ANALYSIS OF IMPULSE RESPONSE MEASUREMENT SIGNALS USED IN ROOM ACOUSTIC

Snejana Pleshkova-Bekiarska, Stanislav Filipov

Department of Telecommunications,
Technical University - Sofia, Kliment Ohridski 8, Sofia

snegpl@tu-sofia.com

Abstract

A test measurement signals are used to examine the acoustic ambience in open acoustic spaces where combined noise sources exist. Using different test signals and approaches we can measure the reverberation time in any premises. These test signals can be different type and we will observe which signal is the most suitable for exact room. Early acoustic measurement was made by complicated and heavy tube equipment which not always gives the proper results. Also the measurement microphones were not so sensitive for small acoustic fluctuations and artifacts. Generally one modern acoustic measurement system is built up from isotropic sound source called dodecahedron, music interface, measurement amplifier, Omni microphone and acoustic software. However, using this set up we can measure reverberation time and other acoustic parameters via test signals like periodic or white noise, sweep signal and maximum length sequence - MLS. Further researches shows that using the described signals we can reach very high accuracy. In this topic we will look at the modern test signals and compare them.

Keywords: *acoustic measurement, open acoustic space, impulse response, reverberation time, acoustic ambience, test signals, periodic noise, swept-sine, MLS.*

1. INTRODUCTION

The acoustic parameters of any room can be obtained from impulse response of the room via real-time spectrum analysis and measurement of the frequency response. There are many tools for acoustical measurements which can deliver accurate results but usually they are sophisticated software programs with many features. Each software tool used for measurements of acoustic ambience should have the follow main features or options:

- Option for measurement of impulse response
- Option for measurement of frequency response
- Built in spectrum analyzer

The frequency response measurement is a major feature and it's based on the classical Fourier analysis which states that every time signal with a finite energy has a corresponding Fourier transform [1, 2, 7]. In a system analysis we assume that linear time-invariant (LTI) system – Fig.1 is excited with a signal *x(t)* and on output has a signal *y(t).* Both signals *x(t)* and *y(t)* have corresponding Fourier transforms *X(f)* and *Y(f).*

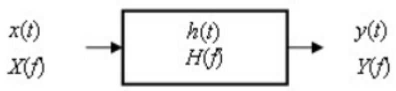

Fig.1. LTI system

The principle of the impulse response measurement is the same as in the Fourier analyzer that we will check later. Only difference is that the impulse response measurement is a non-real time measurement which we will see during our tests.

The spectrum analyzer is convenient option where you can monitor real time fluctuations of the sound pressure level and other events related to frequency domain.

Looking back in the analogous time we will see that the major technique for acoustic measurement uses some type of periodic noise usually pink or white. The benefits from this are that we receive information for the acoustic ambience in whole audio range which is fully enough to analyze the room in frequency or time domain [4,5]. Of course the results in time domain are more complicated and insist more complex calculations where you can easily make a mistake and compromised the results. That's why during the 90's were developed new more precise digital test signals to help downsizing the amount of data and to improve efficiency. Generally two new signals for measurement were implemented – Swept-sine and MLS excitation.

Of course the good old one periodic noise signal was not forgotten. The researches show that he can be adjusted to perfection and more options regarding his spectrum can be added [9,10]. One of these new options is the possibility to generate periodic noise with speech spectrum which is used to measure any premises like open acoustic spaces, offices, rail way terminals and many more.

Let's get back in the digital era and impulse response and to make the conclusion that there are three main types of measurement excitation:

1) Impulse Response Measurement with Periodic Noise Excitation
2) Impulse Response Measurement with Logaritmic Swept-sine technique
3) Impulse Response Measurement with MLS Excitation
4) Impulse Response Measurement with IRS technique.

2. IMPULSE RESPONSE THEORY

The relationship between the input and the output of an LTI system, in the frequency domain, can be expressed as:

$$Y(f) = X(f)*H(f), \quad (1)$$

Where the complex function *H(f)* is called a frequency response:

$$H(f) = \frac{Y(f)}{X(f)} = |H(f)|\, e^{j\varphi(f)}, \quad (2)$$

H(f) is termed magnitude response, and *φ(f)* is termed phase response. The frequency response shows how the system changes the magnitude and phase spectrum of an input signal.

