Measurement and Evaluation of Analog-to-Digital Converters Used in the Long-Term Preservation of Audio Recordings[1]

Ken C. Pohlmann
University of Miami
Frost School of Music

Abstract

The analog-to-digital (A/D) converter lies at the heart of the encoding side of a digital audio system, and is perhaps the most critical component in the entire signal chain. The A/D converter must discretely sample the analog signal, quantify the amplitude of the sample, and represent the measurement as a binary word. Whereas conversions made with the A/D converter's counterpart—the digital-to-analog (D/A) converter—can subsequently be improved for higher-fidelity playback, errors introduced by the A/D converter will accompany the audio signal throughout digital processing and storage and, ultimately, back into its analog state. Thus, the choice of the A/D converter irrevocably affects the fidelity of the resulting signal. For critical applications such as the long-term preservation of historic audio signals, to the greatest extent possible, the A/D converter must exhibit audio *transparency*—that is, it should neither add to nor subtract from the sound. To assess the degree of transparency, the converter's electrical measurements and subjective aural performance, as well as the converter's operating parameters such as sampling frequency and word length, must be considered. Finally, the signal-level input to the converter, converter-component design, and external conditions such as grounding and shielding can greatly affect the fidelity of the resulting file.

Introduction

It is perhaps ironic that although meaningfully audible audio signals exist only in the analog domain, they are best stored in the digital domain. Moreover, the tasks of converting audio signals into the digital domain, and back to the analog domain, are among the most difficult in digital audio technology. Indeed, the only steps in the complete audio signal chain that are more problematic are the transducing of signals from acoustical to electrical, and back again from electrical to acoustical—in other words, what is done by the microphone and loudspeaker. The final irony is that Edison and other audio pioneers did not have to contend with A/D and D/A converters, or even microphones and loudspeakers. Their all-acoustic audio systems were "all natural."

Today, when an analog recording is transferred to the digital domain, the A/D converter is the key component in the signal path (Fielder 1992). The converter's principal operating parameters—sampling frequency and word length—determine the theoretical bandwidth and noise floor of the digital recording as well as other criteria. Traditional audio converter measurements such as frequency response, distortion, jitter, and linearity

---

[1] This paper was written for the roundtable discussion, "Issues in Digital Audio Preservation Planning and Management," held March 10-11, 2006, in Washington. D.C. The meeting was convened by the Library of Congress and Council on Library and Information Resources on behalf of the National Recording Preservation Board.

can be used to evaluate a converter's quality. Factors such as the quality of the prefacing low-pass anti-aliasing filter and practical considerations such as input signal level and grounding and shielding are also important. Choosing an A/D converter must be based on an evaluation of technical measurements and of subjective listening.

Many historic recordings were recorded with low fidelity; for example, audio bandwidth is very limited and the noise floor is high. For this reason, some experts argue that when converting analog recordings for digital storage, the fidelity of the signal chain can be of relatively low fidelity. Others contend that any digitization must use the best-possible signal chain to capture and preserve as much information as possible. Given that in any archival-conversion project, the cost of digitization equipment is trivial compared with the cost of labor, the latter approach seems more prudent. This paper takes the position that an archival conversion signal chain must provide very high fidelity.

The goal in selecting an A/D converter is *transparency*—that is, to select a device that neither adds to nor subtracts from the recorded sound. In other words, an ideal converter has no sound of its own. Converter transparency remains an elusive ideal. Only the best converters can approach transparency. Most converters are certainly not transparent. For some non-critical applications, transparency is not needed. But for critical applications— and the archiving of audio materials for long-term preservation is among the most critical—transparency is vital. This paper describes some of the factors that influence A/D converter fidelity and recommends methods that can be used to identify the best-possible A/D converters for critical audio applications.

Sampling Frequency
Generally, sampling frequencies of 44.1, 48, 96, and 192 kHz are used in high-fidelity recording. The usable audio bandwidth is one-half the sampling frequency, so higher sampling frequencies provide a wider audio bandwidth. This is potentially useful because musical instruments can generate content with wide bandwidths; for example, a cymbal might have response of 90 dB SPL (sound pressure level) beyond 60 kHz, and a violin might have content beyond 100 kHz.

Even so, the use of high sampling frequencies such as 96 and 192 kHz may seem unnecessary. In rare cases, a person may be able to hear frequencies to 24 or 26 kHz, far below the cutoff frequencies of 48 and 96 kHz. In most cases, high-frequency hearing response is below 20 kHz. Thus, for steady-state tones, the higher-frequency response may not be useful. However, a high sampling frequency provides additional benefits beyond wide audio bandwidth. It can be argued that high sampling frequencies improve the binaural time response, leading to improved imaging in multichannel recordings. For example, if short pulses are applied to each ear, a 15-μS difference between the pulses can be heard, and that time difference is shorter than the time between two samples at 48 kHz. Some people can hear a 5-μS difference, which corresponds to the time difference between two samples at 192 kHz. In theory, a high sampling frequency might improve spatial imaging.

