☐ 2700 Series
☑ APx555
☑ APx585 Series
☑ APx525 Series
☑ APx515

Transfer Function Measurements with APx500 Audio Analyzers

by Joe Begin

About This Technote

In this technote we discuss the Transfer Function measurement added to the APx500 control software in version 5.0. We provide background information on transfer function measurements in general, followed by some practical examples of applying this measurement technique to some difficult audio test problems.

One of the key attributes of transfer function analysis is that it provides a means of measuring the frequency response of a device using any broadband signal, including speech and music. This makes it an ideal choice for analyzing devices used for speech communication (i.e., smart speakers, smartphones, headset microphones, etc.). Many of these devices incorporate DSP algorithms that require the use of speech signals, and some are designed to block sinusoidal signals altogether. Transfer function analysis greatly simplifies measuring the frequency response of such devices.

Background

The term transfer function is used in System Analysis, which seeks to characterize the response of a *System* to a time-varying stimulus (Figure 1). We stimulate the *System* with input signal *x(t)* and measure its response - output signal *y(t)*.

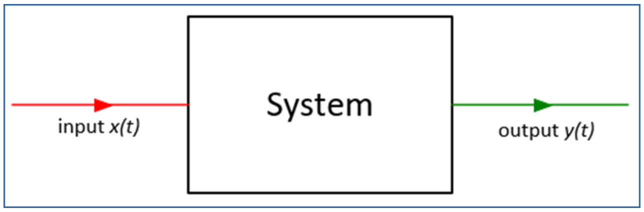

Figure 1. Block diagram of a system.

In the general case, any physical device or mathematical system could be tested or analyzed. However, in the application space of audio and electroacoustic test, the system or device under test (DUT) will most likely be an electronic circuit, an audio device, a filter, a loudspeaker, a microphone, etc.

As audio engineers, we are most often interested in characterizing a System in the frequency domain. For example, a classic, fundamental measure of audio quality is a device's frequency response, which usually refers to its output per unit input (i.e., its gain) as a function of frequency. As it turns out, using a mathematical technique known as the Fourier transform, it is easier to characterize a system in the frequency domain than in the time domain. Figure 2 illustrates how the Fourier transform (denoted by $\mathcal{F}$) is used to transform signals and the system response from the time domain to the frequency domain. There is also an inverse Fourier transform that works in the opposite direction.

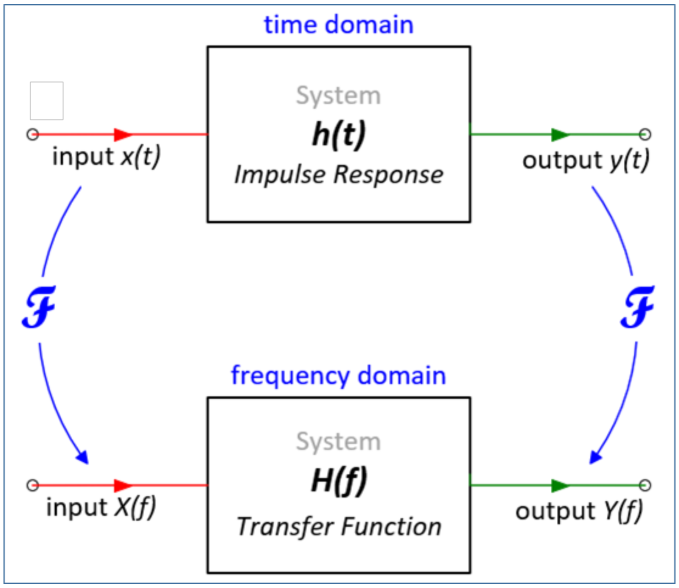

Figure 2. System response in the time and frequency domains.

As Figure 2 shows, the system response in the frequency domain is known as the Transfer Function, denoted by *H(f)*. The term "*Frequency Response Function (FRF)*" is often used interchangeably with *Transfer Function*. The transfer function is complex, meaning it has both magnitude and phase. Its magnitude represents the system's output per unit input (i.e., the gain of the system) as a function of frequency and its phase response represents the phase between the output and input as a function of frequency. As a complex quantity, the transfer function can be expressed as Real and Imaginary parts as an alternative to magnitude and phase. Note that the Transfer Function can also be used to derive the Impulse Response, which represents the System's response in the time domain.

Transfer Function Analyzers

At Audio Precision, we use the term "Transfer Function Analyzer" to refer to an analyzer capable of measuring system transfer functions by means of the complex Discrete Fourier Transform (DFT). We prefer this term over other terms used in the test and measurement industry (including "Dynamic Signal Analyzer" and "Dual-Channel FFT Analyzer"), because it is unambiguous.

A block diagram of a transfer function measurement is shown in Figure 3. On most analyzers, the measurement works as follows:

- A broadband signal is generated and applied as a stimulus to the system input. A broadband signal is a signal that contains energy at all frequencies over the frequency range of interest. Suitable signals include noise, music, speech, continuous swept sines (chirps), etc. A discrete sine signal such as a single tone or a multitone is not suitable.
- The stimulus signal from the generator is looped back and measured on one of the analyzer input channels.
- The output from the system is acquired on a second input channel (on multichannel analyzers, additional DUT output channels can be acquired and analyzed).
- The analyzer acquires the system input and output signals simultaneously with the same sample clock and calculates the system's transfer function (magnitude and phase).

