Optimization of the Impulse Response Length: Application to Noisy and Highly Reverberant Rooms*

LAURENT FAIGET AND CLAUDE LEGROS,** *AES Member*

L.A.U.T.M, Université de Toulouse–le Mirail, 31058 Toulouse Cedex 1, France

AND

ROBERT RUIZ, *AES Member*

L.A.R.A, Université de Toulouse–le Mirail, 31058 Toulouse Cedex 1, France

Early-to-late energy ratios are used to predict speech intelligibility. The computation is based on the measured impulse response for a source–microphone combination. When measurement is corrupted by extraneous noise, the upper limit of integration cannot be the total time of acquisition, but a new time limit value defined as the useful length of the impulse response. This limit is validated with the reverberation time computed from the Schroeder backward integral. Results are shown for three highly reverberant rooms.

0 INTRODUCTION

Reverberation time itself cannot describe completely the acoustical perception of a listener in a room. It has to be associated with other objective parameters in order to really predict a subjective feeling. Concerning speech intelligibility, Bradley [1] developed the concept of useful-to-detrimental energy ratios computed from the measured impulse response defining a new function expressed in decibels, called $U(\tau)$, where τ is the separating time limit between early and late energies. This new function is well correlated with intelligibility tests when τ is equal to 50 ms [2] or 80 ms [1]. A new global model (excluding the frequency aspect) has been developed [3], based on the concept of Bradley and modified in order to take into account the characteristics of both the room and the loudspeaker. The separated limit is stated to be 50 ms in the case of highly reverberant rooms. The useful-to-detrimental energy ratio consists now in calculating a well-defined acoustical criterion called definition D_{50}, which is included in the mathematical model. The computation of D_{50} is made from the impulse response measurement. The most often used technique is based on the maximum-length sequence (MLS) generation because its autocorrelation is an impulse. So when applying such a signal to a linear system, the cross correlation of the system output and the original sequence gives the system impulse response. Schroeder has first used this technique for measuring sound decay in concert halls [4]. The amount of computations is lowered when using the fast Hadamard transform to compute the cross correlation [5]. Advantages of this method in terms of distortion immunity were discussed in [6], [7]. Some conclusions were stated to avoid time aliasing by choosing an MLS period long enough or by optimizing the number of averages. Practical considerations are also given by Bradley in [8] to increase the decay range of the impulse response measurement.

But in factory halls the measured impulse responses are often corrupted by extraneous noise. Sometimes the previous requirements to increase the decay range cannot be fulfilled. It can be seen that in such a situation, computation of D_{50} can lead to erroneous values due to the added noise. The reverberation time computed from the backward integrated impulse response [9] is not the same as that measured with analog devices. The influence of noise has already been discussed in [10]–[12]. But these considerations were developed for reverberation time measurement not for the computation of energy ratios.

Therefore this paper discusses a simple and quick way to compute the energy ratio D_{50} in the presence of extraneous noise in the impulse response leading to a temporal limit defining the useful length of the impulse response.

* Manuscript received 1996 October 28; revised 1997 December 9 and 1998 June 30.

** Now with Laboratoire d'Acoustique Appliquée, 06400 Cannes, France.

This limit is studied with an exponential decay with added noise, and the theoretical errors on energy ratios are computed. It is validated by reverberation time measurement, and energy ratios are then computed from noisy impulse responses and cleaned impulse responses. The comparison between model and measurements shows good agreement.

1 THEORETICAL APPROACH

The sound decay in an enclosure has an exponential form when the sound field is supposed to be diffuse. In most real cases, like concert halls, absorption is not the same on all the walls and almost all located in the seating area. In spite of this consideration, it was found [13] that the sound field decay level takes a linear form in most of rooms. Factory halls or large-volume enclosures, which are the rooms of interest in this study, are of the shoe-box type with a more or less homogeneous distribution of absorption. In this first approach, the sound field is assumed to be diffuse. Thus the ideal squared impulse response will be modeled as follows:

$$h^2(t) = E_0 \cdot \mathrm{e}^{-kt} \tag{1}$$

where E_0 is the energy at $t = 0$, and $k/2$ is the damping constant.

1.1 Pure Exponential Decay

In the case of a pure exponential decay, the sound level has a linear form expressed from Eq. (1) as

$$L(t) = -4.34k \cdot t + 10 \log(E_0) \tag{2}$$

where $-4.34k$ is the slope of the decay curve. Considering a dynamic range of 60 dB [14], the reverberation time is

$$T_{60} = \frac{13.82}{k}. \tag{3}$$

It is also possible to obtain the reverberation time from the Schroeder equation [9]. This means that the ensemble average of many squared decays $\langle s^2(t) \rangle$ for a particular loudspeaker–microphone combination is equal to the integration of the squared impulse response,

$$\langle s^2(t) \rangle = N \int_t^{\infty} h^2(\tau)\, \mathrm{d}\tau = N \frac{E_0}{k} \cdot \mathrm{e}^{-kt} \tag{4}$$

where N is a constant proportional to the power-spectral density of the white noise in the frequency range of measure.

Expressing Eq. (4) in decibels leads to

$$L_s(t) = -4.34k \cdot t + 10 \log\left(N \frac{E_0}{k}\right). \tag{5}$$

From Eqs. (2) and (5), slopes of the decay curves are equal and reverberation times too ($T_{60} = 13.82/k$). Only their signal-to-noise ratios differ.

