"Reshaping Digital Audio: DSD Encoding as a Viable Alternative to PCM"

Honors Thesis
Research Project
Fall 2004

Brandon M. Schexnayder

Acknowledgements

I would like to thank all the folks who helped out with this research-- from guidance to moral support, enthusiasm to criticism.

A special thanks goes out to:
Dr. James Piekarski, Thesis Advisor
Dale Brown, Technical Advising for Listening Tests
Nathan Adams and the MTSU AES Chapter
Professor Chris Haseleu, Department Head
Keith Schexnayder, Electronics "Tutor"
As well as Tyler, Elizabeth, Bryan and Mom for moral support.

Contents

Abstract

This study presents an in-depth analysis of the one-bit digital audio encoding method known as Direct-Stream Digital, specifically focusing on its viability as an audio recording format in comparison to the current standard of Pulse Code Modulation. The assessment of this relatively new format's viability requires the validation of two distinct parameters. The first is an objective assessment of the encoding process's merits in recording and reproducing audio waveforms as accurately as possible. Examples of such parameters are the system's frequency bandwidth and signal-to-noise ratio. A second assessment needed is a subjective analysis, which includes individual human perceptions of the reproduced audio of each system. While a piece of audio technology can never reproduce an analog audio signal completely accurately, it is the inherent inaccuracies that characterize each system, thus a subjective look at them is required. This study explores DSD as a viable alternative to PCM under the scrutiny of both of these independent factors.

I. Introduction

Analog-to-digital conversion is a meticulous process, combining mathematical calculations and intricate electronic designs, to transform the infinitely complex and time-continuous values of analog audio into a time-discrete sequence of finite binary values that digital systems can comprehend, process, and store. Limited by the performance of computer hardware and data storage devices, converted audio is essentially stripped of an immense amount of naturally occurring properties and values that cannot be reintroduced to the signal when converted back to analog for listening applications. While an infinite number of conversion principles exist, "there are only a small number of conversion principles (architectures) that have been demonstrated successfully" based on accuracy standards and hardware limitations (Story 145). Pulse Code Modulation (PCM) is one of these conversion architectures.

Developed by Sony and Phillips for the Compact Disc (CD) over two decades ago, PCM has been overwhelmingly prominent in the world of digital audio. Increases in computer technologies have even allowed the PCM architecture to improve through the utilization of higher sampling rates and larger bit depths than the traditional standards. As a result, audio quality has increased tremendously. Despite this immense progress though, the "incremental improvements in PCM digital audio [are] becoming smaller and smaller," and less apparent to the ear ("Super Audio Compact Disc: A Technical Proposal" 1). By acknowledging this problem of diminishing returns, it is clear that a "fundamental shift in our approach to digitizing and delivering audio is in order" if digital audio is going to improve further (Smithers 1). Rather than focusing on future improvements to the multi-bit PCM structure, Sony and Phillips recognized the need for a shift in ideology for digitized music and developed a single bit digital encoding structure called Direct-Stream Digital (DSD).

The purpose of this new technology is clear—convert audio into the digital domain with greater quality than PCM. In order to validate the success of this purpose though, two uniquely independent elements of the technology must be examined: the objective quality concerning conversion accuracy, frequency response, and dynamic range; as well as, the subjective quality of the encoded music with regards to the aural preferences of consumers. By analyzing any recording technology from these two perspectives it becomes clear whether or not the quality of reproduced audio supports the technology as a viable form of reproduction, specifically in respect to the older technology it aims to replace. This paper discusses DSD through these two determinants, attempting to ascertain whether or not DSD encoding is a viable alternative to PCM.

II. Objective Analysis of DSD

A. Framework of A/D Conversion

Both PCM and DSD analog-to-digital converters (ADC) fulfill a single purpose: convert an alternating current (AC) voltage, which electronically represents analog audio, into a binary stream of data. While both architectures effectively complete this task, differences exist between them due to two main reasons. First, the procedure used to convert an audio waveform into binary data is uniquely dependent on the type of ADC and coding method being used; and second, the output sequence of binary data used to represent the original audio can only be utilized within its respective system. It is important to understand these two differences, because they are the main factors contributing to the quality and unique characteristics of the digitally encoded audio.

The methods used for PCM encoding, and the resulting data stream are more complicated than those used, and produced, by a DSD ADC. Theoretically, PCM encoded audio only undergoes a simple two-step process for conversion. First the waveform is introduced into the ADC and combined with an AC pulse, or sampling frequency (F_s), produced by the ADC. The audio waveform modulates this pulse and creates samples of the original audio at time-discrete intervals based on the ADC's respective sampling frequency (Fig. 1). Once a sample has been created, it is then assigned a value during the second process—quantization. This value is a binary word, with a length determined by the ADC's bit depth, or resolution (Fig. 2). After the completion of this quantization process, the ADC has completed conversion and output a sequence of binary data that contains time-discrete binary words representing an approximated amplitude value of each sample. This somewhat simplistic, two-step, theoretical approach becomes much more complex in reality though, as additional components must be used in order to achieve higher reproduction accuracy.

Fig. 1 "PCM Sampling"

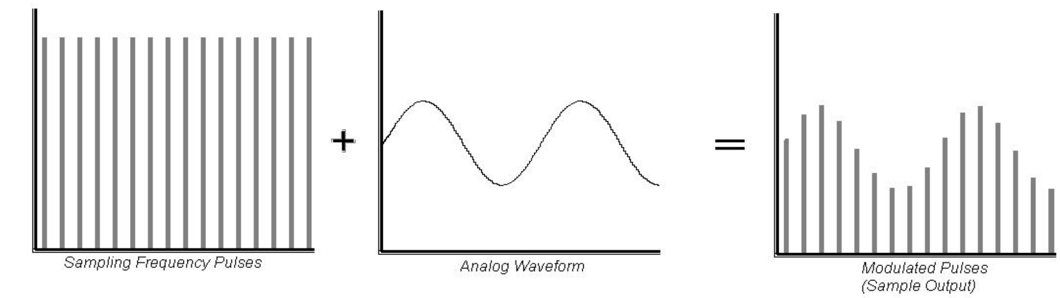
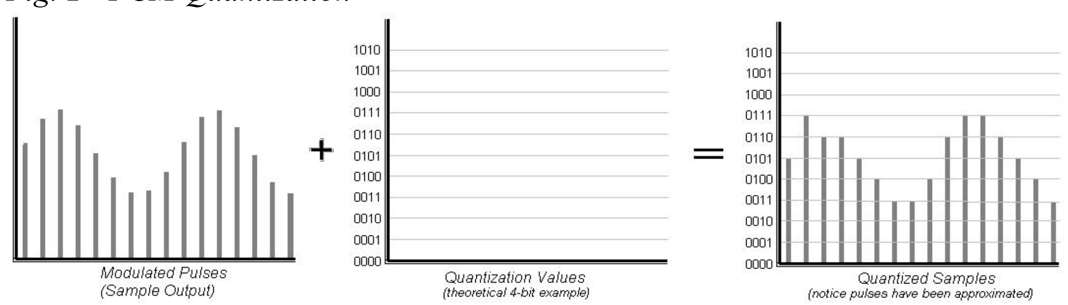
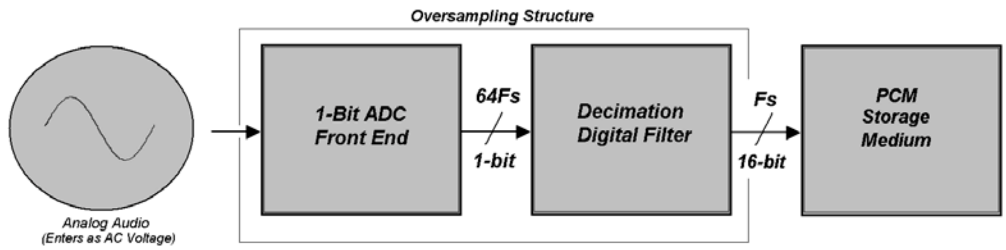


Fig. 2 "PCM Quantization"



While the actual output of a typical CD quality PCM converter would contain a sequence of 44,100 samples per second, the original audio was not actually sampled at 44.1 kHz. Instead, an over-sampling filter is used to sample audio sixty-four times the amount required by the output specifications. During this time, the over-sampling filter quantizes the signal, producing a sequence of samples represented by only one-bit. This quantization step, actually done within the over-sampling filter, does not assign a concrete value to the sample, but instead feeds the single bit sequence into the decimation filter to carry out this process. During decimation, excess samples are discarded and the remaining samples are assigned their multi-bit value as defined by the ADC's output expectations (Fig. 3).