Similarly, higher sampling frequencies provide improved temporal response. For example, the sampling interval at 44.1 kHz is 22.7 μS; at 192 kHz, it is 5.2 μS. Musical instruments can generate transients with rise times of less than 10 μS. As another example, room reverberation comprises a large number of reflections arriving at high rates. For example, reverberation might comprise 500,000 arrivals per second; spaced regularly, this time interval is less than 2 μS. Human subjects are sensitive to interaural time delays of between 2 and 10 μS. Subjects have differentiated between a regular pulse train and one with deviations of 0.2 μS. Higher sampling frequencies clearly preserve temporal response (Woszczyk 2003). In addition, higher sampling frequencies allow greater latitude in the design of the anti-aliasing low-pass filter. For example, a lower-order slope may be employed, providing improved time-domain response; this is further described below. Generally, high sampling frequencies can promote improved filter and signal processing performance in the traditional audio (0 to 20 kHz) band. Ultimately, because the limit of human hearing acuity is not yet known, the point of transparency of a recording system cannot be known. In some cases, such as the conversion of monaural speech recordings, a lower sampling frequency of 48 kHz may be used. However, for highest audio fidelity, higher sampling frequencies of 96 or 192 kHz are recommended.

Quantization Word Length
An A/D converter must represent the amplitude subtlety of the analog audio signal as a series of discrete binary words. The precision of amplitude required is considerable: 15 parts per million for 16-bit resolution, and 1 part per 16 million for 24-bit resolution. If sheets of typing paper representing a 16-bit system were stacked to a height of 22 feet, a single sheet of paper would represent one quantization level. In a 24-bit system, the stack would tower at 5,632 feet—more than a mile high. The quantizer could measure that mile to an accuracy equaling the thickness of a piece of paper. If a single page were removed, the least significant bit would change from 1 to 0. A high-quality digital audio system thus requires components with similar tolerances—not a trivial feat. Such tolerance is necessary because it corresponds to the acuity of the human hearing system (Pohlmann 2005).

The word length of the converter describes the length of the output digital word and hence the number of bits used to represent the amplitude of the audio samples. For example, a converter may convert 16 or 24 bits. The word length is often mistakenly referred to as the converter's resolution. This figure only theoretically determines the level of the quantization noise floor in the digital recording; measured in dB, the noise floor level is $6.02N + 1.76$ where N is the number of bits converted. However, this is a theoretical figure. The number of converted bits is a misleading figure of merit because of converter errors. A more effective benchmark is ENOB (effective number of bits) where ENOB = (dynamic range – 1.72)/6.02. For example, a 16-bit converter with a measured dynamic range of 90 dB provides only 14.7 bits of resolution. A method used to measure dynamic range is described below.

There is debate regarding the converter resolution required for transparency. Twenty-four bits is the highest word length generally available and offers the potential for the highest fidelity. If a 24-bit converter could provide a theoretically perfect noise floor level of -

146 dB, it could be argued that it is overkill. No converter can accomplish this. However, a well-designed 24-bit converter will provide a noise floor that lies at the limits of audibility. It is also important to quantize a word length that is relatively longer than what may be immediately required. Subsequent digital processing effectively decreases useful worth length because truncation and rounding add error to the least significant bit (LSB). Cascading mathematical operations further decrease useful word length. Finally, the larger dynamic range provided by long word length provides greater headroom, which makes level setting less critical. Thus, 24-bit word length is recommended.

Dither

Another important consideration in A/D converter setup is the selection of a dither signal. Dither is a low-amplitude noise deliberately added to the audio signal; dither improves converter linearity, reduces distortion, and greatly extends the converter's dynamic range. An undithered audio signal may be subject to significant converter errors and significantly poor sound quality. Dither signals are characterized by their probability density function (pdf), and the choice of the pdf type and the amplitude of the dither signal are both important. Most experts agree that that triangular pdf dither with amplitude of 1 or 2 LSB peak-to-peak is optimal for music signals. Rectangular pdf dither is often included in test signals when performing low-amplitude audio measurements.

