IMPULSE RESPONSE MEASUREMENTS

In the past, traditional measurements in the time domain were made with impulsive sounds and omnidirectional microphones.
Later, acoustic measurements were carried out based on computer hardware, loudspeaker and an stereo microphones.

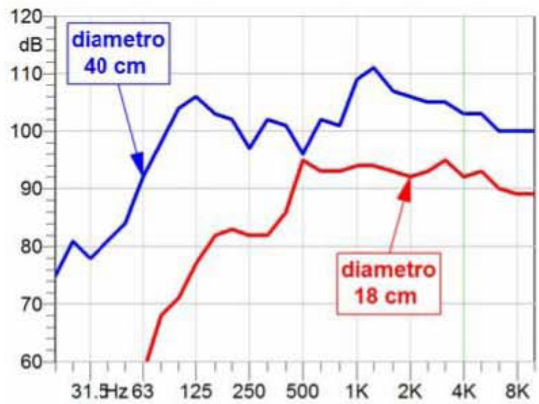
In the present, for electro-acoustical measurements we use a standard sound card or an SD recorder, a dodecahedron omnidirectional loudspeaker, and a Soundfield-like microphone (4 channels).

In the future, it is planned to use microphones arrays for capturing high-order spatial information and artificial sound sources employing a dense array of loudspeakers, capable of synthesizing the directivity pattern of any real-world source.

IMPULSIVE SOURCES

We analyze 4 types of impulsive sources: balloon, firecracker, clapping machine, blank pistol.

Balloon



The balloons have problems of frequency response, because they must be large for radiating low frequencies. But a large balloon is never an omnidirectional source…
In the figure above, we can see how the size influences low frequency spectrum.
Ordinary balloons (with diameter 18 cm) operate from 500 Hz up and so they are not useful for acoustic measurements.

Firecrackers

These are currently the preferred and most reliable impulsive sources available. They are also cheap and practical.

Match-type firecrackers, industrial grade (very repeatable)

The best firecrackers are those of the match type, which are industrially manufactured, and hence provide a very repeatable sound pulse.

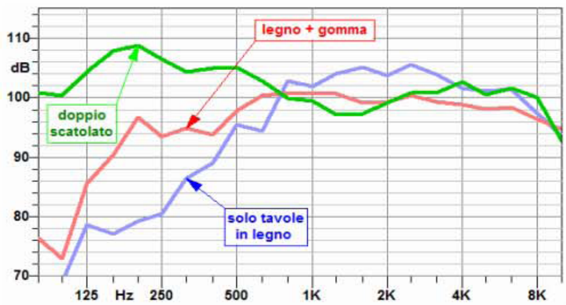
Being very small, they radiate a perfectly omnidirectional wavefront.

Their spectrum is reasonably flat, very powerful, and they radiate a lot of low frequency Energy.

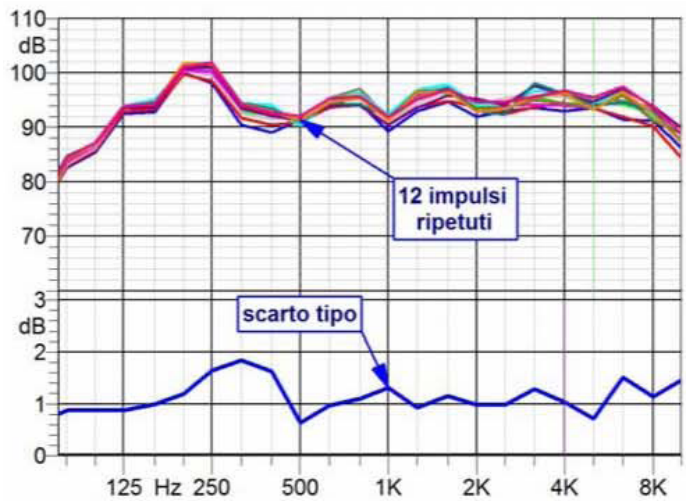
Spectrum of an industrial-grade firecracker

Clapping machine

Fig.7 – Tavole di legno incernierate Clappatore - versione iniziale 2007



The clapping machine has several advantages: <u>good frequency</u> response and <u>repeatability</u>. He has a low cost also. To obtain low frequencies we need “doppio scatolato” (it has a cavity), but for our example and the classroom size we can use also “legno-gomma”.

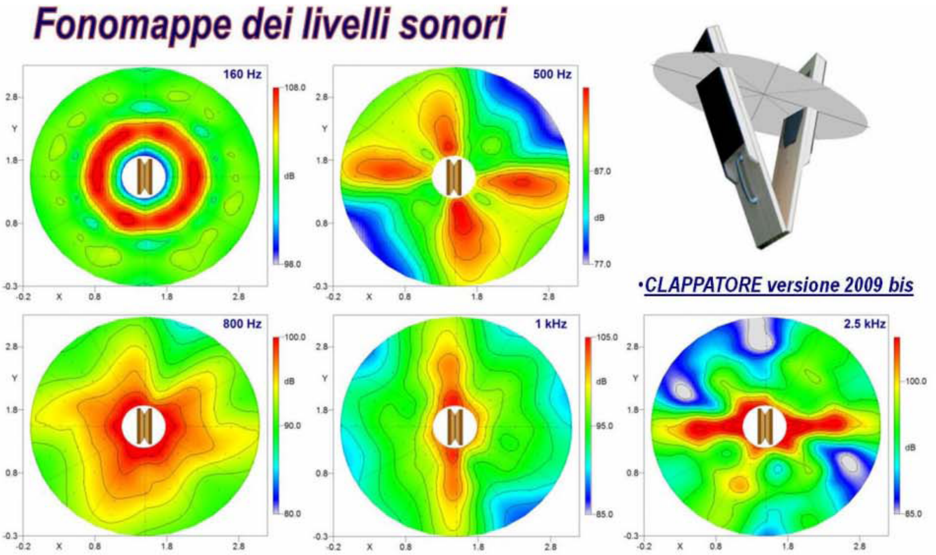


Verification of the repeatability.

