

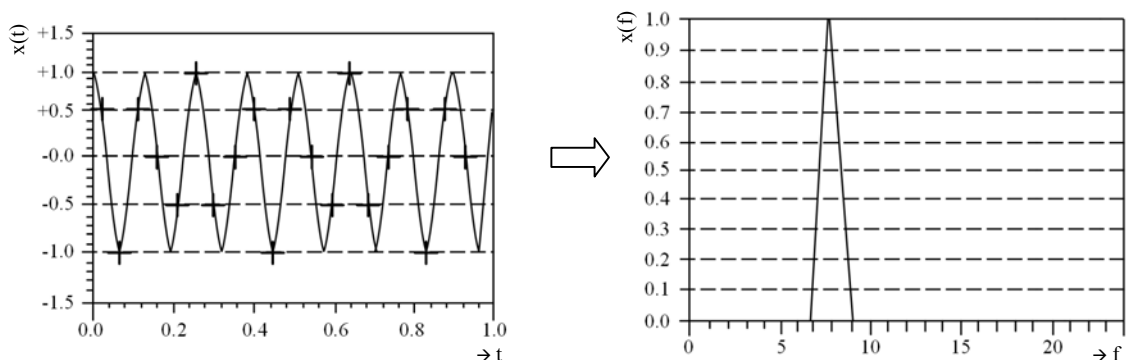
## 2 Technischer Hintergrund

### 2.1 Grundlagen der Analog-Digital-Konvertierung

Um ein analoges, d.h. ein zeit- und wertkontinuierliches Spannungssignal zu digitalisieren, muss es in ein zeit- und amplitudendiskretes Signal in Form einer binären Zahlenfolge umgewandelt werden. Die Genauigkeit dieses Konvertierungsprozesses ist von zwei charakteristischen Werten abhängig: Der Abtastrate und der Quantisierung.

Die Abtastrate ist die Frequenz, mit der dem analogen Signal in gleichmäßigen (zeitlichen) Abständen „Samples“ (Proben, Stützstellen) entnommen werden – man spricht deswegen auch von der Samplingfrequenz. Realisiert wird dieses durch die Faltung des Ursprungssignals mit einer periodischen Folge von Dirac-Impulsen, einem zumindest theoretisch unendlich schmalen Rechteckimpuls mit der Amplitude 1. Dadurch entsteht ein zeitdiskretes Signal, welches lediglich zu bestimmten Abtastzeitpunkten definiert ist. Entsprechend des als allgemein bekannt geltenden Shannon'schen Abtasttheorems von 1949 muss der Abtastvorgang mit mehr als der doppelten maximalen Frequenz des Nutzsignales erfolgen, damit das Signal später wieder einwandfrei rekonstruiert werden kann. Es müssen also bei der höchsten abzutastenden Frequenz minimal mehr als zwei Samples pro Periode genommen werden.

$f_{\text{Abtast}} > 2 * f_{\text{Signal}}$  : bei einer Abtastung gemäß des Shannon'schen Abtasttheorems lässt sich das ursprüngliche Signal wieder einwandfrei rekonstruieren – siehe Grafik Nr.1.



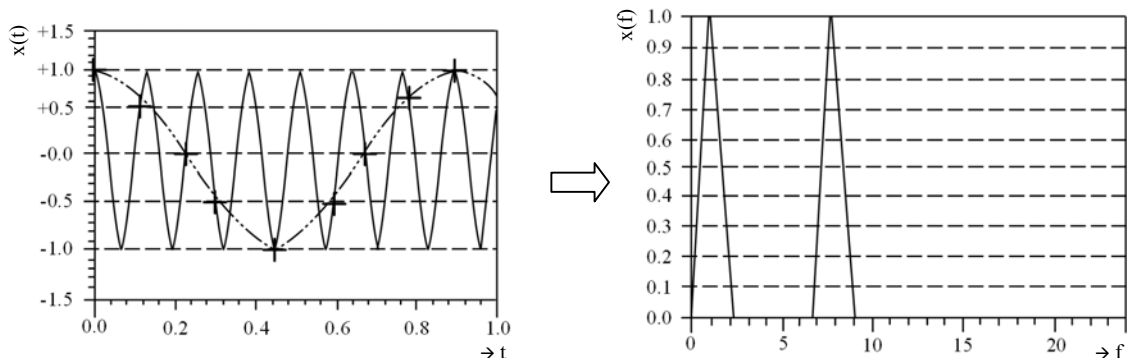
**Grafik Nr.1:** Abtastung gemäß des Shannon'schen Abtasttheorems.<sup>1</sup>

<sup>1</sup> Langen 1997: 7.

## 2 Technischer Hintergrund

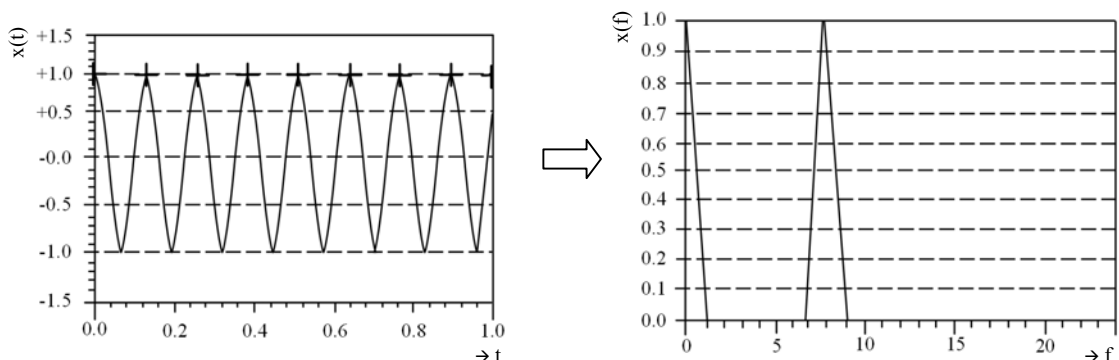
Wird gegen dieses Theorem verstoßen, werden die Samples später falsch interpretiert. Dabei kommen zwei Fälle in Betracht:

$f_{\text{Abtast}} < 2 * f_{\text{Signal}}$  : bei der späteren Rekonstruktion des Signals entsteht eine andere (=Alias-) Frequenz als die Frequenz des Eingangssignals – siehe Grafik Nr.2.



