

Beyond MLS - Occupied hall measurement with FFT techniques

David Griesinger
Lexicon 100 Beaver Street
Waltham, MA 02154

Abstract

The study of hall acoustics continues to suffer from the difficulties of obtaining accurate measurements of halls in the presence of audience and musicians. An effective measurement technique must be fast, accurate, pleasant, and entertaining. It has become routine to measure unoccupied halls using a Maximum Length Sequence (MLS) as a stimulus. Originally chosen because MLS stimuli can be deconvolved without multiplication, there are a number of disadvantages of MLS in occupied hall measurement. Current computers provide fast and efficient deconvolution of arbitrary stimuli, which allows us to design a stimulus specifically for occupied halls. This paper will review currently available measurement methods, present MATLAB routines for generating and deconvolving effective stimuli, and show the results in occupied halls. The routines allow a complete series of binaural hall measurements with three source positions to be completed in less than 1.5 minutes while entertaining the musicians and audience. Good results can be obtained with time-variant enhancement systems, and results can be recorded with ordinary cassette recorders. A laptop computer equipped with a 16 bit sound card can be used for signal generation and analysis.

Introduction

Useful room measurements require a method which will yield an accurate impulse response which has adequate signal to noise ratio when it is analyzed in octave or third octave bands. The standard method of measuring reverberation time requires a Schroeder integral accurate over at least 40dB, and this requires a S/N in the impulse response of at least 45dB. We can fudge - matching curves by eye can give useful results with as little as 20dB S/N, but we would like to avoid this if possible. Achieving high S/N ratios at low frequencies turns out to be surprisingly hard. Bradley (4) has published a careful study of methods for optimizing the S/N ratio of the MLS measurement method, and has shown that with careful choice of stimulus length and the amount of synchronous averaging good results can be obtained in a quiet hall, and useful results in a large hall. However such results take a great deal of care - particularly in a large hall with convection and a possibly non stationary audience. One of the major problems is that even in unoccupied halls there is often a lot of low frequency noise, which means we need a stimulus for the measurement which delivers substantial energy at low frequencies.

Figure 1 shows the background noise in a large occupied concert hall, measured during a measurement session. Note that the noise increases at approximately 4dB/octave as the frequency decreases. The measurement was made with a 1/3 octave analyzer, which means that compared to white noise the hall background increases inversely with frequency at about 7dB per octave. A stimulus used for hall measurement ideally should have a spectrum which matches the spectrum of this background.

Firearms

Traditionally Acousticians have used a blank pistol as a sound source. These typically have a peak energy output in the 1kHz to 2kHz region. Below this frequency the energy falls off rapidly (~14dB/octave on a 1/3 octave analyzer), as one would expect from a spherical source of very small dimension. See figure 2. This is the opposite of the desired spectrum. To achieve a 20dB S/N at 63 Hz, the energy at 1000Hz must be >120dB! In an effort to improve S/N ratio explosive sources such as yachting cannon have been used, which make an excruciating dinn in a small space. Firearms are also difficult to transport on airlines, frightening to audiences, and not repeatable. It is also not clear that they are omnidirectional. Although the sound source is highly localized, velocity is not uniform in all directions.

Figure 3 shows the spectra from a single blank pistol shot measured in the MIT anechoic chamber from four directions simultaneously. The source was a child's cap pistol with rotary caps. To make a louder sound the powder from 5 caps was loaded into a single shot. The pistol appears to be omnidirectional except at low frequencies, where there is a rise in energy from the side. This particular pistol was open from the side, and was blocked from the front and from the top. The spectrum has more high frequency energy than the spectrum of figure 1, which used a .38 caliber blank pistol loaded with black powder. Note the decrease in energy below 1000Hz is about 10dB/octave (on a 1/3 octave analyzer this would be 13dB/octave.)

Balloons

Another traditional source has been balloons. Compared to firearms a 12" or 15" balloon is a relatively large source, and it bursts with a step function in pressure, which has greater low frequency content than an impulse. As a consequence there can be a useable amount of low frequency energy in a balloon burst even when the peak sound level in the 1000Hz band is not painfully loud. In an effort to increase the low frequency content very large balloons can be used - some consultants use balloons as large as 32" diameter. Balloons are also perceived by audience and musicians as harmless and entertaining - but only before you pop one at close range. They are still very loud. Ear protection is recommended for people within a few feet of the burst.

The major disadvantage of balloons is their extreme lack of repeatability and poor omnidirectionality. Balloons do not pop uniformly. A rip develops and spreads around the balloon in an unpredictable way. The spectrum of the burst depends strongly on direction. We pop balloons with a knife blade pointing down, inserted into the side of the balloon. On the average this method seems to produce a burst of high frequency energy in the direction opposite to the knife, and a burst of low frequency energy in the direction of the knife, with the side and up and down somewhere inbetween. We did a few tests of balloon bursts in the MIT anechoic chamber, and the results are alarming: Figure 4 and 5 show some results. Unfortunately the directionality of the source is important to measurements of intelligibility or clarity. For example, if the source emits frequencies in the 1kHz to 4kHz bands more strongly in the direction of the microphone much higher values of clarity can be obtained than if the source were non directional. We have made many measurements using both a dodecahedral loudspeaker and a balloon, and if the balloon is popped with the knife in the lateral direction the Clarity 80ms value is usually (but not always) within 2dB.

Handclaps

Handclaps are another convenient source of impulsive sounds. However they also suffer from poor low frequency content and poor omnidirectionality. The source is easily available and non threatening to hall management. For this reason recorded handclaps are very useful. Figure 6 shows a set of handclap spectra.

In an effort to make repeatable omnidirectional measurements many researchers use a special loudspeaker with 12 drivers arranged in a dodecahedron. A small dodec loudspeaker is reasonably omnidirectional up to 2kHz, but still should be carefully oriented toward the microphone if repeatable results are desired at 4kHz and above. Because the volume of the loudspeaker divided by the number of drivers is small, the fundamental resonance of a dodec is usually in the range of 200Hz to 300Hz, which means the speaker must be driven with a lot of power to produce useable energy at 100Hz. With a high powered amplifier it is possible to excite such a speaker with a step function click and obtain useful values for clarity and reverberation time, at least in bands above the fundamental resonance. This method is particularly useful when there are time varying elements in the room acoustics, but the signal to noise ratio at low frequencies is poor.

Predicting the S/N of convolution methods

S/N can be improved with convolution techniques - see 1,2,3,4,6,7,9. There is a simple argument for calculating the expected S/N from such techniques once we know the length of the stimulus and the

measurement bandwidth. In this example we will assume we are making a measurement in the octave band centered on 1000Hz. The bandwidth is about 700Hz, and the inverse bandwidth is about 1.4ms. The effect of the deconvolution process is to compresses the energy which came from the loudspeaker during the stimulus period into a time window given by the inverse bandwidth of the analysis filter. If we have an MLS stimulus of 3 seconds length, then when we deconvolve the response to the stimulus the energy received over all 3 seconds is concentrated into 1.4ms, giving a concentration factor of $10 \cdot \log_{10}(3/.0014)$ or 33dB. Only the total length of the stimulus matters - we could get the same concentration by using a 1 second stimulus and synchronously averaging 3 times. The time window of the analysis is determined by the filter. If we were using a block FFT for analysis, then the block time determines the concentration factor, at least for a single reflection.

1. Concentration factor = $10 \cdot \log_{10}(\text{stimulus time}/\text{analysis time})$

To find the expected S/N of the measurement we can measure the loudness of the stimulus at the microphone position (in the 1kHz octave band), and subtract the background sound pressure to get the stimulus/background ratio. If we are using a 1kHz one octave filter for analysis, the expected S/N for our hall measurement would be found by adding 33dB. Alternately we can calculate the expected sound pressure from the source loudspeaker at the measurement distance, and compare this pressure to the values in figure 1 to find the expected S/N.

For example, lets assume we are measuring a room with a microphone at a distance of 10meters from the source. Lets further assume that we are driving the loudspeaker with a power of 10 watts, and the speaker has an efficiency of 0.1%. We can use standard formulas to calculate the SPL at our measurement position, which is about 70dB.

At standard temperature and pressure, (from Beranek "Acoustics" section 10.65) where r is the distance in feet, W is the acoustic power of the speaker, R' is the room constant in feet², and Q is the loudspeaker directivity factor:

2. $SPL = 10 \log_{10}(W) + 130dB + 0.5dB + 10 \cdot \log_{10}((Q/4\pi \cdot r^2) + 4/R')$

3. $R' = S \cdot a / (1 - a)$

For R', S is the total area in square feet, and a is the average absorption

Lets assume our 10 watt excitation is an MLS sequence of 3 seconds length, and we will decode a single repeat (no synchronous averaging.) Further assume that the sampling rate is ~12kHz, giving 4kHz as the highest octave band. (These numbers are typical of current measurement systems.) Equation 2 gives a sound pressure ~70dB SPL.