It can be seen how the <u>12 pulses</u> are <u>identical</u> to each other. Instead, balloons had an intrinsic variability.

The problem of the clapping machine is highlighted by following diagrams. They show radiation diagrams of the clapping machine.

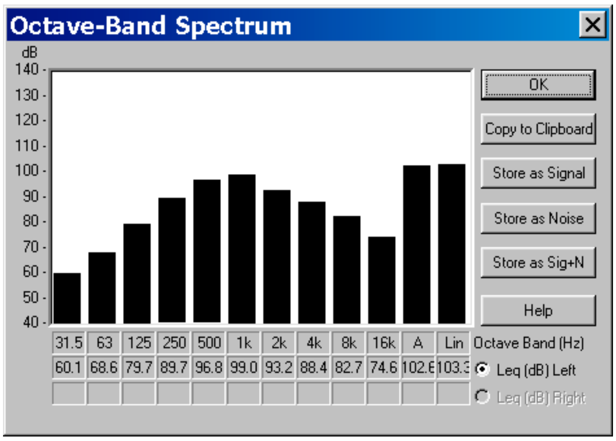
It can be see that at high frequency it is very directive.



This is a big problem, because the response varies depending on how we orient the object.

_Blank pistol

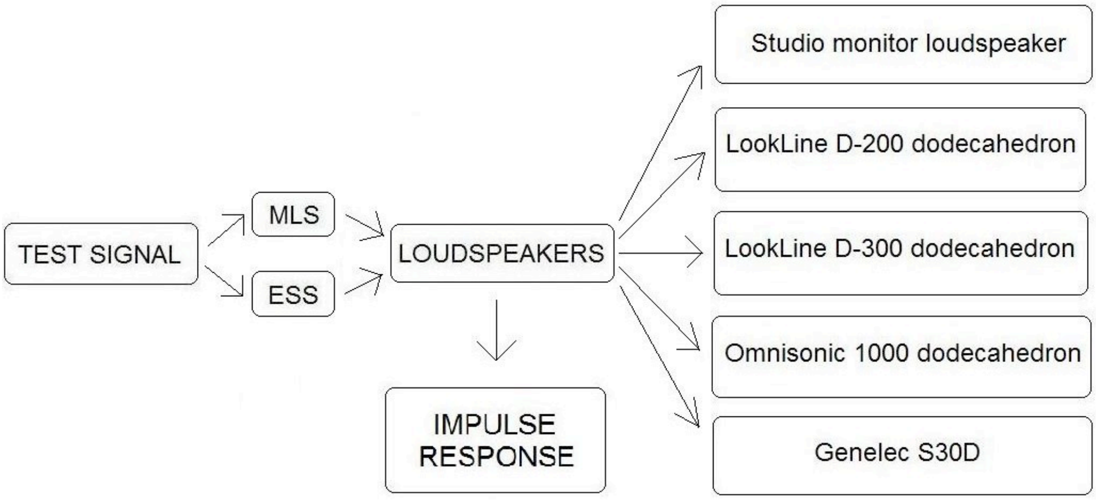
In an octave-band analysis we can see that the blank pistol has the typical bell-shaped spectrum centered at 1 kHz.



By looking at the impulse response, we have seen that at 1000Hz there is a very good signal-noise ratio: the decay curve is perfectly linear.

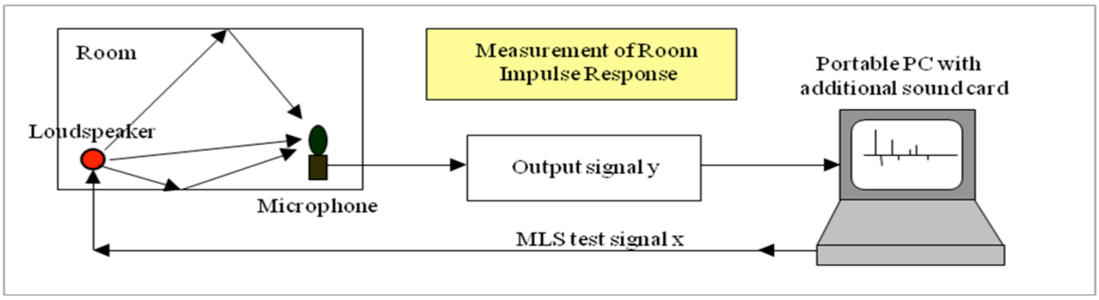
Thanks to the spherical diffuser, the sound radiation is perfectly omnidirectional

Impulse Response measurement with electroacoustic sources (loudspeakers)



The special test signal

A special test signal x(t) is generated by a computer (for example MLS test signal), which feeds the loudspeaker. The special test signal x(t) is not an impulsive signal because loudspeakers are not good in playing a direct delta pulse, they break if you play a loud pulse on a loudspeaker. We get only a limited power from a loudspeaker, so using a true pulse is not the optimal way of employing a loudspeaker. We usually play a long signal that spreads energy over time. The signal reverberates in the room from the loudspeaker while a microphone records the room's response. Thus we have an output signal y(t) but it is not the impulse response h(t) because x(t) is not the impulse signal. x(t) is a continuous signal (it can even be noise). We can determine h(t) if we know x(t) and y(t) by a deconvolution.



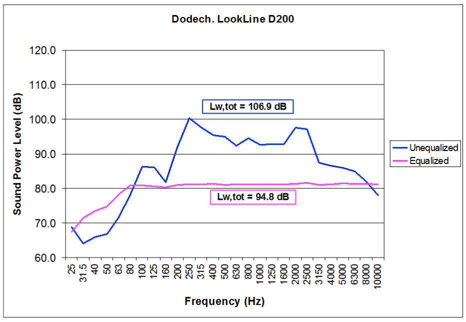
-Fig.01-

Loudspeakers

We have seen the characteristics of different kind of loudspeakers: studio monitor loudspeaker, LookLine D-200 dodecahedron, LookLine D-300 dodecahedron, Omnisonic 1000 dodecahedron and Genelec S30D. The normal studio monitor loudspeaker placed on the desk, has a good quality. It cannot be considered a professional loudspeaker for acoustical measurement because it is a directive source (and not omnidirectional).