**Grafik Nr.2:** Abtastung mit einer Frequenz kleiner als der doppelten Signalfrequenz: Es entsteht eine andere (Alias-) Frequenz als die ursprüngliche Signalfrequenz.<sup>2</sup>

$f_{\text{Abtast}} = 2 * f_{\text{Signal}}$  : bei der späteren Rekonstruktion des Signals werden die Abtastpunkte falsch interpretiert. Da sie alle identisch sind, werden sie als Stützstellen einer Gleichspannung angesehen – siehe Grafik Nr.3.



**Grafik Nr.3:** Abtastung mit Signalfrequenz bewirkt DC-Aliasing. Interpretation als Gleichspannung.<sup>3</sup>

Des Weiteren legt die Abtastrate fest, bis zu welcher Frequenz das zu wandelnde Analogsignal später verarbeitet werden kann, nämlich genau bis zur Hälfte der Abtastrate, der sogenannten Nyquist-Frequenz<sup>4</sup>.

<sup>2</sup> Langen 1997: 6.

<sup>3</sup> Ebd.

<sup>4</sup> Da das menschliche Hörvermögen im optimalen Falle von 16Hz bis 20kHz reicht, sollte für die Wiedergabe des Audiosignals diese Bandbreite weitestgehend ausgenutzt werden. Die Nyquist-Frequenz sollte somit oberhalb von 20kHz liegen.

## 2 Technischer Hintergrund

---

Die im Allgemeinen übliche Samplingfrequenz von 44,1 kHz rührt aus den Anfängen der digitalen Audio-Ära, in denen die „enormen“ Bandbreiten von über 1 MBit pro Sekunde pro Audiokanal gespeichert werden mussten. Um eine brauchbare Aufnahmezeit zu erhalten, griff man auf Videorecorder zurück. Gezwungenermaßen musste die Samplingfrequenz den üblichen Fernsehstandards mit Bild- und Zeilenfrequenzen angepasst werden, so dass ganze Samples pro Zeile gespeichert werden konnten. Bedingt durch den Umstand, dass zwei Fernsehstandards existieren – 525 Zeilen bei 60Hz (M-NTSC) und 625 Zeilen bei 50Hz (BG-PAL) – musste es ein Vielfaches dieser zwei sein, bei gleichzeitiger Eignung als Samplingfrequenz (Abtasttheorem).

Diese Frequenz errechnet sich bei diesem „Pseudo-Videosystem“ aus der Multiplikation von aktiven Zeilen (ohne die ausgetasteten) mit der Bildwiederholfrequenz und der Anzahl der Samples pro Zeile.

Beim NTSC-Verfahren entfallen 35 ausgetastete Zeilen, woraus sich 490 pro Bild, bzw. 245 pro Halbbild ergeben. Bei PAL verringert sich die Anzahl der aktiven Zeilen um 37 auf 588, daraus folgend 294 pro Halbbild. Bei drei Samples pro Zeile ergibt sich:

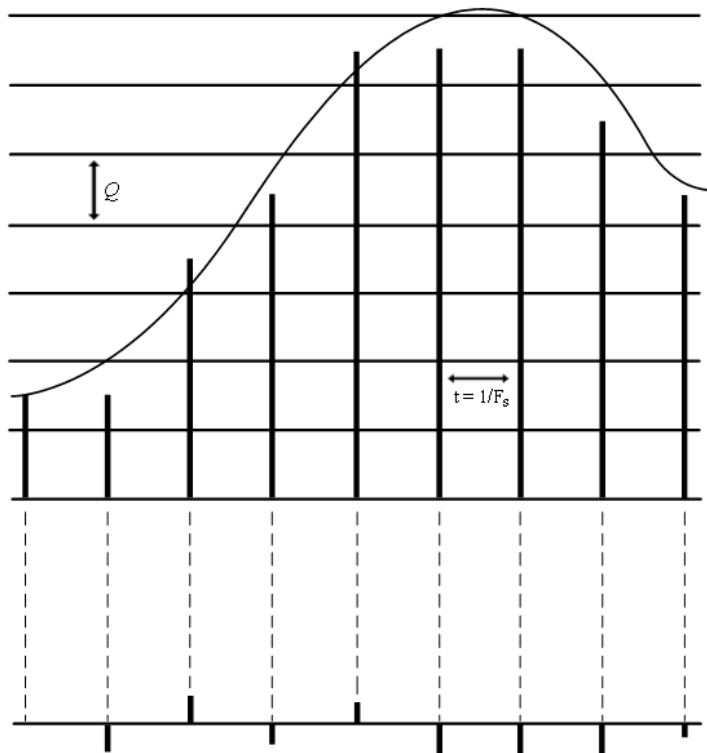
$$\text{NTSC: } 60 * 245 * 3 = 44,1 \text{ kHz}$$

$$\text{PAL: } 50 * 294 * 3 = 44,1 \text{ kHz}$$

Obwohl die CD natürlich keinen Bezug zu Video hat, übernahm man diese Samplingfrequenz, da das Equipment für die Masterherstellung auf Videoequipment basierte.

