Journal for Undergraduate Research in Physics
7 September 2011

Method for Extracting the Frequency Response of an Audio System from a Recording

Herman Orgeron

Abstract. The ability to compare the technical performance of audio systems is necessary to choose the microphone system that best meets operational requirements. A common method to evaluate a system's performance is to generate a transfer function for the system and determine the frequency response. Normally this is done by connecting the system to a network analyzer, however not all systems have this ability. This paper outlines procedures developed to non-traditionally measure the frequency response of audio systems using recorded data files.

INTRODUCTION

There are a variety of systems used to record and transmit audio signals. Each system has its own advantages and disadvantages, typically trading off cost, size, and power consumption for audio fidelity. In order to select the right system for a given situation, it is desirable to use unbiased indicators of performance to determine which system has the best audio quality for a given set of circumstances.

Therefore, it is necessary to develop a simple method to quantify and compare systems that operate in vastly different ways. One of the best ways of quantifying audio quality is to measure a system's frequency response—a gauge of a system's output spectrum given a specific input. The frequency response is used to graphically display how the recording system changes or distorts the input audio signal. The most desirable result is that the plot of the frequency response is a nearly flat line, meaning that at any given frequency the system's response to the input signal is constant.

Historically, an audio system's frequency response was determined by placing the microphone in an anechoic chamber and sweeping through the audio spectrum with a well defined noise source, i.e. a loudspeaker. The system output would then be connected to a network analyzer which would display its response across the desired audio spectrum. This technique is not available to digital systems that record audio for later review—there is no connection to allow audio to be fed into an analyzer. The lack of a real time analog audio stream prevents the use of traditional network analyzers and associated procedures. This complicates extracting the frequency response of the system and must be taken into account when developing a test methodology.

I. APPROACH

The method presented in this paper is designed to generate the frequency response for systems that do not produce continuous audio output—instead they save the audio data into packets or files and all of the analysis must be done post-process. This method utilizes data acquisition and processing programs—in this instance MATLAB was used due to availability—and computer data acquisition cards to acquire the audio from multiple systems and to process the system frequency response.

The basis for this method is that if two systems, one whose properties are known, and one whose properties are unknown, are exposed to the same audio environment, the unknown system's response can be determined based on the known system's response.

In order to determine the frequency response of a particular system, the Power Spectral Density (PSD) of the saved recording is taken. The PSD plots acoustic power (dB) against a specific frequency spectrum (Hz). By comparing two systems with the same inputs, any differences in the PSDs is a difference in the systems.

In this case, a system is defined as anything that takes an input signal, operates on it, and produces an output signal.

$x(t) \rightarrow \text{System } h(t) \rightarrow y(t)$

The transfer response for the system is h(t) and the output defined in the time domain is

$$y(t) = x(t) * h(t) \qquad (1)$$

where * is the convolution operator[1] and is defined as

$$(f * g)(t) = \int f(\tau)g(t-\tau)d\tau. \qquad (2)$$

Convolution is used to mathematically describe the interaction between two signals and how they produce a third, different signal. Typically the third signal created by using the convolution operator is a modified version of one of the original signals.

The convolution theorem can be used to simplify equation (1) by converting it from the time to the frequency domain. The convolution theorem states that the Fourier transform of two convoluted signals is equal to the Fourier transform of each signal multiplied together,

$$F[f * g] = F[f] \bullet F[g] \qquad (3)$$

where F denotes a Fourier transform operator[2]. Fourier transforming x(t) and h(t) converts the signals from the time domain, (t), into the frequency domain, (ω). Thus, equation (1) is simplified from a convolution of signals at a given time to a product of signals at a given frequency:

$$y(\omega) = x(\omega) \bullet h(\omega). \qquad (4)$$

The raw data being fed into the system from the respective microphones is in volts and needs to be converted into a decibel scale, a measure of loudness, using the identity

$$L_{dB} = 10 \log_{10}(A_1^2/A_0^2), \qquad (5)$$

where L_{dB} is the ratio of A_1 to A_0 in decibels, A_1 is the measured amplitude, and A_0 is the reference amplitude[3]. In this case, the reference amplitude is 1 V. Equation (5) can be applied to the x(ω) and h(ω) signals from equation (4), along with the general multiplicative logarithm identity,

$$\log(x \bullet y) = \log(x) + \log(y), \qquad (6)$$

to simplify equation (1) into vector addition, where x(ω) and h(ω) are signals in the frequency domain with the units of decibels. The frequency equation is now represented by

$$\log(y(\omega)) = \log(x(\omega)) + \log(h(\omega)). \qquad (7)$$

The general form to calculate the frequency response for the unknown system based on the known system's input can be written as a system of equations:

$$\log(y_k(\omega)) = \log(x(\omega)) + \log(h_k(\omega)) \quad (8)$$
$$\log(y_u(\omega)) = \log(x(\omega)) + \log(h_u(\omega)) \quad (9)$$

where

$y_k(\omega)$ is the measured PSD of the known system,
$y_u(\omega)$ is the measured PSD of the unknown system,
$h_k(\omega)$ is the transfer response of the known system,
$h_u(\omega)$ is the transfer response of the unknown system,
and $x(\omega)$ is the sound pressure level in the anechoic test chamber.

