Dario D’Orazio e Massimo Garai

THE AUTOCORRELATION-BASED ANALYSIS AS A TOOL OF SOUND PERCEPTION IN A REVERBERANT FIELD

Abstract

A sound in a real space (e.g. a street or a room) is studied by acousticians as the relationship between an anechoic signal and the reverberant sound field. We define ‘anechoic signal’ a signal representing the pressure variation emitted by the sound source, while the ‘reverberant field’ is a field representing the sum of all the sound reflections of the environment in which the sound exists, delayed in time due to the positions of the sound source and the listener. The auditory ability of detecting a sound and assigning it to a source depends both on the anechoic signal and the reverberant sound field. This relationship has been analysed in acoustic literature using the autocorrelation properties of the anechoic signals and objective metrics of the sound field, the last ones being the ‘room criteria’ described in ISO 3382. In this context, the ‘effective duration’ of the autocorrelation function (τ_e) has been proposed as key factor to ‘preferred’ values of several room criteria in relation to different kind of music signals. Relying on some similarities between the definition of ‘auditory objects’ and the sound detection in a reverberant space, this paper proposes the use of τ_e as a potential tool to study and catalogue auditory objects.

1. Introduction

In 1930 Mit Professor of Mathematics Norbert Wiener demonstrated (Wiener 1930) the relationship between Fourier’s transform of power spectra and the autocorrelation function, being the autocorrelation the degree of similarity between a signal with itself delayed in time. The operations involved in carrying out the autocorrelation analysis were different from those involved in the usual frequency analysis. Autocorrelation-based analyses interest different fields of research: cybernetics, psychology, economy, seismology and, of course, acoustics. The interest of acousticians, that had been limited until that time to the frequency analysis, was then focused on the temporal and statistical analysis of the neuronal activity (see Sec. 2).

Rivista di estetica, n.s., n. 66 (3/2017), LVIII, pp. 133-147

Descriptors extracted from the autocorrelation function of a sound signal have been used in perceptual approach to acoustics. Autocorrelation analysis allow studying well known perception issues as the 'missing fundamental' and 'cocktail party effect'. In the field of architectural acoustics the 'effective duration' of the autocorrelation function (τ_e) has been related to subjective criteria as *intimacy* (Ando 1977) and *reverberance* (Ando 1982) (see Sec. 3).

At the beginning, due to technological limitations of the equipment, the range of application of autocorrelation analysis was limited to the music signals whose correlation degree is quite high. In the last years new algorithms allowed to extend the autocorrelation analysis to signals less correlated than music, like speech (D'Orazio 2011; Sato 2011). By using these algorithms the autocorrelation analysis can be applied to a wider range of sound signals – e.g. talking (Kazama 2001) or noises (Soeta 2004) in outdoor environments, i.e. streets, rain forest…

Focusing on monaural and temporal processes only, a potential use of the autocorrelation-based methods is proposed in this work concerning the literature about the 'auditory object', intended as the complex processes of identification and segregation of (one or more) sound sources with respect to the background (Griffiths, Warren 2004). Some considerations about subjective acoustic perception are exposed, in order to propose a potential tool to deal with auditory objects.

2. Autocorrelation-based perception models

Two dominant influences inspire the birth of perception models in the 40's: Helmholtz's resonance-place theory and von Bekesy's work.

Helmholtz allocated each elementary sinusoidal oscillation, in which every stimulus can be decomposed, to its proper position along the length of the basilar membrane. Allocation determines the subjective pitch, the pitch being in one-to-one correspondence with the allocated place. We should note that Helmholtz's formulation suggests a system in which a single compound signal is broken up into elementary parts, which are transmitted in separate channels, spatially distinct from one to another.

Von Bekesy (Von Bekesy 1951) showed that the hydrodynamic action of the cochlea does indeed distribute different frequencies to different locations along the basilar membrane and that this effect is not very sharp. A sinusoidal oscillation of any given frequency, however, excites the vibration of a large part of the basilar membrane. The distribution of vibration amplitude has a maximum, yet, differently than Helmholtz's hypotheses (one single resonant string vibrating and its neighbors quiescent), the distribution is broader. This dullness of the mechanical frequency analysis (an equivalent of blurring in vision) changed in the 30's the conception of auditory system: instead of discrete and distinct channels, von Bekesy's theory proposed a continuum of channels, broad and overlapped.

In 1942 Licklider was writing his doctoral thesis on frequency localization in the auditory cortex of the cat (Licklider 1942), working at Acoustics Laboratory of MIT on speech perception in telecommunications. Following Fletcher's space-time pattern theory (Fletcher 1953), Licklider proposed the *duplex theory* of pitch perception in the 1951 (Licklider 1951). The essence of the duplex theory is that the auditory system employs both frequency analysis and autocorrelation-based analysis, merging Helmoltz's theory on resonance-place and Von Bekesy's approach.

