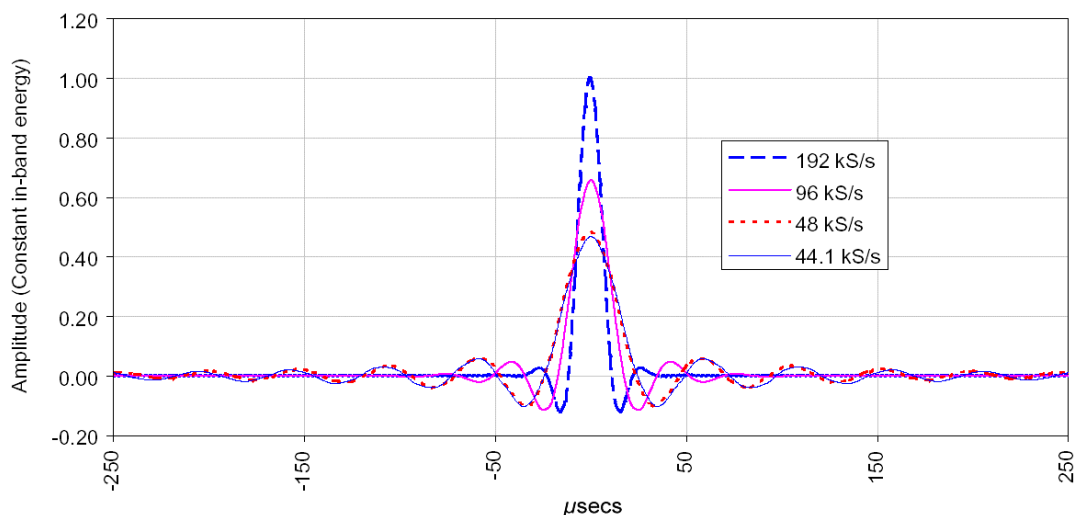


2.3 Vor- und Nachteile von DSD und PCM

Die vor allem technisch orientierten Diskussionen in Bezug auf jeweilige Vor- bzw. Nachteile von DSD und hochauflösendem PCM dauern bereits seit mehreren Jahren an und sind mit diversen Fachaufsätzen, Convention Papers, Artikeln und Interviews belegt. Im folgenden Kapitel soll versucht werden, einen kleinen und neutralen Überblick über die jeweils positiven und negativen Aspekte der unterschiedlichen Aufzeichnungssysteme zu geben. Erklärtes Ziel beider Aufzeichnungsverfahren ist selbstverständlich die Steigerung der digitalen Audioqualität.

Hauptkritikpunkt an PCM sind für die DSD-Befürworter die Anti-Aliasing-Filter. Diese verursachen den sogenannten Filter-Ringing-Effekt: Linearphasige Filter haben eine symmetrische Impulsantwort, sie weisen also nicht nur ein Aus- sondern auch ein spiegelsymmetrisch zur Ankunftszeit des Signals liegendes, ebenso langes Einschwingen auf. Dabei sind beide Schwingungsvorgänge umso länger, je höher die Ordnung, also je steilflankiger das Filter ist (s. Grafik Nr.17). Bereits vor dem Eintreffen eines transienten Impulses reagiert das Filter mit einem Einschwingvorgang. Genau diese Signalverfälschungen sollen nach Aussage der DSD-Befürworter – auch noch bei Abtastraten bis 192 kHz – zu hören sein und bei sehr kurzen Transienten mit einem sehr ausgebreiteten Spektrum zu „Verschmierungen“ führen. Von der „PCM-Seite“ wird diese Aussage (und der sich aus dieser ergebende, vermeintlich *hörbare* Vorteil von DSD) aufgrund der deutlich erhöhten Abtastraten der DVD-A angezweifelt und mit den – um den Faktor der Abtastratenerhöhung – erheblich verringerten Filterlängen und dem damit verbundenen Zeitraum des Vor- und Nachschwingens argumentiert. Ein großer Vorteil der höheren Abtastrate (96 bis 192 kHz) liegt für sie neben der größeren Audiobandbreite vor allem in der Flexibilität, die nun bei der Entwicklung der benötigten digitalen Filter besteht, sowie der Erleichterung bei der digitalen Signalverarbeitung. Ein weiterer Kritikpunkt der DSD-Befürworter sind auf der A/D-Seite die Dezimations- (Downsampling) und auf der D/A-Seite die Interpolationsfilter (Oversampling), da diese Requantisierungsrauschen hinzufügen.