The unknown transfer response, $h_u(\omega)$, can be determined by solving equation (8) for $x(\omega)$ and substituting into equation (9), resulting in

$$\log(y_u(\omega)) = \log(y_k(\omega)) - \log(h_k(\omega)) + \log(h_u(\omega)). \quad (10)$$

We can now determine $h_u(\omega)$ as a function of the two measurements, $y_k(\omega)$ and $y_u(\omega)$, and the given transfer response, $h_k(\omega)$:

$$\log(h_u(\omega)) = \log(y_u(\omega)) - [\log(y_k(\omega)) - \log(h_k(\omega))] \quad (11)$$

Expression (11) can be simplified down into its final form,

$$h_u(\omega) = [y_u(\omega) \bullet h_k(\omega)] / y_k(\omega) \quad (12)$$

using logarithmic identities and then taking the inverse logarithm of both sides of the equation.

The transfer response of the unknown system is determined by subtracting the test chamber sound pressure level from the measured system response. The sound pressure level in the test chamber is determined using the measured response of the known system, from here on referred to as the reference microphone, and the published frequency response taken from the manufacturer's specifications. The final accuracy of the transfer response for the unknown system, from here out referred to as the System Under Test (SUT), depends on the accuracy of the reference microphone, so only high quality laboratory microphone systems should be used.

II. EXPERIMENTAL PROCEDURE

In order to acoustically isolate the systems, all measurements were made using a Brüel & Kjær Type 4232 Portable Anechoic Chamber[4]. The reference microphone and the microphone for the SUT were placed inside the chamber. Using the chamber's internal loudspeaker, a frequency sweep, or 'chirp' was played. The chirp spanned from 500 to 5000 Hz and lasted 5 seconds. This was repeated with one second of silence between each sweep. The chirp was generated using Sound Forge, a computer software program designed for digital audio editing. The digital signal was then converted into analog via computer sound card and output to a loudspeaker. A sound card with a flat frequency response is necessary to avoid distorting the signal and changing the results. Based on this requirement, an Echo Gina 3G card[5] was used.

The reference microphone was an ACO Pacific Type 7012 microphone[6], paired with a matched preamplifier, which was already calibrated. Its manufacturer-provided frequency response is extremely flat from 10 to 10000 Hz (Fig 1), making it a good reference choice. The output of the reference microphone was recorded using a Tektronix TDS6804B Digital Oscilloscope. The SUT audio was recorded using the equipment and software provided with the system. Both recordings were then transferred into MATLAB for post-processing (Fig 2).

This test method was used to compare three audio systems. The first system tested was another microphone with a known frequency response (Tibbetts Model 251-01). If the methodology was accurate, the transfer response generated would be the same as (or within a reasonable error to) the theoretical frequency response published for the microphone.

The second SUT was a digital proprietary wireless microphone. This system used a microphone connected to a wireless radio frequency (RF) transmitter. The data from the microphone was digitized and transmitted to a receiver attached to a computer. Software included with the system captured the signal and recorded the audio.

The final system tested was a wireless microphone system which communicated using a cellular network, GSM (Global System for Mobile communications). The system used Bluetooth to transmit the audio signal from the microphone to a base unit. The base unit then sent the signal over the cellular network to a computer in order to be recorded.

Both the RF and the GSM cellular systems provided true unknowns to test this method because there is no frequency response data published for either system, and neither provides a real-time audio stream for measurements using conventional audio test equipment.

Audio recordings of the chirp were made simultaneously in the chamber with the SUT and reference microphone and saved as .wav files. The data taken from the RF and GSM systems was done using the best-case scenario—short range transmission with no interfering obstacles. A MATLAB program was written to process the data and to calculate the frequency responses of the systems.

In MATLAB the two time domain data sets from each system were cut to the same size, approximately six seconds—the length of the full sweep including silence in the beginning and the end. Then the PSD via periodogram was taken for each data set.

