Audio Engineering Society

Convention Paper

Presented at the 122nd Convention
2007 May 5–8 Vienna, Austria



# Advancements in impulse response measurements by sine sweeps

Angelo Farina[1]

[1] University of Parma, Ind. Eng. Dept., Parco Area delle Scienze 181/A, 43100 PARMA, ITALY
farina@unipr.it

## ABSTRACT

Sine sweeps are employed since long time for audio and acoustics measurements, but in recent years (2000 and later) their usage became much larger, thanks to the computational capabilities of modern computers. Recent research results allow now for a further step in sine sweep measurements, particularly when dealing with the problem of measuring impulse responses, distortion and when working with systems which are neither time invariant, nor linear.

The paper presents some of these advancements, and provide experimental results aimed to quantify the improvement in signal-to-noise ratio, the suppression of pre-ringing, and the techniques employable for performing these measurements cheaply employing a standard PC and a good-quality sound interface, and currently available loudspeakers and microphones.

## 1. INTRODUCTION

At AES-Paris in 2000 a paper of the author [1] did disclose some "new" possibilities related to sine sweep measurements, triggering a wave of enthusiasm about this method. The usage of exponential sine sweep, compared with previously-employed linear sine sweeps, provided several advantages in term of signal-to-noise ratio and management of not-linear systems. Furthermore, the deconvolution technique based on convolution in time domain with the time-reversal-mirror of the test signal allowed for clean separation of the harmonic distortion products. And the release of the Aurora software package [2] made it possible to perform these measurements easily and cheaply for everyone.

In reality, nothing was really new, as other authors (Gerzon [3], Griesinger [4]) did already discover these possibilities. The fact that this approach was not successfully employed before is mainly due to the lack of computers with enough computational power and of easily-usable software tools.

In the following 6 years, many research groups and professional consultants started using sine sweeps, and a lot of papers were published (particularly remarkable were the JAES papers of Muller/Massarani [5] and of Embrechts et al. [6]). The tradeoffs of this technique were understood much better, and it was recognized the need of further perfecting the measurement technique for dealing with some problems.

- pre-ringing at low frequency before the arrival of the direct sound pulse

- sensitivity to abrupt pulsive noises during the measurement

- skewing of the measured impulse response when the playback and recording digital clocks were mismatched

- cancellation of the high frequencies in the late part of the tail when performing synchronous averaging

- time-smearing of the impulse response when amplitude-based pre-equalization of the test signal was employed

All of the problems pointed out here have been investigated, and several solutions have been proposed.

This paper presents these "refinements" to the original exponential sine sweep technique, and divulgates the results of some experiments performed for assessing the effectiveness of these techniques.

The methods analyzed include:

- post-filtering of the time-reversal-mirror inverse filter for avoiding pre-ringing

- "exact" deconvolution by division in frequency domain with regularization

- development of equalizing filters to be convolved with the test signal for pre or post equalization.

- counter-skewing of the measured impulse response when the playback and recording digital clocks are mismatched

- employing running-time cross-correlation for performing proper synchronous averaging without cancellation effects

The experiments for assessing the behavior of these "enhanced" measurement techniques were performed employing a state-of-the-art hardware system, including a multichannel sound interface, a powerful PC, and modified versions of the Aurora plugins [2]. Three rooms were chosen for the test: a small listening room equipped with a professional surround-sound monitoring system, a concert hall employing a wide-band, two-way dodechaedron loudspeaker, and the passenger's compartment of a car.

Various kinds of microphones were employed too, with the goal of assessing if the measurement of certain acoustical quantities, such as the "spatial parameters" described in ISO 3382, and namely LF, LFC and IACC, can be reliably measured with currently available top-brand microphones.

The results show that, whilst some of the proposed methods really improve substantially the sine sweep measurement method, solving the problems shown above, on the other hand the weak part of the measurement chain is still about transducers, and namely loudspeakers and microphones, which do not act always along our expectations, and which can cause severe artifacts in the measured quantities.

It is therefore concluded that any impulse response measurement chain can be used with confidence only after a set of careful preliminary tests and alignments. Without this, the results are prone to be at least suspicious, and significant errors have been found in the experimental tests. Of consequence, it appears necessary to further improve the current measurements standards, and mainly ISO 3382, for ensuring reliable and reproducible measurements employing this (and other) methods of measuring impulse responses.

## 2. QUICK REVIEW OF THE EXPONENTIAL SINE SWEEP (ESS) METHOD

This chapter is recalling the theory already presented in [1], so the reader has a consequential presentation of the “basic” method, before discussing problems and possible enhancements. The reader already knowing this method can skip directly to chapter 3.

When spatial information is neglected (i.e., both source and receivers are point and omnidirectional), the whole information about the room’s transfer function is contained in its impulse response, under the common hypothesis that the acoustics of a room is a linear, time-invariant system.

This includes both time-domain effects (echoes, discrete reflections, statistical reverberant tail) and frequency-domain effects (frequency response, frequency-dependent reverberation).

The following figure shows how a room can be seen, under these hypotheses, as a single-input, single-output "black box".

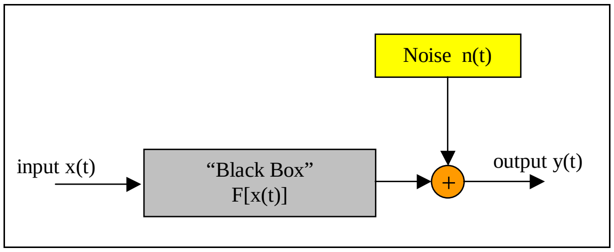


Fig. 1 – A basic input/output system

The system employed for making impulse response measurements is conceptually described in fig. 2. A computer generates a special test signal, which passes through an audio power amplifier and is emitted through a loudspeaker placed inside the theatre. The signal reverberates inside the room, and is captured by a microphone. After proper preamplification, this microphonic signal is digitalized by the same computer which was generating the test signal.

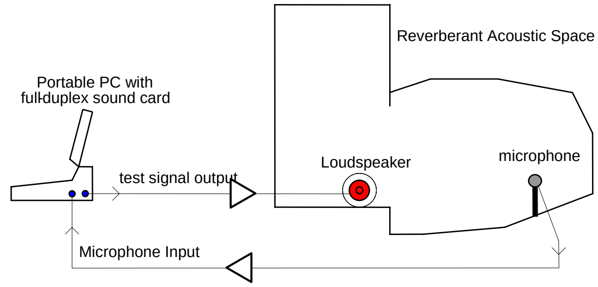


Fig. 2 – schematic diagram of the measurement system

A first approximation to the above system is a "black box", conceptually described as a Linear, Time Invariant System, with added some noise to the output, as shown in fig. 1.

In reality, the loudspeaker is often subjected to not-linear phenomena, and the subsequent propagation inside the theatre is not perfectly time-invariant.

The quantity which we are initially interested to measure is the impulse response of the linear system h(t), removing the artifacts caused by noise, not-linear behavior of the loudspeaker and time-variance.

The method chosen, based on an exponential sweep test signal with aperiodic deconvolution, provides a good answer to three above problems: the noise rejection is better than with an MLS signal of the same length, not-linear effects are perfectly separated from the linear response, and the usage of a single, long sweep (with no synchronous averaging) avoids any trouble in case the system has some time variance.

The mathematical definition of the test signal is as follows:

$$x(t) = \sin\left[\frac{\omega_1 \cdot T}{\ln\left(\frac{\omega_2}{\omega_1}\right)} \cdot \left(e^{\frac{t}{T}\cdot\ln\left(\frac{\omega_2}{\omega_1}\right)} - 1\right)\right] \qquad (1)$$

This is a sweep which starts at angular frequency $\omega_1$, ends at angular frequency $\omega_2$, taking T seconds.

When this signal, which has constant amplitude and is followed by some seconds of silence, is played through the loudspeaker, and the room response is recorded through the microphone, the resulting signal exhibit the effects of the reverberation of the room (which "spreads" horizontally the sweep signal), of the noise (appearing mainly at low frequencies) and of the not-linear distortion.

These "distorted" harmonic components appear as straight lines, above the "main line" which corresponds with the linear response of the system. Fig. 3 shows both the signal emitted and the signal re-recorded through the microphone.

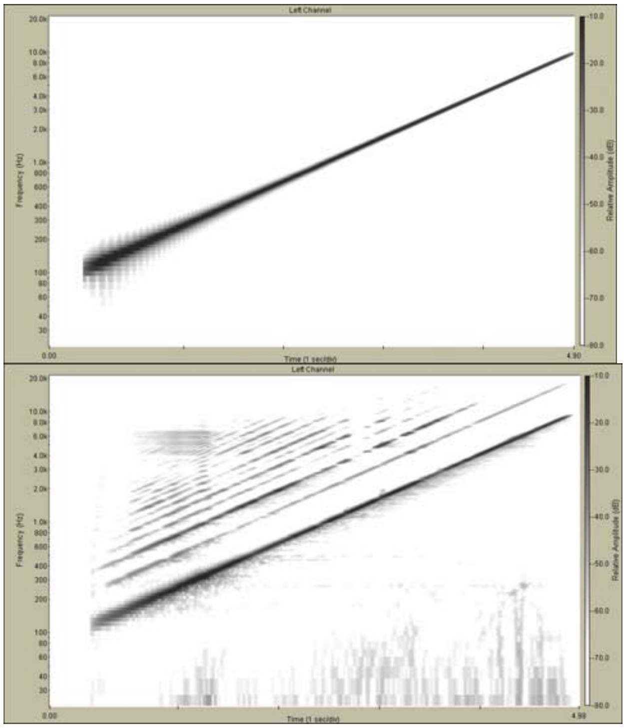


Fig. 4 – sonograph of the test signal x(t) and of the response signal y(t)

Now the output signal y(t) has been recorded, and it is time to post-process it, for extracting the linear system's impulse response h(t).