Bei der Quantisierung entsteht aus dem wertkontinuierlichen ein wertdiskretes (also amplitudendiskretes) Signal. Die Anzahl der Quantisierungsstufen errechnet sich mit  $2^n$ , wobei n die Wortbreite (Anzahl der Bits) ist. Daraus ergibt sich die Genauigkeit der vertikalen Auflösung, mit der die Amplitude der Samples erfasst wird. Die analogen Werte, die zwischen diese Stufenrasterung fallen, werden dem nächstgelegenen diskreten Wert angenähert. Daraus entsteht der sogenannte Quantisierungsfehler, also ein Spannungsfehler, der sich aus der Differenz der Stufenübertragungsfunktion zur analogen Signalkurve errechnet. Dieser Fehler beträgt maximal  $\pm \frac{1}{2}Q$  (Q = Quantisierungsstufe) – siehe Grafik Nr.4a:

## 2 Technischer Hintergrund



**Grafik Nr.4a:**

Quantisierungsfehler bei einem willkürlichen Signal: Die Abweichung beträgt maximal  $\pm \frac{1}{2}Q$ . Die Spitzen der Abtastwerte sitzen jeweils in der Mitte des Quantisierungsintervalls.<sup>5</sup>

**Grafik Nr.4b:**

Die Quantisierungsfehler kann man sich wie ein eigenes, nicht erwünschtes Signal vorstellen, dass dem ursprünglichen hinzugeaddiert wird.<sup>6</sup>

Arbeitet ein Wandler rauschfrei, so bestimmt die Quantisierung auch den Signal/Rauschabstand (SNR = Signal to Noise Ratio) der Digitalisierung. Im Allgemeinen gilt hier die Faustregel:

$$\text{SNR [dB]} = 20 \log (2^n)$$

Bei einem komplexen und großen Eingangssignal wird sich die Größe des Quantisierungsfehlers mit gleicher Wahrscheinlichkeit zwischen den Grenzen  $\pm \frac{1}{2}Q$  verteilen. In diesem Fall entspricht das ungewollte Signal (s. Grafik Nr.4b) einem breitbandigen, spektral weißen (Quantisierungs-)Rauschen, welches mit dem Eingangssignal unkorreliert ist. Es lässt sich folgender SNR berechnen:

Wie oben erwähnt existieren  $2^n$  Quantisierungsintervalle mit einem ebensolchen Spitze-Spitze-Wert der Amplitude (ohne Clipping). Der Spitzeneffektivwert beträgt demzufolge  $\frac{2^{n-1}}{\sqrt{2}}Q$ . Der Quantisierungsfehler mit einer Amplitude von  $\frac{1}{2}Q$  Spitze entspricht einem Effektivwert von  $\frac{Q}{\sqrt{12}}$ .

<sup>5</sup> Watkinson 1995: 117.

<sup>6</sup> Ebd.

## 2 Technischer Hintergrund

---

Aus der Formel des Nutz-zu-Störsignal-Verhältnisses SNR lässt sich nun folgern:

$$\text{SNR} = 20 \log_{10} \frac{Y_0}{Y_r}$$

mit  $Y_0$  = Effektivwert des Ausgangssignals ohne  
Quantisierungsrauschen

mit  $Y_r$  = Effektivwert des Quantisierungsrauschens

daraus folgt:

$$\begin{aligned} \text{SNR [dB]} &= 20 \log_{10} \frac{\sqrt{12} * 2^{n-1}}{\sqrt{2}} \text{ [dB]} \\ &= 20 \log_{10} (\sqrt{6} * 2^{n-1}) \text{ [dB]} \\ &= 20 \log_{10} \left( 2^n * \frac{\sqrt{6}}{2} \right) \text{ [dB]} \\ &= 20n \log_{10} 2 + 20 \log_{10} \frac{\sqrt{6}}{2} \text{ [dB]} \end{aligned}$$

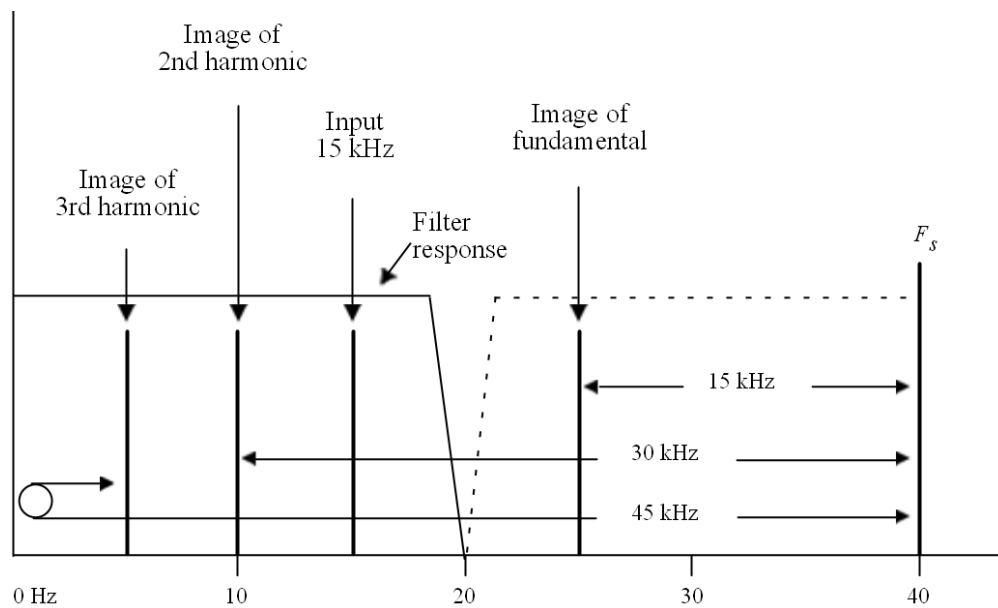
$$\text{SNR [dB]} = 6,02n + 1,76 \text{ [dB]}^7$$

Bei niedrigem Eingangspegel ist die Funktion des Quantisierungsfehlers nicht mehr zufällig, sondern wird von der Wellenform des Eingangssignals abhängig. Das Nutzsinal ähnelt einer Rechteckfolge und der Quantisierungsfehler macht sich als Verzerrung bemerkbar. Diese nichtlinearen Verzerrungen verursachen ein harmonisches Obertonspektrum, welches jedoch erst nach dem Anti-Aliasing-Filter auftritt. Es bilden sich Spiegelfrequenzen (Aliasfrequenzen) und es kommt zu tonalen Artefakten (granulation noise) und Rauschmodulation. – s. Grafik Nr.5:

---

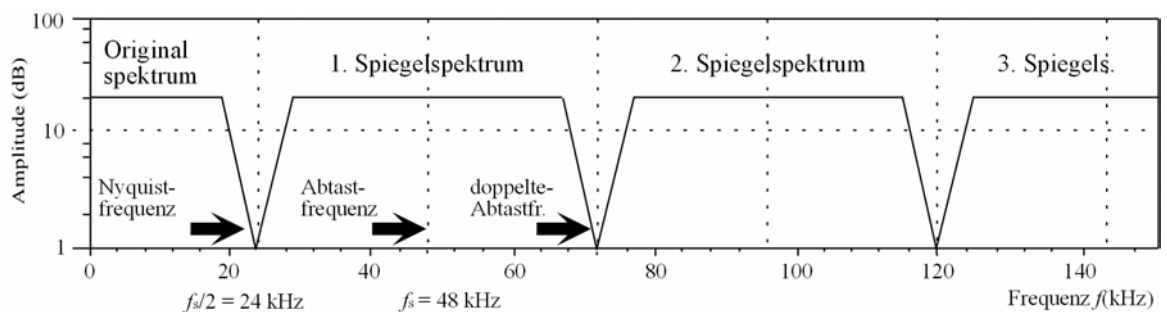
<sup>7</sup> Watkinson 1995: 118.

## 2 Technischer Hintergrund



**Grafik Nr.5:** Bei hohen Pegeln macht sich der Quantisierungsfehler als Rauschen bemerkbar. Je niedriger der Eingangspegel jedoch wird, umso korrelierter ist er mit dem Eingangssignal. Das Ergebnis sind Verzerrungen nach dem Anti-Aliasing-Filter. Bei einem Eingangssignal von 15 kHz bilden sich 2. und 3. harmonische Verzerrungen mit Spiegelfrequenzen bei 10 und (-)5 kHz.<sup>8</sup>

Die Anti-Aliasing-Filter (Tiefpass) bewirken, dass die analogen Eingangssignale vor der A/D-Konvertierung auf die maximal zulässige Bandbreite begrenzt werden. Diese ist durch die Nyquist-Frequenz (= halbe Abtastfrequenz) gegeben. Ein periodisches, diskretes Abtasten des Signals im Zeitbereich hat im Frequenzbereich ein ebenso periodisches, diskretes Spektrum zur Folge. Daraus ergibt sich, dass das Originalspektrum sich jeweils mit unterem und oberem Seitenband symmetrisch an der Samplingfrequenz und deren Vielfachen spiegelt – s. Grafik Nr.6:



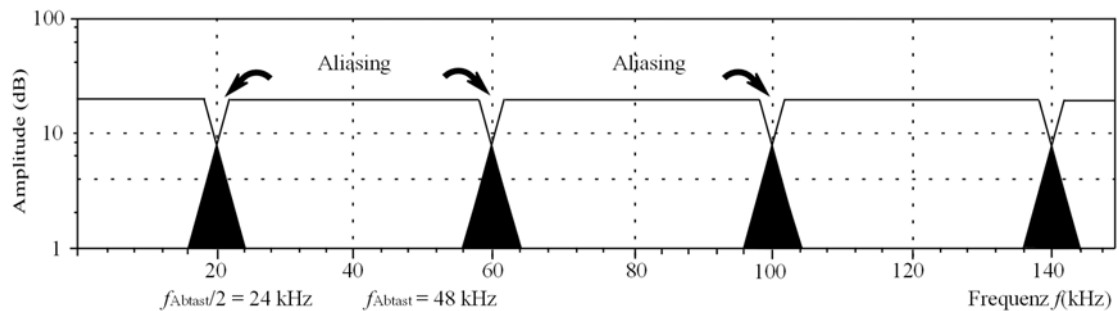
**Grafik Nr.6:** Abtastung mit doppelter Frequenz des Sperrbereichs des Anti-Aliasing-Filters. Es tritt gerade kein Aliasing auf. Durch periodisches Abtasten des Signals wird das Originalspektrum an der Abtastfrequenz und deren Vielfachen periodisch gespiegelt.<sup>9</sup>

<sup>8</sup> Watkinson 1995: 119.

<sup>9</sup> Langen 1997: 8.

## 2 Technischer Hintergrund

Es wird deutlich, dass – sobald gegen das Abtasttheorem verstoßen wird – die jeweils äußeren Seitenbänder ineinander greifen und es u.a. zu einer Überlagerung des Originalsignals mit dem gespiegelten Spektrum kommt. Um die Nyquist-Frequenz und deren Vielfache entstehen Aliasingfrequenzen. Der in Grafik Nr.7 aufgezeigte Fall tritt auf, wenn die Samplingfrequenz kleiner als die größte doppelte Signalfrequenz ist:



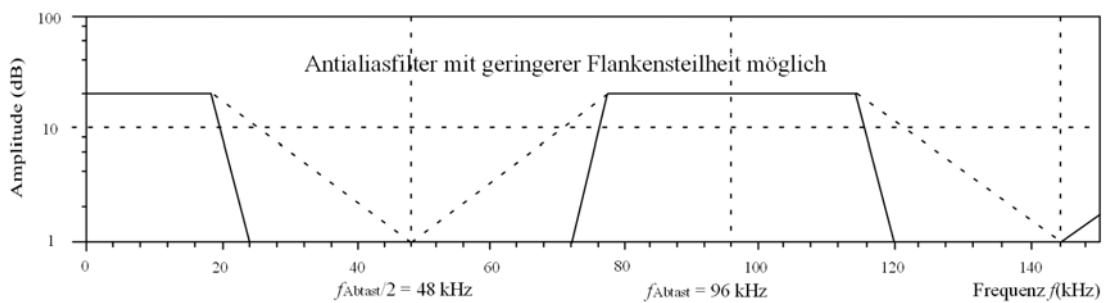
**Grafik Nr.7:** Abtastung mit einer kleineren Frequenz als der größten doppelten Signalfrequenz.<sup>10</sup>

