



## EFFECT OF STIMULUS SPEED ERROR ON MEASURED ROOM ACOUSTIC PARAMETERS

PACS: 43.20.Ye

Hak, Constant<sup>1</sup>; Hak, Jan<sup>2</sup>

<sup>1</sup>Technische Universiteit Eindhoven, Department of Building and Architecture, Unit Building Physics and Systems, Laboratorium voor Akoestiek, P.O. Box 513, NL-5600 MB, Eindhoven, The Netherlands; [c.c.j.m.hak@tue.nl](mailto:c.c.j.m.hak@tue.nl)

<sup>2</sup>Acoustics Engineering; Groenling 43-45, NL-5831 MZ, The Netherlands; [j.hak@acoustics-engineering.com](mailto:j.hak@acoustics-engineering.com)

### ABSTRACT

Impulse response measurements based on deconvolution techniques normally require a connection between the stimulus generator and the response recording device. This is inconvenient for long distance measurements, in particular of room acoustics. Playing the excitation signal from a CD generally introduces errors due to a speed mismatch between the CD player and the response recorder. It is investigated which speed differences are to be expected, and how common room acoustic parameters are affected. It is found that MLS stimuli are not really usable. On the other hand, sweeps allow large speed errors and can be used in a broad range of situations. Parameters based on short initial energy intervals are prone to errors when measured with stimuli played too fast rather than too slowly.

### INTRODUCTION

Many room acoustic parameters are derived from the room impulse responses. Examples of such parameters are the reverberation time, which is related to the energy decay rate, the clarity and definition, which both are related to early to late energy ratios, and the speech intelligibility, which is related to the energy modulation transfer characteristics of the impulse response.

Impulse response measurements are often based on convolution techniques [1][2], using maximum length sequences (MLS) [3] or sweeps [4] as stimulus signals. These techniques basically require the room response to be captured synchronously with the generated excitation signal. In some cases it is necessary to play the stimulus signal and capture the response asynchronously. For instance with speech intelligibility measurements in railway stations, the source may be a CD player in an announcer's booth in one city, while the receiver is a microphone connected to a PC on a platform in another city. Normally, the stimulus playback speed will then differ slightly from the response recording speed, which affects details of the resulting impulse response, hence the derived acoustic parameter values. It is investigated how and to what extent this effect occurs.

### THEORY

#### Equivalent excitation pulse

Hereafter all stimuli and (de)convolutions are assumed to be periodic, with a period length exceeding the measured impulse response length. A system impulse response  $h$  is obtained from its response  $y$  to an excitation signal  $s$  through deconvolution:

$$h = y \otimes s \quad (\text{Eq. 1})$$

If stimulus  $s$  is played by an external device at a speed slightly deviating from the correct speed, resulting in stimulus  $s'$ , equation 1 turns into:

$$h' = y' \otimes s \quad (\text{Eq. 2})$$

where

$$y' = s' * h \quad (\text{Eq. 3})$$

and \* denotes a convolution. Substituting equation 3 into equation 2 yields:

$$h' = s' * h \otimes s = (s' \otimes s) * h = d * h \quad (\text{Eq. 4})$$

Equation 4 expresses that the impulse response  $h'$  from a convolution using stimulus  $s'$  instead of  $s$ , equals the response to a certain approximation  $d$  of the ideal impulse  $\delta$  (the Dirac delta function), where:

$$d = s' \otimes s \quad (\text{Eq. 5})$$

In other words,  $d$  is the equivalent excitation pulse of  $s'$ . For the sake of simplicity, hereafter it is assumed that both  $s'$  and  $s$  start at time  $t = 0$  or, in the discrete time domain, at sample 0. Actually,  $s$  and  $s'$  are asynchronous, resulting in an unknown time shift of the measured impulse response, which however is usually irrelevant for the calculation of room acoustic parameters.

If in equation 5  $s' = s$ , then  $d = \delta$  and all energy of  $d$  is lumped at  $t = 0$ . Now, when  $s'$  is time-compressed by slightly increasing the playing speed, the original energy peak of  $d$  is 'smeared out' in negative time. If  $s'$  is time-expanded by slightly decreasing the playing speed, the energy of  $d$  is smeared out in positive time. To show this effect, the equivalent excitation pulse  $d$  is calculated in the discrete time domain for three kinds of stimuli and several stimulus speed errors  $E_s$  (in parts per million, ppm), using the PC mathematics program MATLAB. The results are depicted in Fig. 1, where  $n$  is the sample number.