What is done, is to convolve the output signal with a proper filtering impulse response f(t), defined mathematically in such a way that:

$$h(t) = y(t) \otimes f(t) \qquad (2)$$

The tricks here are two:

- to implement the convolution aperiodically, for avoiding that the resulting impulse response folds back from the end to the beginning of the time frame (which would cause the harmonic distortion products to contaminate the linear response)
- to employ the Time Reversal Mirror approach for creating the inverse filter f(t)

In practice, f(t) is simply the time-reversal of the test signal x(t). This makes the inverse filter very long, and consequently the above convolution operation is very "heavy" in terms of number of computations and memory accesses required (on modern processors, memory accesses are the slower operation, up to 100 times slower than multiplications).

However, the author developed a fast and efficient convolution technique, which allows for computing the above convolution in a time which is significantly shorter than the length of the signal. [7]

It must also be taken into account the fact that the test signal has not a white (flat) spectrum: due to the fact that the instantaneous frequency sweeps slowly at low frequencies, and much faster at high frequencies, the resulting spectrum is pink (falling down by -3 dB/octave in a Fourier spectrum). Of course, the inverse filter must compensate for this: a proper amplitude modulation is consequently applied to the reversed sweep signal, so that its amplitude is now increasing by +3 dB/octave, as shown in fig. 5.

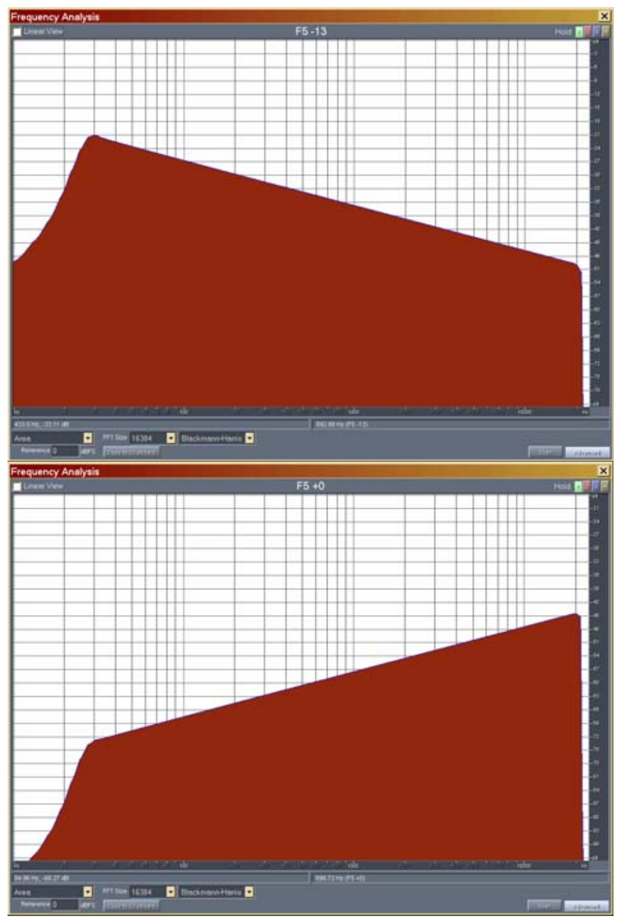


Fig. 5 – Fourier spectrum of the test signal (above) and of the inverse filter (below)

When the output signal y(t) is convolved with the inverse filter f(t), the linear response packs up to an almost perfect impulse response, with a delay equal to the length of the test signal. But also the harmonic distortion responses do pack at precise time delay, occurring earlier than the linear response. The aperiodic deconvolution technique avoids that these anticipatory response folds back inside the time window, contaminating the late part of the impulse response.

Fig. 6 shows a typical result after the convolution with the inverse filter has been applied.

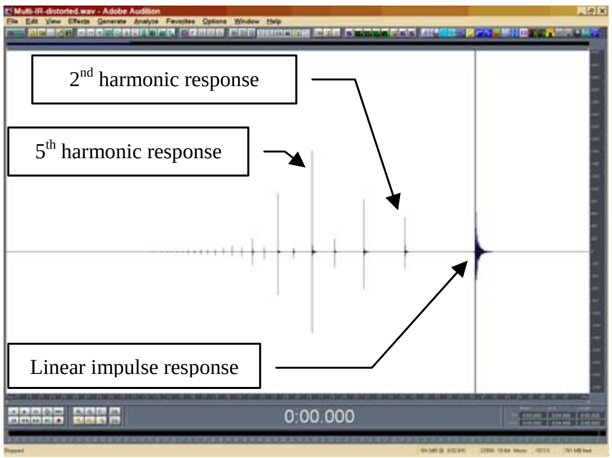


Fig. 6 – output signal y(t) convolved with the inverse filter f(t)

At this point, applying a suitable time window it is possible to extract just the portion required, containing only the linear response and discarding the distortion products.

The advantage of the new technique above the traditional MLS method can be shown easily, repeating the measurement in the same conditions and with the very same equipment. Fig. 7 shows this comparison in the case of a measurement made in a highly reverberant space (a church).

It is easy to see how the exponential sine sweep method produces better S/N ratio, and the disappearance of those nasty peaks which contaminate the late part of the MLS responses, actually caused by the slew rate limitation of the power amplifier and loudspeaker employed for the measurements, which produce severe harmonic distortion.

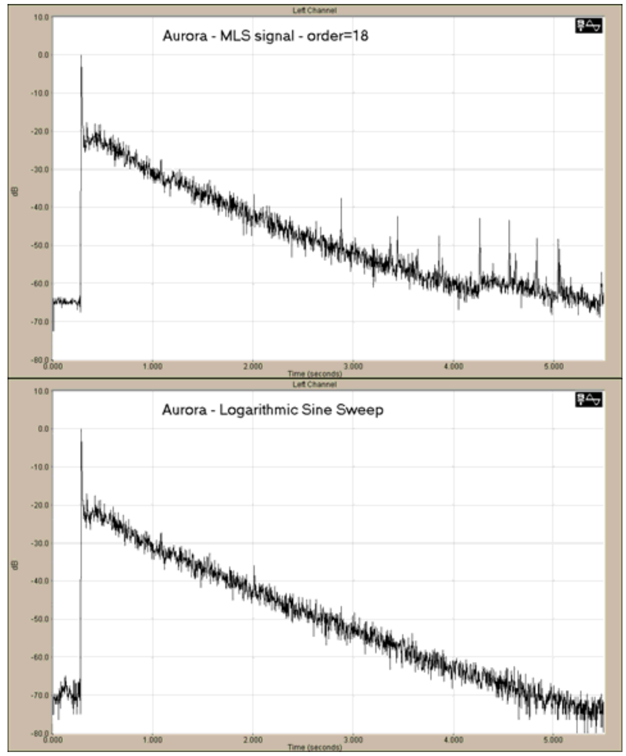


Fig. 7 – comparison between MLS and sine sweep measurements

This method has nowadays wide usage, and is often employed for measuring high-quality impulse responses which are later employed as numerical filters for applying realistic reverberation and spaciousness during the production of recorded music [8].

## 3. PROBLEMS WITH THE ESS METHOD

Despite the significant advantages shown by the ESS method in comparison with all the other previously-employed methods, some problems can still be found, as already pointed out in chapter 1.

In the following subchapters, each of these problems is analyzed, and proper workarounds are presented.

### 3.1. Pre-ringing

The measured impulse response often shows some significant pre-ringing before the arrival of the direct sound.

This is easily shown performing directly the deconvolution of the IR from the original test signal, without having it passing through the system-under-test.

This way, one should get a theoretically-perfect Dirac's delta function. The old MLS method is perfect in this case, providing exactly a theoretical pulse. The following figure shows instead what happens with the standard ESS method.


Fig. 8 – pre-ringing artifact with fade-out

As shown in fig. 8, the peak is in reality some sort of Sync function, and it shows a number of damped oscillations both before and after the main peak. This is due to the limited bandwidth of the signal (22 Hz to 22 kHz, in this case) and to the presence of some fade-in and fade-out on the envelope of the test signal (0.1s in this example, employing a 15s-long ESS). These two factors define substantially a trapezoidal window in the frequency-domain, which becomes the Sync-like function in time domain.

However, the situation ameliorates significantly if we remove the fade-out. The following figure show the results obtained with exactly the same settings as in the previous case, but with a length of the fade-in set to 0.0s (fade-in is still 0.1s).

Albeit the appearance of the waveform looks the same (due to the "analogue waveform" display of Adobe Audition), looking carefully at the digital values (the small squares along the waveform) one now sees that the results are very close to a theoretical Dirac's Delta function, and that no pre-ringing or post-ringing are anymore significantly present.


