Slovak University of Technology in Bratislava
Faculty of Mechanical Engineering
14th Conference on Applied Mathematics **APLIMAT 2015**

Proceedings

EARPOINTS: TIMESCALES IN LISTENING

TEDESCHINI-LALLI Laura (I)

Abstract. Our act of listening incorporates an ability to decode signals in real time and extract several time-scales. We review timescales intrinsic to spatial hearing. Space is encoded, in the auditory signal, in a very small timescale. We review here some of the very small time-scales involved in audio-location, and make the point that the soundwalk, advocated by Murray Schafer to investigate soundscape, is indeed a tool that enhances appreciation of these small scale.

Key words. Time-scales, hierarchies of scales, Nyquist Shannon theorems, auditive. localization.

Mathematics Subject Classification: Primary, 54C70.

1 Introduction: Viewpoints and Earpoints, an Operative Definition

Sound happens in time: it is energy crossing spaces. In mathematics and in visual art, it has been pointed out that the “point of view” of the artist and of the scientist has evolved in history with some parallelism that allowed an evolution of treatments of “space”. Science has relatively recently started to deal with “time” and its descriptions. In our research, we sketched the parallelism of viewpoints in art and mathematics, in order to start talking about “time”, and earpoints. Spatial variables, in mathematics, have as many meanings and uses as the word “spaces”: a “space” is in fact nowadays understood to be a set of elements together with some relations. Much often, relations themselves are the elements, and the space consists of their combinations, …and relations. I think we have nothing quite so sophisticated regarding the word “time”, as a variable and/or as a structured container. This subtly hinders studies about sound, especially when it comes to the complex time organization of aural information.

Earpoints also deal with the time-scale we focus upon while listening. It is well known, and everybody’s experience, that we have the ability to switch our attention from the general to the detail, and the opposite, much as in the visual case. Unlike the visual analogous concept, though, focusing is not well understood in listening. For the sake of our research we introduced the concept of “earpoint” in order to discuss auditive attention, with particular reference to timescale [1]. In particular in sound-scape many sounds, otherwise seemingly unobtrusive, have the quality to draw attention to them. It is the case for example of water of fountains, of birds in the air. They are feeble sounds that give a sense of depth to the perceiver. And, they are sounds that elude precise mathematical treatment, because their short duration makes them impulsive phenomena, rich in transients. We review some research done in this realm, from the mathematical modeling point of view, and as to the method of surveying of existing soundscapes. Since spatial information is encoded in the audio signal in a very small time-scale, it is interesting to see what implications on recorded documentation come from considering appropriate timescales. Moreover, from a mathematical point of view, it is very interesting to consider hierarchical models, incorporating interplay among different scales.

2 Some known time-scales in listening and music.

Musicians, musicologists, and more recently computer musicians, have all always dealt with hierarchies of time-scales. Much as for other aspects, also in this case musical practice and musical theory can inspire scientists to apply their methods of hierarchical organization to time processing of complex messages. As just one deep example, one of the most well-known and widely accepted analytical methods for tonal Art Music is that proposed by Heinrich Schenker [2], which is deliberately hierarchical in method. Even more to our point, is that many celebrated performers of western Art Music, advocate this method of analysis to both understand the piece of music they are to perform, and to retain it in memory, for example one of the greatest conductors of all times, Wilhelm Furtwängler. This means, in particular, that acknowledging and studying the hierarchical organization of music facilitates its understanding, its performance, and the memory of it. In turn, accepting this method from the musicians, we acknowledge that memory itself is organized over different time-scales.
In sound electronic or digital processing, defining such scales becomes really crucial, of course [5].
Here, we are rather interested in our human cognitive ability to process sound in real time, and to assess our expectation regarding it. As definition of listening, just for the purpose of my research (and thus limited to the possibility of modelling), I propose the following [6]
Definition: listening= establish relevant timescales in a signal, and assign phenomena to the right scale, as the sound flows in.

We will draw from the book by Moravcsik [7] the principal scales involved in hearing, and in cognitive memory. According to Moravcsik, who reports several experiments, memory used in listening to music falls in the following five different time-scales[1]:

-Supershort:	1/30—1/10000 sec,	i.e. 10^{-2} --- 10^{-4} sec
-Short:	1—1/10 sec,	i.e. 10^{0} --- 10^{-1} sec
-Medium:	1---5 sec	i.e. order of 10^{0} sec
-Long:	5---1 minute	i.e. 10^{0}---10^{2} sec
-Very long:	1 minute-hours	i.e. 10^{0}---10^{3} sec and beyond

Phenomena on each scale are perceived and assigned to different features, with different names:
-SuperShort scale, intonation is typical, in music, of what we refer to as *pitch*: events repeating at this scales are perceived as one “note”.s

[1] We refer the interested reader to the original book for more specifications. A short description of his table can also be found on Wikipedia, under the entry “rhythm”.

-Short scale is what usually, in music, lets us perceive the *tempo*, or the *beat*
-Medium scale: the *rhythm*, i.e. repetitions of patterns constituted by beats and accented
-Long scale: *phrasing*, or stanzas, any repeating structure just larger than rhythm
-Very long scale: entire compositions.