In Licklider's model the frequency analysis is performed by the cochlea, the autocorrelation-based analysis by the neural part of the system. The latter is an analysis of trains of nerve impulses: a highly non-linear process of neural excitation intervenes between the two analyses. If the lengthwise dimension of the uncoiled cochlea is designated as the x-dimension, the cochlea transforms the stimulus time function $f(t,x)$ into a running spectrum. The running spectrum, a spatial array of time functions, is transmitted by the neurons of the auditory nerve.

After Licklider's work, the model of monaural perception based on the autocorrelation function has been developed by several authors. Patterson (1986) introduced the pulse-ribbon and the *Strobed Auditory Integration* (SAI), as a cross correlation between the neural response within each channel and a strobe of single pulse per stimulus. Meddis (1997) summed the ACF across channels in the frequency domain and obtains a *summary autocorrelation function* (SACF) based on the probability of all-order interspike intervals. With the hair cell transduction model, Meddis introduced the pitch dominance region and permits to compute the relative weight of low-versus-high frequency channels and the breakdown of neural firing synchrony at high frequency. Cariani and Delgutte (Cariani 1996) showed the relation between the shape of the autocorrelation histograms (ACH) (*all-order interspike interval* histograms for each channel) and the ACF. The raw waveform proposed by Yost (Yost 1996) is also similar to SACF. Despite recent improvements to the original Licklider's model, at the present time the words of de Cheveigne are still valid: «autocorrelation model remains a good first-order model, attractive in terms of simplicity, explanatory power and physiological plausibility» (De Cheveigné 1998).

3. Subjective preference of intimacy and reverberation

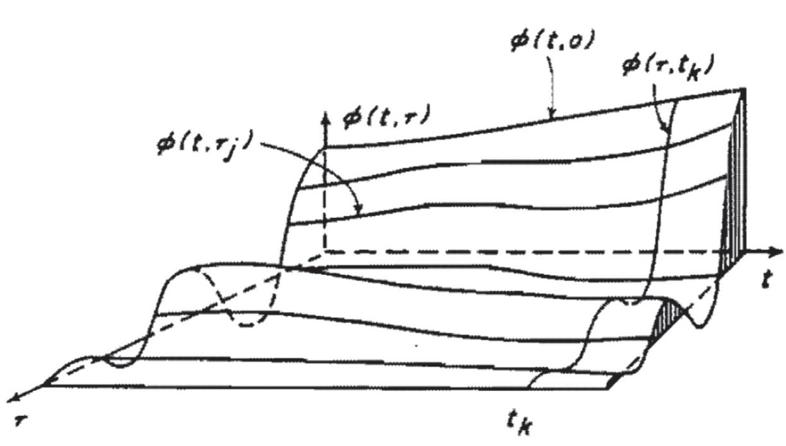
In 1965 the Russian researcher Fourdouiev (1965) proposed that the *optimal* reverberation time depends on the autocorrelation of the musical signal or speech signal. In Fourdouiev's proposal the degree of coherence between a musical signal and the same delayed signal is related to the correlation between the direct sound and each reflection of the listening environment. In Fourdouiev's study a musical motif with a high degree of coherence (such as Bach's *Corale for Organ*) can be supported by a high reverberation of the hall. Instead of this, a music motif with a low degree of coherence (such a Frescobaldi's *Passacaglia for Clavicembalo* or a Grieg's *Lieder for Voice and Piano*) needs low reverberation of the hall.

The acoustic group of Drittes Physikalisches Institut of the University of Göttingen generalized Fourdouiev's point of view from Fourdouiev's *optimal* to *preferred* value of some acoustic criteria. The Göttingen group studied the follow room criteria as defined in the international technical standard Iso 3382-1:2009 [Iso 3382]:

– Reverberation time *T*;
– Initial Time Delay Gap (*ITDG*), time delay between arrival time of the direct sound and first reflection of the hall;
– Definition *D*, ratio between energy of early reflections and energy of the whole room;
– Inter Aural Crosscorrelation Coefficient *IACC*, a binaural criterion related to the spatial perception of the listener.