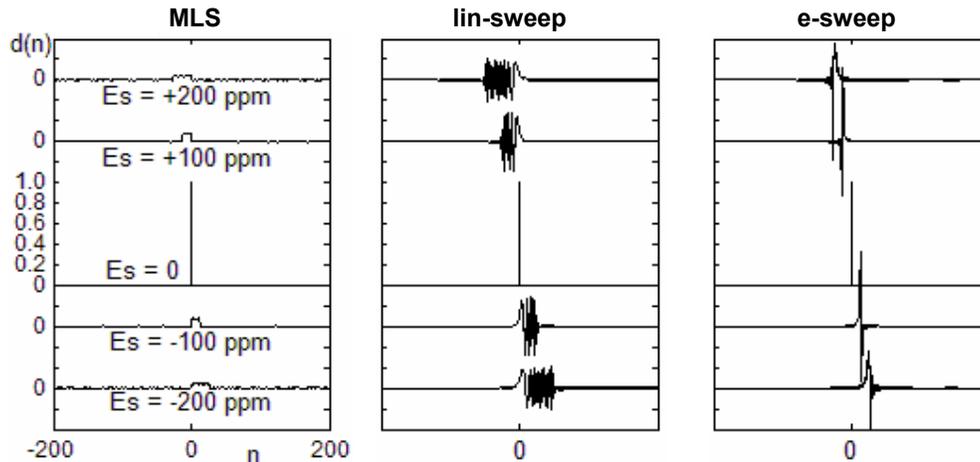


Figure 1.- Equivalent excitation pulses  $d(n)$  of 3 different stimuli at several speed errors  $E_s$ .

The stimuli have nearly equal frequency ranges and are of order  $N = 17$  (corresponding with a period time of 2.7 s at a sample rate  $F_s = 48$  kHz):

1. MLS, period  $L = 2^N - 1$  samples.
2. Linear sweep (lin-sweep), period  $L = 2^N$  samples, sweep range  $f = 0 \dots f_e$  ( $f_e = 0.5 \cdot F_s$ ).
3. Exponential sweep (e-sweep), period  $L = 2^N$  samples, sweep range  $f = f_0 \dots f_e$  ( $f_0 = 2 \cdot 10^{-6} \cdot F_s$ ).

When  $E_s = 0$ , the total energy  $\sum(d^2(n))$  over all samples is contained in sample 0 and equals 1. The total energy does not vary with  $E_s$ , unlike the pulse width, which is roughly proportional to  $E_s$  and  $L$ . If a sweep stimulus is played too slowly, the equivalent excitation pulse  $d$  is a short rising sweep; if played too fast,  $d$  is a short dropping sweep. In all cases, the frequency spectrum of  $d$  is flat, hence the magnitude of the frequency spectrum of  $h'$  equals that of  $h$ .

### Impact on room acoustic impulse responses

To arrive at the measured impulse response using a stimulus with a certain speed error, the corresponding equivalent excitation pulse  $d$  has to be convolved with the room impulse response  $h$ , thereby affecting its shape in several ways:

1. The pulse width of  $d$  causes a measured impulse response to be 'smeared out' in time, and is ideally 1 sample (causing no smearing effect). The lin-sweep shows the largest smearing effect due to speed errors, as observed in the example of Fig. 1. A typical broadband equivalent excitation pulse width, of a linear sweep of period 5 s and a speed error of  $\pm 1000$  ppm, is roughly 10 ms.
2. The signal to noise ratio of  $d$  relates to the SNR of the measured impulse response, and is ideally infinite. With MLS, speed errors cause pulse energy to be transformed into noise energy, thereby strongly decreasing the signal to noise ratio SNR and the impulse response to noise ratio INR [7] which relates to the decay range.
3. A time shift of  $d$  causes the same shift of  $h'$ . Although the broadband version of  $d$  will never show a time shift, with linear sweeps the sweep nature of  $d$  makes band filtering equivalent to time windowing, causing frequency band dependent time shifts.
4. If preaveraging is applied (the process of cyclically recording and averaging multiple room responses to increase the INR), a speed error will cause the periods not to coincide, hence worsen rather than improve the result. Fig. 2 shows an example of a calculated room impulse response, starting from a preaverage over 4 periods and a very large stimulus speed error for the sake of clarity. The repetitive impulse response is useless for parameter calculations, so unless special techniques are applied, preaveraging should not be used in asynchronous measurements.