The inverse Fourier transform of frequency response *H(f)* is a function *h(t)* and it is a system impulse response. We show it by following reasoning. The product *X(f)*H(f)* has a Fourier pair in the time domain defined by the convolution $x(t)\otimes h(t)$. This convolution is equal to the output signal *y(t):*

$$y(t) = x(t) \otimes h(t) = \int_{-\infty}^{\infty} h(\tau)x(t-\tau)d\tau, \quad (3)$$

Where the function *h(t)* is called impulse response of the system, as it is a system response to an impulse *δ*-function excitation. It is obvious, as by analyzing the convolution $\delta(t)\otimes h(t)$, we get:

$$h(t) = \int_{-\infty}^{\infty} h(\tau)\delta(t-\tau)d\tau, \quad (4)$$

The system frequency response is usually estimated by using the input-output cross-spectrum and the input auto-spectrum. By rewriting the expression for the transfer function in the following form:

$$H(f) = \frac{Y(f)}{X(f)} = \frac{Y(f)X'(f)}{X(f)X'(f)} = \frac{S_{xy}(f)}{S_{xx}(f)}, \quad (5)$$

We can get the frequency response by dividing an input-output cross-spectrum with an input auto-spectrum (star denotes the complex conjugate value). This equation is usually called *H1* estimator. Fourier transform pairs of the cross-spectrum *Sxy(f)* and the input auto-spectrum *Sxx(f)* are the cross-correlation *Rxy(t)* and the auto-correlation *(Rxx(t))*, i.e.

Rxy(t) ⇔ *Sxy(f) – cross-correlation*

Rxx(t) ⇔ *Sxx(f) – auto-correlation*

If the system input has a white spectrum *(Sxx(f)=1)*, then *Rxx(t)=δ(t)*, the impulse response is equal to the input-output cross-correlation.

$$h(t) \approx Rxy(t), \quad (6)$$

if the input has white spectrum.

Using the *H1* estimator for the frequency (and impulse) response estimation is important, as it will be shown that this estimator has good properties in reducing the influence of the noise and distortions..

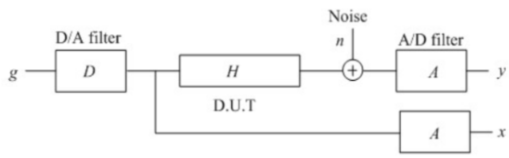

Fig. 2. Measurement setup of dual channel system

Fig.2 shows the measuring system that is typical for acoustical measurements. The computer generate signal – g and after D/A filtering with transfer function D, the signal is applied to the test system that has the transfer function H. Note that H represents the best linear fit of the possible nonlinear transfer function. The noise generator is neglected. The output from the test device (D.U.T), together with the additive system noise n, is acquired by the computer as a discrete signal sequence y. The acquisition process implies the use of an antialiasing filter that has the transfer function A.

3. GENERAL STRUCTURE AND EXCITATIONS TYPES IN DUAL CHANNEL SYSTEMS

3.1. Continuous noise excitation

In a classical Fourier analyzer the excitation is a random noise and the frequency response is estimated by dividing the averaged cross-spectrum *X*Y* with the averaged auto-spectrum *X*X* of *N* input and output discrete sequences *xi* and *yi*. We define the *H1* estimator as:

$$H_e(w) = \frac{\sum_{i=1}^{N} Yi(f)X'(f)}{\sum_{i=1}^{N} Xi(f)X'(f)} = \frac{< Sxy(f) >}{< Sxx(f) >}, \quad (7)$$

Where *He(w)* denotes the estimated frequency response and the brackets < > denote the averaged value. The (7) describes the dual channel system with continuous noise excitation. The *H1* estimator gives a biased estimate of the real transfer function *H(f)*, which is dependent on the noise, distortions and the delay between input and output channel. When only the noise contributes to the bias, the effect of averaging can be expressed by the equation:

$$H_e(f) \cong H(f) + \frac{\sqrt{n} < Ns(f)A(f)X'(f) >}{n < X'(f)X(f) >} \cong$$

$$\cong H(f) + \frac{1}{\sqrt{n}} \frac{< Ns(f)G'(f) >}{< G(f)G'(f) >} \frac{D'(f)}{|D(f)|^2}, \quad (8)$$

Note that signal term is summed coherently, while the stochastic part of the noise is power summed.