To find the S/N in octave bands, we must consider the spectrum of the source. The unequalized MLS spectrum is white, which means half the energy is in the highest band (4kHz), 1/4 the energy is in the 2kHz band, 1/8 in 1kHz, down to 1/128 of the energy in the 63Hz band. Thus the loudness at 4kHz is 70db -3dB or 67dB, and the loudness in the 63Hz band is 70dB - 21dB or 49dB. These loudnesses are increased by the time compression ratio. In this case the factor is 33dB for the 1000Hz octave band with no window averaging. Since the analysis filter width decreases with frequency the compression factor decreases by 3dB per octave, so for 63Hz the compression factor is 21dB. Thus for this measurement we would expect "loudness" numbers of 94dB at 1kHz, going down 6dB per octave to 70dB at 63Hz. (Note that for a block FFT analysis system the block size does not decrease with frequency and the compression factor is constant.)

From figure 1, if we add 4.7dB to convert from 1/3 octave noise levels to octave band noise levels, we find the noise level in dB SPL in 1 octave bands, occupied hall:

1kHz=39dB, 200Hz=46dB, 100Hz=50dB, 50Hz=55dB If we compare this to the expected "loudness" values above we find expected S/N values of:

Expected hall S/N: 56dB at 1000Hz
 15dB at 63Hz.

As can be seen, a major problem with an MLS measurement is the white spectrum of the source, which does not match the noise spectrum of the hall. A better match occurs with the red spectrum of appendix 2. However, 56dB at 1000Hz is quite adequate in practice. This calculation indicates that a higher source power or a longer stimulus length would be useful at low frequencies, as is well known in practice.

A number of factors conspire to make the observed S/N values lower than predicted. This issue is discussed well in a recent paper by Bradley. 50dB S/N is difficult to exceed in the field. For example, using a red stimulus and a high quality electronic reverberation unit yields S/N values at 1000Hz of about 65dB even with no added noise. This limit appears to be due to the 16bit resolution in the A/D and D/A converters used in the computer sound card. The apparent S/N of a single reflection is also reduced by a measurement system which uses block FFTs. Our analysis system uses a block size of 43ms, which results in a concentration factor of 18dB, not 33dB. (In practice the reduction in S/N is less with the block system, since multiple reflections add together.) In theory we can raise the S/N by increasing the stimulus length, either by using a longer sequence or by synchronously averaging. As Bradley shows, this may work well at low frequencies, but can have difficulties at high frequencies. At high frequencies the major limits to the S/N of MLS measurements are distortion and non time invariant effects.

Disadvantages of MLS

As can be seen above, a major disadvantage of MLS is that the source spectrum is white, and the room spectrum is closer to red (increasing at 7dB per octave at low frequencies.) This problem appears to be easily overcome by equalizing the spectrum to pink or red before the loudspeaker amplifier, and then correcting this equalization when the signal is decoded. (In our simulations we use FFT deconvolution, which allows us to use the red stimulus directly to deconvolve the red response.) However when we equalize the MLS sequence we change its waveform drastically. The MLS stimulus is no longer a sequence of ones and zeros. It becomes a pseudorandom noise signal, with a crest factor of about 10dB. Amplifier power is usually rated for sine waves, which have a crest factor of 3dB, so once we equalize the signal we need about 7dB more peak amplifier power than the desired average output. We need a 100 watt amplifier to deliver a clean 20 watts average into the loudspeaker. This is possible, but heavy. It is not likely that a battery operated amplifier would be practical.

There are other problems. It is desirable in practice to record the response of the room using DAT recorders, and to decode the results later in the laboratory. If we want maximum signal to noise ratio at high frequencies this requires that the clocks of the computer generating the signal and the clocks in the DAT machines to be stable to ± 1 sample over the entire averaging period, which may be 3 to 12 seconds. This stability can be achieved by small DAT recorders - but only if the identical recorder is used both for record and playback. You have to keep track of which recorder was used to make every tape, and you must have access to that recorder at playback time. This strict requirement of time invariance causes severe problems when measuring halls which employ a time-variant reverberation enhancement. In fact, such an enhancement is invisible to MLS or any noise based measurement method. Switching on the enhancement alters only the S/N of the measurement.

Distortion

With an MLS stimulus loudspeaker distortion produces artifacts which can either raise the noise floor or produce what look like spurious reflections. We can study these distortion effects with MATLAB. Appendix 1 gives matlab scripts for generating a 2^{16} element MLS sequence. Additional scripts equalize the sequence to red spectrum. Appendix 3 is a script which upsamples the sequence a factor of three, and then applies a cubic distortion. The upsampling is essential to prevent distortion products from aliasing back into the passband. We can use the distorted sequence as a stimulus in a reverberator, and see the effect of the distortion on the measurement.

The effect of distortion on an MLS sequence is well known. Figure 7 shows the result of applying the distortion in appendix 3 to the MLS sequence in appendix 1, and then decoding with the script in appendix 4.

The result is both an increase in the noise level and spurious reflections. The positions of the spurious reflections depend on the value used as a seed. When the distorted stimulus is used to make a measurement on a reverberation unit there is considerable alteration of the impulse response. Figure 8 shows the effect of the distortion on a measurement. The increase in the noise level occurs primarily at high frequencies, with frequencies below 1000Hz showing little increase in noise, but considerable distortion of the impulse response.

The reflection-like artifacts depend on the equalization applied between the stimulus and the distortion. If we convert the MLS signal to red spectrum before the distortion is applied, and then decode with the undistorted red spectrum MLS, the distortion no longer generates the spurious reflections. However, there is a substantial increase in the noise level, and this increase is only weakly dependent on frequency. See figure 9. A similar experiment with an equalization to a pink spectrum produces noticeable artifacts, but at a much lower level than with the unequalized signal. See figure 10.

Time variance

In addition any property of the room which causes the impulse response to vary during the measurement will interfere with an MLS system. There are many such properties in an occupied hall. A full audience generates a lot of heat, as does the lighting. There is considerable air convection as a consequence. The audience is not stationary. They squirm, and the details of the reflections off individuals is important to an MLS measurement. In our experience S/N ratios greater than 40dB are difficult to achieve in a large occupied hall (with any measurement technique, including MLS with averaging.) As we would expect, the high frequencies are more affected by time variance than the low frequencies.

Systematic error in RT from averaging MLS

Worse still, the lack of time invariance creates a systematic error in the reverberation time, particularly at high frequencies. See Bradley. The reason for this is that synchronous averaging works well in the early part of the decay, where small perturbations in the boundary conditions in the room have not affected the phase of the impulse response very much. As the sound decays these boundary conditions become much more important, and the sound appears to decay quicker when several repeats of the stimulus are summed in a phase sensitive way.

A better way: - sine sweeps

Fortunately, the computational efficiency of the Hadamard transform in decoding MLS sequences is no longer very important. MLS as suggested by Schroeder is thus largely obsolete. Modern computers and Fourier transform techniques allow us to use nearly any broad band signal as a stimulus for measuring rooms. Is there an optimum choice for a stimulus for occupied hall measurements?

Swept frequency sinusoids have a long history in room measurement. The Time Delay Spectrometry (TDS) system from Heyser exploited slow linear sweeps, which are easy to decode into an impulse response. Linear sweeps decoded with FFTs were used by Berkhout, and by many Japanese researchers.

Dana Kirkegaard and Sunil Puria have suggested using a 2 second logarithmic sweep, going from high frequencies to low frequencies. They have tried this system in rooms, and have reported some very good signal to noise figures. (Presentations at Acoustical Society meetings.) Their work inspired our current work into sweep stimuli.

Distortion effects with sweep stimuli

Sweep stimuli have many advantages over MLS stimuli for room measurement, and these will be discussed below. However, broadband sweeps are not immune from the effects of distortion. For example, if we pass the sweep used by Kirkegaard and Puria through the non linear distortion in appendix 3, the waterfall of figure 11 results.

Distortion of a sweep stimulus does not manifest itself as noise, but as a spurious reflection somewhere in the deconvolved impulse response. This is because a harmonic of the input frequency appears identical to a time delayed reflection. When the sweep rate is logarithmic the apparent delay for any given distortion product is independent of frequency. The apparent time delay depends only on the harmonic order. The resulting serious distortion of the impulse response resembles the problems with a white spectrum MLS stimulus. As can be seen from figure 11 and 12, sweep stimuli are capable of excellent S/N ratios in room measurement, but it is not clear you can believe the results!

There are work-arounds. Anders Gade developed a measurement system which uses sweep stimuli which cover only two octave bands at a time. Harmonics of the stimuli are allowed to decay before the next frequency range is measured. This system is not sensitive to harmonic distortion, but it takes a long time to run.

After considerable trial and error, the best stimulus to emerge in our work is similar to the solution by Gade. It is based on sinewave sweeps, and includes a pause to allow the worst harmonics from the loudspeaker to decay. The stimulus is designed to match the properties of the loudspeaker and the room.

The drivers used in our dodecahedral loudspeaker produce about 30% harmonic distortion at 100Hz when the array is driven with 10 watts. (A better one is coming soon.) Above the resonant frequency of 300Hz distortion is quite low. Our stimulus sweeps from 70Hz to 200Hz in 2 seconds, with a time variation which produces an energy boost of 12dB per octave as frequency decreases. The 12dB per octave spectrum was chosen to match the response of the room and the loudspeaker, which has poor bass response. The stimulus then pauses for 1 second to let the distortion products decay. We then sweep from 200Hz to 11kHz exponentially in 1.25 seconds, producing a pink power spectrum over this range. This is followed by 1.75 seconds of silence. The silence at the end rounds out the stimulus to 2^{17} samples, and allows measurement of the weaker reflections from high frequencies to occur when the hall is not being excited by the stimulus. (The time limit of 1.75 seconds limits the maximum RT which can be measured at high frequencies unless the stimulus is repeated without pause.)