In a comparative study of European concert halls (Schröder 1974) the group introduced a new method to evaluate the subjective experience in a concert hall: an anechoic registration is convolved with the binaural impulse response of the concert hall and the resultant two signals are proposed to the listener after a crosstalk cancellation. Moreover, authors introduced in the acoustic literature a new formalism: the *preference space*, where the listeners' judgements are evaluated. In the preference space each vector represents the listeners' judgement

Figure 1. Running autocorrelation surface. Each instant tk corresponds to an autocorrelation Phi(tau.tk), evolving in time. After Licklider (1951).



on a particular hall: the projection on the abscissa axis represents the mean consensus on the hall, the projection of the ordinate axis reflects individual preference disparities. The preference space is used also to represent the correlation of each objective criterion with its subjective perception. Two geometric parameters obtained from the drawings of the hall (volume V and width W of the hall) and four acoustic criteria obtained from measured impulse responses (for example *Itdg*, T, D, *Iacc*) are evaluated in figure 2. The results show that the greater the reverberation time T is, the greater the consensus preferences for these halls (high abscissa value, low ordinate value). The definition D has a highly negative correlation with the preference: lower values of definition corresponds to higher preference values (high negative abscissa value, low absolute value of the ordinate). Sound diffusion is also strongly correlated to preference: a high sound diffusion is expected to have a better judgement.

After his PhD held in Göttingen, Ando set up listener tests using the Göttingen listening configuration. He introduced the *listener temporal preference* of a certain music motif in a certain hall, extending the work of M. Schröder and the Göttingen group. Following Fourdouiev's degree of coherence of a music motif, Ando introduced the *effective duration of the autocorrelation function* (τ_e), defined as the decay of the autocorrelation function of the anechoic sound signal, evaluated on 10 dB of decay (see figure 3). τ_e plays a key factor in *subjective preference*, being the subjective preference the overall evaluation of the perception of a music piece in a concert hall. A representation of sound sources in a τ_e-duration space is shown in figure 4.

In the first experiment an anechoic motif was proposed frontally to the listener in addition to a delayed version of the same motif, performed by lateral loudspeakers (Ando 1977). Changing the delay time of the delayed signal (a virtual *Itdg*) and using Thurstone's law (Thurstone 1927), different levels of listener's preference were evaluated. Being *Drr* the *direct-to-reverberant ratio* – similar to the Dry/Fx parameter in a digital processor – the listener's preference decreases when:

$$0 < \mathrm{Itdg} < [1 + 2 \log (\mathrm{Drr})]\tau_e \qquad (1)$$

due to interference phenomena in the coherent region, i.e. to the coloration effects of the reflections. When:

$$[1+2\log(\mathrm{Drr})]\tau_e < \mathrm{Itdg} \qquad (2)$$

the preference also decreases due to the echo disturbance on the listener's perception. It follows that $[1+2\log(\mathrm{Drr})]\tau_e$ is the preferred value of *Itdg* for concert halls. From a perceptual point of view this means that each signal has a proper value of delay of the first reflection, and this value is related to the autocorrelation function. As a value of time delay corresponds to distance travelled by the sound wave, and considering that an open space has a high *Drr* value and a closed space a low *Drr* value – both depending also on the distance – it follows that the localization process has a preferred focusing area for each sound signal: shorter for impulsive sound (low τ_e) and higher for correlated sound (high τ_e).

In a study of 1981, led with a method similar to that described in Ando (1977), Ando studied the listeners' preference as a function of the late reverberation, expressed by *T*. He found by regression the relationship between the preferred value of the late reverberation and the effective duration of autocorrelation function:

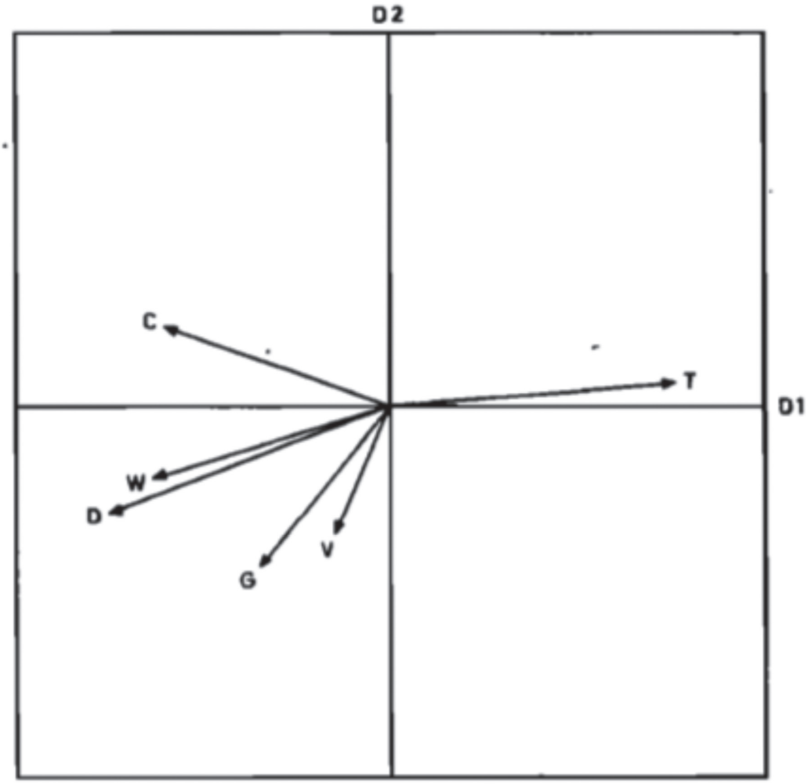
$$T = 23\tau_e \tag{3}$$

Below this limit value, the preference is still quite high, above the value of $23\tau_e$ the preference decreases quickly. It follows that for each signal there is a maximum degree of decorrelation due to the environment and this degree depends on the characteristics of the autocorrelation of the signal (as a first approximation represented by τ_e). As shown in figure 5, a high correlated signal – e.g. a continuous sine wave – when is listened in a reverberant space is still a sine wave, and it is detectable. On the contrary, a low correlated signal – e.g. an impulse – when is listened in a reverberant space is masked by the environment, and thus it is undetectable. An organ sound (high τ_e value) is still an organ sound in a cathedral, and it is still distinguishable, while in the same environment a speech (low τ_e value) is undistinguishable. In other words, in a reverberant space some kind of signals are most focused than others and their degree of focusing depends on τ_e.

With some analogy to the *preference space* of Schröder *et al.*, cross related analyses were done in the eighties (Ando 1982) to evaluate the listener's judgement preference. These studies found a high degree of total *subjective preference* when the preferred value of each factors is obtained and a high degree of independence among these factors that authorize to postulate the orthonormality of each factor in relation to the others. Using a statistical technique called factor analysis, Ando and co-workers showed that a sound signal can be specified in a four-dimension signal space: these coordinates are orthonormal and may represent uniquely the subjective preference of the listener. Moreover, brain activities corresponding to the subjective preference were observed: recording auditory brainstem responses, slow vertex response, MEG and EGG from right and left hemisphere; these studies show that temporal factors are associated with the left hemisphere and the spatial factors are associated with the right one (Soeta 2002).

The Schröder-Ando approach was extended by several authors. Recently Lokki *et al.* (2012) defined a more complex space of perception, focusing on new categories as proximity. It is interesting to note that the original Schroeder's approach of testing in a listening room is still used, convolving anechoic signals with environmental impulse responses: the former using only two audio channels, the latter using a more complex immersive reproduction system.

Figure 2. Consensus *vs.* Preference disparities for different room criteria. Correlations of objective criteria evaluated in 11 concert halls with their subjective positions. The projection of the room criteria on the abscissa represents consensus preference (for high – positive or low – negative values); the projection on the ordinate reflects individual preference disparities. See details in the text. After Schröder (1974).

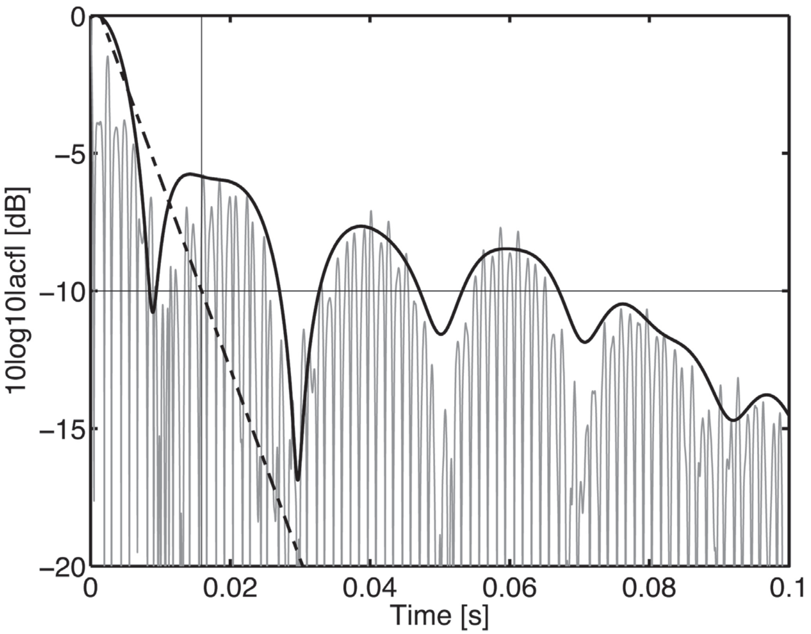


3.1. Experimental results

The aforementioned researches on subjective intimacy and reverberation time show the role of τ_e as objective parameter for characterizing the audio signal that excites a sound field. In this regard, the most important role of the minimum of the running τ_e needs adequate methods of measurement.

In 1965 Fourdouiev used an exponential weighted running autocorrelation function ('fonction d'autocorrelation evolutive') where a time constant of about 30 ms is assumed long enough to approximate a short–time autocorrelation. The whole algorithm was implemented by an analog setup, using overlapped reels for autocorrelation and diodes and condenser for integration. He studied only excerpts of music, spanning from Grieg's *Lieder* to Bach's *Corale*.

