracht. Nach seiner Funktion wird diese auch *Deglitcher* genannt.

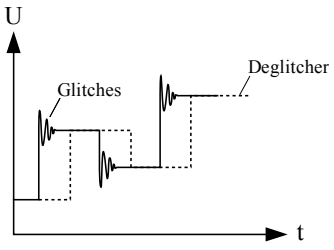


Abb. 18

Danach filtert das *Reconstruction-Filter* die bei der Abtastung entstandenen Alias-Anteile wieder aus, stellt das ursprüngliche Spektrum wieder her und macht so aus dem zeitdiskreten wieder ein zeitkontinuierliches Signal.

11.3.3 Quantisierung

Erst bei der Quantisierung findet die eigentliche Analog-Digital-Wandlung statt. Aus dem zeitdiskreten Analogsignal wird ein zeit- und wertediskretes Digitalsignal. Jedem Abtastwert wird ein, seiner Größe entsprechender, Zahlenwert zugewiesen. Die Anzahl der zur Verfügung stehenden Zahlenwerte wird durch die *Wortbreite* des digitalen Systems bestimmt. Der höchste Zahlenwert entspricht dabei immer $0dB_{FS}$. Eine Aussteuerung über $0dB_{FS}$ führt sofort zu einer krassen Verzerrung des Signals (→ Clipping). Die folgende Abbildung zeigt die Quantisierung am Beispiel eins 2 Bit Wandlers.

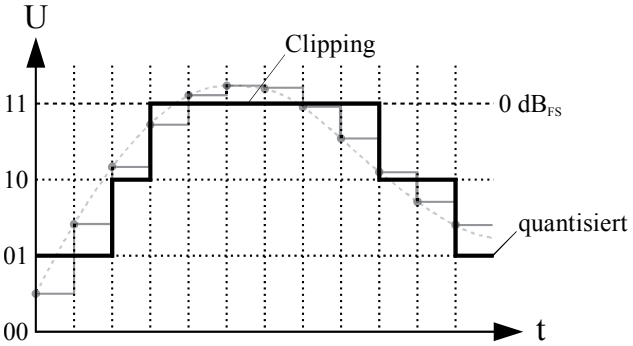


Abb. 19

Zahlensysteme

Normalerweise rechnen wir immer im Dezimalsystem, benutzen also zur Darstellung von Zahlen 10 Ziffern. Digitale Systeme benutzen lediglich 2 Ziffern (bzw. hohe Spannung, niedrige Spannung) um Zahlen darzustellen oder mit ihnen zu rechnen. Man nennt dieses Zahlensystem Dual- oder Binärsystem.

dezimal	binär	hexadezimal
0	0	0
1	1	1
2	10	2
3	11	3
4	100	4
5	101	5
6	110	6
7	111	7
8	1000	8
9	1001	9
10	1010	A
11	1011	B
12	1100	C
13	1101	D
14	1110	E
15	1111	F
16	10000	10

Da eine Zahl im Binärsystem sehr unübersichtlich ist wird zur Darstellung auch oft das Hexadezimalsystem (16 Ziffern) verwendet.

Beispiel:

$$11111111 \rightarrow 255$$

$$FF \rightarrow 255$$

Wortbreite

Die Anzahl der möglichen Spannungsstufen wird durch die Wortbreite in Bit bestimmt. Berechnen lässt sich die Anzahl der Zustände wie folgt.

$$2^{\text{Bitanzahl}} = \text{Anzahl der Zustände}$$

Bitanzahl	Anzahl der Zustände
1 Bit	2
2 Bit	4
3 Bit	8
4 Bit	16
8 Bit	256
16 Bit	65536
24 Bit	16777216

Kennlinie der Quantisierungsstufe

Ebenso wie für ein Regelverstärker kann man für die Quantisierungsstufe auch eine Kennlinie zeichnen, also die Ausgangsspannung über die Eingangsspannung auftragen. Die folgende Abbildung zeigt die Kennlinie wieder am Beispiel eines 2 Bit-Wandlers.