The normalized PSD for the reference microphone, $y_k(\omega)$, was then subtracted from the published frequency response, $h_k(\omega)$. This provides a measure of the audio signal that is present in the chamber at the input of the SUT. With the input known, the SUT transfer response could then be determined by subtracting the SUT measured response from the input, $x(\omega)$.

III. RESULTS

The results from the tests were very promising. The Tibbetts microphone performed closely to its published data[7] (Fig 3). The published frequency response variance from unit to unit for the microphone was ±3 dB. The absolute difference between the frequency response measured using this technique and the published response was less than 6 dB at all frequencies.

The digital wireless microphone displayed an almost perfectly flat frequency response graph. The GSM system did not perform as well, displaying a very broken, choppy curve (Fig 4).

These results were further verified by plotting the time domain audio data received from each of the units (Fig 5). The plots for the Tibbetts microphone and the digital wireless microphone are both very smooth, where the GSM plot is very distorted.

The poor response from the GSM system is most likely due to the way that the cellular network transmits the data. The data is compressed for efficiency using adaptive codecs before sending it through the network. Unfortunately, the codecs can result in the loss of audio quality, especially if there is not a high signal to noise ratio.

IV. CONCLUSION

The experimental method described has proven to be a useful tool for evaluating and comparing audio systems. The accuracy of the technique was validated on a known system before being used to test two unknown systems. The response output from the two unknown systems was validated by comparing the frequency response to their audio output—as expected, the system that output smooth audio had a far better response than the system that produced choppy audio.

This method provides a good comparison because it does not just account for the acoustic capabilities of a recording device, but the entire system, including losses from other causes such as encoding and transmission.

The next step for this method would be to fine tune it so that it would be able to extract an averaged frequency response from any common, broad-spectrum audio source, i.e. a voice recording or white noise. This would add robustness to system testing where the operator would not need the system to be physically present, just data that it had recorded. Additionally, the frequency responses generated from this method could be used to correct recorded signals from these systems in post-processing.

V. ACKNOWLEDGEMENTS

Every paper gains from the critical comments of its author's friends and colleagues. I would like to thank in particular Mr. John Latess, without whose guidance this project would never have been realized.

VI. REFERENCES

1. Weisstein, E. W. (n.d.). Convolution. *MathWorld—A Wolfram Web Resource*. Retrieved on October 13, 2009 from http://mathworld.wolfram.com/Convolution.html

2. Weisstein, E. W. (n.d.). Convolution theorem. *MathWorld—A Wolfram Web Resource*. Retrieved on October 13, 2009 from http://mathworld.wolfram.com/ConvolutionTheorem.html

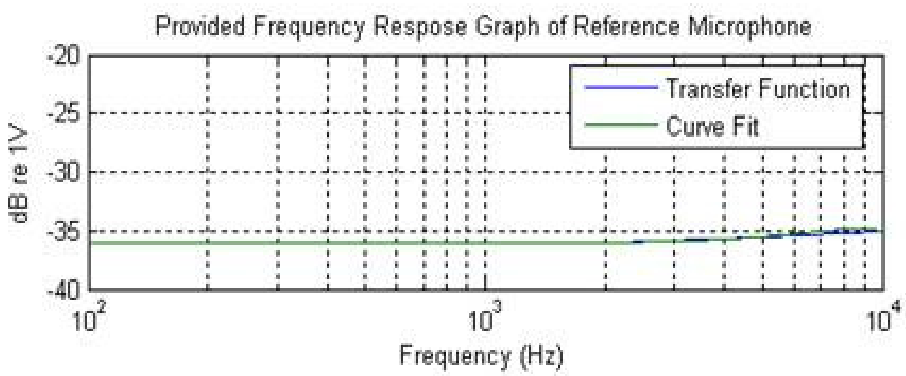
3. Everest, F. A. (2001). *Master handbook of acoustics, Fourth edition*. New York: McGraw-Hill.

