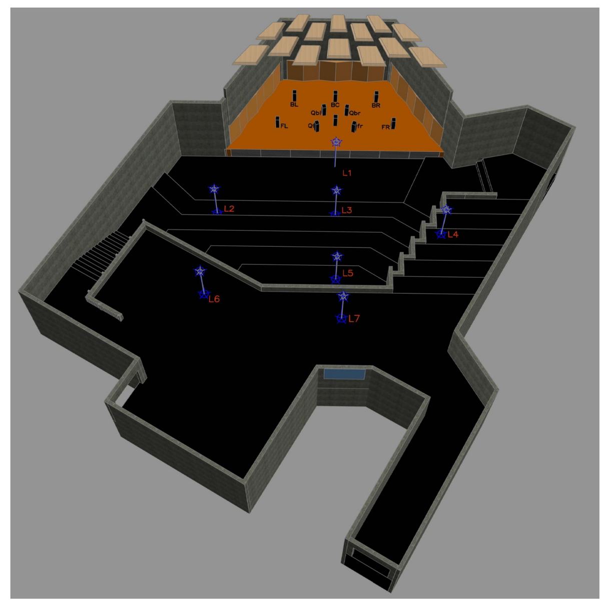
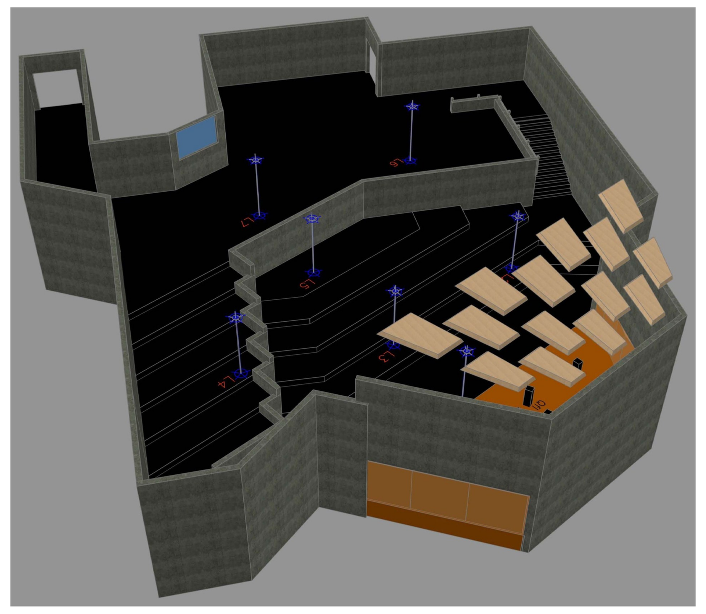
Surround Sound Impulse Response
Measurement and Application in Convolution Reverb

Madeline Carson, Hudson Giesbrecht, Tim Perry
Supervisor: Dr. Peter Driessen
University of Victoria, Victoria, BC, Canada





What is Reverb?

Clap your hands in a large room; the resulting reflections and ambience that you hear is reverb. Reverberation (reverb) is the indirect sound that reaches a listener from a source. From any source there is sound that reaches the listener in a direct path, as well as sound that reaches the listener indirectly through reflections in the acoustic space. This reflected sound reaches the listener later then the direct sound due to its longer traveled path. It has been diffused by certain surfaces, and lost energy due to propagation through air and absorption in room . The reflected/diffused sound continues to interact with its surroundings, until it has been fully absorbed.

Many musical recordings are done in a way that mostly the direct sound is recorded, with little natural reverb. Playback of such a recording sounds dry and lacks the spaciousness that listeners are accustomed to hearing. Although many filtering algorithms have been developed to simulate an acoustic space, the most accurate method is convolution reverb.