The conclusion is that averaging lowers the noise level proportionally with a square root of number of averages, thus improving the measurement S/N by 10log(n). If nonlinear distortions are present, then part of the system noise is coherent with a generated signal and a better measure for the proportionality of the noise + distortion and a number of averages are $1/\gamma \sqrt{n}$, where γ is the input-output coherence function, defined as:

$$\gamma^2 = \frac{|< Sxy(f) >|^2}{< Sxx(f) >< Syy(f) >}, \quad (9)$$

The coherence function is a measure of the proportion of the power in *y* that is due to linear operations on the signal *x*. When estimating the transfer function, the coherence function is a useful check on the quality of data used. The maximum value of coherence is 1. Sometimes we can display the coherence, so it is possible to check the coherence associated with "double channel" measurements. Practically, we must have γ^2 close to 1 to ensure the good estimation, but we must keep in mind that coherence has a sense only if the number of averages is greater than 1.

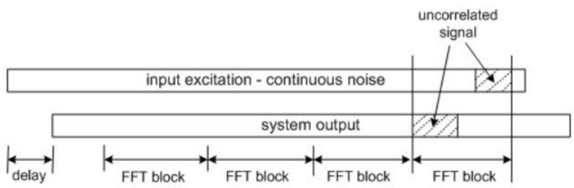

Fig. 3. Uncorrelated estimation in a classical Fourier analyzer

Of course there are some problems in a classical Fourier analyzer with the continuous noise excitation because of:

- The excitation signal does not have constant spectrum at all frequency bins. This gives the frequency selective noise bias. It is high at frequencies where generator spectrum has notches. This resolution bias can be greatly reduced by increasing the number of averaging cycles. It is recommended to make at least 6 spectrum averages and monitor the coherence function.
- In a system with a large delay between the input and the output (Fig.3), i.e. when measuring the loudspeaker in room response, there will be low correlation between measured input and output signals. In some tools it is possible to delay the acquisition of the input channel, so this kind of error can be eliminated. But, if we measure the frequency response in the highly reverberant environment, it is not possible to compensate for all possible delays.

Both problems can be eliminated by using the periodic noise excitation.

3.2. Periodic noise excitation

If the excitation is done with *N* different periodic noise sequences, the frequency response estimator can also be of the form:

$$H_e(f) = \frac{1}{N}\sum_{i=1}^{N}\frac{Yi(f)X'(f)}{Xi(f)X'(f)}, \qquad (10)$$

This type of averaging is called the frequency domain asynchronous averaging. Theoretically it has the same power in reduction of the noise and distortions as the *H1* estimator, but the use of the *H1* estimator is preferred as it enables us to monitor the coherence function.

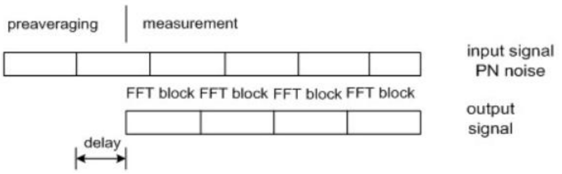

Fig.4. Signal generation during the time domain synchronous averaging process

If the excitation is done with a single periodic sequence, repeated *N* times (Fig.4), the estimator can be of the form:

$$y^{-}(t)\sum_{i=1}^{N} yi(t)$$
$$x^{-}(t)\sum_{i=1}^{N} xi(t) \qquad (11)$$
$$H_e(w) = \frac{{}^{-}Y(f)\,{}^{-}X'(f)}{{}^{-}X(f)\,{}^{-}X'(f)}$$

This type of averaging is called the time domain synchronous averaging. This estimator reduces the system random noise, but it can't reduce nonlinear distortions and the stationary noise that is periodic within the excitation period.