The stimulus is generated in the time domain by the Matlab script in Appendix 4. To find the equation for the rate of increase of phase with time we solve the differential equation for the phase increment, which yields an exponential function for a pink spectrum, an inverse square root function for a red spectrum, and an inverse cube root function for the desired 12dB per octave curve. The solutions for the initial conditions are in appendix 4. In Appendix 4 some care is taken to make sure that phase is maintained between the fade down at the end of the low frequency sweep and the fade up at the beginning of the high frequency sweep. It is not clear such coherence is needed, but I was trying to avoid a dip in the total power at the crossover. The sweep is 2^{17} samples long, which is about 6 seconds at 22050Hz sample rate.

In practice, we convert the stimulus to a .wav file using the script in appendix 5, and then play it from the computer using a standard 16 bit sound card. We record the stimulus on a DAT recorder so it repeats every 10 seconds. (Ideally the stimulus should repeat with no pause. This would allow reverberation times > 2 seconds to be accurately measured at high frequencies.) We then walk around the hall with another DAT recorder, making measurements wherever needed. In an occupied hall it is better to have numerous recorders scattered around the audience and stage. We have found that this stimulus is quite tolerant of speed variation and wow and flutter. A standard cassette recorder gives results nearly identical to a DAT. (Turning off the Dolby B circuit is recommended. See figure 14.)

The responses from the recorders are entered into the computer through the same sound card. It is usually easy to see the beginning of the second sweep in a windows sound editor, and one can edit the computer file

so this starts just after 3 seconds. Once the file is in the computer additional Matlab scripts read the .wav files, and decode the results into the impulse response. The deconvolution of a stereo impulse response takes less than one minute on a 100MHz 486 laptop. 12 mbytes or more memory is recommended. Once the impulse responses have been deconvolved they can be analyzed with Matlab routines, or they can be used as input to another analysis system. The one we wrote for our own use was designed to work in conjunction with the Hypersignal program, but we are rewriting it to work with matlab and a commercial analysis system. The current version is available from the author.

The sweep measurement system in its current form allows complete hall measurements to be made and analyzed using a minimum of equipment: A laptop computer, one or more portable DAT machines, a binaural microphone, a battery powered amplifier, and an omnidirectional speaker. All these components can be packed in small handbag and carried on an airplane. I have only once been questioned about the peculiar x-ray image of the loudspeaker. (However baggage checking the speaker and amp is recommended.)

The procedure for decoding a response in matlab is: First generate the stimulus with appendix 4. Then take the FFT of this using appendix 5. This will store the fft on disk as fstim.mat. Now read the .wav file from the response with Appendix 10, trig.m. (This assumes a stereo .wav file.) The result will be the vectors dl and dr. Now run bdcnv.m. The output will be the two impulse responses il and ir. Now run snipit.m. This will find the direct sound, and edit the impulses to 2 seconds length. The result will be the vectors il and ir, which can be written in Hypersignal format with Appendix 11, hwrite.m.

Distortion

Figure 13 shows the effect of cubic distortion on the our sweep stimulus. The distortion products from the low frequency sweep are decoded into the time period beyond 3 seconds, where they can easily be edited away and ignored. The third harmonic distortion from the high frequency log sweep decodes about 300ms before the direct sound. In practice this means that most of the distortion in an impulse response will decode to a time before time zero, where again it is easy to ignore. Distortion in reflections which are more than 300ms delayed will decode into the region of interest, but if the distortion of the loudspeaker is small enough these artifacts will be masked by the strength of the real reflections and direct sound in this time range.

Short sweeps

A disadvantage of lengthy stimuli such as MLS is that the stimulus is turned on during the measurement. You must rely on the linearity and the stability of the analysis system to extract very small amounts of reflected energy in the presence of the continuing stimulus. This is not a problem for impulsive stimuli. With these we can measure the reflections directly. It is also not a problem if the system under test varies with time - we will accurately measure a snapshot of the system. Small amounts of distortion in the measurement system do not affect the results. Some of the advantages of an impulsive measurement can be retained with a sweep measurement if sweep is short enough. If we only briefly excite the hall, we can measure the low-level late reflections without having to extract them from the continuing stimulus. Thus the sweeps we have chosen are deliberately short - a compromise between the need for high linearity and the need to deliver significant power into the room. An inherent property of a sweep stimulus is that the system under test need only be time stationary over a time period equal to the time it takes for the stimulus to sweep through the frequency range of our analysis filters. The HF section of our sweep takes about 200ms to sweep one octave. Thus the sweep system is inherently about 10 times less sensitive to time variance than a 2 second MLS. In the presence of time variance the reflections are not eliminated - they are broadened. The total energy in a time varying reflection is preserved, but the peak amplitude is lowered. An analysis system which is sensitive to total energy (as is human hearing) will give meaningful results even if the reflections change during the 200ms period. Thus by combining short sweeps with the inherent properties of a sweep system our measurement system allows useful measurement of time varying systems.

Battery power

The sweep stimulus is a constant amplitude sine wave. A small battery powered amplifier can drive the loudspeaker with 10 watts average output, and the combination of a DAT recorder and the amplifier can be carried in one hand, or dropped into a pocket. This allows a single person to hold the loudspeaker and move it to different positions on the stage during an occupied hall measurement. The peak loudness of the speaker/amplifier combination is about 93dB at 1 meter. This is loud, but not intolerable. Musicians do not complain when you carry this source around on stage, and the audience finds it entertaining. We made an occupied hall measurement with three loudspeaker positions with the orchestra on stage in less than 1 minute and 20 seconds using this system. The audience applauded.

Figures 14-16 show some of the results we have obtained with this system, using our FFT analysis system (43ms block time). In spite of the large block time signal to noise ratio of the measurements is quite good.

Conclusions

We have shown that it is possible through sweep techniques to make a hall measurement system which is well suited to the needs of occupied hall measurement. Previous systems - particularly the MLS measurement systems - take a long time to set up, involve heavy equipment and lots of cables, are loud and unpleasant sounding, and take a long time to make each measurement. Sitting through a measurement with these systems is a trial for the audience, and gives the concept of occupied measurements a bad reputation. The new system is small, light, very easy to move around, pleasant sounding, fast, and yields high quality data. Most importantly, it is perceived as quick, harmless, and entertaining by the audience. With DAT recorders the system can be accurately calibrated.

References and Bibliography:

1. H. Alrutz, M. Schroeder 'A Fast Hadamard Transform Method for the Evaluation of Measurements using Pseudorandom Test Signals' 11th ICA Paris 1983
2. J. Barish (Borish) 'Electronic Simulation of Auditoria' Center for Computer Research in Music and Acoustics - Department of Music Report # STAN-M-18 CCRMA Stanford University
3. A. J. Berkhout, D. de Vries and M. M. Boone 'A new method to acquire impulse responses in concert halls' J. Acoust. Soc. Am. 68(1), July 1980 p179
4. J. S. Bradley 'Optimizing the Decay Range in Room Acoustics Measurements using Maximum-Length-Sequence Techniques' JAES V44 #4 p266-274 April 1996
5. D. Griesinger 'Measuring room response with All Pass Deconvolution' in 11th International Conference of the Audio Engineering Society: Audio Test and Measurement, May 1992.
6. P. Kovitz 'Two maximal-length sequence devices for measuring room acoustics parameters' in 11th International Conference of the Audio Engineering Society: Audio Test and Measurement, May 1992.
7. P. Kovitz 'Extensions to the image method model of sound propagation in a room' Phd thesis Penn State University 1994
8. M. R. Schroeder 'New Method of Measuring Reverberation Time' J. Acoust. Soc. Am 1965 p409
9. M. R. Schroeder 'Integrated-impulse method measuring sound decay without using impulses' JASA 66(2) p497-500 Aug. 1979

Appendix 1 - mls.m

```
% script for mls generation in matlab
% setup for 2^16 at present
% warning - this is very slow
x = 1:16;
a = zeros(size(x));
a(1) = 1;
for i = 1:2^16
    a(17)=xor(a(1),a(12));
    a(17)=xor(a(17),a(14));
    a(17)=xor(a(17),a(15));
    a(x)=a(x+1);
```

```

    out(i)=a(1);
end

```

Appendix 2 - red.m

```

% script for a single integrator - turns white to red
% input is the vector in
% output is the vector out
% .9915 ~30Hz at 22050
% 60Hz is .9830 - these differences are important to the distortion
% the value below of .99217 is thus about 28Hz
a = [1 -0.99217]; % scale, first feedback (positive here)
b = [0 1]; % input to output of zero, drive to first delay of 1
out=filter(b,a,in);

```

Appendix 3 - hfdist2.m

```

% scrip to generate distortion of a stimulus
% it first upsamples to 3X the sample rate,
% then distortion is applied - here cubic, ~15%
% signal is filtered at 0.28*new Nyquist frequency
% and then down sampled again. Takes a while to run!
% watch for the t = n statements to see the progress