2 Technischer Hintergrund



Grafik Nr.17: Symmetrische Impulsantworten bei verschiedenen Samplingraten.²⁸

Je höher die Samplingfrequenz, umso niedriger sind die Artefakte durch die Zeitquantisierung. Deswegen propagierten Sony und Philips mit DSD den einbittigen, ungefilterten Datenstrom, der direkt nach dem Komparator des Sigma-Delta-Modulators abgegriffen werden kann. Bei 64-facher Standardabtastrate, also 2,8224 MHz Samplingfrequenz (2,8224 Megasamples pro Kanal pro Sekunde), gilt im Audioband mit Hilfe von Noise Shaping ein Mindestdynamikumumfang von 120 dB und eine Bandbreite bis 100 kHz als SACD-Standard.

Aufgrund der hohen Nyquistfrequenz stellt DSD sehr genügsame Anforderungen an die Anti-Aliasing-Filter. Deren Sperrdämpfung muss erst bei ungefähr 1,4 MHz erreicht werden. Bedingt durch den weichen Filterverlauf oberhalb von 95 kHz ist das Filter-Ringing im Zeitbereich erheblich geringer gegenüber 96 oder auch noch 192 kHz, DSD zeigt hier die geringsten Artefakte.

Der hierfür zu zahlende Preis ist das enorme Quantisierungsrauschen, welches durch die extreme Überabtastung mit 2,8224 MHz Samplingfrequenz entsteht. Da es jedoch nur hochfrequente Komponenten enthält, die mit dem Audiosignal unkorreliert sind, ist es nicht wahrnehmbar.²⁹

Von der Gegenseite wird dahingegen herausgestellt, dass eben dieser neben dem eigentlichen Audiosignal existierende „Quantisierungsschmutz“ wirkungsvoller Unterdrückungsmaßnahmen bedarf. Das starke Noise Shaping verursacht schon ab 20 kHz einen äußerst steilen Anstieg des Rauschpegels, der bei ca. 70 kHz ein extrem hohes

²⁸ Story 1997: 4.

²⁹ Reefman/Janssen 2002: 11.

2 Technischer Hintergrund

Niveau erreicht – insgesamt beinhaltet das DSD-Signal deutlich mehr Energie (acht mal so viel) als ein Sinussignal, welches mit Vollaussteuerung übertragen wird.³⁰ Inwieweit bei diesem enormen Störpegel Subtilitäten und schwachpegelige Details innerhalb des Oberwellenspektrums und von Transienten bei Musikinstrumenten noch zuverlässig wahrgenommen werden können, bleibt auf der Basis des aktuellen Forschungsstandes unklar.

Die Notwendigkeit der Erhöhung der Übertragungsbandbreite begründet sich u.a. durch das breite Spektrum von sehr kurzen Transienten. Bei Entfernen von deren höherfrequenten Anteilen würde sich eine Verbreiterung in eine gauss-förmige Kurve ergeben. Um dieses zu vermeiden, benötigt man eine Bandbreite von 50 bis 100 kHz; für das Unterdrücken der oben erwähnten Ringing-Effekte ganz allgemein sehr breite Übertragungsbänder. Beide Phänomene zusammengenommen, ist ein Übertragungsbereich bis 50 kHz und ein sich daran anschließender Übergangsbereich von 100 kHz – also ein Frequenzbereich bis 150 kHz – nötig, um sämtliche eventuell hörbaren Artefakte zu beseitigen.³¹