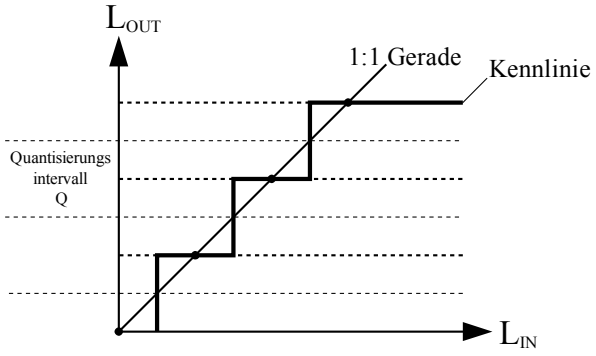


Abb. 20

Wie in der vorhergehenden Abbildung zu erkennen, erzeugt die Quantisierungsstufe ein Quantisierungsfehler, welcher abhängig vom Eingangspegel ist. Dieser Quantisierungsfehler wird hörbar als *Quantisierungsrauschen*. Er beträgt maximal ½ Q. Soll er kleiner werden muss die Anzahl der Stellen vergrößert und somit das Quantisierungsintervall verkleinert werden. Mit jedem zusätzlichen Bit verdoppelt sich die Anzahl der Stellen und der Quantisierungsfehler halbiert sich. Also wird das Quantisierungsrauschen bei gleich

bleibender Aussteuerung mit jedem zusätzlichen Bit um 6 dB leiser und die Systemdynamik um 6 dB größer. Die Dynamik eines digitalen Systems lässt sich wie folgt berechnen.

$$S/N = 6 \cdot \text{Wortbreite}$$

Beispiel:

CD (16Bit)	96 dB
DVD (24Bit)	144 dB
Interne Signalverarbeitung (32Bit)	192 dB

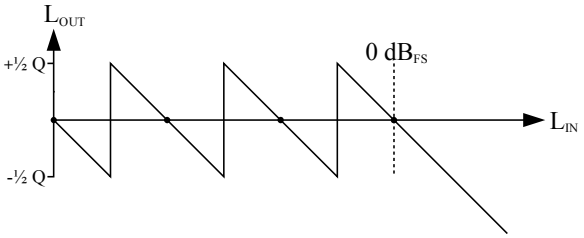


Abb. 21

11.3.4 Wortbreitenreduktion

Eine Verringerung der Anzahl der Bits ist bei der digitalen Signalverarbeitung ist an verschiedenen Stellen notwendig, beispielsweise um von der internen höheren Berechnung mit 32 Bit auf 16 Bit/24 Bit zu kommen.

Truncation

Hier werden zur Requantisierung einfach die LSBs (least significant bit) weggelassen.

Beispiel:

6 Bit

1	0	0	1	0	1

$$\rightarrow \frac{37}{63} = 0,587$$

$$20 \lg 0,587 = -4,62 dB_{FS}$$

4 Bit

1	0	0	1		

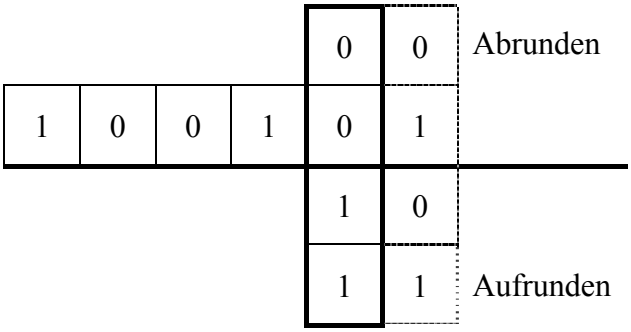
$$\rightarrow \frac{9}{15} = 0,600$$

$$20 \lg 0,600 = -4,43 dB_{FS}$$

Rounding

Bei dieser Technik fallen die letzten Stellen nicht einfach weg sondern es wird zusätzlich noch gerundet. Um dies zu erreichen muss man einfach das höchstwertigste der LSBs hinzuaddieren.