Auf der Grundlage einer Samplingfrequenz von 44,1 kHz muss ein solches Filter also bis 20 kHz möglichst linear übertragen, andererseits aber bis zur Nyquist-Frequenz von 22,05 kHz eine genügend hohe Sperrdämpfung erreichen. Es ergaben sich analoge, extrem steilflankige Filterdesigns achter oder höherer Ordnung mit dementsprechend großen Nachteilen<sup>11</sup>. Aus diesem Grund erhöhte man die Abtastfrequenz um das 2-, 4- oder 8-fache der Standardabtastrate. Dieser Prozess des sogenannten Oversamplings bewirkt, dass die Spiegelspektren erst erheblich höher einsetzen, somit deutlich weichere analoge Tiefpässe (zweiter oder dritter Ordnung) als Anti-Aliasing-Filter eingesetzt werden konnten. Ein wirksamer Nebeneffekt ist, dass aufgrund der erhöhten Abtastrate der Signal/Rauschabstand (bei gleicher Quantisierung) steigt. Das zuvor erwähnte breitbandige, vom Eingangssignal unkorrelierte Rauschen bei höheren Pegeln verteilt sich auf ein breiteres Spektrum bis zur neuen Nyquist-Frequenz, demzufolge existiert weniger Störleistung im Audioband. Innerhalb der gleichen Bandbreite verringert sich der Störpegel bei jeder Verdoppelung der Samplingfrequenz um 3dB, die spektrale Leistungsdichte halbiert sich – s. Grafik Nr.8:

<sup>10</sup> Ebd.

<sup>11</sup> Bei derartig steilflankigen Filtern wird die Phase stark gedreht und die Gruppenlaufzeit ist immer frequenzabhängig. Deswegen benötigen die unterschiedlichen Frequenzen unterschiedliche Zeiten, um die Filter zu passieren (Problem der Dispersion). Daraus resultierend kann der Klang sich eventuell verfärben. Des Weiteren tendieren analoge Filterdesigns solch hoher Ordnung zu Instabilitäten (Schwingungen) und sind natürlich von Temperatur, Toleranzen der Bauteile sowie deren Alterungserscheinungen abhängig.

## 2 Technischer Hintergrund



**Grafik Nr.8:** Abtastung bei wesentlich größerer Frequenz als der höchsten vorkommenden Signalfrequenz, so dass keine Aliasingfrequenzen entstehen.<sup>12</sup>

Des Weiteren ergab sich im Gegensatz zu analogen Tiefpässen bei den verwendeten digitalen FIR-Filtern (*Finite Impulse Response-Filter*) in den Oversamplern der Vorteil der Linearphasigkeit, daraus resultierend konstante, frequenzunabhängige Gruppenlaufzeiten, d.h. alle Frequenzen verzögern sich gleichmäßig in den Filtern. Im Zuge der derzeitigen (v.a. technischen) Diskussionen um DSD bzw. PCM gerieten genau diese linearphasigen Anti-Aliasing-Filter ins Kreuzfeuer der Kritik. Auf diesen Punkt wird in Kapitel 2.3 genauer eingegangen.

Die nichtlinearen Verzerrungen, die bei niedrigem Eingangssignal auftreten, betreffen den Bereich des LSB, des *Least Significant Bit*, die letzte Stelle des digitalen Ausgangswortes – bei der CD also den ungefähren Bereich bei -90dBFS (dB Fullscale). Damit ein Wandler auch hier vollkommen verzerrungsfrei arbeitet, wird dem Eingangssignal vor der Quantisierung ein Dither-Rauschen mit bestimmten statistischen Eigenschaften hinzuaddiert, um die treppenstufenartige Übertragungsfunktion, die bei geringen Pegeln entsteht, zu glätten. Ziel ist, dass der Quantisierungsfehler einen rauschartigen und signalunabhängigen Charakter annimmt.

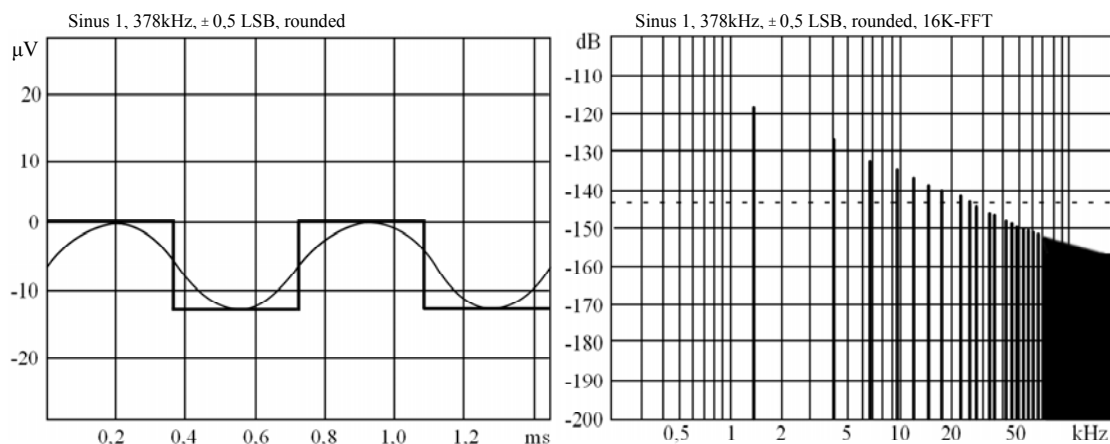
Einwandfreies Dithering ist ebenfalls bei jedem Requantisierungsschritt innerhalb der DSV-Kette, spätestens bei der Wortlängenkürzung (der sogenannten Truncation) der D/A-Konvertierung notwendig: Hierbei dürfen auf keinen Fall die niederwertigeren Bits einfach „abgeschnitten“ werden, weil dies wiederum zu Verzerrungen bei kleinen Pegeln führen würde, sondern auch hier muss der auftretende Quantisierungsfehler durch Hinzufügen von Dither vom Signal dekorreliert und in ein signalunabhängiges weißes Rauschen transformiert werden.

<sup>12</sup> Langen 1997: 9.

## 2 Technischer Hintergrund

Als Pioniere und bedeutende Vertreter auf diesem Gebiet sind sicherlich Stanley Lipshitz, John Vanderkooy und Robert Wannamaker von der University of Waterloo, Kanada zu nennen, die als erste die Notwendigkeit von Dither mit dreiecksverteilter Wahrscheinlichkeitsdichte (TPDF = Triangular Probability Density Function) zur Unterdrückung von Quantisierungsverzerrungen und Rauschmodulation propagierten. Praktikabel sind für die digitale Audiotechnik allerdings lediglich „nicht-subtraktive“ Ditherverfahren.