Converter Chip Architecture

Most modern A/D converters employ an integrated circuit as their basis; in most cases, this chip uses a sigma-delta (also known as delta-sigma) architecture. This is sometimes referred to as a 1-bit converter; however, in practice, several bits may convey the audio signal. All sigma-delta converters differ from traditional multi-bit (ladder or R-2R) converter architectures that use many bits (16 or 18) to convey the audio signal, and sigma-delta converters generally provide better audio quality than traditional systems do. Although the sigma-delta architecture is widely used, other A/D converter architectures are available (Putzeys 2003). Sigma-delta converters operate at a high internal oversampling rate; for example, an output sampling frequency of 48 kHz may use an internal frequency of 64 times, or 3.072 MHz. This obviates the necessity for a brickwall anti-aliasing low-pass filter. Instead, in many cases, a simple, low-pass filter comprising a series resistor and capacitor to ground provide sufficient attenuation of any input aliasing frequencies. The "gentle" slope of this low-pass filter promotes good time-domain response but does not provide a flat response at ultrasonic frequencies. Only a brickwall filter will provide the latter. Discussion continues on the trade-offs in anti-aliasing filter type and potentially audible effects (Craven 2003). Sigma-delta converters use noise shaping to reduce noise in the audio band of interest, at the expense of higher noise levels out of band; this technique performs well in high-fidelity applications.

Converter Component Design

The A/D converter chip is the heart of any A/D converter component, but the design of circuitry supporting the converter chip is critical in ensuring that the chip performs to its design limits. In other words, a good chip can be compromised by the external design. Numerous aspects must be considered. The converter's power supply may use a

switched-mode or a linear design. The former may cause interference and must be filtered and shielded. The following are required: separate analog and digital ground planes, independent analog power supplies for the input buffer and analog section of the converter chip, and a digital supply for the digital section. All supplies should be properly decoupled. Similarly, the converter's voltage reference must be stable and free of interference. For example, the harmonics from an interfering square-wave clock could corrupt this reference. Careful converter designers avoid using clocks that are close to the input sampling frequency or its multiples.

Proper circuit board layout, good power supply design, correct grounding, accurate clocks, and other design criteria are required so that the A/D converter chip can operate optimally. The converter's input buffer must be designed for low noise. The A/D converter chip should not be mounted in a socket because this adds inductance. The sensitivity of the converter to external conditions such as clock jitter may be low or high depending on the converter design.

In summary, while the chip specifications are important when evaluating an A/D converter, it is more important to evaluate the specifications of the component that houses and supports it.

Input Audio Preamplifier and Signal Levels
In some conversion systems, the audio preamplifier is contained within the A/D converter. In other systems, a separate preamplifier is needed after the playback device and before the A/D converter. The preamplifier is used to adjust signal gain, optimizing it for input to the converter. Because all analog signals are fragile, the preamplifier must be high quality, with low noise and distortion—with throughput specifications commensurate with those of the A/D converter. Adjusting signal level is an important consideration. An audio level that is too low will under-utilize the converter's resolution. For example, a too-low level applied to a 24-bit converter may exercise only 20 of the converter's bits, effectively reducing its resolution by 4 bits. Conversely, a signal that is too high will overload (clip) the converter, resulting in severe distortion. Signal levels should be adjusted in the analog domain; preferably, levels should not be normalized after conversion. Meters on the A/D converter must be carefully observed and the playback device or preamplifier appropriately adjusted for optimal signal level. In some cases, levels can be adjusted on the converter.

The decision of where to set audio signal levels must be carefully considered. For example, a recording might exhibit a low signal level, and high-level transient pops and clicks. If the level is set so the pops and clicks are within the converter's dynamic range, the dynamic range available for the audio signal is decreased. If the level is set so the clicks and pops are beyond the available dynamic range, the defects will be clipped. It could be argued that the defects contain useful information and should be accurately recorded. For optimal level setting, it may be necessary to play the recording twice— once to determine maximum signal level and, after level setting, a second time to make the conversion. Perhaps more efficiently, a recording could be played once through two independent conversion channels, with perhaps a 10-dB level difference; the highest-

level resulting digital file that is free of clipping is selected as the final version. (Some converters can independently process two or more audio channels simultaneously). Archivists will recall that in many cases, a disc can be visually observed, and its loudest sections determined (by large groove excursions); these can be played to quickly set approximate levels. Another question is how often to optimize levels. For example, should levels be optimized for every track, each side, or just once for the album. Clearly, any changes in level should be documented.

Electrical Measurements
Harmonic distortion and dynamic range are two fundamental A/D converter measurements. Harmonic distortion is a familiar and useful way to characterize audio linearity. A single, pure sine tone is input to the device under test, and the output is examined for spurious content other than the sine tone. In particular, spectral analysis will show any harmonic multiples of the input frequency. Total harmonic distortion (THD) is the ratio of the summed RMS (root mean square) voltage of the harmonics to that of the input signal. To account for noise in the output, this measurement is often called THD+N. The figure is usually expressed as a decibel figure or a percentage; however, visual examination of the displayed spectral output is also a valuable diagnostic strategy. In most analog systems, THD+N decreases as signal level decreases. The opposite is true in digital systems. Therefore, THD+N should be specified at both high and low signal levels (Frindle 1997). THD+N should be evaluated versus amplitude and versus frequency, using FFT (Fast Fourier Transform) analysis.