```

```

stim=stim/max(stim);
l = length(stim);
a = 1:l*3; a((1:l)*3) = stim(1:l);
s = [1,l];
clear stim; a((1:l)*3+1) = zeros(s);
a((1:l)*3+2) = zeros(s); t = 1

f = fir1(50,0.28); t = 2 % fir1 generates a 50 tap fir at frequency
% given by 0-1, where 1 is the nyquist
a = filter(f,1,a); t = 3
a(1:25)=[];
a = 3*a;
a((length(a)+1):(length(a)+50))=zeros(size(1:50));
a = a - 0.5*a.^3; %distortion applied here!
t = 4
a = filter(f,1,a); t = 5
a(1:25)=[];
b(1:l) = a(3*(1:l)); clear a;
t = 6
resp = b;
clear b n t;

```

Appendix 4 - dstimt.m

```

% MATLAB script for a stimulus for room testing.
% modified 4/16/95 for a super red lf portion
% modified 4/13/95 for a longer bass sweep
% it starts with a linear sweep from 20 to 70Hz.
% in 0.1sec, then a 12dB per octave sweep from 60 to 200Hz.
% this fades down, and lasts about 2 seconds.
% the 200Hz tone is then continued at a constant frequency but
% zero amplitude until the second sweep starts. It fades up
% at constant frequency just before the sweep, then sweeps up

% this insures phase is coherent regardless of sample rate on
% playback
% the period of zero amplitude is to let the
% harmonic distortion from the LF decay away.

% then a second sweep starts, this time 1.25 sec long log sweep from
% 200Hz (fade up) to the Nyquist frequency.

% start linear:
splice = 400;
sr = 22050; %sample rate
sl = 0.1; %length of sweep in seconds
fmin = 20; % the minimum frequency

```

```

fmax = 70; % the maximum frequency
ns = sr*sl;
% the method is to compute the jump in phase for each sample.
% first we compute omega - w here
% then we compute a vector which is the integral of w delta t
% finally we take the sine of this vector, wt
n = (1:ns);
% w is the radial frequency as a function of sample number
w = 2*pi*(fmin + (n).*(fmax-fmin)/ns)/sr; % linear sweep
clear n;

% wt is the integral of this, or the integrated phase

% now do the 12dB per octave sweep

% df/f^4 = cdt solved by f = C/t^(1/3)
% tmin and c solved below

sr = 22050;
sl = 2^16;
ss = 2.0*sr; % try 2 second

fmin = 70;
fmax = 200;

tmin = ss*((fmin)^3)/(fmax^3-fmin^3);
c = (tmin^(1/3))*fmax;

n = 1:ss;
% 4/14/95 - use super red 12dB/octave
wh = 2*pi*c./(((ss-n+tmin).^(1/3)).*sr); clear n;
n = length(wh);
c = length(w);
w(c+1:c+n) = wh(1:n); clear wh

% create the constant frequency portion:
n = 1:(sl-ns-ss+1); % the number of samples to fill out
wc = zeros(size(n)); clear n;
wc = wc+w(length(w)); % final frequency is held for wc
n = length(wc);
c = length(w);
w(c+1:c+n) = wc(1:n); clear wc;

wt = cumsum(w); clear w c n;
wfinal = wt(length(wt));
stim = sin(wt); clear wt;

% to create the join between the low and the high we must start
% the fade down and the fade up at identical parts of the
% waveform. They must be both in phase and complementary
% lets start the fade down at a positive zero crossing

% now fade down
n = ns+ss; % the sample number at the end of the first sweep
% find the first positive value after rising from negative
for x = n:n+sr/200
    if stim(x) > stim(x-1)
        if stim(x-1) <= 0
            if stim(x) > 0
                break
            end
        end
    end
end
n = x;
c = 1:splice;
stim(n+splice-c)=stim(n+splice-c).*(c/splice);

```

```

% now find a similar phase to start the fade up
n1 = sl-sr/200-splice; %start one period before the end
% find the first positive value after rising from negative
for x = n1:n1+sr/200
    if stim(x) > stim(x-1)
        if stim(x-1) <= 0
            if stim(x) > 0
                break
            end
        end
    end
end
end

n1 = x;
c = 1:splice;
stim(n1+c)=stim(n1+c).*(c/splice);

nz = n+splice:n1;
stim(nz)=zeros(size(nz));

% do the fade up
%n = 1:splice;
%stim(n+sl-splice)=stim(n+sl-splice).*(n/splice);

%save stim1 stim;    %save the low stim for decoding
%clear stim n;
clear nz n1

% generator for a log stimulus created in the time domain.
% Be sure to modify file to set stimulus length in
% seconds, as well as the sample rate.

fmin = 200; % the minimum frequency
fmax = sr/2; % the maximum frequency
ss = 1.25*sr; % 1.25 second
b = (log(fmax/fmin))/ss;
c = fmin*2*pi/(sr);
% the method is to compute the jump in phase for each sample.
% first we compute omega - w here
% then we compute a vector which is the integral of w delta t
% finally we take the sine of this vector, wt
n = (1:ss);
% w is the radial frequency as a function of sample number
w = (exp((n).*b)).*c; % log sweep - backwards
clear n;
%wt is the integral of this, or the integrated phase
wt = cumsum(w)+wfinal;
clear w;
stimh = sin(wt); clear wt;
% now pad in zeros
n = ss+1:sl;
stimh(n)=zeros(size(n));
clear n;
stim(sl+1:sl+sl) = stimh;
clear stimh wh;
save stim stim;

Appendix 5 - bstim.m
% fft the stim
n=length(stim);
n1=n/2+1; clear n;
xc=fft(stim); clear stim;
S=xc(1:n1); clear xc;
save fstim S;

Appendix 6 - dcnv.m
% to save memory dcnv is not a function but a script. It takes as

```

```

% input "stim" and "resp". It leaves the vector "r" as
% the resulting impulse response

% save tr resp; clear resp;
% lets "pink" the spectrum by whitening S
% S=(0.1913)*S.*(exp((1:nl)/(nl/8))); clear nl % works better without
% save ts S; clear S;
% 1
% load tr;

% fft resp
n=length(resp);
nl=n/2+1; clear n;
xc=fft(resp); clear resp;
B=xc(1:nl); clear xc;

% fstim is made from stim by bstim.m, and contains S
load fstim;
2
% divide stim and resp

% t = atan(imag(S)./real(S)); clear S;
% SN = 10^6*(cos(t)+i*sin(t)); clear t;
R=B./S;
R(1)=1e-20; clear S B;
3
% transform back
n=length(R);
nl=n; clear n;
I = conj(R(nl-1:-1:2));
r=length(R); l=length(I);
R(r+1:r+1)=I(1:l); clear I % Y=[R,I]; clear R l r;
r=real(ifft(R)); clear R nl;

Appendix 7 - bdcnv.m
% script to decode a stereo data pair
% uses the file stim.mat as the stimulus
% expects the data dl and dr to be in memory, and be
% an even power of two long - else it is very slow!
% data is returned as il and ir

save tmpr dr
clear dr
resp = dl; clear dl
% load stim;
dcnv
save tmpl r; clear;
load tmpr;
resp = dr; clear dr;
% load stim;
dcnv
ir = r; clear r;
load tmpl;
il = r; clear r;
[m,a]=max(abs(il(1:30000)));
a = a-44; % find the position of the maximum, and go 1ms before it
a = 0; % but lets no do this now
% snip
% hwrite

Appendix 8 - snip.m
% snips off the front of ir and il by ammount a
% and then chops at 2 seconds

% function [r2,r3] = snip(ir,il,a)
s = size(il);
s1 = s(2);
n = 1:a;

```

```

il(n) = [];
ir(n) = [];
s = size(il);
s1 = s(2);
il(44101:s1)=[];
ir(44101:s1)=[];

Appendix 9 snipit.m
% snipit - find value greater than trig and snip on it in stereo
% uses value in il as a trig for il and ir

trigv = 50;

s = size(il); % find the number of points
for n = 0:floor(s(2)/1000)
    [m,i] = max(abs(il(1000*n+1:1000*n+1000)));
    if m > trigv
        break
    end
end
n
il(1:1000*n) = [];
ir(1:1000*n) = [];

for i = 1:1000
    if abs(il(i)) > trigv;
        break
    end
end

i = i-43;
if i < 0
    i = 0;
end
a = i;
clear i n trigv
snip

Appendix 10 - trig.m
% read a wave file, trigger
% input is x = 'path and filename' with .wav extention
% output for stereo is dl, dr 2^17 length

%function d=trig(x)

trigv = 0;

file = fopen(x,'r');
% throw away the header
header = fread(file,32,'short');
length = 2^19; % try to read lots of data
d = fread(file,length,'short');
% now throw away the start before the trigger
% because matlab is so slow for for loops, this uses a max

s = size(d); % find the number of points
for n = 0:floor(s(1)/1000)
    [m,i] = max(abs(d(1000*n+1:1000*n+1000)));
    if m > trigv
        break
    end
end
n
d(1:1000*n) = [];
% now i points to a block which has a value in it > trigv
% we have to narrow it down some more
s = size(d); % find the number of points
for n = 0:floor(s(1)/100)