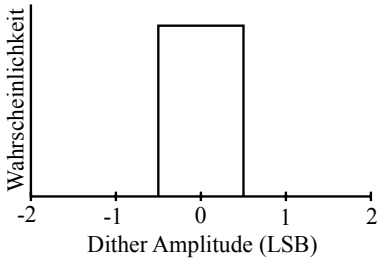
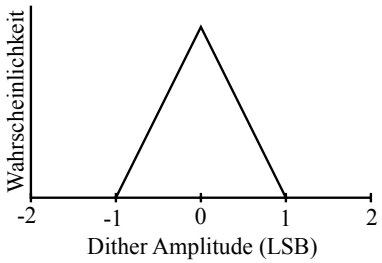
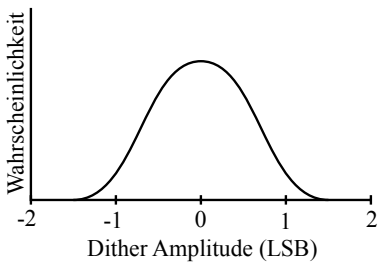
Beispiel:



Dithering

Der Quantisierungsfehler macht sich bei hohen Pegeln als Rauschen bemerkbar. Desto niedriger der Pegel wird, umso mehr bekommt das Rauschen einen tonalen Charakter. Das Spektrum des Quantisierungsgeräusches ändert sich also mit der Aussteuerung. Dadurch macht sich das Quantisierungsgeräusch viel störender bemerkbar als ein Rauschen statistischer Natur. Um diese Korrelation zwischen Eingangssignal und Quantisierungsgeräusch aufzuheben, wird dem Signal ein weißes Rauschen (→ Ditherrauschen) in Größenordnung der Zielwortbreite hinzugefügt. Durch das Rauschen nimmt eine Spannungsstufe nicht mehr nur einen Wert an, sondern zufällig den höheren oder tieferen Wert. Damit ist die Abhängigkeit von der Aussteuerung aufgehoben.

Amplitudenverteilung des Dither-Rauschens

 Abb. 22	Rechteck (RPDF: rectangular probability density function) + 3dB noise
 Abb. 23	Dreieck (TPDF: triangular probability density function) + 4,77dB noise
 Abb. 24	Gauß (GPDF: gaussian probability density function) + 6dB noise

Noise Shaping

Rauschformung (engl. noise shaping) bezeichnet ein Verfahren bei dem die Rauschenergie in Frequenzbereiche verschoben wird die vom Gehör nicht bzw. nicht so laut wahrgenommen werden. Die Rauschenergie wird dabei nicht abgeschwächt.

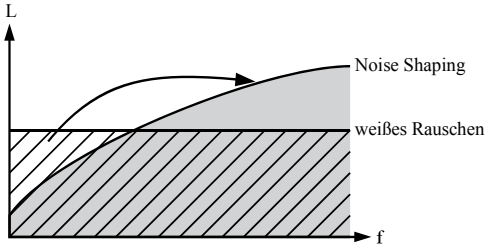


Abb. 25

Noise Shaping-Verfahren

Abkürzung	Verfahrensname, zusätzliche Parameter	Hersteller
UV22/HR	Normal, High	Apogee
SBM	Super-Bit-Mapping	Sony
ANR	Advanced Noise-Shaping Redither	Weiss Engineering
SNS	Super Noise-Shaping (1, 2, 3, 4)	Prism
POW-r	Psychoacoustically Optimized Wordlength-Reduction	Millennia Media, Weiss Engineering, Z-Systems, Lake DSP
IDR	Increased Digital Resolution Off, Moderate, Normal, Ultra	Waves

11.3.5 Jitter

Durch einen schwankenden Sampletakt werden die Amplituden fehlerhaft abgetastet bzw. wiedergegeben. Der dadurch entstandene Fehler wird als Jitter bezeichnet. Die Jitteramplitude ist der Abstand zwischen dem größten und kleinsten Sampletakt. Durch Jitter bzw. eine zu hohe Jitteramplitude nimmt das Rauschen zu hohen Frequenzen hin zu und das Stereobild verschlechtert sich.