An Overview of Convolution Reverb

Convolution reverb is based on the knowledge that if the impulse response (IR) of linear time invariant (LTI) system is known, the response of the system to any input can be calculated through the discrete time convolution sum.

$$\sum_{k=-\infty}^{+\infty} x[k]h[n-k] \quad \text{where} \quad \begin{cases} x[n] \text{ is the input signal} \\ h[n] \text{ is the IR of the LTI system} \end{cases}$$

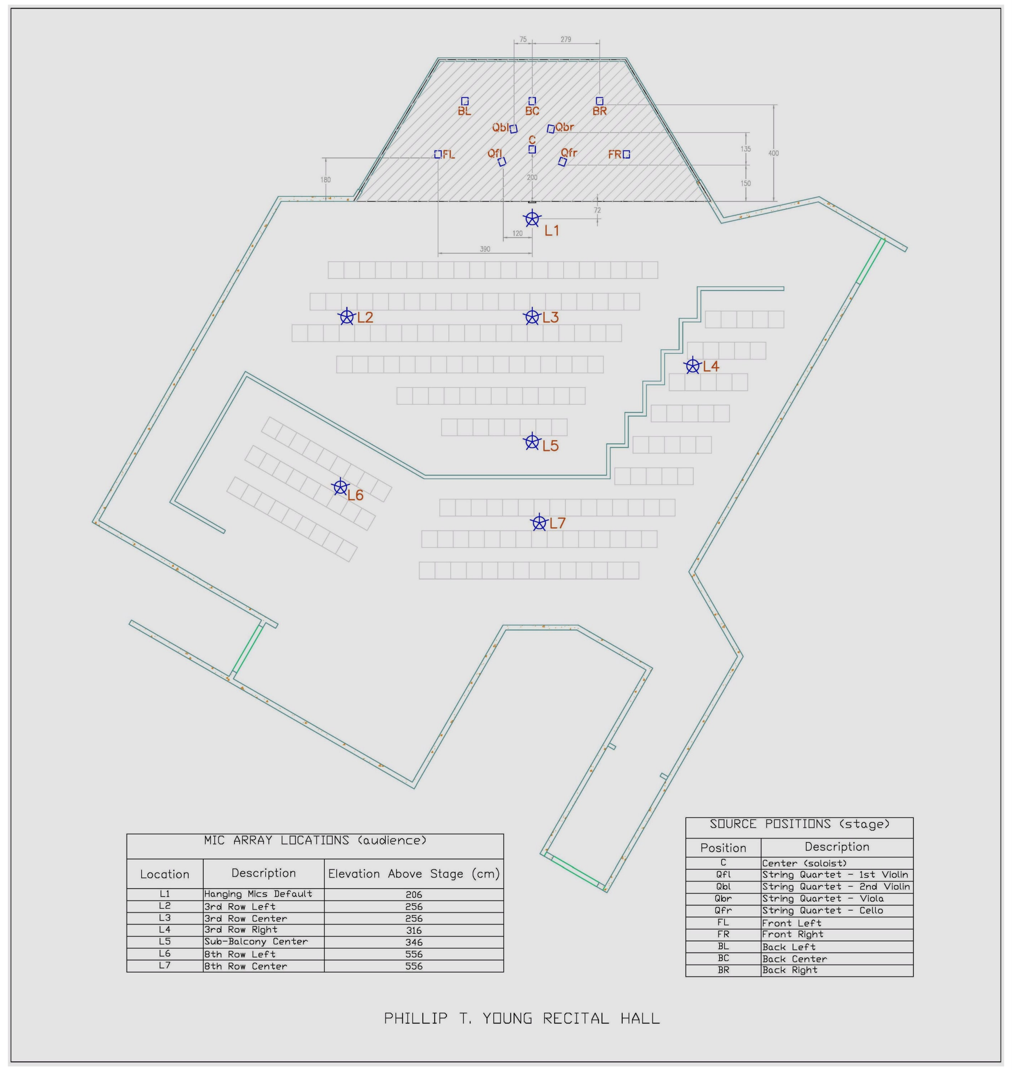
Although an acoustic space is not a perfectly LTI system it can safely be considered to be so for the purpose of the impulse response measurements. A real acoustic environment can be convincingly modeled by its impulse response. In order to add reverberation to an input signal and effectively place the source into the sampled acoustic space, the input signal is convolved with the impulse response of the room. The resulting convolution reverb sounds natural to a listener, as if they are listening to a sound that recorded in the particular room where the impulse response was measured.