```

```

[m,i] = max(abs(d(100*n+1:100*n+100)));
if m > trigv
    break
end
end
d(1:100*n) = [];

% now i points to a block which has a value in it > trigv
% we have to narrow it down some more
s = size(d); % find the number of points
for n = 0:floor(s(1)/10)
    [m,i] = max(abs(d(10*n+1:10*n+100)));
    if m > trigv
        break
    end
end
d(1:10*n) = [];

if rem(d,2) != 0
    d(1) = [];
end
% there is some evidence for a 256 sample trailer in a .wav file
% so lets get rid of it
s = size(d);
d(s(1)-256:s(1))=[];

s = size(d);
if s(1) < 2^18
    n = s(1)+1:2^18;
    d(n) = zeros(size(n));
end
s = size(d);
n = 1:s(1)/2;
dl(n) = d((n-1)*2+1);
dr(n) = d(n*2);
fclose(file);
clear n s header m i d file length trigv

Appendix 11 - hwrite.m
% write a hypersignal file - here a stereo file with input data
% il and ir for the stereo channels

% data is normalized to 32767 - comment this out
% if unneeded
% where data is a vector. Modify this file
% to set header values as indicated.
% these values for 22.050k sample rate stereo data

%function count = hwrite(data)

s = size(il);
n = 1:s(2);
data(2*n-1)=il(n);
data(2*n)=ir(n);
%clear il ir;
s = size(data);
length = s(2);
m = max(abs(data));
data = 32767.*data./m;

header(1) = 32767; % max amplitude
header(2) = 2205; % framesize
header(3) = 22050; % fs_lo, sampling frequency low - up to 32767
header(4) = 0; % fftord

header(10) = fix(length/16384); % numelen_hi - number of samples * 16384
header(5) = length-16384*header(10); % numelen_lo
% number of samples up to 16384

```

```

header(6) = 0;           % overlap
header(7) = 0;           % typedata
header(8) = 1;           % (number of interleaved channels)-1 (0 for mono)
header(9) = 0;           % fs_hi - sample rate * 32768

```

```

ofile = fopen('t.tim','w');
count = fwrite(ofile,header,'ushort')
count = fwrite(ofile,data,'short')
%count = fwrite(ofile,data,'short')
fclose(ofile);

```

Appendix 12 - wave.m

```

% script to meld a new data block into a wave file
% must have a .wav file with zeros of greater length
% we will use one called 'chirp.wav' assumed mono
% new data is the vector d - assumed mono
% output is the file test.wav - mono

```

```

x = 'chirp.wav';
file = fopen(x,'r');
d = d.*30000/max(d); % not all the way up!
s = size(d)

```

```

d1 = fread(file,2^19,'short');
s1 = size(d1)
s = size(d);
if s1(1)-256<s(2)
    s(2) = s1(1)-257;
end

```

```

d1(32+1:32+s(2)) = d(1:s(2));
fclose(file)
x = 'test.wav';
file = fopen(x,'w');
count = fwrite(file,d1,'short');
fclose(file);

```

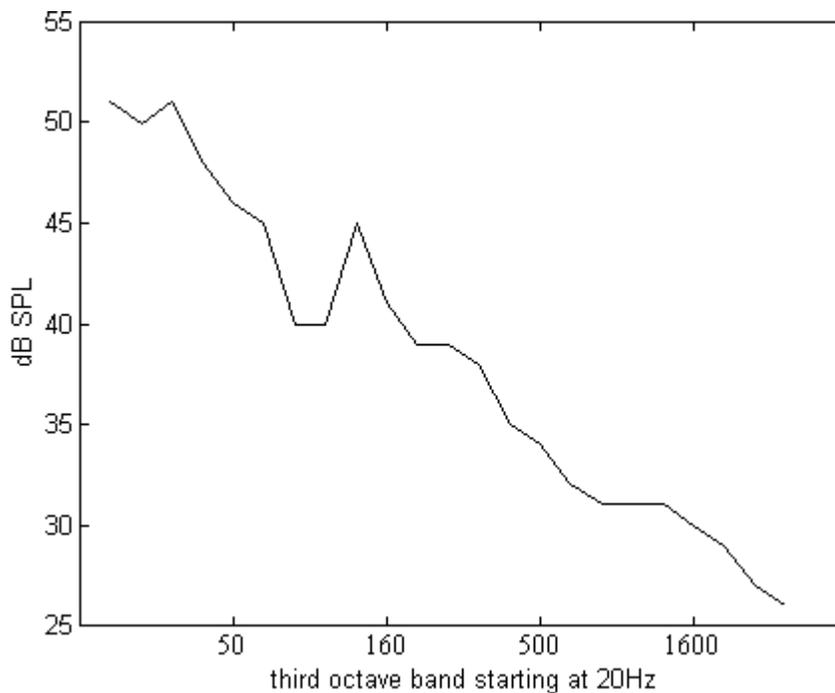


Figure 1: Hall background noise during pauses between sweep measurements, 3000 seat hall, fully occupied with audience and orchestra. Audience requested to be quiet. Microphone on-stage, concertmaster position. Note the wide range 4dB per octave slope.

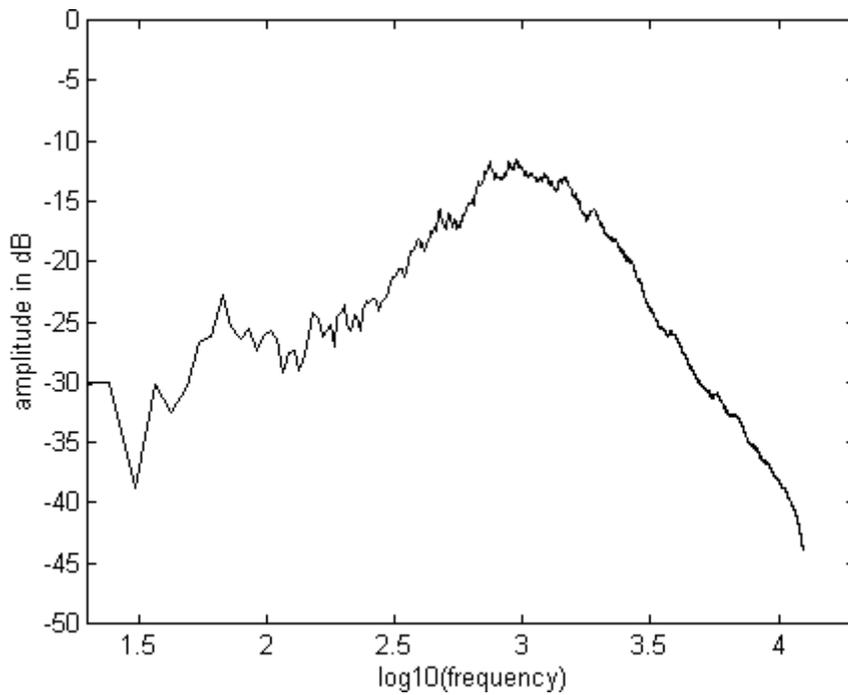


Figure 2: Spectrum of a pistol shot in a large hall. This is a smoothed FFT. To be comparable to figure 1 there should be an additional decrease of 3dB per octave as frequency drops. Note the approximate 10dB per octave decrease below 1000Hz, which corresponds to a 13dB per octave drop on a 1/3 octave scale. When we compare this to figure 1 we see the spectrum is suboptimal by ~17dB per octave below 1000Hz. The cause of the rise below 100Hz is not known.

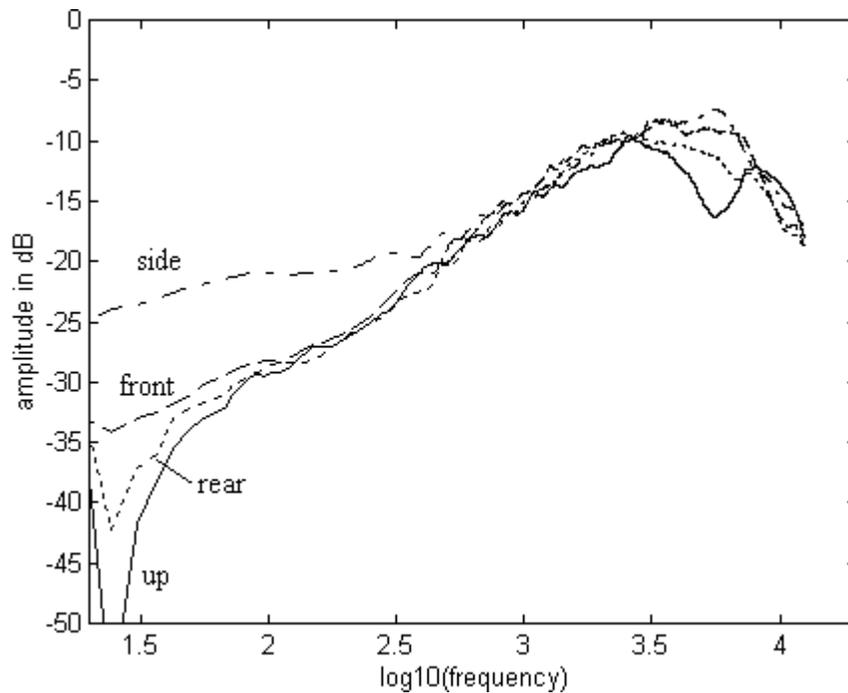


Figure 3: Anechoic spectra from a cap pistol in 4 simultaneous directions. This is a smoothed FFT. On this graph a white spectrum would appear flat. Note the apparant lack of omnidirectionality at LF.

