The Plenacoustic Function and Its Sampling

Thibaut Ajdler, *Member, IEEE*, Luciano Sbaiz, *Member, IEEE*, and Martin Vetterli, *Fellow, IEEE*

Abstract—The spatialization of the sound field in a room is studied, in particular the evolution of room impulse responses as a function of their spatial positions. It was observed that the multidimensional spectrum of the solution of the wave equation has an almost bandlimited character. Therefore, sampling and interpolation can easily be applied using signals on an array. The decay of the spectrum is studied on both temporal and spatial frequency axes. The influence of the decay on the performance of the interpolation is analyzed. Based on the support of the spectrum, the number and the spacing between the microphones is determined for the reconstruction of the sound pressure field up to a certain temporal frequency and with a certain reconstruction quality. The optimal sampling pattern for the microphone positions is given for the linear, planar and three-dimensional case. Existing techniques usually make use of room models to recreate the sound field present at some point in the space. The presented technique simply starts from the measurements of the sound pressure field in a finite number of positions and with this information the sound pressure field can be recreated at any spatial position. Finally, simulations and experimental results are presented and compared with the theory.

Index Terms—Interpolation, plenoptic function, room impulse response, sampling, sound pressure field sampling.

I. INTRODUCTION

ASSUME you are in a concert hall, and you want to faithfully describe the sound pressure field at any location in the hall. If you record the acoustic event with an array of microphones, how many do you need to be able to reproduce the sound pressure field at any point?

Conversely, assume a virtual acoustical environment, where sources are moving, while the sound pressure is measured in a particular spot. How finely do you need to simulate the acoustic impulse responses to be able to place the source at any location?

The answers to the above questions, as well as related ones, lie in the spatio–temporal acoustic sound pressure field and its properties. We call this field the *plenacoustic function* (PAF) in reference to the plenoptic function introduced by Adelson and

Bergen [1] and which defines "all views in a room." More precisely, the plenoptic function is given by a seven dimensional function $f(x, y, z, \theta, \Omega, \lambda, t)$, which describes the intensity of the light field seen at location (x, y, z) when looking in direction (θ, Ω), at wavelength λ and time t. Note that, in this case, one can define a very precise direction (θ, Ω) of arrival of a light ray. This is due to the very small wavelength of light (on the order of 100 nm). Therefore, it is possible to build devices measuring a good approximation of the plenoptic function. In the acoustical case, the wavelengths are 10^6 time larger. This makes it unfeasible in practice to construct a device that would measure the sound pressure at any arbitrary position and the component of its gradient due to a source originating from a precise direction. Taking into account these physical differences existing between the optical and the acoustical case, we define the PAF $p(x, y, z, t)$ as the sound pressure recorded at location (x, y, z) and time t given an acoustic event in a room. Note that from the knowledge of the sound pressure in every point in the three dimensions, one can calculate gradients of pressure in all directions and obtain the particle velocity vector for every point [2]. From this information, one can transform the omnidirectional PAF into a directional PAF. This is achievable since the PAF measures the instantaneous pressure containing the phase information. This information is missing in the intensity measurement given by the plenoptic function. Note that the velocity particle can only be derived along all directions when the PAF is known in all three spatial dimensions. In a practical case, one would measure the PAF along a line. The knowledge of the particle velocity as well as the directionality would then only be partial. In the rest of the paper, the PAF is to be considered as an omnidirectional PAF. It is the solution of the scalar acoustic-wave equation given by

$$\nabla^2 p - \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2} = s(x, y, z, t) \qquad (1)$$

where c represents the speed of sound propagation and $s(x, y, z, t)$ represents a distribution of sources located in space.

Often, one will be concerned with the case of a single point source. Namely, for a given source S, we denote the room impulse response (RIR) at location (x, y, z) by $h(x, y, z, t)$, and then, if the source generates a signal $s(t)$, the PAF is

$$p(x, y, z, t) = \int_{-\infty}^{\infty} s(\tau) h(x, y, z, t - \tau) \mathrm{d}\tau. \qquad (2)$$

When the emitted sound is a Dirac pulse, the PAF becomes simply the spatio–temporal RIR. The PAF is then called the Green's function [2]. Assuming that our system is linear and time invariant, we consider the presence of multiple sources as the superposition of single sources. The sound pressure field can then be regarded as the sum of all point sources convolved with

their spatio–temporal RIR. Without loss of generality, the analysis of the PAF will be conducted by considering the excitation signal as a Dirac impulse.

From the view of the physicist, the PAF is simply the solution of the wave equation with appropriate boundary conditions, and a given driving function. From the point of view of the numerical analyst, the system would be very complex for any reasonable room, even for very simple cases. For the signal processor, acoustic RIRs have been studied, measured, and simulated for many scenarios, and it is thus natural to study the PAF globally. A natural question for a signal processor is of course the sampling question: Is there a discrete set of points in time and space from which the full PAF can be reconstructed? The equivalent question for the plenoptic function was posed and solved by Chai *et al.* [3] with further results from Zhang *et al.* [4]. For the time dimension, bandlimited sources are assumed to allow sampling in time. The more interesting question is of course sampling in space, which directly relates to the number of microphones necessary to acquire the PAF. When considering far field assumptions, it can be shown [5] that the support of the PAF is bandlimited to a spatial frequency ϕ, which is related to the temporal frequency ω in a linear manner[1]

$$\phi = \frac{\omega}{c}. \qquad (3)$$

Thus, if the time domain signal is bandlimited to ω_0, then the spatial frequency is limited to ω_0/c, and the far field PAF can be sampled with any spatial distance Δx satisfying

$$\Delta x > \frac{c\pi}{\omega_0}. \qquad (4)$$

Without the far field assumption, the PAF is not perfectly bandlimited anymore and contains energy in the region $\phi \geq (\omega_0/c)$. In this paper, the PAF is studied without the far-field assumption and its spectral characteristics are derived both in the case of free field and in reverberating rooms. From this analysis, we derive a precise quantitative sampling theorem which trades off spatial sampling density for signal reconstruction quality.

The name "plenacoustic function"[2] has been mentioned for the first time in [6]. The first analysis of the function has been given in [7]–[9]. As mentioned earlier, previous literature exists on the bandlimitedness of the spatial sound pressure field. In [5], [10], [11] using microphone arrays, the sound pressure field is studied both along the temporal and the spatial dimensions. The spectrum of the spatio–temporal wave equation is studied under the far field assumption which simplifies the obtained sampling results [5], [10], [11]. Recently, and in parallel to our work, Coleman [12], [13] has investigated the wideband electromagnetic impulse response in far field, deriving sampling results under this assumption for linear arrays and planar arrays of sensors. In [14], Berkhout studies in detail the extrapolation of waves in the field of seismic wave theory. Results on the sampling and the extrapolation of wave fields are obtained using

the representation of the wave field by its spatio–temporal spectrum. Also practical aspects such as the aperture size of the array are discussed. Recent techniques have shown interesting extrapolation results using one-dimensional (1-D) microphone arrays (mostly circular arrays), but limitations occur when trying to extrapolate real three-dimensional (3-D) RIRs [16]. Similarly to the extrapolation of the sound field, a technique called wave field synthesis (WFS) [15] has been presented. From the knowledge of the sound field in a region of space, the WFS reproduces the sound field in other regions of space. The WFS is based on the Huygens' principle stating that the propagation of a wave through a medium can be described by adding the contributions of all secondary sources positioned along a wave front. Measuring the sound field on an infinite plane of microphones would allow us to reproduce it accurately at any point of the source-free half space. The theory has been adapted to consider 1-D loudspeaker arrays and still be able to accurately reproduce the sound field in the largest possible area [17]. Dual to the WFS is the wave field analysis [18] where the ray parameter versus intercept time-domain representation is used. This representation is shown to be equivalent to taking the linear Radon transform of the spatio–temporal RIRs. The method gives an improved insight in the structure of complex sound pressure fields by being able to separate the different contributions of the sound pressure field (direct path and each reflection on the walls).

Our contribution to the existing theory is the detailed study of the decay of the spatio–temporal sound pressure field spectrum along the spatial and temporal frequency axes both in free field and in rooms with reverberation. Given this characterization, sampling theorems and interpolation formulas can be derived. In extension of the results considering far field assumptions, we introduce a signal-to-noise ratio (SNR) on the reconstruction of the sound pressure field at the different spatial positions. The theoretical results are verified both through simulation experiments and through measurements in actual rooms.

Beyond the fundamental interest of characterizing precisely the PAF and its sampling, the results are useful in spatial audio applications. For example, it indicates to what extent a microphone array can be used to interpolate any spatial location. Or conversely, how many spatial positions of a source are needed to synthesize arbitrary positions for a virtual source. This principle can be applied to the interpolation of head-related transfer functions (HRTFs) [19]. HRTFs are filters measuring the effects of head, pinnae and torso of a person on the sound measured at the entrance of the ear canal. HRTFs are measured in anechoic chambers using circular arrays of loudspeakers. The impulse responses as functions of time and loudspeaker positions correspond to the PAF defined on circles and is studied in [20].

The outline of the paper is the following. In Section II, the PAF and its construction is presented. Section II-A reviews RIRs and how they can be simulated, while Section II-B constructs the space–time representation. Section III studies the spectrum of the PAF on a line in a room. Its spatial and temporal frequency decay are described in Section III-A and III-B, respectively. Section IV studies the sampling of the PAF. The sampling theorem of the PAF is presented in Section IV-A followed by a discussion on different sampling patterns in Section IV-B. Based on the sampled function, one would like to reconstruct the

---

[1]In acoustics as well as in some array signal processing books, the spatial and temporal frequencies are expressed as wave numbers. Then, ω/c is usually denoted as k and the spatial frequency ϕ as k_x.