Fig. 9 – reduced pre-ringing artifact without fade-out

However, it is not a good idea to remove completely the fade-out: at the end of the sweep, the final value computed could be not-zero, and consequently the sound system will be excited with a step function, which spreads a lot of energy all along the spectrum.

A solution alternative to removing the fade-out is to continue the sweep up to the Nyquist frequency (22050 Hz, in our example, as the sampling rate was 44.1 kHz), and cutting it manually at the latest zero-crossing before its abrupt termination. This way, no pulsive sound is generated at the end, and the full-bandwidth of the sweep removes almost completely the high-frequency pre-ringing.

However, in some cases, also low frequencies can cause a significant pre-ringing. This is shown easily employing a "loopback" connection, that is, connecting a wire directly from the output to the input of the sound card.

The following figure shows the result of a "loopback" measurement, employing the same parameters as for the previous example (fs=44100 Hz, sweep from 22 Hz to 22050 Hz, 15s long, 0.1s fade-in, no fade-out).


Fig. 10 – low-frequency pre-ringing artifact

Removing the fade-in does not provide any benefit, in this case. So, the way of controlling this type of pre-ringing (due to the analog equipment) is to create a proper time-packing filter, and to apply it to the measured IR.

A packing filter is a filter capable of compacting the time-signature of the impulse response. Various methods for creating a numerical approximation to an ideal packing filter have been proposed in the past. The method employed here is the one developed by Ole Kirkeby, when working at the ISVR with prof. Nelson [9]. Although Kirkeby did propose this method for multichannel inversion (cross-talk cancellation), it can be successfully employed also just for the purpose of packing in time the transfer function of a single-input, single-output system.

The Kirkeby algorithm is as follows:

1) The IR to be inverted is FFT transformed to frequency domain:

$$H(f) = FFT\,[h(f)] \qquad (3)$$

2) The computation of the inverse filter is done in frequency domain:

$$C(f) = \frac{Conj[H(f)]}{Conj[H(f)] \cdot H(f) + \varepsilon(f)} \qquad (4)$$

Where ε(f) is a small regularization parameter, which can be frequency-dependent, so that the inversion does not operates outside the frequency range covered by the sine sweep

3) Finally, an IFFT brings back the inverse filter to time domain:

$$c(t) = IFFT\,[C(f)] \qquad (5)$$

Usually the regularization parameter ε(f) is choosen with a very small value inside the frequency range covered by the sine sweep, and a much larger value outside that frequency range, as shown in the following figure:

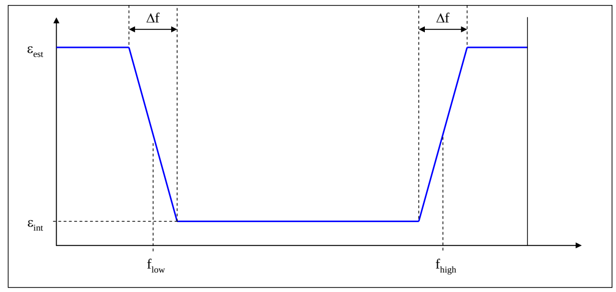


Fig. 11 – frequency-dependent regularization parameter

The following figure shows the inverse filter computed for compacting the "loopback" IR shown in fig. 10:

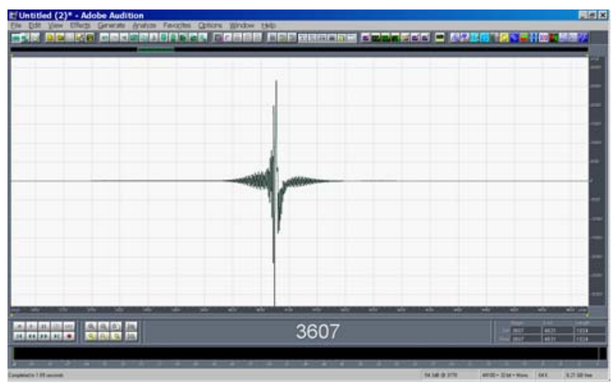


Fig. 12 – "compacting" inverse Kirkeby filter

When this filter is convolved with the measured "loopback" IR shown in fig. 10, the result is the one shown in the next figure:


Fig. 13 – "loopback" IR convolved with the "compacting" inverse Kirkeby filter

It can be seen that the usage of the inverse filter managed to re-pack the measured IR back to an almost perfect Dirac's Delta function.

In conclusion, pre-ringing artifacts can be substantially avoided by combining the usage of a wide-band sweep running up to the Nyquist frequency, without any fade-out, and the usage of a suitable "compacting" inverse filter, computed with the Kirkeby method from a "reference" impulse response.

In the example shown here, the "reference" measurement for computing the inverse filter has been performed electrically, so it does not contain the effect of power amplifier, loudspeaker and microphones. This makes sense if the goal of the measurement is to get information about the behaviour of these electroacoustics components (in most cases, for measuring the performances of the loudspeaker).

### 3.2. Equalization of the equipment

In other cases, in which the goal of the measurement is just to analyze the acoustical transfer function between an "ideal" sound source and an "ideal" receiver, also the effect of the electroacoustical devices should be removed. In this case, the "reference" measurement is a complete anechoic measurement including power amplifier, loudspeaker and microphone, and the Kirkeby inverse filter will remove any time-domain and frequency-domain artifact caused by the whole measurement system.

For example, the following figure shows the anechoic measurement of the transfer function of a loudspeaker+microphone setup:

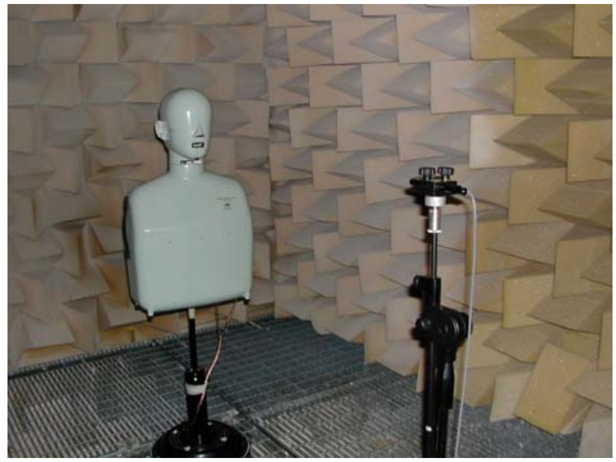

Fig. 14 – measurement of the "reference" IR of an artificial mouth and an omnidirectional microphone

This example refers to a small, limited-range loudspeaker, employed in a head-and-torso simulator. The measured IR and its frequency response are shown in the following pictures:


Fig. 15 – measured IR of the artificial mouth system


Fig. 16 – measured frequency response of the artificial mouth system

Again, a Kirkeby inverse filter is computed, for correcting the transfer function of the whole measurement system (this time the usable frequency range has been narrowed to 10-11000 Hz):


Fig. 17 – "equalizing" inverse Kirkeby filter

When this inverse filter is applied (by convolution) to the measured IR of this artificial mouth system, we get an IR and a frequency response as shown here below:

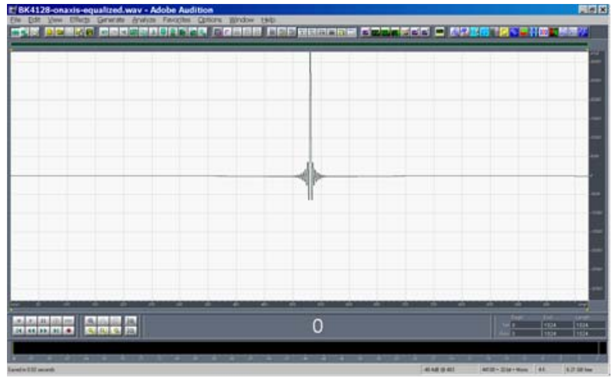


Fig. 18 – measured IR of the artificial mouth system after equalization with the inverse filter

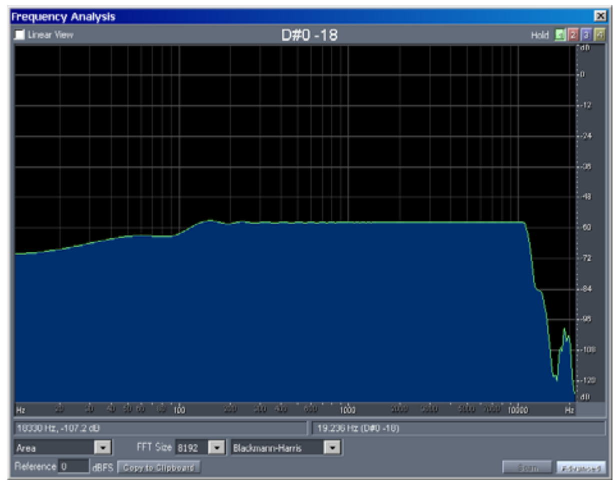


Fig. 19 – measured frequency response of the artificial mouth system after equalization

Although in this case the inverse filter did not manage to provide a "perfect" result, it still caused the transfer function of the system to closely approach the "ideal" one. This way, the electroacoustical sound system can be employed for measurements without any significant biasing effect.

The latter point to be discussed is if it is better to apply this equalizing filter to the test signal before playing it through the system, or to the recorded signal (indifferently before or after the deconvolution).