Similarly, one could hierarchize a verbal spoken signal in phonemes, words, sentences…
And, in fact, while there is no discussion that there is a hierarchy, the levels themselves are subject of longstanding for continuous research in various branches of Linguistics.

And…, there is life below the super-short scale!
In fact, on times shorter than these, lies the scale of "transients of attack", full of information and, musically speaking, a distinctive feature of each particular performer. Here lies the scale for "timbre", or sound-colour, the study of the waveform, usually discussed in terms of spectra, and therefore of frequencies. I would like to stress that at this level, there is already an interesting connection between different timescales, that of timbre, and that of the particular choice of pitches specific to a musical culture. It has been an object of research, in fact, how much the particular structure of sound-colour of a musical culture affects the musical scale that is projected along the centuries in that musical culture (independently discussed in [8],[9]). In short, strings and hollow canes, mathematically speaking, share model justifying a harmonic spectrum, quite strict in the first few components. This can explain that intervals of octave and fifth, corresponding to the first two harmonics, are very strong in musical cultures whose practice is largely based on orchestras and ensembles of strings and hollow canes. On the other side, as a different example, musical cultures largely based on bronze tuned slabs, such as in insular South-East Asia, project a musical scale where even the interval of fifth and octave are discussed. We have here an example of information passing from one timescale (that of the frequencies in the spectrum of an instrument) to a larger one (that of the repeated pitches).

3 Time Scales in Spatial Listening: well Below 10^{-4} sec

Our ability to locate a source of sound by listening is due to several factors. The first and most evident is our binaurality. Due to binaurality, we are able to discriminate sources on a horizontal plane. We are able to discriminate whether the source of a sound is closer to one or the other of our ears, because the sound reaches either ear at different times. Our brain directly encodes as spatial information about the source, the differences in such arriving times. For the description of these scales, we refer to [12], reporting the measurements of experiments on large groups of perceivers. The time-scales that can be assessed by binaurality are of the order of 100 - 700 μs = 1-7 10^{-4} sec, i.e. order of magnitude 10^{-4} sec.
In reality, experiments show we appreciate and locate spatialization over much smaller scales. In fact, our head is immersed in the fluid carrying the sound, and thus the vibration is modified by this obstacle. The character of the same sound is qualitatively different on the left and the right ear, because it is modified by the head separating the ears. Morover, once in the auricle (the large external ear), sound is diffracted by the many folds in the auricle, and directed toward the auditory channel. By tilting our head, this diffraction changes ever so slightly and the quality of sound entering the auditory channel changes. These changes, usually measured in frequency, are again encoded by our brain as spatial informations to locate the sound source.

This subtle discrimination of directions and distances, involves assessing times differences caused by reflections on the auricle folds and is of 2-80 μs, i.e. order of magnitude 10^{-6} sec.

2 Soundscape: the Appreciation and Documentation of Large Spaces.

The term "soundscape" has been introduced by Raymond Murray Schaffer [13]. While today the word soundscape encompasses many activities, we are interested in the original problem of investigating and documenting the auditive aspects of our surrounding. Schaffer introduced the "sound-walk", as a method for investigating soundscape. The method of soundwalking is today accepted and used by the large variety of professions today interested in soundscape and soundscaping. One could do different soundwalks, but they all have in common the fact that the perceiver actually changes place in space, by walking. Obviously soundwalking has to do with travelling the space one wants to investigate, displacing, evaluating differences, and checking by ear, personally. But there is more to it than it is generally discussed, and I think this is one of the reasons for its success. Walking allows comparison between different time-scales, relating the time-scale of the displacement of the perceiver to the time-scale of the sound which needs to be assessed. We know moving facilitates appreciation of depth in visual assessment, by allowing comparison between spatial scales of the object and spatial scales of one's displacement; much in the same way, walking facilitated appreciation of spatial depth in sound. Walking allows listening times that sitting, and facing a source, does not allow. Remembering that in art music we are generally expecting a "central focus" listening [14], this is a quite different experience.
Documentation of soundscape is particularly delicate. In fact, most of our recording devices are still tuned to relatively small spaces. Steven Feld, a world known ethnomusicologist, has deviced some years ago a way to wear microphones over his head. He then records while walking, so that both his walking toward or away from a source, and his moving head, captate very small time differences. This technique allowed him unprecedented clarity in his world wide documentation of bells [15]. Bells are notoriously very difficult to record, as it is easy to end up with distortions.
On the other hand, it is often heard that the modern digital recording deprives sound of some depth. People usually say this in a metaphor, but we now have arguments that it is really the timescale of depth the one that gets cancelled. In documentation of soundspace, we should either have a sampling rate allowing for the small timescales involved in spatial appreciation, or else include them directly in the recorded stream, by means of binaural microphones worn on the head of a human, moving perceiver.
In fact, the sampling rate of digital recording is established in application of Nyquist Shannon theorem:

Theorem (NS): If a function $x(t)$ contains no frequencies higher than B Hz, it is completely determined by giving its ordinates at a series of points spaced $1/(2B)$ seconds apart.