[2]Note that the plenacoustic and plenoptic functions are expressions mixing Greek and Latin roots. The Latin expression would be "the plenaudio function" while the Greek expression would be "the panacoustic function."
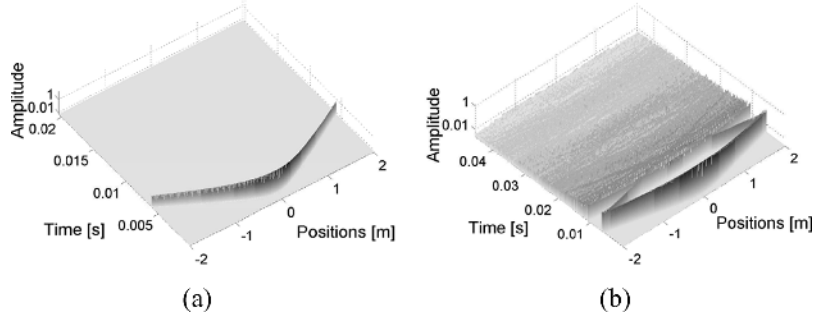
Fig. 1.   PAF in time and space: (a) in free field and (b) inside a room.

field in every possible position which is shown in Section IV-C. Limitations due to the finite length of the array are taken in account in Section V. The theory presented in this paper is then verified using simulations in Section VI-A and measurements done in real environments in Section VI-B. Section VII is devoted to the generalization of the PAF to multidimensional spatial positions. The conclusions are drawn in Section VIII.

II. CONSTRUCTION OF THE PLENACOUSTIC FUNCTION

To study the sound pressure field along a line of microphones in a room, one needs to study the sound pressure field from every possible source position in the room to any possible microphone position on the line. For simplicity, the technique is presented for a single source but it will be shown later that the technique works as well for multiple sources. Consider a source S emitting a signal $s(t)$. The microphones located on the line will not record exactly $s(t)$. The sound at microphone m_1 is $s(t)$ convolved with the RIR corresponding to the direct path between S and m_1, followed by a possibly infinite number of reflections on the walls (each microphone will receive a sum of delayed and attenuated versions of $s(t)$). At another microphone position m_2, the recorded signal will be different since the RIR from S to m_2 is different than the RIR from S to m_1. The only parameter changing between the different spatial positions is the RIR. Therefore, the rest of the analysis of this paper will be focused on the spatial evolution of the RIRs. Also for that reason, the name of the PAF will be used for the rest of the paper as the spatio–temporal RIRs.

A. Modeling the Room

In order to construct the PAF in a room, the RIRs at any point in the room need to be known. The image method discussed in [21] for the simulations of RIRs is used. 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), the RIRs are given by [21]

$$p(t, S, M) = \sum_{p=0}^{7} \sum_{v=-\infty}^{\infty} \frac{\delta\left(t - \|d_p + d_v\|/c\right)}{4\pi\|d_p + d_v\|} \qquad (5)$$

where $d_p = (x_s \pm x_m, y_s \pm y_m, z_s \pm z_m)$, $d_v = (2lL_x, 2nL_y, 2oL_z)$, (l, n, o) is an integer vector triplet. The RIR is a function of time and is dependent on the source $S = (x_s, y_s, z_s)$ and the microphone position

$M = (x_m, y_m, z_m)$. The first sum shows that in a 3-D field, seven virtual sources are created in addition to the original source due to the first reflections on three walls. 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 [21]. Special attention has to be given to the problem of the quantization rounding in the computation of the RIR. In [21], the delay corresponding to each virtual source was rounded to the closest sample in time in order to reduce the complexity of the simulations. This leads to aliasing in time and space. In our simulations, each dirac has been replaced by a sinc function of appropriate bandwidth delayed with the exact non integer delay. This removes the aliasing effect. However, as the sinc functions have a very slow decay in time, one has to consider long enough RIRs to allow the sincs to sufficiently vanish.

B. Space–Time Representation

With the RIRs as defined in (5), the PAF is constructed for a line in the room. In that case, one can construct a two-dimensional (2-D) function by gathering all the RIRs at any position on the line, leading to a 2-D continuous function of space and time. Space represents the position, time being the duration of the RIR. This representation is shown in Fig. 1(a) when a pulse is recorded on a line of microphones in free field and in Fig. 1(b) for the case of a room.

III. SPECTRUM OF THE PAF ON A LINE

In this section, the PAF on a line and its associated spectrum are studied. An analytical expression of the 2-D Fourier transform (2D-FT) of the PAF is given. Further, the spatial and temporal frequency decay of the spectrum of the PAF are studied in the case of a single source and further in the case of a rectangular room.

To study the decay of the PAF, the following notations are used:

- $f(x) \sim g(x)$ means that $\lim_{x \to \infty} (f(x)/g(x)) = 1$;
- $f(x) \in O(g(x))$ means that there exist positive constants c and k, such that $|f(x)| \le cg(x), \forall x \ge k$.

A. Spatial Decay of the Spectrum of the PAF

An analytical expression for the 2D-FT of the PAF is given. This is first reviewed for the free-field case, followed by a general formula in the case of a rectangular room.
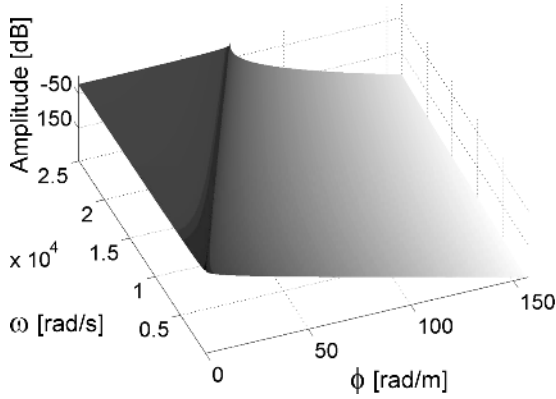
Fig. 2. Theoretical 2-D spectrum of the PAF according to (8).

1) Free-Field Case: The evolution of the RIR along the x axis is studied. The PAF in space and time domain is given by the following formula [2]:

$$p(x,t) = \frac{\delta\left(t - \frac{\sqrt{(x-x_s)^2+(y_m-y_s)^2+(z_m-z_s)^2}}{c}\right)}{4\pi\sqrt{(x-x_s)^2+(y_m-y_s)^2+(z_m-z_s)^2}}. \quad (6)$$

The x component of the microphone is only varied. For simplicity, the subscript in the variable x_m was removed, denoting it by x. The variables y_m, z_m, x_s, y_s, and z_s are constant. Calling $d^2 = (y_m - y_s)^2 + (z_m - z_s)^2$, (6) is rewritten as

$$p(x,t) = \frac{\delta\left(t - \frac{\sqrt{(x-x_s)^2+d^2}}{c}\right)}{4\pi\sqrt{(x-x_s)^2+d^2}}. \quad (7)$$

The 2D-FT of this function is calculated in Appendix I. The obtained result for $\omega \in \mathbb{R}^+$ and $\phi \in \mathbb{R}$ is[3]

$$P(\phi,\omega) = -\frac{j}{4}e^{-j\phi x_s}H_0^*\left(d\sqrt{\left(\frac{\omega}{c}\right)^2 - \phi^2}\right) \quad (8)$$

with ϕ and ω being, respectively, the spatial and temporal frequencies. The magnitude of (8) is plotted for $d = 1$ in Fig. 2 for positive frequencies. H_0^* represents the complex conjugate of the zero-order Hankel function of the first kind. This function is infinite at zero. Therefore, when either $d = 0$ or $|\phi| = \omega/c$, the plenacoustic spectrum becomes infinite. When $d = 0$, it corresponds to the situation where the source is located on the line of the microphones. For the case of $|\phi| = \omega/c$, a more intuitive explanation is given. Consider a sinusoid of temporal frequency ω rad/s emitted from a certain position. The signal acquired by the microphones located at positions tending to infinity is at one instant an attenuated sinusoid of spatial frequency ω/c rad/m. For these microphone positions, the source appears as being on the line. This explains the infinite value of the spectrum for this spatial frequency.

The values where $|\phi| > (\omega/c)$ correspond to the evanescent mode of the waves. The waves lose their propagating character to become exponentially fast decaying waves [14], [22]. There-

[3]Since $p(x,t)$ is a real function, one has that $P(-\phi, -\omega) = P^*(\phi, \omega)$, with P^* the complex conjugate of P. To simplify the notation, all further derivations are done for $\omega \in \mathbb{R}^+$.

fore, considering $\omega, \phi \in \mathbb{R}$, most of the energy is contained in the part of the spectrum satisfying

$$|\phi| \le \frac{|\omega|}{c}. \quad (9)$$

This result will be used later in the sampling of the PAF. As the spectrum is decaying very fast along the spatial frequency axis, one will be able to derive a sampling theorem to sample and reconstruct the PAF along the spatial axis (see Section IV). Note also that for $|\phi| \ge (\omega/c)$, the argument of the Hankel function in (8) becomes imaginary and (8) can be rewritten as

$$P(\phi,\omega) = \frac{1}{2\pi}e^{-j\phi x_s}K_0\left(d\sqrt{\phi^2 - \left(\frac{\omega}{c}\right)^2}\right) \quad (10)$$

where K_0 is a modified Bessel function of the second kind and order zero. The modified Bessel function of the second kind has the following asymptotical behavior (see [22]):

$$K_0(x) \sim \sqrt{\frac{\pi}{2x}}e^{-x}. \quad (11)$$

Note that further numerical computations show that

$$K_0(x) \le \sqrt{\frac{\pi}{2x}}e^{-x} \quad \text{for} \quad x > 0. \quad (12)$$

For large ϕ, (10) can be rewritten using (11) as

$$P(\phi,\omega) \sim \frac{1}{2\pi}e^{-j\phi x_s}\sqrt{\frac{\pi}{2d\sqrt{\phi^2 - \left(\frac{\omega}{c}\right)^2}}}e^{-d\sqrt{\phi^2 - \left(\frac{\omega}{c}\right)^2}}. \quad (13)$$

For a fixed $\omega = \omega_0$, (13) asymptotically behaves as

$$P(\phi,\omega_0) \sim \frac{e^{-j\phi x_s}}{2\sqrt{\pi}}\frac{e^{-d\phi}}{\sqrt{d\phi}}. \quad (14)$$

The decay along the spatial frequency axis is faster than exponential.

2) Rectangular Room: In the case of a rectangular room of size (L_x, L_y, L_z) with perfectly reflecting walls, all the reflections are considered as virtual sources as explained in Section II-A and the superposition principle is applied. The expression for the PAF is then given by (5). Each virtual source leads to a spectrum that follows (8). The total spectrum of the PAF is the sum of the spectra of each virtual source taken separately, leading to an infinite sum. One would like to know how this sum is decaying for large spatial frequencies.

Results on the decay of the spatial frequency are presented in the easier case of all the virtual sources located in the plane. Similar results can be obtained in the general case of sources located in space.