Considering Eq. (4), one can see that the upper limit of integration reaches infinity. In real cases, impulse response acquisition cannot have this value but a finite time called T. Therefore the Schroeder equation expressed in decibels becomes

$$L_T(t) = 10 \log\left(N \frac{E_0}{k}\right) + 10 \log(\mathrm{e}^{-kt} - \mathrm{e}^{-kT}). \tag{6}$$

The slope behavior is given by

$$\frac{\mathrm{d}L_T(t)}{\mathrm{d}t} = -\frac{10}{\ln 10} N \frac{h^2(t)}{\langle s_T^2(t) \rangle}. \tag{7}$$

A plot of Eq. (7) for given values of k and T (Fig. 1) shows a constant behavior for large values of T. For example, if $T = 8$ s and $k = 2$ ($T_{60} \cong 7$ s) one can consider that the slope is constant until $T_{60} \cong 6$ s, while for $T = 3$ s, the slope cannot be taken as a straight line (Fig. 1). So the longer the acquisition time, the smaller is the error in the estimate of the reverberation time. Hence the calculation of reverberation time using $L_T(t)$ and $L_s(t)$ will lead to close values when T is long enough.

1.2 Exponential Decay and Background Noise

A real room impulse response measurement is always corrupted by noise (acoustical and electrical). This effect can be simulated theoretically by adding noise energy to Eq. (1). The noisy squared impulse response can now be written as

$$h_N^2(t) = E_N + E_0 \cdot \mathrm{e}^{-kt} \tag{8}$$

where E_N is the additional noise energy.

Calculation of the reverberation time using a logarithmic plot of $h_N^2(t)$ is always possible. But the dynamic range may not be sufficient and leads to inaccurate val-

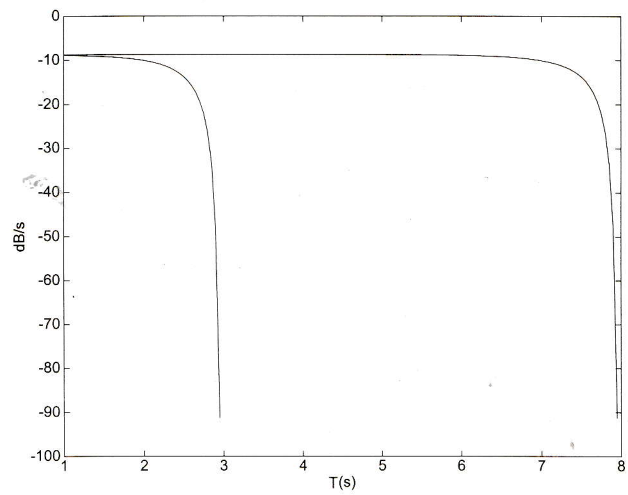


Fig. 1. Slope behavior of pure exponential decay for $k = 2$ from Schroeder's backward integral. Upper curve—$T = 8$ s; lower curve—$T = 3$ s.

ues of T_{60}. In real cases the upper limit of integration is not infinity but a finite value of T. The Schroeder equation becomes

$$\langle s_{N,T}^2(t)\rangle = \langle s_T^2(t)\rangle + NE_N(T - t) . \tag{9}$$

Eq. (9) shows that the noisy backward integral is equal to the pure exponential backward integral plus a noisy window depending on time. Furthermore, this window has a greater effect considering the very first part of the squared impulse response (for small values of t). Eq. (9) leads to

$$L_{N,T}(t) = 10 \log\left[NE_N(T - t) + N\frac{E_0}{k}(\mathrm{e}^{-kt} - \mathrm{e}^{-kT})\right] . \tag{10}$$

Opposite to the case of a pure exponential decay where the acquisition time needs to be long, it is now necessary to compute the backward integral with a smaller time T to make slight differences between $L_{N,T}(t)$ and $L_T(t)$.

The two slopes of $L_T(t)$ and $L_{N,T}(t)$ differ considerably, depending on the noise level. A plot of Eqs. (6) and (10) is given in Fig. 2. Typical values are taken for the main variables, $k = 2$ ($T_{60} \cong 7$ s), $T = 8$ s, $E_0 = 4 \times 10^{-2}$ V^2 [that is, $L_0 = 80$ dB (re 2×10^{-5} Pa)]. E_N is varying from 4×10^{-8} to 4×10^{-3} V^2 [that is, 20–70 dB (re 2×10^{-5} Pa) by 10-dB increments]. N is chosen to be equal to 1 (offset 0 dB). It can be seen from Fig. 2 that when the noise level is high, it is very difficult to have a linear decay over a sufficiently large dynamic range to compute the reverberation time. Furthermore with the highest values of noise, one cannot even have a straight line. For example, the lowest curve labeled 0 dB, which is a noiseless backward integral, gives a T_{60} of about 7 s. This result corresponds to a real reverberant enclosure. For example, for the plot labeled 60 dB, the curve leads to an overvalued T_{60}, which is equal to

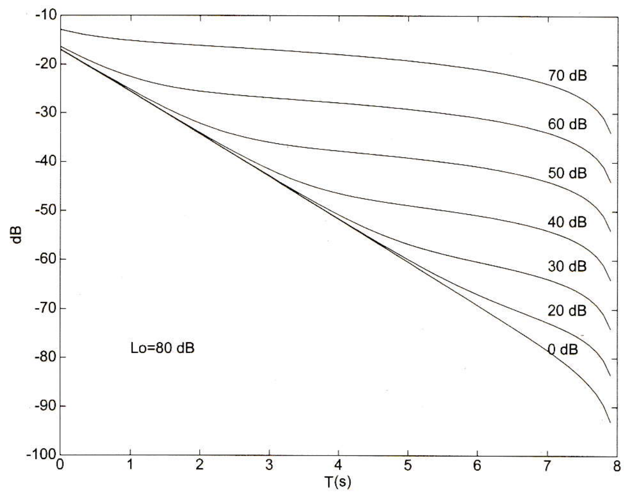


Fig. 2. Logarithmic plot of theoretical Schroeder integral [Eq. (12)] for different signal-to-noise ratios, where signal is equal to $L_0 = 80$ dB and noise is varying from 0 to 70 dB. $T = 8$ s, corresponding to total time duration of theoretical impulse response.

approximately 10 s. Due to the background noise, the term $NE_N(T - t)$ affects the computation of T_{60} considerably.