An added feature of the APx500 transfer function measurement implementation is that in many applications, the stimulus signal does not have to be acquired as one of the analyzer input channels. This is because on APx audio analyzers the analog input and output systems are precisely calibrated, and both analog and digital IO systems are time aligned. Therefore, in many cases, the signal loaded into the generator or an audio file on disk can be used as the Reference (i.e., as the system input signal). This frees up one analyzer input channel for measurement. It is also a very useful feature for testing smart devices, tablets and media players that must be tested in an open loop configuration.

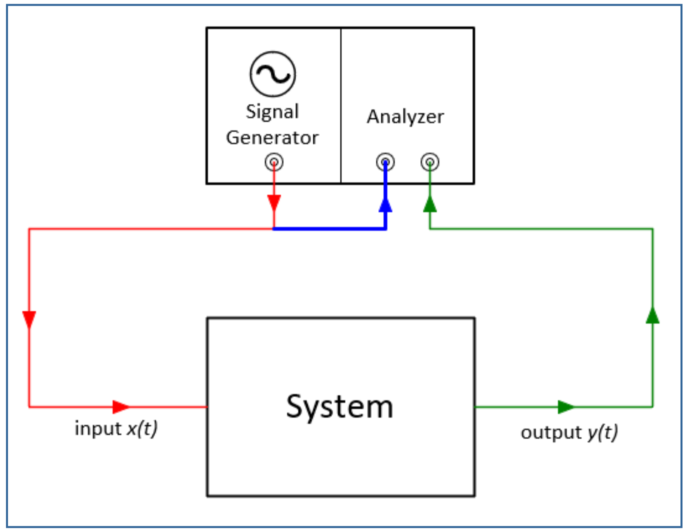

Figure 3. Block diagram of a transfer function measurement

Transfer Function Details

To apply the APx Transfer Function measurement, a basic understanding is required of how the transfer function is derived.

Definitions

First, some definitions: Consider the system in Figure 2 with input signal x(t) and output signal y(t).

The ***Cross Spectrum*** G_{xy} estimates the relationship (including phase) between x and y as a function of frequency. G_{xy} is complex (i.e., has magnitude and phase).

The ***Auto Spectrum*** G_{xx} or G_{yy} represents the power (level-squared) of the signal x(t) or y(t) as a function of frequency. The Auto Spectrum is also known as the Power Spectrum. G_{xx} & G_{yy} are not complex (they have magnitude only, but no phase).

Calculation Modes (Derivation)

There are three "Calculation Modes" available in the APx Transfer Function measurement: H1, H2 and Magnitude Only. These specify how the transfer function is calculated.

In ***H1*** mode, the transfer function is calculated as the Cross Spectrum between the output and input divided by the Auto Spectrum of the input (as shown in equation 1). This minimizes the effect of noise introduced at the system output.

$$H1: \quad H_{xy} = \frac{G_{xy}}{G_{xx}} \quad (1)$$

In ***H2*** mode, the transfer function is calculated as the Auto Spectrum of the output divided by the Cross Spectrum between the output and input (as shown in equation 2). This minimizes the effect of noise introduced at the system input.

$$H2: \quad H_{xy} = \frac{G_{yy}}{G_{xy}} \quad (2)$$

Note: For most audio test applications, the H1 calculation mode is more appropriate than H2, because the input signal comes from a precision signal generator and is likely to have a higher signal-to-noise ratio than the DUT output signal.

In ***Magnitude Only*** mode, the transfer function magnitude is calculated as the magnitude of the output spectrum divided by the magnitude of the input spectrum (as shown in equation 3). Phase is not available in this case. This mode is useful when the input and output signals have sample clocks that are not precisely synchronized (for example, when testing Bluetooth devices).

$$Magnitude\ Only: \quad |H_{xy}| = \sqrt{\frac{G_{yy}}{G_{xx}}} \quad (3)$$

The Coherence Function

A benefit of using the H1 or H2 calculation mode is that the ***Coherence Function C^2*** is also calculated (equation 4).

$$C^2 = \frac{|G_{xy}|^2}{G_{xx}G_{yy}} \quad (4)$$

Coherence is a metric of measurement quality. The Coherence Function indicates the degree to which there is "coherence" or a causal relationship between the output Y(f) and the input X(f) at each frequency. Its value varies between 0 (no relationship) and 1.0 (perfect relationship). Low Coherence is often caused by a poor signal to noise ratio.

A Measurement Example

As a first example of a transfer function measurement, consider the graphs in Figure 4, which display results from the measurement of a loudspeaker in an acoustic test box (a single driver intended for speech measurements in the frequency range from 100 Hz to about 10 kHz). In this case, the input (Reference) signal was the voltage of a pink noise signal applied to the loudspeaker and the output signal was the sound pressure in Pa as measured at the test position in the box. Figure 4 shows the Magnitude response (measured using H1 and H2 calculation modes), the Phase response and the Coherence. Note that:

- The FRF Magnitude (bottom graph) has units of Pa/Vrms. This is the gain or sensitivity of the loudspeaker as a function of frequency. For example, from 1 to 2 kHz, the sensitivity of the driver is approximately 1 Pa (94 dBSPL) per volt.
- The Coherence (top graph) is close to unity at all frequencies from about 30 Hz to 16 kHz, indicating good measurement quality within this range.