Fig. 3 "16-Bit/44.1kHz PCM Over-Sampling Architecture"



In essence DSD's design is a simpler approach to this complex process of an over-sampling PCM ADC. By eliminating the re-quantization and decimation phases of a PCM ADC, audio stays at a sample frequency of $64F_s$ (2.8224 MHz) and is only filtered and quantized once—by the over-sampling filter, technically named the Sigma-Delta Modulator (Fig. 4). As the name implies, a Sigma-Delta modulator operates by comparing the sum (sigma) of a sample's amplitude to the change (delta) in amplitude of the previous sample (Fig. 5). Rather than recording the exact change in amplitude, one basic principle guides the conversion—if an analog waveform within a given bandwidth

Fig. 4 "DSD Conversion Architecture"

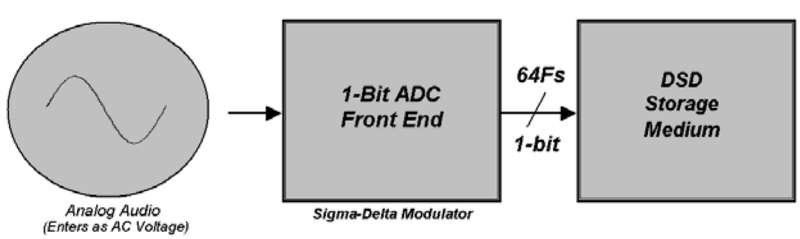
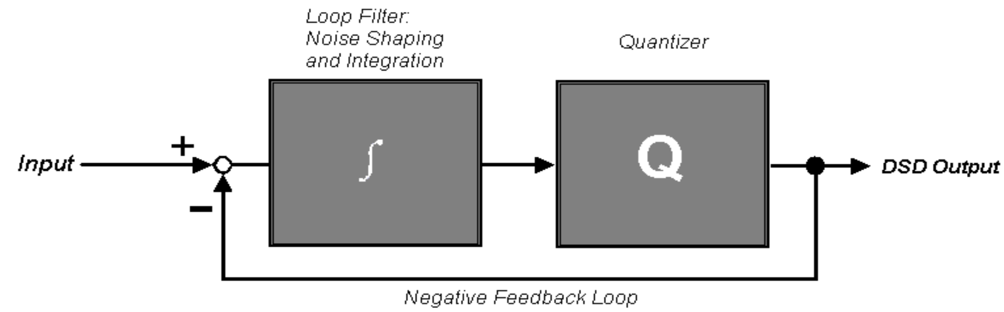


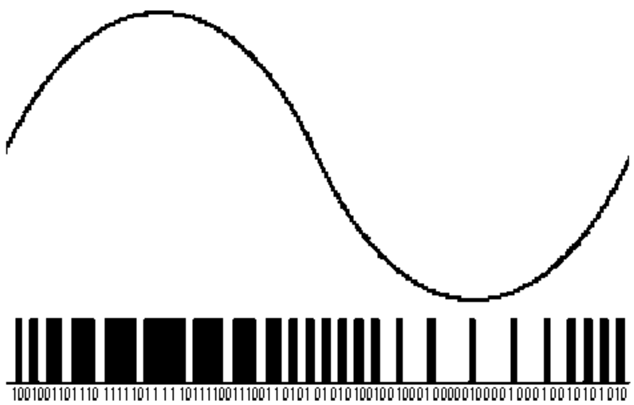
Fig. 5 "DSD Sigma-Delta Modulator"



is sampled at a high enough rate, all changes in amplitude between samples average out to be approximately the same value. Thus, sampling at $64F_s$ only records positive or negative changes in the amplitude between samples, not the value of the samples or their relative change. The actual architecture of the SDM, as presented in Fig. 5, accepts the analog input into the ADC at the filter stage. This filtered signal is then fed to the quantizer where it is sampled. The samples are then returned to the beginning of the ADC through the negative feedback loop, where the original sample's amplitude is compared to the sample immediately following. This process continues the rotation as long as audio enters the ADC. The resulting output of the DSD ADC is a binary data stream where only one bit is assigned to each sample: a "one" for positive amplitude change, and a "zero" for a negative (Fig. 6).

This SDM architecture is not unique to DSD though, as a similar version of the SDM design is employed in a PCM over-sampling filter. The difference between the PCM SDM design and DSD design is precisely where the output data stream is fed, as well as what processes will be implemented concerning this output within the rest of the analog-to-digital conversion. In PCM structures this stream of data is decimated

Fig. 6 "DSD Representation of a Sine Wave"



following the SDM, down-sampling the one-bit output of the SDM to a multi-bit data stream to a form compatible with the needs of that specific PCM system or storage device (Fig. 3). In contrast, the DSD structure utilizes the SDM's one-bit output without any further processing. This essentially makes the SDM the only component in DSD's entire ADC. In these regards, the encoding process used for DSD seems to be not only a simpler approach to analog-to-digital conversion, but also the source of a much simpler data stream than that of a PCM ADC. Consequently though, this simpler data stream actually retains more information from the original analog waveform than either of the PCM encoding procedure.

B. Basic Properties of Digitally Encoded Audio

The differences in the conversion procedures and types of binary output of PCM and DSD directly affect the basic properties of the audio being reproduced. Such attributes as dynamic range, quantization error, frequency bandwidth, stereo localization, transient response and signal distortion are all affected by the parameters inherent to the encoding method used in analog-to-digital conversion. In order to compare these differences between DSD and PCM though, a distinction must be made between two different types of PCM. The first is the standard PCM found on the CD (44.1kHz/16-bit PCM), and the second is the more modern, high-end PCM found on the DVD-A (192 kHz/24-bit PCM). Both of these types are accepted as consumer formats and must be analyzed in comparison to DSD in order to ascertain DSD's complete viability in regards to all forms of PCM.

Dynamic Range and Quantization

One finite difference between the PCM and DSD encoding methods is the way each represents the possible dynamic range, or amplitude, of reproduced audio. In PCM encoding, the bit depth used during conversion is the direct determinant of the possible dynamic range a signal is able to utilize; whereas, in DSD the dynamic range is determined by the sampling rate. The accepted principle guiding a PCM system's dynamic range is that a single bit is capable of representing approximately 6 dB of amplitude. Thus, a 16-bit system can theoretically encode up to 96 dB of dynamic range and a 24-bit system is theoretically capable of 144 dB. Realistically though, it is impossible to recreate 144 dB of dynamic range due the quality of current physical hardware converters. Due to both the immense dynamic range, and its logarithmic nature, converters would theoretically be forced to handle voltages as low as 10 nV (one-billionth of a volt). Noise generated by modern converters mask signals this low in voltage though, and in order to avoid this problem, the incoming line level is boosted by approximately 7 dBu. Unfortunately, this level boost does infringe on the 24-bit capability of the system, limiting the system's performance to a dynamic range that can only fully utilize approximately 20-bits. The result is that 24-bit PCM's dynamic range is approximately 120 dB.