Exponential Sine Sweep Method for IR Measurement

For IR measurement applications, the exponential sine sweep (ESS) provides several distinct advantages over traditional methods linear sine sweep, as well as the MLS and IRS methods [1],[2],[3],[5]. The ESS method uses an exponentially swept sinusoid for room excitation, and aperiodic deconvolution to extract the IR from the recorded room response. Unlike the linear sine sweep method, which relies on synchronous averaging of multiple linear-sweeps to increase the signal to noise ratio of an IR, the ESS method uses one long sweep. The ESS method offers:

- *Better noise rejection than the MLS method, given a signal of the same length*
- *Near-perfect separation of non-linear effects from the desired linear response*
- *For systems that may have time variance, such as those involving sound propagation in air, a long sweep utilizing no synchronous averaging is useful for avoiding artefacts in the reverberant tail and high frequency phase errors.*

Spatial Microphone Array and Recording Environment



MIC ARRAY LOCATIONS (audience)

Location	Description	Elevation Above Stage (cm)
L1	Hanging Mics Default	206
L2	3rd Row Left	256
L3	3rd Row Center	256
L4	3rd Row Right	316
L5	Sub-Balcony Center	346
L6	8th Row Left	556
L7	8th Row Center	556

SOURCE POSITIONS (stage)

Position	Description
C	Center (soloist)
Qfl	String Quartet - 1st Violin
Qbl	String Quartet - 2nd Violin
Qbr	String Quartet - Viola
Qfr	String Quartet - Cello
FL	Front Left
FR	Front Right
BL	Back Left
BC	Back Center
BR	Back Right

The surround sound impulse responses were recorded in the Phillip T. Young (PTY) Recital Hall at the University of Victoria. A location map is shown above, outlining microphone array and speaker placements. To allow for placement of the sound source and choice of listening position (to be utilized in a convolution reverb), 10 source locations were each recorded at 7 listener locations, creating 70 distinct source-listener combinations.

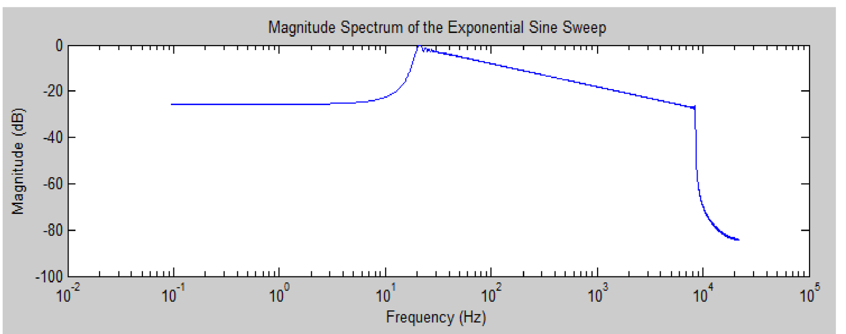
A 1st order circular array of 5 near-coincident hypercardioid microphones (*Schoeps CMC 6-U*) was used for recording the response to ESS signals in the Philip T Young Recital Hall. The circular microphone array is designed to capture surround information through amplitude and timing differences between the microphones. The 5 microphones were oriented with each capsule tangential to a circle in the horizontal plane, with equal spacing between the microphones (an angle of 72° between each pair of neighbouring capsules). Two shotgun microphones oriented vertically (one pointing up, one pointing down) complete the microphone setup, allowing 3D spatial IR recordings. These vertical microphones were not used for the IR measurements in the surround sound surround convolution reverb; however, they will be employed as an extension to this project in the very near future, and for ongoing research at the University of Victoria.