The slope behavior is

$$\frac{\mathrm{d}L_{N,T}(t)}{\mathrm{d}t} = -\frac{10}{\ln 10} N \frac{h_{N,T}^2(t)}{\langle s_{N,T}^2(t)\rangle} . \tag{11}$$

The influence of the upper limit of integration T has been pointed out with or without the presence of background noise. Without background noise T must be long, and with it T must be short. From theoretical considerations, it can be seen that T must be sufficiently long to avoid time aliasing on the impulse response [6]. A new temporal limit on the logarithmic plot of the squared impulse response has to be defined keeping all the useful information of the impulse response and also reducing the influence of noise.

2 USEFUL DURATION OF THE DECAY CURVE

2.1 Procedure

The time limit T is revised, which is now called t', defined on the logarithmic squared impulse response, which allows one to avoid the last part where noise is preponderant. Usually in most room acoustic software, this limit is chosen by the user [15]. Another approach is taken from ISO 3382-1975 [16], which specifies 5 dB above background noise. The aim is now to have a systematic tool able to make a truncation of the logarithmic decay curve, leading to the useful part of the energy decay. In order to achieve this goal, regression lines based on the computation of correlation coefficients are made on both background noise and decay [17].

2.2 Limits of the Regression Lines

In the following it is supposed that the impulse response acquisition exceeds the useful duration by a sufficiently large time to avoid time aliasing and also to estimate the background noise level. It is necessary to determine four points to delimit the two parts where regression lines can be computed, two points for the decay (P_1 and P_2) and two points for the background noise (P_{N1} and P_{N2}) (Fig. 3).

Point P_1: Direct sound arrival and first high-level reflections must be avoided in the computation of the reverberation time. So according to [18], the mean free path is used to have a systematic determination of P_1. It can be seen that after three times the mean free path value, the diffuse field is reached [19]. Thus in large reverberant rooms this consideration must be adopted. To be in the diffuse field, point P_1 is equal to the time of direct sound arrival added to the time for the sound to travel three times the mean free path.

Point P_2: P_2 can be obtained from the background noise level. As in ISO 3382-1975, P_2 will be the point 5 dB above the mean value of the background noise level.

Point P_{N1}: P_{N1} is chosen just after the decay seems to reach the noise. The requirement is that the acquisition exceed the real decay to have a large noise window.

This point is chosen visually by the user by a mouse click on the plot.

Point P_{N2}: This point is taken at time T, end of acquisition.

The crossing of the two regression lines gives the time t' when the decay curve reaches the background noise.

The method is implemented on a computer. From a measured impulse response, the decay curve is plotted on screen. The user chooses point P_{N1}. The software computes t' and T_{60} from the decay regression line and from the Schroeder equation and, finally, various energy ratios.

It is now necessary to study the influence of the choice of t' on the estimation of the reverberation time.

3 ERROR ON REVERBERATION TIME

The introduction of time t' as the upper limit of integration in the Schroeder equation leads to the following, expressed in decibels:

$$L_{N,t'}(t) = 10 \log \left[NE_N(t' - t) + N\frac{E_0}{k}(\mathrm{e}^{-kt} - \mathrm{e}^{-kt'}) \right]. \tag{12}$$

The value of t' is less than the time of acquisition T. The term $NE_N(t' - t)$ will affect the slope of $L_{N,t'}(t)$ in a different way than $NE_N(T - t)$ in $L_{N,T}(t)$. It is necessary to quantify the errors made on the slopes if we consider T or t'. Because $L_{N,t'}(t)$ is defined from 0 to t' and $L_{N,T}(t)$ from 0 to T, the comparison must be made on the overlap time, that is, from 0 to t'. Two errors are computed. The first, $\boldsymbol{\varepsilon}_T(t)$, is the difference between the noisy backward integral $L_{N,T}(t)$ computed with the upper limit T and the pure exponential backward integral $L_S(t)$ showing the deviation depending on time. The second, $\boldsymbol{\varepsilon}_{t'}(t)$, is the difference between the noisy backward integral $L_{N,t'}(t)$ computed with the upper limit t' and the pure exponential backward integral $L_S(t)$. The errors in decibels can be written

$$\boldsymbol{\varepsilon}_T(t) = 10 \log \left[1 + \frac{E_N(T - t) + \dfrac{E_0}{k} \cdot \mathrm{e}^{-kT}}{\dfrac{E_0}{k} \cdot \mathrm{e}^{-kt}} \right] \tag{13}$$

$$\boldsymbol{\varepsilon}_{t'}(t) = 10 \log \left[1 + \frac{E_N(t' - t) + \dfrac{E_0}{k} \cdot \mathrm{e}^{-kt'}}{\dfrac{E_0}{k} \cdot \mathrm{e}^{-kt}} \right] \tag{14}$$

with $0 < t < t'$. For a given time t, $E_N(T - t)$ is always greater than $E_N(t' - t)$ and e^{-kT} is less than $\mathrm{e}^{-kt'}$. So $\boldsymbol{\varepsilon}_{t'}(t) \leq \boldsymbol{\varepsilon}_T(t)$ for $0 < t < t'$. The reverberation time computed from Eq. (12) will be closer to the theoretical one [see Eq. (5)].

The previous sections show clearly that care must be taken when impulse responses are corrupted by extraneous noise. Two parameters were pointed out in the computation of the reverberation time using the Schroeder equation: the background noise level and the upper limit of integration. Fig. 2 illustrates the influence of the noise on the slope of the decay curve. In fact, a good value of the reverberation time is found when the measurement is noiseless. But this kind of situation is very difficult to obtain in real measurements. Often impulse responses have a noise component. So a way to deal with noise is to change the upper limit of integration of the Schroeder equation. In a sense, changing this limit is like truncating the impulse response and excluding the contribution of the noise. Section 3 explains the method to obtain the limit t' in a systematic way. The theoretical developments show also that without background noise, integrating over T or t' will lead to the same result. So t' will always be used in each case, with or without background noise.