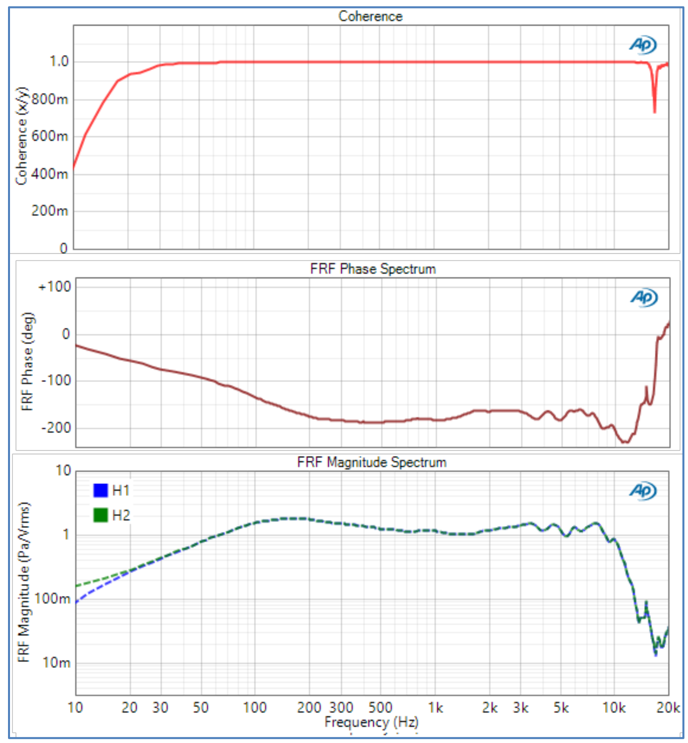

Figure 4. Transfer Function coherence, phase and magnitude (loudspeaker measured in an acoustic test box).

- The Coherence falls significantly below 30 Hz. This is likely due to a combination of factors at play when conducting acoustic measurements at low frequencies, including (1) low output from the driver, (2) the presence of higher ambient noise levels and (3) the need for longer measurement times.
- The Coherence also has a dip from about 16 to 18 kHz, likely due to the notches in the driver response in this range, resulting in a low-level output signal.

- There is not much difference between the H1 and H2 Calculation mode results, except at very low frequencies and at 16 to 18 kHz, where the Coherence is significantly less than 1.

APx500 Implementation

As of version 5.0, a measurement named Transfer Function is available in Sequence Mode. We'll describe the measurement's Generator and Analyzer controls before showing some measurement examples.

Generator Controls

The Generator control section of the Transfer Function measurement is shown in Figure 5. Using the Waveform control, you can select the built-in Noise signal, or load any .wav file containing a signal to be used as a stimulus. When Noise is selected, the Noise Shape control allows you to specify White or Pink noise.

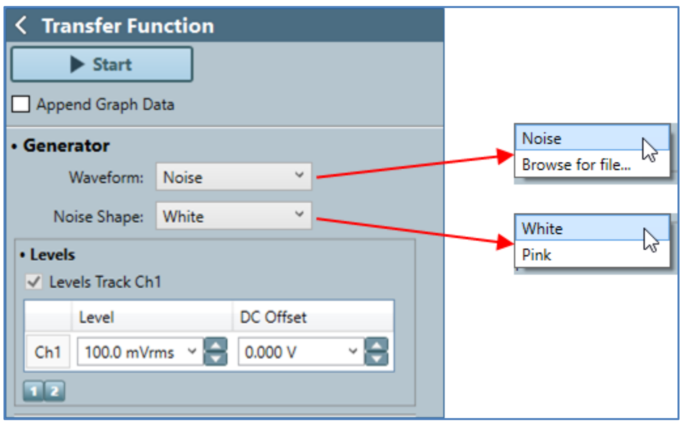

Figure 5. Transfer Function Generator controls.

Analyzer Controls

As shown in Figure 6, the Analyzer control section of the Transfer Function measurement has many controls. We'll describe them in the order they appear from top to bottom in the Analyzer Control group.

- The **Reference** control is used to specify the signal being applied to the input of the DUT. Choices are Generator or Input Channel. When a physical output connector has been selected, the default value for Reference is Generator, meaning the input signal for the transfer function analysis is taken directly from the APx signal generator buffer, eliminating the need for the stimulus signal to be measured on one of the analyzer input channels. The other choice for Reference is Input Channel, with a channel selector. This allows the Reference (or input) signal to be measured directly, if required.

- The **Match** control is used as a type of trigger. As the input signals are acquired, the system uses cross-correlation between the start of the reference signal and successive signal blocks of the specified input channel to try to find a matching signal. **Match** has choices of High, Medium and Low, which specify how closely the acquired signal must match the beginning portion of the reference signal for the trigger condition to be satisfied. The signal before the match is excluded from the transfer function analysis. The **Match** control also has a choice of None, meaning the entire incoming signal is included in the analysis.