According to the ISO 3382 standard the typical loudspeaker used to create an omnidirectional source is the dodecahedron. The LookLine D-200 dodecahedron in Fig.02 it is not just a dodecahedron because the base contains both the power amplifier and a subwoofer that radiates very low frequencies, while the mid-high frequencies are radiated by the sphere with the twelve loudspeakers. This kind of source generates the blue line which indicates the frequency response visible in Fig.03.

-Fig.02-



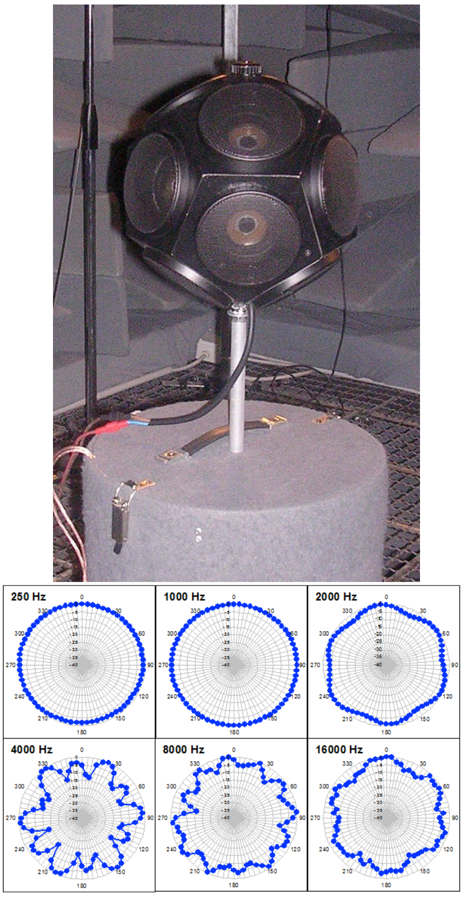
-Fig.03-

If we do not apply digital equalization the distribution is uneven, which is not good. We should have a flat response. If we apply digital equalization we get the flat pink line. The consequence of getting a flat power spectrum is that we loose 12dB. The radiation efficiency of this kind of loudspeaker is very poor because it irradiates only 1 acoustical Watt with an electrical power which can be 200W. The radiation efficiency of normal loudspeakers is lower than one percent. The loss of decibels can be really a trouble, so we need some tricks to improve signal to noise ratio, using special test signal, not just like white noise. If we use a normal test signal, like white noise, the signal to noise ratio will be bad.

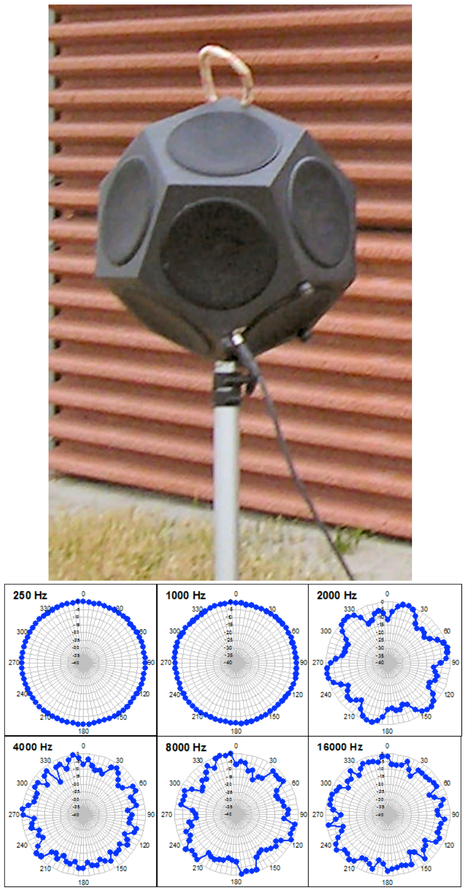
Directivity

We think that a spherical source radiates perfectly at every frequency, but this actually is not true. In fact the LookLine D-200 dodecahedron in the Fig.04 does not radiate perfectly the sound at every frequency: in effect it is uniform up to 2kHz but from 4kHz upwards it becomes distorted. The LookLine D-300 dodecahedron is bigger than th LookLine D-200 and already at 2kHz we can see the deterioration in radiation - Fig.05.

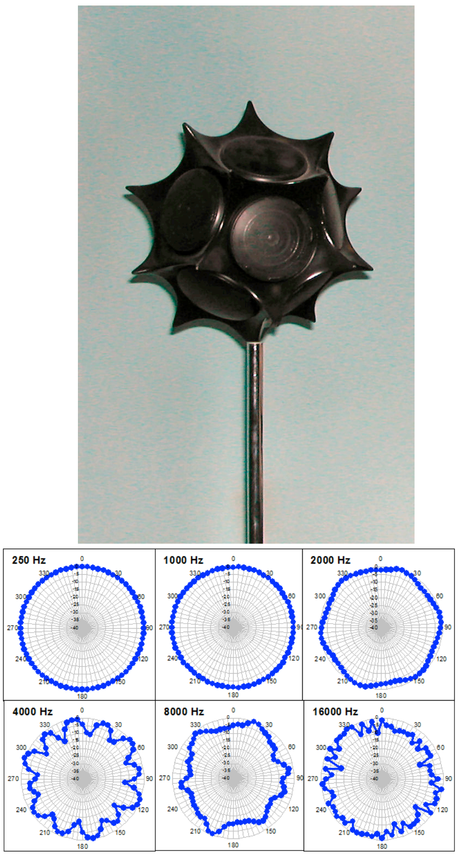
The smaller the loudspeaker, the more omnidirectional it is. The bigger the loudspeaker, the more directive it will be, so not omnidirectional. Omnisonic 1000 dodecahedron is a new type of loudspeaker, it has a particular shape like a flower. This loudspeaker makes sound radiated by different units to merge together, recreating a continuous spherical waveform. The American company who produces Omnisonic 1000 claims that its shape makes the radiation omidirectional at every frequency, but this is not the case, see Fig.06. Genelec S30D is not an omidirectional source, as shown in Fig.07. It can be used to reproduce real world sounds from some sources such as horn instruments, brass instruments and human voices.