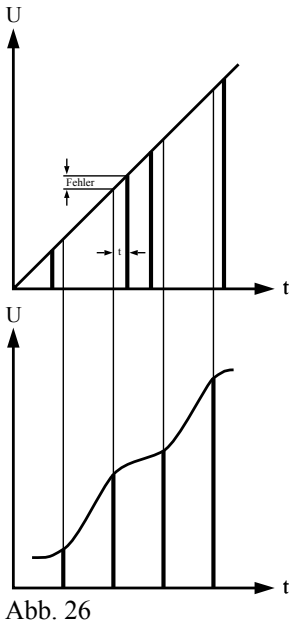


Abb. 26

11.3.6 Oversampling

Um Aliasing komplett zu verhindern müsste der analoge Anti-Alias-Filter, wie die nachfolgende Rechnung zeigt eine extrem hohe Flankensteilheit (z.B. 480dB/Oktave ≙ Filter 80. Ordnung) besitzen. Solch ein analoges Filter kann aber praktisch überhaupt nicht sinnvoll gebaut werden, da mit zunehmender Flankensteilheit Phasenverschiebungen immer mehr zunehmen und der Frequenzgang im Übergangsbereich äußerst ungleichförmig wird.

Beispiel:

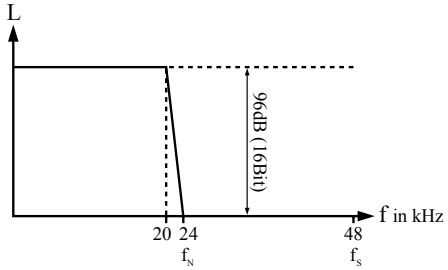


Abb. 27

$$\frac{24\text{kHz} - 20\text{kHz}}{20\text{kHz}} = \frac{1}{5}\text{Oktave}$$

$$\frac{96\text{dB}}{\frac{1}{5}\text{Oktave}} = 480\frac{\text{dB}}{\text{Oktave}}$$

Um dieses Problem zu lösen wird beim Oversampling das Signal zuerst mit einem Vielfachen der Samplefrequenz (z.B. 2-fach) abgetastet. Dadurch kann das analoge Anti-Alias-Filter viel flacher verlaufen (weniger Phasenverschiebungen, bessere Impulswiedergabe, günstiger, …). Nach dem A/D-Wandler sitzt dann noch ein weiterer Anti-Aliasing-Filter. Dieser ist jedoch als digitaler FIR-Filter realisiert, welche keine frequenzabhängigen Phasenverschiebungen erzeugen. Danach wird die Samplefrequenz direkt wieder reduziert.

