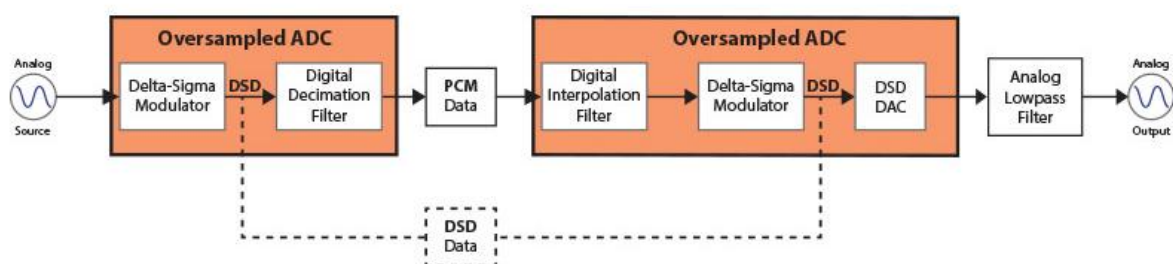


PCM im Vergleich zu DSD

Seit Jahren wird in der Fachpresse eine mehr oder weniger hitzige Debatte bzgl. der vermeintlichen Überlegenheit des einen oder anderen digitalen Modulationsverfahrens (PCM im Vergleich zu PDM=DSD) geführt. Inzwischen gibt es eine Reihe von Digital/Analog-Konvertern von sehr seriösen Herstellern, die nicht nur DSD-Dateien nativ ins Analoge konvertieren, sondern auch alle PCM-Dateien zuerst in das DSD-Format wandeln, bevor sie sie analog konvertieren, weil es vermeintliche klangliche Vorteile mit sich bringt (z.B. Nagra, PS Audio).

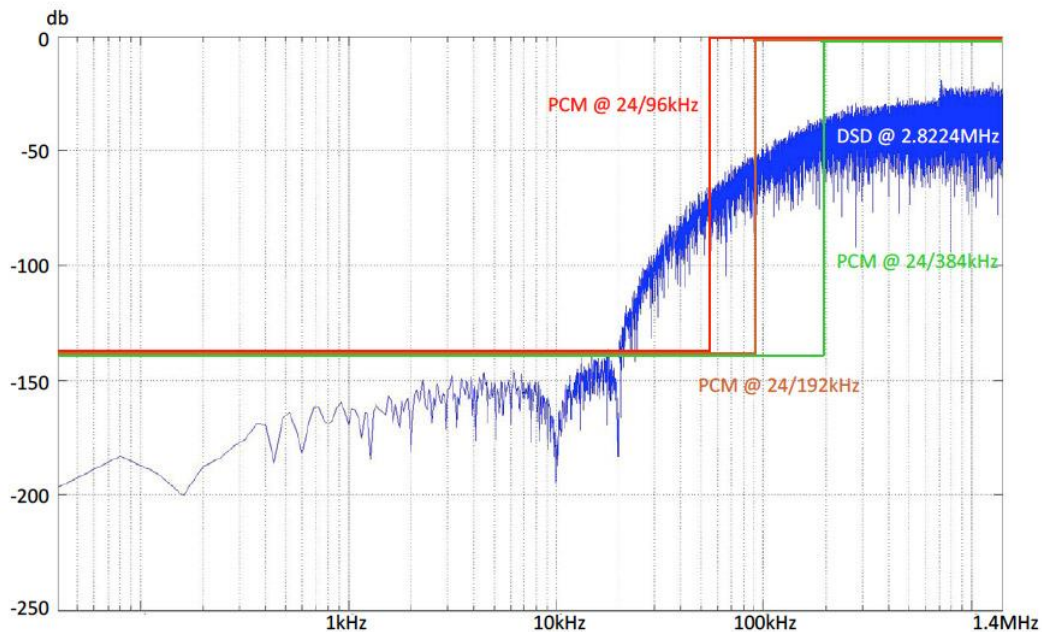
Die Gründe, warum das sein könnte, sollen hier näher beleuchtet werden. Unabhängig von allen theoretischen oder technischen Überlegungen gilt natürlich, dass am Ende das Ohr entscheidet.