Dynamic range is the amplitude range between a maximum-level signal and the noise floor. Dynamic range (using the EIAJ (Electronic Industries Association of Japan) specification) is typically measured by reading THD+N at an input amplitude of -60 dB; the negative value is inverted and added to 60 dB to obtain dynamic range. Signal-to-noise ratio (examining idle channel noise) can be measured by subtracting the idle noise from the full-scale signal. For consistency, a standard test sequence such as the ITU CCITT 0.33.00 (mono) or CCITT 0.33.01 (stereo) can be used; these comprise a series of tones and are useful for measuring parameters such as frequency response, distortion, and signal to noise. A series of test signals acting as a "Rosetta Tone" should be generated in the analog domain and converted and recorded along with the audio content. These tones would allow future analysis to better understand the performance of the converter, and evaluation of the condition of the recording medium. Also, such tones may permit future archivists to reverse engineer the performance of the converter, and thus potentially apply signal processing to undo converter errors.

Another useful converter measurement is amplitude linearity, which compares output and input linearity. Ideally, the output value should correspond exactly with input level, regardless of level. To perform the test, a series of tones of decreasing amplitude, or a fade-to-zero tone, is input to the converter. The tone is dithered with rectangular pdf dither. A plot of device gain versus input level will reveal any deviations from a theoretically flat (linear) response.

An A/D converter is susceptible to jitter, defined as a variation in the time base of the clocking signal. Random-noise jitter can raise the noise floor and periodic jitter can create sidebands, thus raising distortion levels. Generally, the higher the specified dynamic range of the converter, the lower the jitter level required (Harris 1992). A simple way to test an A/D converter for jitter limitations is to input a 20-kHz, 0-dBFS (full-amplitude) sine tone, to observe an FFT of the output signal, and then to repeat the process with a 100-Hz sine tone. An elevated noise floor at 20 kHz compared with 100 Hz indicates a potential problem from random-noise jitter; discrete frequencies at 20 kHz indicate periodic jitter. Jitter is inherent in any digital device and it can never be eliminated. High-quality A/D converters contain internal clocks that are extremely stable, or when accepting external clocks, have clock-recovery circuitry to reject jitter disturbance. It is incorrect to assume that one converter using a low-jitter clock will necessarily perform better than another converter using a high-jitter clock; actual performance depends very much on converter design. Even when jitter causes no data error, it can cause sonic degradation. Its effect must be carefully assessed in measurements and listening tests.

Noise modulation is another useful measurement (Cabot 1997). This test measures changes in the noise floor relative to changes in signal amplitude. Ideally, there should be no correlation; however, in practice, because of low-level nonlinearity in the converter, there may be audible shifts in the level or tonality of the background noise that correspond to changes in the music signal. Because such shifts are correlated to the music, they are potentially much more perceptible than benign, unchanging noise. In one method to observe noise modulation (Cabot 1991), a low-frequency sine tone is input to the converter, the sine tone is removed at the output, and the spectrum of the output signal is examined in 1/3-octave bands. The level of the input signal is decreased in 5 dB steps and the test is repeated. Deviation in the noise floor by more than a decibel in any band across the series of tested amplitudes may indicate potentially audible noise modulation.

When measuring an A/D converter or any other digital device, test tones should not use frequencies that are submultiples of the sampling frequency. Submultiples exercise only a few code values; for example, a 1 kHz test tone and 48 kHz sampling frequency will employ only 48 code values. This does not fairly represent the converter's performance. Rather, tone frequencies should be relatively prime to the sampling frequency (Finger 1986). For example, 997 kHz should be used instead of 1 kHz. The International Standards Organization (ISO) has standardized a set of recommended test tone frequencies.

The bit stream output from an A/D converter may be measured. A digital interface (such as AES3, also known as AES/EBU) should be checked for data waveform characteristics, to ensure that data can be transmitted reliably. Checks must be done to ensure that the waveform's amplitude falls within the allowed minimum and maximum levels. The jitter in the output waveform should be checked; generally, this interface jitter is not a concern. A larger system should be designed to ensure proper clocking among connected components. In some cases, it is advisable to synchronize components using an external reference clocking signal. A larger system should avoid long cable runs and should

observe good engineering practices for grounding and shielding. The uncompressed PCM output from the converter should follow a standard interface format such as AES3; the output serial data format should be documented. Pre-emphasis is not recommended; if it is used, it should be documented. In some cases, it may be desirable to save the output in both PCM and bitstream formats.

Practically speaking, accurate measurements of A/D converter performance require special equipment and skill. For example, whenever possible, optical cables should be used in lieu of coaxial cables to avoid interference. During testing, other equipment should be turned off; for example, radio-frequency signals from a PC and its monitor can affect the measurement.

