FINDING THE AUDIBILITY OF THE TEMPORAL DECAY RATE OF A LOW FREQUENCY ROOM MODE

Andrew Goldberg

Genelec Oy
Sepänkatu 11, 00150 HELSINKI, Finland.
andrew.goldberg@genelec.com

ABSTRACT

A listening test system designed to measure the threshold of audibility of the decay time of low frequency resonances is described. The system employs the Parameter Estimation by Sequential Testing (PEST) technique and the listening test is conducted on calibrated headphones to remove factors associated with the listening environment. Program signal, replay level, and resonance frequency are believed to influence decay time threshold. A trial listening test shows that the system reveals realistic results but the temporal resonance modelling filter requires some adjustment to remove audible non-modal cues. The system is robust but transducer limitations still affect the test at low frequencies and high replay levels. Factors for a future large-scale listening test are refined. Early indications are that temporal decay thresholds rise with reduced frequency and SPL.

1 INTRODUCTION

The aim of this related series of studies is to gain a greater understanding of the sensitivity of human hearing to temporal smearing (longer decay times at certain frequencies due to room or other system resonances) particularly at low frequencies. The hypothesis is that the absolute hearing threshold limits the human capacity to recognise both narrow-band resonances/antiresonances and temporal smearing at low frequencies, and secondly that the sensitivity to hear either magnitude response errors or temporal smearing is further decreased in real room conditions.

The study series firstly describes the design of a suitable listening test for identifying the audibility of resonances, secondly conducts the listening test and analyses the results, and finally tests existing audibility models for applicability to the listening test results. If existing models prove to be inadequate, modifications will be proposed. Initially, the aim is to understand how the human auditory system performs in this respect under ideal listening conditions (headphone listening) by eliminating the room as a variable. Later research may seek to understand how the auditory system performs in typical room listening conditions.

In earlier investigations, upward deviations in the magnitude response have been shown to be more audible than downward ones [1]. It has also been demonstrated that signals with continuous spectra, such as white noise, are more revealing of resonances than signals having discrete spectra, such as solo instrument music, and can be explained by a requirement for the test signal to excite the resonant frequency for it to be audible. Even so, listeners needed extensive training to learn to detect resonances. With the exception of two test frequencies at 85 Hz and 150 Hz, the resonant frequencies in these tests were located above 200 Hz. Other research has also mainly concentrated on mid-to-high frequencies [2].

Focusing more on the time domain, modal equalisation has recently grown in popularity as a research topic for reducing modal decay time [3-5]. The decay time is identified [6] and a modal equalisation technique reduces the pole radii of the modes in the overall transfer function [5].

Whilst objective improvements are readily quantified, recent subjective listening tests suggest that below 100 Hz the threshold of audibility of temporal decay time is high [7]. Those listening tests were conducted at low listening levels and with a sparse frequency resolution, so the question raised is, what is the threshold of audibility of resonances with respect to frequency, level, signal type, and the listening environment? In addition, closely spaced resonances can cause additional effects such as beating in the overall decay pattern, which may alter the audibility threshold.

In summary, objective system equalisation is no longer the limiting factor in improving the listening experience, so the focus has moved towards a listener's ability to perceive an audible sound quality improvement [8]. The subjective research to date shows that the sensitivity of hearing at frequencies below 200 Hz to narrow-band resonances and notches (changes in magnitude response level) or to narrow-band decay (a lengthening of the room decay time at resonant frequencies) has not been studied sufficiently to precisely understand the capability of the human hearing system to these two factors.

This paper opens with a method for generating artificial room modes in audio signals. Possible test signals are classified prior to discussing their reproduction. A description of two methods to ascertain the threshold in the psychometric function is followed by a short description of the listening test system. Recommendations for the full listening test follow a test run of the system to examine its effectiveness.