Both approaches have some advantages and disadvantages. Applying the equalizing filter to the test signal usually results in a weaker test signals being radiated by the loudspeaker, and in clipping at extreme frequencies (where the boost provided by the equalizing filter is greater).

On the other hand, the usage of the filter after the measurement is done results in “colouring” the spectrum of the background noise, which can, in some case, become audible and disturbing.

In practice, has it often happens, the better strategy revealed to be hybrid: the test signal is first roughly equalized, employing one of the standard tools provided by Adobe Audition (for example Graphic Equalizer). This allows to limit the boost at extreme frequencies and the gain loss at medium frequencies, but however the radiates sound becomes already almost flat.

Then, as usual, a reference anechoic measurement is performed (employing the pre-equalized test signal); a Kirkeby inverse filter is thereafter computed, with the goal of removing the residual colouring of the measurement system. This inverse filter is applied as a post-filter, to the measured data, ensuring that the total transfer function of the measurement system is made perfectly flat. This is the approach successfully employed in the Waves project, as described in more detail in [8].

### 3.3. Pulsive noises during the measurement

When long sweeps are employed for improving the signal-to-noise ratio, the risk that some pulsive noise occurs during the measurement increases, as it is difficult to keep people perfectly still for more than a few seconds. Typical sources of pulsive noise are objects falling on the floor, seats being moved, or “cracks” caused by steps over wooden floors.

The following sonogram shows a recorded sweep contaminated by an evident spurious pulsive event (the vertical line), caused by an object falling on the floor.

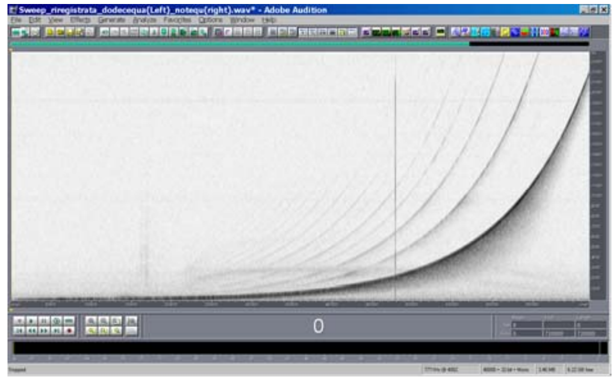


Fig. 20 – pulsive event contaminating an ESS measurement

After convolution with the inverse filter, this pulsive event causes a quite evident artifact on the deconvolved IR, as shown here:

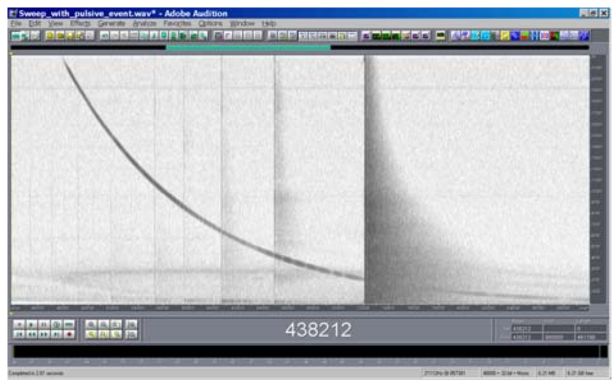


Fig. 21 – Artifact caused by a pulsive event

In practice, the artifact is a sort of frequency-decreasing sweep, starting well before the beginning of the linear impulse response, and continuing after it. The first part is practically irrelevant on the linear IR, as it will be cut away together with the harmonic distortion responses.

However, the part of this spurious sweep occurring in the late part of the measurement can cause severe problems. In particular, when analyzing the reverberant tail, this artifact is causing large errors on the estimate of the reverberation time and of the other acoustical parameters computed according to ISO 3382. The following figure shows a comparison between the octave-band-filtered IR with and without contamination by the spurious pulsive noise.

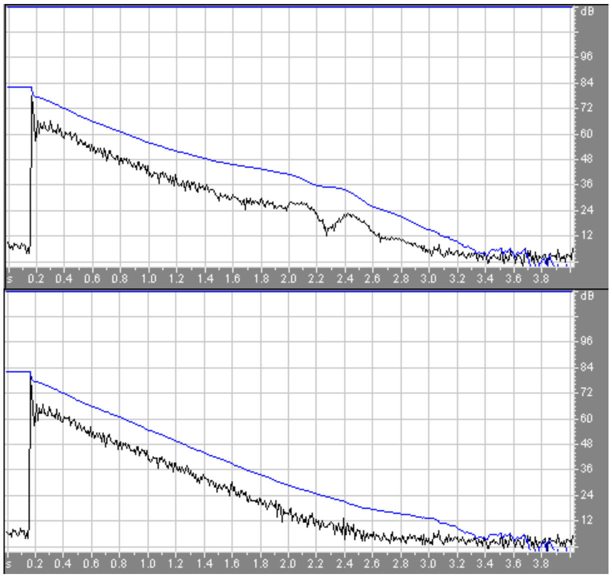


Fig. 22 – octave-band filtered IR (at 1 kHz) contaminated from pulsive noise (above) and without contamination (below)

The presence of the spurious effect generated by the pulsive noise is causing an overestimate of T30 (2.48 s instead of 2.13 s). Also Clarity C80 and Center Time are affected, but more lightly.

One way of removing this artifact consists in silencing the recording signal in correspondence of the pulsive event, as shown in the following figure:

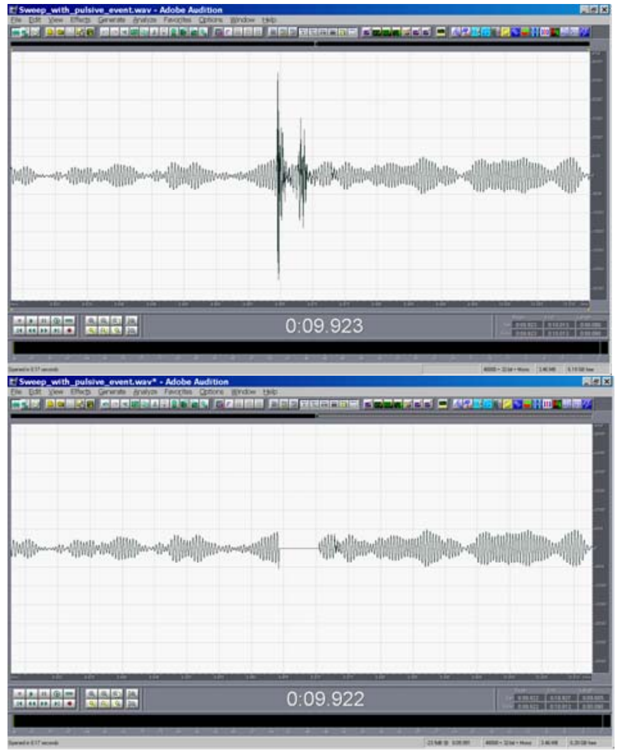


Fig. 23 – silencing the spurious event

After deconvolving the edited signal, the following IR is obtained:

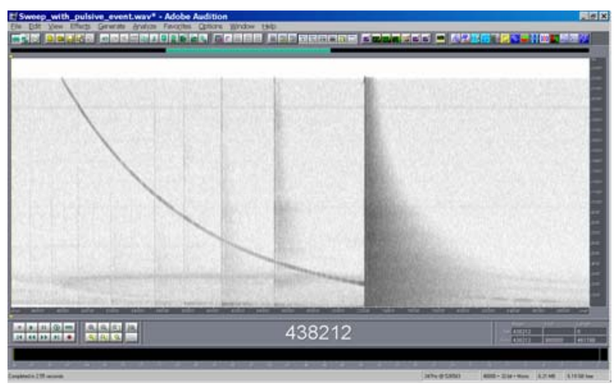


Fig. 24 – effect of the silenced pulsive event on the deconvolved IR

Despite silencing the event, the artifact is still there, albeit with reduced amplitude. The analysis of the reverberant tail still shows some effect of the pulsive artifact, as shown here:


Fig. 25 – octave-band filtered IR
with silenced pulsive event

A much better removal of the pulsive event is obtained by employing the Click/Pop Eliminator provided by Adobe Audition. The following picture shows how it works:

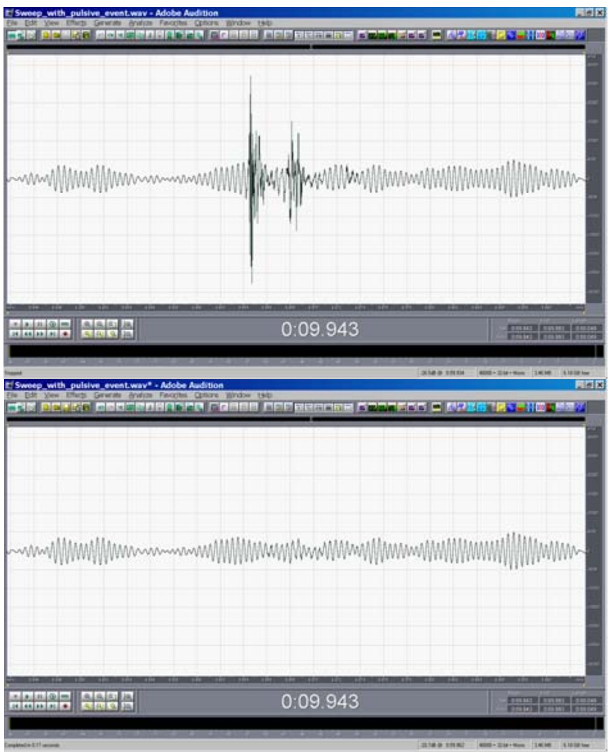


Fig. 26 – effect of the Auto Click/Pop Eliminator

In this case, the result of the deconvolution is the following:

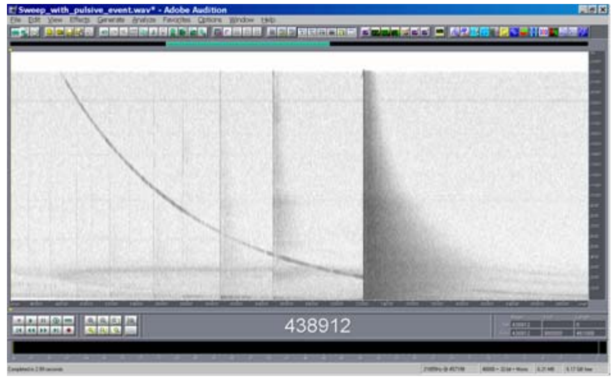


Fig. 27 – effect of the pulsive event
on the deconvolved IR after click/pop Eliminator

The artifact has been further reduced, but it is still there.