4 ERROR ON ENERGY RATIOS

In room acoustics most criteria come from the measured impulse response. The concept of the early-to-late energy ratio taken from the square impulse response was introduced many years ago [20], [21]. It is useful to describe both speech and music clarity and also intelligibility. The general expression of an energy ratio (ER) is

$$\mathrm{ER} = \frac{\int_{t_0}^{t_1} h^2(t)\,\mathrm{d}t}{\int_{t_2}^{t_3} h^2(t)\,\mathrm{d}t} \tag{15}$$

where t_0, t_1, t_2, and t_3 are temporal limits. Generally speaking, and for most parameters, $t_0 = 0$ and $t_3 = \infty$. Regarding the previous considerations (Sections 2 and 3), attention must be paid to t_3. Integration to infinity is not possible in real cases. So with a finite duration, some errors are made in the computation of the energy ratio [22]. Background noise, like in the calculation of reverberation time, is also a main cause of error. A new temporal limit t' has been introduced, leading to a noise truncation. The aim is now to calculate errors made on

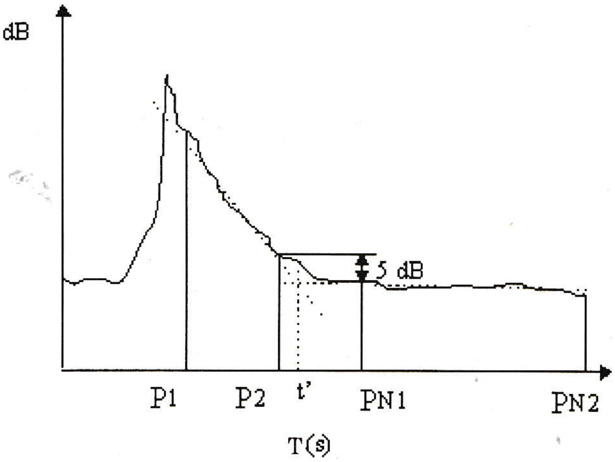


Fig. 3. Logarithmic plot of squared impulse response showing different points used in regression lines. Crossing of two regression lines leads to t', the useful length of the impulse response.

energy ratios integrating over $t_3 = T$ or $t_3 = t'$, compared to the theoretical energy ratio for a pure exponential decay. For convenience, t_0 and t_2 will be equal to zero. The value of t_1 is not very important and can be just a variable, generally denoted by τ. The denominator of Eq. (15) is to be studied to show the influence of the upper limit of integration.

Considering the pure exponential decay, the energy ratio can be expressed as

$$\mathrm{ER} = \frac{\int_0^\tau h^2(t)\,\mathrm{d}t}{\int_0^\infty h^2(t)\,\mathrm{d}t}\,. \tag{16}$$

To simplify the notations, ER will be written in a basic form showing only the influence of the upper limit of integration of the denominator,

$$\mathrm{ER} = \frac{\mathrm{Num}}{\mathrm{Den}_\infty} = 1 - \mathrm{e}^{-k\tau} \tag{17}$$

where Num is the numerator and Den_∞ the denominator of Eq. (16). Now integrating the denominator of Eq. (16) to T yields

$$\mathrm{ER}_T = \frac{\mathrm{Num}}{\mathrm{Den}_T} = \frac{1 - \mathrm{e}^{-k\tau}}{1 - \mathrm{e}^{-kT}}\,. \tag{18}$$

The aim is now to compare the previous ratio to the theoretical one. It can be seen that the difference is found in the denominators Den_∞ and Den_T of Eqs. (17) and (18). A comparison of these two denominators can quantify the increase or decrease of the different energy ratios. The variation can be simply expressed as

$$\alpha = \left(\frac{\mathrm{Den}_\infty}{\mathrm{Den}_T} - 1\right) = \left(\frac{\mathrm{e}^{-kT}}{1 - \mathrm{e}^{-kT}}\right)\,. \tag{19}$$

For large values of T, ER_T will be very close to ER (α is negligible) but with a smaller value. So the upper limit of integration is not a real problem in the case of a pure exponential decay.

Considering now the noisy theoretical squared impulse response $h_N^2(t)$, the energy ratio will take the form

$$\mathrm{ER}_N = \frac{\mathrm{Num}_N}{\mathrm{Den}_{N,\infty}}\,. \tag{20}$$

Integrating the denominator of Eq. (20) over infinity, ER_N will converge to zero. So integrating over T, the time of acquisition, yields

$$\mathrm{ER}_{N,T} = \frac{\mathrm{Num}_N}{\mathrm{Den}_{N,T}} = \frac{E_N\tau + \dfrac{E_0}{k}(1 - \mathrm{e}^{-k\tau})}{E_N T + \dfrac{E_0}{k}(1 - \mathrm{e}^{-kT})}\,. \tag{21}$$

The variation α_T due to the amount of added background noise energy compared to the noiseless ratio ER_T can be estimated in the same way,

$$\alpha_T = \left(\frac{\mathrm{ER}_{N,T}}{\mathrm{ER}_T} - 1\right) = \frac{E_N(\tau - T) + E_N(T\cdot \mathrm{e}^{-k\tau} - \tau \cdot \mathrm{e}^{-kT})}{\left[E_N T + \dfrac{E_0}{k}(1 - \mathrm{e}^{-kT})\right](1 - \mathrm{e}^{-k\tau})}\,. \tag{22}$$

The numerator of Eq. (22) is always negative for large values of T ($T > \tau$). So α_T will always be negative and $\mathrm{ER}_{N,T}$ is less than the theoretical energy ratio integrated over T.