2 TEMPORAL DECAYS

Resonances in rooms (room modes) are damped and therefore exhibit a decay rate, hence the term temporal decay. Damped resonances with extended decay times are prevalent in listening environments, even in high quality listening rooms. There are often one or more dominant room modes, which are visualised in a waterfall plot or reverberation time curve (Figure 1). It can be seen in this case that a 1/3 octave band T_{60} curve has insufficient resolution to show the fine structure of the temporal decay.

A temporal decay can be simulated in the z-domain transfer function by an allpass complex conjugate pole pair, with poles inside and zeros outside the unit circle,

$$H_{\text{mode}}(z) = \frac{\left(1 + re^{j\theta}z^{-1}\right)\left(1 + re^{-j\theta}z^{-1}\right)}{\left(1 - re^{j\theta}z^{-1}\right)\left(1 - re^{-j\theta}z^{-1}\right)} \tag{1}$$

where r, the pole radius, is defined by the decay time, T_{60},

$$r = e^{\frac{(\log 0.001)/T_{60}}{f_S}} \tag{2}$$

and θ, the pole angle, is defined by the resonance frequency, f_{res},

$$\theta = \frac{2\pi f_{\mathrm{res}}}{f_S} \tag{3}$$

A minimum-phase all pole resonator may also be used and various filter topologies are possible. Figure 2 shows an example of synthetic temporal decays added to artificial test signals. In this paper the test signal is simply feed through the all pass filter described above.

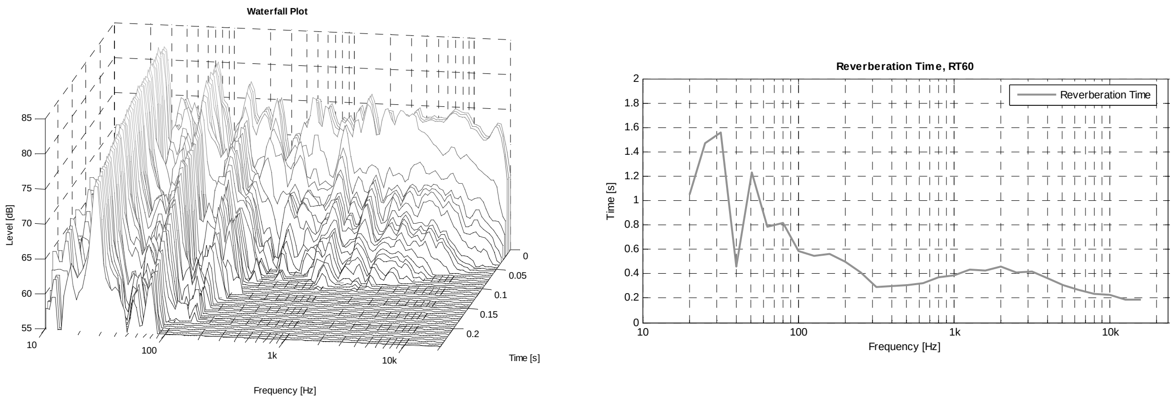

Figure 1. A measurement of a room with insufficient low frequency acoustical treatment. Left: A third-octave smoothed cumulative spectral decay showing strong room modes at 25 Hz and 75 Hz. Right: Third-octave reverberation time.