IMPULSE RESPONSE (IR) MEASUREMENT

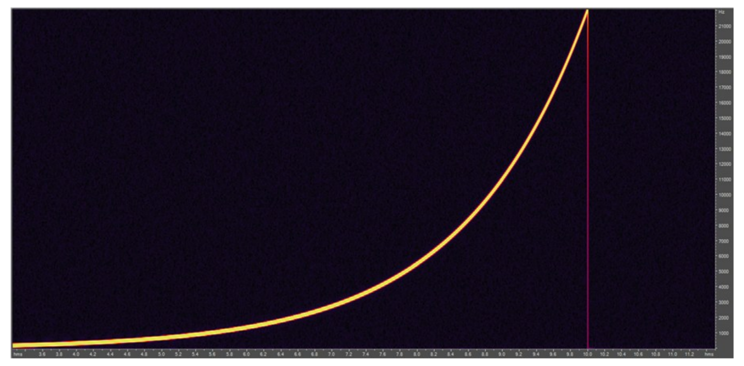
1) Exponential Sine Sweep (ESS) Signal Computation

The exponential sweep of duration T, bounded by frequencies ω_1 and ω_2, can be expressed in the continuous time domain as:

$$s(t) = \sin[\theta(t)] = \sin[K \cdot (e^{-t/L} - 1)] \qquad \text{where} \qquad K = \frac{\omega_1 T}{\ln\left(\frac{\omega_1}{\omega_2}\right)}, \quad L = \frac{T}{\ln\left(\frac{\omega_1}{\omega_2}\right)}$$

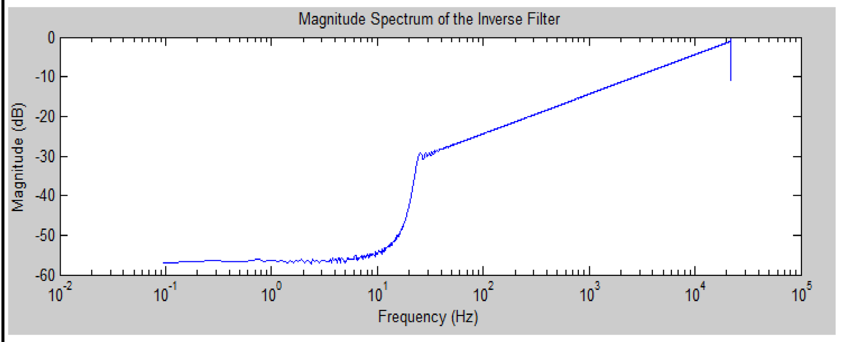


Magnitude response of the ESS signal (-3dB per octave slope in energy spectrum)



Time-Varying spectrum of ESS signal (frequency vs. time).

2) Inverse Filter Computation



Magnitude response of the inverse filter (+3dB per octave slope in energy spectrum)

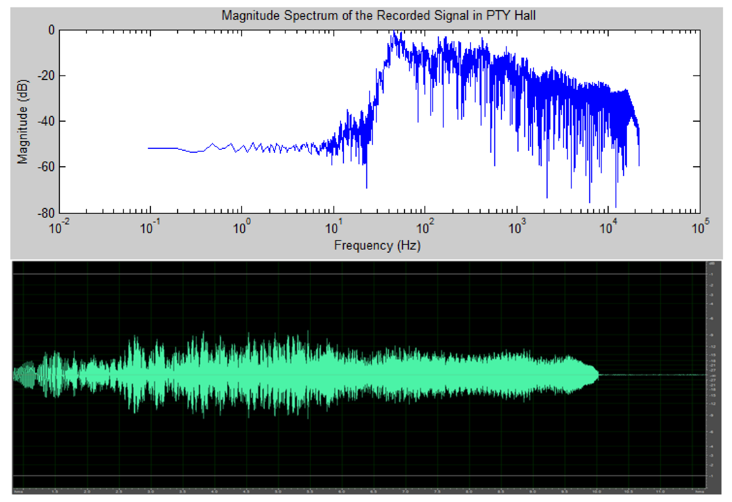
Suppose that *r(t)* denotes the room response to an exponential reference sweep, *s(t)*. To extract the IR from *r(t)*, a deconvolution is performed with the reference signal. The result of the deconvolution must achieve a flat spectrogram; however, the time-reversed ESS signal exhibits a -3dB/octave decrease in its energy spectrum. As a result, an **amplitude modulation** *m(t)* must be applied to the spectrum of this time reversed reference signal [3].

$$m(t) = \frac{A}{\omega(t)} = A\left[\frac{K}{L} \cdot e^{t/L}\right]^{-1}$$

The objective is to create an **inverse filter *f(t)*** which is to be convolved with *r(t)*, yielding the impulse response *h(t)*.