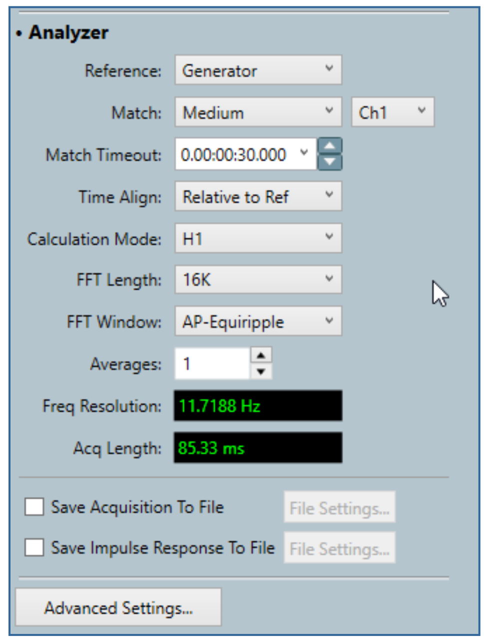

Figure 6. Transfer Function Analyzer controls.

- The **Match Timeout** specifies the time after which the analyzer aborts the measurement if a matching signal is not found. The default value is 30 seconds.

- The **Time Align** control specifies how the input signals are time-aligned with the Reference signal. It has choices of Relative to Ref and Relative to Ch.
  - When Time Align is set to Relative to Ref, all input channels are aligned to the Reference signal, and each channel's *Delay* is reported as the time between the first input sample and the match condition on that channel.
  - When Time Align is set to Relative to Ch, the specified channel's delay is removed from all input

channels such that the specified channel is time-aligned with the reference. In this case, the specified channel's Delay is reported as 0.0 seconds and for all other channels the delay reported is the time between the signal match on that channel and the specified channel. This time alignment is ideal for microphone arrays, because it preserves the phase relationship between channels.

- The **Calculation Mode** control has choices of H1, H2 and Magnitude Only (explained in the previous section).
- The **FFT Length** control specifies the number of samples analyzed for each average accumulated in the transfer function analysis. It has choices from 1k to 1.2M.
- The **FFT Window** control specifies the type of window used in the analysis, with a default value of AP-Equiripple.
- The **Averages** control specifies the number of data blocks that are accumulated to the transfer function averaged results.
- When **Averages** is set to greater than 1, the Overlap control becomes visible. This specifies the overlap between successive data blocks analyzed, with choices of 0, 25, 33, 50, 67 and 75%. Overlap processing is more efficient when analyzing signals that have an envelope which varies with time (like speech or music).
- **Freq Resolution** and **Acq Length** are indicators provided for user convenience.
- **Freq Resolution** displays the current frequency resolution, which depends on the sample rate and the FFT Length.
- **Acq Length** displays the total length of the signal to be analyzed in seconds. This depends on the sample rate, FFT Length, Number of Averages and Overlap.
- The **Save Acquisition to File** checkbox is used to direct the system to save the acquired signal(s) as one or more .wav files.
- The **Save Impulse Response to File** checkbox is used to direct the system to save the calculated Impulse Response result(s) as one or more .wav files. Note that each Impulse Response Result is arbitrarily scaled such that the largest positive or negative peak in the impulse response is scaled to +1.0 D or –1.0 D (i.e., digital full scale).

Measurement Results

The Transfer Function measurement has 16 primary results, as shown in Figure 7. Note that when the Calculation Mode is set to Magnitude Only, the Phase, Coherence and Impulse Response results are not available.

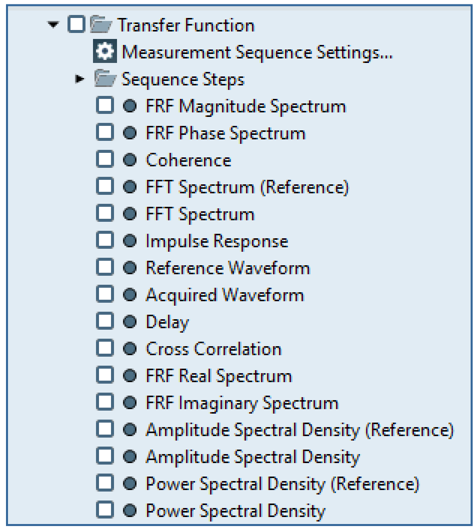

Figure 7. Transfer Function measurement results.

Measurement Examples

One of the most challenging tasks in audio test is measuring the frequency response of a smart speaker's input and output systems. Smart speakers are unique in the sense that audio input to the device is in the form of a spoken command that begins with a "wake-up word." The acoustic command signal is digitized and transmitted via the web to a server for processing. Output from the device is in the form of a speech or music signal that originates on a server and is played over the device's powered speaker(s).

In the next sections we provide examples of testing the input and output systems of a smart speaker.

Smart Speaker Input System

By input system, we refer to the part of a smart speaker system that records a spoken command and sends it to the "cloud", where the system back-end uses speech recognition to interpret the command and respond accordingly. Some systems, (for example, Amazon Alexa) record the spoken commands as audio files that can be played back and listened to, from the user's web portal to the service. In this case, the recorded audio can be retrieved as a .wav file. Using this .wav file and a .wav file containing the original spoken command, the Transfer Function measurement in APx can be used to determine the frequency response of the system.