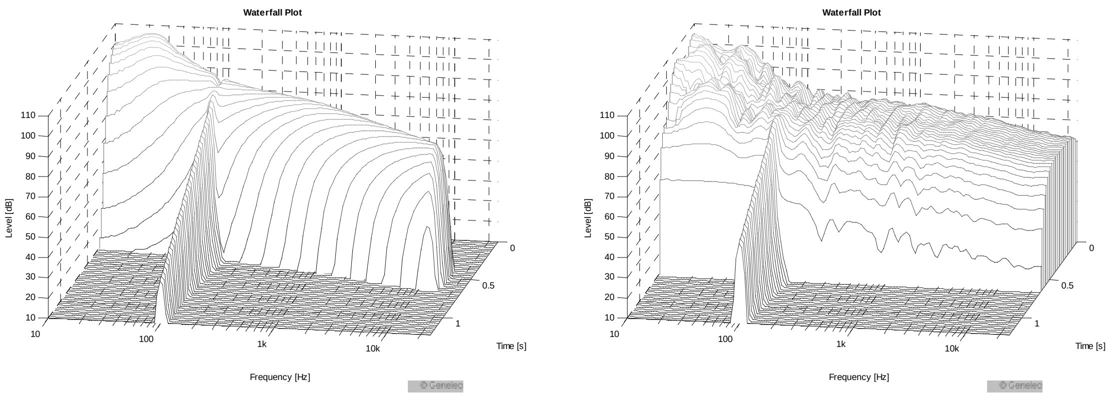

Figure 2. Left: 1/3 octave smoothed cumulative spectral decay of a 0.5 second logarithmic sine sweep (20 – 20000 Hz). It is headphone response equalized with a 1 second temporal decay at 100 Hz added. Right: The same but for 0.5 seconds of pink noise.

3 TRIAL LISTENING TEST

A trial test was conducted as an assessment of the listening test system efficacy in determining the threshold of audibility of temporal decays. Large-scale listening tests to gain a fuller understanding of temporal decay thresholds follow at a later date.

3.1 Trial Listening Test Design

PEST [9] is used to determine the temporal decay threshold of audibility and listening test system has been programmed using Matlab [10] (Figure 3). The factors selected for this listening test are single temporal decays at 32, 50, 80, 125 and 200 Hz, at signal replay levels of 70 and 85 dB SPL. 100 Hz controls replayed at 70 dB SPL are used to check the listeners' performance. These factors are fixed for each PEST run and the decay time (stimulus level) is adjusted for each trial according to PEST rules. The signal is a 10-1000 Hz, 0.5 s, upwards log sweep and the starting decay rate is based on previous informal tests. A minimum step size based on the required accuracy and number of listeners is used to determine the threshold.