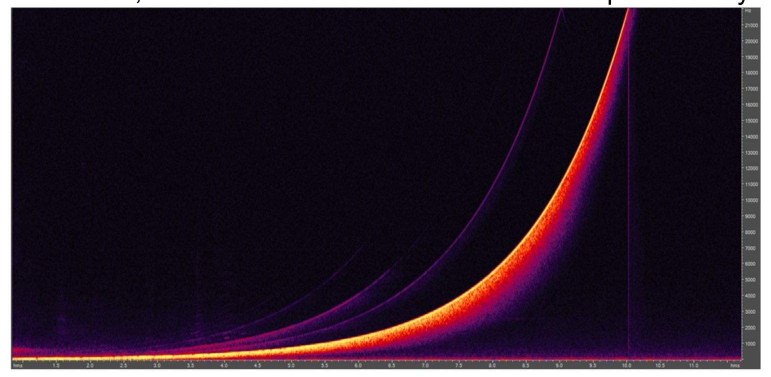
$$h(t) = r(t) * f(t)$$

3) Room Excitation with ESS Signal



Magnitude spectrum (top) and waveform (bottom) of ESS response as recorded in PTY recital hall (center microphone channel, location L1).

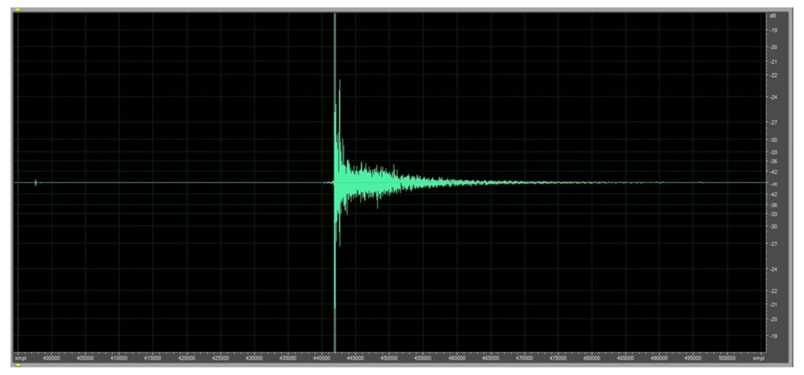
The exponential sweep is played through a speaker in an acoustic environment; the results are recorded with the microphone array.