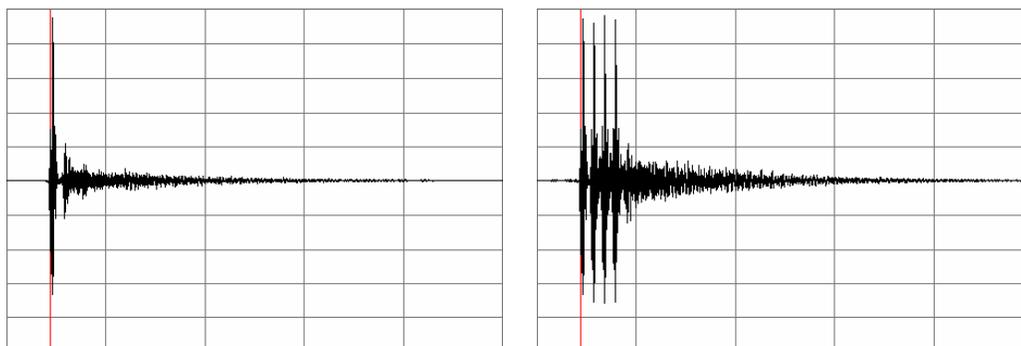


Figure 2.- Effect of preaveraging on room acoustic impulse response.  
Left: not preaveraged; right: preaveraged over 4 periods.

### IMPACT ON ROOM ACOUSTIC PARAMETERS

#### Simulation

To gain insight in the effect of a speed error on room acoustic parameters, such as those mentioned in [2], [5] and [6], the equivalent excitation pulses of several stimuli are first convolved with a simple room impulse response, modeled by an exponentially decaying pink noise signal with 'reverberation time' 2 s and without background noise. From the resulting impulse response, the room acoustic parameters are then calculated. All necessary operations, i.e. the stretching and compression of  $s$  to get  $s'$ , the deconvolution of  $s$  and  $s'$  to get  $d$  (Eq. 5), the convolution of  $d$  and  $h$  to get  $h'$  (Eq. 4) and the calculation of all acoustic parameters are carried out using the PC program DIRAC 4.0 (B&K/Acoustics Engineering Type BZ5449).

Table 1 shows the resulting maximum room acoustic parameters errors over the 125 Hz to 8 kHz octave frequency bands. To enable judgement of the parameter errors, the parameter's JND values are included as well, partly taken from [5], and partly based on estimations (these values are marked by an asterisk \*).

Table 1.- Errors in modeled room acoustic parameters at several stimulus speed errors.

Param	MLS -100 ppm	MLS +100 ppm	lin-sweep -1000 ppm	lin-sweep +1000 ppm	e-sweep -1000 ppm	e-sweep +1000 ppm	JND	Unit
EDT	> 1000	> 500	-0.1	1.1	-0.1	0.1	5	%
T <sub>20</sub>	> 500	> 100	0.0	-0.6	0.1	0.1	10*	%
T <sub>30</sub>	> 200	> 100	0.0	-0.6	0.1	0.1	10*	%
T <sub>s</sub>	> 2000	> 2000	2.8	9.3	-2.7	2.1	10	ms
C <sub>80</sub>	-17	-17	-0.2	-0.7	0.1	0.3	1	dB
D <sub>50</sub>	-0.29	-0.29	-0.01	-0.05	0.04	-0.02	0.05	-
ST <sub>early</sub>	2.3	-1.1	1.2	24	-1.4	2.0	1*	dB
ST <sub>late</sub>	18	17	1.1	23	-1.3	1.9	1*	dB
STI	-0.28	-0.28	< 0.005	< 0.005	< 0.005	< 0.005	0.1*	-

\*estimated

Referring to Table 1 the following can be noted:

- Small MLS speed errors cause large parameter errors, which is caused by the pulse energy being transformed into noise energy. This does not happen with sweeps. It is found that even at MLS speed errors within  $\pm 20$  ppm, most resulting parameter errors significantly exceed the JND values .
- Positive lin-sweep speed errors affect the stage parameters ST<sub>x</sub> much more than negative lin-sweep speed errors. This mainly holds at low frequencies, and can be explained from the down-sweep nature of the equivalent excitation pulse (see Fig. 1): the low frequency energy is delayed, and shifted out of the initial 10 ms time interval stage parameters are based upon.
- More generally, the positive lin-sweep speed error causes larger parameter errors than the negative lin-sweep speed error.
- The speech intelligibility is affected significantly by MLS speed errors, but negligibly by sweep speed errors. This is explained as follows [8][9]. With MLS the increased noise level reduces the modulation in each frequency band, while a sweep speed error can affect the modulation only by its equivalent excitation pulse width, which is normally very short (in this case 10 ms) compared to the shortest speech modulation period of interest (80 ms).