Anhand der unten gezeigten Abbildungen wird deutlich, wie notwendig das Addieren von Dither ist: Bei einem 20-Bit-D/A-Konverter wird von einem Sinussignal mit einer Amplitude von  $\pm 0,5$  LSB (bezüglich des D/A-Konverters) und 8-fach Oversampling ausgegangen. Durch eine einfache Rundung (von einer zuvor erheblich höheren Wortbreite) entsteht eine reine Rechteckschwingung mit einer Amplitude von 1 LSB; mit der Frequenz des ehemaligen Sinussignals schaltet das niederwertigste Bit alternierend hin und her, dementsprechend entsteht im Frequenzbereich (rechts) das typische Oberwellenspektrum einer Rechteckschwingung. Eingezeichnet ist als gestrichelte Linie außerdem der Pegel des analogen Ausgangsrauschens eines guten 20-Bit-Konverters. Sämtliche Oberwellen, die ins Audioband fallen, liegen oberhalb des Rauschens und sind tatsächlich wahrnehmbar<sup>13</sup> – s. Grafik Nr.9:



**Grafik Nr.9:** Digitalsignal ohne korrektes Dithering beim Übergang auf eine geringere Wortbreite.<sup>14</sup>

Wird nun vor der Rundung ein geeignetes Dither-Rauschen hinzugefügt (im Fall der Messung in Abbildung Nr.10:  $\pm 0,5$  LSB mit gleichverteilter Amplitudendichte<sup>15</sup>), lassen

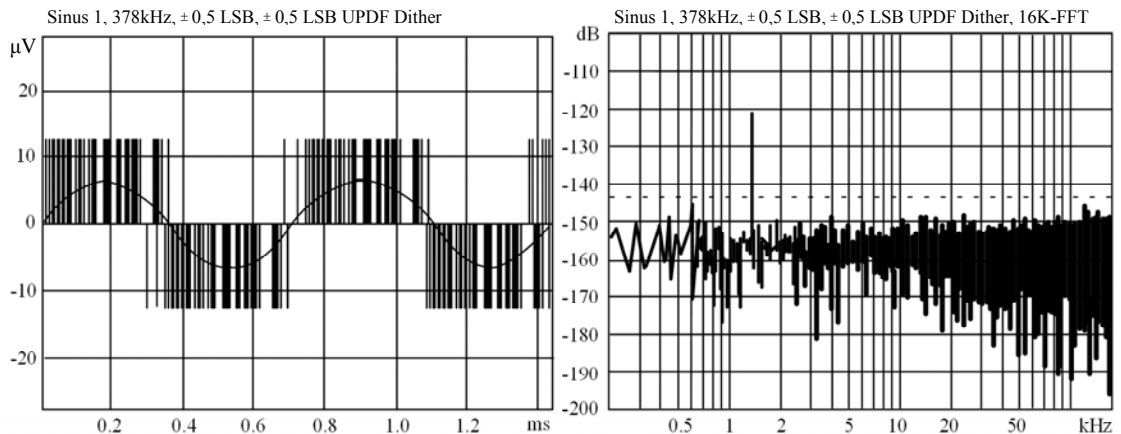
<sup>13</sup> Müller 2002: 37.

<sup>14</sup> Ebd.

<sup>15</sup> Zur vollständigen Beseitigung aller korrelierten Artefakte wäre streng genommen  $\pm 1$  LSB dreiecksverteilter Amplitudendichte erforderlich.

## 2 Technischer Hintergrund

sich dadurch die Oberwellen eliminieren. Man erhält stattdessen ein weißes Grundrauschen, welches deutlich unter dem analogen Ausgangsrauschen des 20-Bit-Konverters liegt. Bei einem korrekt gedithertem Digitalsystem versinken verstummende Töne ebenso verzerrungsfrei im Rauschteppich wie bei einem analogen System<sup>16</sup> – s. Grafik Nr.10:



**Grafik Nr.10:** Digitalsignal mit korrektem Dithering beim Übergang auf eine geringere Wortbreite.<sup>17</sup>

Die technische Realisierung von Analog-Digital-Wandlern findet in der heutigen digitalen Audiotechnik hauptsächlich durch das Sigma-Delta-Verfahren statt. Die Idee des Oversamplings weiterführend, wird mit 64-facher Standard-Abtastfrequenz gearbeitet ( $64 \cdot 44,1 \text{ kHz} = 2,8224 \text{ MHz}$ ), allerdings nur mit einem Bit. Der resultierende Signal/Rauschabstand eines einfachen Ein-Bit-Wandlers wäre aber so schlecht – obwohl sich die Rauschleistung bis zur Nyquistfrequenz bei 1,4112 MHz verteilen würde – dass man den gewünschten Dynamikumfang nur durch das sogenannte Noise-Shaping-Verfahren erreichen kann. Hierbei wird das statistisch gesehen weiße Rauschen im Spektrum so gelegt, dass davon möglichst wenig auf das Audioband entfällt.