Now integrating the denominator of Eq. (20) to t' yields a new energy ratio, called $\mathrm{ER}_{N,t'}$, which has the same formulation as Eq. (21), but with T replaced by t',

$$\mathrm{ER}_{N,t'} = \frac{\mathrm{Num}_N}{\mathrm{Den}_{N,t'}}\,. \tag{23}$$

The variation $\alpha_{t'}$, made using Eqs. (23) and (18), is

$$\alpha_{t'} = \left(\frac{\mathrm{ER}_{N,t'}}{\mathrm{ER}_T} - 1\right)\,. \tag{24}$$

α_T represents the decrease in the noisy energy ratio $\mathrm{ER}_{N,T}$ computed until T compared to the pure energy ratio ER_T computed until T shows just the influence of noise. In the same manner, $\alpha_{t'}$ represents the decrease in the noisy energy ratio $\mathrm{ER}_{N,t'}$ but now computed until t', compared to the pure energy ratio ER_T computed until T. $\alpha_{t'}$ shows the influence of both the noise and the upper limit of integration. A comparison between α_T and $\alpha_{t'}$ will yield the energy ratio closest to the theoretical one ER_T. It can be seen that only $\mathrm{ER}_{N,T}$ and $\mathrm{ER}_{N,t'}$ must be compared and, more precisely, their two denominators $\mathrm{Den}_{N,T}$ and $\mathrm{Den}_{N,t'}$. It is clear that $\mathrm{Den}_{N,t'}$ is less than $\mathrm{Den}_{N,T}$ (because $t' < T$). In conclusion,

$$\alpha_T < \alpha_{t'}\,. \tag{25}$$

Because α_T and $\alpha_{t'}$ are already negative values,

$$|\alpha_{t'}(\%)| < |\alpha_T(\%)|\,. \tag{26}$$

Integrating the denominator of Eq. (20) over t' rather than T leads to a smaller variation compared to the theoretical energy ratio. For a real squared impulse response, an energy ratio calculated with t' will give a value that can be interpreted as being noiseless, that is, truncated to eliminate the noisy last part, but with no noise subtraction.

5 MEASUREMENTS

Acoustical measurements have been made in three different rooms. Room 1 (R1) and room 2 (R2) are two highly reverberant rooms. Room 3 (R3) is a typical factory hall with a lower reverberation time. R1 and R2 have shoe-box shapes, R3 has a more complex shape.

5.1 Impulse Response Measurements

The impulse responses have been measured at different source-to-microphone distances. The device used for the measurements was a maximum-length-sequence (MLS) equipment. For rooms R1 and R2, the sampling frequency was $f_s = 8000$ Hz and the number of averages was $N = 4$. The MLS generated was a 16th-order one, leading to a time of acquisition $T = 8.192$ s. T was large enough to exceed the real duration of the impulse response, avoiding time aliasing [6]. In R1 and R2 a single loudspeaker was chosen with regard to its directivity properties close to those of human voice. In R1, measurements were made on the diagonal at 2, 4, and 6 m. The first point (2 m) is considered to be in the direct field and the last point (6 m) in the reverberant field. The 4-m point is an intermediate point. Two points were measured in R2 at distances of 7 and 8.5 m. In R1 and R2, loudspeaker and microphone were on axis.

The electroacoustic device was different in R3. The measurements were made using the existing sound reinforcement system composed of the same six loudspeakers. The sampling frequency was $f_s = 16\,000$ Hz and the number of averages was $N = 4$. The order of the MLS was 15, leading to $T = 2.048$ s. Three points were chosen in R3 in order to cover the entire surface. Each point had at least one loudspeaker in direct sight. In the first part of the experiments, impulse responses were measured without noise. Some other measurements have been performed to check the background noise influence. An additional noise source, including a loudspeaker driven by a white noise source, was added to simulate a real machine. Three other impulse responses were measured with noise.

5.2 Reverberation Time Measurements

Table 1 shows the mean of the reverberation times in the three rooms with their standard deviations and the number of measurements n. Reverberation times have been measured in each room using a sound level recorder (Brüel & Kjaer type 2307) and a white noise source (Brüel & Kjaer type 1405). The different T_{60} were computed without filtering over a 30-dB dynamic range. In room R1, the number of measurements ($n = 96$) corresponds to the characterization of the entire volume. In this reverberant room, reverberation times appear to be globally equal everywhere in the room.

6 EXPERIMENTAL RESULTS

6.1 Validation of t' from Measurements of Reverberation Time

From the logarithmic plot of the squared impulse response, the crossing of the regression lines of both the decay and the background noise gives the value of t'.

Fig. 4 shows differences in integration of the Schroeder backward integral over 1) T (total time of acquisition) and 2) t' (useful length of the decay curve) to compute the reverberation time. The plot is made for an impulse response measured in R3 (point 2). The upper curve of Fig. 4 is obtained by integration over $T = 2$ s. It can be seen that no straight line can be found easily in the first part of the curve to compute the reverberation time (the last part is due to background noise). Only the very first part can be considered as a straight line, but here the lack of dynamic range can lead to consistent errors. In contrast, the lower curve of Fig. 4 can be considered as a straight line until a value close to $t' = 0.5$ s. Here the dynamic range is large enough (about 15 dB) to compute the reverberation time.

The reverberation time is computed by means of Schroeder's backward integral with the upper limit of integration equal to t'. But considering Section 3, Eq. (12), one can see that $L_{N,t'} \rightarrow -\infty$ when $t = t'$. $L_{N,t'}$ cannot be approximated by a straight line over $0 < t < t'$. To have a systematic determination of the reverberation time, an algorithm computes a backward correlation coefficient based on the $L_{N,t'}$ curve [17]. First a correlation coefficient is computed from 0 to t' from all the decay points. If the correlation coefficient is greater than -0.99, t' is decreased by 10 ms and another correlation coefficient is computed. These operations are made until the correlation coefficient is less than -0.99, defining a time value denoted by t'' with the highest dynamic range (Fig. 5). So a new reverberation time is found, computed from 0 to t'' from Schroeder's decay curve. Results for the three rooms are shown in Table 2.

In these tables, columns 1 and 2 describe the room

Table 1. Linear analog measured reverberation times.