Finally, an even better way of removing the artifact is based on the knowledge of the frequency of the sine sweep at the moment in which the pulsive event did happen. In the case presented here, the instantaneous frequency was 2159 Hz. So, applying a narrow-passband filter at this exact frequency, all the wide-band noise is removed, and a "clean" sinusoidal waveform is restored, as shown in the following figures:

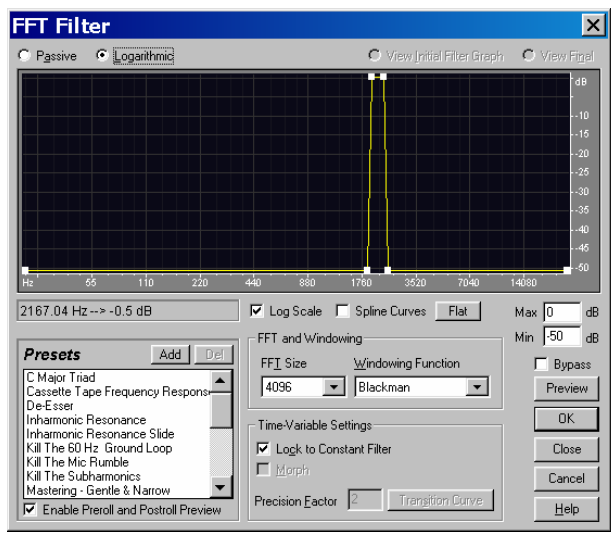


Fig. 28 – usage of FFT Filter for removing the pulsive artifact

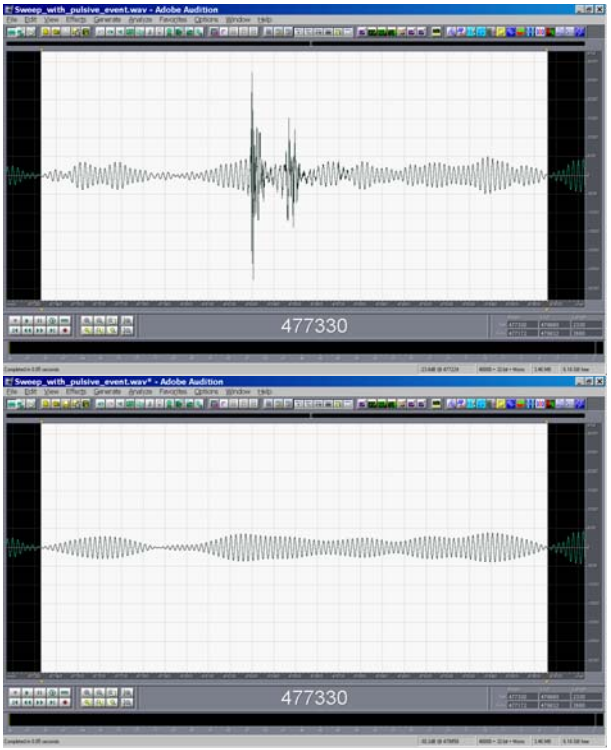


Fig. 29 – effect of FFT filter for removing the pulsive artifact

After deconvolution, the measured impulse response is as follows:

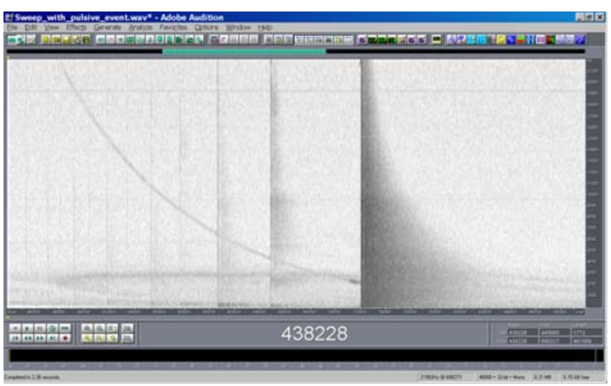


Fig. 30 – result of the FFT filter

Now the artifact amplitude has been reduced so much that there is no more distortion of the reverberant tail, as shown here:

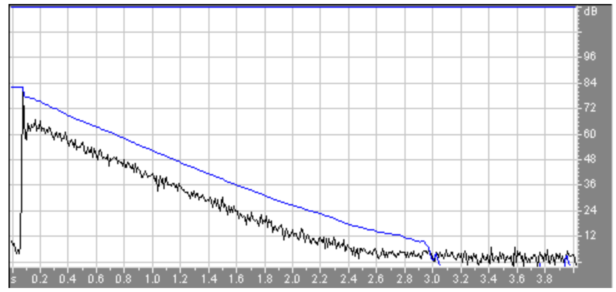


Fig. 31 – octave-band filtered IR with pulsive event removed with FFT filter

So it can be concluded that the best way of removing a pulsive artifact from a sweep measurement is to apply a narrow-band filter just around the instantaneous frequency at which the event occurred.

### 3.4. Clock mismatch

One of the great advantages of the ESS method over other methods for measuring the impulse response is that a tight synchronization between the playback clock and the recording clock is not required.

In fact, even if two completely independent hardware devices are employed, and no clock synchronization is employed, usually the impulse response obtained is perfectly clean and without observable artifacts. However, when the mismatch between the two clocks becomes significant, the deconvolved impulse response starts to be “skewed” in the frequency-time plane.

For example, the following figure shows the result of a purely-electrical measurement, obtained playing the test signal with a portable CD player, directly wired to a computer sound card, employed for recording.

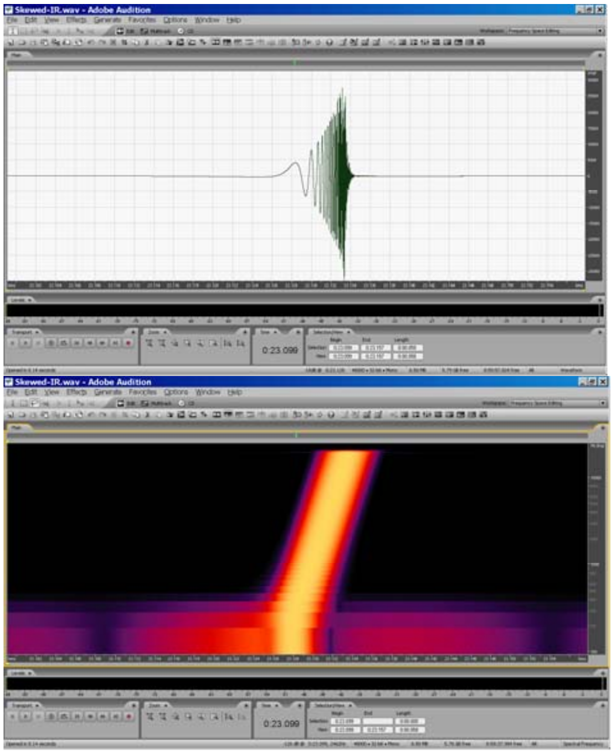


Fig. 32 – a skewed IR

The waveform clearly shows that low frequencies are starting earlier than high frequencies, and the sonograph demonstrates that, with a logarithmic frequency scale, the IR does not have a vertical (synchronous) appearance, but a sloped (skewed) appearance.

Various methods can be applied for re-aligning the clocks. For example, if a “reference” measurement can be performed, we could try to use a Kirkeby inverse filter for fixing the mismatch, as already shown in chapters 3.1 and 3.2.

The following figure show the result of such an inverse filter applied to the electrical measurement performed.