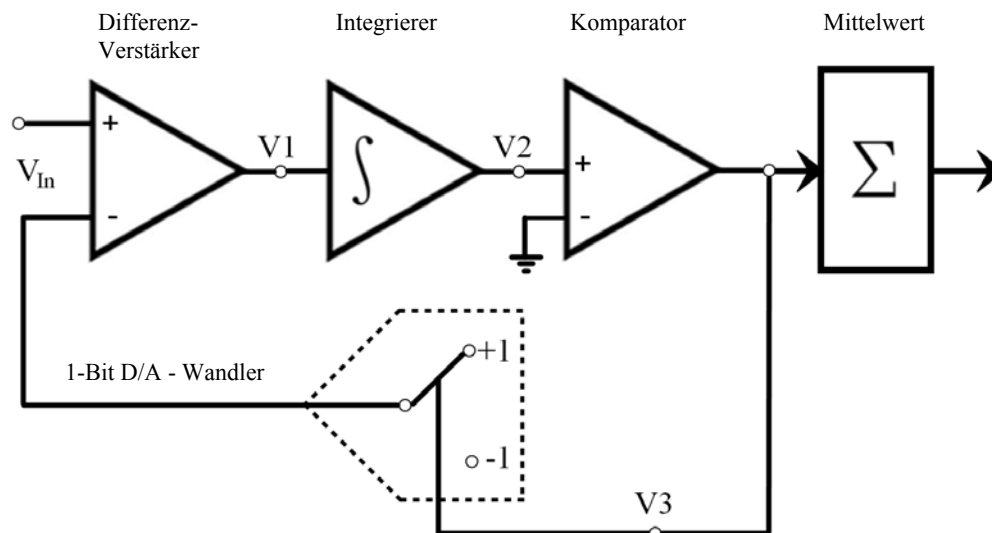
Ein Sigma-Delta-Wandler ( $\Sigma\Delta$ -Wandler, s. Grafik Nr.11) besteht grundsätzlich aus zwei Stufen, einem analogen Modulator und einem digitalen Filter. Der analoge Modulator setzt sich prinzipiell zusammen aus einem in Reihe geschalteten Eingangsdifferenzverstärker, einem Integrierer ( $\Sigma$ ) und einem Komparator ( $\Delta$ ), der als 1-Bit-Quantisierer fungiert, und einem Ein-Bit-D/A-Wandler im Gegenkopplungsweig. Das

<sup>16</sup> Müller 2002: 38.

## 2 Technischer Hintergrund

nachgeschaltete Filter (digitales Dezimierfilter) dient der anschließenden Mittelwertbildung aus einer Anzahl von Abtastwerten.

Die grundsätzliche Funktionsweise lässt sich am besten anhand eines Sigma-Delta-Wandlers erster Ordnung zeigen (die Ordnung wird bestimmt durch die Zahl der Tiefpässe oder Integrierer vor dem Komparator): Das Eingangssignal soll dafür  $\pm 1V_{SS}$  betragen und an den Punkten V1, V2 und V3 betrachtet werden.



**Grafik Nr.11:** Prinzipieller Aufbau eines Sigma-Delta-Wandlers.<sup>18</sup>

Durch den Eingangsdifferenzverstärker wird vom Eingangssignal das gegengekoppelte Ausgangssignal des 1-Bit-D/A-Wandlers subtrahiert, zu messen am Punkt V1. Dieses wird im Integrierglied zu dessen vorherigem Ausgangssignal hinzuaddiert (Messpunkt V2) und anschließend im Komparator mit einer Spannung von 0V verglichen. Ist V2 größer als 0V nimmt er am Ausgang (=V3) einen Spannungswert von +1V an, sollte er kleiner als 0V sein, legt er den Ausgang auf -1V. Im Gegenkopplungszweig schaltet der getaktete Ein-Bit-D/A-Wandler entsprechend des Messpunktes V3 auf + bzw. -1 Volt und gibt diesen Wert an den Eingangsdifferenzverstärker weiter, wo er wiederum vom nun anliegenden eventuell veränderten Eingangssignal  $V_{In}$  abgezogen wird. Dieser Zyklus wiederholt sich pro Abtasttakt.

<sup>17</sup> Ebd.

<sup>18</sup> Langen 1997: 14.

## 2 Technischer Hintergrund

Die nachfolgende Tabelle in Grafik Nr.12 soll als Berechnungsbeispiel diesen Konvertierungsvorgang mit Hilfe eines Gleichspannungswertes von +0,6V verdeutlichen:

Takt – Nummer	V1	V2	V3	Mittelwert
0	±0,0	±0,0	±0,0	$\frac{1}{8} \sum P3$ = 0,5 Volt
1	+0,6	+0,6	+1,0	
2	-0,4	+0,2	+1,0	
3	-0,4	-0,2	-1,0	
4	+1,6	+1,4	+1,0	
5	-0,4	+1,0	+1,0	
6	-0,4	+0,6	+1,0	
7	-0,4	+0,2	+1,0	
8	-0,4	-0,2	-1,0	

**Grafik Nr.12:** Berechnungsbeispiel des Ausgangssignals eines  $\Sigma\Delta$ -Wandlers 1.Ordnung bei einer anliegenden Gleichspannung von  $V_{in} = +0,6V$ .<sup>19</sup>

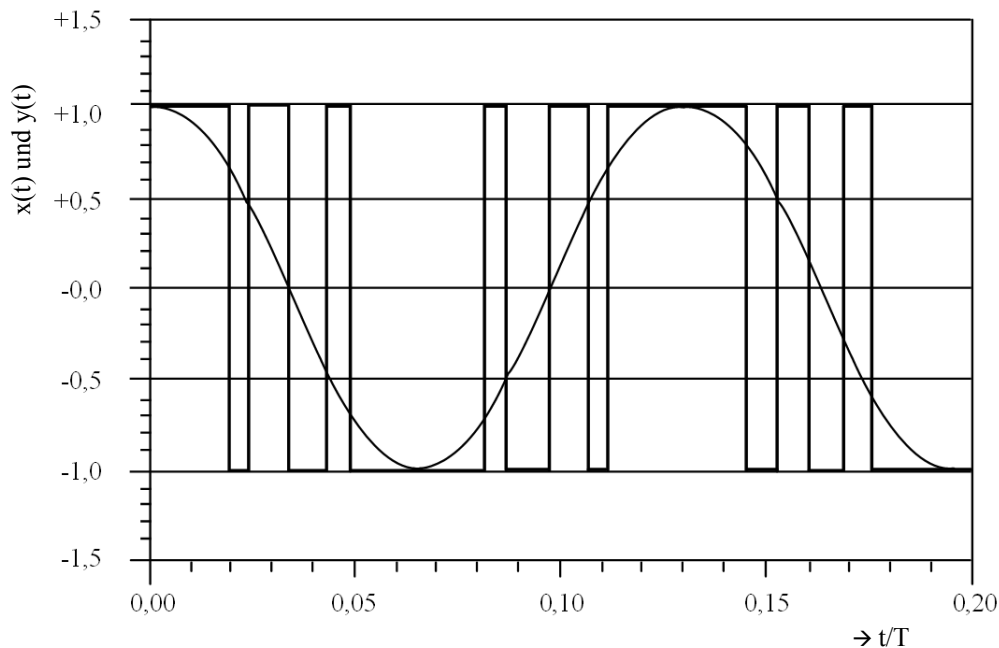
Zunächst liegen alle Messpunkte auf 0V. Gibt man die Gleichspannung von +0,6V nun im ersten Abtasttakt auf den Eingangsdifferenzverstärker, übernimmt er diesen Wert auf den Ausgang an V1, da der D/A-Wandler noch auf 0V liegt. Das Integrierglied addiert diese 0,6V zu seinem am Ausgang liegenden Wert von noch 0V. An V2 ist daraufhin eine Spannung von ebenfalls 0,6V zu messen. Der Komparator legt nun seinen Ausgang V3 auf +1V, da seine Eingangsspannung größer Null ist. Der D/A-Wandler gibt diesen Ausgangswert weiter an den Differenzverstärker, der, beginnend mit dem zweiten Abtasttakt, die +1V von der weiterhin als Eingangssignal anliegenden Gleichspannung von +0,6V abzieht. Es ergeben sich daraus -0,4V an V1, usw.