The image model given in Fig. 3 was used. Our original source is s_1 (with coordinates (x_{s_1}, y_{s_1})) and in its immediate neighborhood, one can see three other virtual sources (s_2, s_3 and s_4). These four mother sources will create an infinite number of repetitions to form all the virtual sources in the plane. These four mother sources will be repeated in the x and y direction with a periodicity of $2L_x$ and $2L_y$, respectively.

As the room has finite size, an infinite line of microphones cannot be considered. The microphone line is considered to
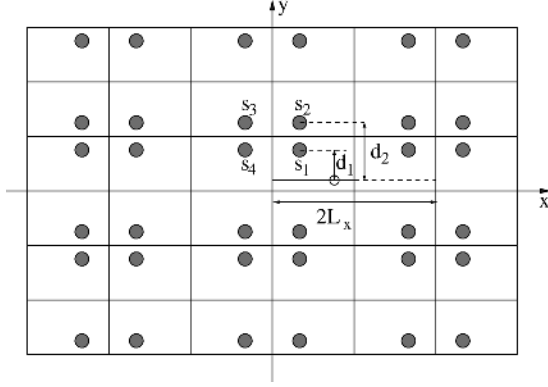
Fig. 3. Image source model with the original source s_1 and all the other virtual sources.

cover the whole length of the room. In Fig. 3, the line is parallel to the x axis.

Define the distances from sources s_1 and s_2 to the line of microphones as d_1 and d_2, respectively (with $d_1 \leq d_2$). The other distances for the sources s_4 and s_3 to the line are in this case also d_1 and d_2.

An interesting aspect of this construction is that by the $2L_x$ periodicity of the source positions along the x axis, the sound recorded on an infinite line is also periodic with period $2L_x$. Further, using the symmetry of the construction, one realizes that the sound heard at position a is the same as the one heard at position $2L_x - a$ with $a \in [0, L_x]$.

By sampling the spectrum of the PAF on the ϕ axis on the multiples of $\phi_0 = \pi/L_x$, one exactly obtains the Fourier series of the sound pressure field recorded on a line from 0 to $2L_x$ when this sound pressure field is $2L_x$ periodic.

Considering the four mother sources (s_1 to s_4) with their periodic repetitions along the x axis, the discrete spectrum of the PAF can be expressed as follows (for large $n\phi_0$):

$$P(n\phi_0, \omega_0) \sim \left(\frac{e^{-jn\phi_0 x_{s_1}}}{2\sqrt{\pi}} + \frac{e^{-jn\phi_0 x_{s_2}}}{2\sqrt{\pi}} \right) \times \left(\frac{e^{-d_1 n\phi_0}}{\sqrt{d_1 n\phi_0}} + \frac{e^{-d_2 n\phi_0}}{\sqrt{d_2 n\phi_0}} \right). \quad (15)$$

Therefore

$$P(n\phi_0, \omega_0) \sim C_1(n) \left(\frac{e^{-d_1 n\phi_0}}{\sqrt{d_1 n\phi_0}} + \frac{e^{-d_2 n\phi_0}}{\sqrt{d_2 n\phi_0}} \right) \quad (16)$$

with $C_1(n)$ a bounded function in n.

Consider now the $2L_y$ periodic repetitions of the sources along the y axis. Call the sources $s_{1,i}$ the sources with coordinates $(x_{s_1}, y_{s_1} + i2L_y)$ and similarly $s_{2,i}$ the sources with coordinates $(x_{s_2}, y_{s_2} + i2L_y)$. Call $D_{1,i}$ the distances between the line of microphones and the sources $s_{1,i}$, and $D_{2,i}$ the distances between the line of microphones and the sources $s_{2,i}$.

$$D_{1,i} = |d_1 + i2L_y| \quad (17)$$
$$D_{2,i} = |d_2 + i2L_y|. \quad (18)$$

When considering all the source repetitions in the x and y directions, the spectrum becomes

$$P(n\phi_0, \omega_0) \sim C_1(n) \sum_{i=-\infty}^{\infty} \left(\frac{e^{-D_{1,i} n\phi_0}}{\sqrt{D_{1,i} n\phi_0}} + \frac{e^{-D_{2,i} n\phi_0}}{\sqrt{D_{2,i} n\phi_0}} \right). \quad (19)$$

The right-hand side of (19) can be upperbounded by

$$C_1(n) \sum_{i=0}^{\infty} \left(\frac{e^{-(d_1 + i2L_y)n\phi_0}}{\sqrt{d_1 n\phi_0}} + \frac{e^{-(d_2 + i2L_y)n\phi_0}}{\sqrt{d_2 n\phi_0}} \right.$$
$$\left. + \frac{e^{-(d_1' + i2L_y)n\phi_0}}{\sqrt{d_1' n\phi_0}} + \frac{e^{-(d_2' + i2L_y)n\phi_0}}{\sqrt{d_2' n\phi_0}} \right) \quad (20)$$

with $d_1' = 2L_y - d_1$ and $d_2' = 2L_y - d_2$. Since $d_1 \leq d_2 \leq L_y$, (20) can be upperbounded by

$$\frac{C_1(n) \left(e^{-d_1 n\phi_0} + e^{-d_2 n\phi_0} + e^{-d_1' n\phi_0} + e^{-d_2' n\phi_0} \right)}{(1 - e^{-2L_y n\phi_0})\sqrt{n d_1 \phi_0}}. \quad (21)$$

Asymptotically for large n, the above expression is of the following order:

$$P(n\phi_0, \omega_0) \in O\left(\frac{e^{-d_1 n\phi_0}}{\sqrt{n}} \right). \quad (22)$$

This shows that for a reverberant room, the decay is faster than exponential when the line of microphones is parallel to a wall.[4]

B. Temporal Frequency Decay

The study of the temporal frequency decay is of interest to fully characterize the plenacoustic function. Nevertheless, in most cases sounds are bandlimited along the temporal frequency due to the bandwidth of the emitters and receivers. Therefore, the results of this section will be briefly presented since they only are interesting from a theoretical point of view. The more detailed analysis of the presented results can be found in [23]. Similarly to Section III-A, results on the temporal frequency decay for the free-field case are first presented before generalizing them for a rectangular room.

1) Free Field: The spectrum of the PAF is given by expression (8). The asymptotic behavior of the Hankel function is given by [22]

$$H_0(x) \sim \sqrt{\frac{2}{\pi x}} e^{j(x - \pi/4)}. \quad (23)$$

For large ω, (8) can be rewritten using (23) as

$$P(\phi, \omega) \sim -\frac{j e^{-j(\phi x_s - \pi/4)}}{2\sqrt{2\pi}} \frac{e^{-jd\sqrt{\left(\frac{\omega}{c}\right)^2 - \phi^2}}}{\sqrt{d\sqrt{\left(\frac{\omega}{c}\right)^2 - \phi^2}}}. \quad (24)$$

---

[4]The case where the line of microphones is not covering the whole length of the room can be seen as a windowing and is discussed in Section V. The case where the line is not parallel to the wall is studied in [23]. There, the line is extended along the periodic repetitions of the room and it is shown that the measured sound field has a spectrum decaying slightly faster than an exponential. The restriction of this infinite line inside the room can then again be seen as a windowing as explained in Section V.

Considering a fixed $\phi = \phi_0$, (8) asymptotically behaves as

$$P(\phi_0, \omega) \sim -\frac{j\sqrt{c}}{2\sqrt{2\pi}} e^{-j(\phi_0 x_s - \pi/4)} \frac{e^{-jd\frac{\omega}{c}}}{\sqrt{d\omega}}. \qquad (25)$$

Therefore

$$P(\phi_0, \omega) \sim \frac{C_2(\omega)}{\sqrt{\omega}} \qquad (26)$$

with $C_2(\omega)$ a bounded function of ω. This last relation shows that the decay of the PAF spectrum along the temporal frequency is, up to a constant, asymptotic as $1/\sqrt{\omega}$.

2) Rectangular Room: In the case of a rectangular room, one follows the same construction as in Section III-A-2. Considering the four mother sources (s_1 to s_4) in Fig. 3 with their periodic repetitions along the x axis, the discrete spectrum of the PAF can be expressed as follows (for large ω and a finite $n = n_0$):

$$P(n_0\phi_0, \omega) \sim \frac{-j\sqrt{c}e^{j\pi/4}}{2\sqrt{2\pi}} \left(e^{-jn_0\phi_0 x_{s_1}} + e^{-jn_0\phi_0 x_{s_2}} \right)$$

$$\times \left(\frac{e^{-jd_1\frac{\omega}{c}}}{\sqrt{d_1\omega}} + \frac{e^{-jd_2\frac{\omega}{c}}}{\sqrt{d_2\omega}} \right) \qquad (27)$$

$$\sim C_3(n_0) \left(\frac{e^{-jd_1\frac{\omega}{c}}}{\sqrt{d_1\omega}} + \frac{e^{-jd_2\frac{\omega}{c}}}{\sqrt{d_2\omega}} \right) \qquad (28)$$

with $C_3(n_0)$ being independent of ω.

Considering the $2L_y$ periodic repetitions of the sources along the y axis, one obtains

$$P(n_0\phi_0, \omega) \sim C_3(n_0) \sum_{i=-\infty}^{\infty} \left(\frac{e^{-jD_{1,i}\frac{\omega}{c}}}{\sqrt{D_{1,i}\omega}} + \frac{e^{-jD_{2,i}\frac{\omega}{c}}}{\sqrt{D_{2,i}\omega}} \right). \qquad (29)$$

It can be shown [23] that this sum converges and that

$$P(n_0\phi_0, \omega) \sim \frac{C_4(\omega)}{\sqrt{\omega}} \qquad (30)$$

with $C_4(\omega)$ a bounded function of ω.

IV. SAMPLING AND RECONSTRUCTION

In the previous sections, the decay of the 2-D spectrum of the PAF has been studied both along the temporal and the spatial frequency axes. One has observed that the spectrum of the PAF lies on a support that is almost bandlimited. This result is valid for a single source, but also for a finite number of sources. In the scope of this paper, we are mostly interested in the sampling of the sound pressure field along the spatial axis. We consider that the sound pressure field positions are recorded with omnidirectional point microphones.[5] Since no spatial antialiasing filter can be applied in the spatial direction, the speed of the spatial frequency decay of the 2-D spectrum of the PAF is the key factor for the quality of the reconstruction. Along the temporal direction, the signal can be filtered in order to avoid aliasing.