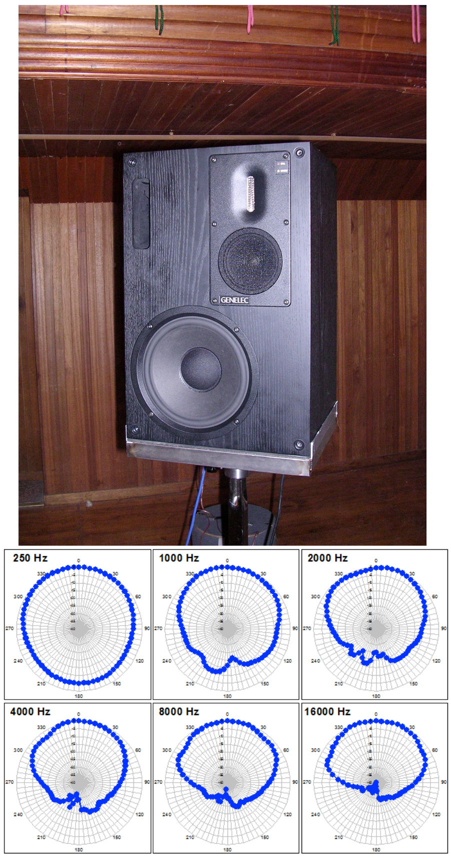
-Fig.04-



-Fig.05-



-Fig.06-



-Fig.07-

Measurement process

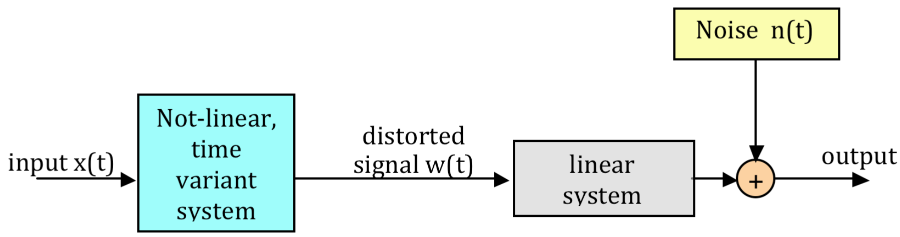
The measurement process, outlined in Fig.08, starts when the test signal x(t) passes through a loudspeaker, which in general is not a linear device. There is always some kind of distortion. It generates a distorted signal w(t) that reverberates inside the room, where the reverberation process occurs; reverberation is usually a perfectly linear process. Then we also capture the environmental noise, n(t), but of course we do not want this noise and we do not want the bad effects caused that are caused by the not linearity effects of the loudspeaker.

When measuring with a loudspeaker, the S/N ratio is always a problem, because the loudspeaker is not as powerful as the firecracker or the balloon. The radiation of the firecracker is 30dB more powerful than a typical loudspeaker.

If we try to squeeze out more power from the loudspeaker, driving it at an higher voltage, we boost its nonlinearities.

An optimal deconvolution technique has hence two goals:

- To improve the signal/noise ratio.
- To reject the not linear effect of the loudspeaker.



-Fig.08-

Test signals

In time different types of test signals have been developed, providing good immunity to background noise and easy deconvolution of the impulse response:

- MLS (Maximum Length Sequence, pseudo-random white noise).
- TDS (Time Delay Spectrometry, which is a linear sine sweep, also known in Japan as “stretched pulse” and in Europe as “chirp”)
- ESS (Exponential Sine Sweep) The most modern one invented by Prof. Farina.

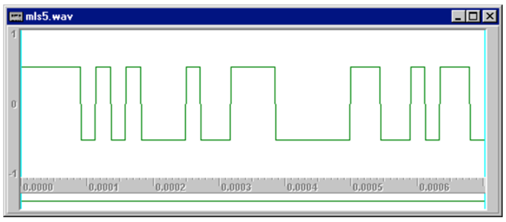
Each of these test signals can be employed with different deconvolution techniques, resulting in a number of “different” measurement methods. Today, the preferred choice is the ESS with non-circular deconvolution.

Short story of impulse response measurements

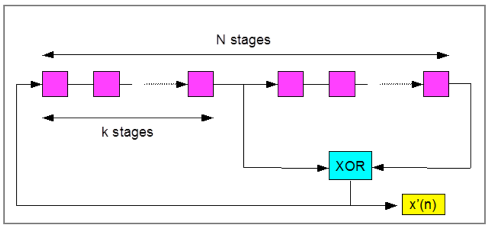
Before 1989 you could only measure the impulse response with a pistol or a balloon. The MLS signal was the first test signal invented in 1989, where an impulse response could be measured electroacoustically. It is generated by a long board inserted in a big computer and the board captured the sound from only one channel. MLSSA was the first apparatus for measuring impulse responses with MLS. A more modern version of MLSSA signal was the CLIO system, invented in Florence. Techron TEF 10 was the first apparatus for measuring impulse responses with TDS and it supported Iso MLS. We now use a normal PC with a good esternal Firewire soundcard, where we connet our microphone, loudspeaker, and we can play our test signal. It is considerably more powerful, with 8-10 input channels available, and it works up to 24bit, 192kHz. Some of the Adobe Audition Plugins called Aurora were developed here in Parma. It was the first measurement system based on standard sound cards and employing the Exponential Sine Sweep method. It also works with traditional TDS and MLS methods, so the comparison can be made employing exactly the same hardware.

MLS (Maximum Length Sequence)

The MLS signal is a periodic binary signal (one or zero), as shown in Fig.09, obtained with a suitable shift-register, shown in Fig.10, configured for maximum length of the period (L).