When measuring in reverberant environment the period of the multi-sine must be greater than the reverberation time - T60. The following reasoning can approve this requirement. The room acoustical response has the bandwidth of resonance peaks equal to 2.2/T60. If we choose that a frequency difference between two multi-sine component is less than half of this value, to allow built up of all room resonances, we can conclude that period of the periodic noise have to be equal or greater than the reverberation time. Also, it follows that length of the pre-averaging cycle must be greater or equal to the reverberation time.

3.3. MLS and IRS techniques

The impulse response measurements using the MLS technique were first proposed by Schroeder [!] in 1979 and have been used for more than thirty years. The measurement with MLS technique is based upon the excitation of the acoustical space by a periodic pseudo-random signal having almost the same properties as the white noise. The number of samples of one period of an *m* order MLS signal is: $L = 2m - 1$.

With the MLS technique, the impulse response is obtained by circular cross-correlation between the measured output and the determined input (MLS sequence). Because of the use of circular operations to de-convolve the impulse response, the MLS technique delivers the periodic impulse response *h'[n]* which is related to the linear impulse response by the following equation:

$$h'[n] = \sum_{l=-\infty}^{\infty} h[n + lL], \qquad (12)$$

This equation reflects the major problem of the MLS technique: the time-aliasing error. This error is significant if the length *L* of one period is shorter than the length of the impulse response to be measured.

While the MLS is fighting with the distortion artifacts caused by distortion peaks the IRS technique uses sequence with a *2L* samples period *x[n]* is defined from the corresponding MLS sequence of period *L(mls[n])* by the following relation:

$$x[n] = \begin{cases} mls[n],\ \text{if } n \text{ is even},\ 0 \le n \le 2L \\ -mls[n],\ \text{if } n \text{ is odd},\ 0 < n < 2L \end{cases}, \qquad (13)$$

The deconvolution process is exactly the same as for the MLS technique.

3.4. Impulse response measurement with logaritmic swept-sine technique

The MLS, IRS and Time-Stretched Pulses methods are based on the assumption of LTI systems and cause distortion artifacts to appear in the deconvolved impulse response when this condition is not fulfilled.

The Swept-Sine technique was developed by Farina ([!]) eliminate these limitations. It is based on the idea by using an exponential time-growing frequency sweep. This means that it is possible to simultaneously deconvolve the linear impulse response of the system and to selectively separate each impulse response corresponding to the harmonic distortion orders considered. The harmonic distortions appear prior to the linear impulse response. Therefore, the linear impulse response measured is assured exempt from any non-linearity and, at the same time, the measurement of the harmonic distortion at various orders can be performed.

In mathematical manner the Swept-sine signal is defined as a sine signal *g(t)* with a time varying phase function *φ(t):*

$$g(t) = \sin(2\pi\varphi(t)), \qquad (14)$$

where the frequency of this signal is defined as:

$$f(t) = \frac{d\varphi(t)}{dt}, \qquad (15)$$

If we have logarithmic swept-sine according to Farina the phase function of this signal is:

$$\varphi(t) = \frac{f_1 T}{\ln \frac{f_2}{f_1}}\left(e^{\frac{t}{T}\ln f2/f1} - 1\right), \qquad (16)$$

Where *f1* is the start frequency, *f2* is the end frequency and *T* is the duration of the signal.