Room	V (m^3)	T_{60} (s)	Standard Deviation (s)	n
R1	1100	6.1	0.5	96
R2	318	7.2	0.2	16
R3	3000	1.9	0.1	30

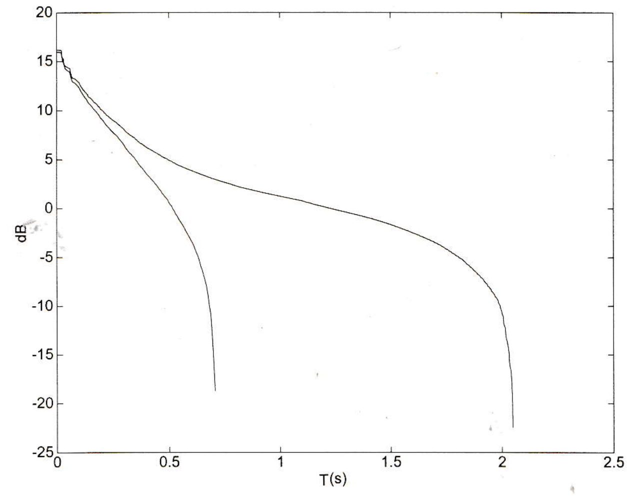


Fig. 4. Backward integrated impulse response (R3, $d = 6.8$ m). Upper curve–plot of $L_{N,T}(t)$. Here no straight line can be found to compute reverberation time due to noise energy. Lower curve–plot of $L_{N,t'}(t)$. A straight line can be defined until $t \approx 0.5$ s, giving a dynamic range of 15 dB.

and the point considered, column 3 is the reverberation time calculated from Schroeder's backward integral from 0 to t'', column 4 is the useful duration of the impulse response t', and column 5 gives the new time value, which allows the computation of the reverberation time from the logarithmic backward integral.

These results must be compared to the analog ones with the corresponding standard deviations. Looking at Table 2 one can see that for the three points of R1, values of $T_{60,\mathrm{Sch}}$ are quite different (Table 1). Although the values are different, the corresponding deviations are not so important and the values are in the confidence interval for the analog measurements.

Results for R2 lead to the same conclusion, but the values are closer. This can be explained because the standard deviation is smaller for the analog measurements than for R1.

R3 is a factory hall filled with machines. The reverberation time is lower than the previous values. Furthermore, the standard deviation of the measurements is small. Only the position $d = 5.5$ m is just out of the confidence interval. Good results are even found when background noise is present in R3 during the measurements. The difference is just in t', which is less than that for the same points without noise.

The results of the different experiments in the three rooms validate the choice of t'. In all the cases the values of the computed reverberation times are equal to the measured ones (within the confidence interval). In addition the influence of noise has been checked, showing also satisfactory results. The introduction of t', whatever the measurement (long or short reverberation, with or without noise), allows a good computation of the reverberation time using Schroeder's backward integral taken from the measured impulse response.

6.2 Real Error in Energy Ratios

As D_{50} is used as the main variable to estimate speech intelligibility, it will also be used to compute energy ratios [3]. It consists in summing the energies from the direct sound and the first reflections arriving within 50 ms and dividing by the total energy,

$$D_{50} = 100 \frac{\int_0^{50} h^2(t)\,\mathrm{d}t}{\int_0^{\infty} h^2(t)\,\mathrm{d}t}. \tag{27}$$

The inequality of Eq. (26) shows that an energy ratio computed from 0 to t' will yield a value that is closer to the theoretical one than an energy ratio computed with T when noise is present. To confirm the developments of Section 5, the measurements described in this section have been used.

In R3 two kinds of impulse responses have been recorded, the first in a classical way, and the second with an additional background noise generated by an external source. From these two types of impulse responses, energy ratios are computed [22]. First the upper limit of integration of the denominator of Eq. (27) is not equal to infinity but equal to $T = 2.048$ s, the total time of acquisition. Then T is replaced by t', the useful length of the impulse response. Results of these calculations are gathered in Table 3, columns 3 and 4, for the three points of interest. It can be seen that energy ratios computed with T give very different values, depending whether or not noise is considered. For example, for the point $d = 4.8$ m, using T leads to a relative error of

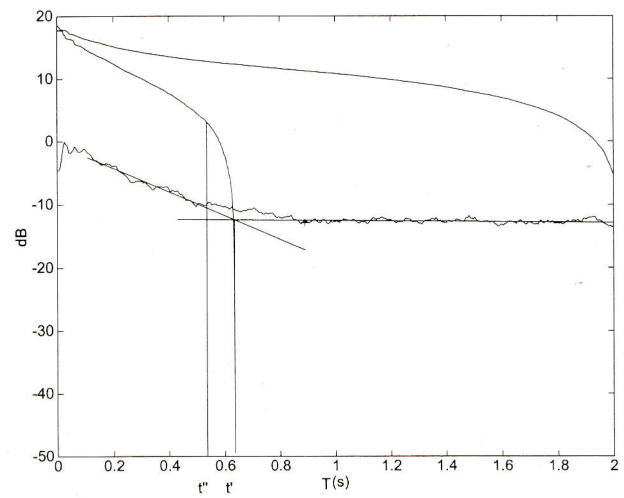


Fig. 5. Final plot of software treatment showing t' and t'' limits leading to maximum dynamic range to compute T_{60} from Schroeder's integral.

Table 2. Results of reverberation time calculation t' and t'' for rooms R1, R2, and R3 using Schroeder's backward integral.

Room	Distance (m)	$T_{60,\mathrm{Sch}}$ (s)	t' (s)	t'' (s)
R1	2	5.9	3.20	3.16
	4	5.6	3.01	2.83
	6	5.7	3.36	3.31
R2	7	7.2	1.85	1.60
	8.5	6.8	1.80	1.53
R3 without noise	4.8	2.0	0.88	0.74
	5.5	2.2	1.04	0.89
	6.8	2.0	0.78	0.63
R3 with added noise	4.8	1.9	0.53	0.39
	5.5	2.0	0.45	0.25
	6.8	2.0	0.52	0.39

Table 3. Results of the calculation of D_{50} (in %) using T, t', and after the cleaning.