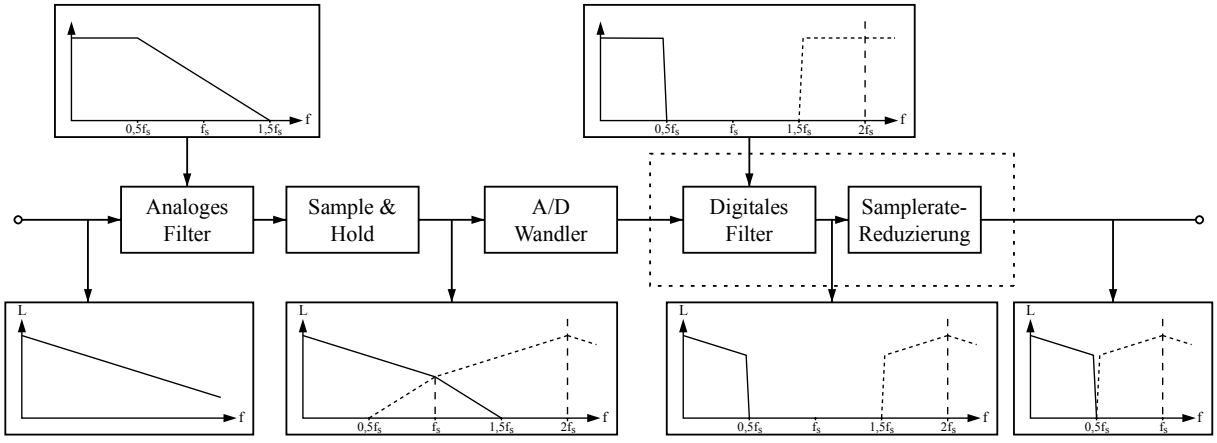


Abb. 28

Ein weiterer Vorteil des Oversamplings ist es das sich das Quantisierungsrauschen auf einen größeren Frequenzbereich verteilt, welcher aber durch das digitale FIR-Filter wieder teilweise herausgeschnitten wird. Das Quantisierungsrauschen wird deshalb pro Frequenzverdopplung um 3dB leiser. Wird Oversampling zusammen mit Noise-Shaping verwendet, kann das Ditherrauschen auch oberhalb der ursprünglichen Nutzbereichs (z.B. >20kHz) gelegt werden.

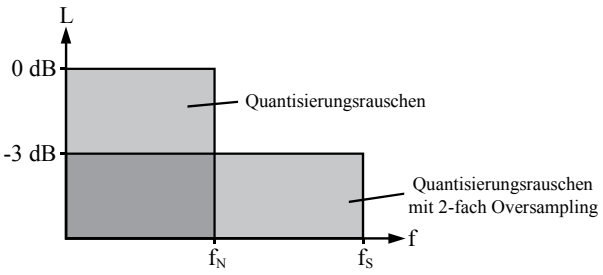


Abb. 29

11.3.7 Delta-Sigma-Wandlung

Im Gegensatz zur A/D-Wandlung mit dem PCM-Verfahren, bei dem jedes Sample einzeln als Zahlenwert gespeichert wird, der für eine bestimmte Aussteuerung steht, wird bei der Delta-Sigma-Wandlung lediglich die Different zum vorhergehenden Wert gespeichert und mit 1 Bit quantisiert. Da nur mit einem Bit gewandelt wird, verwendet man zusätzlich ein vielfaches Oversampling und ein Noise-Shaping Verfahren. Eingesetzt wird dieses Verfahren z.B. bei der SACD.

Super Audio CD (SACD)
• DSD: Direct Stream Digital • 64-faches Oversampling • “radikales” Noise-Shaping

11.4 Digitale Schnittstellen

11.4.1 Übersicht

Name	Kanäle	Stecker	Leitungsführung	Länge
S/P-DIF ”consumer” 48kHz/24Bit	2	Coaxial: Cinch	Unbalanced	2m-5m
		Optisch: Toslink	-	2m-10m
AES/EBU AES-3 ”professional” 48kHz/24Bit 96kHz/24Bit 192kHz/24Bit	2	XLR	Balanced	100m
ADAT 48kHz/24Bit	8	Toslink	-	5m-max. 10m
T-DIF 48kHz/24Bit	8	25-Pol-D-Sub	Unbalanced	5m(- max. 15m)
MADI 48kHz/24Bit 96kHz	56/64	2x BNC	Unbalanced	50m-100m

11.4.2 Wordclock

Die Wordclock ist der Taktgeber in einem digitalen System. Es wird zwischen synchroner und asynchroner Übertragung unterschieden. Bei der synchronen Übertragung wird der Takt mit übertragen und zwar entweder über eine separate Taktleitung (BNC) oder ebenfalls über die Datenleitung (Genlock). Der Empfänger generiert bei der asynchronen Übertragung den Takt (z.B. MIDI, Start- und Stopbit).