4. *Product data: Anechoic Test Box – Type 4232*. (2005). Brüel & Kjær. Retrieved on November 28, 2008 from http://www.bruel.ru/UserFiles/ File.4232.pdf

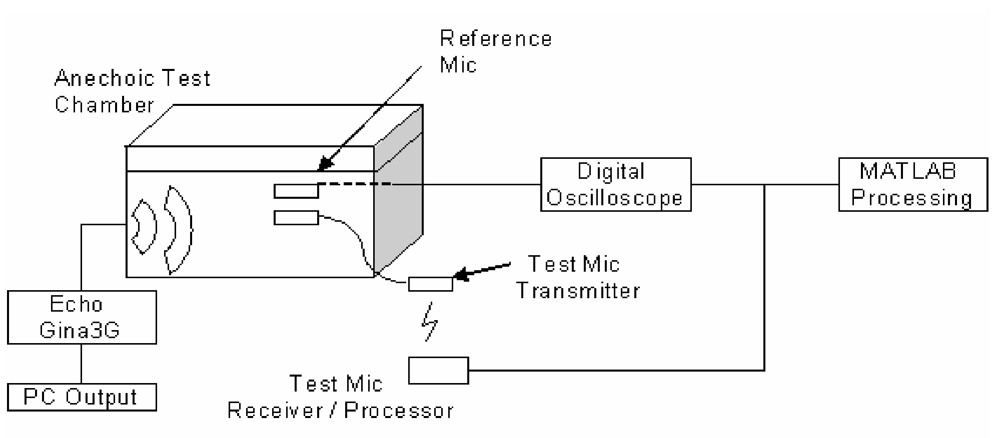
5. *Layla 3G/Gina 3G owner's manual version 1.0.1 for Windows.* (2007, April 23). Echo Audio. Retrieved on November 26, 2008 from http://echoaudio.com/Downloads/Manuals/Echo3G%20Windows%20Manual%201.0.1.pdf

6. *Precision microphones and systems for end users and OEM.* (2002). Aco Pacific. Retrieved on December 1, 2008 from http://www.acopacific. com/acopaccat.pdf

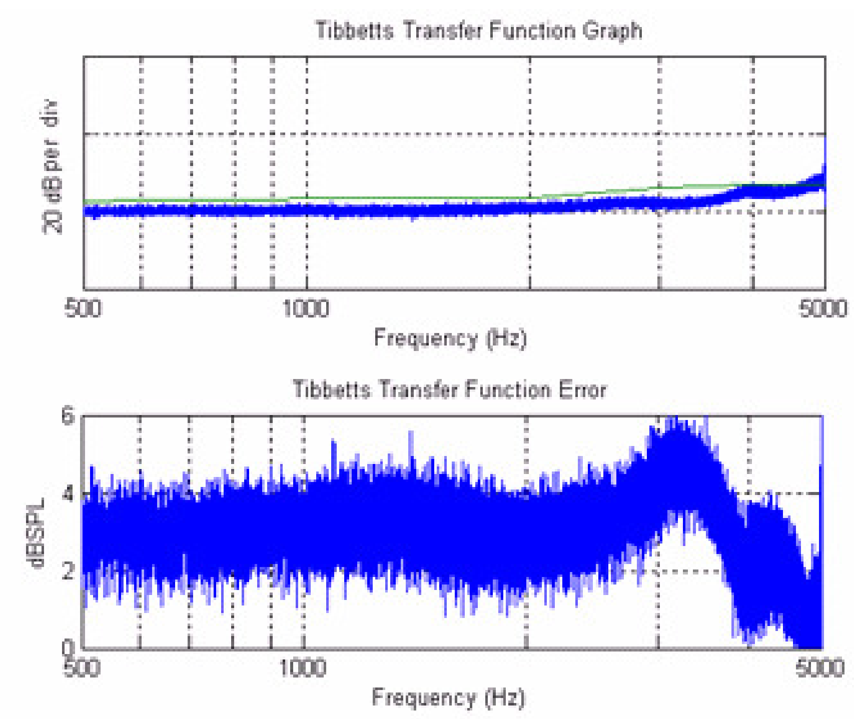
7. *Microphone products 251 Series technical data sheet*. (2004). Tibbetts Industries. Retrieved on November 28, 2008 from http://www.tibbetts industries.com/pdf/108257224.pdf



