LETTERS TO THE EDITOR

Room acoustic transition time based on reflection overlap (L)

Cheol-Ho Jeong,[a)] Jonas Brunskog, and Finn Jacobsen
Department of Electrical Engineering, Acoustic Technology, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark

A transition time is defined based on the temporal overlap of reflected pulses in room impulse responses. Assuming specular reflections only, the temporal distance between adjacent reflections, which is proportional to the volume of a room, is compared with the characteristic width of a pulse at time t, which is mainly controlled by the absorption characteristics of the boundary surfaces of the room. Scattering, diffuse reflections, and diffraction, which facilitate the overlapping process, have not been taken into account. Measured impulse responses show that the transition occurs earlier in a room with nonuniform absorption and furniture than in a room that satisfies the underlying assumptions. © *2010 Acoustical Society of America.* [DOI: 10.1121/1.3365251]

PACS number(s): 43.55.Br [NX]   Pages: 2733–2736

I. INTRODUCTION

It is well known that the modal density of an enclosure, dN_f/df, is a quadratic function of the frequency. A similar quadratic relationship between the reflection density, dN_t/dt, and the time after an impulse has been emitted in a room can be derived on the assumption of specular reflections. The two expressions are

$$dN_f/df = 4\pi \frac{V}{c^3} f^2,\qquad(1)$$

$$dN_t/dt = 4\pi \frac{c^3}{V} t^2,\qquad(2)$$

where V is the volume of the room, and c is the speed of sound.[1] The latter expression was first derived by Cremer for rectangular rooms.[2] However, the expression can be extended to rooms of a more general shape.

On the basis of the average spacing between modal frequencies, $\Delta f_{res} = c^3/(4\pi V f^2)$, and the average half-power bandwidth of a mode, $\Delta f_w = 2.2/T$, where T is the reverberation time at the frequency of interest, the Schroeder frequency was proposed as the frequency at which $\Delta f_w = 3\Delta f_{res}$.[3] This leads to a well-known expression for the Schroeder frequency, $f_s = 2000\sqrt{T/V}$ (with T in seconds and V in m^3), above which the modal overlap is sufficient to justify a statistical approach. Note that Δf_{res} is related only to the physical dimensions of the room, whereas Δf_w is exclusively dependent on the damping of the room.

One might expect a phenomenon similar to the modal overlap in the frequency domain to occur also in the time domain; after a certain time has elapsed, the impulse response of the room can be expected to be dominated by overlapping reflected impulses, and it would seem logical to define a transition time by analogy with the Schroeder frequency as the time at which the width of a reflected pulse, Δt_w, equals, say, three times the average spacing between reflections, Δt_{refl}. However, no such definition has been found in the literature. On the other hand, many different approaches to overlapping pulses in the time domain are reported in the literature.

Most authors concerned with the phenomenon have been interested in whether the reflections are audible as discrete reflections or not. After a certain time has elapsed, the reflection density is high and human ears can no longer discriminate one reflection from another because of the limited human auditory time resolution. The maximum audible reflection densities reported in the literature vary between 2000 and 10 000 reflections per second.[4] Based on psycho-acoustic tests, transition times for perceived statistical reverberation have been suggested. Cremer suggested a "beginning time" to statistical reverberation as $t_{st} = \sqrt{(dN_t/dt)_{max} \cdot V/(4\pi c^3)}$ (with all quantities in SI units), and, using a maximum audible reflection density of 2000 reflections per second, derived the expression $t_{st} = 2\sqrt{V}$ (with V in m^3 and the time in ms).[5] Reichardt and Lehmann[6] argued that $\sqrt{V}$ would be more appropriate. Polack proposed a criterion based on ten reflections within the time resolution of the human auditory system (supposed to be 24 ms), which leads to a "mixing time" of $t_{mix} = \sqrt{V}$ (still with V in m^3 and the time in ms),[7] which coincides with the statistical time suggested by Reichardt and Lehmann.[6] This expression was recently referred to by Defrance and Polack, but—somewhat confusingly—presented as an objective measure.[8]

a)Author to whom correspondence should be addressed. Electronic mail: chj@elektro.dtu.dk

The question of whether discrete reflections are audible or not is obviously important in room acoustics, but it would also seem to be useful to have a physical, objective transition time, after which reflected impulses overlap. One possible candidate could be the mixing time. According to Polack, the mixing time is the time at which the energy is equidistributed in the room.[9] Blesser defines the quantity as the time it takes for a sound field to have "no memory of its initial state,"[10] whereas Defrance and Polack defines it as "the time it takes for initially adjacent sound rays to spread uniformly across the room."[8] Referring to personal communication with Polack, Blesser stated that the mixing time is the time it takes for a sound wave to travel three times the mean free path.[10] The mean free path is a quantity known to be $4V/S$, where S is the total surface area of the room.[1] Note that this quantity is independent of the damping of the room.