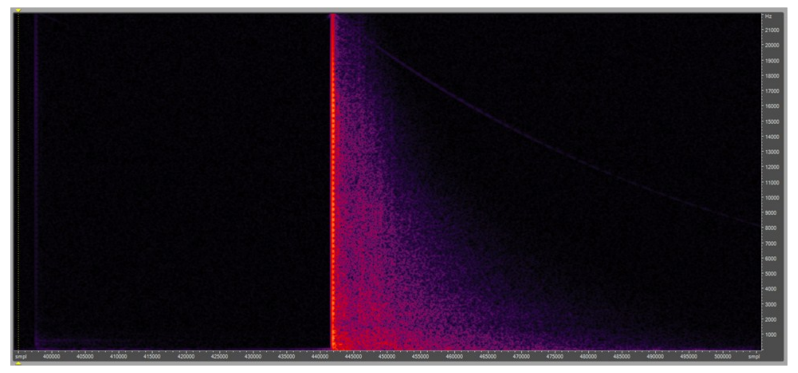


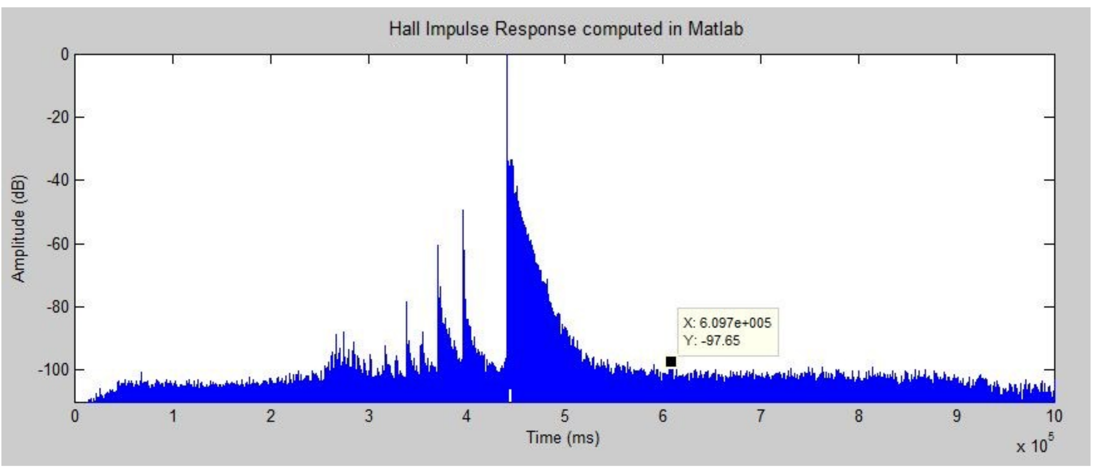
Time-Varying spectrum of ESS response in PTY Rectial Hall (frequency vs. time). Harmonic distortion can be seen in the form of curved lines above the sweep tone.

4) Deconvolution of Impulse Response

By convolving the room response *r(t)* with the inverse filter *f(t)*, the impulse response is extracted. The nonlinear response appears as a collection of vertical spikes (distortion harmonics) to the left of the linear response, occurring earlier in time. Therefore, the harmonic distortion is cleanly separated from the desired linear response.



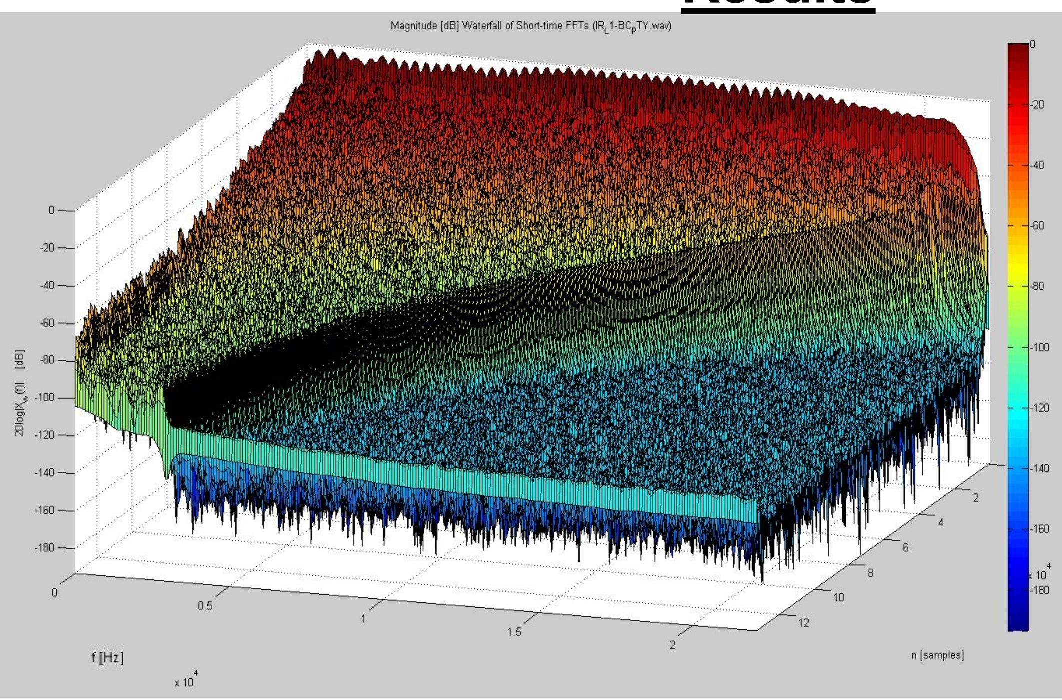




Impulse Response waveform (top left) and spectrum as a function of time (top right) as viewed in Adobe Audition. Viewing the spectrum as a function of time: an impulse clearly occurs, followed by rapid attenuation of energy in the high frequency end of the spectrum, and a more gradual attenuation in the lower octaves.

