ACOUSTIC MEASUREMENT TECHNOLOGY IN THE AUDIO PRECISION SYSTEM ONE

R. C. Cabot, L. R. Eagle

Audio Precision, Beaverton, Oregon, USA

INTRODUCTION

MLS.DSP is a program for the Audio Precision™ System One + DSP™ or the System One Dual Domain™. It uses maximum length sequence (MLS) testing to characterize the linear response of acoustical devices. This permits time-selective measurements in which one signal, such as the direct sound from a loudspeaker, may be separated from another similar signal, such as a room reflection. The time window may be adjusted to allow measurement of any arrival in a complex reverberation pattern. These arrivals may be examined in the time domain (showing pressure or energy as a function of time) or in the frequency domain (amplitude and phase vs frequency). Impulse responses may be saved to disk for later download to the DSP and further analysis.

The program generates a special digital noise signal called a maximum length sequence (MLS). The signal is available with selectable source impedance, balanced/unbalanced selections, and wide-range controllable amplitude. The System One analog inputs perform balanced-to-unbalanced conversion, automatic gain ranging, and drive the DSP's A/D converters. The DSP module and MLS software then perform a cross-correlation between the received and transmitted signals to obtain the impulse response which is stored into DSP memory and may be displayed on the computer. A portion of the impulse response may be transformed into the frequency domain to study both magnitude and phase response versus frequency. As with any other System One test data, limits may be applied for go/no-go testing and tests may be incorporated into procedures to totally automate testing. The DSP hardware and MLS software are both capable of two-channel operation. The input signal selection capability of the analog interface permits assigning a signal to one DSP channel and another related or independent signal to the other channel. Both may be acquired simultaneously or one previously acquired signal may be stored in memory while the other is acquired.

These properties are of obvious use when measuring loudspeakers or other electroacoustic devices. The time-selective capability permits separating the device-under-test response from that of the room in which the measurements are made. Alternately the room itself may be measured, studying the reflection characteristics of each surface in the room or of the room taken as a whole. The analysis technique depends on time invariance and is therefore not suited to measuring time varying systems such as the Lexicon Assisted Reverberation System (LARS) which uses low frequency modulation of the reverberation delays to decorrelate the outputs of the enhancement speakers and increase the gain before feedback.

ACOUSTIC MEASUREMENT TECHNOLOGY IN THE AUDIO PRECISION SYSTEM 1

QUASI-ANECHOIC MEASUREMENTS AND LOW FREQUENCY LIMITATIONS

Other techniques exist which provide measurement capability similar to maximum length sequence analysis. These include Time Delay Spectrometry (TDS) and impulse testing. They all provide quasi-anechoic frequency response measurements of the loudspeaker alone, unaffected by room reflections. All share the limitation that this anechoic response is useful only above a critical frequency determined by the physical dimensions of the test environment.

The principal advantage of MLS over impulse testing is in signal-to-noise ratio improvement. The 32k point sequence used in the Audio Precision implementation is equivalent to averaging 32,768 individual impulses of the same amplitude, and thus produces a signal-to-noise advantage of about 45 dB. To obtain low-frequency response data, TDS must sweep very slowly but MLS has the same testing time whether evaluating only the anechoic portion (first arrival) or a longer portion of the signal for accurate bass response. With MLS, one acquisition and correlation produces an impulse response which may then be evaluated over and over to look at anechoic response, response of any selected reflection, integrated room response, etc. The TDS technique requires that the generator-to-bandpass filter frequency offset (delay time) be reset to equal the acoustical propagation delay for each acoustical path to be measured. Sweep speed is critical for the TDS technique, and an operator may easily select a speed which produces erroneous data without knowing he is making an error.