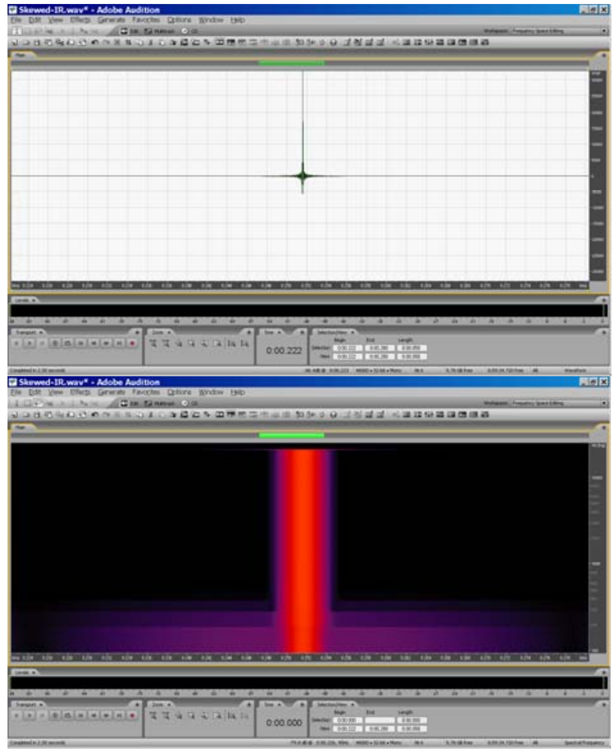


Fig. 33 – correction of a skewed IR employing a Kirkeby inverse filter

The result obtained employing the inverse filter is quite good; and it is also correcting for the magnitude of the frequency response of the system, not only for the frequency-dependent delay.

Nevertheless, this approach requires the availability of a clean reference measurement, performed either electrically (as in this example) or under anechoic conditions.

Whenever a reference measurement is not available, the inverse filter approach cannot be employed. Another possible solution is the usage of a pre-strecthed inverse filter for performing the IR deconvolution.

For example, in this example it can be seen how the original inverse filter is too short. If we now create an inverse filter slightly longer than the original one, we can correct for the skewness of the sonograph.

Looking again at fig. 32, we see that the skewness is approximately 8.5 ms long. So we generate a new sine sweep, and its inverse sweep, 8.5 ms longer than the original one.

When we convolve this longer inverse sweep with the recorded signal, the deconvolution produces the following result:

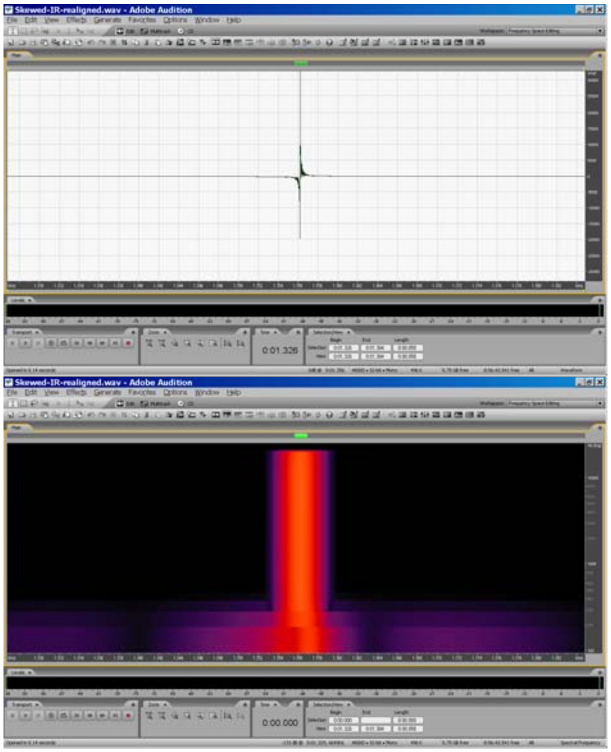

Fig. 34 – correction of a skewed measurement employing deconvolution with a longer inverse sweep

This result is not so clean as the one obtained with the Kirkeby inversion, but now we have got a quite good clock realignment without the need of a reference measurement.

It must be said, however, that a skewed impulse response, although bad to see and to listen, is still quite usable for computing acoustical parameters. It is nevertheless always useful to correct for the clock mismatch, as this significantly improves the peak-to-noise ratio. For example, with the data presented here, the usage of the longer inverse sweep for the deconvolution provides an amelioration of the peak-to-noise ratio by 12.45 dB, which is quite significant.

### 3.5. Time averaging

The usage of averaging several impulse responses for improving the signal-to-noise ratio is a deprecated technology when working with the ESS method.

Synchronous time averaging works only if the whole system is perfectly time-invariant. This is never the case when the system involves propagation of the sound in air, due to air movement and change of the air temperature. So, the preferred way for improving the signal to noise ratio is not to average a number of distinct measurements, but instead to perform a single, very long sweep measurement, as clearly recommended in the ISO 18233/2006 standard.

However, in some cases the usage of long sweeps is not allowed (for example, when the method is implemented on small, portable devices equipped with little memory), and so time-synchronous averaging is the only way for getting results in a noisy environment.

Unfortunately, even a very slight time-variance of the system produces substantial artifacts in the late part of the reverberant tail, and at higher frequencies.

This happens because the sound arriving after a longer path is more subject to the variability of the time-of flight due to unstable atmospheric conditions. Furthermore, a given differential time delay translates in a phase error which increases with frequency.

The following picture compares the sonographs of two IRS, the first comes from a single, long sweep of 50s, the second from the average of a series of 50 short sweeps of 1s each.

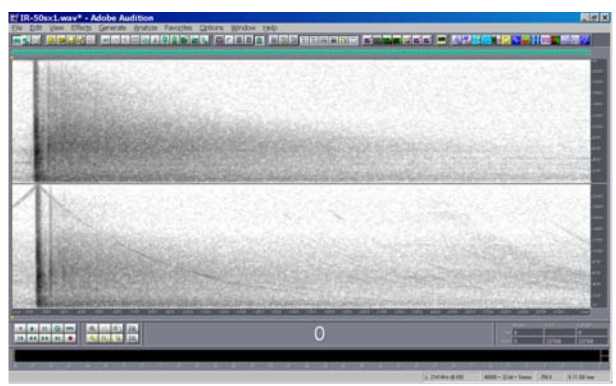

Fig. 35 – single sweep of 50s (above) versus 50 sweeps of 1s (below)

Although from the above picture it is not very easy to see the difference, it can be noted that the energy of the reverberant tail is significantly underestimated, at high frequency, in the second measurement. This can be seen easily displaying the spectrum of the signal in the range 100 ms to 300 ms after the direct sound, as shown here:

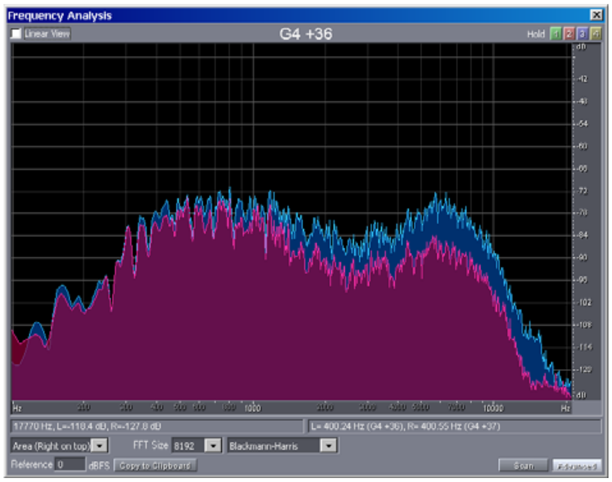


Fig. 36 – spectrum of single sweep of 50s (above) versus 50 sweeps of 1s (below)

It can be seen how, above 350 Hz, the synchronously-averaged IR is systematically underestimated. Around 5-6 kHz the underestimation is more than 10 dB.

This of course affects also the slope of the decay curve, and the estimate of reverberation times. The following figure shows the comparison between the octave-band filtered impulse response and decay curves at 4 kHz:

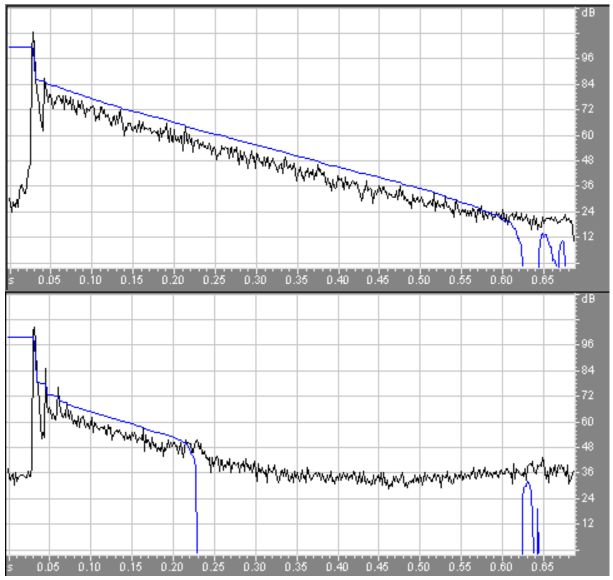

Fig. 37 – octave-band-filtered impulse response of a single sweep of 50s (above) versus 50 sweeps of 1s (below)

It can be seen how the single-sweep measurement is providing a perfectly linear decay with quite good dynamic range (63 dB), whilst the synchronously-averaged IR exhibit strong underestimate of the energy of the reverberant tail, and simultaneously a much worst signal-to-noise ratio (43 dB).

It can be concluded that synchronously-averaging a number of subsequent IRs obtained with the ESS method is causing unacceptable artifacts.