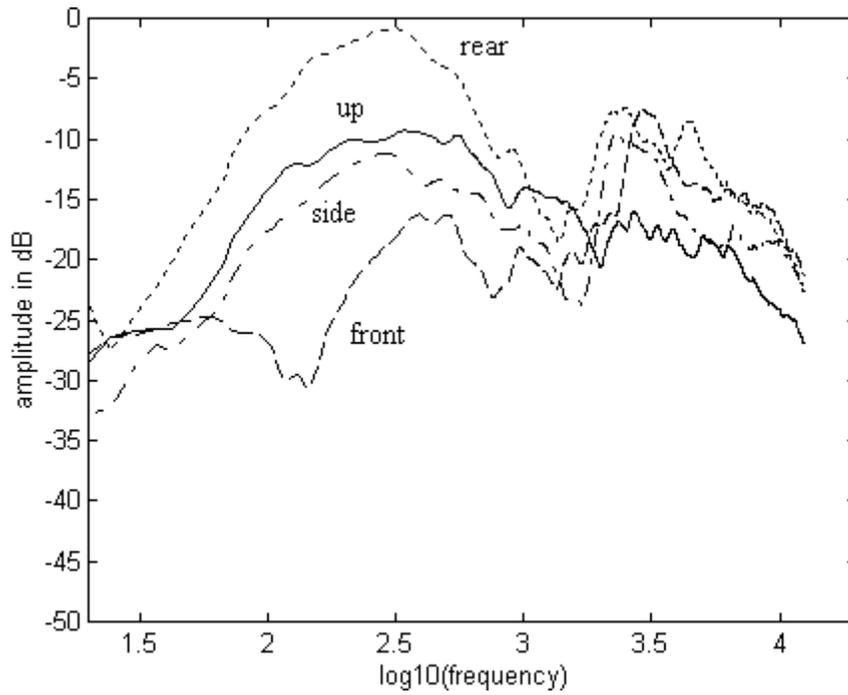


Figure 4: Spectra from a balloon in several directions. Balloon burst with a knife from the rear direction, moving toward the front direction. Smoothed FFT.

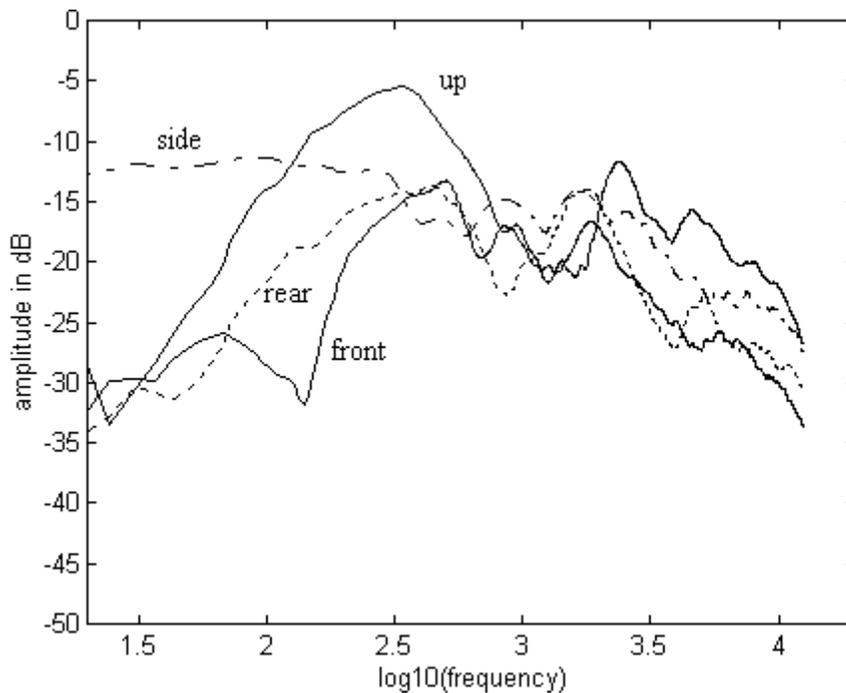


Figure 5: The same as figure 4, but a different balloon. - Note the low frequency directional behavior is quite different.

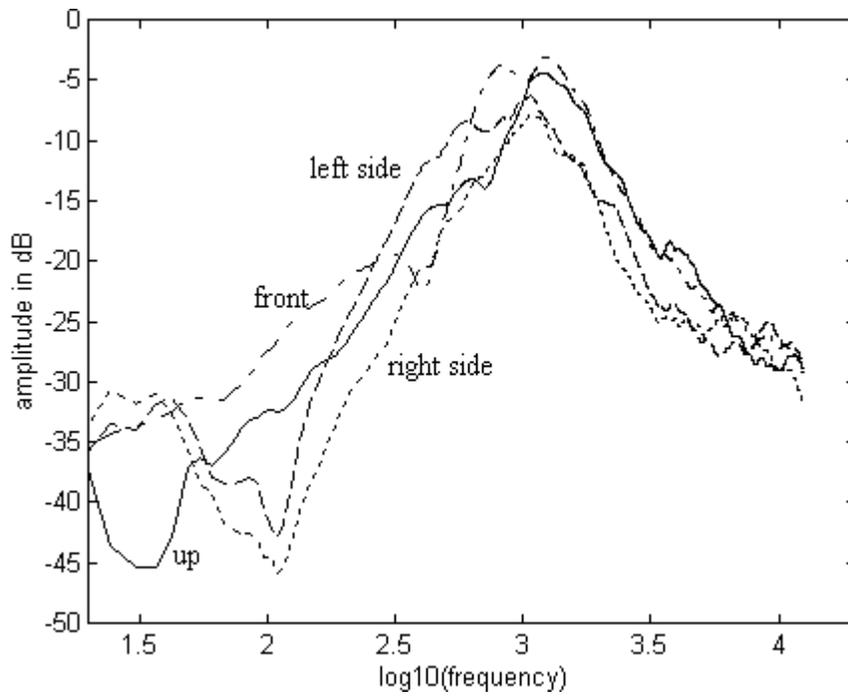


Figure 6: Spectra of a single hand clap in four directions. Smoothed FFT

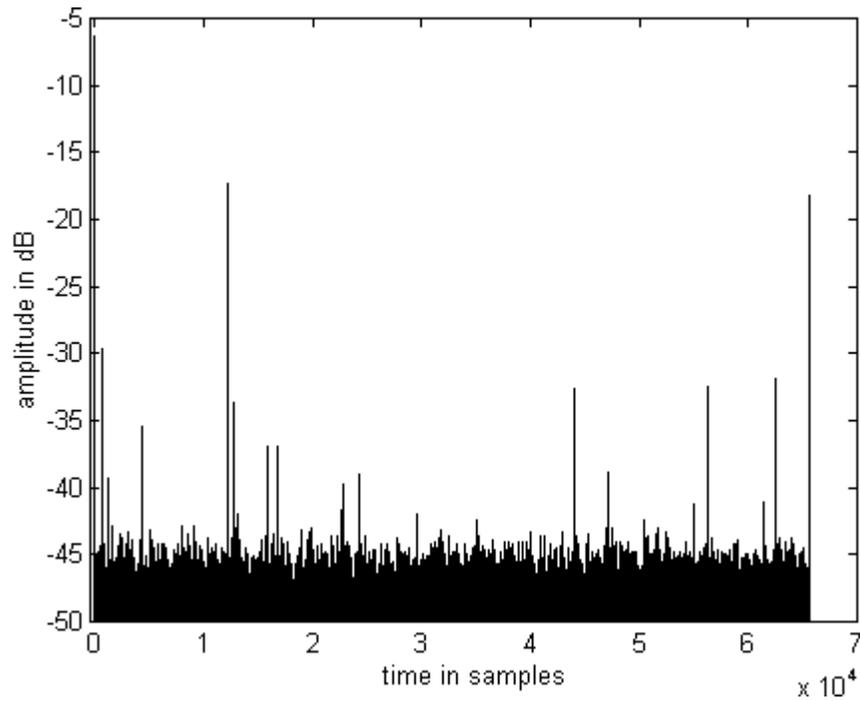


Figure 7: Deconvolution of a 2 second MLS sequence with distortion. Note the numerous spurious reflections and noise.

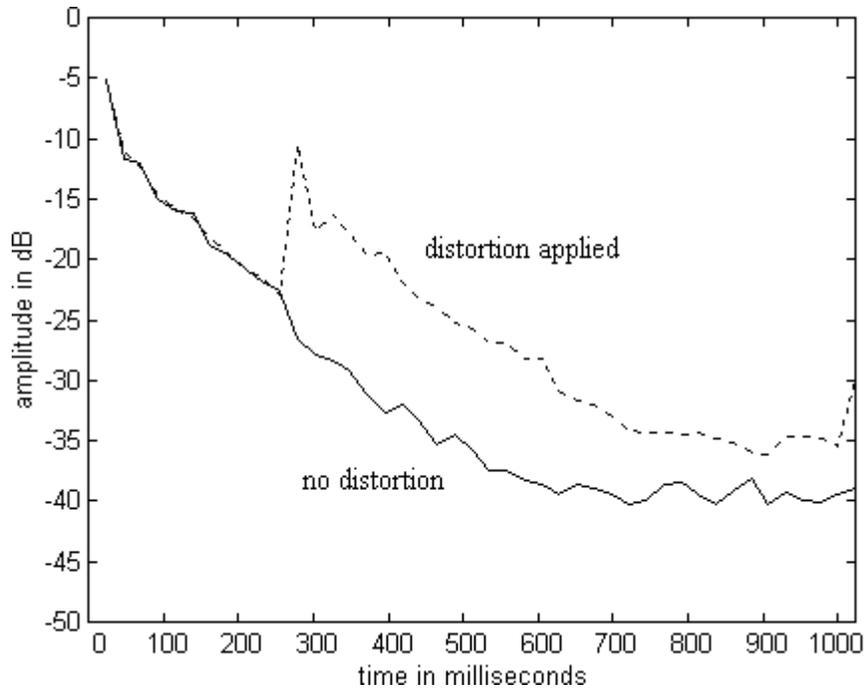


Figure 8: Distorted MLS signal used for the measurement of a reverberator, along with the same reverberator measured with undistorted MLS. 1kHz octave band, integrated in a 43ms window. Note the effect of the distortion is seen mostly in the distortion of the impulse response, not in a decrease in S/N.