Von der Verschwendung an Kanalkapazität bei der Speicherung des nicht unterdrückten Quantisierungsrauschens und der daraus folgenden Unterlegenheit gegenüber des vergleichbaren Multibit-PCM-Modus der DVD-A³² einmal abgesehen, ist bei DSD wiedergabeseitig in den nachfolgenden Analogstufen wegen der Gefahr von Intermodulationsprodukten unterhalb von 20 kHz, Instabilitäten der Endverstärker oder der Zerstörung der Hochtöner eine effektive Tiefpassfilterung absolut notwendig. Neben den digitalen Filtern in den D/A-Konvertern kommen deswegen zusätzlich moderate, analoge Tiefpässe in den SACD-Playern mit Eckfrequenzen zwischen 40 bis 80 kHz zum Einsatz, wodurch die behauptete Bandbreite von 100 kHz und die hohe Impulstreue natürlich zu einem gewissen Grad egalisiert werden. Die hohe Abtastrate von 2,8224 MHz bei DSD liefert somit in der Praxis letztendlich nur ein Passband, dessen linearer Frequenzgang nicht über 40-50 kHz hinausgeht. Zusätzlich nimmt das Signal-Rauschverhältnis oberhalb von 15 kHz bereits deutlich ab.

Bereits in mehreren und allseits bekannten Beiträgen³³ auf AES-Conventions haben Stanley Lipshitz und John Vanderkooy auf deutliche Nachteile der 1-Bit-Sigma-Delta-Modulation hingewiesen. Sie vertreten vehement den Einsatz von Multi-Bit-Modula-

³⁰ Müller 2001: 33.

³¹ D.Reefman in Müller 2001: 45.

³² Gerade bei Gebrauch von effektivem Noise Shaping (Müller 2001: 33).

³³ Hier sind u.a. die AES-Convention Preprints 5188, 5395 und 5398 zu nennen, s. Literaturverzeichnis.

2 Technischer Hintergrund

toren, da sich deren einbittige Brüder bei korrektem Dither permanent im Clipping befinden würden. Dies liegt in der Tatsache begründet, dass allein die Amplitude des TPDF-Dithers (Triangular Probability Density Function, s. Kapitel 2.1), der für die vollständige Linearisierung des Quantisierers zuständig ist, Clipping bewirkt. Wird dann noch das eigentliche Eingangssignal und das Signal der Fehler-Rückkopplung (Error-Feedback) des Noise-Shaping-Quantisierers addiert, trägt dies unweigerlich zu einem weiteren Pegelanstieg bei. Die Konsequenz wären Rauschmodulation, Instabilitäten und Verzerrungen. „Selbst wenn der 1-Bit-Konverter so weit gedithert wird, wie es eben möglich ist, können sich Grenzyklen-Artefakte bis zur Nyquist-Frequenz herauf ausbilden [...]“.³⁴

Des Weiteren ergibt sich ein Problem bei der Bearbeitung von 1-Bit-Signalen, da bei jedem noch so trivialen Signalverarbeitungsschritt (wie z.B. einer Pegeländerung) das 1-Bit-DSD-Datensignal in ein Multi-Bit-Wort konvertiert und anschließend wieder requantisiert werden muss. Sollte nun eine Sequenz von DSV-Schritten notwendig sein, würde der Signalpfad sehr viele Requantisierungen enthalten. Hierzu äußert sich Derk Reefman:³⁵

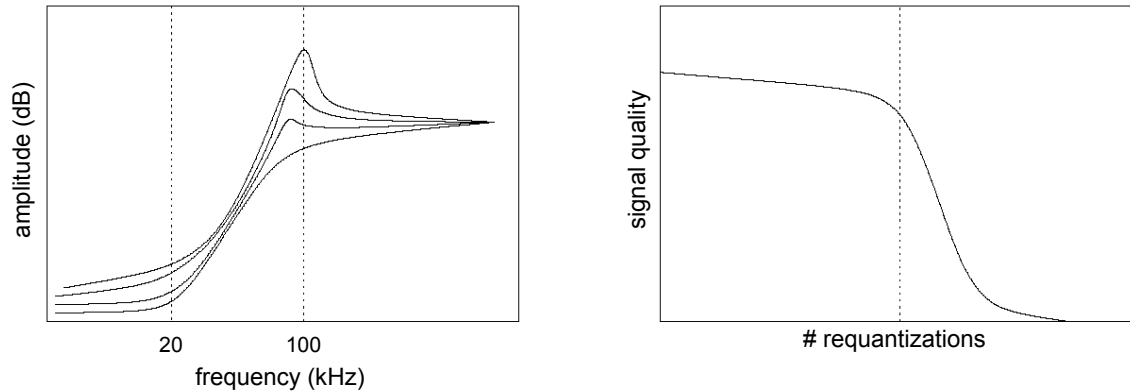
„[...]the total signal path will contain multiple requantizations. As a result of this, build-up of HF-noise will occur. This effect is illustrated in Fig.21 (=Grafik Nr.18, linkes Bild), where schematically the effect of multiple requantizations is displayed. This figure can be explained as follows. If we have a DSD signal, its noise starts to rise above 20-30 kHz, and reaches an almost flat level at about 90 kHz. If, in a subsequent requantization, the bandwidth of DSD is maintained, the signal is low pass-filtered at a frequency of about the same value (90 kHz). If this signal is fed to a next SDM (Sigma-Delta-Modulator), its output signal will contain both its own quantization noise, as well as the quantization noise that has been input to it. If this cascade is repeated, it is easy to see why there will be a build-up of HF-noise in the area of about 80-90 kHz. Eventually, this signal will be large enough to drive the SDM into its clippers, or, worse: instability. This effect is shown in the right of Fig.21 (=Grafik Nr.18); as the number of requantizations increases, the signal quality drops slowly. At the moment that the HF noise is large enough to activate the clippers, the signal quality drops rapidly.