DSD systems, which include only a single bit per sample, are not guided by the same dynamic range standards of PCM. Rather than achieving dynamics through bit resolution, the range is attributed to the high sampling rate of 2.8224 MHz. As previously discussed, each single-bit sample within the DSD signals represents a change in the signal's amplitude rather than a finite value. This allows for the availability of 2.8224×10^6 quantization levels. In contrast, a 16-bit PCM system contains 65,536 and a 24-bit system contains 16.777×10^6 quantization levels. While this may seem to prove that the resolution of DSD encoded audio is much smaller, DSD is also affected by the same electronic faults as 24-bt PCM, but similarly, it can still produce a dynamic range of approximately 120 dB. The inherent flaw with the DSD amplitude quantization however, is that it is much noisier than the quantization noise in a PCM system. This problem is addressed later though, with respect to each systems frequency bandwidth, dither and noise-shaping abilities—all tools in diminishing the effects of quantization error noise to a practically inaudible level.

Despite the altered context for the creation of dynamic range and quantization procedures, both high end PCM and DSD are equally qualified to encode modern audio signals. Both of these architectures greatly outperform the ability of a CD quality PCM system, allowing a larger dynamic range and greater resolution to low-level signals that can be lost due to limited quantization levels. The superior dynamic range also allows such musical events as snare hits and other transient based instruments. More amplitude headroom allows audio engineers to exploit greater peak ratios without digital clipping or the need for compression, while retaining superior digital resolution. These extended dynamics from louder transients do not affect the overall perceived volume, but do allow an instrument's timber and overall dynamic nuances to have a more audibly realistic quality within the music.

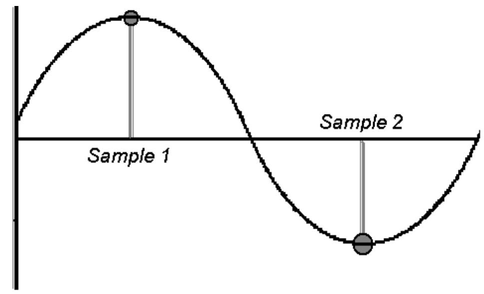
Aside from audio amplitude, several other factors concerning audio quality are affected by the dynamic range and quantization process discussed here; however, both are in conjunction within the second major element of digital: the frequency bandwidth.

Frequency Bandwidth

Just as all digital systems are restricted to a finite dynamic range, the frequency response of an ADC must also conform to specific parameters. Both DSD and PCM architectures produce this set frequency response with respect to the Nyquist Theorem, which states, "in order to digitally represent a given frequency, the sampling rate must be at least twice that frequency" (*Audiomedia* 1). For example, a frequency as high as 20 kHz requires a minimum sampling rate of 40 kHz, allowing two samples to represent the positive and negative peaks of the waveform (Fig. 7). Since all frequencies above the

Nyquist entering a system will cause an alias tone, or a false representation of that tone as a frequency lower than the Nyquist (Fig. 8), each ADC must utilize a low-pass filter (LPF) prior to any encoding. Traditionally, CD quality PCM converters filter all frequencies above 22.05 kHz.

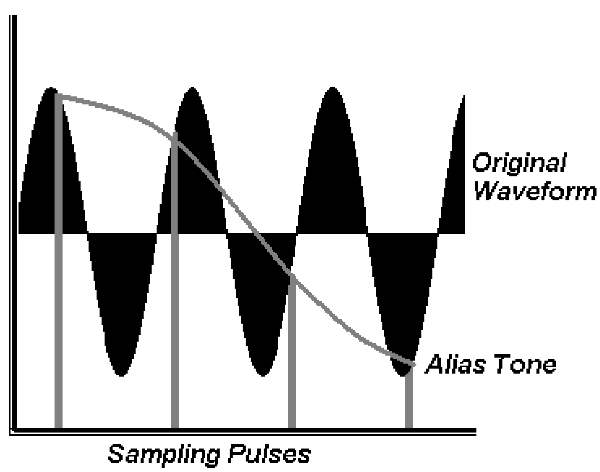
Fig. 7 "Nyquist Frequency at ½F_s"



With respect to the typical human aural frequency perception, this standard may seem enough; however, a LPF in these instances is forced to roll off frequencies very close to the audible range.

The results of this low-Q, or "brick-wall," filtering are several audibly apparent anomalies such as oscillation, phase shift and high frequency loss. In addition to these results, the "steep low-pass filters create pre-echoes which the ear interprets as a loss of transient response, obscuring the sharpness or clarity of the sounds" (Katz, 225). Due to these common affects found in low-Q filters required for PCM, it is a logical step to reference the steepness of the anti-aliasing filter's slop when addressing the quality of a digital system. Of course, this slope is able to have a much lower Q when

Fig. 8 "Effects of a Frequency Above the Nyquist"



an ADC is able to utilize a larger frequency response, allowing the least amount of artifacts to be introduced to the audible incoming signal in audible regions.

The audio band entering an ADC can be interpreted as three separate frequency bands in respect to the LPF: pass band, guard band, and stop band. The pass band is the frequency spectrum allowed to enter the encoding process without any filtering. It is in the guard band where a gradual slope of attenuation begins, terminating at the stop band where all frequencies are attenuated at least 100 dB. This attenuation is well below the dynamic range of both PCM and DSD converters. The contrast between traditional PCM, high-end PCM, and DSD converters, with respect to the low-pass filtering, is the width of each of these audio bands. Since a 44.1 kHz PCM ADC allows a total frequency response of approximately 5-22,050 Hz, the pass band includes all frequencies below 20 kHz, the guard band includes frequencies between 20-22.05 kHz and the stop band consists of everything above 22.05 kHz (Fig. 9). High-end PCM sampled at 192 kHz has a much broader range to work with, as the Nyquist frequency is 96 kHz, allowing a much larger guard band with a more gradual filtering slope. While this high-end PCM undoubtedly allows greater quality, by introducing fewer artifacts from steep filtering, the nature of DSD in regards to filtering is slightly different (Fig. 10).

Fig. 9 "Freq. Spectrum of 16-Bit/44.1 kHz PCM" (Not to Scale)

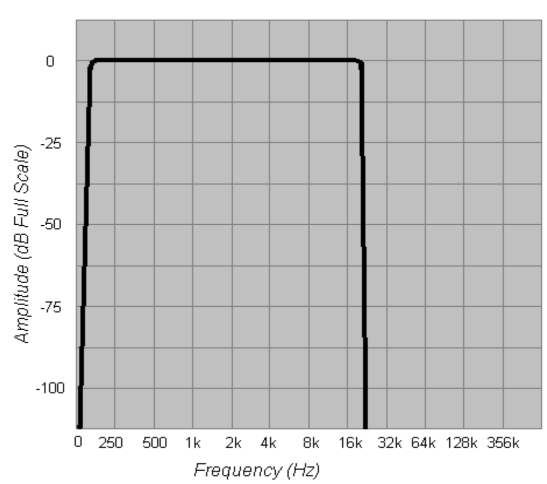
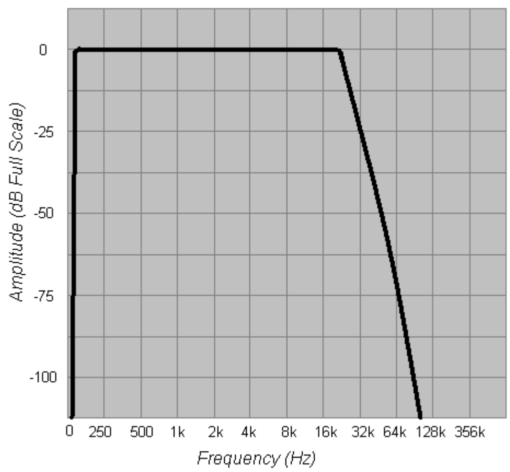
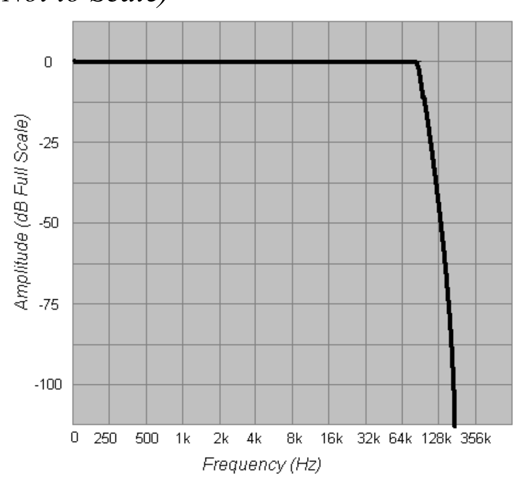


Fig. 10 "Freq. Spectrum of 24-Bit/192 kHz PCM" (Not to Scale)



Unlike PCM filtering methods that utilize the entire bandwidth capable under the Nyquist theorem, DSD does not use all frequencies available within its system. Instead of encoding all frequencies below the Nyquist frequency of 1.4 MHz, a DSD ADC creates a flat frequency pass-band from DC-100 kHz, avoiding frequencies higher than 200 kHz. The guard-band and low-pass filtering then begin, following 100 kHz with an extremely gradual slope that does not reach -100 dB until approximately 200 kHz (Fig. 11). Not only does this gradual sloping LPF produce less artifacts as the steeper filters used in PCM ADCs, but the fewer artifacts created only affect frequencies more than four times higher than the highest audible frequency. These attribute alone makes DSD a stronger option during digital encoding.

