

# THE PLENACOUSTIC FUNCTION, SAMPLING AND RECONSTRUCTION

Thibaut Ajdler\* and Martin Vetterli\*<sup>◊</sup>

\*Laboratory for Audio-Visual Communications (LCAV)

Swiss Federal Institute of Technology Lausanne (EPFL), 1015 Lausanne, Switzerland

<sup>◊</sup>Department of EECS, University of California at Berkeley, Berkeley CA 94720, USA

Email: {thibaut.ajdler, martin.vetterli}@epfl.ch

## ABSTRACT

In the present paper, we study the spatialization of the sound field in a room, in particular the evolution of room impulse responses as function of their spatial positions. The presented technique allows us to completely characterize the sound field in any arbitrary location if the sound field is known in a certain finite number of positions. The existing techniques usually make use of room models to recreate the sound field present at some point in the space. Our technique simply starts from the measurements of impulse responses in a finite number of positions and with this information the total sound field can be recreated. An analytical solution of the problem is given for different cases of spaces. Further, we determine the number and the spacing between the microphones needed to perfectly reconstruct the sound field up to a certain temporal frequency. The optimal sampling pattern for the microphone positions is given. Applications are also discussed.

## 1. INTRODUCTION

We introduce a function called the plenacoustic function. The plenacoustic function characterizes the sound field in space, e.g. inside a room. The idea of the plenacoustic function is similar to the one introduced by the plenoptic function already studied in the image processing world [1]. There, the idea is to consider the light field in the neighborhood of a scene. The plenoptic function records the light intensity of all the rays of light in this neighborhood. This function requires a huge amount of memory to be stored. Sampling of the plenoptic function is studied in [2].

We study a similar idea in the area of audio. We focus on the plenacoustic function in rooms. Therefore, we are interested in the room impulse responses to characterize what one would hear at any point in the room. Knowing this information, we can simply calculate the convolution of the sound produced by the source at some point in the room with the room impulse response from the source's position to the listener's position. This gives precisely what the sound is at the listener's position. The plenacoustic function is thus parameterized by the following factors:

$$P_a(S \text{ and } R \text{ position, characteristics of room, time})$$

with  $S$  and  $R$  being the source and receiver and the characteristics of the room being for example the reflection factors of the walls of the room.

The plenacoustic study answers the following question: "How many microphones do we need to place in the space in order to completely reconstruct the sound field at any position in the space?"

The name of the plenacoustic function has been for the first time mentioned in [3]. The first analysis of the function has been given in [4]. In the present paper, we briefly review some aspects explained in [4] and we develop new sampling and reconstruction techniques. Furthermore, we develop an analytical analysis of the problem in different cases. We finally use measurements in order to compare the theoretical and practical results.

The outline of the paper is the following. We briefly summarize the description of the plenacoustic function in Section 2. Section 2.1 introduces the model for the room impulse responses. We then explain the construction of the plenacoustic function in Section 2.2. In Section 2.3, we study the sampling of this function. Different sampling methods are compared to achieve better reconstruction in Section 2.4. An analytical solution of the plenacoustic function in different cases is then given in Section 3. Section 4 shows the experimental results obtained by measuring the room impulse responses in a room. The conclusions are finally drawn in Section 6.

## 2. PLENACOUSTIC FUNCTION

### 2.1. Room impulse response

In order to calculate the plenacoustic function in a room, we need to know the room impulse response at any point in the room. We use the image method discussed in [5]. The method is based on the creation of virtual sources in order to simulate the effect of the reflections on the walls. In the case of a rectangular rigid-walls room of size  $(L_x, L_y, L_z)$ , we get the following expression:

$$p(t, S, R) = \sum_{p=0}^7 \sum_{v=-\infty}^{\infty} \frac{\delta[t - |d_p + d_v|/c]}{4\pi|d_p + d_v|}, \quad (1)$$

where  $d_p = (x_s \pm x_r, y_s \pm y_r, z_s \pm z_r)$ ,  $d_v = (2nL_x, 2lL_y, 2mL_z)$  and  $(n, l, m)$  being an integer vector triplet. The room impulse response is a function of time and is dependent on the source's position  $S = (x_s, y_s, z_s)$  and the receiver's position  $R = (x_r, y_r, z_r)$ . The first sum shows that in a 3-dimensional field, 7 virtual sources are created in addition to the original source. The second sum shows that sound between two parallel rigid walls is infinitely reverberated. More general formulas taking into account the reflection factors of the walls are given in [5].