However, an alternative technique can be used, in these cases, for processing the data.

It is necessary to create a stereo file, containing the test signal in the left channel, and the recorded signal in the right channel, as shown here:

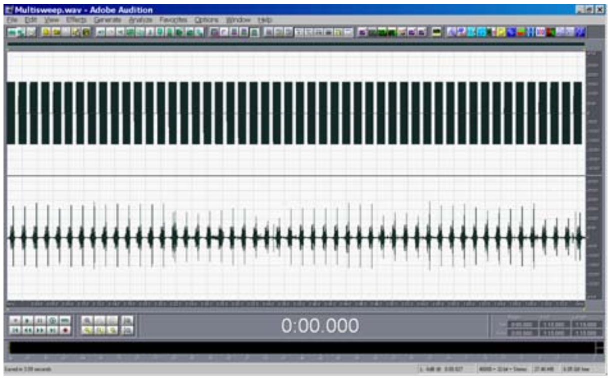


Fig. 38 – multisweep signal (test and response)

Now this stereo waveform is processed with the new Aurora plugin named Cross Functions, which is employed for computing the transfer function $H_1$, by performing complex averaging in spectral domain:

$$H_1(f) = \frac{G_{LR}}{G_{LL}} \qquad (5)$$

Where GLR and GLL are the averaged cross-spectrum and autospectrum, respectively

This is the user's interface of this plugin:

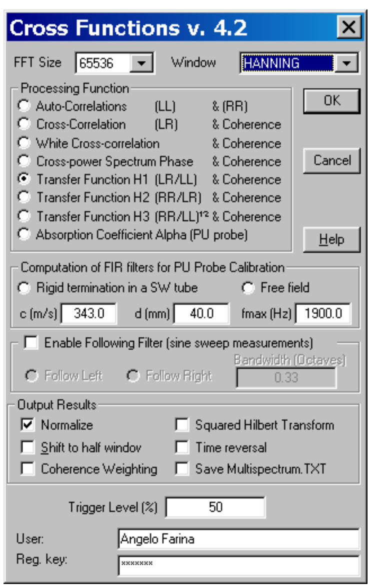


Fig. 39 – Computation of $H_1$

Only the first half of the resulting transfer function is kept, for removing most of the effects of the Hanning window. The following figure shows the recovered impulse response, compared with the single-sweep one:


Fig. 40 – single sweep of 50s (above)
versus 50 sweeps of 1s (below)
processed with the Cross Functions module

Analyzing the octave-band-filtered impulse response (at 4 kHz), the following is obtained:

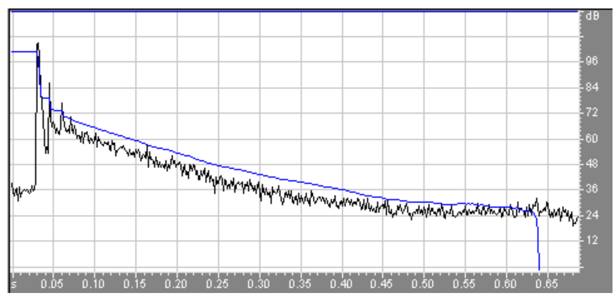


Fig. 41 – octave-band-filtered impulse response
of a 50 sweeps of 1s (Cross Functions)

It can be seen that the situation is now significantly better than with "standard" time-synchronous averaging: the frequency-domain processing provided an impulse response with better signal-to-noise ratio and with a reverberant tail only slightly underestimated. The single sweep method is still better, but now the difference is not so large, and the measurement result is still usable.

So, in practice, the employment of a number of independent sweeps can provide almost acceptable results, provided that the deconvolution and averaging of the impulse response are performed in reversed order (first averaging, then deconvolution), and in the frequency domain.

## 4. PERFORMANCE OF ELECTROACOUSTIC TRANSDUCERS

For room acoustics measurements, it is common to employ:

- An omnidirectional loudspeaker (dodecahedron)
- An Omni + Figure of Eight microphone
- A binaural microphone (dummy head)

In the previous chapter it has been already discussed how to measure the impulse response and frequency response of a measurement chain containing also loudspeakers and microphones, and how to reasonably equalize it. However, the problem still arises of the spatial properties (directivity) of these transducers.

It will be shown here that the measured directivities of loudspeakers and microphones differ significantly from the nominal ones, causing errors which are orders of magnitude greater than those described in the previous chapter.

### 4.1. Dodechaedron loudspeakers

These loudspeakers are usually employing single-way, wide-band transducers, and require heavy equalization fro providing flat sound power response. However, the equalization cannot correct the polar patterns of these loudspeakers, which deviate significantly from omnidirectional starting at frequencies above 1 kHz.

Here we present the results of polar patterns measured in anechoic conditions for three dodechaedrons. The first one is a standard-size (40cm diameter) employing for building acoustics measurements (LookLine D-300); the second one is a smaller version (25 cm diameter) specifically developed for measurement of impulse responses in theaters and concert halls (Look Line D-100). Finally, the third one employs waveguides for reconstructing a more uniform spherical wavefront (Omnisonics 1000).

The following figure shows the three dodechaedrons analyzed:

Fig. 42 – 3 dodechaedron loudspeakers

The above loudspeakers have been measured inside an anechoic chamber over a turntable, so the horizontal polar patterns have been obtained, in octave-bands.

The following three figures compare these polar patterns at 1000, 2000 and 4000 Hz.

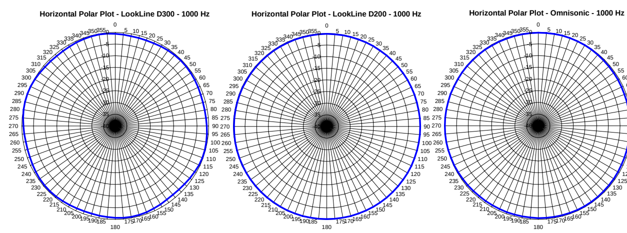


Fig. 43 – directivity patterns at 1 kHz

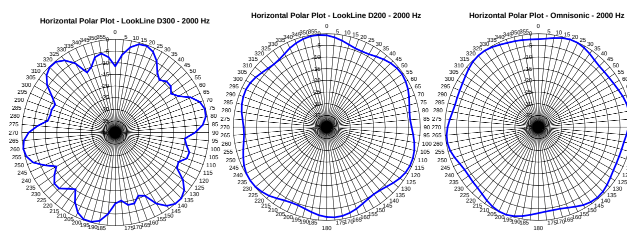


Fig. 44 – directivity patterns at 2 kHz

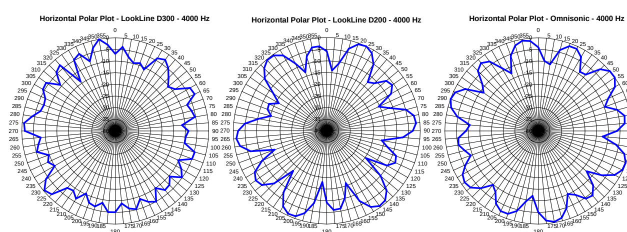


Fig. 45 – directivity patterns at 4 kHz

It can be seen how all three these dodecaedrons exhibit quite irregular polar patterns at medium-high frequency.

## 4.2. Omni + Figure of 8 mics

Although the usage of small-size measurement microphones does not pose any significant problem (as a B&K ½" capsule is almost perfectly omnidirectional and with flat frequency response up to 20 kHz), when spatial parameters such as LE, LF or LFC need to be measured it is necessary to employ a variable-directivity-pattern mike, providing both omnidirectional and figure-of-8 patterns.

For this purpose, it is common to employ not-measurement-grade probes, often manufactured by top-quality makers such as Neumann or Schoeps. However, the values of spatial parameters measured with different microphonic probes are often quite unreproducible.

So it was decided to perform a comparative experiment among 4 of these dual-pattern probes, including these mikes:

- Soundfield ST-250
- Bruel & Kjaer sound instensity kit type 3595
- Schoeps CMC5
- Neumann TLM 170R

The following image shows some of the probes being compared, during the measurements performed inside the Auditorium of Parma:

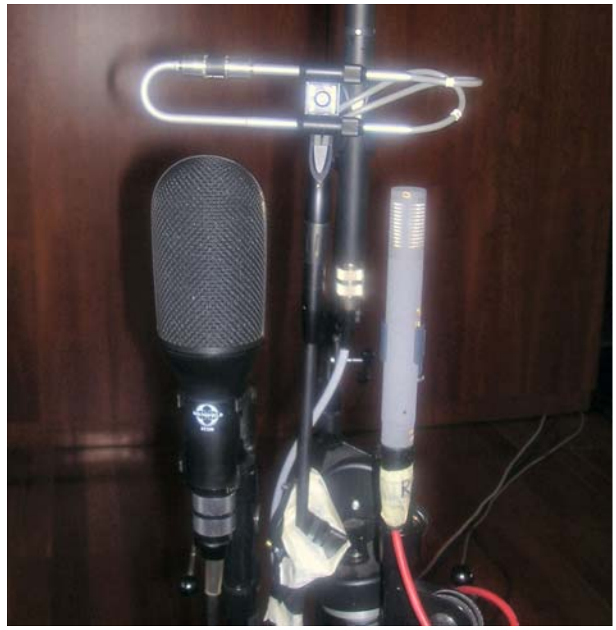

Fig. 46 – 3 microphonic probes

A stereo impulse response has been measured with each probe, containing the Omni response on the left channel, and the figure-of-8 response in the right channel. Each of these 2-channels IRs have been processed with the Aurora plugin named Acoustical Paramaters, specifying the type of probe being employed, as shown here:

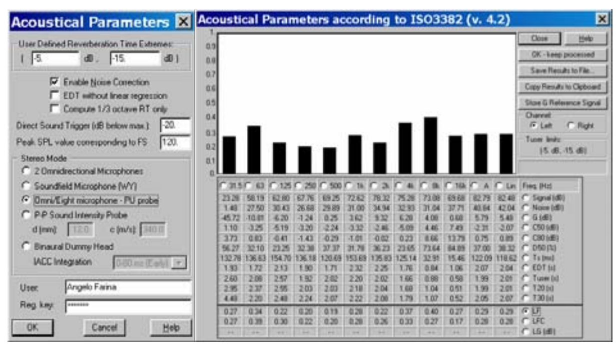


Fig. 47 – the Acoustical Parameters plugin

This way, the LF parameter has been measuring for all 4 probes, in octave bands, and at two distances from the sound source (7.5m and 25m). The following figure shows the results at 25m:

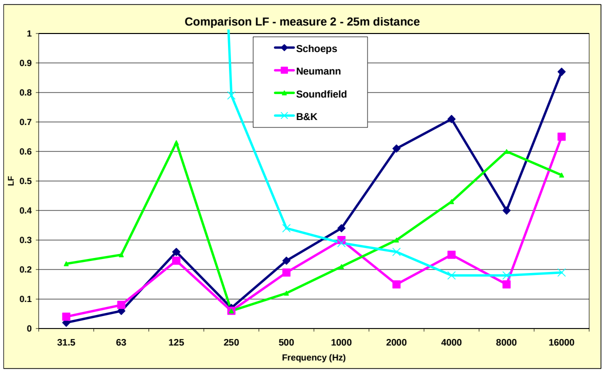


Fig. 48 – LF measured at 25m

It can be seen how the results are completely diverging; it is impossible to establish what of the 4 probes was measuring correctly, albeit the Schoeps looks more “reasonable” than the other three.

These deviations are caused by the polar patterns of the probes. As an example, here we report a couple of polar patterns of the Soundfield ST-250, measured on a turntable inside an anechoic room:

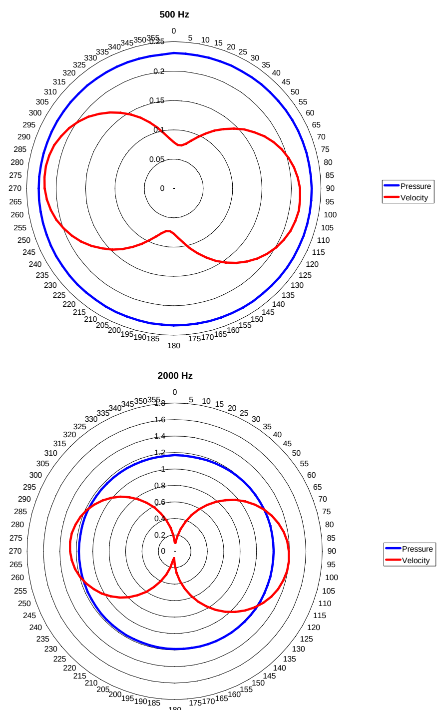


Fig. 49 – ST-250 – polar patterns at 500 Hz and 2 kHz

It can be seen that, even at medium frequencies, the figure-of-8 pattern is distorted, and is not properly gain-matched with the omnidirectional one. These deviations are even greater at very low and very high frequencies, as shown here:

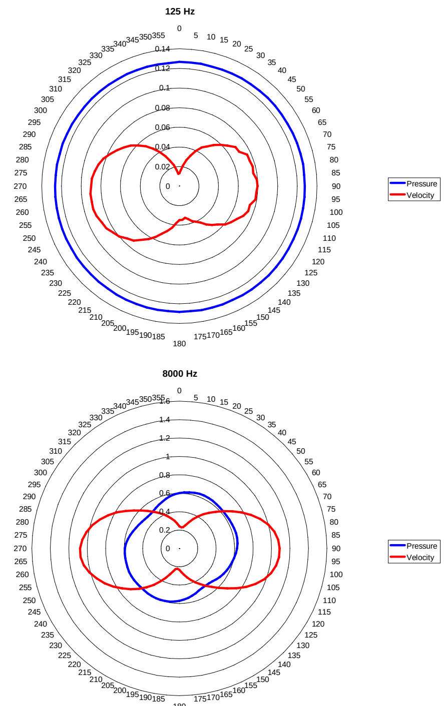


Fig. 50 – ST-250 – polar patterns at 125 Hz and 8 kHz

It can be concluded that actually no available microphonic system can be used for assessing reliably the values of spatial acoustical parameters such as LE, LF or LFC.

### 4.3. Binaural microphones

Another way of assessing the spatial properties of a room is by means of the IACC parameter (inter aural cross correlation), also defined in ISO-3382, and measurable employing a binaural microphone and the Aurora Acoustical Parameter plugin.

However, various makers of dummy heads produce quite different microphone assemblies. For checking comparatively their performances, a set of impulse response measurements have been performed in a large anechoic chamber, employing a turntable controlled by the sound card, as shown in the following figure:

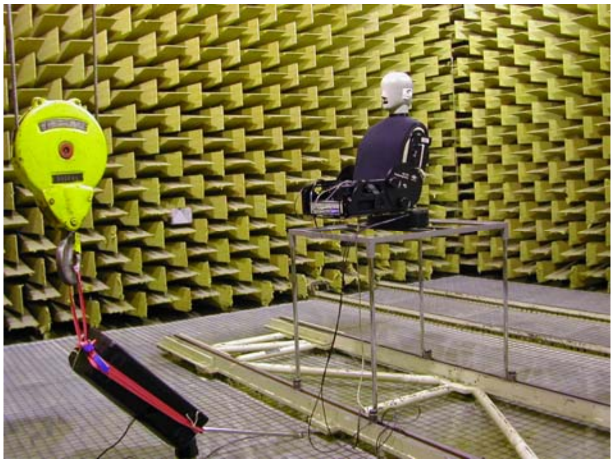

Fig. 51 – anechoic measurements on dummy heads

Also in this case 4 different binaural microphones have been tested:

- Bruel & Kjaer type 4100
- Cortex
- Head Acoustics HMS-III
- Neumann KU-100

A synthetic diffuse sound field has been generated, employing a number of loudspeakers surrounding the dummy head and feeding them with uncorrelated pink noise.

In principle, given the fact that the sound field was exactly the same, all the dummy heads should have given the same value of IACC. Instead, as shown in the following figure, the results have been quite diverging:

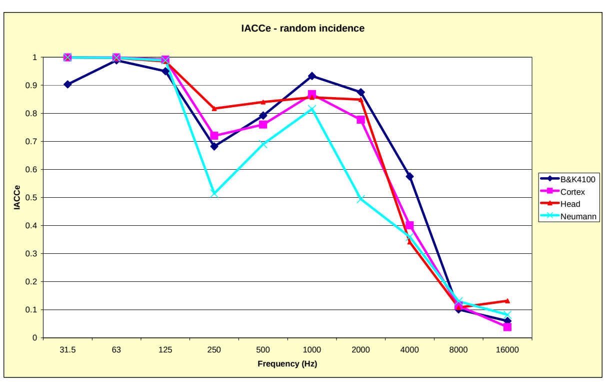


Fig. 52 – IACC measured with the 4 dummy heads

The deviations, however, are not so bad as those obtained in the previous chapter for the measurement of LF. It can be concluded that, with currently available systems, the measurement of IACC is slightly more reproducible than that of LF.

## 5. ACKNOWLEDGEMENTS

This work was supported by LAE (www.laegroup.org).

## 6. REFERENCES

[1] A.Farina – “Simultaneous measurement of impulse response and distortion with a swept-sine technique”, 110th AES Convention, February 2000.

[2] www.aurora-plugins.com

[3] P.Craven, M.Gerzon - "Practical Adaptive Room And Loudspeaker Equaliser for Hi-Fi Use" - 92nd AES Convention, March 1992

[4] D.Griesinger - "Beyond MLS - Occupied Hall Measurement With FFT Techniques" - 101st AES Convention, Nov 1996

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

[6] G. Stan, J.J. Embrechts, D. Archambeau – “Comparison of Different Impulse Response Measurement Techniques”, *JAES* Vol. 50, No. 4, p. 249, 2002 April.

[7] A. Torger, A. Farina – “Real-time partitioned convolution for Ambiophonics surround sound”, *2001 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics* - Mohonk Mountain House New Paltz, New York October 21-24, 2001.

[8] A. Farina, R. Ayalon – “Recording concert hall acoustics for posterity” - 24th AES Conference on Multichannel Audio, Banff, Canada, 26-28 June 2003

[9] O. Kirkeby, P. A. Nelson, H. Hamada, “The "Stereo Dipole" - A Virtual Source Imaging System Using Two Closely Spaced Loudspeakers” – *JAES* vol. 46, n. 5, 1998 May, pp. 387-395.