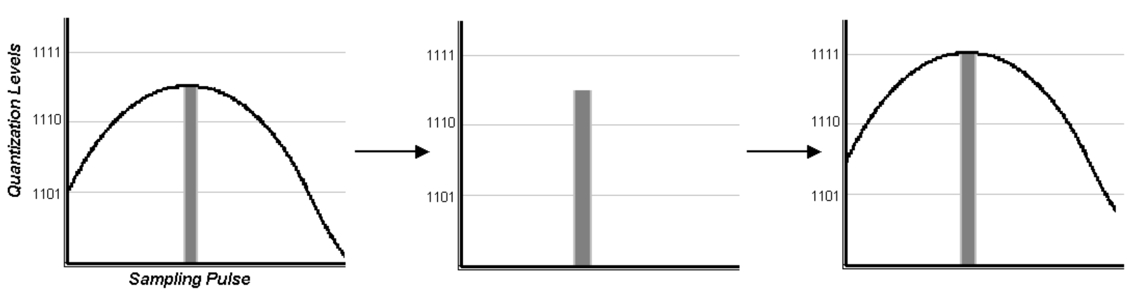
Fig. 11 "Freq. Spectrum of DSD" (Not to Scale)



Noise Shaping

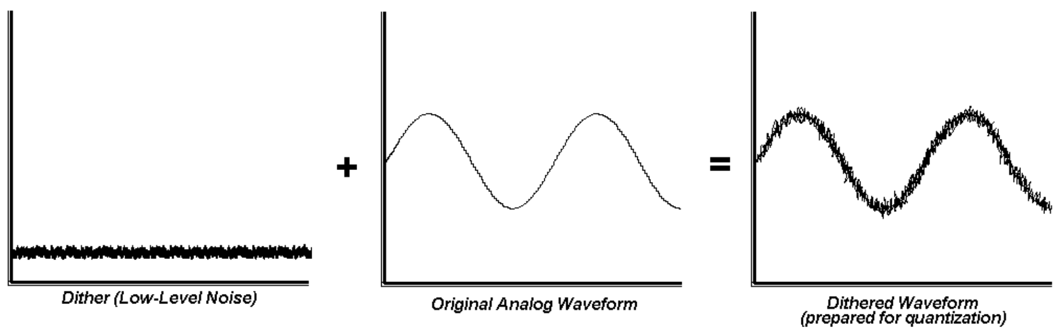
The large frequency response of both high-end PCM and DSD allows inherent noise issues, such as dither and quantization error, to be addressed. One of the largest problems to occur when encoding analog audio during analog-to-digital conversion is the presence of audible errors created during the quantization procedure. As previously discussed, quantization is the process of assigning amplitude values to samples by approximating the original amplitude to the nearest finite value the ADC recognizes. While these analog amplitude values have an infinite number of significant figures, digital values do not and compromises must be made during encoding. The quantization error that ensues is the difference between the assigned voltage value and the original waveform's voltage at that sampling instant. The worst case scenario occurs when the analog amplitude value falls directly between two quantization levels (Fig. 12). Whichever value the system assigns will obviously be a distortion of the signal, and the ensuing "quantization error is audible as a rough, granular sound" (Harley 538). DSD unfortunately suffers from more of this effect than PCM; however, as seen later, is able to combat the issue much more effectively.

Fig. 12 "Quantization Error Example"



Both PCM and DSD utilize a similar practice to prevent quantization error and its audibly intrusive results. The ADC adds a constant stream of random noise to the analog waveform prior to the quantization process. This addition, called dither, is a fixed low-level noise that helps replace "highly undesirable distortions entirely," by allowing greater quantization resolution (Fig. 13). Unfortunately, dither is also the result of a lower signal-to-noise ratio. While this may not be a flawless way of reducing digital distortions, the dither noise is less audibly intrusive than those audible effects of

Fig. 13 "Addition of Dither to a Waveform"



quantization error. In addition, since the dither is added by the ADC, it can be filtered in specific ways to ensure that it rests at intervals within the frequency spectrum where it is aurally masked. This technique, called noise-shaping, "re-equalizes the spectrum of the dither while retaining the power, effectively moving the noise away from the areas where the ear is most sensitive (circa 3 kHz), and into the high frequency region," as seen in Fig. 14 (Katz 55).

Both high-end PCM and DSD are at an advantage during the process of noise-shaping because of the size of their extensive bandwidth. Since most frequencies under the 22 kHz Nyquist frequency are utilized by recorded audio in the PCM system, noise-shaping must follow psychoacoustic principles to place the dither signal where it is least audible (mainly 10-22 kHz); however, 192 kHz PCM and DSD can displace the majority of this noise above the audible spectrum completely, reducing the audible effects by much more than the 3 dB that lesser quality PCM produces. As previously mentioned, one major pitfall during DSD quantization is that "the quantization error and noise associated with sigma-delta modulation is very large, resulting in significant quantization noise;" however, unlike the PCM structures this "noise can be shaped such that virtually all noise power falls outside the range 0-20 kHz," well above the range of human aural perception (Nuijten 27). An even further distinction separating DSD and PCM is the ability of DSD filters to shape the dither with a much less complicated linear equalization curve (Fig. 15) that is not forced to rely on less audible frequency regions below 20 kHz.

Fig. 14 "44.1 kHz PCM Noise-Shaping Example"

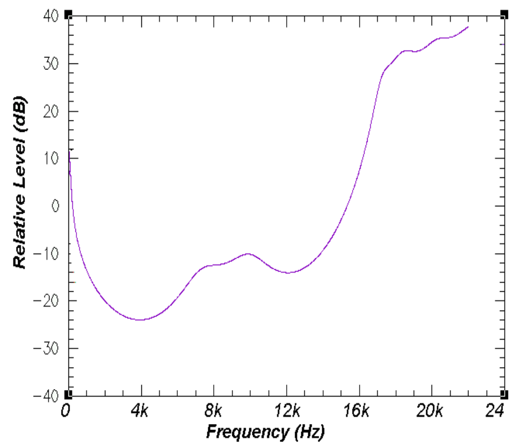
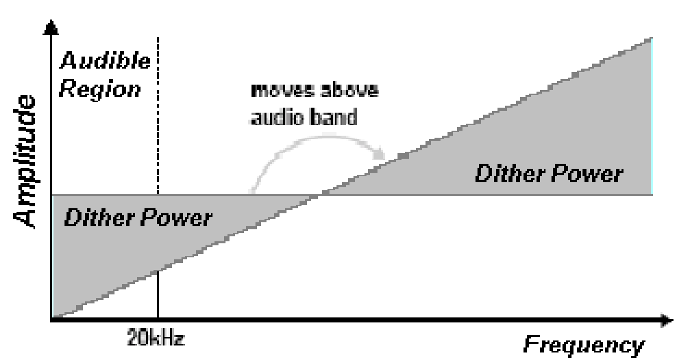