For high-quality performance, an external professional A/D converter is required. The converter on a sound card in to a PC, for example, cannot perform adequately for critical applications. A manufacturer's published specifications are shown below. These represent a high-quality A/D converter (circa 2006):

- Frequency response (1 Hz to 48 kHz): -1 dB
- Total harmonic distortion and noise (1 kHz at 0 dBFS): < -108 dB (0.0005%) (unweighted RMS)
- Dynamic range: > 130 dB (unweighted RMS)
- Intermodulation distortion: <-90 dB
- Spurious aharmonics: <-130 dBFS
- Crosstalk (50 Hz, 0 dBFS in opposite channel): < -130 dB
- Crosstalk (15 kHz, 0 dBFS in opposite channel): <-140 dB
- Linearity (at -144 dBFS): < 3 dB
- Intrinsic jitter: < 18 pSec RMS
- Phase linearity: < 1°
- Internal clock accuracy: ±5 ppm

Limitations of Electrical Measurements
Although electrical measurements such as frequency response, noise level, and distortion can correlate to perceived sound quality, they are far from an exact match to human perception. Whereas these traditional measurements are adequate for evaluating lower-fidelity audio equipment, they are less applicable to evaluating the performance of more modern, higher-quality audio equipment. Moreover, such tests are not applicable to audio equipment that uses psychoacoustic perceptual models to code the audio signal.

A THD distortion figure of 0.01% might be audible under some conditions, while a distortion figure of 10% might be inaudible under others. An audio signal can undergo considerable signal processing (such as AAC [Advanced Audio Coding]) with relatively little or no audible change, even though the measured distortion is high. New measures of distortion are clearly needed, and work continues on developing new testing methods and on identifying ways to correlate the results of analytical tests to human perception (Geddes and Lee 2003). For example, new metrics may take into account psychoacoustic effects such as masking on the audibility of distortion. Meanwhile, while still very useful,

traditional measures such as THD are increasingly seen as unreliable measures of perceptual distortion. For example, it has been shown (Fielder 1989) that an A/D/A system with a dynamic range of 92 dB, full-level distortion of 0.008%, and distortion of 0.3% at -40 dB generated audible noise and distortion components not immediately discernible using traditional metrics. Observation revealed idle channel noise, modulation noise, and narrow-band distortion above the threshold of hearing. Any factor that affects audio transparency can be measured. However, we sometimes do not know a priori what should be measured. In many cases, listening tests identify a defect, and a measurement is then devised to quantify it.

Subjective Evaluation
As John William Rayleigh said in 1877 (the year in which Thomas Edison first shouted into his tinfoil recorder), ''All questions connected with this subject must come for decision to the ear, as the organ of hearing; and from it there can be no appeal.''

Subjective evaluation uses humans to assess auditory performance. Subjective testing is relevant for assessing both the highest-quality audio signals and the lower-quality perceptually coded signals. This suggests that subjective evaluation is more perceptually relevant than electrical measurement. Subjective evaluation must be part of a total component evaluation.

Any audio recording comprises a useful signal, such as music and speech, as well as unwanted noise and distortion. It is difficult for any measurement to assess the level of the unwanted signals; for example, there is no strictly analytical method to differentiate between music and noise (Burkhard 1992). Thus, while analytic measurements are extremely helpful, particularly for quickly identifying poor fidelity, human evaluation is needed to distinguish among the highest levels of fidelity.

Audio lore, and some literature, is replete with unexplained discrepancies between theory and analytical measurement and subjective opinion. For example, some have claimed that different pressings of numerically identical Compact Discs can sound different. Although this claim is sometimes accepted as fact, one study (Dennis et al. 1997) preliminarily showed it is not supported by scientific evidence. On one hand, many subjective comments are simply apocryphal and incorrect, or perhaps based more on marketing hopes than on audible reality. On the other hand, most audio practitioners, even those most analytically inclined, will admit that perceptual acuity is vast and that its scope is not yet fully charted or understood.

Listening Tests
With careful control, human hearing remains the optimal method for sound analysis. Without careful control, human evaluations are meaningless and potentially misleading. Subjective testing cannot substitute for analytical measurement, but it does play a very important accompanying role.

One way to subjectively evaluate the performance of two (or more) A/D converters is to simultaneously record a live music performance through the converters to separate files

and to then perform an ABX test. In such a test, the listener is presented with the known A and B sources, and an unknown X source that can be either A or B; the assignment is made pseudorandomly for each trial. The listener must identify whether X has been assigned A or B. The response reveals whether the listener can hear a difference between A and B. Statistical analysis is then used to select a qualified listening panel of individuals who can reliably hear a difference between the files.