-Fig.09-



-Fig.10-

$$L = 2^N - 1$$

The deconvolution is performed with the Hadamard matrix, obtained by L permutations of the original MLS sequence.

$$\tilde{M}(i,j) = m\left[(i + j - 2)\bmod L\right] - 1$$

The result is the required impulse response h(i), if the system is linear and time-invariant.

$$h = \frac{1}{L+1} \cdot \tilde{M} \otimes y$$

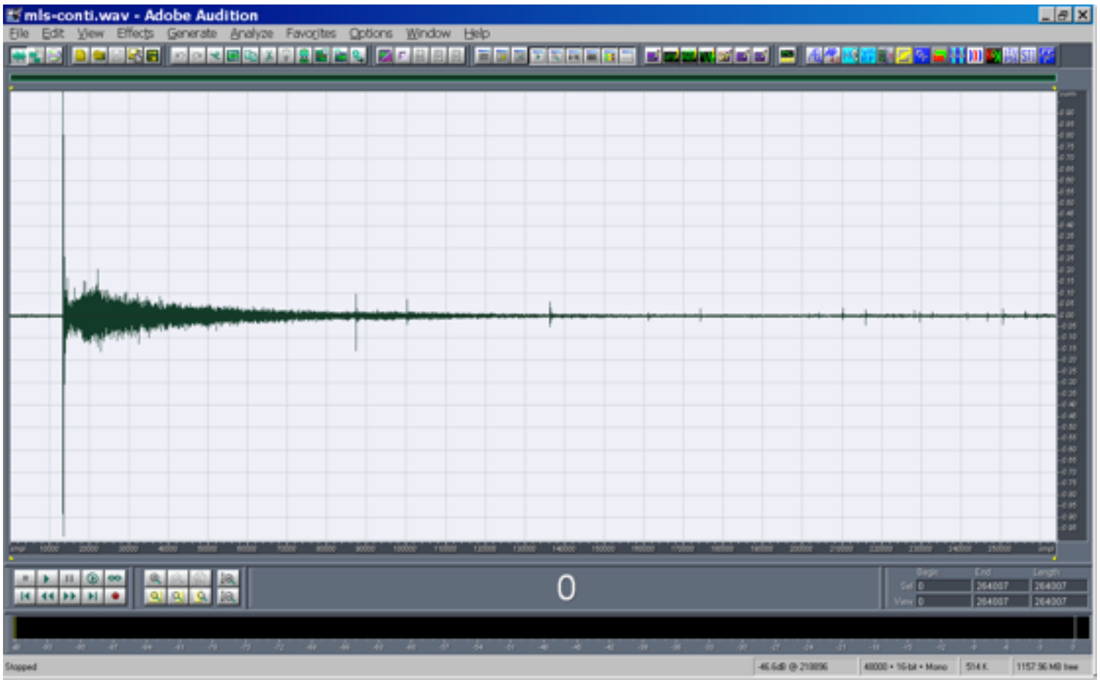
The Hadamard transform is very efficient computationally because it is fast, but does not well in a non linear system. Another advantage of the Hadamard transform is that the signal to noise ratio improves a lot with MLS method. The two big advantages are:

- ► Little computational load for performing the deconvolution.
- ► Very good improvement in signal to noise ratio.

The main disadvantage of MLS are the bad artifacts due to non linearity of the loudspeakers, and its sensitivity to clock mismatch between player and recorder.

MLS measurement procedure

We have seen an example of measurement of impulse response in a church with MLS signal, the test signal was played and the system response was recorded simultaneously. The signal to noise ratio was quite good.
However in the last part of the measurement there were artifacts, sounding like echoes, as visible in fig. 11. The artifacts are due to the non-linearities of the loudspeaker, which are wrapped inside the impulse response due to the "circular" nature of the Hadamard transform.



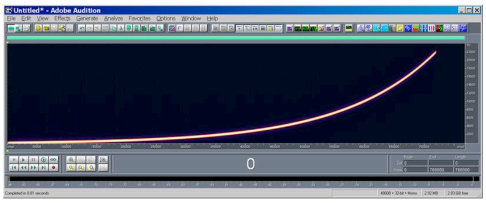
-Fig.11-

Exponential Sine Sweep method

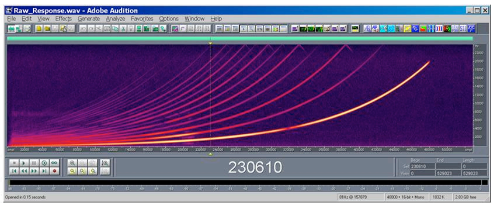
x(t) is a band-limited sinusoidal sweep signal, where frequency is varied exponentially with time, starting at f_1 and ending at f_2.

$$x(t) = \sin\left[\frac{2\cdot\pi\cdot T\cdot f_1}{\ln\left(\frac{f_2}{f_1}\right)}\cdot\left(e^{\frac{t}{T}\cdot\ln\left(\frac{f_2}{f_1}\right)}-1\right)\right]$$

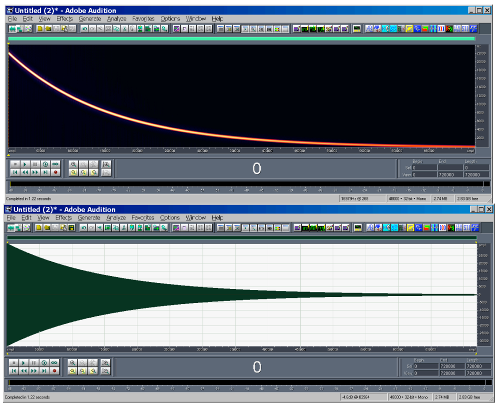
The instantaneous frequency of the test signal x(t) does not increase lineary, but it is increases exponentialy with time, as shown in Fig.11. When we play it through a loudspeaker, there will be some distortion. At every instant the not linear loudspeakers not only radiates the signal at the frequency of the test signal, it will also radiate at 2kHz, 3kHz, 4kHz etc., as shown in Fig.12. This phenomenon is called harmonic distortion and causes the artifacts of MLS.