Stellvertretend für die Argumentation der "prinzipiellen Überlegenheit des DSD-Formates" sei ein Artikel von Andreas Koch, einem der wichtigsten Entwickler der DSD-Technologie, "DSD – the new Addiction" aus dem Jahre 2012 herangezogen. In diesem Artikel argumentiert Herr Koch, dass die Delta-Sigma-Modulation als Lösung für Linearitäts- und Phasenprobleme analoger Filtertechnologien, seit den späten 80ern in allen PCM-Wandlerchips eingesetzt wird und die Welt seit dem ohnehin in der einen oder anderen Form DSD nutzt, ohne es zu wissen. Dabei ergaben sich jedoch neue Probleme in der DSD-zu-PCM- und PCM-zu-DSD-Wandlung, dem sog. Filter-Ringing (Pre-Ringing und Post-Ringing), das für den vielgeschmähten "digital sound" verantwortlich sei. Mit der von Phillips entwickelten DSD-Technologie sei es möglich, durch Ausschaltung der PCM-Wandlungsschritte und damit durch Verkürzung und Vereinfachung des Signalverarbeitungspfades, dieses Filter-Ringing zu eliminieren und den Klang des Wandlungsprozesses dramatisch zu verbessern (Graphik 1):



Graphik 1: Moderner PCM und DSD AD- und DA-Wandlungsprozess, gem. A. Koch

Als Beleg führt er folgende Graphik 2 an, die eine Gegenüberstellung von PCM- und DSD-Technologien zeigen soll. Im Audioband zwischen 20Hz und 20kHz beträgt der Geräuschabstand von DSD über 150dB und der Frequenzumfang reicht bis in den MHz-Bereich (wobei das Marketingmaterial von Philips und Sony "nur" von einem Frequenzumfang von bis zu 100kHz spricht). Im Gegensatz dazu bringt PCM 24Bit/96kHz es "nur" auf 144dB Geräuschspannungsabstand und einen Frequenzumfang von 48kHz. PCM 24Bit/192kHz erweitert den Frequenzumfang bestenfalls auf 96kHz um den Preis deutlich vergrößerter Datenmengen.



Graphik 2: Frequenz- und Dynamikumfang von PCM und DSD-Verfahren, gem. A. Koch

Diese Argumentation kombiniert Herr Koch mit dem Argument, dass Musik zunehmend über Downloadkanäle vertrieben wird und die DSD-Dateien bei hoher Qualität relativ wenig Speicherplatz und Übertragungsbandbreite verbrauchen, was mit folgendem Vergleich für die Aufzeichnung von einem 3-Minuten 2-Spur-Audiosignal demonstriert werden soll:

Format	Dateigröße	Download Zeit
Redbook CD (16/44.1kHz)	32MB	1.0 min.
PCM 24/88.2kHz	95MB	2.6 min.
PCM 24/96kHz	103MB	2.8 min.
PCM 24/176.4kHz	190MB	5.0 min.
PCM 24/352.8kHz (DXD)	380MB	10.0 min.
DSD 2.8224MHz	127MB	3.4 min.

Herr Koch zieht nun den Schluss, dass das DSD-Format in der Größe mit dem PCM (24/96kHz)-Format vergleichbar wäre, aber wesentlich bessere Geräuschabstände und Frequenzumfang bietet und somit "Bit-effizienter" sei und außerdem frei von Ringing-Problemen.

Da das DSD-Format bisher in seiner Verbreitung durch die ausschließliche Verwendung in SACDs beschränkt war, äußert er abschließend die Erwartung, dass mit der zunehmenden Verlagerung weg vom physischen Datenträger SACD hin zum Download, sich das DSD-Format zum dominanten Digitalformat entwickeln wird.