In this section, the sampling theorem of the PAF is presented. The quality of the reconstruction is studied when sampling the sound pressure field in space. Further, interpolation techniques are discussed in order to reconstruct the signal from the available samples.

[5]If the membrane of the microphone has a finite dimension, a spatial filtering is applied on the recorded sound.
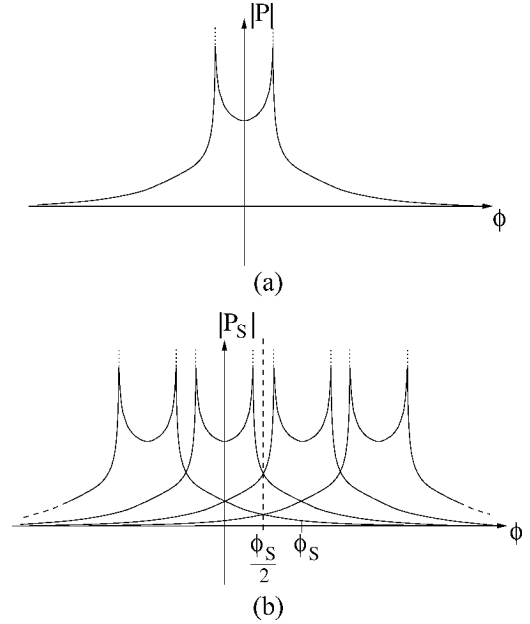
(a)

(b)

Fig. 4. Magnitude of the spectrum of the PAF. (a) Cut of the spectrum of the PAF for a particular temporal frequency. (b) Spectrum of the sampled PAF (denoted as P_S) for a particular temporal frequency consisting of spectral repetitions of the original PAF spectrum.

A. Sampling Theorem for the PAF

In order to uniformly sample the PAF along the spatial direction, a uniformly spaced infinite number of impulse responses are considered. Call ϕ_S the spatial sampling frequency defined as $2\pi/\Delta x$ where Δx is the sampling interval between two consecutive positions of the measured impulse responses. Next to the spatial sampling, the RIRs also need to be sampled at a certain temporal sampling rate depending on the desired audio bandwidth. Call ω_S the temporal sampling frequency, or $\omega_S = 2\pi/\Delta t$ with Δt the sampling period of the impulse responses.

Consider now the spectrum of the PAF given by (8) at a particular temporal frequency ω_0. It has approximately the shape given in Fig. 4(a). When the PAF is sampled, repetitions of the spectrum occur as shown in Fig. 4(b). As the spectrum is not perfectly bandlimited, the repetitions will affect the reconstruction. A theorem quantifying the SNR of the reconstruction of the PAF is presented for one source emitting in free field. Call $\text{SNR}(\phi_S, \omega_0)$ the SNR of the reconstruction for a sinusoid emitted at frequency $\omega = \omega_0$ with the microphones positioned with a spatial sampling frequency ϕ_S. In the present case, the SNR is defined as follows:

$$\text{SNR}(\phi_S, \omega_0) = \frac{\int_{\phi=-\infty}^{\infty} |P(\phi, \omega_0)|^2 \, d\phi}{4 \int_{\phi=\frac{\phi_S}{2}}^{\infty} |P(\phi, \omega_0)|^2 \, d\phi}. \qquad (31)$$

The numerator in (31) corresponds to the energy of the spectrum of the PAF at temporal frequency ω_0. The denominator in (31) corresponds to the energy contained in the spectral repetitions that will contaminate the reconstruction in the spectral domain of interest. Two different kinds of energy are present in this denominator: the "in band" and the "out of band" energy. The "in band" energy corresponds to the energy of all the spectral repetitions in the domain of interest, namely $[-(\phi_S/2), (\phi_S/2)]$.

The "out of band energy" is the energy present in the spectrum that is outside of the domain of interest. It can be shown that the "in band" and the "out of band" energies are equal in the case of an infinite line of microphones. This explains the factor 4 in the denominator of (31).

Theorem 1: Assume one single source emitting in free field at a frequency $\omega = \omega_0$. When sampling the PAF at a spatial sampling frequency of ϕ_S, for a particular $\omega = \omega_0$, and reconstructing it using an ideal interpolator, the SNR of the reconstructed signal in the band $[-(\phi_S/2), (\phi_S/2)]$ can be expressed as

$$\text{SNR}(\phi_S, \omega_0) = \frac{1}{2d \int_{\phi=\frac{\phi_S}{2}}^{\infty} \left| H_0 \left(d\sqrt{\left(\frac{\omega_0}{c}\right)^2 - \phi^2} \right) \right|^2 d\phi} \tag{32}$$

where H_0 is the zero-order Hankel function of the first kind.

When considering $\phi_S > (2\omega_0/c)$ and using the exponential integral function defined as $E_i(x) = \int_x^{\infty}(e^{-t}/t)dt$, the SNR can be lowerbounded, as follows:

$$\text{SNR}(\phi_S, \omega_0) \geq \frac{\pi}{4E_i \left(2d\sqrt{\left(\frac{\phi_S}{2}\right)^2 - \left(\frac{\omega_0}{c}\right)^2} \right)}. \tag{33}$$

Proof: The numerator in (31) can be rewritten using the Parseval's relation as follows:

$$2\pi \int_{x=-\infty}^{\infty} |\tilde{p}(x, \omega_0)|^2 \, dx \tag{34}$$

with $\tilde{p}(x, \omega_0)$ the inverse Fourier transform of $P(\phi, \omega_0)$ along the spatial axis. We have that $\tilde{p}(x, \omega_0) = (e^{-j}(\omega_0\sqrt{x^2+d^2}/c)/4\pi\sqrt{x^2+d^2})$, and therefore the numerator in (31) is

$$\frac{2\pi}{16\pi^2} \int_{x=-\infty}^{\infty} \frac{1}{x^2 + d^2} dx = \frac{1}{8d}. \tag{35}$$

Using (8) and (35) in (31) leads to (32).

When considering $\phi_S > (2\omega_0/c)$, the denominator in (31) is

$$\frac{1}{\pi^2} \int_{\phi=\frac{\phi_S}{2}}^{\infty} \left| K_0 \left(d\sqrt{\phi^2 - \left(\frac{\omega_0}{c}\right)^2} \right) \right|^2 d\phi. \tag{36}$$

Using (12), (36) can be upperbounded by

$$\frac{1}{2\pi} \int_{\phi=\frac{\phi_S}{2}}^{\infty} \frac{e^{-2d\sqrt{\phi^2 - \left(\frac{\omega_0}{c}\right)^2}}}{d\sqrt{\phi^2 - \left(\frac{\omega_0}{c}\right)^2}} d\phi. \tag{37}$$

Using the change of variable $z = \sqrt{\phi^2 - (\omega_0/c)^2}$, (37) can be rewritten as

$$\frac{1}{2\pi d} \int_{z=\sqrt{\left(\frac{\phi_S}{2}\right)^2 - \left(\frac{\omega_0}{c}\right)^2}}^{\infty} \frac{e^{-2dz}}{\sqrt{z^2 + \left(\frac{\omega_0}{c}\right)^2}} dz. \tag{38}$$
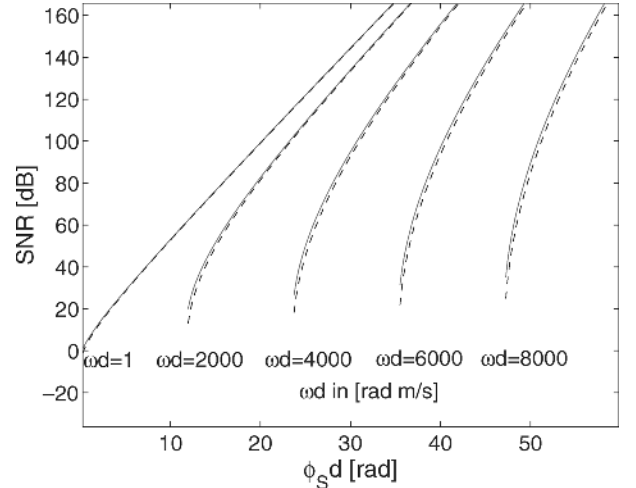

Fig. 5. In full lines, SNR on the reconstruction signal for different $\phi_S d$ for normalized temporal frequencies ranging from 1 rad m/s to 8000 rad m/s. In dotted line, the corresponding SNR lower bounds are given.

Then, (38) can be upperbounded by

$$\frac{1}{2\pi d} \int_{\sqrt{\left(\frac{\phi_S}{2}\right)^2 - \left(\frac{\omega_0}{c}\right)^2}}^{\infty} \frac{e^{-2dz}}{z} dz$$

$$= \frac{1}{2\pi d} E_i \left(2d\sqrt{\left(\frac{\phi_S}{2}\right)^2 - \left(\frac{\omega_0}{c}\right)^2} \right) \tag{39}$$

where E_i represents the exponential integral function. Using (39) and (35), (33) is obtained. ∎

As a numerical check of the tightness of the bound, the SNR has been computed as a function of the spatial sampling frequency and the temporal frequency. To take into account the distance d, it was observed that (32) and (33) can be expressed as a function of normalized frequencies $\phi_S d$ and ωd. This allows to obtain a numerical evaluation for these two equations independently of d. The results are shown in Fig. 5 in full lines for (32) and in dotted lines for (33). In order to avoid numerical instability due to the infinite value of the spectrum at the position $\phi_S = (2\omega/c)$, the simulations start for each ωd at a value of $\phi_S d$ larger than $\phi_S d = 2\omega d/c$. At one normalized temporal frequency, it can be observed that the SNR increases for larger normalized spatial sampling frequencies. Note that the lowerbound follows tightly the SNR obtained numerically.

As an example, it can be observed from Fig. 5 that to reconstruct the soundfield up to 8000 rad/s (or 1.3 kHz), a spatial distance of 12.35 cm is necessary to achieve a reconstruction quality of 100 dB when considering a unitary distance between the source and the line of microphones. Note that when making the far-field assumption, using (4) would lead to a spacing of 13.35 cm.