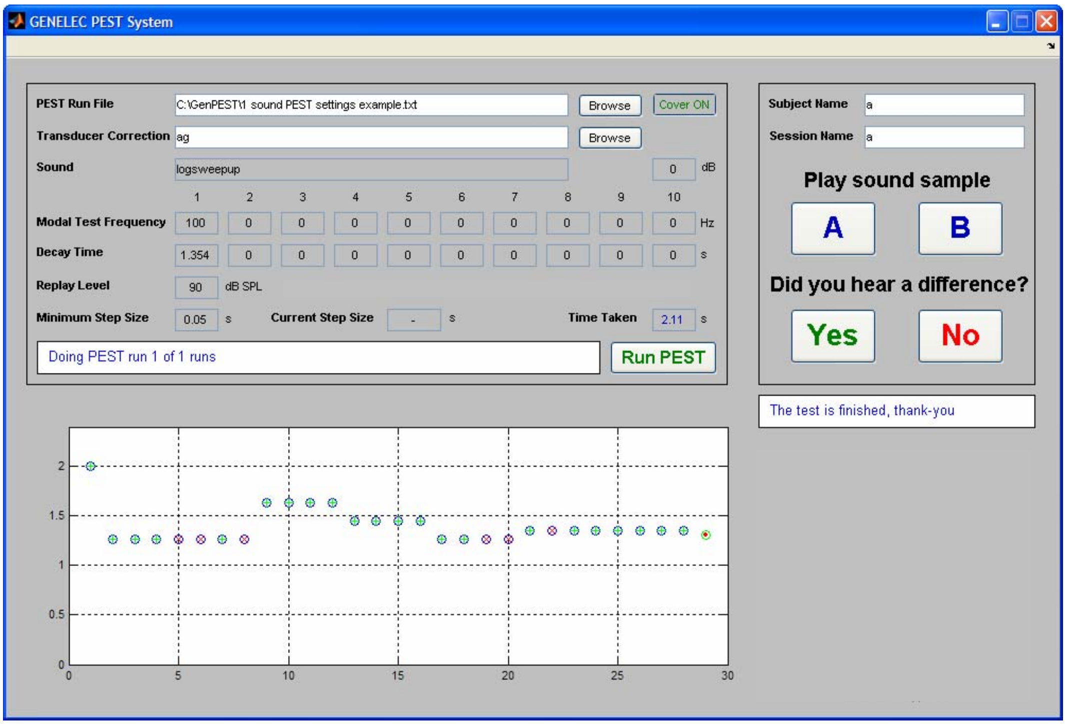

Figure 3. Listening test system graphical user interface. The left side is covered during the listening test to avoid listener bias. A ○ *indicates the stimulus level being tested. A* + *indicates a positive response (audible difference between A & B). An* x *indicates a negative response (no audible difference between A & B). The* ⦿ *indicates the result, which in this example is 1.354 s.*

3.2 Trial Listening Test Results

Six listeners with a median and standard deviation age of 32 ± 13 years, are included in this trial run. Two of the listeners, who are associated with this research program (but not the author), have participated in other listening tests, and all listeners can be described as critical listeners. The least experienced listener reported some known noise induced hearing loss and coincidentally gave the least consistent results. Inconsistent results were also seen for a listener who reported an additional “delay” cue in the test signal during the listening test. In all but one case (the least experienced listener described above), the listeners submitted lower values for the second control signal than the first (100 Hz at 70 dB SPL). The listening tests were conducted in an acoustically well-treated room (low background noise level, T_{60} = 0.2 s) and took 84 ± 17 minutes, with approximately another 30 minutes for listener instruction and two practice PEST runs. The median number of trials in each PEST run was 42 ± 13 and the median number of trials per listening test was 479 ± 80. The number of times each sample (A and B) was played in a trial was not recorded, however it was observed that some listeners were faster in their decision making than others.

To build up the threshold of audibility curves as a function of resonance frequency, SPL, and signal type, the data for each listener is pooled into the factor categories. Taking the median of the pooled results reveals the threshold of audibility (Figure 4) and standard deviation error bars show the data spread. The general pattern is that lower frequency temporal decays have a higher threshold and spread of results, however above 80 Hz SPL appears not to be a factor. The standardised standard deviation (STD std dev) shows increased data spread at higher frequencies. In addition, the lower SPL results show a higher threshold at lower frequencies. Good intra-listener consistency is seen down to 100 Hz at both SPL's but at lower frequencies there is an increased spread of threshold values. Even with an initial stimulus level of 3 s, half the listeners could not hear the 32 Hz resonance in the 70 dB SPL signal. Some listeners noted audible distortion for the lower frequency decays, which is function of the transducer. Other listeners reported a cue they called "delay before the ringing", which was later traced to the initial phase of the temporal decay. In some cases, listener performance changed during the test when this additional cue was perceived.

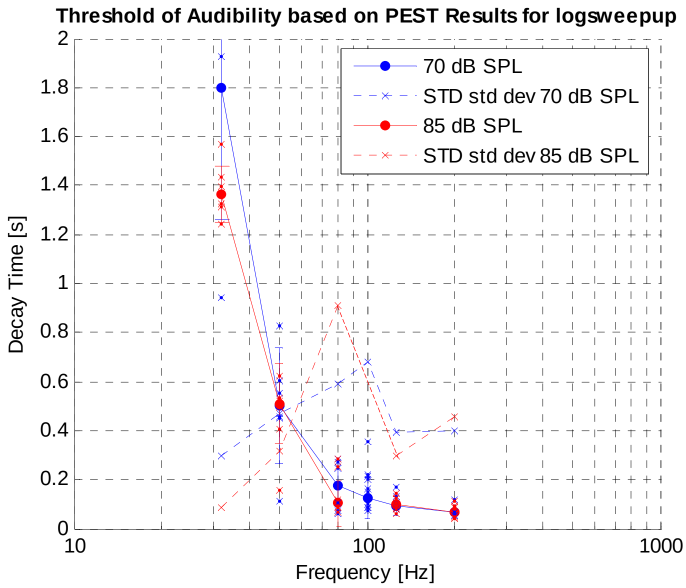

Figure 4 – Threshold of audibility at different SPL's derived from the PEST output (left) and threshold derived from a psychometric function fitted to all the listener's responses pooled together (right).