Ambient acoustic noise tends to be greatest at low frequencies due to heating and air conditioning systems, traffic noise, etc. To improve signal-to-noise ratios under these typical conditions, the Audio Precision software provides pink spectral shaping (high-frequency attenuation) of the generated pseudo-random noise. A complementary filter in the analysis process produces overall flat response. Additionally, pink noise is similar in spectral distribution to voice and music, so the heating effect on individual loudspeakers of a multi-way loudspeaker system with pink MLS noise is similar to that which occurs during normal operation. The TDS technique (or any swept-sine method) concentrates all the input energy into a single loudspeaker driver at any single moment during the sweep, which can change driver characteristics during the test. This will often result in upward TDS sweeps yielding different response curves from downward TDS sweeps.

MLS.DSP performs measurements very rapidly. Stimulus generation, acquisition, cross-correlation, and fast Fourier transform operations are all performed within the System One and are thus independent of computer speed. On the order of two seconds total is typically required for these functions. Transmission time of the data to the computer for graphing and/or limits comparison is computer-speed-dependent, and typically takes one to two seconds with a 386-based computer. The operating speed of MLS.DSP is identical whether evaluating a short section of the record for anechoic response or longer portions for integrated room response. The signal need not be re-acquired or re-correlated in order to evaluate response of various impulse response portions with various window or processing functions.

IMPULSE RESPONSE OF LINEAR SYSTEMS

Any linear device or system may be completely characterized by knowing its impulse response, that which results when the device under test is stimulated with an infinitely narrow pulse of infinite amplitude. In practice, if the stimulus pulse width is short compared to the length of the impulse response the measurements will still be accurate. However, data must be acquired for the full length of time occupied by the impulse response of the device under test. This time is dictated by the lowest frequency desired in the measurement. An additional, and usually more stringent limitation, is that the reverberation time of the measurement space must allow the signal to adequately decay before another impulse may be generated.

The impulse amplitude is limited by the peak signal-handling capability, but the energy in the test is represented by the rms signal value. The energy directly sets the signal-to-noise ratio for any given background noise level. A single impulse has infinitesimal energy content, and so has very poor signal-to-noise ratio. Because of this, most engineers working with impulse test signals average the results of many impulses to improve signal-to-noise ratio. However, the impulse repetition rate is limited by the required acquisition time and reverberation time. If the impulses are coherently averaged and the noise is not synchronous with the impulse repetition rate, the impulses will reinforce each other and the interference will tend to zero.

Psuedo-random noise can be viewed as a random sequence of impulses, some positive and some negative. This sequence repeats at a specific rate called the repetition rate of the noise. If the device under test is linear, the response to the psuedo-random noise will be the sum of the responses to the individual impulses. When these impulses arrive to be measured, the DSP effectively shifts each one in time to align them at the same point and averages them together. As long as the interfering noise is asynchronous to the signal, the noise will average out toward zero. The averaging operation creates a single impulse response which has lower noise than any of the individual impulse responses. The signal-to-noise ratio improvement is proportional to the square root of the number of impulse responses averaged. For the 32767-point psuedo-random sequence used in the DSP, the noise improvement relative to a single impulse is 181 times or 45 dB. MLS.DSP offers a selection of four different maximum length sequences, each 32k long. Each will cross-correlate to approximately -45 dB against any of the other three. This feature permits up to four different System One equipped test stations to be located near one another on the production test floor with no significant interference effects.

In the frequency domain, psuedo-random noise has a flat spectrum, with components spaced at the repetition rate of the noise. For a 32,767 point sequence operating at a 48 kHz sample rate the psuedo-random sequence will repeat every 0.68 seconds, or a rate of 1.46 Hz. This is called white noise and has equal energy per unit of bandwidth when analyzed on a linear frequency scale. The ear hears on a logarithmic scale (in fractions of an octave) and the spectrum of most interfering noise is also flat on a logarithmic frequency scale. Consequently, white noise produces more energy than is necessary at high frequencies and, conversely, less energy at low frequencies than is desirable. For example, in the octave band from 40 Hz to 80 Hz there will be 27 frequency components of the noise. In the octave band from 10 kHz to 20 kHz there will be 6849 components, 254 times as many or a power level about 24 dB higher. To compensate for this effect, the DSP software filters the test signal with a digital pinking filter, plotted in Figure 1, which attenuates the higher frequency components in direct proportion to their number. The result is equal power in each octave band, providing a more constant signal-to-noise ratio across the measurement frequency range.