Instead of considering the SNR at a specific temporal frequency, the previous result for the SNR on the reconstruction can be generalized when the signal is in a frequency band
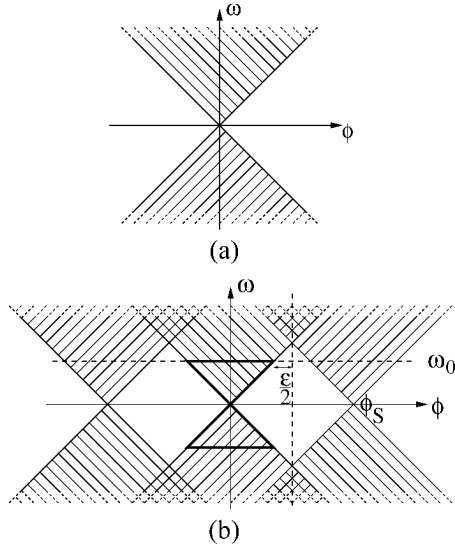
Fig. 6. PAF spectrum. (a) Top view of the PAF spectrum. (b) Top view of the PAF spectrum with its repetitions due to the spatial sampling.

$[\omega_1, \omega_2]$ with power spectral density $|\beta(\omega)|^2$. The SNR on the reconstruction is then given by the following formula:

$$
\text{SNR}\,(\phi_S, [\omega_1, \omega_2])
$$

$$
= \frac{\int_{\omega_1}^{\omega_2} |\beta(\omega)|^2 \, d\omega}{2d \int_{\omega_1}^{\omega_2} |\beta(\omega)|^2 \left(\int_{\frac{\phi_S}{2}}^{\infty} \left| H_0 \left(d\sqrt{\left(\frac{\omega}{c}\right)^2 - \phi^2} \right) \right|^2 d\phi \right) d\omega}. \quad (40)
$$

When considering $\phi_S > (2 \max(\omega_1, \omega_2)/c)$, (40) can be lower-bounded as follows:

$$
\text{SNR}\,(\phi_S, [\omega_1, \omega_2])
$$

$$
\geq \frac{\pi \int_{\omega_1}^{\omega_2} |\beta(\omega)|^2 \, d\omega}{4 \int_{\omega_1}^{\omega_2} |\beta(\omega)|^2 \, E_i \left(2d\sqrt{\left(\frac{\phi_S}{2}\right)^2 - \left(\frac{\omega}{c}\right)^2} \right) d\omega}. \quad (41)
$$

Generalization of the sampling theorem for the cases of multiple sources in free field or inside a room is matter of current research.

B. Plenacoustic Sampling

In this section, the effect of the sampling in time and space of the PAF is represented in the spectral domain. This representation takes into account the results of the previous section. Sampling in space is done considering a margin taking into account the evanescent decay of the PAF spectrum. The sampling of the PAF will first be shown in details for the rectangular sampling pattern. Further, it will be shown that quincunx sampling can also be used to sample the PAF.

1) Rectangular Sampling: The schematic top view of the spectrum of the PAF is shown in Fig. 6(a). When sampling the PAF along the spatial dimension with a spatial sampling frequency of ϕ_S, repetitions of the spectrum occur as shown in Fig. 6(b). When one considers the far field approximation of the
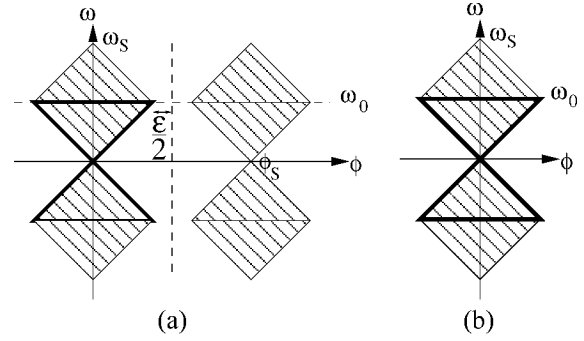

Fig. 7. PAF spectrum. (a) PAF spectrum with its spectral repetitions along the temporal and the spatial frequencies. (b) PAF spectrum with its spectral repetitions along the temporal frequencies. In both figures, the region in bold corresponds to the original spectrum bandlimited along the temporal frequency without spectral repetitions.

sound pressure field, the spatial sampling frequency needed to reconstruct the PAF up to a temporal frequency ω_0 is given by

$$
\phi_S > \frac{2\omega_0}{c}. \quad (42)
$$

When one does not consider any far field approximation, Fig. 5 has to be considered. In that figure, the spatial sampling frequency is given for each possible temporal frequency for a wanted reconstruction SNR. When one wants to ensure that the highest frequency of the signal is still reconstructed with a given SNR, the spatial sampling frequency needs to be slightly increased compared to (42). To reconstruct the sound pressure field up to ω_0 with a given SNR_0 at that frequency, the spatial sampling frequency is given by

$$
\phi_S > \frac{2\omega_0}{c} + \varepsilon(\text{SNR}_0, \omega_0) \quad (43)
$$

where $\varepsilon(\text{SNR}, \omega)$ is obtained from Fig. 5 when considering a specific temporal frequency and a given SNR. It represents the difference between the value read on the graph and the far field approximation (42). Inside the region in bold in Fig. 6(b), it can be observed that for a particular ϕ_S chosen according to (43):

$$
\forall |\omega| \leq \omega_0 \quad \text{SNR}(\phi_S, \omega) \geq \text{SNR}_0. \quad (44)
$$

When $|\omega| \geq \omega_0$, due to the periodic repetitions along the spatial axis, the $\text{SNR}(\phi_S, \omega)$ decreases.

The temporal sampling of the PAF is now considered. The signal is first bandlimited to ω_0 satisfying (43) and sampled with a temporal sampling frequency of $\omega_S = 2\omega_0$. Repetitions of the spectra occur now also along the temporal frequency. The obtained spectrum for the PAF sampled in space and time is shown in Fig. 7(a).

Conversely, considering temporal sampling first, one can say that if the maximal temporal frequency present in the signal is ω_0, then sampling the signal at a temporal sampling frequency $\omega_S = 2\omega_0$, one obtains the signal whose spectrum is shown in Fig. 7(b). When sampling this signal along the spatial dimension, it is necessary to choose ϕ_S as in (43) in order to avoid aliasing as shown in Fig. 7(a). Using the rectangular sampling, the sampling of the PAF happens in space and time domain as shown in Fig. 8(a). The corresponding spectrum is shown in
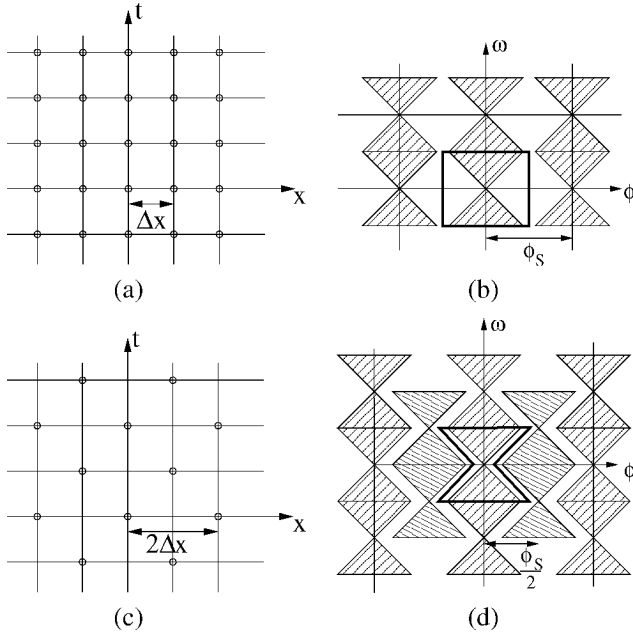
Fig. 8. Sampling and interpolation of the PAF. (a) Rectangular sampling grid. (b) Plenacoustic spectrum with its repetitions for a rectangular sampling grid with the interpolation filter in bold. (c) Quincunx sampling grid. (d) Plenacoustic spectrum with its repetitions for a quincunx sampling grid with the interpolation filter in bold.

Fig. 8(b). The final expression for our sampled PAF 2-D spectrum (denoted as P_S) becomes

$$P_S(\phi, \omega) = \frac{1}{\Delta x \Delta t} \sum_{k_1, k_2 = -\infty}^{\infty} P\left(\phi - \frac{2\pi k_1}{\Delta x}, \omega - \frac{2\pi k_2}{\Delta t}\right). \tag{45}$$

2) Quincunx Sampling: A tighter packing of the spectrum can be achieved by using quincunx sampling. In time domain, the grid to be used is shown in Fig. 8(c). In the corresponding spectrum, the spectral repetitions are placed such that they fill better the whole frequency space as shown in Fig. 8(d). In the quincunx sampling the spatial sampling frequency is now only $\phi_S/2$. This corresponds to a distance between two samples on the space axis of $2\Delta x = 4\pi/\phi_S$. This shows that using quincunx sampling one only needs to sample the even microphones at even times while the odd microphones are sampled at odd times. This leads to a gain of factor 2 in the processing. However it does not reduce the number of necessary microphones. Similar approaches have been used in [13] in the study of the far field electromagnetic field. The precise value of ϕ_S to be used in the case of quincunx sampling to ensure a SNR of reconstruction at a certain temporal frequency is not given in this paper but could be derived similarly to the expression (32).

C. Reconstruction by Interpolation

Knowing the sound field at every point of the sampling grid, the usual interpolation techniques [24], [25] are applied in order to reconstruct the sound pressure field at any location. The interpolation filter to be used is dependent on the sampling grid,
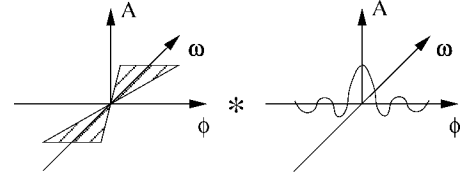

Fig. 9. Effect of the windowing due to the finite length of the array.

and may be separable in time and space. When the samples have been obtained by rectangular sampling, the interpolation filter is a lowpass filter with support $[-(\phi_S/2), (\phi_S/2)]$. The support of the filter is shown in bold in Fig. 8(b). When the samples have been obtained by quincunx sampling, the filter used for interpolation is a "fat fan filter" [25]. The support of the filter is shown in bold in Fig. 8(d). One can compute the ideal filters in closed forms, however these ideal interpolation filters are not realizable in practice. Therefore, one needs to slightly increase the spatial sampling frequency to take into account the decay of the real filter. For their design, standard techniques can be used [25].

V. FINITE-LENGTH APERTURE

In this section, the field is not measured along an infinite line but on a finite interval inside the room. This can be seen as a windowing of the PAF in the spatial domain. Consider a rectangular window $w(x)$. In the present case, the window is simply a function of the spatial position. Calling the windowed PAF $q(x, t)$. It satisfies $q(x, t) = p(x, t)w(x)$. In frequency domain, this is written as

$$Q(\phi, \omega) = P(\phi, \omega) * W(\phi, \omega) = P(\phi, \omega) * (W(\phi)\delta(\omega)). \tag{46}$$

The situation is schematically shown in Fig. 9.