-Fig.11-

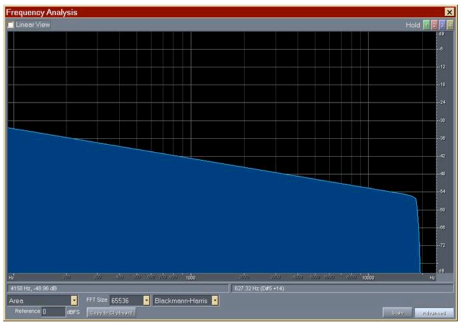


-Fig.12-

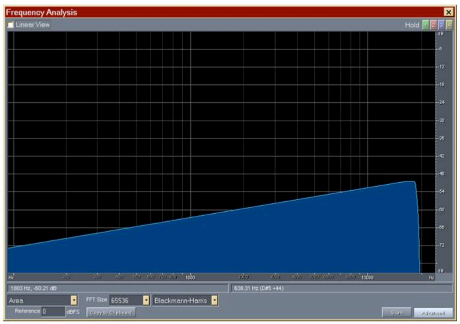


-Fig.13-

To get back the impulse response h(t) we have to define an inverse sweep signal z(t), shown in Fig.13. It is the time reversal of the original sweep signal played backward from high frequency to low frequency. Also the amplitude has been changed, because the spectrum of the ESS was not white, as shown in Fig.14, it was pink. By applying the amplitude modulation we have reversed the slope of the spectrum of the inverse filter z(t), which is shown in Fig.14.



-Fig.14-



-Fig.15-

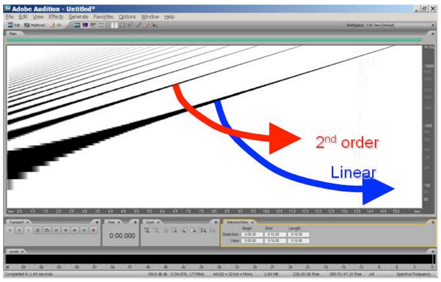
The inverse sweep signal is very loud at high frequency and reduces proportionally towards low frequency. When we convolve the test signal x(t) with inverse filter z(t) the result is flat (in frequency domain) and an almost perfect Dirac's Delta function in time domain.

A measurement made with ESS profits of the pink spectrum of the test signal, and this gives two advantages:

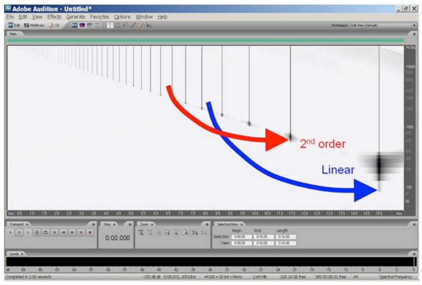
- Much better signal to noise ratio at a low frequency.
- The energy at high frequency is weak, so it does not destroy the tweeter. The tweeter is delicate, the woofer is not.

Deconvolution

After recording the output signal y(t), we get a signal containing the linear response but also the distorted response at second order, third order etc., as shown in Fig.16. Convolving the recorded signal y(t) and the inverse filter z(t), the result of convolution is the rotation of the sonogram because high frequency will not be delayed and low frequency will be delayed. So the linear response, which was stretched over time, packs at a precise time in the rotated frequency-time plane, as shown in Fig.17. After the convolution the last impulse (on the right) is the wanted impulse response (the linear part of the response of our system).



-Fig.16-

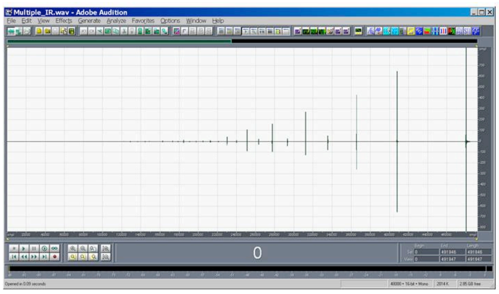


-Fig.17-

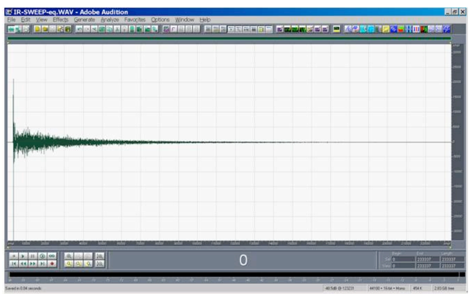
Result of convolution

Coming back from the sonogram in Fig.17 to the normal waveform shown in Fig. 18, there are a number of impulse responses. The last impulse response on the right is the linear impulse response which we want to measure. All the other peaks are the distortion products. These peaks contaminate the impulse response, so we must get rid of them to avoid any contamination with the non linear part of the transfer function of the system. Non-linearity of the loudspeakers is no longer a problem. This is the big advantage of the exponential sine sweep against the MLS. The loudspeaker can be pushed in heavy distortion and that will not be a problem because the distortion is separated in time domain, coming before the linear impulse response. So at this point we can see only the linear impulse response, without any artifacts in the tail, as shown in Fig.18. Considering the same example used for the MLS signal in the church, using the ESS has the following advantages:

- Very good signal to noise ratio, around 20dB better than the MLS signal with the same loudspeaker and the same microphone.
- Any problem of nonlinearity was removed.



-Fig.17-



-Fig.18-

Impulse response from a microphone array

The experimentation described in this section was realized using the Eigenmike^{TM} microphone array produced by MH acoustics [4]. As shown in Figure 19, the Eigenmike^{TM} is a sphere of aluminium (the radius is 42 mm) with 32 high quality capsules placed on its surface; microphones, pre-amplifiers and A/D converters are packed inside the sphere and all the signals are delivered to the audio interface through a digital CAT-6 cable, employing the A-net protocol.

The audio interface is an EMIB Firewire interface; being based on the TCAT DICE II chip, it works with any OS (Windows, OSX and Linux through FFADO). It provides to the user two analogue headphones outputs, one 8-channels ADAT digital output and the word clock ports for syncing with external hardware.

The preamplifier's gain control is operated through MIDI control; we developed a GUI (in Python) for making easy to control the gain in real-time with no latency and no glitches.