This qualified panel is asked which file (and hence which converter) it prefers; results are statistically analyzed. To evaluate subjective preferences, a five-point impairment scale devised by the CCIR (International Radio Consultative Committee) can be used. Panels of listeners rate the impairments they hear on a 41-point continuous scale in categories from 5.0 (transparent) to 1.0 (very annoying impairments). In some cases, subjective tests are conducted using guidelines in ITU-R Recommendation BS.1116-1. These guidelines address selection of audio materials, performance of the playback system, listening environment, assessment of listener expertise, grading scale and methods of data analysis. Other issues in sound evaluation are described in ITU-R BS.1534, ITU-T P.800, P.810, and P.830; ITU-R BS.562-3, BS.644-1, BS.1284, BS.1285 and BS.1286, among other standards.

Ideally, a listening panel should use expert listeners, in a controlled, double-blind environment. However, for practical reasons, it may be necessary to distribute listening test materials on CD or DVD (Isherwood 2003). This may provide a larger number of trials and access to a variety of reproduction systems. The test disc contains a reference track (with good fidelity but not necessarily possessing the best sound quality). Other tracks are compared with the reference using a numerical score to evaluate preference relative to the reference track. Various types of music are used (choice of listening material is an important consideration) and tracks are scrambled uniquely on each disc to discourage collaboration among subjects.

Because long-term memory of human hearing is extremely poor, any evaluation must provide fast comparisons. When comparing different signals, signal levels must be matched as closely as possible. Differences of as little as 0.1 dB may influence an evaluation. The effect is especially insidious because the difference itself may not be directly audible, but could lead to misjudgments about signal quality (Frindle 1997).

Analysis of Listening Test Results
Given the relatively high quality of top A/D converters, it is unlikely that listeners will be unanimous in their preference of converter. . Thus, the mixed results must be statistically analyzed. To be meaningful, and not misleading, listening test results must be carefully interpreted. For example, in an ABX test, if a listener correctly identifies the reference in 12 out of 16 trials, has an audible difference been noted? Statistical analysis provides the answer, or at least an interpretation of it. In this case, because the test is a sampling, we define our results in terms of probability. The larger the sampling, the more reliable the result. A central concern is the significance of the results. If the results are significant, they are due to audible differences; otherwise, they are due to chance. In an ABX test, producing a correct score 8 of 16 times indicates that the listener has heard no

differences; such a score could be arrived at by guessing. A score of 12 of 16 might indicate an audible difference, but could also be due to chance. Care must be taken to generate a valid analysis that has appropriate statistical significance. The number of listeners, the number of trials, the confidence interval, and other variables can dramatically affect the validity of the conclusions (Burstein 1988). If test results are to be valid and helpful, care must be taken in the design of subjective listening tests and in the analysis of their results.

Comments on Analog Media Playback
When transferring audio from a historical medium for long-term preservation, the two most important factors are the condition of the original medium and the mechanical reproducer used to play back the medium. For the latter, it is assumed that original reproducers will not be used. Their deficiencies will degrade the quality of the audio signal. Furthermore, acoustic reproducers (wax-cylinder players, for example) require the use of a microphone to capture the acoustic output and convert it into an electrical signal. Although such a system offers "authentic" playback, its limitations are evident. If acoustic playback is attempted, selection of microphone, its placement, recording environment, use of a horn or a direct coupler, and choice of microphone preamplifier are all critical. Modern mechanical reproducers with modern electrical pickups (phonograph cartridges) or experimental optical pickups are required. These reproducers play the original media and output an electrical audio signal. This signal can be input into the A/D converter; however, care must be taken to ensure that the amplitude of the audio signal matches the converter's input requirements. For example, the highest signal amplitude must not clip the converter's input.

Most analog disc and tape media are recorded with frequency equalization to compensate for media deficiencies. For example, LP records are recorded with RIAA (Recording Industry Association of America) equalization, and many tapes use IEC (International Electrotechnical Commission) or NAB (National Association of Broadcasters) equalization. Consideration should be given as to whether the corresponding playback equalization should be performed in the analog domain prior to A/D conversion, or in the digital domain after conversion (Davies 2002). Both can restore correct spectral balance. Prior A/D processing is more expedient because the recording is immediately available and no further processing is required. However, use of an incorrect standard (for example, NAB instead of IEC) may be difficult to correct, particularly if it is followed by other processing. Post A/D processing may offer more accurate results and the "flat" recording can be marked with metadata. However, because an equalization curve may effect a gain change of perhaps 20 dB, postprocessing essentially can decrease resolution by 2 or 3 bits in the conversion; also, special nonequalized playback hardware is required. Furthermore, similar consideration should be given to whether nonstationary hardware signal processing such as Dolby. A companding should be performed before or after A/D conversion.