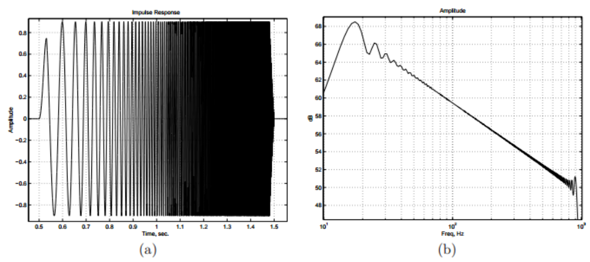

Fig. 5. (a) Time domain representation of logarithmic swept-sine signal, (b) Amplitude representation

4. ACOUSTICAL CHARACTERISTICS, MEASUREMENT SETUP AND COMPARASION OF METHODS AND ALGORITHMS USED IN IMPULSE RESPONSE MEASUREMENTS

4.1. Measurement block diagram

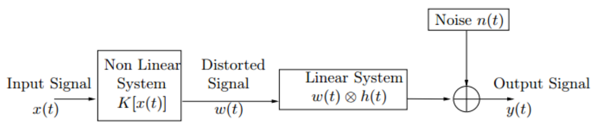

Fig. 6. Nonlinear system (speaker) and LTI measurement system

As for nonlinear system is used Philips AH587 MFB Studio monitor and for LTI is used omnidirectional microphone, Roland professional interface UA-25 and Arta with 64 bit FFT.

Fig. 7. The legendary Philips 22AH587

Fig. 8. Cost effective solution Behringer ECM-8000

4.2. Measurement Setup

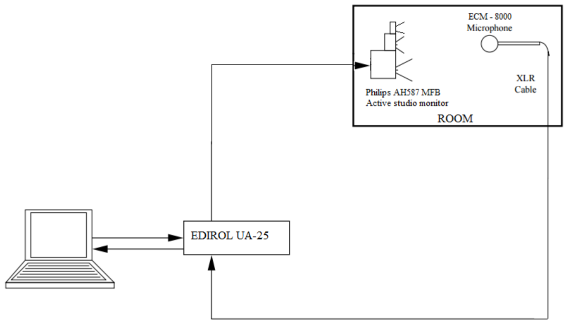

Fig. 9. Measurement setup in casual room

5. RESULTS

Our measurements and tests will be performed in non-anechoic room to see the practical side of the impulse response measurement. According the expression in the previous chapter we will extract the reverberation time from impulse response of all measurements to see what the differences are. Also don't forget that Philips 22AH587 studio monitor is closed box with MFB driver and this can control the room resonances in normal limits.

The test conditions for all measurements will be:

- Sampling frequency – 48 kHz
- Bit resolution – 24bit
- Time constant - 5461.31msec (duration of emitted signal)
- Sequence length – 16k, 64k, 256k, 262k only for MLS
- Number of Averages – 3 (number of times the emitted signal will be send to the monitor)
- Noise level at mic position – MLS-60dBA, Swept-Sine – 80dBA, Periodic noise – 78dBA.