³⁴ St. Lipshitz in Müller 2002: 41.

³⁵ Reefman, Janssen 2002: 33.

2 Technischer Hintergrund

Hence, *all* signal processing should be done in a multi-bit domain; only after the final signal processing step the conversion to $64f_s$ 1-bit signals should be made.”



Grafik Nr.18: Schematische Darstellung des Effektes mehrerer Requantisierungsschritte bei einem DSD-Signal.³⁶

Bereits im Vorwort desselben Aufsatzes stellen die beiden Autoren deutlich heraus.³⁷

“[...] In this respect, it is essential to realize that DSD at $64f_s$ is a *consumerformat* – hence, not necessarily the format that is used in the studio which can be in principle any format as long as it is of equal or better quality compared to standard DSD.”

Da ein Multi-Bit-Signal prinzipiell ohne Verluste bearbeitet werden kann, empfiehlt Philips bis zur endgültigen 1-Bit DSD-Konvertierung die Verwendung von 32-Bit-PCM mit 352,8 kHz Abtastrate bei der gesamten Verarbeitung und Speicherung für die professionelle SACD-Produktion. Auch Sony rät zum Multi-Bit-Ansatz, wobei ihrer Meinung nach jedoch das sogenannte DSD-Wide verwendet werden soll, welches auf einer 8-Bit Version von DSD basiert.

Von den DVD-A-Befürwortern wird die SACD mit ihrer Bandbreite und Dynamik oftmals mit einer DVD-A von 20 Bit und 96 kHz Samplingfrequenz verglichen, welche jedoch den geringen Noise-Floor über die gesamte Bandbreite bis 48 kHz beibehalten kann. Gerne wird darüber hinaus betont, dass auch bei der DVD-A aufnahmeseitig Noise Shaping verwendet werden und somit bei einer Quantisierung von nur 16 Bit der Rauschpegel der SACD (bei einem deutlich geringeren Gesamttrauschpegel) mühelos unterboten werden kann.

³⁶ Ebd.

³⁷ Reefman, Janssen 2002: 7.

2 Technischer Hintergrund

Abschließend sei natürlich der effektive, mehrfache Kopierschutz der SACD genannt, der durch die sukzessiv steigenden Zahlen an Raubkopien in Mitleidenschaft gezogenen, gesamten Audiobranche eventuell zur Regeneration verhelfen könnte. Zum Einsatz kommt hierbei eine Kombination von physikalischen und kryptographischen Methoden. Das oft benutzte Argument der Hybrid-SACD trägt selbstverständlich nicht, da diese Sandwich-Herstellung natürlich bei der DVD-A – wie ja auf dem Markt bereits gezeigt wurde – ebenso möglich ist.

Bei allem Gesagten bleibt noch zu erwähnen, dass bezogen auf die Rezeptivität schneller Transienten, einer erweiterten Audiobandbreite, „Verschmierungen“ bei Filtern, etc. – obwohl auf diesem Gebiet auch schon Versuche durchgeführt wurden – noch immer viele Fragen offen bleiben bzw. auditive Nachweise noch immer ausstehen.