The cross-correlation operation required to shift the individual impulses in time and average them together is accomplished with a Fast Hadamard Transform. For a description of the Fast Hadamard Transform and its application to MLS testing see Borish and Angell 1983. For a technique to simplify computation of the Hadamard Transform see Borish (1985).

ACOUSTIC MEASUREMENT TECHNOLOGY IN THE AUDIO PRECISION SYSTEM 1

TIME WINDOWS FOR TIME-TO-FREQUENCY TRANSFORMS

The FFT algorithm is used to transform a segment of the impulse response into the frequency domain in order to see the frequency response. This segment is selected from the original impulse response by setting to zero all data outside the region of interest. If signal in the data record being transformed does not naturally decay to zero at the beginning and end of the segment, there will be sharp discontinuities introduced by the selection of this segment for transforming. These discontinuities in the waveform produce large amounts of high frequency energy in the transformed result. This high frequency energy results in ripples on the displayed frequency response curve.

To alleviate this problem, a window may be applied to the data. The idea behind a window is to gradually taper the data at both ends of the record toward zero so that it will always make a smooth transition with the following and preceding repetitions of the record. This is accomplished by multiplying each point in the data record by a mathematical function which is near unity (1.000) in the center of the data record and small at the ends of the record. The simplest such function is an inverted cosine wave raised above zero with an added DC offset so that its negative peaks just reach zero. After multiplication by the window function the data record goes to zero at the ends and so smoothly meets each data record on either side of the one being transformed. Multiplying the data by the window function does alter the spectrum of the original signal. As might be expected by visualizing the envelope of the repeating windowed data record, the frequency response will be smoothed. However, the spurious high frequency components produced by the sharp discontinuities will have been eliminated. The raised cosine window described is called the HANN window after its inventor, Austrian meteorologist Julius von Hann. (It is often incorrectly called a Hanning window due to confusion with the Hamming window, named after its inventor Richard Hamming.)

The term window is used because it restricts the view of the FFT to the central portion of the data record in much the same way a window restricts the view of a person looking through it. There are a wide variety of windows developed which trade off resolution in frequency against the ultimate attenuation of the spurious energy created by the ends of the data record.

Transformation of impulse responses of loudspeakers is a special case, since the typical impulse has a fast rise and slow decay. Thus, it is desirable to use an asymmetrical window function which also has a fast rise and slow decay in order to taper values at the two ends of the selected portion to zero with minimal effect on the important information in the impulse. The window selection is therefore made of two sections as shown in Figure 2. The first portion is a "half-window" beginning at the selected START time and a second "half-window" ending at the selected STOP time. The nominal time selections (<5%, <10%, <20%, <30%) refer to the percentage of the selected portion of the impulse across which the half-window makes its full transition. The less than sign (<) indicates that the actual percentage may be less than specified, since the actual number of samples for the transition from zero to unity (or vice-versa) will always be an exact binary power such as 4, 8, 16, etc. The DSP therefore rounds down from the selected value to the largest exact binary power within that value. For example, assume a START time of 4 ms and a STOP time of 8 ms, resulting in a 4 ms time span. If the 48 kHz sample rate (20.8 microsecond sample period) is in use, there will be approximately 192 samples in the selected span. A START selection of 5% would nominally make the transition in the first 5% of the 192 samples, or 9.6 samples; the actual transition time will be rounded down to 8 samples as an exact binary power. Similarly, if

30% is chosen for the STOP selection, the nominal transition would take place during the final 30% of the 192 samples, or 57.6 samples. The actual transition time will be rounded down to 32 samples, the next lower exact binary power. All the data between the 8th sample from the beginning and the 32nd sample from the end will be unattenuated. For an excellent technical discussion of windows and their characteristics see Harris (1978).