5.1. Periodic noise speech spectrum – sequence length 16k

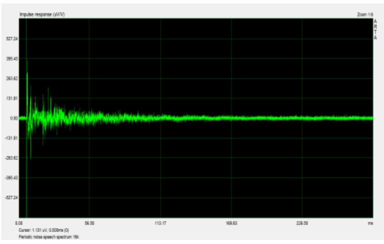

Fig. 10

5.2. Periodic noise speech spectrum – sequence length 64k

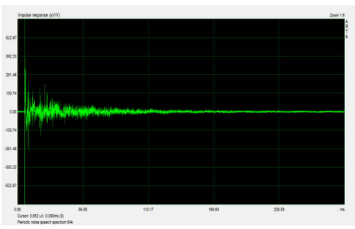

Fig. 11

5.3. Periodic noise speech spectrum – sequence length 256k

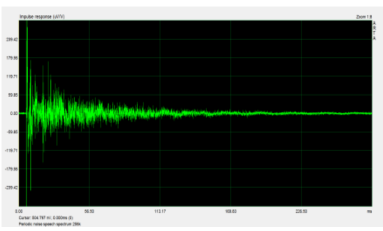

Fig. 12

5.4. MLS – sequence length 16k

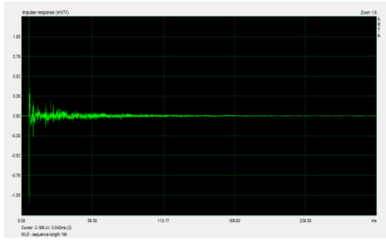

Fig. 13

5.5. MLS – sequence length 65k

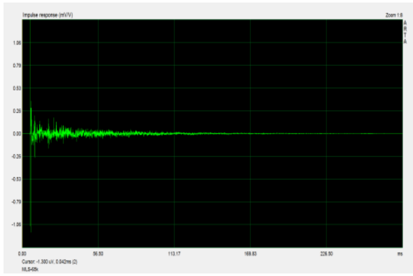

Fig. 14

5.6. MLS – sequence length 262k

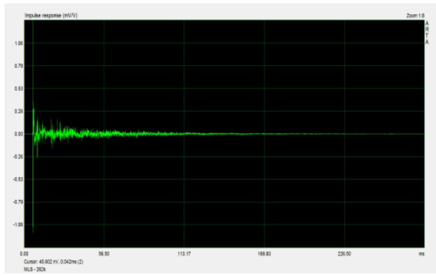

Fig. 15

5.7. Swept-sine – sequence length 16k

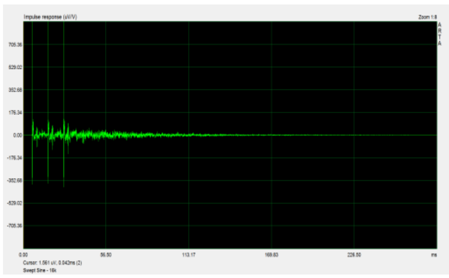

Fig. 16

5.8. Swept-sine – sequence length 64k

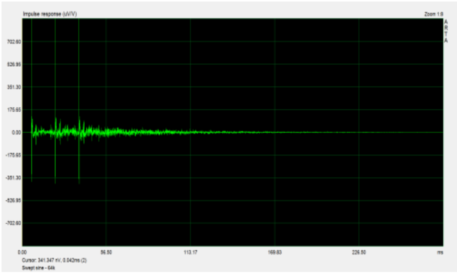

Fig. 17

5.9. Swept-sine – sequence length 256k

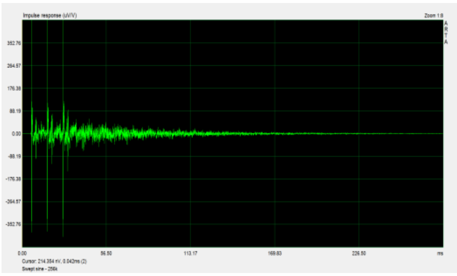

Fig. 18

Analysis

The MLS method seems the hardest method when the measurements have to be performed in a casual non soundproofed room due to its strong defense to all kinds of noise (impulse or background). If we look at Fig.13, 14, 15 we can find that the three measurements are almost identical without big differences. Also the tail of the impulse response compared to the other two methods is very short which shows the protection of any others ambient noises. However, its major drawback is the laggard calibration that has to be done to receive optimal results and the second drawback is in the appearance of distortion peaks due to the inherent non linearity of the measurement system.

That's why if we want to extricate the distortion peaks we can use periodic noise method which avoids the appearance of the distortion peaks. Anyway, the remaining non-linear artifacts can possibly be overlay with the de-convolved "linear" impulse response. The presence of a residue of the excitation signal in the de-convolved impulse response is a result of such superposition problem. This residue can be completely eliminated with a precise calibration of the measurement microphone which we done with pulsar calibrator. However, the main disadvantage in this method comes from its timbre and the high value of the output signal level (see the big tail Fig.10, 11, and 12) needed to mask out the ambient noise. This makes it unusable in occupied rooms.

At last the perfect and complete rejection of the harmonic distortions prior to the "linear" impulse response and the excellent signal-to-noise ratio of the Swept-Sine method make it the best impulse response measurement technique in an non occupied and quiet room. Moreover, unlike the preceding methods, it does not insist a complicated calibration in order to receive good results (no compromise between the signal-to-noise ratio and the superposition of nonlinear artifacts in the room impulse response). Anyway, as for the periodic noise method, the Swept-Sine technique is not recommended for measurements in occupied rooms.