Fig. 15 "DSD Noise-Shaping (Approximation)"

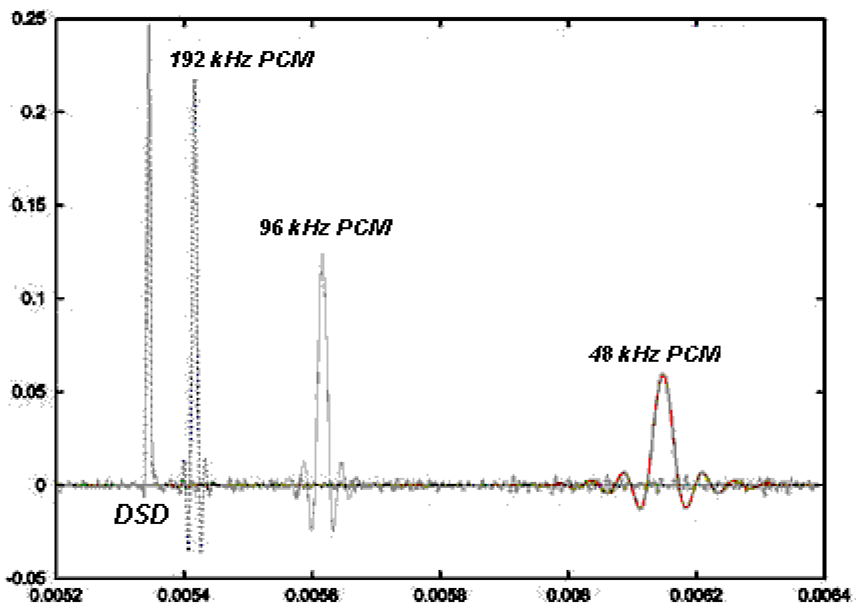


Inherent Effects of Sampling Rate

Aside from the sampling frequency's ability to create a large frequency response in a digital system, this digital audio parameter also affects several other elements of encoded audio, particularly the stereo localization and impulse response. These two attributes are tied together because they are both affected by the time-domain spread created by the sampling rate. Since the sampling process is essentially a quantization of samples on a time scale rather than on an amplitude scale, signals occurring between those instances are cropped. In a 44.1 kHz system the time between each sample is fairly large, with approximately 22.68 μs between each sample. On the other hand, 192 kHz PCM samples are only 5.1 μs apart, and DSD has an even greater time domain resolution with a mere .357 μs between samples. As a result, the 192 kHz PCM and DSD systems are better suited to respond to transients accurately. Figure 16 demonstrates how this response time actually affects the system's ability to reproduce a transient, as each system was fed a –6 dB block input (click) of a 3 μs duration. The resulting graph shows that the DSD and 192 kHz PCM systems respond the fastest and most accurately; whereas the 48 kHz PCM system not only distorts the signal, but also takes a much longer time to even react. This distinction, most visible in the large width of the 48 kHz sampling frequency reproduction, is audibly apparent in transient events as a ringing or "bell-like" sound.

The binding factor between a digital system's transient response and stereo localization abilities is a psychoacoustic principle relating that "most people can hear a time delay of 15 milliseconds or more" (Moorer 1). If two sounds are played more than 15 milliseconds apart, they will be audibly perceived as two distinctly separate sounds. Obviously a 44.1 kHz system does not have a large enough sample rate time to utilize this principle to its fullest extent, as seen in the smearing of the audio reproduced in a 48 kHz system, causing an imaging issue that results in a blurring of the stereo soundstage. This is a direct result of the same factors that cause the "ringing" sound in transient events. In contrast to the extremely inaccurate width of the 48 kHz reproduced sample seen in Fig. 16, both DSD and 192 kHz PCM are able to accurately reproduce a 3µs click almost instantaneously, without any blatantly apparent inaccuracies.

Fig. 16 "Impulse Response of a –6 dB Block Input of 3µs Duration"



C. Analysis of Objective Research

It is clear that CD quality PCM is outdated in respect to both DSD and 24-bit/192 kHz. However, the properties and attributes of both DSD and high-end PCM are remarkably similar. They each offer the same dynamic range, similar frequency response, and extremely practical options in regards to noise shaping and error correction. While each utilizes its own unique methods, resulting in a very similar output, DSD still proves to be a viable alternative to even 24-Bit/192 kHz PCM. As previously discussed, DSD is simply a simpler approach to the PCM conversion process. Since it is merely the utilization of the output of a SDM without a decimation filter employed, its bitstream has the unique ability to be decimated at any time in order to be used as a 192 kHz, 96 kHz, or even 44.1 kHz PCM stream without any complex mathematical equations that can degrade audio quality. This is because more analog information can be digitally encoded and recorded by a DSD system than by any current PCM ADC, while maintaining a sample rate that is a direct multiple of other professional standard sampling rates. Thus, objectively DSD is theoretically a completely viable alternative to PCM A/D conversion.

III. Subjective Analysis of DSD

A. Purpose of Subjective Testing

The objective analysis of any type of audio technology is capable of revealing telling attributes about the quality of the reproduced audio; however, "critical observational listening can reveal aspects of audio equipment quality not exposed by traditional measurement techniques" (Harley 551). The largest dilemma with objective test results is that one factor is typically not assessed: the subjective response consumers have in regards to the reproduced music. Despite the preceding research's conclusion that objectively DSD is a viable alternative to PCM, the true test of the encoding system is how it sounds and how consumers respond to it.

In order to assess this consumer response, sixty MTSU Recording Industry students and several faculty members participated in a series of blind A/B listening sessions in which they were asked to make subjective conclusions in regards to several digital formats they were presented with. The goal was not to define which encoding format was the best, but instead it to provide insight into the consumers' ability to actually distinguish between these formats and ascertain if the possible audible differences between them were great enough to prove the viability of DSD in the subjective domain. Essentially, the results of this listening test weigh much more greatly than those found in the preceding objective analysis.

B. Basic Structure of Listening Survey

The research took place in MTSU Critical Listening Lab B on October 6, 2004 within the context of three separate, identical listening sessions. During each session, participants were explained the expectations, purpose and goals of the listening tests. Each was given a pre-fabricated survey containing questions relating to their own personal opinions regarding musical recordings they were listening to. Participants were then asked to listen through the selections on headphones four separate times. During the first, all three digital formats were available to switch through: Compact Disc (44.1 kHz, 16-Bit PCM), DVD-Audio (192 kHz, 24-bit PCM) and Super Audio CD (DSD). Each participant had personal control, through the use of a headphone distribution amp at each station, of switching between these formats with the knowledge of what format they were listening to. During this test only one question was asked: "which of the three formats do you feel is the most aurally pleasing?" This allowed participants to not only formulate their own opinions regarding the formats' individual strengths and weaknesses, but also allowed them to make an assumption as to whether or not they actually felt there was an audible difference between these mediums. Once the "open-eye" test concluded, and participants were prepared, three blind tests followed in which only two formats were presented at a time. During these three tests participants were completely unaware of what format combinations they were assessing, and were asked to complete a survey based only on what they heard. Following the completion of the tests, each survey sheet was tallied up and results were analyzed to form a conclusion based on their individual, subjective responses to the musical recordings presented.

C. Preparation Factors

In order to make the most accurate scientific assessment of such a subjective experience as musical preference, while maintaining the scope of this research, several factors had to be addressed in regards to the test's implementation. Numerous variables were accounted for: identical signal paths for each test format, matched volume levels between formats, participants' placement in relation to both the listening room and recording's stereo field, the participants' control over instantaneous format switching and the presentation of identical musical material on all formats. It was with utmost importance that each of these factors was addressed with extreme accuracy so that the only variable left to affect the participants' subjective interpretations was their own perception of the musical formats.