### 2.2. Construction of the plenacoustic function

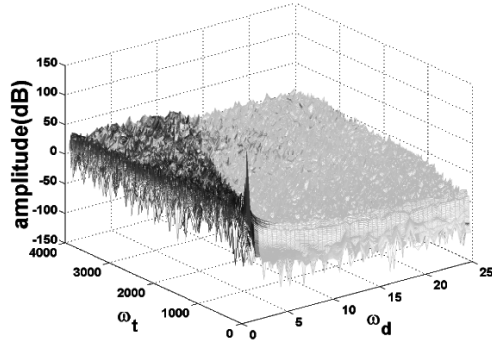
With these room impulse responses, we construct the plenacoustic function for a room. Let us consider (without loss of generality)

that we study the plenacoustic function along a line in the room. We can construct a 3-dimensional graph formed by all the room impulse responses at any position along the axis. In such a graph the two base axes are space and time. Space corresponds to the listener's position, while the time corresponds to the duration of the room impulse responses. Studying the plenacoustic function along a line in the room would require to store the room impulse responses for every single point along this line. The time continuous impulse responses should also be known. This data is impossible to store, therefore we need to sample the plenacoustic function.

### 2.3. Plenacoustic sampling

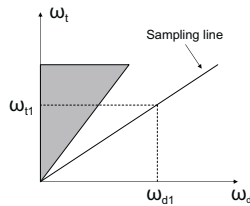
In order to sample the plenacoustic function, we need to sample the room impulse responses at a certain temporal sampling rate depending on the desired audio bandwidth. Further, by taking an evenly spaced finite number of impulse responses, we uniformly sample the plenacoustic function in space. When the plenacoustic function is sampled, repetitions of the spectrum occur [6].

The 2-dimensional Fourier transform (2D-FT) of the obtained plenacoustic function is shown in Fig. 1. In this figure, the two



**Fig. 1.** Two dimensional Fourier transform of the plenacoustic function.

base axes represent the corresponding axes of time and space, being a temporal and a spatial frequency. Temporal and spatial pulsations are defined as  $\omega_t = \frac{2\pi}{\Delta t}$  and  $\omega_d = \frac{2\pi}{\Delta d}$ , respectively.  $\Delta t$  is the sampling period of the impulse responses and  $\Delta d$  the sampling interval between the different positions of the measured impulse responses. The third axis is the amplitude (in dB) of the 2D-FT. We clearly see a triangular shape in this 2D-FT of the plenacoustic function. We can furthermore observe that the part outside of the triangle is of much lower amplitude than the part inside the triangle. The contents of the 2D-FT of the plenacoustic function can be schematized as in Fig. 2. As it can be seen in Fig. 2, the spatial fre-



**Fig. 2.** Schematic view of plenacoustic function in frequency domain.

quency support grows for increasing temporal frequencies. Spatial sampling of the plenacoustic function at some spatial sampling frequency (e.g.  $\omega_{d1}$  in Fig. 2) will lead to aliasing for all the temporal frequency above some corresponding temporal frequency ( $\omega_{t1}$ ). Perfect reconstruction becomes then impossible above this temporal frequency. Similarly, the sampling line at the right of the triangle in Fig. 2 shows for any temporal frequency, the minimal spatial sampling frequency to avoid aliasing. It can be shown that the sampling line with minimal coefficient of direction is given by:

$$\omega_d = \frac{2\omega_t}{c}, \quad (2)$$

with  $c$  the sound propagation speed in the air.

In order to understand physically the meaning of the triangle we have to consider a certain temporal frequency of  $v$  Hz. The wavelength of this frequency is  $c/v$  meters (m). Similarly to the Shannon theorem in time domain, we can say that in order to perfectly reconstruct a sine of  $c/v$  m we need to sample it at a period of maximum  $c/2v$  m in space. The spatial sampling frequency corresponding to this maximal period is  $u_s = \frac{1}{c/2v} = \frac{2v}{c}$ . The maximal spatial frequency that we can obtain is

$$u = u_s/2 = v/c. \quad (3)$$

We can verify this in Fig. 1 where the plenacoustic function has been sampled every 2 cm along a line in the room and where for a spatial frequency of  $12 \text{ m}^{-1}$  the corresponding temporal frequency lies around 4 kHz.