Figure 3. Calculation of effective duration of the autocorrelation function (τ_e). The τ_e value is the delay time, in s or ms, corresponding to a 10 dB decay of the autocorrelation function, fitted on the first 5 dB of decay only. After D'Orazio (2011).

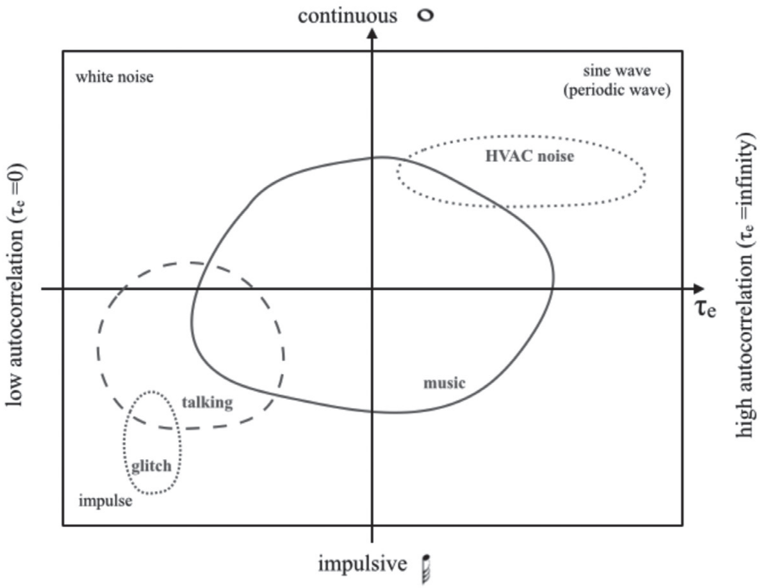


Ando used a digital autocorrelator, analysing monophonic anechoic recordings from BBC (Burd 1969). They consisted of compositions for small orchestra, recorded by one point microphone, with a low dynamics. The minimum value of τ_e computable from this equipment was about 30 ms, adequate for the music but inadequate for less correlated signals as talking or noise. Same approach was used by Hidaka (1988); also in this case symphonic music was analyzed.

D'Orazio *et al.* (2011) and Sato and Wu (2011) proposed algorithms capable of dealing with smaller values of τ_e. Iterating and optimizing some operations, these algorithms may extract low τ_e values (less than 10 ms), e.g. to analyze the talking processes in different languages, or categorize different kind of sounds in soundscape recordings.

The algorithm for envelope extraction proposed by D'Orazio *et al.* has been extended also to statistic signals (D'Orazio 2012; De Cesaris 2015), such the room impulse responses. Concerning the analysis of music pieces, more detailed musical anechoic recordings have been done (Pätynen 2008; D'Orazio 2016) in order to extend the latest model of perception (Lokki *et al.* 2012) to complex closed environments like, e.g., historical theatres (Garai 2015).

Figure 4. Representation of sound signals in a τ_e/duration space. On the abscissa axis the degree of correlation of the signal, expressed by the effective duration of the autocorrelation function (τ_e). On the ordinate axis the duration of the sound signal, from impulsive to continuous. In the corners three ideal signals: the impulse (uncorrelated and impulsive), the white noise (uncorrelated and continuous without any periodicity) and the sine wave (correlated and periodic). Some categories are sketched qualitatively, e.g. the noises coming from air conditioning systems (HVAC noises) or the digital glitches used in contemporary music compositions.

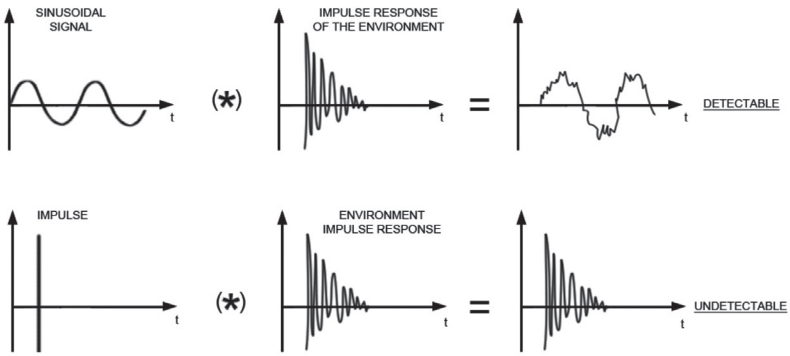


4. Discussion

In the field of applied acoustics each sound may be decomposed in the direct wave (the sound as generated) and the contribution of the environment. The first one may be evaluated by recording the sound in anechoic conditions, the second one measuring the so-called impulse response of the environment (Iso 2009), the impulse response being the sum of each acoustic reflection of the acoustic environment, delayed in time. Indeed, under certain conditions, it is possible to render a sound by convolving the anechoic signal with the impulse response of the environment (real or virtual). It must be noted that also outdoor environments return an impulse response, due to scattering effect of all surfaces: this is the case of squares, ancient Greek theatres, streets and so on.