WAVEFORM DISPLAY

Several display modes are available for graphing results in time or frequency domain. If INTERPOLATE is selected the DSP will compute the data value corresponding to the exact time or frequency value specified during a sweep. If the specified time or frequency value exists the DSP will return the corresponding measurement value from memory. If the value does not exist it will be interpolated from the nearby measured values. Interpolation computes a value from the 15 FFT bins centered at the requested frequency, thus smoothing out the stair-step appearance of frequency response curves at low frequencies where the bin width (usually 2.93 Hz at the 48 kHz sample rate) occupy a significant portion of the screen.

If NORMAL is selected the DSP will return the closest actual measured value and will never alter the data. This may produce adequately faithful reproduction of impulse responses when the signal changes slowly relative to the sample rate. Each signal peak is then represented by many samples and the display will be relatively smooth. When the impulse variation is fast relative to the sample rate, however, each peak is represented by only a few points.

The PEAK mode will return the largest value between the last requested sweep point and the current one. This avoids missing signal peaks which might exist between the plotted points (graphic aliasing). If the vertical axis of the graph is set to a linear scale the largest data point will be returned without regard to its sign. If a LOG scale (or linear with a dB unit) is selected the absolute value of the data will be returned, allowing proper log computations in the computer. This is helpful in viewing low level reflections in impulse responses.

SETTING TIME SPANS FOR FREQUENCY DOMAIN ANALYSIS

The section of the impulse response which is to be transformed into the frequency domain is selected by the most recent set of START and STOP time values used in a time domain display sweep. See Figure 3 for an illustration of the first 30 milliseconds of an impulse response of a loudspeaker. If it is desired to limit the analysis to only the direct sound, for example, the START and STOP values must set to the beginning and end of the direct sound arrival. If the characteristics of the impulse response are not known in advance, the START and STOP values can be set to a wide span which includes the whole impulse response. The direct sound and the first reflections can be identified and the graph "zoomed in".

The setting of START and STOP values limits all subsequent frequency domain analysis to the portion of the time record bounded by these values. An impulse response will oscillate or be active for a length of time roughly proportional to the period of its lowest frequency component. The time record must be long enough to include this oscillation to obtain meaningful information about the lowest frequency component. For accurate measurements this oscillation must be allowed to die down for several cycles, requiring a time record approximately three times the period of the lowest frequency of interest. For example, the period of a 200 Hz signal is 5 ms, implying a 15 ms required time span. For loudspeaker measurements this also implies an approximate 5 meter (16 foot) distance between arrival paths of the direct sound and the interfering reflections, which requires an unusually large space.

The REF TIME field is used to compensate for the effect of time delay on the measurements. This is of concern only when making phase measurements since a time delay becomes an increasing phase shift with frequency. This field is normally set to a value slightly greater than the START time value. These values are usually not identical since the impulse being measured is spread out in time by the device under test and the START time must be set to include the build-up of the impulse, while the REF time is set to the exact peak of the impulse. In practice, the REF time is "fine tuned" by examining the unwrapped phase response and adjusting the value for the flattest curve. It is also possible to use the COMPUTE LINEARITY function of System One software to determine the phase deviation from a linear phase (constant time delay) measurement. This reduces the need to adjust the REF TIME for the exact transit delay.

DISPLAYING ENERGY-TIME RESPONSE

The time domain behavior may be examined using the impulse response (IR) or the energy-time response (ETim). The term energy-time response is a misnomer since a true computation of energy requires knowledge of both kinetic and potential energy and a microphone signal can only supply one of these. A more accurate term sometimes used in technical papers is the analytic signal magnitude. However, the term energy-time has become common usage and, to avoid confusion, will be used here. ETim curves may also be considered as similar to the envelope of the impulse response.

The software computes an estimate of the energy by applying a Hilbert transform to the frequency domain data before transforming back to the time domain. This produces an estimate of the imaginary portion of the complex impulse response. The vector magnitude of this and the original real portion of the impulse response is the energy estimate. The resulting trace will not show the negative excursions of the impulse response. This display is useful for determining arrival times and relative energy distribution in time.