4 DISCUSSION

4.1 Listening Test System Design and Conditions

The listening test system described above works as realistic results are obtained with experienced listeners. In fact, it is sensitive enough to reveal deficiencies in the signal-resonator combination – see later comments. As has been observed in many other listening tests, practice improves listener performance and reduces listening test time. There is objective evidence, in the form of lower results for the second control signal, that the listeners' audibility performance improved during the test. This indicates that more careful listener selection and a longer training period are required. In addition, despite the use of mandatory breaks, there is some anecdotal and objective evidence of listener fatigue due to the test length, so it appears that 1 hour and breaks every 20 minutes is the maximum tolerable test session time.

The initial stimulus level (temporal decay time) does not affect the threshold result and so the decay time may be reduced. However, it must long enough to be clearly audible at a lower SPL for low frequency decays, but not so long that the PEST run length becomes unnecessarily long. Results show that initial decay times at higher frequencies can be significantly shortened to reduce the overall listening test time. Termination step sizes should be reviewed for the same reason, however this affects accuracy of the results. Listener lapses due to inaudibility of the resonance lengthen the test time, but, as the PEST adaptive algorithm is robust to lapses, it recovers quickly.

The headphones are an open-back design as so exhibit high out-to-in sound leakage. Although not tested formally, there is some evidence that listening room background noise may bias results. Impulsive noise probably has little effect, because the listener may replay the sounds as required, however tonal noise may influence the results and should be minimised.

4.2 Test Signal and Temporal Decay Modelling

It is not desirable to have audible non-modal cues. One such cue is distortion at low frequencies, which some listeners reported. This distortion can be reduced by increasing the sweep speed, starting the sweep at a higher frequency, and/or reducing the SPL. Another audible non-modal cue comes from the initial phase of the resonator being non-zero (270°). When the signal excites the resonance, the temporal decay starts with a negative step and so it interferes with the signal that is still sweeping upwards in frequency. This anomaly explains most of the spread in threshold values as half the listeners perceived a “delay before the ringing” or “click” whereas the others did not, thereby creating two populations of results at lower frequencies. This phase problem in the decay onset must be addressed to remove the audible non-modal cue and hence give more consistent results. An all-pole minimum-phase resonator configured in various topologies is currently under consideration for creating the modal resonance in the test signal. Other test signals will be required for the large-scale listening test.

4.3 Threshold Values

The results indicate that the threshold of the decay time of a temporal decay is higher at lower frequencies. Currently it is suspected that this is related to the increase in absolute hearing threshold with lower frequency, but this will be tested later in the program of studies. Thresholds may also be higher for lower SPL, however fixing the problem with the initial phase of the temporal resonance filter should yield a more definitive answer. Many rooms, including those used for critical listening, have decay times in excess of the audibility thresholds measured in this trial listening test, especially at higher frequencies. Rooms may start to “sing” when sound sources are loud because the room modes become audible, so modal equalisation is probably required for sound systems capable of high SPL, especially if installed in smaller rooms suffering from room modes above 80 Hz.

There is a knee point in both curves below 80 Hz (Figure 4) so an increased density of resonant frequencies is required. The addition of 63 Hz and 40 Hz gives a one-third octave resolution below 80 Hz and a two-thirds octave resolution above 80 Hz. Adding a 400 Hz frequency point extends the data set into the mid band to include higher frequency resonances observed in some listening rooms and to allow comparison against other research which mainly focuses on this region.