11.4.3 Digitale Synchronisation

Damit der Empfänger auch immer zur selben Zeit ein Sample erwartet zu der der Sender dieses auch sendet, muss die Wordclock des Empfängers synchron zu der des Senders laufen. Zu diesem Zweck wird immer ein Gerät zum Wordclockmaster und alle weiteren Geräte in dem Digitalen System zum Wordclockslave. Im folgenden Beispiel soll ein DAT-Recorder via S/P-DIF an ein PC angeschlossen werden. Da ein DAT-Recorder schon durch Drücken von „Play“ als Master eingestellt wird, muss der PC als Slave fungieren. Damit er dies tut, muss in den Einstellungen der Soundkarte die Syncquelle auf External (auch Coaxial, Optical oder Lock to Input A genannt) gestellt werden.

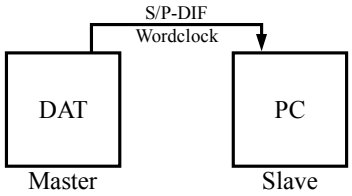


Abb. 30

Um von einem PC auf einen DAT-Recorder Audiodaten zu überspielen, muss der PC zum Master werden (Einstellung: internal). Der DAT-Recorder wird durch Auswahl des Digitaleingangs und Drücken der Record-Taste automatisch als Slave geschaltet.

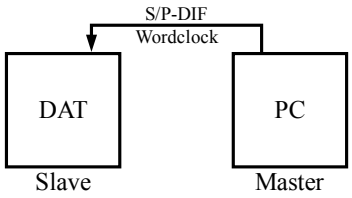


Abb. 31

Um ein komplettes Studio zu verkabeln ist jedoch eine sternförmigeverkabelung am sinnvollsten. Damit dies funktioniert müssen jedoch alle Geräte extern synchronisierbar sein. Ein so genannter Haustakt (*Synchronizer*) stellt den Master dar. Die Kabellänge sollte 6m nicht überschreiten.

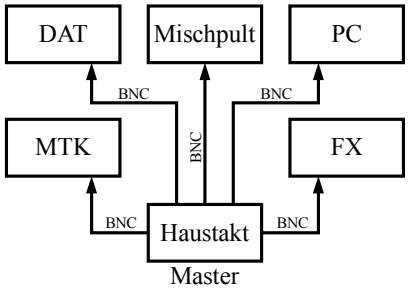


Abb. 32

Um auch Geräte die sich nicht extern synchronisieren lassen oder deren Sampletakt von dem des Haustakts abweicht in das digitale System einbinden zu können, benötigt man einen *Sample Rate Converter*.

11.4.4 Kanalmodulation/Kanalcodierung

Die Bitfolge des digitalen Audiosignals wird durch die Kanalcodierung an die Eigenschaften des Übertragungs-/Aufzeichnungskanals angepasst. Wird z.B. eine lange Folge von Einsen gefolgt von einer Folge von Nullen über ein Kabel übertragen, würde es bei Erreichen der Nullfolge erst eine gewisse Zeit dauern bis sich das Kabel entladen hat.

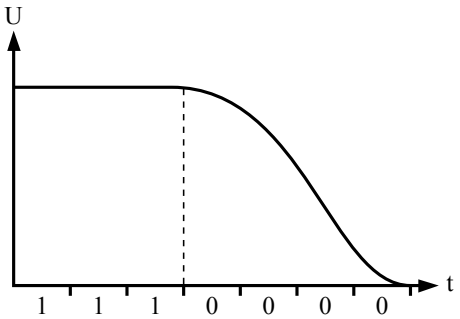


Abb. 33

Biphase-Mark-Code (Biphase-Manchester)

Der Leitungscode der z.B. bei AES-3 und S/P-DIF zum Einsatz kommt wird Biphase-Mark-Code genannt. Der Binärzahl "1" sind die Bitfolgen "10" & "01" und der Binärzahl "0" die Bitfolgen "11" & "00"zugeordnet. Außerdem ist beim Biphase-Mark-Code festgelegt, dass nach jeder codierten Binärzahl ein Flankenwechsel erfolgen muss.

Vorteile:

- Wordclockinformation mit im Datenstrom
- Kabel können sich nicht mehr "aufladen" → größere Kabellängen möglich
- kein Gleichspannungsanteil
    - Übertrager sind einsetzbar
    - Filter unterdrücken Störungen
- Signal ist verpolsicher