The size of the aperture has an influence on the decay of the windowed PAF. Taking measurements from positions between $-L/2$ and $L/2$ leads to a convolution of the PAF spectrum with the following sinc function:

$$W(\phi) = \int_{x=-\frac{L}{2}}^{\frac{L}{2}} e^{-j\phi x} \mathrm{d}x = L \sin c\left(\frac{\phi L}{2\pi}\right). \tag{47}$$

At a given ϕ, the larger the value of L, the faster the decay will be. This fact can be observed in Fig. 10. A section of the 2-D spectrum of the PAF at a particular temporal frequency $(2000\pi \text{ rad/s})$ is presented. One can observe that for larger aperture sizes the spectrum decays more rapidly as given in (47).

The decay of the spectrum of the PAF along the spatial frequency will now be slower than the one in (10). For a particular temporal frequency $\omega = \omega_0$, the decay is

$$Q(\phi, \omega_0) = \frac{e^{-j\phi x_s}}{2\pi} K_0\left(d\sqrt{\phi^2 - \left(\frac{\omega_0}{c}\right)^2}\right) * W(\phi). \tag{48}$$

In the case of a rectangular window, this decay is the convolution of a sinc with a modified Bessel function K_0.
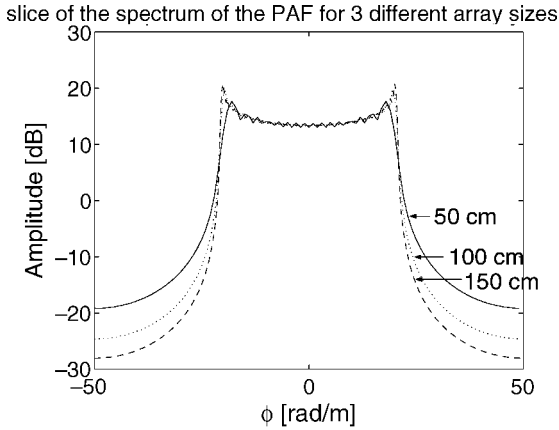
Fig. 10. A section of the PAF at a particular temporal frequency (2000π rad/s). The curves represent data acquired on intervals of different lengths: 50 cm (full line), 100 cm (dotted line), and 150 cm (dashed line). A larger interval leads to a faster decay. For this graph, we used $d = 1$ in (8).

Fig. 11. Interpolation error for different array sizes. We use the same spacing between the microphones, namely 2 cm.

Combining the finite aperture effect with the sampling of the PAF, the following expression for the 2D-FT of the sampled windowed PAF (denoted as Q_S) is obtained:

$$Q_S(\phi, \omega) = \frac{1}{\Delta x \Delta t} \sum_{k_1, k_2 = -\infty}^{\infty} Q\left(\phi - \frac{2\pi k_1}{\Delta x}, \omega - \frac{2\pi k_2}{\Delta t}\right). \quad (49)$$

VI. SIMULATIONS AND MEASUREMENTS

In this section, simulation results are presented for the interpolation of RIRs. These results are then compared with real measurements.

A. Simulation Results

RIRs have been simulated on a line in a room using the image source model. For simulation purposes, one derives a dense set of impulse responses, keeps a subset, and interpolates the missing ones. In the simulations, the case of rectangular sampling of the sound field was considered. To compare the interpolated with the simulated RIRs, the normalized mean-square-error (MSE) criterion was used:

$$\text{MSE} = \frac{\sum_{i=1}^{N} (r[i] - \hat{r}[i])^2}{\sum_{i=1}^{N} r^2[i]} \quad (50)$$

with N the length in samples of the simulated RIRs, r the simulated RIR and $\hat{r}$ the interpolated RIR.

In the presented simulations, $2n + 1$ RIRs were simulated every cm along a line in the room. From these simulations, we kept $n + 1$ measurements spaced with 2 cm to interpolate the n "in-between" positions. With the spacing of 2 cm, the RIRs were reconstructed up to a temporal frequency of $16\,000\pi$ rad/s. In this setup, due to the limited length of the array, the decay of the spectrum is mostly determined by the decay of the windowing function as discussed in Section V. We chose ε large enough in (43) to ensure a sufficient decay in the spectrum of the finite length aperture PAF. Note that this ε can be substantially larger than the one obtained in Fig. 5. The MSE for those n positions is
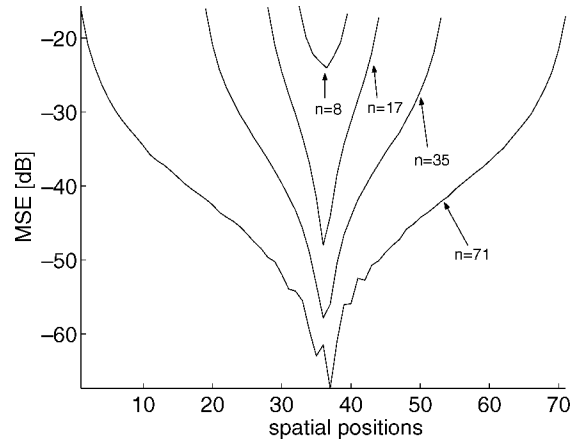
shown in Fig. 11. Different lengths of the array centered around the same spatial position were considered, namely $n = 8$, 17, 35 and 71. Note that the MSE is only given for the interpolated positions since the positions used to apply the interpolation are perfectly reconstructed. One can observe that using the array corresponding to $n = 71$ leads to a very small error (less than -60 dB) for the interpolation in the middle of the array. When using the same spacing between the microphones but reducing the number of RIRs, the interpolation error increases due to the border effects introduced by the finite length of the array. Note also that the curves on Fig. 11 are not symmetric. This is due to the fact that the microphone array was not symmetrically located inside the room.

B. Experimental Results

Experimental results were carried out in a partially sound insulated room with RIRs measured at different spatial positions. One loudspeaker (Genelec 1029A) was used together with a microphone array (composed of eight Panasonic *WM*61A). A frequency logarithmic sweep [26] was used to measure the RIRs. 71 RIRs were measured with a microphone spacing of 2 cm along a line in the room.[6] The spectrum of the measured PAF is shown in Fig. 12(a). Similarly to the technique presented in the simulations, we kept 36 measurements spaced with 4 cm to interpolate the 35 "in-between" positions. The RIRs were low-pass filtered to 8000π rad/s. Considering (43), ε was chosen large enough to ensure a sufficient decay in the spectrum of the windowed PAF. The MSE on the 35 interpolated RIRs measurements is shown in Fig. 12(b). Note that the MSE shown in Fig. 12(b) is obtained when using only the first 100 ms of the RIRs. When considering the full RIRs (1 s) poorer results are obtained (on the order of -25 dB). This results leads to the thinking that only the beginning of the RIRs is well interpolated. To justify this conjecture, successive measurements were performed at the same spatial positions with a 30-s interval. We studied the MSE between two successive measurements using a sliding window of 25 ms. The results were averaged over 100

---

[6]All the measurements are available online at http://lcavwww.epfl.ch/~thibaut/Measurements/Acoustic.html
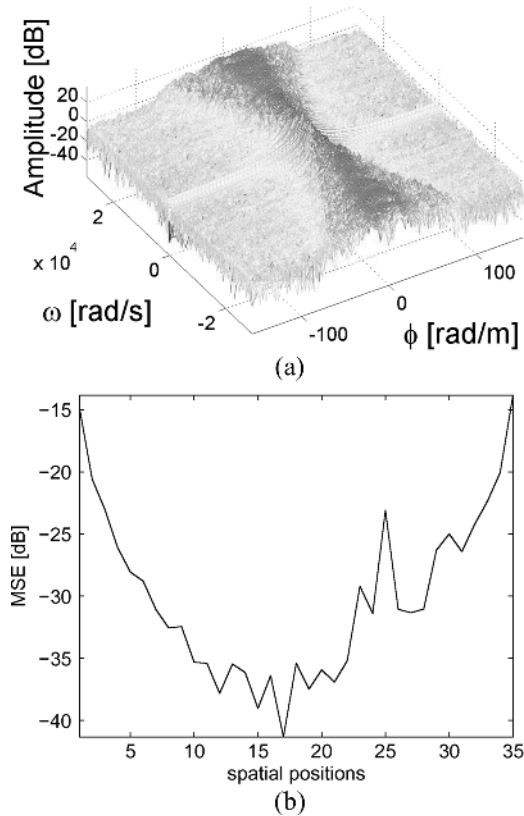
Fig. 12.   Experimental PAF. (a) 2D-FT of the measured plenacoustic function. (b) Interpolation error on measured RIRs.
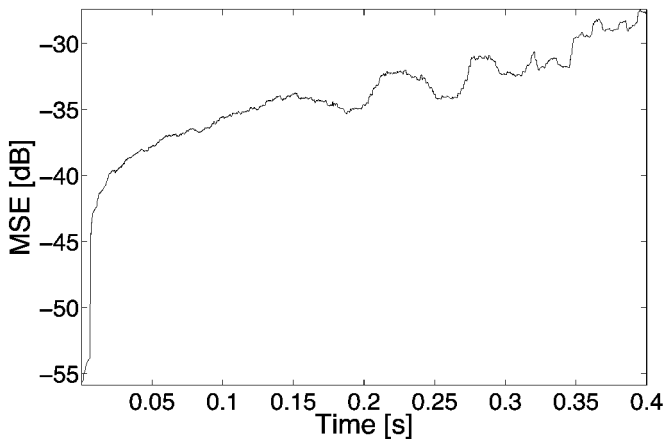

Fig. 13.   MSE on successive RIRs along a sliding window of 25 ms. The results were averaged over 100 pairs of successive measurements.

measurements. It can be observed in Fig. 13 that the MSE is very low at the beginning of the RIR and increases with time. After 100 ms, the MSE between consecutive measurements is already of about -38 dB. This fact can probably be understood by a variation of the speed of sound propagation over time. This effect will be the most severe for waves traveling over a wider area. Therefore, the reverberant part of the RIR will undergo the largest relative timing changes. The speed of sound fluctuation will also have a larger impact for larger microphone to loudspeaker distances. This speed of sound variation can be the consequence of variation of different parameters such as tem-

perature, humidity, and pressure in the room. In [27], similar observations have been done and attributed to variation of temperature. They showed that a variation of $0.1°$ can create a misalignment between RIRs of more that 25 dB.