Impulse Response waveform in dB (bottom left), illustrating the dynamic range (and corresponding S/N ratio), as well as the separation of harmonic distortion from the linear response.

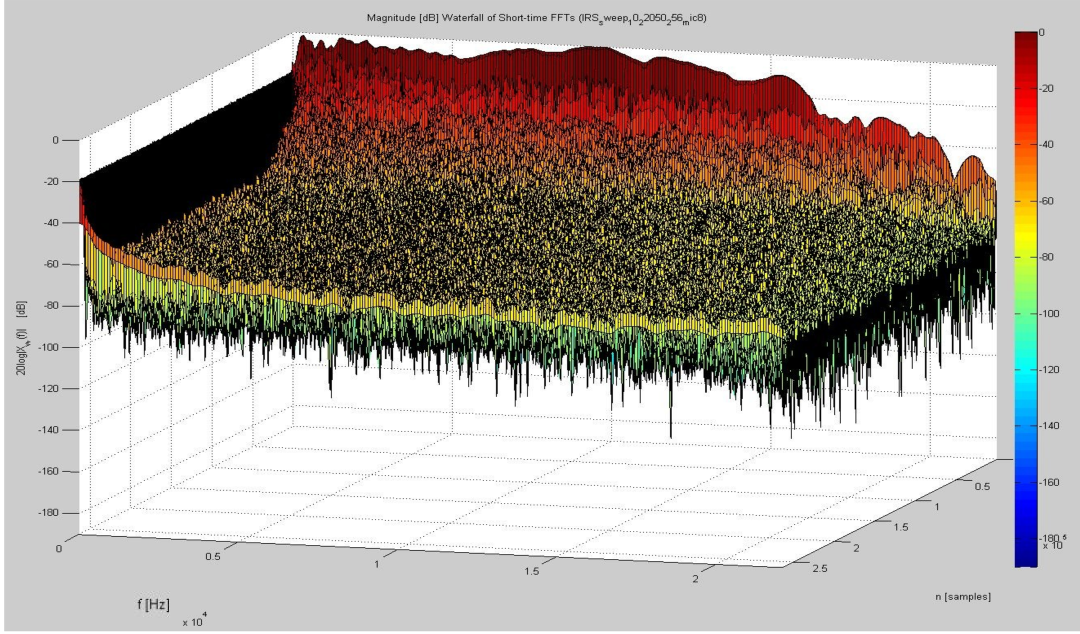
Results



Time-frequency (waterfall) representation of center microphone channel IR, array location L1, source position C.

Reverb Across the Spectrum

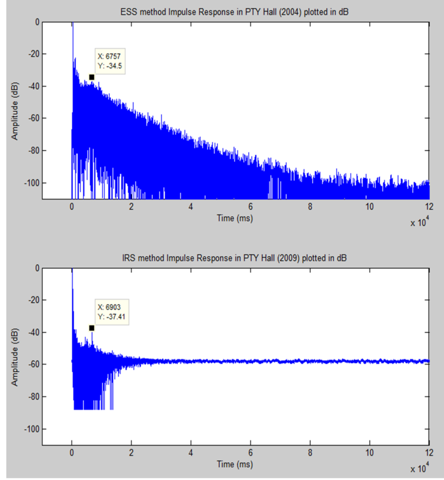
Viewing the time-frequency (waterfall) representation of a measured IR, it is clear that an impulse (energy at all frequencies) occurs, followed by rapid attenuation of energy in high end of the spectrum. Shortly after the impulse, the *first reflections* occur (at an instant called the *initial time delay gap*, an important acoustic parameter for human perception of proximity to source relative to reflective surfaces). As the reverb tail settles, it changes in timbre, losing richness as the mid and high frequencies die off in energy. The *reverb time, RT60,* which is the time it takes for the signal to decay by 60dB, can be measured at each frequency. Due to the size of the hall, resonant modes and modal decay do not have a significant impact on the IR within the spectrum that humans can hear.