The impulse response itself may be divided in two parts, respectively the correlated (or deterministic) one and the statistical part. The correlated part

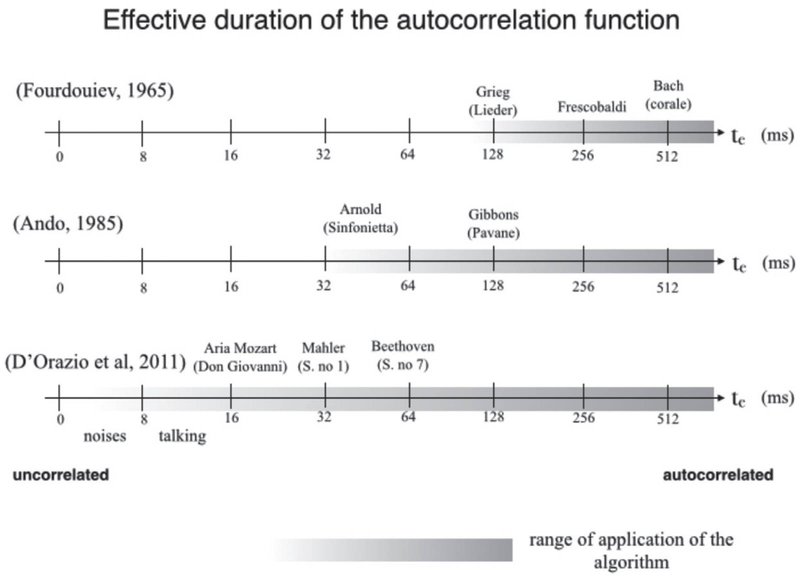
Figure 5. Detectability of a sound source in a reverberant environement. Anechoic signals (on the left) are convolved (*) with an impulse response of the environement. In the right column the results of the convolutions: how the sound sources are perceived. A high correlated signal (a sine wave) is still detectable when a low correlated signal (an impulse) is undetectable.



corresponds to the first specular reflections near to the source of the sound, sparse in time. This first part is responsible for the perceptual effects of source localization, intelligibility, etc. The statistical part is due to the subsequent reflections coming from all the surfaces of the environment. Due to multiple acoustic effects these reflections are dense in time and contribute to 'late reverberation' of the environment. The correlation of the reflections of the early part of the impulse response means that these reflections have similarities with the direct sound and among each other. Moreover, each one of these reflections has a predominant arrival direction to the listener. On the contrary, in the statistical part of the impulse response the reflections cannot be detected individually and can be treated only with a statistical approach. So the correlated part contributes to the focus of the object, the statistical one is related to the background, borrowing some syntax from the field of auditory objects (Bizley and Coehn 2013).

Figure 7 shows a model of impulse response. The initial time delay gap (*ITDG*) is the time delay between the arrival time of the direct sound and the arrival time of the first reflection: it depends on the deterministic part of the impulse response. In other words it depends on the reciprocal positions of the sound source and the listener in the environment. In wider terms, the localization process depends on the deterministic part of the impulse response of the environment. The reverberation is the decay of sound pressure level, fitted over the statistical part of the impulse response only. In other words the reverberation does not

Figure 6. Schematic representation of the progress of the algorithms for detecting the the minimum of the running τ_e over time.



depend on the sound source and listener's position, but rather it depends on the acoustic properties of the environment.