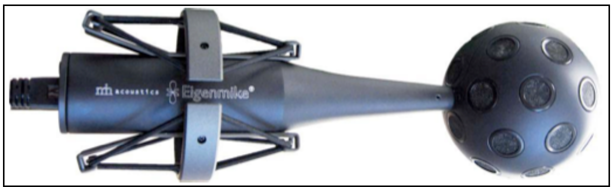


Figure 19 Eigenmike^{TM} probe

Our idea is to derive a set of 32 directive virtual microphones in the directions of the capsules employing a set of digital filters. In our case the *M=32* signals coming from the capsules need to be converted in *V=32* signals yielding the desired virtual directive microphones: so we need a bank of *M*×*V* filters. As always, we prefer FIR filters.

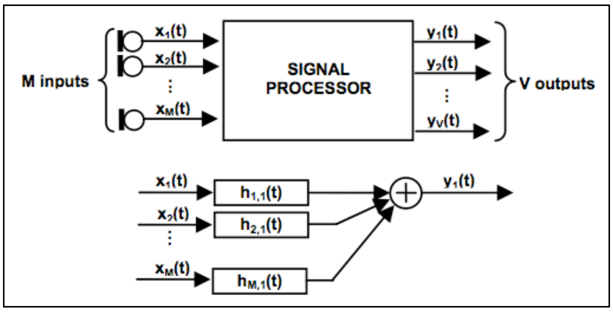


Figure 10 Scheme of the signal processing

Assuming $\boldsymbol{x_m}$ as the input signals of *M* microphones, $\boldsymbol{y_v}$ as the output signals of *V* virtual microphones and $\boldsymbol{h_{m,v}}$ the matrix of filters, the processed signals can be expressed as:

$$y_v(t) = \sum_{m=1}^{M} x_m(t) * h_{m,v}(t) \qquad \textbf{(1.)}$$

Where * denotes convolution, and hence each virtual microphone signal is obtained summing the results of the convolutions of the *M* inputs with a set of *M* proper FIR filters.

Most of the techniques employed for deriving filter banks for generating virtual microphones usually operate following one of several complex mathematical theories, based on the solution of the wave equation, often under certain simplifications, assuming that the microphones are ideal and identical. In some implementations the signal of each microphone is processed through a digital filter for compensating its deviation, with a heavier computational load.

In this novel approach no theory is assumed: the set of filters *h* are derived directly from a set of measurements, made inside an anechoic room. A matrix of measured impulse response coefficients ***c*** is formed and the matrix has to be numerically inverted (usually employing some approximate techniques, such as Least Squares plus regularization) [5,6].; in this way the outputs of the microphone array are maximally close to the ideal responses prescribed. This method also inherently corrects for transducer deviations and acoustical artefacts (shielding, diffraction, reflection, etc.).

In order to derive the matrix of filters, a Matlab script was produced. This script employs 2048 samples of each impulse response and it needs as inputs the number of virtual microphones to synthesize, their directivity, their azimuth and elevation. From these inputs, according with the theory and the procedure previously described [8], it is possible to invert the matrix of impulse responses obtaining the matrix of filters to be applied to the capsule signals. The convolution of the FIRs matrix with the 32 signals coming from the capsules of the array should give as outputs the signals of virtual microphones with the desired characteristics. Some experimental results are shown in one of our previous paper [8], showing some of the different directivity patterns obtained. For this research only the 4th order cardioid was used and the polar pattern of its directivity is shown in Figure 9.

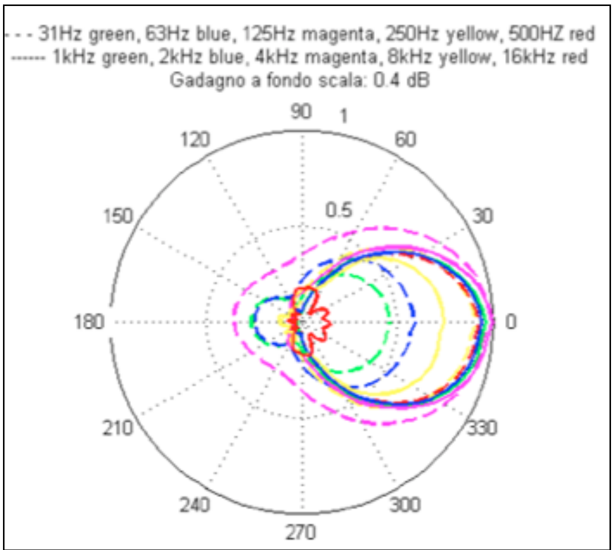


Figure 21 - plot of the 4th order cardioid polar pattern obtained with the multichannel inversion technique

Measurements

A session of measurements was performed inside the Opera House “La Scala” (Milan – Italy).

The Eigenmike^{TM} probe, connected to a laptop, was used for recording the sine-sweeps generated by a dodecahedron; all the convolutions for obtaining the impulse responses were made using Aurora Convolver [10].

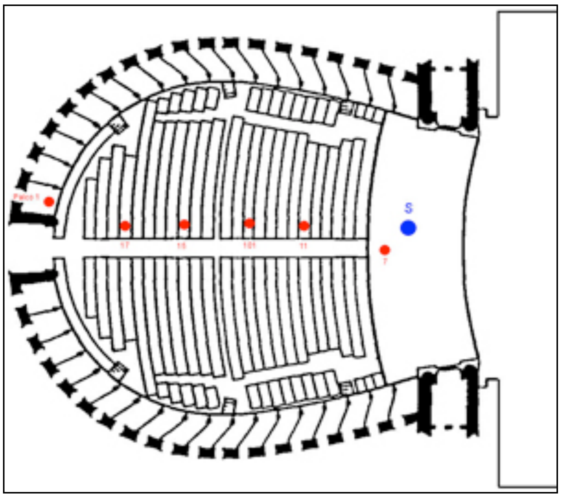


Figure 22 Map of “La Scala” (Milan)

In “La Scala Theatre” the dodecahedron source was placed 1m off the center of the stage and a set of points was chosen for the receiver (Figure 22). In correspondence of the receivers, a panoramic 360° picture was taken with the aim of creating on it the dynamic map of the sound levels.