When testing devices, it is good practice to adhere to industry standards. Although there is no standard governing the test of smart speakers, we can borrow from test standards for similar devices, such as ETSI TS 103 738, which covers tests of speakerphones. This standard recommends that tests be conducted in an anechoic chamber, with the DUT positioned on a table 40 cm from the table edge. A mouth simulator (a special loudspeaker) located at the table edge and 30 cm above the DUT's microphone system is used to generate the input signal to the DUT. A diagram of the test is shown in Figure 8.

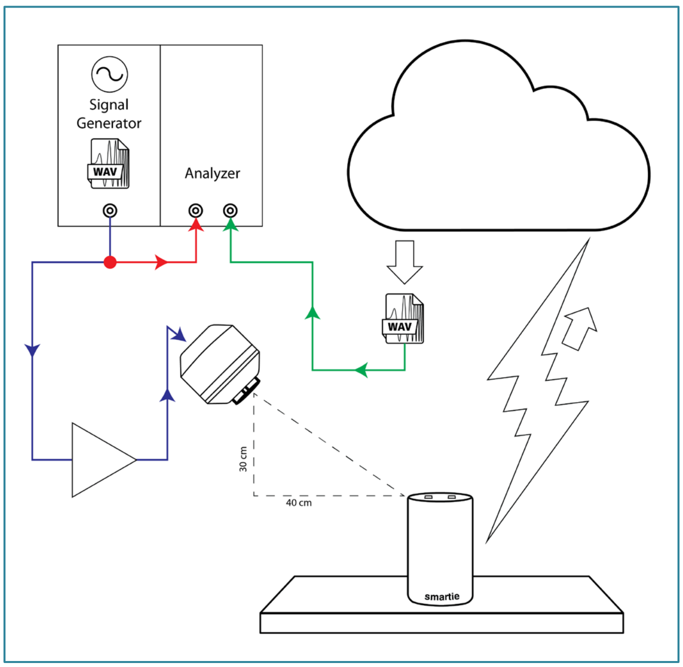

Figure 8. Schematic diagram: Transfer function measurement of a smart speaker's input system

We used the following procedure to test the smart speaker microphone function:

1. To calibrate the mouth simulator, a measurement microphone is required at a special point called the Mouth Reference Point (MRP). First, this microphone must be calibrated using a sound level calibrator (or TEDS with the APx1701) so that the APx analog input system displays results in dBSPL.
2. Position the reference mic at the MRP and calibrate and EQ the mouth simulator so that its frequency response is flat ± 0.5 dB over the frequency range of interest (100 Hz to 16 kHz)– see Figure 9.
3. Select a speech stimulus signal that includes broadband energy from 100 Hz to 8 kHz. (We used a recording identified as "nightclub chatter" available from freesound.org).

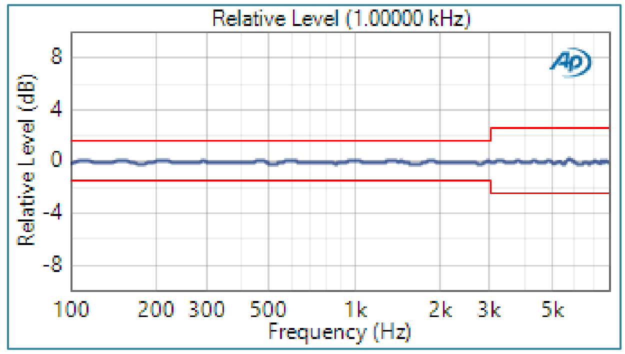

Figure 9. Mouth simulator frequency response after EQ applied.

4. Using an audio waveform editor, prepend the smart speaker wake-up word followed by a short period of silence to the speech stimulus signal (Figure 10).

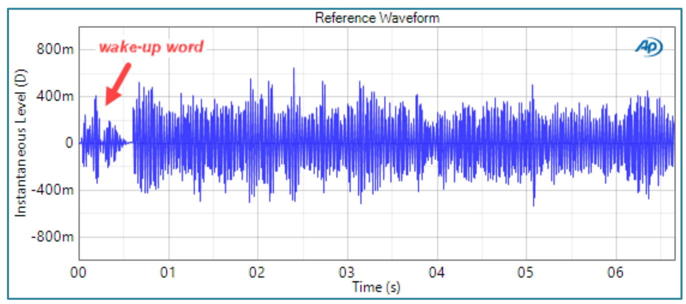

Figure 10. Waveform of wake-up word followed by speech chatter used to stimulate the smart speaker input system.

5. To conduct a transfer function analysis, the stimulus and response signals must have matching sample rates. Using the audio editor, resample the stimulus signal with wake-up word to match the sample rate of the audio file recorded by the smart speaker system (16 kHz in the case of this example), and save it as a .wav file. This becomes the Reference waveform for the transfer function analysis.
6. Load the .wav file as a Generator waveform in a measurement that can be used to measure the overall RMS level of the signal (e.g., Noise Recorder with an RMS Level derived result). Measure the RMS level at the MRP with the reference mic and adjust the generator level until the target sound pressure level (SPL) is measured. Note: For speakerphones, the ETSI standard specifies -4.7 dBPa (89.3 dBSPL) at the MRP.