T30 - Reverberation time

The reverberation time – T30 is collected from the measurements of above impulse responses and plotted separately, but why T30 and not T60? In many books is written that the reverberation time is defined as a time interval required for the sound energy to decay 60 dB after the excitation signal has stopped. This is because in ARTA the T30 is defined according to the ISO 3382 and the ISO defines that the T30 is the reverberation time determined from the average slope of the energy decay curve obtained from part of the decay curve between -5dB and -35dB.

5.10. Periodic noise speech spectrum – sequence length 16k

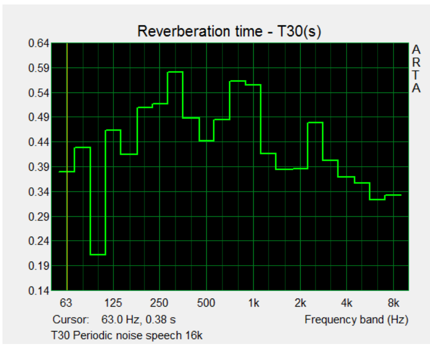

Fig. 19

5.11. Periodic noise speech spectrum – sequence length 64k

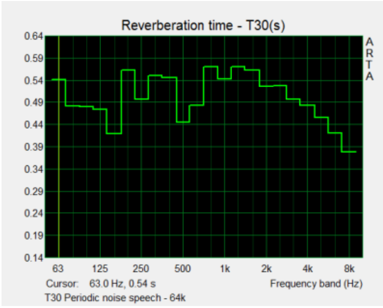

Fig. 20

5.12. Periodic noise speech spectrum – sequence length 256k

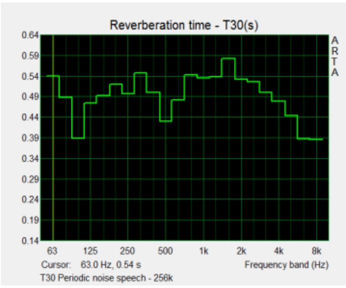

Fig. 21

5.13. MLS – sequence length 16k

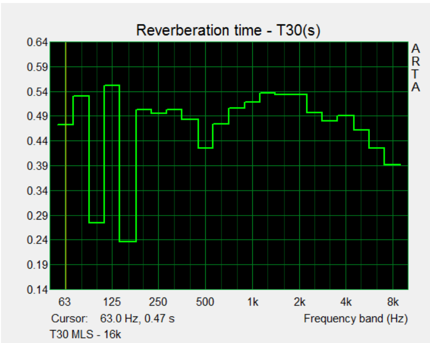

Fig. 22

5.14. MLS – sequence length 64k

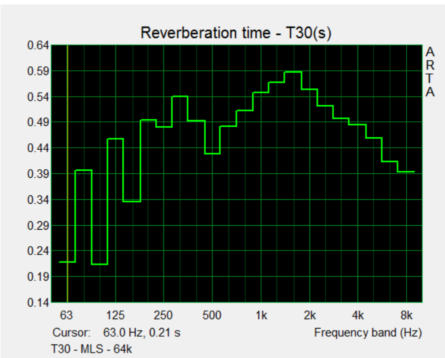

Fig. 23

5.15. MLS – sequence length 262k

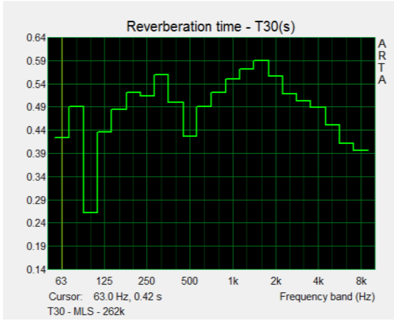

Fig. 24

5.16. Swept-Sine – sequence length 16k

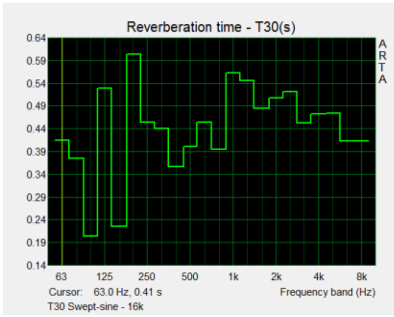

Fig. 25

5.17. Swept-Sine – sequence length 64k

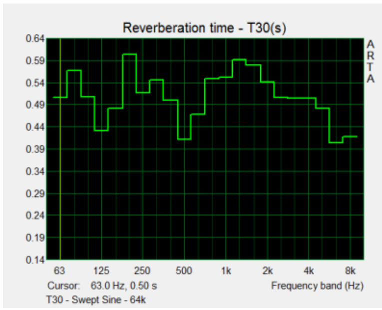

Fig. 26

5.18. Swept-Sine – sequence length 256k

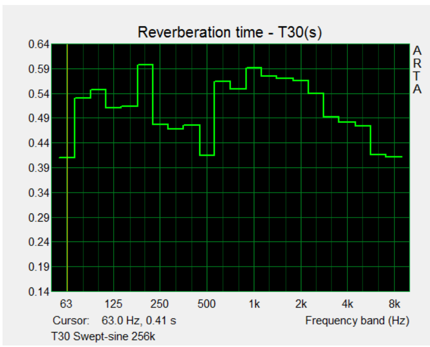

Fig. 27

5.19. T30 comparison - sequence length 16k

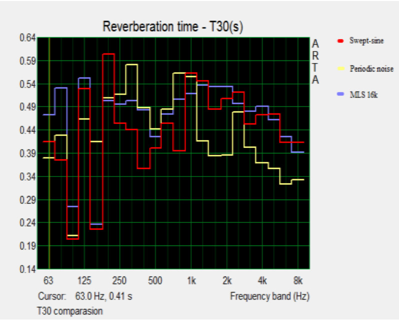

Fig. 28

Analysis

On Fig.19 to Fig.27 are shown test results from T30 extracted from impulse responses. As you can see the T30 from different responses has different shape that's why we plotted all together group by the sequence length. On Fig.28 we see the T30 measured with 16k and it's obvious that the graphs are identical but there are some differences. We can conclude that above 1 kHz the Swept-sine and MLS are similar and the behavior of the curve is due to ray distribution of the high frequency.