The selection of a proper musical recording was the first variable addressed in preparation for the research. The recording had to fulfill several requirements in order to address the research goals intended. First, the musical recording must have been released in stereo on all three formats (while multi-channel releases are available on the SACD and DVD-A, they presented within the scope of the project). Second, since different mastering practices exist between the CD, DVD-A and SACD, the music on each of these formats must have all been taken from the same identical analog source without any type of digital processing until the final encoding phase for the respective digital format. Any type of PCM processing prior to the analog to DSD conversion could greatly affect the pure presentation of the quality of DSD encoding. The recording chosen, one of the few available that fulfilled all of these requirements, was the Diana Krall song "S'Wonderful," from her 2003 release, *The Look of Love*.

The signal path and playback devices were the second most important aspect of presenting the most controlled environment for listening evaluations. Aside from the initial "eyes-open" test, each format was played on an identical universal player. The model chosen was a Pioneer DV-578A-S, specifically because it utilized the same digital-to-analog converter for both PCM and DSD recordings. Had different players been used to playback different formats, a variable concerning the player's impact on digital-to-analog conversion would have arisen and negated all results. The RCA analog outputs of each player were then fed into a small pre-fabricated box containing several potentiometers for matching playback level between formats. The need for this box will be discussed later. These potentiometers then fed a balanced amplifier that distributed the signal to every headphone station in the listening lab. From there every participant was given an identical set of Sony MDR-7506 headphones to listen to the material on. The key to this entire signal path is that every format was affected by the same electronic circuits, avoiding any possible variables between formats that could have affected the final audio output heard by each individual participant.

The importance of line-level matching between digital formats is possibly the largest concern regarding the accuracy of results. In order to ensure this standard, the analog outputs of the playback device were recorded for each of the formats into a Cubase SX computer system with the same output volume set. The computer software then analyzed the RMS, or average volume level, of each of these selections revealing a slight difference of approximately .01dB between the CD and DVD-A, and a much larger difference of approximately 7.98 dB between the CD and SACD. This vast difference in volume level had to be matched, and was done so through a series of potentiometers. In order to approach this properly, pink noise was output to each potentiometer (assigned to both the left and right channel for each format), which then attenuated the CD and DVD-A formats to a volume level approximately 0.01 dB away from the SACD's RMS level. After once again measuring the output of the respective formats, all RMS playback levels were accurately matched when they were output to the power amplifier during the research sessions.

The final factor addressed in preparation for testing was the volume level that each participant received through their headphones. By sending pink noise through the power amplifier, each station measured approximately 74 dB on an SPL meter outfitted with an earpiece. This allowed the dynamic music on the digital formats to never peak above approximately 77 dB, and average at about 72 dB—a comfortable listening level with enough volume to perceive audible differences in selections, as well as accurate resolution. The output voltage of each headphone distribution box was then set to the same level, and the final preparation for testing was complete. Each of these factors was addressed with the most accuracy available within the scope of this project.

D. Test Results

The data collected involved several aspects. The first was a two-part question regarding whether or not the participant felt there was an audible difference between the two formats, as well as a clear cut question asking which format the participant preferred. The second section asked each individual to relate specific subjective attributes they felt influenced this choice. Finally, participants were asked to relate how deeply they felt the audible differences were between the formats. This was done on a scale of one to five.

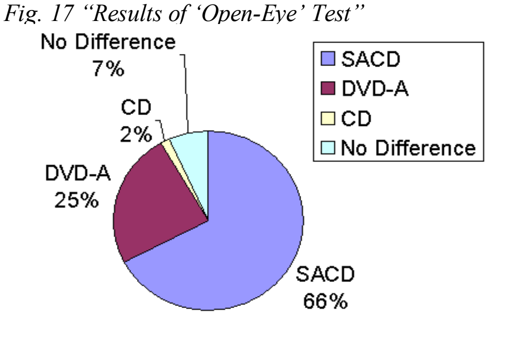
In regards to this audibility scale participants were clearly asked to judge their answer based on difference, not preference. Several other questions were asked in regards to how participants felt about their choices; however, these were done in order to assess patterns between listeners, not assess overall viability or quality.

Making critical assumptions from this collected data was done through placing specific importance of each of these aspects independently, with specific attempts to focus mainly on those questions that directly relate to the purpose of the research—an assessment of the *viability* of DSD. With this in mind, three determinants weighed the heaviest in analyzing the data: is there an audible difference between the PCM formats and DSD, what is the intensity of this difference, and does any format overwhelmingly outperform the others. Since it is the viability of DSD that needs to be ascertained, not whether it is a higher quality than PCM, it is not important if the DSD media was the most popular format. The resulting raw data collected is located in Appendix C.

CD/DVD-A/SACD Open Eye Test

A total of 61 people took part in the initial test. With the knowledge of what formats they were listening to, the results weighed heavily in favor of the SACD: forty-one chose the SACD, fifteen chose the DVD-A, only one chose the CD, and the remaining four participants felt there was no difference in audio quality (Fig. 17). No additional questions were asked during this test time.

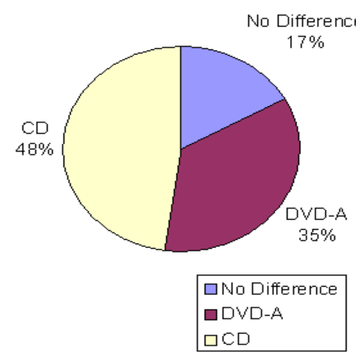
Fig. 17 "Results of 'Open-Eye' Test"



CD/DVD-A Blind Test Results

Out of the sixty-five participants blindly comparing the CD and DVD-A formats, only eleven felt that there was no difference between the CD and DVD-A (an audibility score of one), twenty-three preferred the DVD-A, and thirty-one preferred the CD (Fig. 18). Despite this fact that 48% of participants preferred the CD, the group as a whole felt that there was only a small difference between the two formats, rating an average of 2.5 on a scale of one-to-five concerning how audible the difference was. In addition, only eight participants that chose the CD actually felt it was a dramatic enough increase in quality to warrant complete preference over the DVD-A

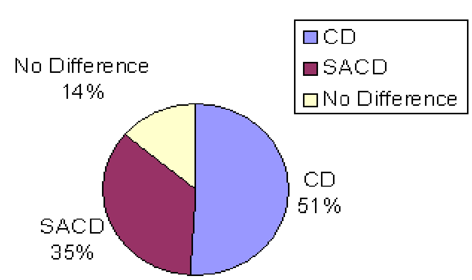
Fig. 18 "Results of CD/DVD-A Blind Test"



CD/SACD Blind Test Results

During the CD/SACD test, there was much more of a preference for the CD than seen in the CD/DVD-A test. Out of sixty-five participants: nine participants heard no difference, thirty-three preferred the CD, and only twenty-three preferred the SACD (Fig. 19). While the majority preferring CD was greater during this test than the DVD-A comparison, the average score concerning how audible the difference between the formats were, showed much less of an audible difference—

Fig. 19 "Results of CD/SACD Blind Test"

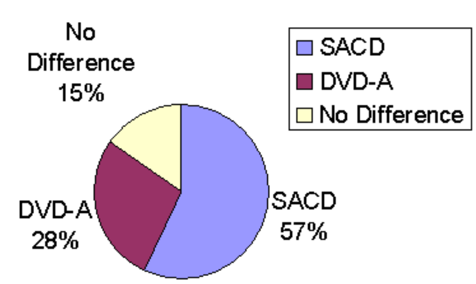


2.23 out of five. In addition, only six participants that chose the CD actually felt it was a superior format for their own personal use; whereas, seven participants that chose the SACD felt this way about their choice.

DVD-A/SACD Blind Test Results

Unlike the results seen in the blind comparisons with the CD, the SACD and DVD-A results showed a stronger preference towards the SACD than the same participants felt with the CD. Thirty-seven participants chose the SACD, eighteen chose the DVD-A, and only ten heard no difference (Fig. 20). Similarly, the audibility scale showed only an average of 2.4 in relation to the difference between the SACD and DVD-A.