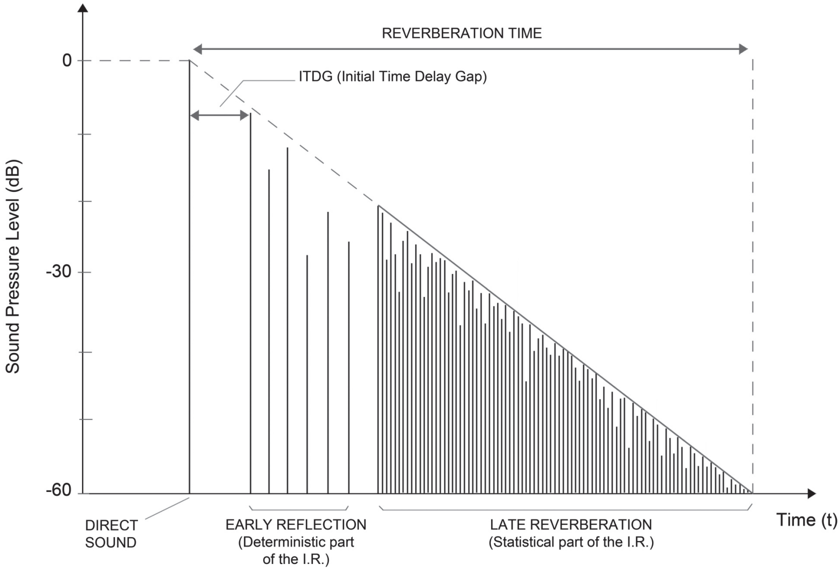
In the previous section it has been shown that the preferred value of the initial time delay gap (the objective criterion for the intimacy of sound) depends both on τ_e and on the direct-to-reverberant ratio, while the preferred value of the reverberation depends on τ_e only. Merging the results on preferred values for intimacy and reverberation, with the analysis of an impulse response, we may conclude that τ_e value, as representative of the degree of correlation of the anechoic signal, is related to a threshold of discrimination of an auditory object with respect to the uncorrelated background. In a low reverberating space (like an open space) the discrimination process is due to the deterministic part of the impulse response: depending on their τ_e value some signals may be localized better than others. In a high reverberating space (e.g. a church) the discrimination process is due to the statistical part of the impulse response: depending on their τ_e value some signals may be more distinguishable than others.

Some mid-reverberant spaces are able to emphasize both the deterministic and the statistical parts of the impulse response: for example, in historical Italian opera houses the vocal parts are well distinguishable and localized, due to the position of the singer on the stage, the reflections from the proscenium arch and the geometry of the cavea. On the other hand the orchestra is less localized, as its position in the pit de-emphasizes the deterministic part of the sound from the orchestra, but the reverberation of the hall contributes to balance the perception of the listener: on the correct focus the singer, in the background the orchestra, thanks to a proper reverberation.

5. Conclusions

Autocorrelation models of acoustic perception have often been applied to the evaluation of the environment. In this work a review of applications of the effective duration of the autocorrelation (τ_e) was presented. Recent algorithms allow extending the use of the τ_e to cover a wider range of acoustic signals: it is now possible to extend the autocorrelation analysis from music only to other typologies of sound. It follows that τ_e extraction can be a potential tool for studying auditory objects, assuming valid the room acoustics model in which a sound is a convolution between an anechoic sound and an impulse response of the environment.

Figure 7. Qualitative representation of an impulse response, measured recording the time-behaviour of the reflections in an environment. After the direct sound the main surfaces around source and listener contribute to the early reflections. These are deterministic because they depends on the geometry of the environment and on the mutual position between sound source and listener. After some reflections the reverberation turns on a statistical process (late reverberation), depending on the environment only.



Literature says that the temporal perception is related to τ_e; thus, the temporal perception is related to room acoustics criteria extracted from impulse responses: the objective parameters of intimacy – initial time delay gap (*ITDG*) – and reverberance – reverberation time (*T*) – depend on τ_e. All acoustic environments may be described using an impulse response, so *ITDG* and *T* may be generalized to objective parameters related, respectively, to the localization and to the de-focusing of an acoustic source (with some mutual influences). With this generalization the two criteria may be discussed from a point of view of auditory objects.

The key factor τ_e may be taken into account as the intrinsic ability of a sound signal to be focused. A sound signal with a low τ_e may be focused only in a low reverberant space, otherwise it is shifted in the background perception. A sound signal with a high τ_e may be focused also in a high reverberant space.

Furthermore, τ_e value is also related to the localization of a sound source both in indoor and outdoor environments. In the first case the relationship between the τ_e of the sound signal and the *ITDG* of the hall is related to the subjective categories of intimacy and proximity.

6. Acknowledgements

The authors are grateful to Sofia for her helpful comments.

References

ISO
– 2009, ISO 3382-1: Acoustics - Measurement of room acoustic parameters Part 1: Performance spaces, Geneva, International Standards Organization.

ANDO, Y.
– 1977, *Subjective preference in relation to objective parameters of music sound fields with a single echo*, "J. Acoust. Soc. Am.", 62: 1436-1441.

ANDO, Y.
– 1982, *On the preferred reverberation time in auditoriums*, "Acustica", 50: 134-141.
– 1985, *Concert Hall Acoustics*, New York, Springer.

BIZLEY, J.K., COHEN, Y.
– 2013, *The what, where and how of auditory-object perception*, "Nature Rev. Neurosci.", 14: 693-707.

BURD, A.
– 1969, *Nachfaller Musik fur akustische Modelluntersuchungen*, Rundfunktechn. Mitteilungen, 13: 200-201.

CARIANI, P., DELGUTTE, B.
– 1996, *Neural correlates of pitch of complex tone. i. Pitch and pitch salience*, "J. Neurophysiol.", 76: 1698-1716.