As hordes of experiments show that the threshold of audibility for humans is 20 kHz, applying Nyquist Shannon theorem would guarantee that audio signals sampled at 40 kHz are completely faithful to the original, within the limits of human ability to discern. This yields a sampling time of 1/40000 sec, order of magnitude 10^{-5} sec, which is rougher than the timescale involved in spatial appreciation. So we are cutting out all spatial information except for that linked to distance and the absorption of sound energy by the medium. It seems that sampling at this rate we quite literally "flatten out" the recorded sound.

This train of reasoning for this widely accepted application of Shannon's theorem, relies on experiments performed by playing some definite frequencies to humans. But we can hear and appreciate more than frequencies. Impulsive and faint or feeble sounds are not easily described in terms of frequencies, for instance, while we can appreciate them well. What happens to impulsive sounds when recorded? By applying NS theorem to auditive signals, we are cutting out all non-periodic sounds. Non periodic sounds are pieces of signal where Fourier theorem is of no avail: we would need infinitely many frequencies to approximate them. Such sounds encompass all the clicks, swishing and rustlings. Sounds we all know, can hear, and appreciate. In fact, these sounds are used by the visually impaired to audiolocate obstacles. They emit such a sound, and the sound will come back to them subtly altered when reflected on an obstacle. Because these sounds are rich in transients, their distortion can be full of information, and the visually impaired learn to interpret this distortion as distance, material, and shape of the reflecting ostacle.

Conclusions

Study of timescales involved in perception of sound are a relatively recent subject. Exeperiments are underway. The first recommendation following form these studies, is for the cognitive psychologists to avoid immobilization of the head of the perceiver. Such immobilization is performed in analogy with experiments on vision, but timescales are larger, here. Allowing a head to tilt is analogous to allow an eye to blink!

A mathematical model of incorporation of different timescales in a coherent decodification was introduced in [6], in turn following the hierarchic model introduced in [16].

Acknowledgements

The research on the topics of this paper was supported, along the years, by GNFM-INDAM.
Aplimat has provided a very precious and rare forum of discussions from different realms.

References

[1] GIURIATI, G. TEDESCHINI LALLI, eds, *Spazi sonori della Musica* L'Epos 2010

[2] TEDESCHINI LALLI, L. "Punti di vista e punti di udito" in: *Spazi Sonori della Musica,* op. cit,

[3] SCHENKER, H. *Ein Beitrag zu Ornamentik*. Vienna: Universal Edition, 1904. Revised 1908.

[4] SCHENKER, H. *Beethovens neunte Sinfonie: eine Darstellung des musikalischen Inhaltes unter fortlaufender Berücksichtigung auch des Vortrages unter der Literatur*. Vienna: Universal Edition, 1912.

[5] ROADS, CURTIS (2001). *Microsound,* p.vii and 20-28. Cambridge: MIT Press. ISBN 0-262-18215-7.

[6] TEDESCHINI LALLI, L. "Sound Stream as a Clock" *Int. J. Mod. Phys. B,* **18**, 793 (2004). DOI: 10.1142/S0217979204024410

[7] MORAVCSIK, M. J. *Musical Sound An Introduction to the Physics of Music. Kluwer Academic* 2001 SBN: 978-0-306-46710-3 (Print) 978-1-4615-0577-8 (Online)

[8] SETHARES, W. *Timbre, Tuning, Spectrum*, Scale Springer (1998) ISBN-13: 978-1852337971

[9] TEDESCHINI LALLI, L. “Matematiche e Musiche: un modello per gli strumenti musicali in bronzo” in: *Matematica e Cultura* 1999, Emmer, M. ed Springer

[10] TEDESCHINI LALLI, L. “Mathematics And Music: A Mathematical Model For Tuned Bronze Instruments” in *Proceedings Of The E.M.S. Forum Diderot Of The European Mathematical Society, Vienna Dec.1999*, pp 325-333

[11] SETHARES, W. A. *Rhythm and Transforms*. Springer Verlag, 2007. ISBN 9781846286391

[12] CANÉVET, G. “La localization auditive des sons dans l’espace” in *Le Son et l’Espace* Grame Aléas (1998) revised and updated version in G. Canévet “La localizzazione uditiva dei suoni nello spazio” in *Spazi sonori della Musica*, (2010) op.cit p.53

[13] MURRAY SCHAFER, R. 1977 and 1993 *The Soundscape; Our Sonic Enviroment and the Tuning of the World* Kopf

[14] VIDOLIN, A. “Spazi fisici e spazi acustici nella musica d’oggi” in : *Spazi Sonori*, op cit.

[15] FELD , S. *The Time of Bells* n.1,2,3,4,5 VoxLox (2005-2012)

[16] BAFFIONI, C. GUERRA, F. TEDESCHINI LALLI, L. « Music and Aleatory Processes » in *Proceedings of the "Interdisciplinary Colloquium on Stochastic Differential Equations*", Bielefeld, 12-13 October 1981

Current address

Laura Tedeschini Lalli, Professor of mathematical physics
Dipartimento di Architettura
Via della Madonna dei Monti 40 00184 Rome, Italy
e-mail: tedeschi@mat.uniroma3.it