Fig. 20 "Results of SACD/DVD-A Blind Test"



E. Analysis of Subjective Testing

The main conclusion drawn from the data collected is that no single format was overwhelmingly chosen as the participants' preference. Most participants felt as if audible differences were scarce, and assessed that these differences were not enough to even warrant a migration from their current personal listening format to the one they felt was more impressive. Conversely, no format was proven to be inferior, despite the objective research's assessment that DSD was capable of greater quality audio reproduction. In fact, the Compact Disc, with low-end PCM, showed a slightly higher preference among participants when individually paired with the other two high resolution formats. With this in mind, it can be deduced that the DSD encoded audio on the SACD is effectively a viable alternative to PCM, since their differences are nearly inaudible to most consumers, and slightly preferred in comparison to high resolution PCM.

IV. Conclusion

Without any regards to research done here, DSD has already proven to be a viable option within the field of audio. Enthusiastically embraced by more and more mastering engineers and audiophiles, the technology is still too new for its implementation in all aspects of audio recording. Similarly, it still is not compatible with most of the current audio equipment standards in the music production industry. Despite all proof that DSD is theoretically a viable option, the true test in regards to its success will be both how much the technology will be able to saturate into the industry and how well engineers will be able to utilize it to better recording.

Just as any other technology within the field of audio recording, the quality of DSD encoding lies more in the hands of those utilizing it than in the actual technology itself. While it is true that the objective attributes of DSD encoding produce a better and more versatile reproduction of analog audio, much of that audio's quality relies on how it sounds prior to ever entering analog-to-digital conversion. The result is that some DSD recordings may sound better than their high-resolution PCM counterparts, and others may not. Similarly, this is also a matter of personal preference. DSD is definitely a viable and higher quality alternative to PCM in the professional recording field, but the one assumption that cannot be made from this research is that DSD will always meet the needs of every music consumer in the same way that a good analog or PCM recording might.

Appendix A
Works Cited

"DSD/SACD" *Audiomedia* Oct. 1999.

Harley, Robert. *The Complete Guide to High-End Audio.* Tijeras, New Mexico: Ascapella Publishing, 2004.

Katz, Bob. *Mastering Audio: The Art and the Science.* New York: Focal Press, 2002.

Janssen, Erwin and Derk Reefman. "Enhanced Sigma Delta Structures for Super Audio CD Applications" *Audio Engineering Society Convention Preprint* presented at the 112th Convention 2002 May 10-13 Munich Germany.

Masciarotte, Oliver. "The Other Optical Format" *Mix.* Jan 2001.

Moorer, James A. "New Audio Formats: A Time of Change, and a Time of Opportunity" <http://www.jamminpower.com/PDF/New%20Audio%20Formats.htm> 5 Aug 2004.

Nuijten, Peter and Derk Reefman. "Multichannel High-Density Audio Format: Does DSD (Bitstream) Have the Right Equation?" *Optical Disc Systems*. Jan. 2001 pp. 66-73.

Ranada, David. "Hold Your 1-Bit Horses!" *Sound & Vision.* Oct 2001, pp.35-36.

Reefman, Derk and Erwin Janssen. "One-Bit Audio: An Overview." *Journal of Audio Engineering Society*, Mar 2004, Vol. 52 No. 3.

Rumsey, Francis. *Desktop Audio Technology*. New York: Focal Press 2004.

Smithers, Brian. "A Better Mousetrap." *Electronic Musician,* May 2003.

Story, Mike. "Audio Analog-to-Digital Converters." *Journal of the Audio Engineering Society.* Mar 2004 Vol. 52, No. 3

"Super Audio Compact Disc: A Technical Proposal." Sony Electronics Inc. and Phillips Electronics N.V., ©1997, <http://www.sel.sony.com/SEL/consumer/dsd/dsd.pdf> 10 June 2004.

Appendix B
Survey Used for Listening Test

(participants received three identical copies like this)

1. *Was there a noticeable difference in audio quality when switching between from Selection A and Selection B?*	___ YES ___ NO
2 *Which audio selection is the better sounding of the two?*	___ Selection A ___ Selection B
3. *What is it that sounds better about the format you chose?(select as many as appropriate)*	__ Clearer __ Warmer __ More Lifelike OTHER:________________________
4. *What qualities do you dislike about the lesser format? (select as many as appropriate)*	__ Sounds Thin __ Tin-like Sound __Not as Clear __Brittle Sounding OTHER: _________________________
5. *How noticeable is the difference in quality between the two formats? (circle one response)*	1 (None) 2 3 4 5 (Extreme)
6. *Do you feel like the better sounding format is good enough to buy a completely new stereo system?* *Would you give up CDs for the better sounding system?*	___ YES ___ NO

Appendix B
Survey Used for Listening Test

Blind A/B Test #1: Compact Disc & DVD-Audio

			Positive qualities			*Negative qualities*				*Audible*	*Good*
	Audible	*Format*	*of your format choice?*			*of the "lesser" format?*				*Difference*	*enough to*
	Difference	*Choice*	*Clearer*	*Warmer*	*Lifelike*	*Thin*	*Tin-like*	*less clear*	*brittle*	*(scaled 1-5)*	*switch?*
1	Y	DVD-A	Y		Y		Y	Y		3	N
2	Y	DVD-A								2	N
3	N	-	-	-	-	-	-	-	-	1	N
4	N	-	-	-	-	-	-	-	-	1	N
5	Y	CD	Y	Y				Y	Y	2	Y
6	N	-	-	-	-	-	-	-	-	1	N
7	N	-	-	-	-	-	-	-	-	1	N
8	Y	CD			Y	Y				2	N
9	Y	CD		Y						2	N
10	Y	DVD-A	Y		Y			Y		3	Y
11	Y	CD	Y	Y		Y			Y	3	Y
12	Y	CD	Y	Y			Y			2	N
13	N	-	-	-	-	-	-	-	-	1	N
14	Y	DVD-A		Y	Y		Y			4	N
15	N	-	-	-	-	-	-	-	-	1	N
16	Y	CD		Y					Y	2	N
17	Y	CD	Y							4	N
18	Y	DVD-A	Y	Y	Y	Y				2	N
19	N	-	-	-	-	-	-	-	-	1	N
20	N	-	-	-	-	-	-	-	-	1	N
21	Y	DVD-A	Y	Y				Y		3	N
22	Y	DVD-A	Y					Y		2	N
23	N	-	-	-	-	-	-	-	-	1	N
24	N	-	-	-	-	-	-	-	-	1	N
25	Y	DVD-A	Y							3	Y
26	Y	DVD-A	Y		Y			Y		4	Y
27	Y	CD		Y	Y	Y		Y		2	N
28	Y	DVD-A	Y					Y		2	N
29	Y	CD			Y		Y			5	Y
30	x	-	-	-	-	-	-	-	-	-	-
31	Y	CD		Y	Y	Y				5	N
32	Y	DVD-A	Y					Y		4	Y
33	Y	CD		Y			Y			2	N
34	Y	CD		Y		Y	Y			2	Y
35	Y	DVD-A	Y		Y	Y				2	N
36	Y	DVD-A	Y					Y		3	N
37	Y	CD							Y	3	Y
38	Y	DVD-A						Y		2	Y
39	Y	DVD-A	Y					Y		3	Y
40	Y	CD		Y		Y				3	Y
41	Y	CD	Y					Y		2	N
42	Y	DVD-A		Y				Y		2	N
43	Y	DVD-A	Y	Y	Y	Y		Y		4	Y
44	Y	DVD-A	Y							4	Y
45	Y	CD	Y	Y	Y			Y		2	N
46	Y	CD		Y	Y	Y		Y		2.5	Y
47	Y	CD		Y				Y		3	Y
48	Y	CD			Y	Y				4	N
49	Y	DVD-A			Y					3	N
50	Y	CD		Y	Y	Y				4	Y
51	Y	CD		Y	Y	Y		Y		2.5	N
52	Y	CD	Y					Y		2	N
53	Y	DVD-A		Y		Y				3	N
54	Y	CD	Y		Y	Y				2	N
55	Y	CD	Y	Y		Y		Y		3	N
56	Y	DVD-A		Y	Y				Y	5	Y
57	Y	CD	Y	Y		Y		Y	Y	3	N
58	Y	CD	Y	Y	Y					3	Y
59	Y	DVD-A		Y	Y					4	N
60	Y	CD	Y			Y				2	N
61	Y	CD	Y					Y		2	N
62	Y	CD		Y						2	N
63	Y	DVD-A					Y		Y	2	N
64	Y	CD			Y			Y		2	N
65	Y	CD	Y		Y	Y		Y		3	N