Since the energy-time graph is computed with transforms, a window must be applied to the data to prevent alias behavior. The MLS.DSP choices are no_window, half_Hann, Hann, 240_8kHz, and 120_16kHz. These windows operate in the frequency domain and are plotted in Figure 4 and Figure 5. The no_window selection will perform the required transformations with all frequency components of the signal included in the computations. The deviations from a flat frequency response create ripples in the time domain energy response. The Hann selection is the one window found on software from other manufacturers. This reduces both high and low frequency energy, concentrating on arrivals at the center of the frequency range. Since the processing occurs on a linear frequency scale, this will focus analysis on signals around one quarter of the sample rate. At 48 kHz this will result in the 12 kHz energy dominating the energy-time display, producing very attractive displays which are very wrong.

The half_Hann selection is a window suggested by Lipshitz and Vanderkooy (1990) which only reduces the contribution of high frequencies. The low frequency information remains unchanged. When operating at the 48 kHz sample rate this window filters out energy above 12 kHz. Audio Precision developed the remaining two windows for even more accurate measurements of typical audio signals. The 240_8kHz window filters energy below 240 Hz and above 8 kHz, producing equal sensitivity to signals over a 5 octave range. The 120_16kHz window spreads the analysis over a 7 octave range. Both windows produce much more accurate results than the Hann window with only minor increases in alias behavior.

RESOLUTION AND ACCURACY

The impulse response measurement length is always the maximum which can be acquired into the available memory. This is done to reduce the risk of errors which occur when the impulse response being measured is longer than the measurement length. When this happens the portion of the impulse response which exceeds the measurement buffer will fold over to the beginning of the buffer. This is a form of time domain aliasing as illustrated in Figure 6. The measurement accuracy will be limited by the amplitude of the portion which folds relative to the desired signal. If the largest amplitude in the aliased portion of the impulse response is 1% (-40 dB) of the non-aliased portion, frequency response errors will be approximately 1% or 0.1 dB. Loudspeaker impulse responses are seldom long enough to create a problem on System One. However, when a loudspeaker is measured in a non-anechoic room, the room impulse response will dominate and can create errors. A System One will provide 32787 samples of impulse response before aliasing. This is 683 ms at a 48 kHz sample rate or 1.024 seconds at a 32 kHz rate. The room reverberation time is a rough indicator of when trouble might occur since it is the time required for the room energy to decay by 60 dB. If measurements to 1% accuracy are desired the energy must decay by 40 dB within the time represented by the measurement buffer. Assuming the reverberation is a single exponential decay, the allowable reverberation time (at a 48 kHz sample rate) would be 1.024 seconds.

The equivalent frequency resolution of an FFT in the general case is obtained by dividing the sample rate by the number of waveform samples. The MLS.DSP transform length is 16,384 samples. At a 48 kHz sample rate, the resulting FFT will consist of 8192 spectral lines (bins) evenly spaced from zero Hz to the folding frequency (1/2 sampling rate). The resulting bin width and frequency resolution is therefore approximately 2.93 Hz.

With MLS.DSP, actual resolution is further limited by the fact that only a specific time span of the acquired signal is usually selected for analysis. This is selected by the START and STOP time parameters. The remainder of the transform buffer is padded out with zeros. Actual resolution is determined by the number of non-zero samples. That actual resolution is then effectively interpolated, with the interpolation resolution (apparent resolution) being the 2.93 Hz value of the FFT. Typical selected time spans are on the order of several milliseconds for anechoic measurements (with limited low-frequency response) up to tens of milliseconds for full frequency-range measurements. At the 48 kHz sample rates, these time spans may thus include on the order of 150 to 1500 data samples and produce actual resolutions on the order of 320 Hz (48000/150) to 32 Hz (48000/1500). Interpolation will make it appear that the resolution is better, but the fundamental resolution cannot be greater than the frequency whose period equals the selected time span.

REFERENCES

J. Borish, Self-Contained Crosscorrelation Program for Maximum Length Sequences, Journal of the AES, Vol. 33, No. 11, November 1985, pp. 888-891.