Time-frequency (waterfall) representation of an IR measured in PTY Hall in 2004 using the linear sine sweep

ESS vs. Linear Sine Sweep

Prior IR measurements were been conducted in the PTY Rectial Hall in 2004 by Carola Behrens and Peter Driessen. Linear sine sweeps were used, and the IR was extracted by computing the cross-correlatiion of the recorded response with a synchronized reference sweep. The IR selected for comparison has the greatest energy in the lower octaves of the batch, and was recorded close to the source.



Comparison of early reflections with the linear sine sweep IR that was measured in 2004. The early reflections help indicate the relative microphone positions of these two measurements in PTY hall.

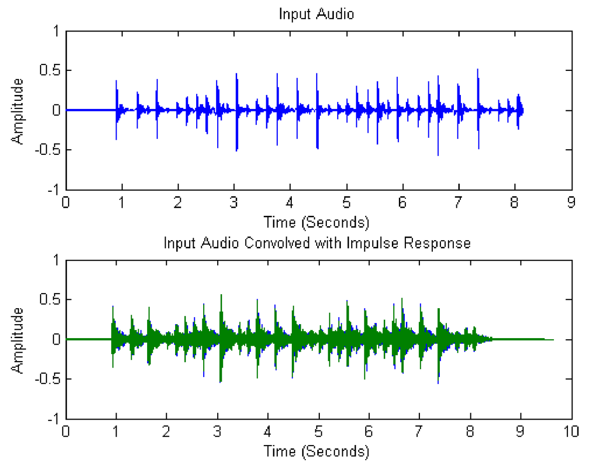
The S/N ratio is significantly superior for the ESS measured impulse response, which, depending on noise floor of the particular measurement, varies between 95dB and 100dB. The S/N ratio for the linear sine sweep measurement in contrast, was estimated to be about 56 dB. While test conditions cannot be directly compared, it is evident that the ESS method does offer a significant improvement in signal to noise ratio, and therefore a better dynamic range.

The low end response is much better captured by using an exponential sweep as the room excitation signal. The ESS signal sweeps slowly through the low end of the spectrum, where harmonics are more tightly spaced if viewed on a linear scale. As a result, the speaker has time to respond to the excitation signal in the lower octaves. Additionally, comparing the above waterfall plots, it is evident that the ESS signal has a more consistent broadband response at the instant of the impulse, leading to a more accurate measurement.

Convolution of Audio with Measured Impulse Response—*Convolution Reverb*

The final step in achieving convolution reverb is to convolve an audio signal with a measured IR. The processed audio sounds as if it was played in the room where the impulse response was measured.

In order to create surround sound convolution reverb, the dry input audio is convolved with each IR from the surround microphone array. Complex *Spatial Impulse Response* Rendering is not necessary for 5 channel playback, thanks to the intuitive nature of the microphone array used. For playback, each microphone channel is sent to a separate speaker of a 5 channel Dolby standard surround sound playback configuration. The result is a surround sound convolution reverb that uses ESS measured impulse responses, achieving a high level of realism with an excellent dynamic range.



Input audio (top) and processed audio (bottom) . Only center channel is shown

For stereo reproduction, only the impulse responses corresponding to the front center microphone and left and right forward facing microphones are used (3 microphones total).

Through superposition, the effect of multiple sources at multiple stage locations can be heard. Phantom sound sources can be virtually moved about on stage by changing the IR that the dry signal is convolved with, giving many options for reproduction of a signal with high

References

CITED REFERENCES

[1] Farina, A., "Advancements in impulse response measurements by sine sweeps", *Presented at the 122ed AES Convention, Vienna, Austria*, May 2007.

[2] Farina, A., "Simultaneous measurement of impulse response and distortion with a swept sine technique", *Presented at the 108th AES Convention, Paris, France,* 2000.

[3] Meng, Q.; Sen, D.; Wang, S.; Hayes, L., "Impulse response measurement with sine sweeps and amplitude modulation schemes," *Signal Processing and Communication Systems, 2008. ICSPCS 2008. 2nd International Conference on* , vol., no., pp.1-5, 15-17 Dec. 2008.

GENERAL REFERENCES

S. Müller, P. Massarani – "Transfer-Function Measurement with Sweeps", *JAES* Vol. 49,
Number 6 pp. 443 (2001).

G. Stan, J.J. Embrechts, D. Archambeau –"Comparison of Different Impulse Response