Herr Koch ist einer der führenden Köpfe in der Digitaltechnologie und seine Kompetenz steht völlig außer Frage. Dennoch stimmt mit der Argumentation in seinem Artikel einiges nicht:

1. Die Topologie der von Herrn Koch beschriebenen PCM-Delta/Sigma-Wandler stammt aus den 90er Jahren und ist schon lange nicht mehr aktuell. Seit der Jahrtausendwende verwenden praktisch alle PCM-Konverter Multibit-Oversampling (zumeist 4Bit), da dies den Geräuschabstand im Gesamtsystem drastisch reduziert und das Ausmaß des nötigen "Noise-Shaping" erheblich reduziert. Zudem ermöglicht das Multibit-Oversampling eine Bearbeitung durch die Tonstudios (Mixing, EQ, Dithering, Level-Changing etc.), ohne beständig (verlustbehaftet) zwischen 1Bit-DSD und PCM hin- und her-

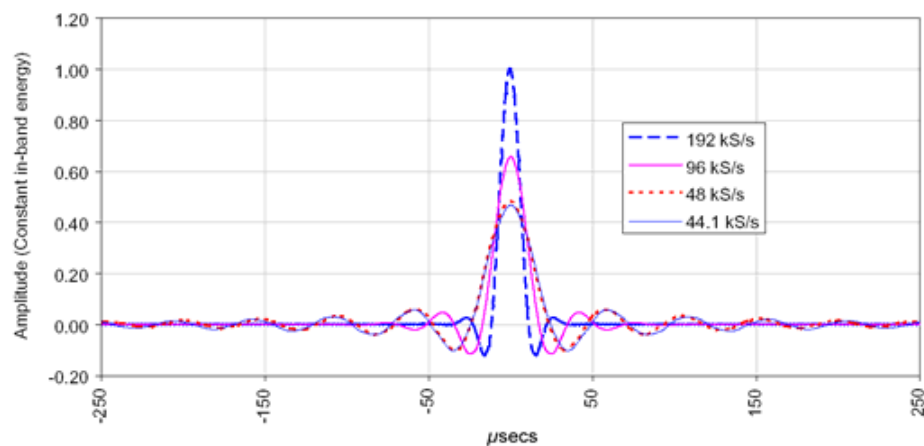
konvertieren zu müssen. Die von Herrn Koch beschriebene Prozessverkürzung und -vereinfachung setzt an beiden Punkten, wo die gestrichelte Linie in den PCM-Prozess einhakt, einen 1Bit-DSD-Prozess voraus, den es heutzutage gar nicht mehr gibt - außer vielleicht in den einfachsten Direkt-to-Disk DSD-Aufnahmen. Sobald bei DSD-Aufnahmen Editierungsschritte vorgenommen werden müssen (was fast immer der Fall ist), erfolgen sie im sog. DSD-wide-Format (Sony). Philips empfiehlt für die DSD-Produktion 32-Bit-PCM mit 352,8 kHz Abtastrate, wobei die Wandlung ins DSD-Format erst ganz am Schluss erfolgen soll. Aber selbst das DSD-wide-Format von Sony ist zumeist ein 4 oder 8Bit-Format mit DSD-Samplerate - in anderen Worten, das ist ein PCM-Format(!). Weshalb das DSD-wide-Format von profilierten Kritikern wie dem Kanadier Stanley Lipshitz gerne despektierlich aber durchaus zutreffend als „PCM-narrow“ bezeichnet wird.