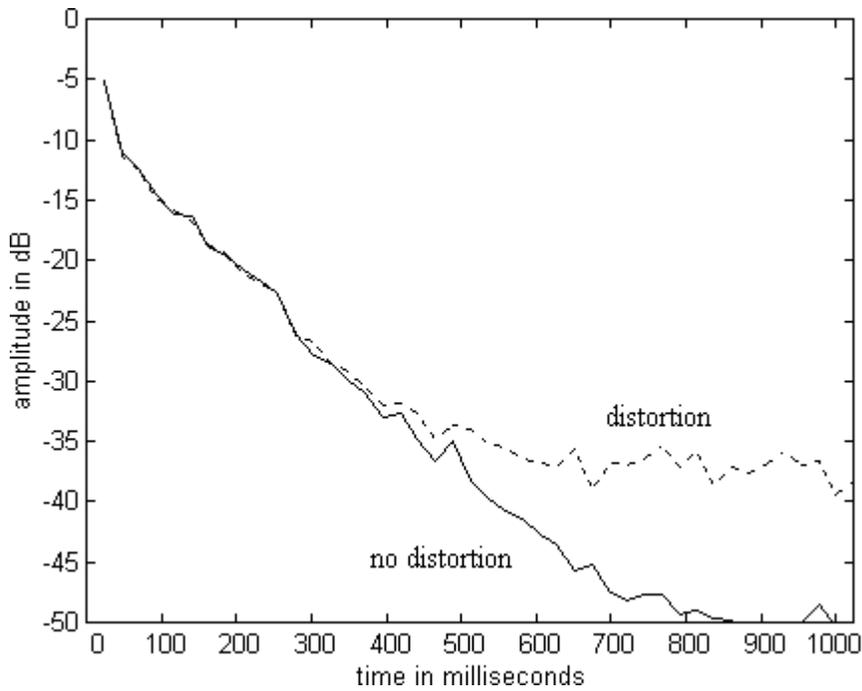


Figure 9: The same reverberator measured with distorted red spectrum MLS.

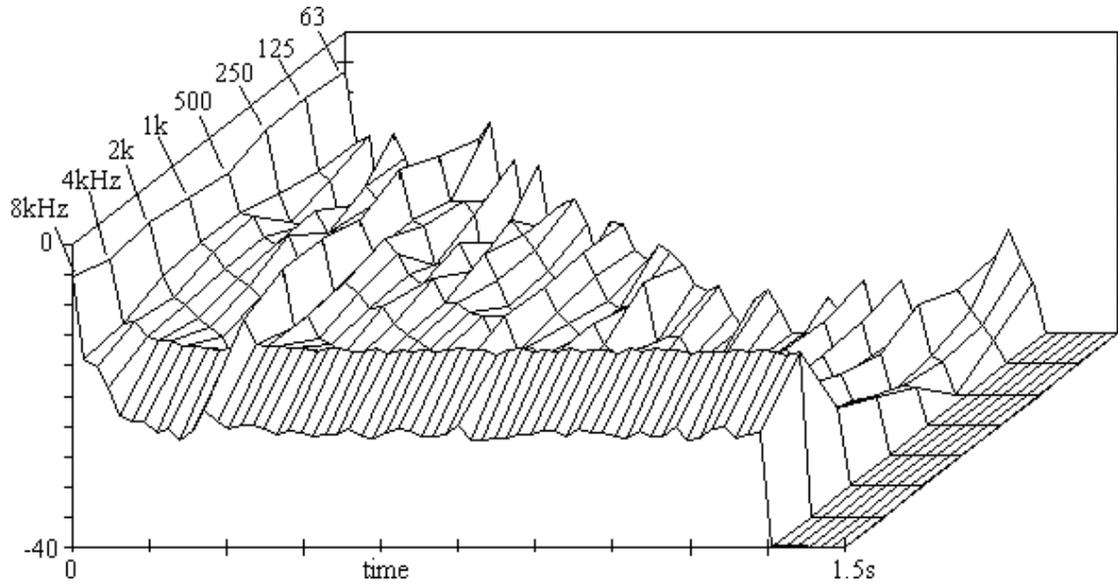


Figure 10: waterfall spectrum from distorted pink spectrum mls. Note the high increase in noise at 4kHz and above, and several reflection-like artifacts.

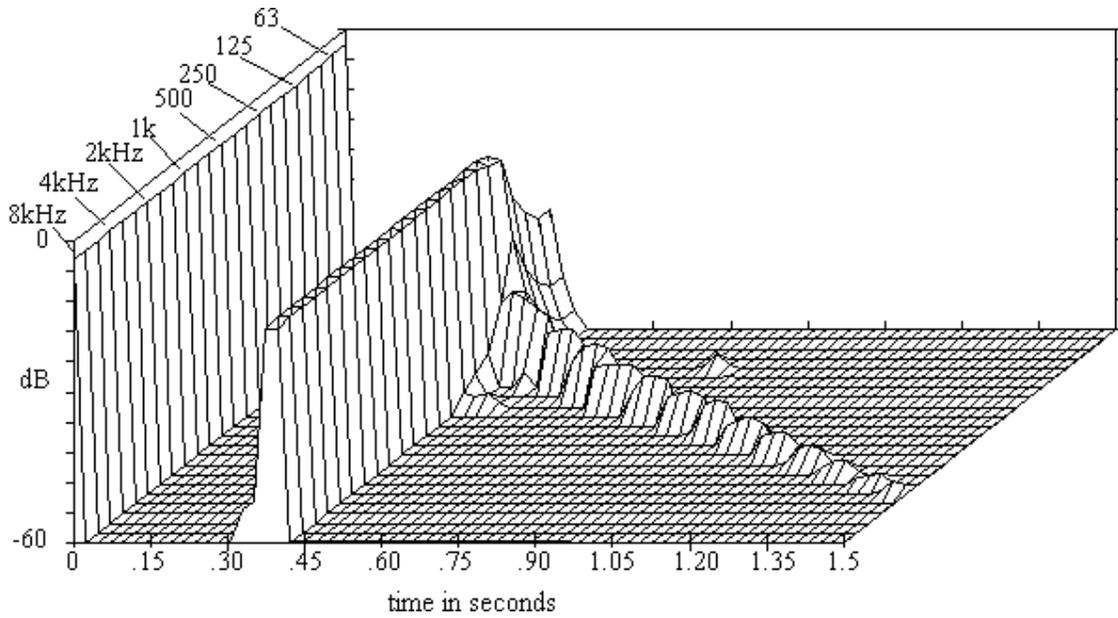


Figure 11: effect of cubic distortion on a logarithmic sweep stimulus going from high frequencies to low frequencies. Note the strong spurious reflection. Plotted with a block FFT analysis using 43ms blocks with 50% overlap.

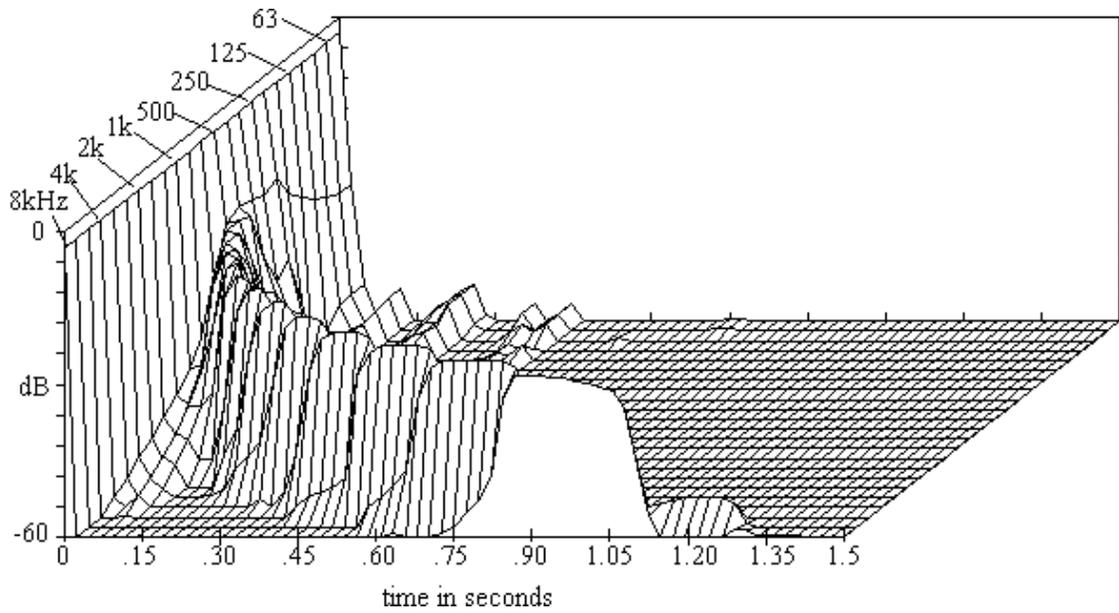


Figure 12: effect of cubic distortion on a linear sweep stimulus. Note the time delay of the spurious reflection is frequency dependent.