De Cesaris, S., D'Orazio, D., Morandi, F., Garai, M.
– 2015, *Extraction of the envelope from impulse responses using pre- processed energy detection for early decay estimation*, "J. Acoust. Soc. Am.", 138, 4: 2513-2523.
De Cheveigne, A.
– 1998, *Cancellation model of pitch perception*, "J. Acoust. Soc. Am.", 103: 1261-1271.
D'Orazio, D., De Cesaris, S., Garai, M.
– 2011, *A comparison of methods to compute the 'effective duration' of the autocorrelation function and an alternative proposal*, "J. Acoust. Soc. Am.", 130, 4: 1954-1961.
– 2012, *Measuring reverberation time using preprocessed energy detection*, "41st International Congress and Exposition on Noise Control Engineering (Inter-noise) 2012", 1: 740-750.
– 2016, *Recordings of Italian opera orchestra and soloists in a silent room*, Proceedings of Meetings on Acoustics "22nd International Congress on Acoustics", 28, 1.
Fletcher, H.
– 1953, *Speech and hearing in communication*, New York, Van Nostrand.
Fourdouiev, V.
– 1965, *Evaluation objective de l'acoustique des salles*, in *Proceedings of 5th International Congress on Acoustics*, Lièg: 41-54.
Kazama, M., Tohyama, M.
– 2001, *Estimation of speech components by acf analysis in a noisy environment*, "J. Sound and Vib.", 241: 41-52.
Garai, M., Morandi, F., D'Orazio, D., De Cesaris, S.
– 2015, *Acoustic measurements in eleven Italian opera houses: Correlations between room criteria and considerations on the local evolution of a typology*, "Building and Environment", 94, 2: 900-912.
Griffiths, T.D., Warren, J.D.
– 2004, *What is an auditory object?*, "Nature Rev. Neurosci.", 5: 887-892.
Hidaka, T., Kageyama, K., Masuda, S.
– 1988, *Recording of anechoic orchestral music and measurement of its physical characteristics based on the auto-correlation function*, "Acustica", 67: 68-70.
Licklider, J.
– 1942, *An electrical investigation of frequency localization in the auditory cortex of the cat*, PhD thesis, University of Rochester.
Licklider, J.
– 1951, *A duplex theory of pitch perception*, "Experientia", 7: 128-134.
– 1956, *Three auditory theories*, in C. Cherry (ed.), *3rd London Symposium on Information Theory*, London, Butterworth's.
Lokki, T., Patynen, J., Kuusinen, A., Tervo, S.
– 2012, *Disentangling preference ratings of concert hall acoustics using subjective sensory profiles*, "J. Acoust. Soc. Am.", 132, 5: 3148-3161.
Meddis, R., O'Mard, L.
– 1997, *A unitary model of pitch perception*, "J. Acoust. Soc. Am.", 102: 1811-1820.
Patterson, R.D., Moore, B.C.J.
– 1986, *Auditory filters and excitation patterns as representations of frequency resolution*, in *Frequency Selectivity in Hearing*, London, Academic.

Patynen, J., Pulkki, V., Lokki, T.
– 2008, *Anechoic recording system for symphony orchestra*, "Acta Acustica united with Acustica", 94, 6: 856-865.
Sato, S., Wu, S.
– 2011, *Comparison of different calculation methods of effective duration of the running autocorrelation function of music signals*, "Acta Acustica united with Acustica", 97: 432-440.
Schroeder, M., Gottlob, D., Siebrasse, F.
– 1974, *Comparative study of Europen concert halls: Correlation of subjective preference with geometric and acoustic parameters*, "J. Acoust. Soc. Am.", 57: 1195-1201.
Soeta, Y.
– 2004, *Annoyance of bandpass-filtered noises in relation to the factor extracted from autocorrelation function*, "J. Acoust. Soc. Am.", 116: 3275- 3278.
Soeta, Y., Nakagawa, S., Tonoike, M., Ando, Y.
– 2002, *Magnetoencephalographic responses corresponding to individual subjective preference of sound fields*, "J. Sound and Vib.", 258: 419-428.
Thurstone, L.L.
– 1927, *A Law of Comparative Judgment*, "Psychol. Rev.", 34: 273-286.
Yost, W., Patterson, R., Sheft, S.
– 1996, *A time domain description for the pitch strength of iterated rippled noise*, "J. Acoust. Soc. Am." 99: 1066-1078.
Von Bekesy, G., Rosenblith, W.A.
- 1951, *The mechanical properties of the ear*, New York, Wiley.
Von Helmoltz, H.
– 1877, *On the sensations of tone*, New York.
Wiener, N.
– 1930, *Generalized harmonic analysis*, "Acta Math.", 55: 117-258.