Mit Hilfe des darauffolgenden, die Abtastrate dezimierenden, digitalen FIR-Filters, das prinzipiell wie ein mittelwertbildender Akkumulator funktioniert, wird das digitale Ausgangswort mit der gewünschten Samplefrequenz und Wortbreite erzeugt.

Das am Ausgang des  $\Sigma\Delta$ -Modulators anliegende Signal ist ein 1-Bit-Datenstrom, der einem pulswertenmodulierten Signal ähnelt – s. Grafik Nr.13:

<sup>19</sup> Langen 1997: 16.

## 2 Technischer Hintergrund



**Grafik Nr.13:** Ausgangssignal eines  $\Sigma\Delta$ -Modulators bei sinusförmigem Eingangssignal.<sup>20</sup>

Wie schon erwähnt, ist die rein aus dem Prinzip des Sigma-Delta-Wandlers (durch Bitanzahl = 1 und 64-faches Oversampling) sich ergebende Dynamik zu gering, um den Ansprüchen von Audiosignalen zu genügen. Das am Ausgang des Modulators anliegende Signal enthält einen extrem hohen Anteil an Quantisierungsrauschen. Um den Signal-Rauschabstand zu erhöhen, wird dieses durch das Noise-Shaping-Verfahren spektral geformt und zum größten Teil aus dem Audioband in den unhörbaren Bereich geschoben. Die Verformung des Quantisierungsrauschens erfolgt bereits im Integrierglied, somit also schon im analogen Bereich. Dies lässt sich mathematisch bei Darstellung des Sigma-Delta-Wandlers als Linearem Modell herleiten<sup>21</sup>. Wenn man den Integrierer als analogen Tiefpassfilter realisiert, ergibt sich ein Ausgangssignal, das aus der Summe des tiefpassgefilterten Eingangssignals und des hochpassgefilterten Rauschens besteht. Somit erfüllt er einerseits die Anti-Aliasing-Filterung, andererseits aber auch die spektrale Verformung des Quantisierungsrauschens.

Das nachfolgende digitale Dezimierungsfiler verrichtet ebenso wie das Integrierglied zwei Aufgaben: zum einen wird das Quantisierungsrauschen mitsamt aller Audioanteile oberhalb der halben Samplingfrequenz herausgefiltert, so dass eine Unterabtastung auf die gewünschte Samplingfrequenz möglich ist (Abtasttheorem), zum anderen dient das

<sup>20</sup> Langen 1997: 15.

<sup>21</sup> Langen 1997: 18.

## 2 Technischer Hintergrund

---

FIR-Filter zur Erzeugung der erforderlichen Wortbreite des Ausgangssignals. Dies geschieht durch das Zusammenfassen mehrerer verzögerter 1-Bit-Samples zu einem  $n$ -Bit-Sample. Die dadurch erzielbare Wortbreite ist nicht von der Dynamik des Sigma-Delta-Modulators abhängig, sondern von der Filterfunktion.

Technisch realisiert wird dieses digitale Filter häufig durch ein Brickwall-Filter, da der Frequenzgang dieses Filters nahezu einer angenäherten Rechteckfunktion entspricht. Wie auch oben schon erläutert, ist der Vorteil dieses FIR-Filters (*Finite Impulse Response-Filter*) v.a. der lineare Phasengang und damit eine frequenzunabhängige Gruppenlaufzeit. Entworfen als nichtrekursives Filter (ohne Rückkopplungspfad), ist es weniger empfindlich gegenüber Quantisierungseffekten als ein IIR-Filter (*Infinite Impulse Response-Filter*).

Auf der D/A-Seite setzte man zu Beginn der digitalen Audio-Ära überwiegend Wandler nach dem R-2R-Prinzip ein. Dieses Konverterprinzip, auf das hier nicht weiter eingegangen werden soll, zeichnet sich zwar auch heutzutage noch durch ein sehr geringes Grundrauschen und demzufolge einen sehr hohen Dynamikumfang aus, konnte sich jedoch aufgrund v.a. sehr hoher Bauteilkosten<sup>22</sup> nicht durchsetzen. Zum Einsatz kommt wiederum der Sigma-Delta-Wandler. Hierfür wird das Audiosignal zuerst auf üblicherweise das 8-fache der Standard-Abtastrate überabgetastet (Oversampling mit Hilfe eines Interpolationsfilters), nach dem  $\Sigma\Delta$ -Wandler steht anschließend wieder das Ein-Bit-Stream-Signal zur Verfügung. Von einem internen Switched-Capacitor-Filter wird es abschließend analog vorgefiltert.<sup>23</sup>

---

<sup>22</sup> Hier sind v.a. möglichst gleiche, deswegen lasergetrimmte Widerstände, daraus resultierend differentielle Nichtlinearitäten und „individuelles“ Klirrverhalten der einzelnen R-2R-Wandler zu nennen.

<sup>23</sup> Müller 2001: 23-24.