The thresholds reported here are low compared to recent research [8] as the test signal with a single resonance is artificial and analytical. Natural signals with complex time and frequency variance may yield higher values due to time- and frequency-domain masking. Conversely, closely spaced multiple resonances can result in beating in the decay curve. This additional cue may lead to lower threshold values. Further listening tests with a variety of signals are required to build a complete low frequency psychoacoustical model.

5 CONCLUSIONS

A listening test system for discovering the audibility of slowly decaying room modes at low frequencies has been described. It is based on the PEST methodology. Headphone equalisation is included, as is the provision for artificial and natural test signals. Artificial temporal decays are added to the test signal and the decay time varied until the threshold of audibility is discovered using the adaptive algorithm. A small-scale listening test shows that the system yields realistic values although, as this was only a trial run, this has not been checked against the literature. More consistent results are expected with increased listener practice and improved temporal decay modelling in the signal processing. The trial listening test revealed deficiencies in the temporal decay model, so additional work is required before commencing large-scale listening tests. The listening test system itself was effective in finding the threshold of audibility of temporal decays and is ready for large-scale listening tests to gather data for testing temporal decay audibility models. Early indications are that low frequency temporal decays have a higher threshold and wider spread of values than high frequency temporal decays. In addition, lower SPL shows a higher threshold at lower frequencies.

6 ACKNOWLEDGEMENTS

The author would like to thank Dr. Aki Mäkivirta for his support and advice throughout the duration of the system development and the writing of this paper. In particular, he is credited with input on the headphone response correction and the chirp test signal algorithm. Prof. Matti Karjalainen has also offered overall guidance throughout this research. The listening test subjects are also thanked. Finally, Genelec Oy is thanked for the provision of technical facilities and support for this research.

BIBLIOGRAPHY

[1] R. Bücklein, "The audibility of frequency response irregularities", *J. Audio Engin. Soc.* **29**, 126-131 (1981).

[2] S. E. Olive, P. L. Schuck, J. G. Ryan, S. L. Sally and M. E. Bonneville, "The detection thresholds of resonances at low frequencies", *J. Audio Engin. Soc.* **45**, 116-127 (1997).

[3] P. Antsalo, M. Karjalainen, A. Mäkivirta and V. Välimäki, "Comparison of modal equalizer design methods", *Proc. of 114th Audio Engin. Soc. Conf.*, preprint 5844 (2003).

[4] M. Karjalainen, P. A. A. Esquef, P. Antsalo, A. Mäkivirta and V. Välimäki,

"Frequency-zooming ARMA modeling of resonant and reverberant systems", *J. Audio Engin. Soc.* **50**, 1012-1029 (2002).

[5] A. Mäkivirta, P. Antsalo, M. Karjalainen, and V. Välimäki, "Modal equalization of loudspeaker-room responses at low frequencies", *J. Audio Engin. Soc.* **51**, 324-343, (2003).

[6] M. Karjalainen, P. Antsalo, A. Mäkivirta, T. Peltonen and V. Välimäki, "Estimation of modal decay parameters from noisy response measurements", *J. Audio Engin. Soc.* **50**, 867-878 (2002).

[7] M. Karjalainen, P. Antsalo, A. Mäkivirta and V. Välimäki, "Perception of temporal decay of low-frequency room modes", *Proc. of 116th Audio Engin. Soc. Conf.* preprint 6083 (2004).

[8] M. Karjalainen, T. Paatero, J. N. Mourjopoulos and P. D. Hatziantoniou, "About room response equalization and dereverberation", *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New York (2005).

[9] M. M. Taylor and C. D. Creelman, "PEST: Efficient estimates on probability functions", *J. Acous. Soc. Am.* **41**, 782-787 (1969).

[10] The MathWorks, "MATLAB (v.7.0.1)", *The MathWorks Inc., Natick,* (2004).