2. Wo immer Multitrack-Aufnahmen gemacht werden, müssen diese in sog. Digital Audio Workstations (DAW) gemischt und bearbeitet werden. Eine Bearbeitung ist in 1Bit-Dateien Prinzip-bedingt nicht möglich. Folglich müssen zur Bearbeitung auch DSD-Aufnahmen in einem PCM-Multibit-Prozess konvertiert werden. Das ist exakt, was die DAWs von z.B. Pyramix oder Sonoma tun: sie verarbeiten PCM-Dateien mit DSD-Sampleraten. Diese Konversion von DSD zu PCM ist verlustfrei, solange die Samplerate nicht geändert wird. Da ja nur die Wortbreite vergrößert wird, gehen keine Informationen verloren.
3. Somit sind fast alle DSD-Projekte (ca. 99%, laut John Siau von Benchmark) in PCM bearbeitet worden, bevor sie in DSD konvertiert wurden.
4. Problematisch wird es erst, wenn die PCM-Dateien ins 1Bit-DSD-Format konvertiert werden, da beim Herunterdithern auf 1Bit riesige Mengen Quantisierungsrauschen erzeugt werden. Das ist schließlich der Grund, warum moderne Digitalverfahren beim Sigma-Delta-Verfahren nur auf 4 oder 8Bit runter modulieren und nicht mehr auf 1Bit. Die Rausch-Situation verschlimmert sich noch, wenn mehrere multi-Bit zu 1Bit-Konvertierungen nacheinander erfolgen (z.B. eine vom A/D-Konverter und eine weitere durch die DAW).
5. DSD-Dateien haben als 1Bit-Dateien definitionsgemäß einen Geräuschabstand von 6dB über das gesamte Spektrum. Um brauchbare Signal-Rauschabstände im Audioband zu erhalten, muss das gesamte Rauschen in den Bereich über 20kHz per "Noise-Shaping" verschoben werden, wobei dennoch ab ca. 15kHz der Rauschteppich stark ansteigt, wie in der Graphik von Herrn Koch gut erkennbar ist. Aber dieses Ultraschallrauschen ist aus den DSD-Dateien Prinzip-bedingt nicht zu entfernen. Erst wenn sie in PCM oder ins Analoge gewandelt werden, kann das Rauschen mit einem Tiefpassfilter herausgefiltert werden. Somit ist in der Realität der Rauschabstand deutlich geringer als die vielleicht theoretisch möglichen 150dB: Mit den besten D/A-Konvertern lassen sich heute bei DSD Rauschabstände von 120dB im Bereich von 20Hz-20kHz erreichen - in etwa entsprechend einem 20Bit PCM-System.
6. So ist die von Herrn Koch verwendete Vergleichs-Graphik (Graphik 2) der Frequenz- und Rauschabstände von DSD- und PCM-Dateien zumindest stark irreführend. Denn er verwendet für die Darstellung der DSD-Technologie eine Fast Fourier Transformation (FFT) und stellt diese einer geraden Linien-Darstellung der PCM-Technik gegenüber, was einem ungültigem Vergleich von Äpfeln mit Birnen entspricht. PCM in FFT würde ganz anders aussehen und viel weiter herunterreichen als die DSD-Graphik und somit dem Argument von Herrn Koch den Boden entziehen.
7. Entsprechend ist auch die Aussage bzgl. des erweiterten Frequenzspektrums der DSD-Technik irreführend: Zwar hat DSD ein in den MHz-Bereich reichenden Frequenzumfang, allerdings besteht dieser nicht aus Nutzsignal, sondern aus enormen Rauschlärm. Wie auf der Graphik von Herrn Koch gut zu erkennen ist, beginnt dieser Rauschlärm gerade oberhalb von 15kHz. Um das Signal z.B. bei der D/A-Konvertierung nutzen zu können, ist ein Lowpassfilter nötig, der das Signal stark Frequenz-

beschränkt. Der SACD-Standard (Scarlet Book) schreibt einen 50kHz Lowpassfilter Dritter Ordnung zwingend vor, um Schaden von Verstärkern und Lautsprechern fernzuhalten. Da der Filter Phasenverzerrungen und Ringing verursacht je näher er an die 50kHz-Marke kommt, heißt das, dass der maximal nutzbare Frequenzbereich der 64xDSD-Technologie bei ca. 45kHz liegt. 128xDSD erweitert den nutzbaren Frequenzbereich zu Lasten deutlich steigender Dateigrößen.

Ungeachtet des - vielleicht aus Marketing-Gründen - etwas tendenziösen Charakters der Aussagen von Herrn Koch, das Argument der DSD-Anhänger basiert im Wesentlichen auf 3 Punkten:

1. Die in der PCM-Technologie verwendeten Anti-Aliasing-Filter weisen ein sog. Filter-Ringing auf: Linearphasige Filter haben eine symmetrische Impulsantwort, sie weisen also nicht nur ein Aussondern auch ein spiegelsymmetrisch zur Ankunftszeit des Signals liegendes, ebenso langes und völlig unnatürliches Einschwingen auf. Dabei sind beide Schwingungsvorgänge umso länger, je höher die Ordnung, also je steilflankiger das Filter ist (s. Grafik 3). 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:



Graphik 3: Symmetrische Impulsantworten bei verschiedenen Samplingraten

Zwar hat das DSD-Format grundsätzlich die gleiche Herausforderung, stellt aber aufgrund der hohen Nyquist-Frequenz 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 bei DSD tatsächlich deutlich geringer. Aus demselben Grund verringert sich das Ringing allerdings auch im PCM-Prozess mit zunehmender Samplingrate. Dies ist der Grund warum auch in DSD-Produktionen das PCM-Format DXD (352,8kHz, 24Bit) verwendet wird.