### Measurements

The effect of stimulus speed errors is also investigated through measurements in the concert hall of Eindhoven ‘Muziekcentrum Frits Philips’ using an audio CD with stimuli of known speed errors. For reference, synchronous measurements are carried out as well.

The measurement equipment consisted of the following components:

- *signal source*: CD-player playing an audio CD with preprocessed stimuli;
- *power amplifier*: (Acoustics Engineering);
- *sound source*: omni-directional (B&K - Type 4292);
- *microphone 1*: omni-directional (Schoeps - MK2);
- *microphone 2*: figure-of-eight (Schoeps - MK8);
- *microphone 3*: head simulator (B&K - Type 4128C);
- *input*: USB audio device (Acoustics Engineering - Triton);
- *software*: DIRAC (B&K/Acoustics Engineering - Type BZ5449) running on a laptop PC.

The stage parameters are measured at a distance of 1 m from the sound source, while the other parameters are measured at a listener position 30 m from the sound source.

The results are given in Table 2. The spaciousness parameters LF and IACC are also included, as well as the parameter’s JND values.

Table 2.- Errors in measured room acoustic parameters at several stimulus speed errors.

Param	MLS -100 ppm	MLS +100 ppm	lin-sweep -1000 ppm	lin-sweep +1000 ppm	e-sweep -1000 ppm	e-sweep +1000 ppm	JND	Unit
EDT	-56	> 3000	2	17	2	1	5	%
T <sub>20</sub>	-94	-95	0.7	-2	-3	-1	10*	%
T <sub>30</sub>	-94	> 100	0.5	-3	-2	0.5	10*	%
LF	0.03	0.05	0.01	-0.01	0.01	-0.01	0.05*	-
IACC <sub>0,80</sub>	0.04	0.08	0.02	-0.03	-0.01	-0.11	0.05*	-
IACC <sub>80,+</sub>	-0.02	-0.05	-0.02	-0.02	-0.01	-0.02	0.05*	-
T <sub>s</sub>	> 2000	> 2000	2.9	8.5	3.6	1.9	10	ms
C <sub>80</sub>	-21	-23	-0.2	-0.7	-0.2	-0.3	1	dB
D <sub>50</sub>	-0.66	-0.67	-0.03	-0.17	-0.02	-0.05	0.05	-
ST <sub>early</sub>	4.7	18	-0.3	1.4	1.0	-0.2	1*	dB
ST <sub>late</sub>	26	42	-0.5	1.1	0.9	0.4	1*	dB
STI	-0.20	-0.37	< 0.005	< 0.005	< 0.005	< 0.005	0.1*	-

\*estimated

Referring to Table 2, compared to Table 1 the following can be noted:

- Small MLS speed errors again cause significant parameter errors.
- Positive lin-sweep speed errors affect the stage parameters ST<sub>x</sub> only little more than negative lin-sweep speed errors. On stage the direct sound peak is apparently short enough to stay substantially within the initial 10 ms, even when delayed due to the speed error applied. The errors in these parameters are nevertheless significant, which also holds for EDT and D<sub>50</sub>.
- The spaciousness parameters LF and IACC are basically quite insensitive to speed errors, because the numerator and denominator energies are based on the same time interval and affected in the same way. The error in IACC<sub>0,80</sub> of -0.11 at an e-sweep speed error of +1000 ppm is caused by an excessive value in the 500 Hz octave frequency band.
- The speech intelligibility is affected again significantly by MLS speed errors, but negligibly by sweep speed errors.

### PRACTICAL STIMULUS SPEED ERRORS

Playback and recording speed errors are usually caused by sample rate deviations. Sample clock circuits are usually based on electronic timing components such as crystals or resonators, the latter being cheaper but less accurate. The stimulus speed error E<sub>s</sub> is the deviation of the stimulus playback device sample rate relative to the deviation of the response recorder sample rate. Therefore, in case of opposite signs, the speed errors of these devices magnify each other in E<sub>s</sub>. In case of equal signs, the errors partly compensate each other.