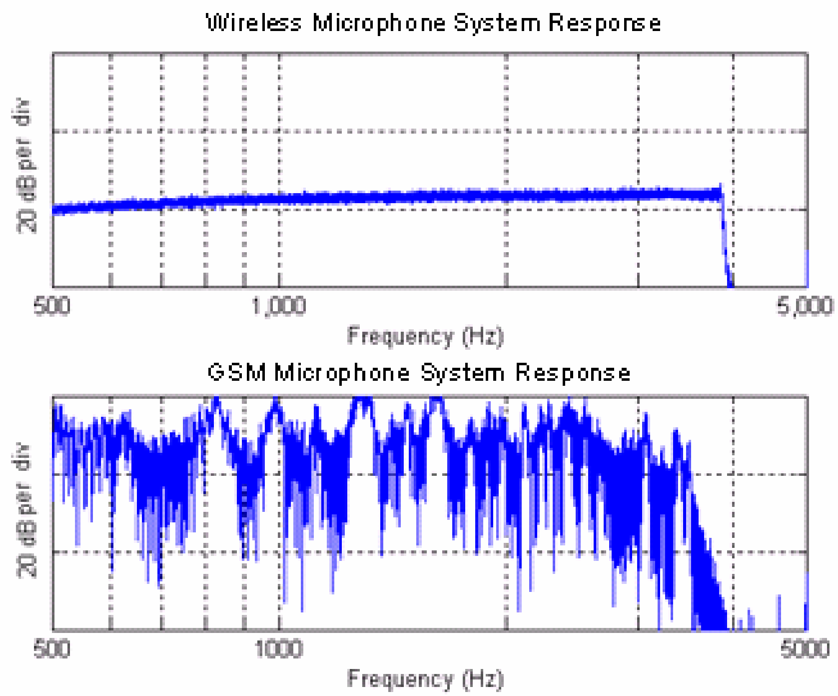
(Fig 1) Theoretical frequency response graph for the reference microphone and the MATLAB curve fit, from 100 to 10000 Hz



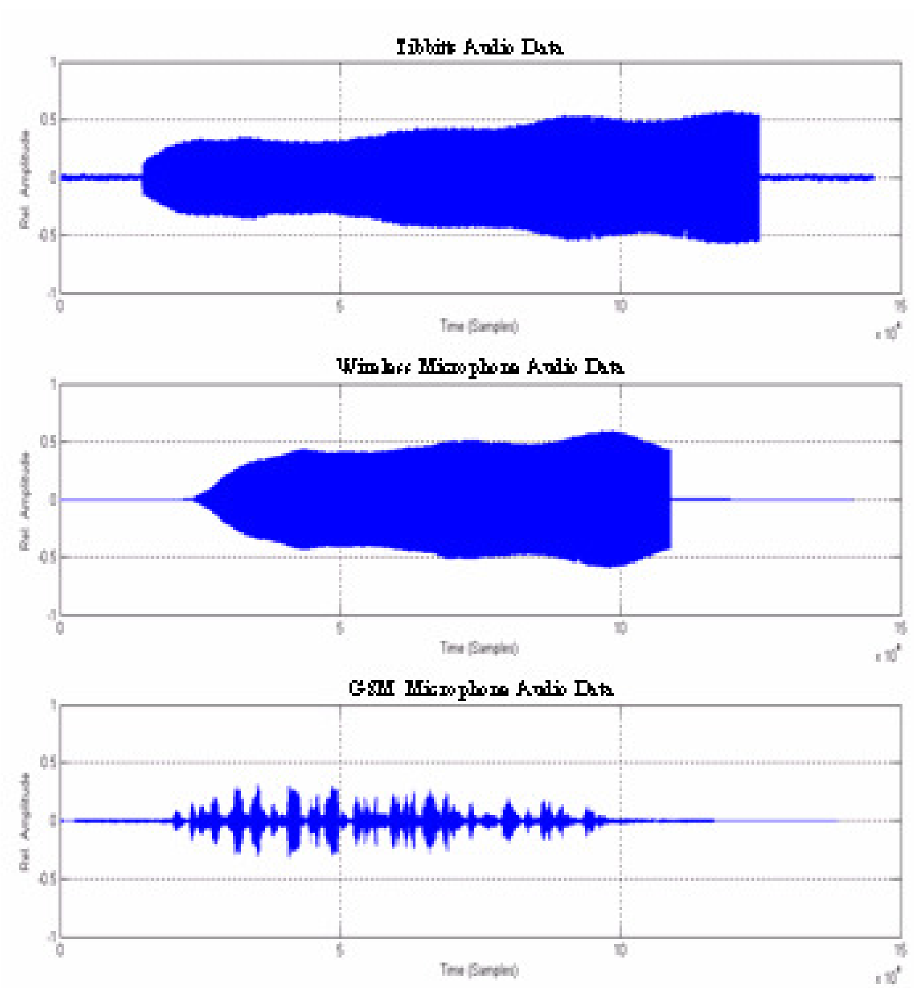
(Fig 2) Diagram of experiment set-up.



(Fig 3) Above: Graph of Tibbetts transfer function (blue) and published theoretical transfer (green). Below: Graph of the absolute error between the experimental and theoretical.



(Fig 4) Above: Frequency response for the wireless microphone. Below: Frequency response for the GSM based microphone.



(Fig 5) Top: Tibbitts audio data. Middle: Wireless microphone audio data. Bottom: GSM microphone audio data.