When playing original historical media, it is important to consider playback speed. A variety of speeds were often employed, and in many cases the speed was inaccurate both in terms of absolute speed and constancy ("wow and flutter"). Ideally, the correct

playback speed must be used when playing the original media. When necessary, speed inaccuracies can be corrected in the digital file.

Selection of the correct playback stylus is another important factor in cylinder and disc playback. Many stylus types were used for playback (different styli are used for cutting and playback) and the correct stylus must be used. Ideally, this would match the type originally intended to play back the media, but in some cases, wear or damage to the grooves may necessitate using a different stylus that avoids worn and damaged areas and instead contacts more useful information stored in the groove walls. It may be useful to transfer an original medium using several types of styli; potentially, each could capture somewhat different information on the groove walls. In any case, the choice and adjustment of the stylus (or tape head) are critical.

Many historical recordings are single-channel monaural and can be transferred to a single-channel monaural digital file. However, it may be useful to play back monaural disc media using an appropriate stereo stylus and to store the output signal as a two-channel monaural file. A stereo stylus may capture a greater amount of information from both groove walls. For example, theoretically, a useful signal from one groove wall could be used to replace and patch a momentarily corrupted signal on the other groove wall. However, undesirable phase differences between groove walls may be present. The file can be formatted as a single two-channel file, or as two separate one-channel files (left and right); the latter offers redundancy in the event that one file is corrupted or lost. Most A/D converters are stereo converters; therefore, output of two-channel monaural files does not entail added setup complexity.

Summary
- Even low-fidelity analog recordings should be converted using a high-quality A/D converter. Given the time and effort needed to accomplish a good transcription, the cost of a high-quality converter is trivial. There is no merit in using a converter or any audio equipment that is inferior to the limits of human hearing. To do so is to risk losing information present on the analog source.
- Generally, only a professional-quality, external A/D converter will suffice. Converters on PC soundcards are inferior and subject to noise and interference.
- Before purchasing a converter, consider all the published electrical specifications of the converter chip and the converter component.
- Before purchasing, if possible, perform independent electrical testing and measurement of the converter component.
- Test tones such as a "Rosetta Tone" should be recorded along with the audio material to permit future analysis of the converter accuracy, evaluate the condition of the recording medium, and potentially remove converter errors.
- Before purchasing, if possible, compare converters in a listening test. A listening test must be carefully designed for validity and reliability. Numerical results should be analyzed with recognized statistical methods.
- When possible, seek additional expert advice on choosing a converter.
- When necessary, a high-quality audio preamplifier should be used to adjust analog signal levels prior to conversion.

- Carefully monitor input signal level with peak-reading meters. If the input signal overloads (clips) the converter, severe distortion results, essentially ruining the fidelity of the digital recording. Input overload cannot be satisfactorily fixed after conversion.
- For the highest audio quality, a 96 or 192 kHz sampling frequency should be selected. This provides a wide audio bandwidth, good temporal response, and allows improved low-pass filter characteristics.
- For the highest audio quality, a 24-bit word length should be selected. This provides a large dynamic range, permits more headroom in level setting, and helps insulate against the effects of rounding in subsequent digital signal processing.
- If available, triangular pdf dither should be selected. Undithered conversion should be avoided.
- For critical conversions, recognizing that a perfect converter is an impossibility, an archivist may perform two conversions, using two different converters.
- In most cases, the output signal will be in PCM format. In some cases, it may be desirable to also save the file in a bitstream format.
- Carefully aurally monitor input levels, and output levels using high-quality D/A converters. Loudspeakers can be used for monitoring, but only if the speakers are of high quality, and ambient room noise is low. Often, high-quality headphones provide a better monitoring alternative.

Conclusion

The analog-to-digital converter plays a critical role in the transfer of analog recordings to digital media. For critical applications such as archival conversion, an operation that provides the greatest-possible degree of audio transparency is absolutely necessary. The architecture of the converter chip, the design of the supporting converter component, and the interfacing of the converter component to other equipment must all be considered. Analytical measurements should be used to evaluate the audio performance of an A/D converter. However, no known electrical test can fully evaluate the perceptual transparency of an A/D converter. Therefore, a carefully designed listening test is an equally important part of the evaluation. Because the full acuity of the human hearing mechanism is not completely understood, we must be generous in our choice of sampling frequency and word length when selecting an A/D converter. Commercially available A/D converters can perform very well and will continue to improve; however, for optimal archival results, it may be necessary to commission a new reference converter. If well chosen and used with skill in a benign environment, modern A/D converters are able to capture virtually all useful information present in an analog audio recording.

References
Burkhard, Mahon. 1992. Merging Subjective and Objective Acoustical Measurements. Paper presented at the 11th International Conference of the Audio Engineering Society, Portland, Oreg., May 29–31, 1992.