Note finally that due to the limit of our 8 inputs microphone array, we had to move the array to the next positions in order to measure the 71 RIRs (our intrusion in the room probably modified the temperature and other factors between two sets of measurements). Better results would be obtained if all the measurements could be captured simultaneously, which was not possible due to hardware limitations.

VII. PLENACOUSTIC FUNCTION FOR MULTIDIMENSIONAL SPATIAL POSITIONS

The previous sections were devoted to the detailed study of the sampling and interpolation of the sound pressure field on the line. This study can obviously be generalized to other spatial positions of microphones (or loudspeakers). Therefore, one wants to study the shape of spectra associated with different microphones setups. Section VII-A studies the spectrum of the PAF associated to a plane of microphones. The optimal sampling pattern for positioning the microphones is studied. Further, the 3-D space filled of microphones is presented in Section VII-B.

Due to lack of space, these different setups are analyzed in less details than the line of microphones. Temporal and spatial decay analysis in rooms as well as spatial windowing are not presented but can be generalized from the study of the line of microphones.

A. Plenacoustic Function on a Plane

1) Study of the Spectrum: In Section III, the shape and the properties of the PAF on a line in the room were studied. In this section, one considers a more general case where the RIRs are studied on a plane. Consider a plane in the space filled with receivers in the x and the y directions. Further, a source is located at position (x_s, y_s, z_s). At any receiver position (x_m, y_m, z_m) the direct path coming from the source is

$$p(x_m, y_m, t) = \frac{\delta\left(t - \frac{a}{c}\right)}{4\pi a} \qquad (51)$$

with $a = \sqrt{(x_m - x_s)^2 + (y_m - y_s)^2 + (z_m - z_s)^2}$. The derivation of the 3D-FT of (51) is done in Appendix II and follows the same approach as the one presented in [14]. The result is

$$P(\phi_x, \phi_y, \omega) = \begin{cases} \dfrac{-j\Phi}{2} \dfrac{e^{-j|z|\sqrt{\left(\frac{\omega}{c}\right)^2 - \phi_q^2}}}{\sqrt{\left(\frac{\omega}{c}\right)^2 - \phi_q^2}} & \text{for } |\phi_q| \leq \frac{\omega}{c} \\ \dfrac{\Phi}{2} \dfrac{e^{-|z|\sqrt{\phi_q^2 - \left(\frac{\omega}{c}\right)^2}}}{\sqrt{\phi_q^2 - \left(\frac{\omega}{c}\right)^2}} & \text{for } \frac{\omega}{c} \leq |\phi_q| \end{cases}. \qquad (52)$$

with $\phi_q^2 = \phi_x^2 + \phi_y^2$ and $\Phi = e^{-j(\phi_x x_s + \phi_y y_s)}$. Note that ϕ_x and ϕ_y represent the spatial frequencies for the microphones in the x and y directions respectively. The obtained spectrum has a conical shape as shown in Fig. 14. The surface of the cone is given by

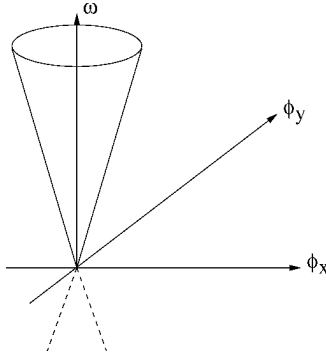$$\frac{\omega}{c} = \sqrt{\phi_x^2 + \phi_y^2}. \qquad (53)$$

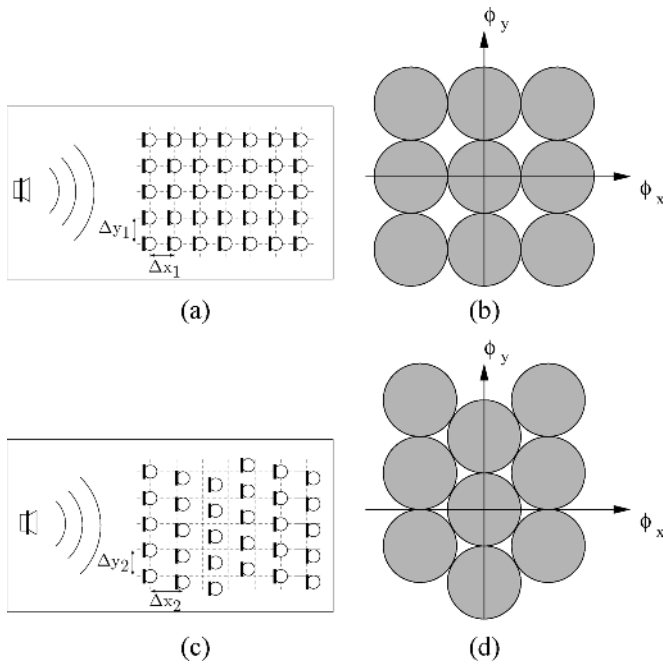Fig. 14. Schematic view of the 3-D spectrum of the PAF.

Fig. 15. Sampling of the PAF on a plane. (a) Placement of the microphones on the plane on a rectangular sampling grid. (b) Plenacoustic spectrum with its repetitions for a rectangular sampling grid. (c) Hexagonal sampling grid. (d) Plenacoustic spectrum with its repetitions for a hexagonal sampling grid.

Similarly to the results obtained with the line of microphones, one sees that the decay of the spectrum is also faster than an exponential outside of the conical shape. Remark that in the specific case of the source located on the plane of the microphones, the decay becomes slower and is, up to a constant, asymptotic as $1/\phi_q$.

2) Optimal Sampling Pattern: Similarly to the analysis presented in Section IV-C, the optimal sampling pattern for the positioning of the microphones on the plane is studied. The first approach is to use the rectangular sampling as shown in Fig. 15(a). A spacing of Δx_1 and Δy_1 was used for the spacing between the microphones in the x and y directions. Fig. 15(b) shows the corresponding packing of the circles in the Fourier spectrum for one temporal frequency (typically the highest frequency present in the emitted signal).

The conical shape of the spectrum allows us to obtain a tighter packing of the circles. The use of an hexagonal sampling pattern leads to a reduction of about 15% in the number of necessary microphones. Fig. 15(c) shows the new positions of the microphones on the plane. In this case

$$\Delta x_2 = \frac{2}{\sqrt{3}}\Delta x_1, \ \Delta y_2 = \Delta y_1. \tag{54}$$

Fig. 15(d) shows the corresponding spectrum with its spectral repetitions. Other packings of the cones can be realized to lower the temporal sampling frequency of the analog–digital (A/D) converters but do not reduce further the number of microphones needed to sample the sound pressure field on a plane [12].

3) Simulation Results: We simulated RIRs on a plane in a room using the image source model. The 3-D Fourier transform was applied on the gathered data. By looking at sections of this spectrum for $\omega = 1500$ rad/s and $\omega = 3000$ rad/s, one obtains Fig. 16(a) and (b) respectively. One can see that with growing temporal frequencies, the support of the PAF spectrum also increases as given by (53).

B. Plenacoustic Function in Space

In this section microphones located in the 3-D space are considered. Similarly to the setup presented in Section VII-A, a source is located at position (x_s, y_s, z_s) and microphones at positions (x_m, y_m, z_m). The PAF is also given by (51). Note that in the present setup also z_m is a variable. Call ϕ_z the spatial frequency of the microphone positions in the z direction. Calculating the fourth-dimensional Fourier transform (4D-FT) of (51) is done in Appendix III. The result is

$$P(\phi_x, \phi_y, \phi_z, \omega) = \frac{e^{-j(\phi_x x_s + \phi_y y_s + \phi_z z_s)}}{\phi_x^2 + \phi_y^2 + \phi_z^2 - \left(\frac{\omega}{c}\right)^2} \tag{55}$$

which represents a cone in four dimensions. For a particular temporal frequency, the section of this cone is a sphere. The size of the sphere obeys the following rule:

$$\frac{\omega}{c} = \sqrt{\phi_x^2 + \phi_y^2 + \phi_z^2}. \tag{56}$$

Note that the decay outside of the sphere is not exponential as it was for the plane and line of microphones. This is due to the presence of the source at one of the microphone positions. At a particular temporal frequency, the optimal packing of the spheres is given by face-centered cubic lattice packing [28]. It allows to reduce the number of samples by a factor of about 29.3% when compared to a normal rectangular sampling pattern.

VIII. CONCLUSION

In this paper, we have introduced and studied the plenacoustic function. It characterizes the sound pressure field at any point in space. This function has been studied and its spectrum for the
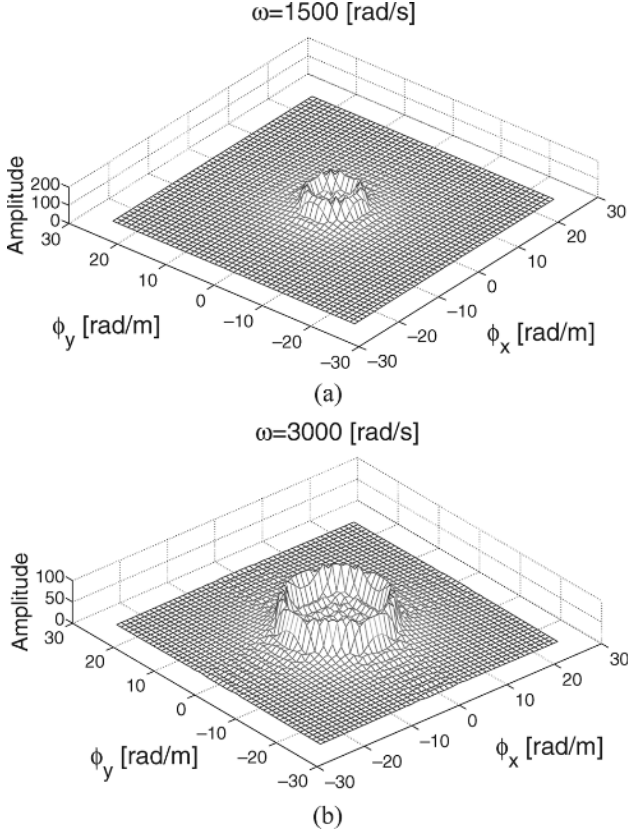
Fig. 16. Spectrum of the PAF obtained by simulations at different temporal frequencies. (a) $\omega = 1500$ rad/s. (b) $\omega = 3000$ rad/s.

linear, planar, and three-dimensional case has been calculated without making any far field assumption. The decay of the spectrum has been studied along both the temporal and spatial frequency axis. Based on the support of the spectrum, the number and the spacing between the microphones needed to reconstruct the sound pressure field up to a certain temporal frequency has been determined. A quantitative sampling theorem, trading off sampling rate for SNR, has been presented. As an example, it has been shown that to reconstruct the soundfield up to 1.3 kHz, a spatial distance of 12.35 cm is necessary to achieve a reconstruction quality of 100 dB for a source located at unitary distance from the line of microphones. Using the far-field assumption, this spacing would be of 13.35 cm. The optimal sampling pattern for the microphone positions has also been given for the linear, planar, and three-dimensional case. Finally, simulations and experimental results were presented and compared with the theoretical results.