7. Turn on the smart speaker and generate the stimulus signal.

8. Log in to the smart speaker cloud portal and retrieve a .wav file of the signal recorded by the smart speaker. Note: Using a web browser in developer mode, we were able to copy the URL of the audio file to a separate web page and then save it as a .wav file.

9. In the APx software, change the Output Connector to None (External) and the Input Connector to File (Digital Units).

10. Add a Transfer Function Measurement to the Signal Path and configure it as follows:

    a. Set the Reference to the .wav file created above.

    b. Add the .wav file retrieved from the server to the File List.

    c. Set the Match to Medium.

    d. Set Time Align to Relative to Ref.

    e. Set the Calculation Mode to Magnitude Only. Note: Magnitude Only is used in this application because there is virtually no chance that the sample rate of the file acquired by the smart speaker system will exactly match the sample rate of the Reference file.

    f. Set the FFT Length as required (for example 16k yields a frequency resolution of just less than 1.0 Hz).

    g. When using a signal such as speech chatter, the smart speaker system will be unable to recognize the speech, and will give up after a few seconds, resulting in a short file. In the example shown here, the recorded audio file was only 7 seconds long (Figure 11). To maximize the use of such a short file, set the Overlap to 75% and the Averages such that the Acq Length indicator is close to the length of the file.

    h. Analyze the transfer function and observe the results.

Figure 12 shows the spectra of the Reference signal and DUT signal on the same graph, and Figure 13 shows the Frequency Response Magnitude, which in this case is calculated as the ratio of the two measured spectra. Note that the measured FRF magnitude in Figure 13 appears to become progressively "noisier" at frequencies greater than about 1 kHz. This is a consequence of the DUT waveform's sample rate being slightly different than that of the Reference waveform. As shown, when the measured FRF magnitude is smoothed (1/12-octave smoothing was used in this case), the smoothed waveform appears to be a good representation of the mean value.

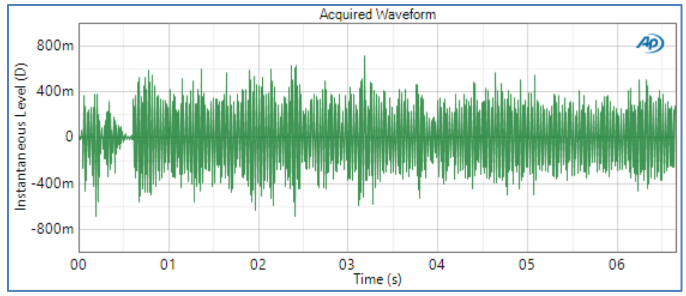

Figure 11. Waveform recorded by smart speaker input system and retrieved from the cloud.

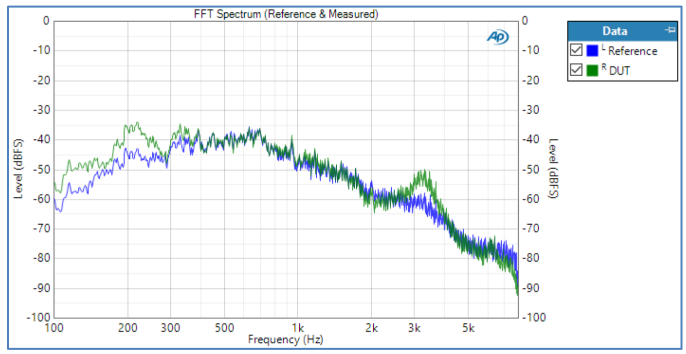

Figure 12. FFT Spectra of stimulus signal (Reference) and signal recorded by smart speaker input system.

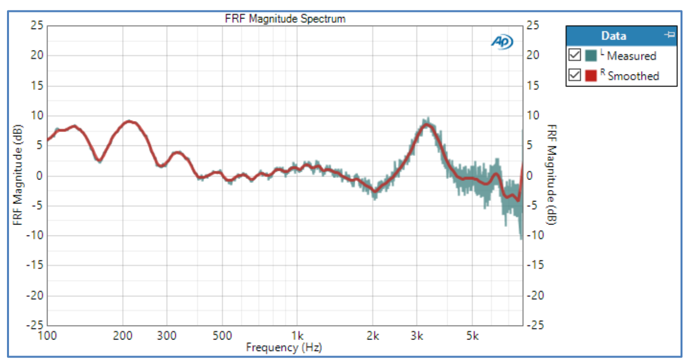

Figure 13. Frequency response magnitude of the smart speaker input system determined using the Transfer Function measurement in Magnitude Only Calculation mode.

For comparison purposes, we measured the frequency response of the same speaker input system using Open Loop Chirp with sweep lengths of 0.35, 1.0 and 4.0 seconds. As shown in Figure 14, the DUT's response due to a chirp signal is quite different, especially as the sweep length is increased. This is probably caused by the DSP in the smart speaker system recognizing the chirp as a sinusoidal signal and attempting to block it. This highlights the importance of using a real speech stimulus for such device.