### 2.4. Reconstruction

Knowing the sound field at every point of the sampling grid, we apply the usual interpolation techniques [6] in order to reconstruct the sound field at any location. First, we need to upsample our time domain signal accordingly to the desired location. We then filter the upsampled plenacoustic function with an appropriate 2-dimensional filter. The value at the location of interest is then obtained by interpolation. The interpolation filter to be used is dependent on the sampling grid.

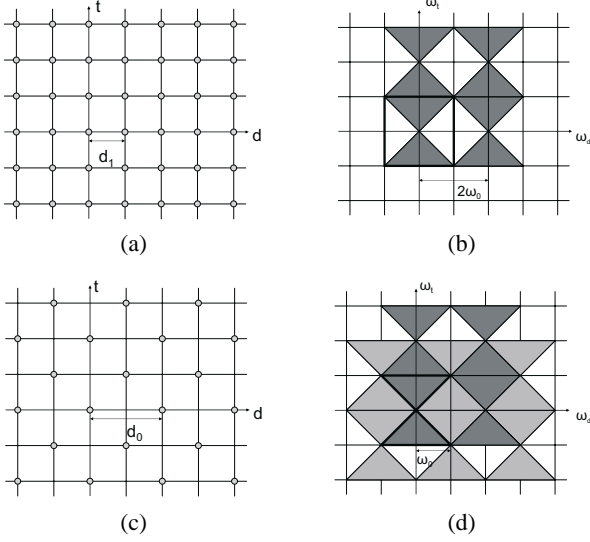
#### 2.4.1. Rectangular sampling

In the case of rectangular sampling, we just sample the plenacoustic function in time domain with a sampling grid shown in Fig. 3a. Convolution of the sampling grid with the spectrum of the plenacoustic function leads us to Fig. 3b. In this figure, we can also see the filter needed. It is just a rectangular filter. In Fig. 3b we observe that the spatial sampling frequency is  $2\omega_0$ . The corresponding spacing between the samples on the space axis is  $d_1 = \frac{2\pi}{2\omega_0} = \frac{\pi}{\omega_0}$ . In order to gain some processing time we need to find a better sampling.

#### 2.4.2. Quincunx sampling

Better performance can be achieved using quincunx sampling. We can actually fill the space lost in frequency domain in Fig. 3b, in order to fill the whole frequency space. This is achieved in Fig. 3d. The corresponding sampling grid in time domain is shown in Fig. 3c. In this case the filter used for interpolation is a fan filter [7, 8]. The filter is shown in Fig. 3d. In the quincunx sampling the spatial sampling pulsation is now only  $\omega_0$ . This corresponds to the distance between two samples on the space axis of  $d_0 = \frac{2\pi}{\omega_0}$ . This

shows us that using the quincunx sampling we need to sample the even microphones at even times while the odd microphones need to be sampled at odd times. This leads to a gain of factor 2 in the processing.



**Fig. 3.** (a) Rectangular sampling grid. (b) Plenacoustic spectrum with its repetitions for a rectangular sampling grid. (c) Quincunx sampling grid. (d) Plenacoustic spectrum with its repetitions for a quincunx sampling grid.

### 3. MATHEMATICAL DERIVATION OF THE PLENACOUSTIC FUNCTION

We have studied mathematically the 2D-FT of the plenacoustic function in different cases. An analytical formula can only be derived in a few cases. We analyze the example of a very long corridor. One axis can then be considered as infinite. We then study the evolution of the room impulse response along this axis. We fix the source's position and move the microphones ( $x_r$ ). The only variables are thus  $x_r$  and the time  $t$ . All the other variables are fixed. The plenacoustic function is:

$$f(x_r, t) = \sum_{p=0}^3 \sum_{m=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \frac{\delta(t - a/c)}{4\pi a}, \quad (4)$$

with  $a = \sqrt{(x_s - x_r)^2 + (y_s \pm y_r - 2mL_y)^2 + (z_s \pm z_r - 2lL_z)^2}$ . The sum on  $p$  models the 4 first images of the source [5]. We can calculate the 2D-FT of this plenacoustic function. The mathematical derivation can be found in [9]. We give a mathematical expression for the part of the data inside and outside of the triangle of the plenacoustic function in the frequency domain:

for  $u < v/c$

$$F(u, v) = -\frac{g}{4} \sum_{p=0}^3 \sum_{m=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} [N_0(s) + jJ_0(s)], \quad (5)$$

for  $u > v/c$

$$F(u, v) = \frac{g}{2\pi} \sum_{p=0}^3 \sum_{m=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} K_0 \left( 2\pi d \sqrt{u^2 - \left(\frac{v}{c}\right)^2} \right), \quad (6)$$

where  $d = \sqrt{(y_s \pm y_r - 2mL_y)^2 + (z_s \pm z_r - 2lL_z)^2}$ ,  $g = e^{-j2\pi u x_s}$  and  $s = 2\pi d \sqrt{\left(\frac{v}{c}\right)^2 - u^2}$ . The part of the expression that represents the inside of the triangle is a sum of Bessel and Neuman (Bessel function of second kind) functions. For  $u > v/c$ , which corresponds to the part outside of the triangle of the plenacoustic function in frequency domain, the 2D-FT is a sum of modified Bessel functions of the second kind. The order of the function is zero and this is a typically decreasing function. We are thus able to determine how the function is decreasing when, for the same temporal frequency, we increase the spatial frequency. Another interesting point is that the parameter  $d$  has an influence on the decreasing part of the plenacoustic function.  $d$  is defined for each source and it gets larger for the virtual sources that are far from the listener's position. So, when  $d$  gets larger, the Bessel function corresponding to the considered virtual source decreases. That means that this virtual source will have a smaller influence on the decreasing part of the plenacoustic function. Only the first peaks of the room impulse response will have an important influence for the part outside of the triangle. Therefore it is also interesting to study the plenacoustic function in free field. In this case the plenacoustic function is simply given by

$$f(x_r, t) = \frac{\delta(t - a/c)}{4\pi a}, \quad (7)$$

with  $a = \sqrt{(x_s - x_r)^2 + (y_s - y_r)^2 + (z_s - z_r)^2}$ . In this case, the analytical derivation is similar. The solution is the same as in (5) and (6) but does not contain any sum. The solution outside the triangle is just one modified Bessel function of the second kind. Therefore, the 2D-FT of the plenacoustic function will decrease faster in the case of large distance between source and receiver. Remark that  $d = 0$  leads to infinite values in (5) and (6). This is simply the consequence of the fact that loudspeaker and microphone cannot have the same position in the image method [5].

### 4. EXPERIMENTS

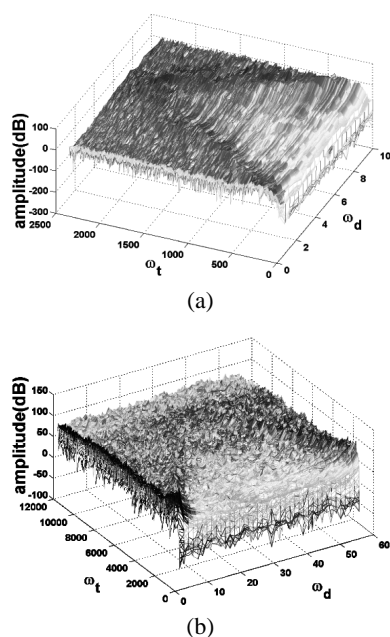
After observing the plenacoustic function by simulating different room impulse responses with different reflection factors using the formulas in [5], we made some experiments in order to check if the triangle shape of the 2D-FT of the plenacoustic function was also observed with measured room impulse responses. We used two sets of measures.

The first set of data comes from measurements made in the varechoic chamber at the Bell Laboratories, Lucent Technologies. These measures were done using a logarithmic sweep signal as excitation signal [10]. The room impulse responses are given for an array of 22 microphones where the spacial spacing between the microphones was 10 cm.

We measured the second set of impulse responses using the Maximum Length Sequences (MLS) [11] technique. The measurements were made in a sound insulated room (LCAV audio room at EPFL). 40 measures were taken with 1.8 cm spatial spacing.