Yet another time has been suggested by Schroeder. Schroeder himself revisited the Schroeder frequency and derived a relation between the crossover wavelength, $\lambda_c = c/f_s$, and the reverberation distance, $r_c = \sqrt{3/8\pi}\lambda_c$.[11] This quantity indicates a transition from a deterministic to a statistical regime in space. By equating the average temporal distance between adjacent reflections [the reciprocal of the temporal density given by Eq. (2)], $\Delta t_{\text{refl}} = V/(4\pi c^3 t^2)$, with "the echo decay time," $T/\log_e(10^6)$, a "diffuse field time interval," $t_c = \sqrt{\log_e(10^6)V/(4\pi c^3 T)}$, was derived, which, moreover, was shown to be identical to the time it takes for a sound wave to travel the "diffuse field distance" or "critical distance" r_c. Note that t_c does depend on the damping of the room.

The purpose of this study is to define an objective measure of the temporal transition. It is apparent that most of the measures proposed in the literature have not taken account of the damping of the room, although damping can be expected to affect the width of the reflected pulses; therefore a new approach is suggested. The proposed transition time has nothing to do with human auditory perception.

II. TRANSITION TIME BASED ON SURFACE ABSORPTION

Consider an impulse response of a room $h(t)$. It consists of direct sound and successive reflections and is mathematically expressed as follows:

$$h(t) = h_{\text{dir}}(\tau) * \delta(t - \tau_o) + \sum_{i=1}^{\infty} h_i(\tau) * \delta(t - \tau_i). \tag{3}$$

Here, $h_{\text{dir}}(\tau)$ is the direct pulse which arrives at τ_o, and $h_i(\tau)$ is the ith reflected pulse shape arriving at τ_i. The symbol $*$ denotes convolution. Consider a Dirac delta impulse as the directly transmitted pulse, which contains all frequency components. Since high frequency energy is more easily absorbed after hitting surfaces, and accordingly less reflected, the frequency spectrum of the reflected pulse is a low-pass filtered version of the original spectrum. Such frequency-dependent absorption leads to a reduction of the initially flat spectrum at high frequencies, which eventually modifies the width of the original impulse. As a consequence of this low-pass-filtering, reflected pulses are likely to broaden in time, and accordingly after a certain time heavy overlap makes the

fine details of the reverberation unimportant. Such effects are accumulated as time progresses.

An analogy between the time and frequency domain is found: broadened pulse widths at later times and broadened half-power bandwidths at higher frequencies. Such broadenings facilitate the overlapping process in both domains. Being analogous to the frequency domain, Δt_{refl} is associated only with the physical room dimensions, whereas the characteristic width is dependent on the damping of the room.

Since not only the width but also the shape of the pulse change with the acoustic properties of surfaces, a general definition of the pulse width for arbitrarily shaped pulses is needed. The most easy and simple way of defining the pulse duration is the full width at half maximum,[12] the time difference between the points on the initial and decaying part of a pulse where the pressure is half of the peak. It will be called simply the "width" and symbolized as Δt_w in what follows. The width changes with the reflection order, because it changes stepwise whenever the pulse hits surfaces. The number of undergone reflections at the time t is calculated as follows:

$$n_t = \left\lfloor \frac{t}{\text{mean free path}/c} \right\rfloor, \tag{4}$$

where the symbol $\lfloor x \rfloor$ means the floor function.

Instead of a Dirac delta impulse as the direct sound, a Gabor pulse is assumed,[13] a modulated Gaussian pulse, due to its general use in acoustics,[14] as follows:

$$h_{\text{dir}}(t) = e^{-\beta^2 t^2} \cos(2\pi f_o t). \tag{5}$$

Its Fourier transform results in

$$H_{\text{dir}}(f) = \frac{\sqrt{\pi}}{2\beta} \left[e^{-\pi^2[(2\pi f + 2\pi f_o)^2/4\beta^2]} + e^{-\pi^2[(2\pi f - 2\pi f_o)^2/4\beta^2]} \right], \tag{6}$$

where $\beta = \sqrt{2\pi}\Delta f$, f_o is the center frequency, and Δf is the frequency bandwidth. The first reflected pulse from a surface is found by convolution of the direct pulse and the temporal pressure reflection coefficient of a surface material, which is simple multiplication in the frequency domain, as follows:

$$h_1(t) = F^{-1}\{H_{\text{dir}}(f) \cdot r(f)\}. \tag{7}$$

Here, F^{-1} denotes the inverse Fourier transform, and $r(f)$ is the reflection coefficient of a surface expressed in the frequency domain. For simplicity, an infinitely large panel and normal incidence is assumed, which corresponds to

$$r(f) = \frac{Z(f) - \rho_o c}{Z(f) + \rho_o c}, \tag{8}$$

where $Z(f)$ is the normal incidence surface impedance, and ρ_o is the air density. After multiple reflections, for instance, n_t, from the surface, the pulse is expressed as

$$h_{n_t}(t) = F^{-1}(H_{\text{dir}}(f) \cdot r(f)^{n_t}). \tag{9}$$

Figure 1 shows how reflected pulses broaden with the reflection order. Figure 1(a) shows surface impedance of a plastic foam, and Fig. 1(b) shows an example of a Gabor pulse as the direct sound, whose parameters f_o and Δf are 10 and 20 kHz, respectively, and reflected pulses from the plas-
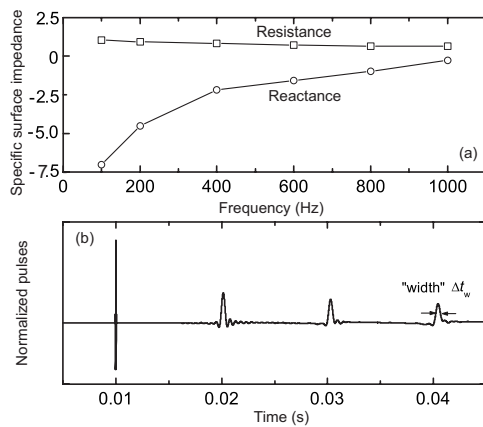
FIG. 1. (a) Specific surface impedance; (b) a series of a direct reflection and broadened reflected pulses up to the third reflection order (even spacing is assumed, and the magnitude is manipulated for clarity).

tic foam. The reflected pulses are broader than the original one. Note that the width of the direct Gabor pulse is 0.01 ms and the widths of the reflected pulses are 0.33, 0.38, and 0.43 ms.

In practice, it is not possible to know reflection coefficients or surface impedance data of all surfaces. What is available in most cases is the reverberation time, from which the average reflection coefficient can be extracted as follows:

$$r(f) = \sqrt{1 - \alpha(f)} = \sqrt{1 - 0.16\frac{V}{T(f) \cdot S}}. \qquad (10)$$

This is only a crude approximation and it holds better for uniformly treated rooms without a dominant absorbing area than otherwise. Once the reflection coefficient is computed from either Eq. (8) or Eq. (10), the reflected pulses are determined by Eq. (9).

Since the width of the Gabor pulse is 0.01 ms, fine resolution is necessary. In order to obtain a temporal resolution Δt of 0.0025 ms, a sampling frequency f_s of 400 kHz is used in what follows. However, since surface properties or reverberation times are normally measured from 125 to 8000 Hz in 1/1 or 1/3 octave bands, the outmost values of the available data set at the lowest and highest frequencies are padded outside the measured frequency bands. Within the given frequency range, the data set is interpolated at 1 Hz interval.

III. TEST EXAMPLES

Consider a room that satisfies most underlying assumptions: flat surfaces that give rise to specular reflections, evenly distributed absorption, and no scattering obstacles. The room is fairly small having a volume of 57.8 m^3 (6.8 m $\times$ 3.4 m $\times$ 2.5 m), and the surface area is 96 m^2. The theoretical mean free path is 2.37 m. The measured reverberation time is given in Table I. The reverberation time is fairly long, up to 1.5 s, due to the low surface absorption. In Fig. 2, the width as functions of the reflection order and the

TABLE I. Reverberation times (s) for the test rooms in octave bands.

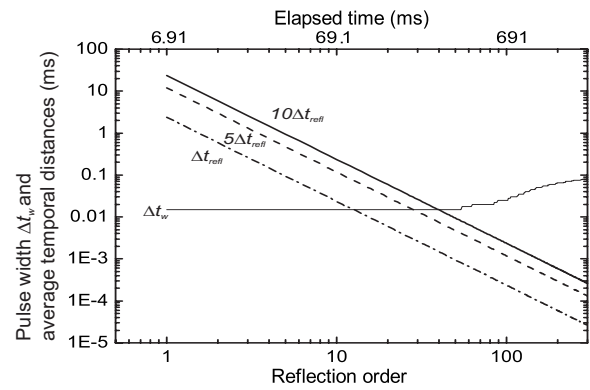| Frequency (Hz)	125	250	500	1000	2000	4000	8000
Office	1.11	1.13	1.35	1.51	1.41	1.09	0.69
Auditorium	0.84	0.80	0.57	0.41	0.37	0.25	0.20