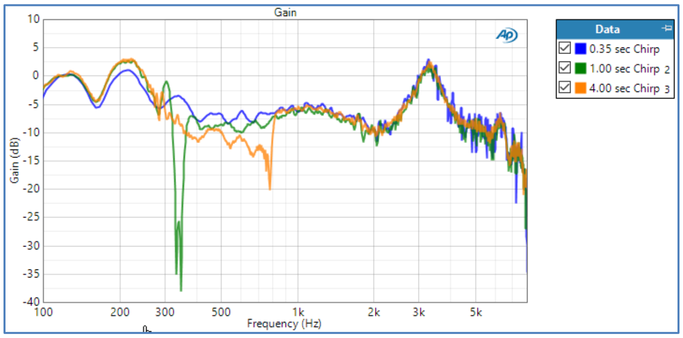

Figure 14. Frequency response magnitude of the smart speaker input system as measured by open loop chirp.

Smart Speaker Output System

The signal output from a smart speaker system originates as an audio file on a server. A music signal can be used to measure the transfer function of this subsystem, but to do so, the analyzer must have a copy of the original file to use as the Reference. The easiest way to accomplish this is to create a Reference waveform file and then load it onto a server which can be accessed from the smart speaker. Some systems allow you to place an audio file directly on the smart speaker's server using your login account. Others will not allow this, in which case you can put the file on a 3rd party server which the smart speaker has access to.

Regardless of which server is used, there will likely be restrictions on the file format and sample rate, etc. For example, for this Technote we used a Plex server with the Amazon Alexa system after installing the Plex "skill." This required an audio file with a sample rate of 44.1 kHz, encoded to MP3 format, with multimedia tags identifying the song and artist. Most waveform editors can be used to change a file's sample rate and save it to MP3 format. And multimedia tags can be edited with programs like Windows Media Player.

As mentioned earlier, the Transfer Function measurement uses a matching algorithm, to time-align the output signal from the DUT with the Reference waveform. This is particularly useful when testing a smart speaker, because after issuing a voice command like, *"Smart speaker – play my_Song by Artist from my_Server"*, if the system recognizes the command and has access to the song, it will respond sometime later with a message like, *"Okay, playing my_Song by artist on my_Server"*, followed by the requested audio track. With the Transfer Function measurement, you can start the measurement at any time before the response and use the Match algorithm to time-align the music signal with the Reference waveform.

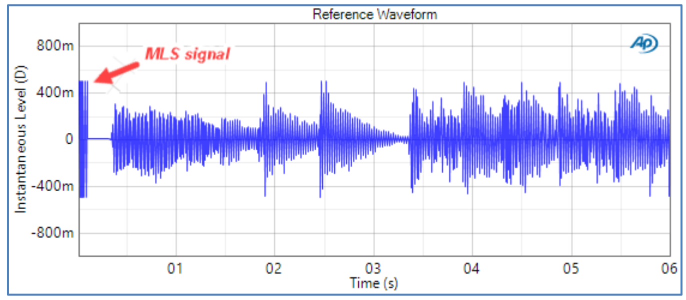

Figure 15. Beginning section of the music waveform with pre-pended MLS signal.

One potential problem with using a music signal is that music is typically quite repetitive (i.e., most songs have rhythm), which increases the likelihood of the algorithm triggering on the wrong part of the signal. We have found that one way to ensure a good match is to prepend to the stimulus signal a short burst of a signal known as a Maximum Length Sequence (MLS). An MLS is a special waveform that is much like white noise, but has properties that make it easy to detect by the match algorithm. A series of .wav files containing short MLS signals at different sample rates are available at AP.com for this purpose. Figure 15 shows the beginning of our music signal with a 100-millisecond burst of an MLS signal at the beginning.

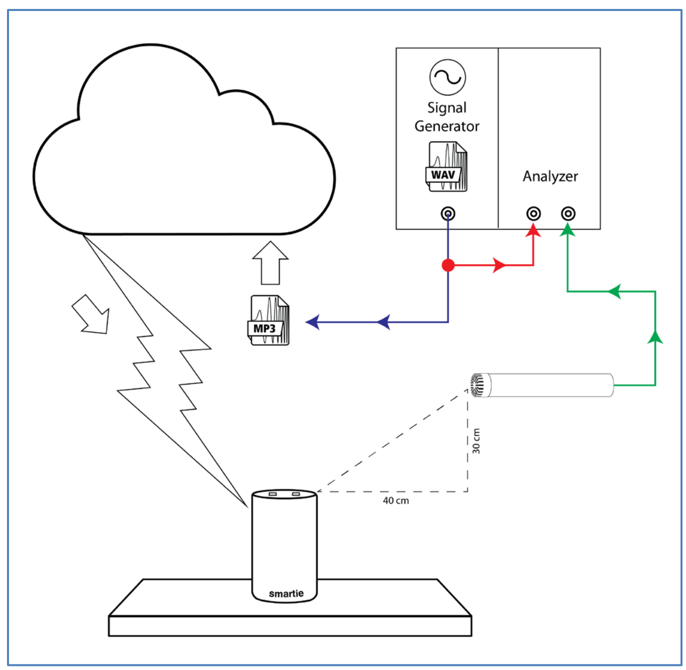

Figure 16. Schematic of smart speaker output system test.

The test setup for a smart speaker's output path is similar to the input path, but instead of a mouth simulator, a measurement micropone is located at the table edge and 30 cm above the DUT, to measure the loudpseaker output. A schematic of the test is shown in Figure 16.