Room	Distance (m)	$D_{50}(T)$	$D_{50}(t')$	$D_{50,\mathrm{clean}}(T)$
R3	4.8, with noise	29.5	38.0	38.9
	without noise	38.4	38.5	39.0
	5.5, with noise	18.5	34.1	35.8
	without noise	36.1	36.2	36.3
	6.8, with noise	28.5	38.6	39.7
	without noise	39.4	39.6	39.8
R1	2	50.7	50.9	50.9
	4	26.4	26.7	26.8
	6	22.6	22.7	22.7
R2	7	6.4	7.9	8.8
	8.5	6.1	7.6	8.3

23%. On the other hand, when t' is used, the relative errors decrease. For the same example, the error takes now the value of 1%. The maximum error encountered is for the point d = 5.5 m, where the error is about 6%. Furthermore, a comparison can be made when t' is used instead of T with noiseless measurements. It can be seen that using T or t' with a noiseless impulse response leads to the same results (the maximum error is 0.4% for the point d = 6.8 m).

To ensure the previous results, a numerical cleaning of the impulse responses has been made. The aim is to compare the real measured squared impulse response with the theoretical one. The method is based on the minimization of the squared errors. To avoid the fluctuations of the measured squared impulse responses, the Schroeder equations [Eqs. (4) and (9) are used instead of Eqs. (1) and (8)]. The minimization is made by an algorithm leading to E_0, E_N, and k. So rearranging Eq. (8) gives

$$h(t) = \frac{1}{\sqrt{1 + E_N/E_0 \cdot \mathrm{e}^{-kt}}} h_N(t) \qquad (28)$$

which is the noiseless impulse response. The noisy impulse response is multiplied by a window, which decreases depending on time. The details of this method are described in [23].

This treatment is applied to all the impulse responses measured in R3. Energy ratios are computed with the upper limit of integration equal to T. The results (Table 3, column 5) show that the noiseless values of D_{50} are the same and that the "cleaning" does not really affect a noiseless impulse response. But it can be seen that the error is less when using t' instead of T in every case with noise, as compared to the cleaned impulse response. The maximum error encountered is about 1%. On the other hand, the $D_{50}(t')$ computed from the noisy impulse responses give also satisfying results, the maximum error being 6%. Results for R1 and R2 are also presented in Table 3.

7 CONCLUSION

This paper has shown the importance of the upper limit of integration in both the reverberation time computed by Schroeder's integral method and the energy ratios such as D_{50}. The first part has pointed out the fact that when using the MLS method to compute the impulse response, the period of measurement should be long enough to avoid time aliasing. The theoretical approach leads to the conclusion that the upper limit cannot be simply this total time duration of the impulse response when the measurement is corrupted by extraneous noise, but the time limit must be revised. So a new time limit is defined, denoted by t', which is the crossing of two regression lines computed from the decay and the last part of the impulse response where noise is present. Results obtained show good agreement in the computation of both reverberation time and energy ratio for measurements made with and without noise.

In the case of energy ratios, which is the main concern of this paper, comparison between noiseless measurements and noisy measurements leads to very similar values. The maximum error encountered is about 6%. So the computation of t' is a simple and quick way to obtain the real value of D_{50} just after the measurement, even if a noise component is present in the impulse response. Furthermore, the numerical cleaning of the impulse response proposed in this paper leads also to good results. The maximum error encountered between noiseless and cleaned impulse responses has now decreased to only 1%. But this method, while satisfying, uses more computation time because of the least mean squared algorithm. This method can be used in the laboratory to obtain the results obtained previously with t'.

Now the energy ratio D_{50} can be used in a mathematical model to predict speech intelligibility in noisy and highly reverberant rooms.

8 REFERENCES

[1] J. S. Bradley, "Predictors of Speech Intelligibility in Rooms," *J. Acoust. Soc. Am.*, vol. 80, pp. 837–845 (1986).

[2] J. S. Bradley, "Speech Intelligibility Studies in Classrooms," *J. Acoust. Soc. Am.*, vol. 80, pp. 846–854 (1986).

[3] L. Faiget, R. Ruiz, and C. Legros, "The Influence of Room and Diffusor on Speech Intelligibility," in *Proc. 32nd Czech Conf. on Acoustics* (Prague, 1995 Sept. 23–26), pp. 37–40.

[4] M. R. Schroeder, "Integrated-Impulse Method Measuring Sound Decay without Using Impulses," *J. Acoust. Soc. Am.*, vol. 66, pp. 497–500 (1979).

[5] J. Borish and J. B. Angell, "An Efficient Algorithm for Measuring the Impulse Response Using Pseudorandom Noise," *J. Audio Eng. Soc.*, vol. 31, pp. 478–488 (1983 July/Aug.).

[6] D. D. Rife and J. Vanderkooy, "Transfer-Function Measurement with Maximum-Length Sequences," *J. Audio Eng. Soc.*, vol. 37, pp. 419–444 (1989 June).

[7] J. Vanderkooy, "Aspects of MLS Measuring Systems," *J. Audio Eng. Soc.*, vol. 42, pp. 219–231 (1994 Apr.).

[8] J. S. Bradley, "Optimizing the Decay Range in Room Acoustics Measurements Using Maximum-Length-Sequence Techniques," *J. Audio Eng. Soc.*, vol. 44, pp. 266–273 (1996 Apr.).

[9] M. R. Schroeder, "New Method of Measuring Reverberation Time," *J. Acoust. Soc. Am.*, vol. 37, pp. 409–412 (1965).

[10] W. T. Chu, "Comparison of Reverberation Measurements Using Schroeder's Impulse Method and Decay-Curve Averaging Method," *J. Acoust. Soc. Am.*, vol. 63, pp. 1444–1450 (1978).

[11] K. Bodlund, "On the Use of the Integrated Impulse Response Method for Laboratory Reverberation Measurements," *J. Sound Vib.*, vol. 56, pp. 341–362 (1978).

[12] M. Vorländer and H. Bietz, "Comparison of

Methods for Measuring Reverberation Time," *Acustica*, vol. 80, pp. 205–215 (1994).