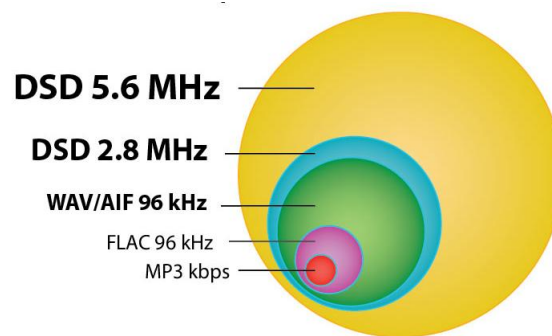
Inwieweit das Filter-Ringing tatsächlich die hörbare Wiedergabequalität beeinflusst ist noch unklar. Das Ringing tritt im Allgemeinen zwischen dem Passband und dem Stopband eines Filters auf, bei einer CD beispielsweise, also zwischen 20kHz und 22,05kHz, bei höheren Samplerraten bei entsprechend höheren Frequenzen. Vieles hängt von der Wahl des spezifischen Filters ab (einige Hardware-Anbieter ermöglichen hier dem Konsumenten eine Auswahl alternativer Filter) und der Qualität der Filterumsetzung.

2. 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. Der Vorwurf ist zumindest für den Dezimationsfilter berechtigt, ist aber zu vernachlässigen im Vergleich zum Quantisierungsrauschen, das beim Herunterdithern auf 1 Bit im DSD-Format entsteht.

- Bei vergleichbarer Datenmenge weist das DSD-Format angeblich einen höheren Rauschabstand und weiteren Frequenzumfang auf. Das DSD-Format ist somit angeblich Bit-effizienter als das PCM-Format. Wie bereits dargestellt, ist weder der Frequenzumfang noch der Geräuschabstand bei DSD größer als bei 20Bit/96kHz, so dass keine Effizienzvorteile auf Seiten des DSD-Formates bestehen. Im Gegenteil, während der Geräuschspannungsabstand bei PCM über den gesamten Frequenzbereich konstant bleibt, ist er bei DSD mit Hilfe des Noise-Shapings nur bis 20kHz aufrecht zu erhalten. Danach steigt der Geräuschteppich dramatisch an.

Will man DSD- und PCM-Format vergleichen, so kommt man, nicht zuletzt auf Basis der von Herrn Koch selber gelieferten Daten zu dem Schluss, dass das 64xDSD-Format in etwa einem 20Bit/96kHz-PCM-Format entspricht - bei ähnlicher Datenmenge im unkomprimierten WAV-Format. Im komprimierten FLAC-Format verschiebt sich die Relation deutlich zugunsten von PCM-Dateien. Dabei weist das PCM-Format den Rauschabstand über das gesamte Frequenzspektrum auf und nicht nur im Bereich bis 20kHz wie das DSD-Format. Effizienzvorteile sind für das DSD-Format damit nicht auszumachen. Unkomprimiert sind beide Formate ähnlich Bit-effizient, komprimiert ist das PCM-Format deutlich effizienter.

Die Dateigröße verschiedener PCM und DSD-Formate ist aus folgender Graphik gut ersichtlich:



Graphik 4: Dateigrößenvergleich verschiedener Audioformate

Das heute übliche Übertragungsformat für verlustfreie Audiodaten ist FLAC. Der Größenunterschied zwischen FLAC 96kHz/24Bit und DSD 2,8 (oder DSD64) liegt, bei etwa vergleichbarer Klangqualität, in 70% zusätzlichem Rauschlärm in den DSD-Dateien. Das Bit-Effizienz-Argument von Herrn Koch ist nicht plausibel.