Burstein, Herman. 1988. Approximation Formulas for Error Risk and Sample Size in ABX Testing. *Journal of the Audio Engineering Society* 36(11): 879.

Cabot, Richard C. 1997. Fundamentals of Modern Audio Measurement. Paper presented at the Audio Engineering Society UK Conference, The Measure of Audio, London, April 28–29, 1997.

Cabot, Richard C. 1991. Noise Modulation in Digital Audio Equipment. Preprint 3021. 90th Convention of the Audio Engineering Society, Paris, February 19–22, 1991.

Craven, Peter G. 2003. Controlled Pre-Response Antialias Filters for Use at 96 kHz and 192 kHz. Preprint 5822. 114th Audio Engineering Society Convention, Amsterdam, March 22–25, 2003.

Davies, S. W. 2002. (2002). Equalisation for Archival Transfer: In the Analogue or Digital Domain? Preprint 5534. 112th Audio Engineering Society Convention, Munich, May 10–13, 2002.

Dennis, Ian, Julian Dunn, and Doug Carson. 1997. The Numerically Identical CD Mystery–A Study in Perception Versus Measurement. Paper presented at the Audio Engineering Society UK Conference, The Measure of Audio, London, April 28–29, 1997.

Fielder, Louis D. 1989. Human Auditory Capabilities and Their Consequences on Digital-Audio Converter Design. Paper presented at the Audio Engineering Society 7th International Conference, Toronto, Ontario, May 14–17, 1989.

Fielder, Louis D. 1992. Introduction: Subjectively Relevant Objective Assessment, Part 1, Part 2. Paper presented at the Audio Engineering Society 11th International Conference, Portland, Oreg., May 29–31, 1992.

Finger, Robert. 1986. A Review of Frequencies and Levels for Digital Audio Performance Measurements. *Journal of the Audio Engineering Society* 34(1/2): 36–48.

Fleming, James. 2001. Forward into the Past–Protecting Our Musical Heritage. *Journal of the Audio Engineering Society* 49(7/8): 675–79.

Frindle, Paul. 1997. Are We Measuring the Right Things? Artifact Audibility Versus Measurement. Paper presented at the Audio Engineering Society UK Conference, The Measure of Audio, London, April 28–29, 1997.

Geddes, Earl R., and Lidia W. Lee. 2003. Auditory Perception of Nonlinear Distortion, Theory. Preprint 5890. 115th Audio Engineering Society Convention, New York, October 10–13, 2003.

Grotke, Robert W. 2004. Digitizing the World's Largest Collection of Natural Sounds: Key Factors to Consider When Transferring Analog-Based Audio Materials to Digital Formats. *RLG DigiNews* 8(1). Available at www.rlg.org.

Harris, Steven. 1992. Measurement Techniques for Debugging ADC and DAC Systems. Paper presented at Audio Engineering Society 11th International Conference, Portland, Oreg., May 29–31, 1992.

Harris, Steven, et al. 1997. Techniques to Measure and Maximize the Performance of a 120-dB, 24-Bit, 96-kHz A/D Converter Integrated Circuit. Preprint 4530. 103rd Audio Engineering Society Convention, New York, September 26–29, 1997.

Isherwood, David, et al. 2003. Augmentation, Application, and Verification of the Generalized Listener Selection Procedure. Preprint 5984. 115th Audio Engineering Society Convention, New York, October 10–13, 2003.

Lee, Lidia W., and Earl R.Geddes. 2003. Auditory Perception of Nonlinear Distortion. Preprint 5891. 115th Audio Engineering Society Convention, New York, October 10–13, 2003.

Metzler, Bob. 1993. *Audio Measurement Handbook.* Beaverton, Oreg.: Audio Precision.

Nishimura, Akira, and Nobuo Koizumi. 2002. Measurement of Sampling Jitter in Analog-to-Digital and Digital-to-Analog Converters Using Analytic Signals. Preprint 5586. 112th Audio Engineering Society Convention, Munich, May 10–13, 2002.

Pohlmann, Ken C. 2005. *Principles of Digital Audio*, 5th ed. New York: McGraw-Hill.

Putzeys, Bruno. 2003. Design Techniques for High-Performance Discrete A/D Converters. Preprint 5823. 114th Audio Engineering Society Convention, Amsterdam, March 22–25, 2003.

Rayleigh, John William. 1877. *The Theory of Sound, vol. 1*. Reprint, New York: Dover, 1945.

Reefman, Derk, et al. 2003. A New Digital-to-Analogue Converter Design Technique for HiFi Applications. Preprint 5846. 114th Audio Engineering Society Convention, Amsterdam, March 22–25, 2003.

Woszczyk, Wieslaw. 2003. Physical and Perceptual Considerations for High-Resolution Audio. Preprint 5931. 115th Audio Engineering Society Convention, New York, October 10–13, 2003.