Below 1 kHz to around Schroeder frequency we see that MLS and Periodic noise.
It's obvious that the shape of the filter derived from simultaneous masking is expanded compared to this one extracted from forward masking.

Acknowledgment

This paper was supported by Technical University – Sofia inner program to research projects under 161ZF0003-07: "Methods and tools for the study of audio signals".

References

[1] Zwicker E, Fastl H, Psychoacoustics, Springer-Verlag, Berlin, Germany, 1990.

[2] Gelfand, S. A. (2004) Hearing- An Introduction to Psychological and Physiological Acoustics 4th Ed. New York, Marcel Dekker.

[3] Moore, B. C. J. (2004) An Introduction to the Psychology of Hearing, 5th Ed. London, Elsevier Academic PressK. Elissa, "Title of paper if known," unpublished.

[4] Moore, B. C. J. (1986) Frequency Selectivity in Hearing, London, Academic Press.

[5] Moore, B. C. J. (1995) Perceptual Consequences of Cochlear Damage, Oxford, Oxford University Press

[6] Moore, B. C. J. (1998) Cochlear Hearing Loss, London, Whurr Publishers Ltd

[7] Chin-Tuan Tan (2005), Estimates of auditory filter shape using simultaneous and forward notched-noise masking, School of Information Science, Japan

[8] Baker S, Baker RJ. (2006). Auditory filter nonliearity across frequency using simultaneous notched-noise masking. J. Acoust. Soc. Am. 119, 454-462.

[9] Mayer, A. M. 1894. Researches in acoustics. Philosophical Magazine 37, 259-288

[10] Carlos Jurado, David Robledano (2007), Auditory filters at low frequencies: ERB and filter shape, Aalborg University.