Blind A/B Test #2: Compact Disc & Super Audio CD

			Positive qualities			Negative qualities				Audible	Good
	Audible	Format	of your format choice?			of the "lesser" format?				Difference	enough to
	Difference	Choice	Clearer	Warmer	Lifelike	Thin	Tin-like	less clear	brittle	(scaled 1-5)	switch?
1	Y	CD	Y	Y			Y	Y		2	N
2	Y	SACD								2	N
3	Y	SACD	Y					Y		2	N
4	Y	SACD	Y		Y			Y		2	N
5	Y	SACD	Y					Y		2	N
6	Y	CD	Y					Y		2	N
7	Y	SACD	Y					Y		2	Y
8	N	-	-	-	-	-	-	-	-	1	N
9	Y	SACD	Y	Y				Y	Y	3	N
10	Y	CD			Y			Y		2	N
11	Y	SACD	Y			Y		Y		3	Y
12	Y	CD			Y		Y			2	N
13	Y	CD	Y	Y				Y		2	N
14	Y	CD	Y	Y			Y		Y	3	N
15	Y	SACD			Y			Y		2	Y
16	Y	SACD	Y					Y		3	N
17	N	-	-	-	-	-	-	-	-	1	N
18	Y	SACD		Y	Y	Y				2	N
19	Y	SACD	Y		Y			Y		2	Y
20	Y	CD		Y		Y				2	N
21	Y	CD	Y	Y				Y		4	N
22	N	-	-	-	-	-	-	-	-	1	N
23	N	-	-	-	-	-	-	-	-	1	N
24	Y	SACD		Y			Y			2	N
25	Y	CD			Y			Y		2	N
26	N	-	-	-	-	-	-	-	-	1	N
27	Y	CD	Y	Y		Y				2	N
28	Y	CD	Y	Y				Y		2	N
29	Y	SACD		Y	Y		Y		Y	4	N
30	Y	SACD								2	N
31	Y	CD	Y	Y	Y	Y				5	N
32	Y	CD			Y			Y		3	Y
33	N	SACD		Y			Y			2	N
34	Y	CD	Y	Y			Y	Y		3	Y
35	Y	CD			Y				Y	2	N
36	Y	CD	Y		Y			Y		3	N
37	Y	CD	Y	Y						3	Y
38	N	-	-	-	-	-	-	-	-	1	N
39	Y	CD	Y					Y		3	N
40	Y	SACD		Y	Y					3	Y
41	Y	SACD		Y						2	N
42	N	-	-	-	-	-	-	-	-	1	N
43	Y	CD	Y	Y	Y	Y			Y	4	Y
44	Y	CD		Y				Y		2	Y
45	Y	SACD	Y	Y	Y			Y		2	N
46	Y	CD		Y	Y					2	Y
47	Y	SACD							Y	3	Y
48	Y	CD			Y	Y				2	N
49	Y	SACD			Y					2	N
50	Y	CD		Y	Y					3	N
51	Y	CD	Y	Y				Y		2	N
52	Y	CD		Y		Y				2	N
53	Y	CD			Y	Y				2	N
54	Y	CD		Y	Y	Y				1.5	N
55	Y	CD	Y	Y	Y			Y		2	N
56	Y	SACD		Y	Y					4	Y
57	Y	CD	Y		Y	Y		Y		2	N
58	Y	CD		Y	Y					2	Y
59	Y	CD	Y		Y			Y		2	N
60	Y	CD	Y					Y		4	Y
61	Y	SACD		Y		Y				2	N
62	N	-	-	-	-	-	-	-	-	1	N
63	N	-	-	-	-	-	-	-	-	1	N
64	Y	CD	Y				Y			2	N
65	Y	SACD		Y				Y		2	N

Blind A/B Test #3: DVD-Audio & Super Audio CD

			Positive qualities of your format choice?			*Negative qualities of the "lesser" format?*				*Audible Difference (scaled 1-5)*	*Good enough to switch?*
	Audible Difference	*Format Choice*	*Clearer*	*Warmer*	*Lifelike*	*Thin*	*Tin-like*	*less clear*	*brittle*		
1	Y	SACD	Y	Y	Y	Y	Y			4	Y
2	Y	SACD	Y	Y						4	Y
3	Y	SACD	Y	Y	Y	Y	Y	Y		3	Y
4	Y	SACD	Y		Y			Y		3	N
5	Y	SACD	Y		Y			Y		3	Y
6	Y	SACD	Y					Y		2	N
7	Y	SACD	Y					Y		5	Y
8	Y	SACD			Y					3	N
9	Y	SACD	Y		Y	Y				3	N
10	Y	DVD-A								2	N
11	Y	SACD	Y		Y			Y		4	Y
12	Y	DVD-A	Y	Y	Y	Y				3.5	N
13	Y	SACD	Y			Y		Y		3	N
14	Y	DVD-A		Y	Y				Y	2	N
15	Y	SACD		Y		Y				3	Y
16	Y	SACD	Y					Y		5	N
17	Y	SACD	Y		Y	Y				5	N
18	Y	DVD-A	Y	Y	Y	Y				3	N
19	Y	SACD	Y		Y			Y		3	Y
20	Y	DVD-A		Y		Y				3	N
21	Y	SACD	Y	Y	Y			Y		3.5	N
22	Y	SACD	Y	Y	Y	Y		Y		4	Y
23	Y	SACD	Y		Y	Y				4	Y
24	Y	DVD-A		Y	Y					3	Y
25	N	-	-	-	-	-	-	-	-	1	N
26	N	-	-	-	-	-	-	-	-	1	N
27	N	-	-	-	-	-	-	-	-	1	N
28	Y	SACD	Y					Y		2	N
29	Y	DVD-A			Y				Y	3	N
30	N	-	-	-	-	-	-	-	-	1	N
31	Y	DVD-A		Y		Y				2	N
32	Y	DVD-A		Y					Y	2	N
33	Y	SACD	Y					Y		1.5	N
34	Y	SACD	Y					Y		2	N
35	Y	DVD-A			Y			Y		2	N
36	Y	SACD			Y			Y		2	N
37	Y	SACD		Y						2	N
38	Y	SACD		Y				Y		2	Y
39	Y	SACD	Y					Y		2	N
40	Y	DVD-A		Y		Y				2	Y
41	N	-	-	-	-	-	-	-	-	1	N
42	Y	SACD	Y			Y				3	Y
43	Y	SACD	Y			Y		Y		2	N
44	Y	SACD		Y			Y			2	Y
45	Y	SACD	Y	Y	Y			Y		2	N
46	Y	SACD	Y					Y		2	N
47	Y	SACD			Y	Y				1	Y
48	Y	DVD-A		Y			Y			3	N
49	N	-	-	-	-	-	-	-	-	1	N
50	Y	SACD		Y		Y				2	N
51	Y	SACD	Y			Y				2	N
52	Y	DVD-A	Y					Y		2	N
53	Y	SACD	Y	Y		Y				2	N
54	N	-	-	-	-	-	-	-	-	1	N
55	Y	SACD	Y					Y		2	N
56	Y	DVD-A		Y						2	Y
57	Y	SACD	Y		Y				Y	4	N
58	Y	SACD	Y							2	Y
59	Y	DVD-A	Y					Y		2	N
60	N	-	-	-	-	-	-	-	-	1	N
61	Y	DVD-A	Y					Y		2	N
62	N	-	-	-	-	-	-	-	-	1	N
63	Y	DVD-A	Y				Y		Y	2	Y
64	N	-	-	-	-	-	-	-	-	1	N
65	N	DVD-A		Y						2	N