We used the following procedure to test the smart speaker output function:

1. Upload the Reference audio file created above to the server and ensure that the smart speaker can access it.
2. In the APx software, change the Output Connector to None (External) and the Input Connector to Analog Unbalanced (or Transducer Interface if using an APx1701).
3. Set the input mode to Acoustic and calibrate the microphone so that input units are Pa (or dBSPL).
4. Add a Transfer Function Measurement to the Signal Path and configure it as follows:
    a. Set the Reference to the .wav file created above.
    b. Set the Match to Medium.
    c. Set Time Align to Relative to Ref.
    d. Set the Calculation Mode to Magnitude Only (again, the system will have a slightly different sample rate than the Reference .wav file).
    e. Set the FFT Length as required (for example 16k at 44.1 kHz yields a frequency resolution of just less than 3 Hz).
    f. Set the Overlap and Averages controls as needed. In this example we used 50% overlap with 322 averages, resulting in a total acquisition length of 60 seconds, (the length of the Reference waveform).
    g. Tell the smart speaker to play the Reference audio file and start the Transfer Function measurement. Note: As an alternative, you can record the voice command and generate it from a different Sgnal Path through a speaker in the test chamber.
    h. The audio analyzer will trigger when it finds a signal match and process the acquired data. Observe the measurement results.

Figure 17 shows the beginning of the acquired input signal. Note the similarity to the Reference signal (Figure 15). In this case the Delay meter result had a value of 1.375 seconds, indicating the time after the measurement start when the matching signal was found.

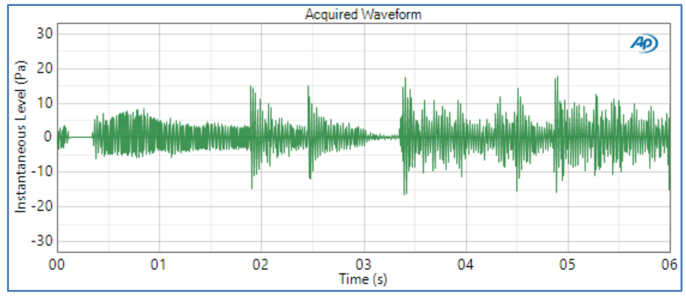

Figure 17. Smart speaker acoustic output signal (note Pa units).

Figures 18 through 20 show the spectrum of the Reference signal, the spectrum of the smart speaker output signal and the transfer function magnitude. In spite of the "peaky" nature of the Reference and response signals, the transfer function magnitude is a relatively smooth curve. Additional smoothing could be applied using a Smooth derived result.

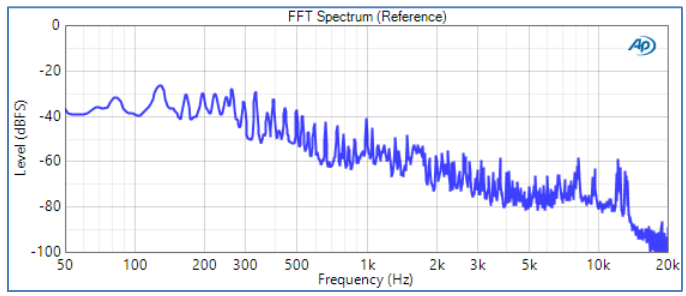

Figure 18. Spectrum of the Reference signal.

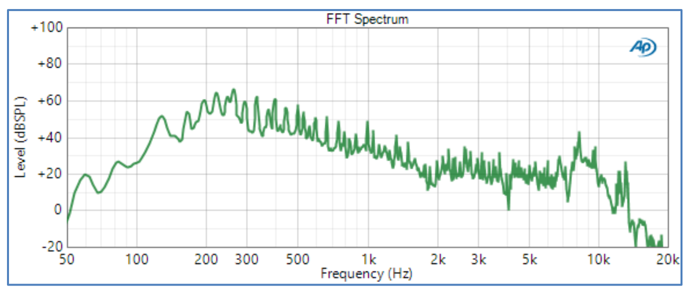

Figure 19. Spectrum of the acquired acoustic signal.

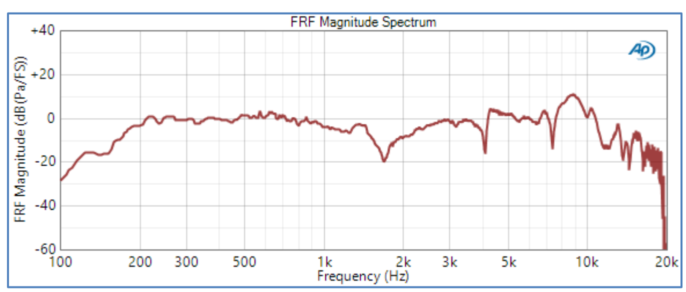

Figure 20. The smart speaker's measured transfer function magnitude in dB(Pa/FS) without smoothing.

Conclusion

The Transfer Function measurement is one more tool in the APx measurement toolbox that can be used for a variety of applications. It is particularly useful for devices that require a speech stimulus and for devices that must be tested open loop.

5750 SW Arctic Drive, Beaverton, Oregon 97005 | 503-627-0832
www.ap.com

MMXIX0125163100