ACOUSTIC MEASUREMENT TECHNOLOGY IN THE AUDIO PRECISION SYSTEM 1

J. Borish and Angell, An Efficient Algorithm for Measuring the Impulse Response Using Psuedo-Random Noise, Journal of the AES, Vol. 31, No. 7, July/August 1983, pp. 478-488.

F. J. Harris, On the use of Windows for Harmonic Analysis with the Discrete Fourier Transform, Proc. of the IEEE, Vol. 66, No. 1, Jan 1978, pp. 51-83.

S. P. Lipshitz and J. P. Vanderkooy, Uses and Abuses of the Energy-Time Curve, Journal of the AES, Vol. 38, No. 11, November 1990, pp. 819-836.

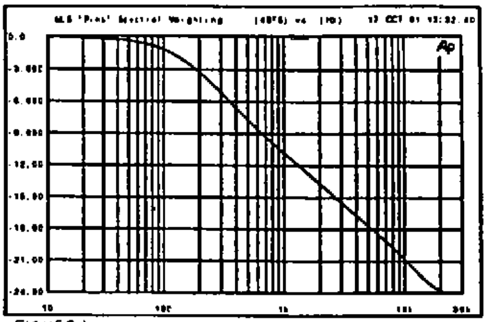
FIGURE 1

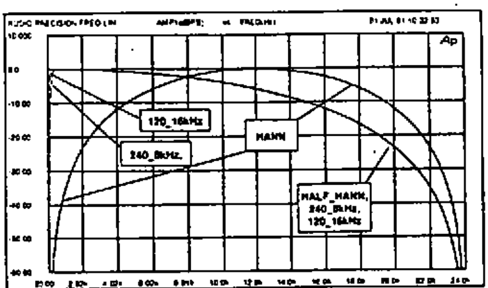

Figure : Frequency Windows for ETim Curves, Linear Frequency Axis

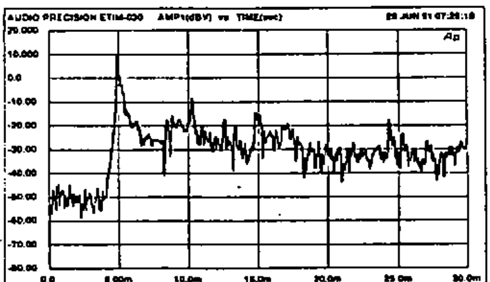

Figure 2 Energy-Time Response of Loudspeaker, First 30 Milliseconds

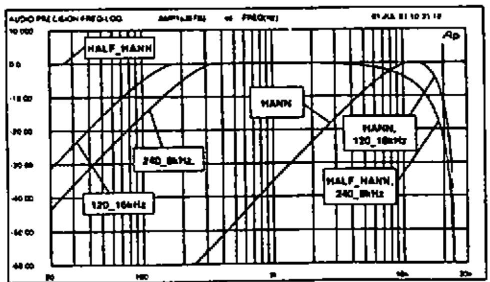

Figure 5 Same Frequency Windows as Previous Figure, Logarithmic Frequency Axis

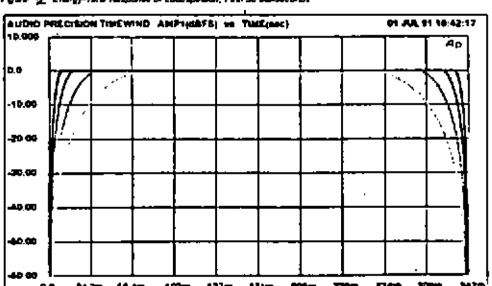

Figure : Time Window Selections for Impulse-to-Frequency Transformation. Typical Selections Are Short "Half-Window" at START, Longer "Half-Window" at STOP.

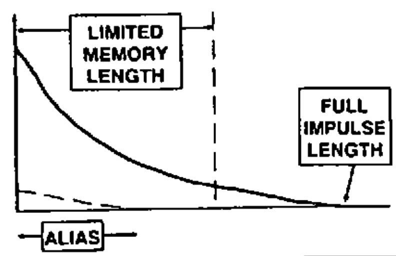

Figure : Time Domain Alias When Impulse Length Exceeds Memory Length