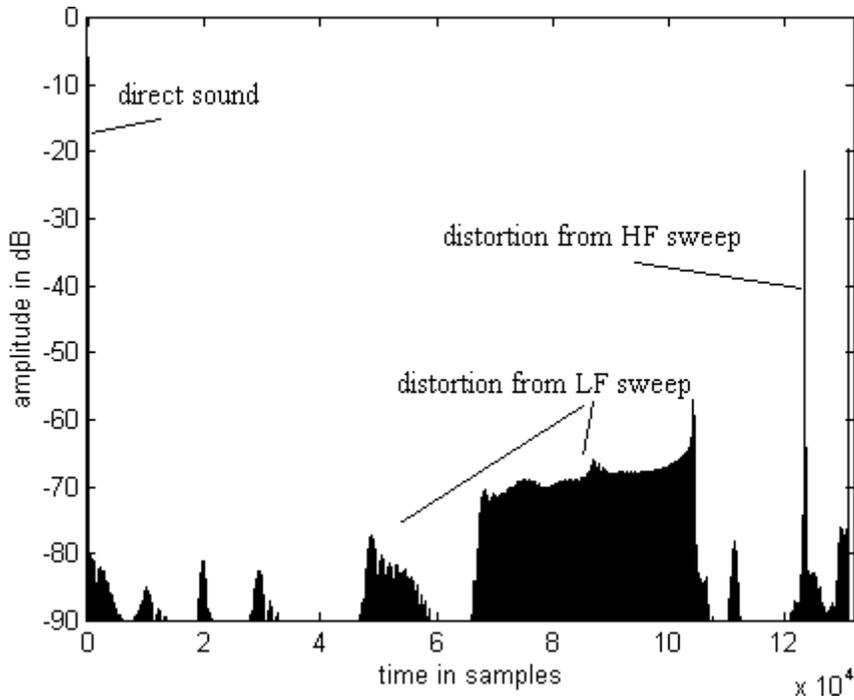


Figure 13: Effect of cubic distortion ($x-0.5*x^3$) on the two-part sweep. Note LF distortion decodes to times > 3 seconds, and the HF distortion decodes to times > 300 ms before the reflection which contains the distortion (in this case the direct sound.) Typically the HF distortion is much lower than in this example. Artifacts in the region of interest are < -70 dB, even with this high distortion.

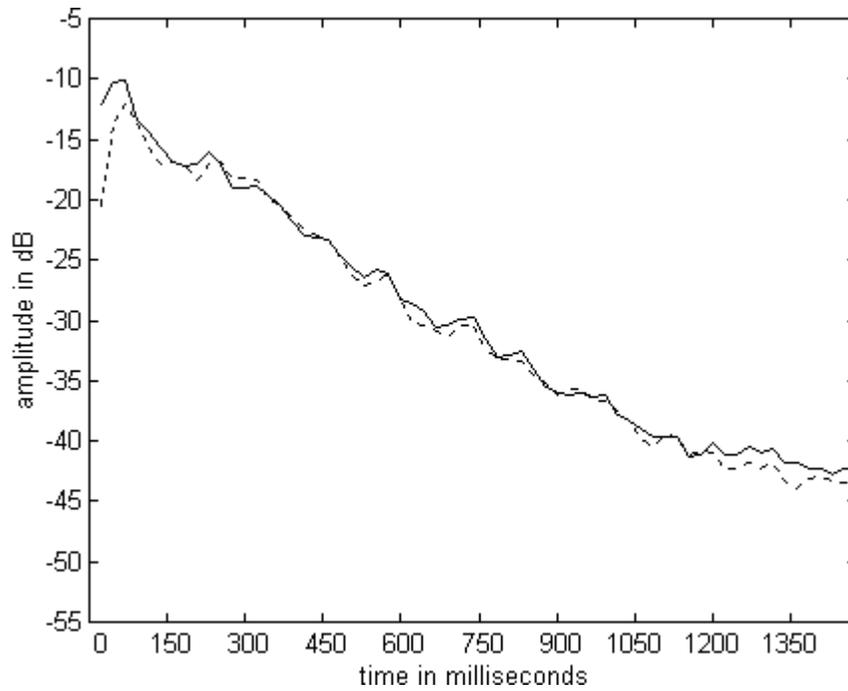


Figure 14: Medial ___ and Lateral ---- response of the hall of figure 1, occupied with the orchestra on stage. Battery operated source hand carried on stage, binaural head connected to a standard cassette recorder in second to last row of the second balcony. 1kHz octave band.

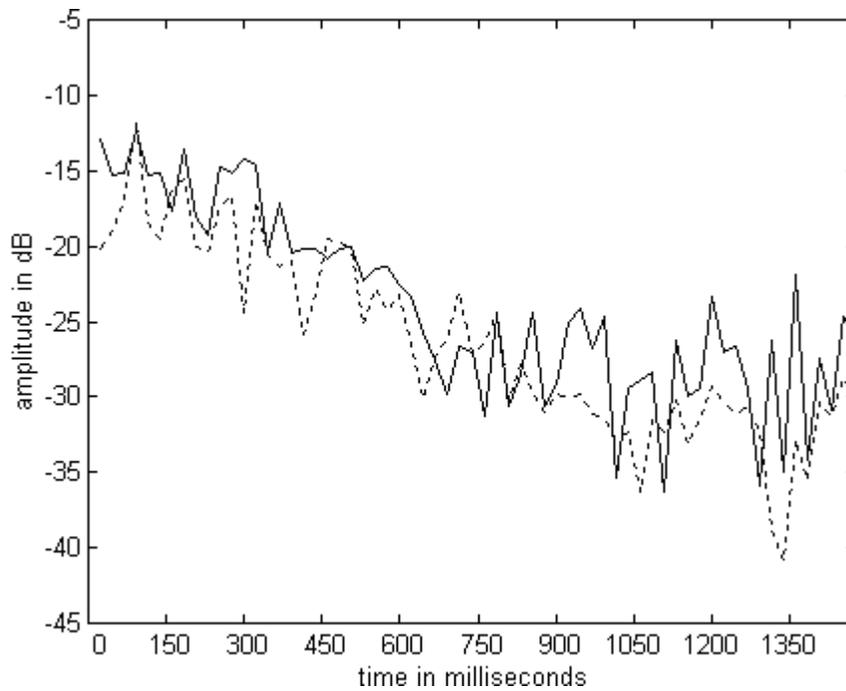


Figure 15: Same measurement as figure 10, but the 125Hz octave band

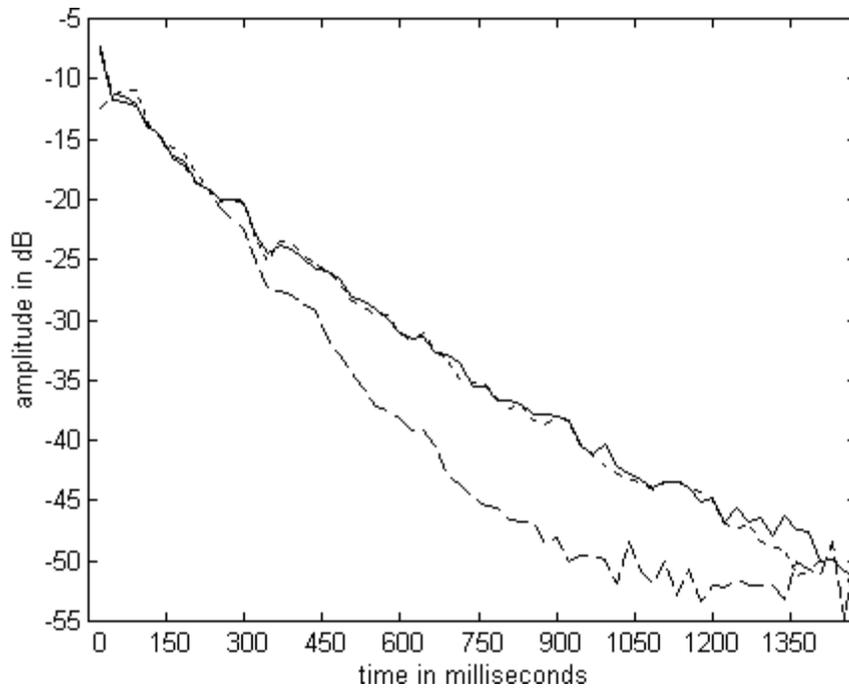


Figure 16: Berlin Staatsoper - source on stage center, microphone in the center of the stalls. ____ = medial, enhancement on. - - - - = Lateral, enhancement on. - . - . - = Medial, enhancement off. 1kHz octave band.