FIG. 2. Graphical solution for an office.

elapsed time is shown. The width Δt_w increases with the reflection order. Figure 2 shows three decreasing curves, indicating Δt_{refl}, $5\Delta t_{refl}$, and $10\Delta t_{refl}$, together with the increasing width curve Δt_w. By overlaying the three curves on the width curve, the solutions can be found graphically. The intersection points indicate that the width equals Δt_{refl}, $5\Delta t_{refl}$, and $10\Delta t_{refl}$, meaning that one, five, and ten individual reflections fall within an average duration of the pulses. The reflection order can easily be converted to the elapsed time by multiplying by the mean free path and dividing by the speed of sound. The elapsed times at which the equalities are satisfied are 86, 196, and 276 ms. The latest distinct peak is found at 216 ms by inspecting the ratio of the instantaneous slope of the Schroeder integrated curve to the mean slope over 30 dB interval. High slope ratios correspond to distinct peaks. The latest peak location is determined as the latest time among the ratios exceeding 13 as shown in Fig. 3(b). The transition time thus lies around fivefold overlap within a pulse width.

Another example is a moderate sized auditorium that does not fulfill the underlying assumptions well: The absorption at the ceiling and the floor is dominant, and it contains a number of heavily upholstered chairs and tables. The reverberation time is listed in Table I. The volume is 298 m^3 and the surface area is 352 m^2. In Fig. 4(a), the transition times using the three criteria are 127, 167, and 192 ms. By the same inspection, the latest peak appears around at 104 ms, which is little shorter than the onefold overlap.
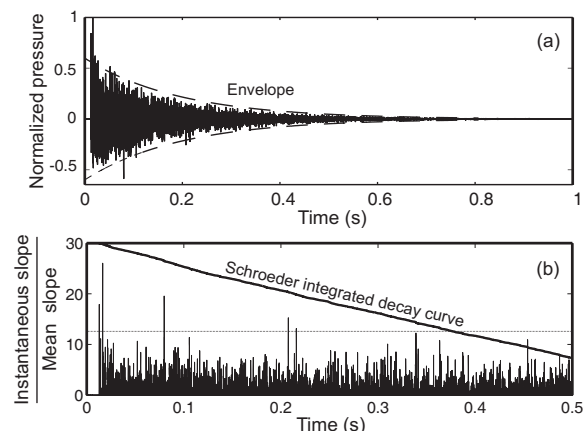

FIG. 3. (a) A measured impulse response; (b) ratio of instantaneous slope to mean slope.
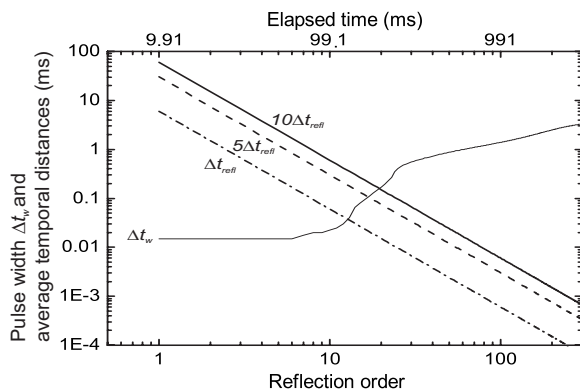
FIG. 4. Graphical solution for an auditorium.

Room	t_{mix} $\sqrt{V}$	t_{st} $2\sqrt{V}$	t_c $\sqrt{\dfrac{\log_e 10^6 V}{4\pi c^3 T}}$	Δt_w	t_{trans} $\left\{\begin{array}{c}1\\5\\10\end{array}\right\}\times\Delta t_{refl}$	Visual inspection
			Transition times (ms)			
Office	7.6(1)	15.2(2)	1.25(0)		86 (13) 196 (28) 276 (40)	216 (31)
Auditorium	18.3(2)	36.5(3)	4.99(1)		127 (13) 167 (17) 192 (19)	104 (10)

IV. DISCUSSION

In Table II, various transition times and corresponding reflection orders are listed for the two tested rooms. The proposed time, t_{trans}, is longest, whereas t_c is shortest in all cases. Physically t_c does not ensure a diffuse field, although it was named the "diffuse field time interval."[11] Note that t_{mix} and t_{st} do not consider the surface properties of the room. A large volume does not guarantee a long t_{trans}, since the acoustic properties of the boundary surfaces have a significant influence. It was found that if the underlying assumptions are well fulfilled, the t_{trans} appears around fivefold overlap, which is equivalent to 25–35 reflections for the studied case. In the auditorium, the transition occurs a little before the time of which the width Δt_w equals Δt_{refl}, being equivalent to ten reflections for the auditorium.