To determine practical sample rate deviations in sound devices, the audio output frequency of 4 desktop PC's, 13 laptop PC's, 10 laptop PC's with USB sound device and 9 portable CD players are measured, while playing a 1 kHz signal. The devices equipped with an audio input are assumed to have approximately the same sample rate errors for recording operations. The measurements are performed using a frequency counter (Philips/Fluke type PM 6685) using a gating time of 10 s. Figure 3 shows the results.

The portable CD players show much larger speed errors than the PC's. Noting that the worst-case speed error E<sub>s</sub> is double the found playback speed errors (because the playback and recording errors may add up), it may be hard to find a CD player that together with a given PC will result in a total absolute stimulus speed error below 100 ppm, while this will become much easier if the speed error restriction can be relaxed to 1000 ppm.

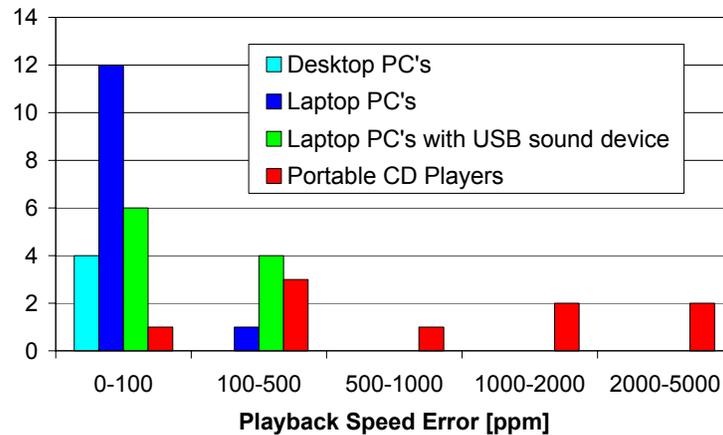


Figure 3.- Distribution of some sound devices over absolute playback speed errors ranges.

## CONCLUSIONS

From the simulations and measurements on the impact of speed errors on asynchronous measurements of room acoustic parameters, the following can be concluded.

Unless special techniques are applied, with asynchronous measurements:

1. Preaveraging is unusable.
2. MLS signals are unusable.
3. STI related parameters can be measured very accurately using sweeps, even with speed errors exceeding  $\pm 1000$  ppm, but very inaccurately using MLS signals, even with speed errors within  $\pm 20$  ppm.
4. For most parameters, sweep signals are usable at stimulus speed errors up to  $\pm 1000$  ppm, currently available with most common sound devices. However, playing back lin-sweep stimuli too fast tends to result in larger parameter errors than playing them back too slowly.
5. For the accurate measurement of parameters that depend on short initial impulse response interval energies, such as the early decay time EDT, the stage parameters  $ST_x$  and the definition  $D_{50}$ , large positive lin-sweep speed errors ( $> 100$  ppm) should be avoided.
6. In general, e-sweeps result in smaller parameter errors than lin-sweeps at the same stimulus speed error.

## REFERENCES

- [1] M.R. Schroeder: Integrated-impulse method for measuring sound decay without using impulses. Journal of the Acoustical Society of America **66** (1979) 497-500
- [2] **ISO 18233:2006** Acoustics - Application of new measurement methods in building and room acoustics
- [3] D.D. Rife, J. Vanderkooy: Transfer-function measurements with maximum length sequences. Journal of the Audio Engineering Society **37** (1989) 419-444
- [4] S. Müller, P. Massarani: Transfer-function measurements with sweeps. Journal of the Audio Engineering Society **49**, No.6 (2001) 443-471
- [5] **ISO/DIS 3382-1:2006 Draft** Acoustics - Measurement of room acoustic parameters – Part 1: Performance rooms
- [6] **IEC 60268-16:2003** Sound system equipment - Part 16: Objective rating of speech intelligibility by speech transmission index
- [7] Acoustics Engineering: Impulse response to Noise Ratio INR. Technical Note 007, <http://www.acoustics-engineering.com/files/TN007.pdf>
- [8] T. Houtgast, H.J.M. Steeneken: The modulation transfer function in room acoustics as a predictor of speech intelligibility. Acustica **28** (1973) 66-73
- [9] M.R. Schroeder: Modulation transfer functions: definition and measurement. Acustica **49** (1981) 179-182