For the first data, we can reconstruct the signal up to a temporal frequency of 1.7 kHz, as can be seen using (3). For the second

set of data we can reconstruct the sound field up to a temporal frequency of 9.2 kHz. The 2D-FT of the plenacoustic functions of the two sets of data are shown in Fig. 4a and 4b. In these two figures, we show the whole spatial axis until the sampling frequency (and not until Nyquist) to see clearer until where we get the triangle shape and when the aliasing begins. We can see in Fig. 4a that the aliasing begins around 1.7 kHz while in Fig. 4b the aliasing begins around 9 kHz. We get thus a good correspondence between theoretical and practical results.



**Fig. 4.** Measured plenacoustic functions in frequency domain in (a) the varechoic Bell Labs chamber, (b) the LCAV audio room.

## 5. APPLICATIONS

First, the plenacoustic function can be used to predict what the sound is at any point for a moving person in a room thanks to fast interpolation techniques. If we study the room impulse response with a supplementary parameter being the directionality, we could convolve the impulse response from a source to the listener's position with the corresponding head-related transfer function (HRTF) [12]. This would allow us to take in account the architectural effect of the room impulse response and the perceptual effect of the HRTF. A second application is the achievement of a better convergence in echo cancellation problems with moving talker. If information about the talker's position is known (use of camera), we can predict the room impulse response at the speaker's position and use this as starting value for the adaptive filter. A third application is the use of the triangular shape and decay of the 2D-FT of high quality plenacoustic functions for storage or coding of the information. After quantization, all information outside of the triangle is likely to be zero. The data to code is thus much reduced. Finally, we are currently working on sampling of the HRTF. We want to answer the question: "What is the minimal sampling elevation and azimuthal angle necessary to be able to reconstruct the full plenHRTF function?"

## 6. CONCLUSION

In this article, we have introduced the plenacoustic function. The plenacoustic function characterizes the sound field at any point in space. We then studied the sampling of this function and showed how to reconstruct the sound field between the microphones using interpolation techniques. We are now able to predict the number of microphones to place in a room to perfectly reconstruct the sound field at any point in the room up to a certain temporal frequency. We also showed the relation between the critical spacing of the microphones and the corresponding maximal frequency before aliasing occurs. An analytical derivation of the plenacoustic function was then given in different cases. An optimal sampling pattern in time and space is given in order to achieve the best performance for reconstruction of the sound field. Finally, we presented experimental results and compared theoretical results with measurements.

## 7. ACKNOWLEDGEMENT

We would like to thank Dr. Aki Härmä who made us available some of his measurements of the varechoic room at Bell Laboratories, Lucent Technologies.

## 8. REFERENCES

- [1] E.H. Adelson and J.R. Bergen, "The plenoptic function and the elements of early vision," in *Computational Models of Visual Processing*, pp. 3–20. MIT Press, 1991.
- [2] J.X. Chai, X Tong, S.C. Chan, and H.Y. Shum, "Plenoptic sampling," in *Proceedings of the conference on Computer graphics*, 2000, pp. 307–318.
- [3] M. Kubovy and D. Van Valkenburg, "Auditory and visual objects," *Cognition*, vol. 80, pp. 97–126, 2001.
- [4] T. Ajdler and M. Vetterli, "The plenacoustic function and its sampling," in *Benelux Workshop on Model Based Processing and audio Coding*, 2002.
- [5] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *JASA*, vol. 65, 1979.
- [6] M. Vetterli and J. Kovačević, *Wavelets and Subband Coding*, Signal Processing. Prentice Hall, A.V. oppenheim edition, 1995.
- [7] P.P. Vaidyanathan, *Multirate Filter Banks and Wavelets*, Prentice Hall, 1992.
- [8] C. Zhang and T. Chen, "Generalized plenoptic sampling," Tech. Rep. Amp 01-06, Carnegie Mellon University, September 2001.
- [9] T. Ajdler and M. Vetterli, "The plenacoustic function," Tech. Rep., EPFL, 2002.
- [10] A. Harma, T. Lokki, and V. Pulkki, "Drawing quality maps of the sweet spot and its surroundings in multichannel reproduction and coding," *21st Conference on Architectural Acoustics and Sound Reinforcement, AES*, pp. 317–325, June 2002.
- [11] J. "Borish and J.B. Angell, "An efficient algorithm for measuring the impulse response using pseudorandom noise," *J. Audio Eng. Soc.*, vol. 31, no. 7, 1983.
- [12] J. Blauert, *Spatial hearing*, MIT press, 2001.