The post-processing software

All the processing useful for deriving the behavior of the sound inside the theatres is implemented in Matlab [11]. The GUI is made of two main parts: a polar plot of the sound levels in the horizontal plane and a mesh map of the levels plotted on a 360°x180° picture of the theater. They are both dynamic plots, a movie of the sound at the receiver. The processing is done in three steps:

- Import of WAV files (multiple files or one single multichannel file) and selection of the right portion
- FFT processing and choice of the number of frames in which perform the calculation
- Dynamic plot of the results

In the second step the user can choose the frequency band of investigation. For a dynamic representation of the sound at the receiver, multiple, overlapped FFTs are performed. In particular for making a slow-motion analysis of an impulse response, we have to increase the number of time steps by increasing the overlapping of the FFT blocks. In this way a 3D matrix (microphone, frequency band, time step) is created. The user can also choose the type of values being charted: sound pressure, SPL in dB, squared pressure (energy) are the possible choices for visualizing all ranges of values. The plot of the values is performed considering the level derived from every single virtual microphone and representing it on a polar or a map in the corresponding coordinates. A cubic interpolation is necessary for giving continuity to the colour maps, whilst a simple linear interpolation is employed for polar plots.

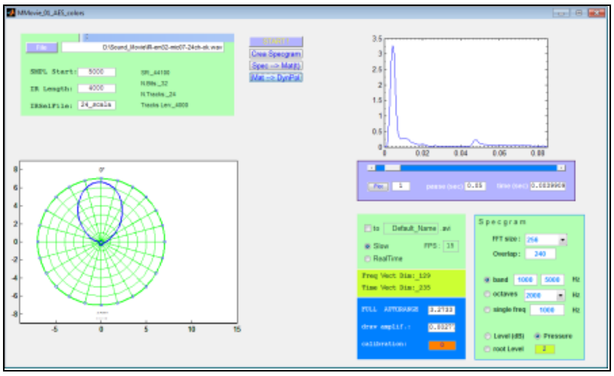


Figure 23 GUI of the post-processing Matlab software

At the end of the processing the user can export a movie (an ".AVI" file), which includes a two-channels impulse response referred to that specific measurement's point.

Analysis of the results

We choose to synthesize 24 virtual microphones regularly spaced on the azimuthal plane and 32 virtual microphones in the direction if the capsules: the first set of microphones is used for a polar plot of the sound at elevation 0°, the second one is useful for creating a color map of the sound on 360° photos.

We can compare the results obtained with the two set of virtual microphones only in the case of measures made with the receiver in the position number 7 (Figure 24) because of the absence of 360°-photos in the other points.

Figure 24 La Scala: receiver in position n.7

As we expected, the direct sound produces in the map a blur that grows in parallel with the increasing of the frequency band of analysis (Figure 25 and Figure 25). Is interesting to notice that in the 500 Hz band, two blurs, one on the stage and one on the back wall, appear before the arrival of the direct sound: this is due to the radiation of the wooden stage under the excitation of the subwoofer while this is playing very low frequencies. The source is close to the receiver so the direct sound covers the first reflection on the floor masking it on the map. The first visible reflection appears after 4 ms coming from the column behind the source, followed by a second reflection on the floor and by a reflection coming from the opposite direction of the first reflection. An interesting behavior is revealed at 57 ms where the sound bounces between the columns of the proscenium arch (Figure 28). Only after these reflections arrives the one coming from the ceiling.

From this dynamic view we can remark some observations:

- The dynamic view of the impulse responses shows as the shape of the theatre refocus the sound in the point of provenience.
- As we expected in the 500 Hz band lobes appears on the map: they are the effect of the diffraction not present at higher frequencies.
- The reflections coming from the walls of the room are displayed on the map with cold colors and this chromatic difference from the direct sound reveals a big amount of absorption in that area of the theatre.

Figure 25 Direct sound in the 500 Hz band

Figure 26 Direct sound in the 4000 Hz band

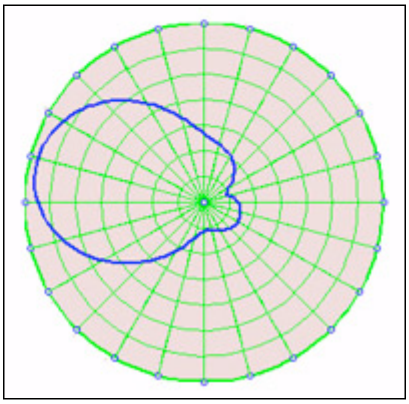


Figure 27 Polar plot of the direct sound in the 4000 Hz band

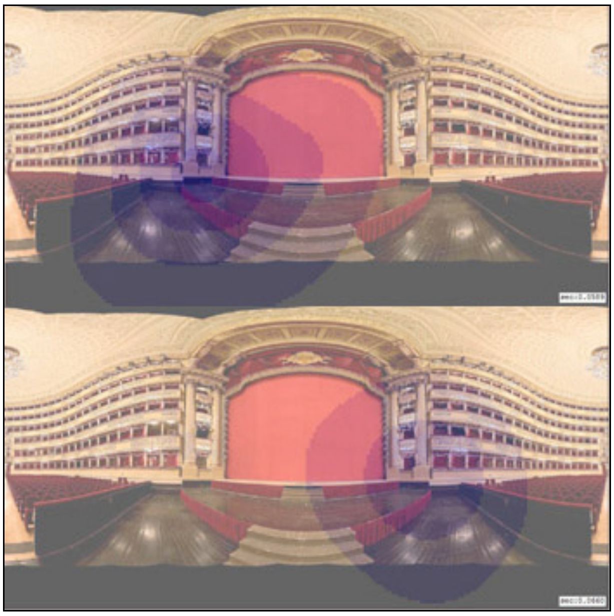


Figure 28 Bounce of the sound between the columns of the proscenium arch