The proposed time t_{trans} can be interpreted such that after this time all the reflected pulses are heavily overlapping, and therefore the envelope becomes more important than the detailed structure. It is assumed that all the impulses from a source undergo the same experience statistically and that many reflections of the same magnitude arrive at a receiver point at an elapsed time t. This is not true for early reflections; however, if many reflections arrive at the receiver, such an assumption becomes valid.

Other major factors to be considered are scattering, diffuse reflections, diffraction, and uneven distribution of absorption, most of which may facilitate the overlapping process. In particular, scattering and diffuse reflections, which disperse the specular energy spatially and temporally, cannot be neglected after three to four reflections,[15] thus fairly short transition time may occur in fitted rooms and performance spaces. Finally note that this theory is the first attempt based on the reflection overlap in the time domain, and may not be practically applied yet in its current form, because it assumes only specular reflections in uniformly treated rooms. However, the concept and the methodology could probably be extended further for practical applicability.

V. CONCLUSIONS

A method for determining a transition time has been proposed. The main difference from previous definitions is that the absorption of the room is taken into account. Such considerations make it possible to define the reflected pulse width, which is compared with the average temporal distance between adjacent reflections. A graphical solution to the equality between the theoretical specular reflection density and the pulse width with a proper factor leads to a transition time. The transition time is therefore based on an assumption of specular reflections in a room where the absorption is uniformly distributed, without considering scattering, diffuse reflection, and diffraction. If the underlying assumptions are satisfied, the transition occurs at around fivefold overlap. For a room with furniture and nonuniform wall absorption, the temporal transition takes place around onefold overlap.

[1] H. Kuttruff, *Room Acoustics*, 4th ed. (Spon, London, 2000), pp.70–98 and 124.
[2] L. Cremer and H. Muller, *Principles and Applications of Room Acoustics* (Applied Science, London, 1978), pp. 503–504.
[3] M. R. Schroeder and H. Kuttruff, "On frequency response curves in rooms: Comparison of experimental, theoretical, and Monte Carlo results for the average frequency spacing between maxima," J. Acoust. Soc. Am. **34**, 76–80 (1962).
[4] P. Rubak and L. G. Johansen, "Artificial reverberation based on a pseudo-random impulse response," in 104th Audio Engineering Society Convention, Amsterdam, 16–19 May, 1998, Paper No. 4725.
[5] L. Cremer, *Statistischer Raumeindruck (Statistical Room Impression)* (Hirzel, Stuttgart, 1961).
[6] V. W. Reichardt and U. Lehmann, "Raumeindruck als oberbegriff von raumlichkeit und halligkeit, Erlauterungen des raumeindrucksmasses R (Room impression as an overall concept of spaciousness and liveness. Examples of room measurement factors R)," Acustica **40**, 277–290 (1978).
[7] J.-D. Polack, "La transmission de l'énergie sonore dans les halles (The transmission of sound energy in the halls)," Ph.D. thesis, University of Maine, Le Mans, France (1993).
[8] G. Defrance and J.-D. Polack, "Measuring the mixing time in auditoria," in the 155th Meeting of the Acoustical Society of America, **123**, p. 3499, Paris, June, 2008, Paper no. 4aAAa11.
[9] J.-D. Polack, "Playing billiard in the concert hall: The mathematical foundation of geometrical room acoustics," Appl. Acoust. **38**, 235–244 (1993).
[10] B. Blesser, "An interdisciplinary synthesis of reverberation viewpoints," J. Audio Eng. Soc. **49**, 867–903 (2001).
[11] M. R. Schroeder, "The Schroeder frequency revisited," J. Acoust. Soc. Am. **99**, 3240–3241 (1996).
[12] R. B. Standler, *Protection of Electronic Circuits from Overvoltages* (Wiley-Interscience, New York, 2002), p. 84.
[13] D. Gabor, "Theory of communication," J. Inst. Electr. Eng. **93**, 429–457 (1946).
[14] J. K. Hammond and P. R. White, "The analysis of non-stationary signals using time-frequency methods," J. Sound Vib. **190**, 419–447 (1996).
[15] H. Kuttruff, "A simple iteration scheme for the computation of decay constants in enclosures with diffusely reflecting boundaries," J. Acoust. Soc. Am. **98**, 288–293 (1995).