Etwaige Nachteile des Filter-Ringings auf PCM-Seite sind im Fall der Oversampling-Wandler zwar vorhanden, aber das PCM-Format ist nicht Prinzip-bedingt mit dem Oversampling und damit mit dem Filter-Ringing verbunden, wie Herr Koch vermittelt. Andere Wandlertechniken, wie NOS („Non-Oversampling“) DACs haben das Problem nicht. Zudem sind etwaige Nachteile des Filter-Ringings für die Wiedergabequalität noch nicht klar. Hier bedarf es der weiteren Forschung. Ein am Erich-Thienhaus-Institut der Hochschule für Musik Detmold 2004 durchgeführte aufwendige ABX-Doubleblind-Vergleichsuntersuchung in 145 Tests mit 110 Versuchspersonen bzgl. der Frage, ob "es möglich ist zwischen den zwei digitalen Kodierungsverfahren DSD und PCM (176,4kHz/24Bit) zu differenzieren" verlief negativ. Laut dieser Studie war es den Probanden nicht möglich zwischen den Kodierungsverfahren zu unterscheiden: "Aufgrund der Resultate des vorliegenden Hörvergleichs hat sich gezeigt, dass selbst mit hochwertigstem Equipment unter optimalen Abhörbedingungen und unterschiedlichster Hörfokussierungen bzw. Hörerfahrungen der Probanden in der Regel keine signifikanten Unterschiede zwischen DSD und High Resolution PCM (24bit/176,4kHz) hörbar sind, sich demzufolge die These aufstellen ließe, dass sich keines der getesteten Systeme durch klangliche Eigenschaften hervorhebt". Insofern scheint die Filterproblematik in der Praxis nicht zu wahrnehmbaren Unterschieden zu führen.

Unabhängig davon, gibt es auch im PCM-Bereich Wandlungs-Methoden, die keinerlei Filter-Ringing aufweisen. So verwenden die Digitale Audio Systeme NOS-DAC-Technologie, die auf den Rekonstruktionsfilter in der Digital/Analog-Wandlung verzichten kann und somit ebenso keinerlei Filter-Ringing erzeugt (Details zur NOS-

DAC-Technologie finden Sie hier: www.digital-audio-systems.com/technischer-hintergrund/digital-analog-converter-techniken).

PCM und DSD sind beides gute Distributionsformate für den Consumer-Bereich. PCM ist deutlich einfacher zu bearbeiten und hat daher viele Vorteile auf der Produktionsseite, wo es auch bei Tonstudios fast ausnahmslos - auch für DSD-Produktionen - zum Einsatz kommt. Auf der Wiedergabeseite beim Konsumenten sind beide Formate bestens geeignet, um für glückliches Musikhören zu sorgen. Es gibt viele SACDs und DSD-Dateien im Umlauf und die Konsumenten sollten sie bestmöglich wiedergeben können. Ansonsten ist recht, was gefällt.

Einen prinzipieller Klangvorteil für das DSD- oder PCM-Format ist aus den technischen Gegebenheiten nicht nachweisbar, zumal es kaum DSD-Produktionen gibt, die nicht in der einen oder anderen Phase ihrer Entstehung eine PCM-Produktion waren.

Die 64xDSD- und 96/24 PCM-Formate sind als Modulationsverfahren in fast jeder Hinsicht den Komponenten selbst der besten Aufnahme- und Wiedergabeketten weit überlegen. Die Bandbreiten beider Verfahren übersteigen bei weitem die verfügbaren Bandbreiten der verwendeten Mikrophone, Verstärker oder gar Lautsprecher. Dasselbe gilt für die Störabstände. Auch hier übersteigen die machbaren Geräuschspannungsabstände beider Systeme bei weitem, was mit modernsten Mikrofonen, Verstärkern oder gar Lautsprechern erreichbar ist. Beide Formate übersteigen auch die Möglichkeiten selbst der besten A/D- und D/A-Wandler, so dass man sagen kann, dass das Modulationsverfahren kein begrenzender Faktor für die Audioreproduktion darstellt. Eine Fokussierung auf die vermeintliche Überlegenheit des einen Systems über das andere lenkt die Aufmerksamkeit nur von den tatsächlich wichtigen Themen in der Aufnahme- und Wiedergabekette ab.

© Alexej C. Ogorek

Quellen:

1. Andreas Koch, "DSD – the new Addiction", 2012
2. John Siau, Benchmark (<http://www.realhd-audio.com/?p=804>) Mai 2013
3. Dominik Blech, Min-Chi Yang: "Untersuchung zur auditiven Differenzierbarkeit digitaler Aufzeichnungsverfahren", Hochschule für Musik Detmold, Erich-Thienhaus-Institut 2004