[13] M. Barron and L. J. Lee, "Energy Relations in Concert Auditorium. I," *J. Acoust. Soc. Am.*, vol. 84, pp. 618–628 (1988).

[14] L. Beranek, *Acoustics* (American Institute of Physics, New York, 1986).

[15] X. Meynial, J. D. Polack, and G. Dodd, "Comparison between Full-Scale and 1:50 Scale Model Measurements in Théâtre Municipal, Le Mans," *Acta Acustica*, vol. 1, pp. 199–212 (1993).

[16] ISO 3382-1975, "Measurement of the Acoustical Characteristics of Rooms for Speech and Music," International Standards Organization, Geneva, Switzerland (1975).

[17] L. Faiget, R. Ruiz, and C. Legros, "The True Duration of the Impulse Response Used to Estimate Reverberation Time," in *Proc. ICASSP 96*, vol. II (Atlanta, GA, 1996 May 7–10), pp. 913–916.

[18] H. Kuttruff, *Room Acoustics*, 2nd ed. (Applied Science Publishers, London, 1979).

[19] H. Kuttruff, "A Simple Iteration Scheme for the Computation of Decay Constants in Enclosures with Diffusely Reflecting Boundaries," *J. Acoust. Soc. Am.*, vol. 98, pp. 288–293 (1995).

[20] L. Cremer and H. Muller, *Principles and Applications of Room Acoustics* (Applied Science Publishers, London, 1982).

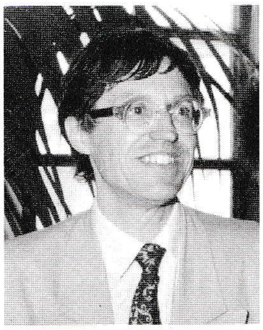
[21] J. P. A. Lochner and J. F. Burger, "The Influence of Reflections on Auditorium Acoustics," *J. Sound Vib.*, vol. 1, pp. 426–454 (1964).

[22] L. Faiget, R. Ruiz, and C. Legros, "Estimation of Impulse Response Length to Compute Room Acoustical Criteria," in *Proc. Forum Acusticum* (Antwerp, Belgium, 1996 Apr. 1–4), p. 148.

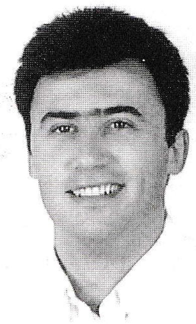
[23] F. Junker, "Présentation d'une méthode de mesure objective de l'intelligibilité de la parole" (in French), contrats d'étude P64L16 and P64L17/1J5682/EP715, Electricité de France, Direction des Etudes et Recherches (1995).

THE AUTHORS

L. Faiget



C. Legros



R. Ruiz

Laurent Faiget was born in Narbonne, France, in 1969. He obtained a bachelor's degree in civil engineering in 1990 from the University of Montpellier and a master's degree in physics applied to building in 1992 from the University Paul Sabatier (Toulouse III). The work described in this paper was part of the research concerning speech intelligibility in noisy and highly reverberant rooms, leading to his Ph.D. obtained in 1997 from the University Paul Sabatier (Toulouse III).

Dr. Faiget's field of research is room acoustics in which he studies the influence of room and sound reinforcement impulse responses on speech intelligibility. He is also involved in sound spatialization and speech perception in difficult environments. He is now working in the Laboratoire de'Acoustique Appliquée in Cannes, France, where he is in charge of applied research and studies in room acoustics and speech intelligibility.

•

Claude Legros was born in Bréhal, France, in 1947. He is Maître de Conférences at the University of Toulouse-le Mirail (Toulouse II). He obtained a Ph.D. from the University of Maine in 1978 for work on automatic vowel recognition. He teaches musical acoustics to music students and statistics to history students at the University of Toulouse II. He also teaches voice properties at the speech therapy school and acoustics at the Conservatoire National des Arts et Métiers and at the Ecole Nationale Supérieure d'Ingénieurs de Constructions Aéronautiques.

Dr. Legros's research work is concentrated on general acoustics, structural acoustics, speech and hearing. His general studies include linear acoustics modeling sound wave propagation in the atmosphere and modeling sound propagation and radiation for jet engines. His work in structural acoustics is on statistical energy analysis to study vibration of satellite panels and particularly the internal loss factor frequency dependence. His speech research involves voice analysis to detect workload influence, and intelligibility in highly reverberant conditions. He is also studying acoustic middle ear muscle reflex to model associated nonlinear phenomena.

Dr. Legros organized the 3rd French Congress on Acoustics in May 1994, and was treasurer of the Société Française d'Acoustique from 1987 to 1997, and is now president of the Société.

•

Robert Ruiz was born in Toulouse, France, in 1964. He received an engineering degree (1987) and a Ph.D. (1991) from the University Paul Sabatier (Toulouse III), both in acoustics. His doctoral dissertation dealt with workload and stress influence on voice acoustical characteristics. In 1994, he joined the University of

Toulouse-le Mirail (Toulouse II) as Maître de Conférences at the Ecole Supérieure d'AudioVisuel where he is head of the sound section. He teaches acoustics applied to audio visual, sound recording and reproducing techniques, and transduction, audio technology, architectural acoustics, and spatial hearing. His research area is the measure of speech perception, quality, and intelligibility. Dr. Ruiz is involved in research on speech analysis to identify stress acoustic indicators and in room and loudspeaker impulse responses processing to study relationships between objective and subjective measurements in room acoustics. He has studied spectral properties of vowels in various stressful situations, robustness of stress acoustic indicators, speech intelligibility models and influence of sound recording on subjective evaluations.

Dr. Ruiz received the Award of Young Researcher from the French Association of Acoustics (SFA) in 1992. He is member of the Audio Engineering Society, the Acoustical Society of America, the European Speech Communication Association and the French Association of Acoustics. His email address is rruiz@univ-tlse2.fr