APPENDIX I
DERIVATION OF THE 2D-FT OF THE PAF ON A LINE

$$p(x,t) = \frac{\delta\left(t - \frac{\sqrt{(x-x_s)^2 + d^2}}{c}\right)}{4\pi\sqrt{(x-x_s)^2 + d^2}}. \qquad (57)$$

The 2D-FT of (57) is

$$P(\phi, \omega) = \int_{x=-\infty}^{\infty} \int_{t=-\infty}^{\infty} p(x,t)e^{-j(\phi x + \omega t)}dt dx. \qquad (58)$$

Call $u = x - x_s$; therefore

$$P(\phi, \omega) = \frac{e^{-j\phi x_s}}{4\pi} \int_{u=-\infty}^{\infty} \frac{e^{-j\left(\phi u + \frac{\omega}{c}\sqrt{u^2+d^2}\right)}}{\sqrt{u^2+d^2}}du \qquad (59)$$

$$= \frac{e^{-j\phi x_s}}{2\pi} \int_{u=0}^{\infty} \frac{e^{-j\frac{\omega}{c}\sqrt{u^2+d^2}}\cos(\phi u)}{\sqrt{u^2+d^2}}du. \qquad (60)$$

Using existing formula in [29], one obtains

$$P(\phi, \omega) = -\frac{j}{4}e^{-j\phi x_s}H_0^*\left(d\sqrt{\left(\frac{\omega}{c}\right)^2 - \phi^2}\right). \qquad (61)$$

APPENDIX II
DERIVATION OF THE 3D-FT OF THE PAF ON A PLANE

$$p(x_m, y_m, t) = \frac{\delta\left(t - \frac{a}{c}\right)}{4\pi a}$$

with $a = \sqrt{(x_m - x_s)^2 + (y_m - y_s)^2 + (z_m - z_s)^2}$. Call $\mathbf{u}_m = [x_m, y_m]^T$ and $\boldsymbol{\phi} = [\phi_x, \phi_y]^T$. One can calculate the spectrum of this function

$$P(\boldsymbol{\phi}, \omega) = \int_{\mathbb{R}^2} \int_{\mathbb{R}} \frac{\delta\left(t - \frac{a}{c}\right)}{4\pi a}e^{-j(\boldsymbol{\phi}^T \mathbf{u}_m + \omega t)}d\mathbf{u}_m dt. \qquad (62)$$

This expression can be rewritten as

$$P(\boldsymbol{\phi}, \omega) = \int_{\mathbb{R}^2} \frac{e^{-j\left(\boldsymbol{\phi}^T \mathbf{u}_m + \omega \frac{a}{c}\right)}}{4\pi a}d\mathbf{u}_m. \qquad (63)$$

Introducing $\mathbf{u} = [x_m - x_s, y_m - y_s]^T$, the integral is rewritten as

$$P(\boldsymbol{\phi}, \omega) = e^{-j(\phi_x x_s + \phi_y y_s)} \int_{\mathbb{R}^2} \frac{e^{-j\left(\boldsymbol{\phi}^T \mathbf{u} + \omega \frac{a}{c}\right)}}{4\pi a}d\mathbf{u}. \qquad (64)$$

Call $r^2 = x^2 + y^2$, $\Phi = e^{-j(\phi_x x_s + \phi_y y_s)}$, $z = z_m - z_s$ and $\phi_q^2 = \phi_x^2 + \phi_y^2$. The integral can be rewritten as [30]

$$P(\boldsymbol{\phi}, \omega) = \frac{\Phi}{2} \int_{r=0}^{+\infty} \frac{rJ_0(\phi_q r)}{\sqrt{r^2+z^2}}e^{-j\left(\omega\frac{\sqrt{r^2+z^2}}{c}\right)}dr. \qquad (65)$$

Call $m = \sqrt{r^2 + z^2}$, we have that $r = \sqrt{m^2 - z^2}$, and also $dm = (r/\sqrt{r^2 + z^2})dr$. The integral becomes

$$P(\boldsymbol{\phi}, \omega) = \frac{\Phi}{2} \int\limits_{m=|z|}^{+\infty} J_0(\phi_q \sqrt{m^2 - z^2})e^{-j(\omega\frac{m}{c})}dm. \quad (66)$$

Using existing formulas in [29], one obtains that

$$P(\phi_x, \phi_y, \omega) = \begin{cases} \dfrac{-j\Phi}{2} \dfrac{e^{-j|z|\sqrt{\left(\frac{\omega}{c}\right)^2 - \phi_q^2}}}{\sqrt{\left(\frac{\omega}{c}\right)^2 - \phi_q^2}} & \text{for } |\phi_q| \leq \frac{\omega}{c} \\ \dfrac{\Phi}{2} \dfrac{e^{-|z|\sqrt{\phi_q^2 - \left(\frac{\omega}{c}\right)^2}}}{\sqrt{\phi_q^2 - \left(\frac{\omega}{c}\right)^2}} & \text{for } \frac{\omega}{c} \leq |\phi_q| \end{cases}. \quad (67)$$

APPENDIX III
DERIVATION OF THE 4D-FT OF THE PAF IN SPACE

$$p(x_m, y_m, z_m, t) = \frac{\delta\left(t - \frac{a}{c}\right)}{4\pi a}$$

with $a = \sqrt{(x_m - x_s)^2 + (y_m - y_s)^2 + (z_m - z_s)^2}$. Call $\mathbf{u}_m = [x_m, y_m, z_m]^{\mathrm{T}}$ and $\boldsymbol{\phi} = [\phi_x, \phi_y, \phi_z]^{\mathrm{T}}$. The spectrum of this function can be calculated

$$P(\boldsymbol{\phi}, \omega) = \int\limits_{\mathbb{R}^3} \int\limits_{\mathbb{R}} \frac{\delta\left(t - \frac{a}{c}\right)}{4\pi a} e^{-j(\boldsymbol{\phi}^{\mathrm{T}}\mathbf{u}_m + \omega t)} d\mathbf{u}_m dt. \quad (68)$$

By introducing $\mathbf{u} = [x_m - x_s, y_m - y_s, z_m - z_s]^{\mathrm{T}}$, the integral can be rewritten as

$$P(\boldsymbol{\phi}, \omega) = e^{-j(\phi_x x_s + \phi_y y_s + \phi_z z_s)} \int\limits_{\mathbb{R}^3} \frac{e^{-j\left(\boldsymbol{\phi}^{\mathrm{T}}\mathbf{u} + \omega\frac{a}{c}\right)}}{4\pi a} d\mathbf{u}. \quad (69)$$

Introducing $\phi_s^2 = \phi_x^2 + \phi_y^2 + \phi_z^2$, the integral becomes [30]

$$P(\boldsymbol{\phi}, \omega) = e^{-j(\phi_x x_s + \phi_y y_s + \phi_z z_s)} \int\limits_{m=0}^{+\infty} \sin(\phi_s m)e^{-jm\frac{\omega}{c}}dm. \quad (70)$$

The solution of this integral can be found in [29]

$$P(\phi_x, \phi_y, \phi_z, \omega) = \frac{e^{-j(\phi_x x_s + \phi_y y_s + \phi_z z_s)}}{\phi_s^2 - \left(\frac{\omega}{c}\right)^2}. \quad (71)$$

REFERENCES

[1] E. Adelson and J. Bergen, "The plenoptic function and the elements of early vision," in *Computational Models of Visual Processing*. Cambridge, MA: MIT Press, 1991, pp. 3–20.

[2] P. Morse and K. Ingard, *Theoretical Acoustics*. New York: McGraw-Hill, 1968.

[3] J. Chai, X. Tong, S. Chan, and H. Shum, "Plenoptic sampling," in *Proc. Conf. Computer Graphics*, 2000, pp. 307–318.

[4] C. Zhang and T. Chen, "Generalized plenoptic function," Carnegie Mellon Univ., Pittsburgh, PA, Tech. Rep. 01-06, Sep. 2001.

[5] D. H. Johnson and D. E. Dudgeon, *Array Signal Processing: Concepts and Techniques*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[6] M. Kubovy and D. Van Valkenburg, "Auditory and visual objects," *Cognition*, vol. 80, pp. 97–126, 2001.

[7] T. Ajdler and M. Vetterli, "The plenacoustic function and its sampling," presented at the IEEE Benelux Workshop on Model Based Processing and Audio Coding (MPCA), Leuven, Belgium, Nov. 15, 2002.

[8] ——, "The plenacoustic function, sampling and reconstruction," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, 2003, vol. 5, pp. 616–619.

[9] ——, "Acoustic based rendering by interpolation of the plenacoustic function," in *Proc. SPIE/IS&T Visual Communication and Image Processing Conf.*, 2003, pp. 1337–1346.

[10] P. S. Naidu, *Sensor Array Signal Processing*. Boca Raton, FL: CRC Press, 2001.

[11] L. J. Ziomek, *Fundamentals of Acoustic Field Theory and Space-Time Signal Processing*. Boca Raton, FL: CRC Press, 1995.

[12] J. Coleman, "Three-phase sample timing on a wideband triangular array of 4/3 the usual density reduces the Nyquist rate for far-field signals by two thirds," presented at the 38th Annu. Asilomar Conf. Signals, Systems, Computers, Pacific Grove, CA, 2004.

[13] ——, "Ping-pong sample times on a linear array halve the Nyquist rate," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, 2004, vol. 4, pp. 925–928.

[14] A. Berkhout, *Applied Seismic Wave Theory*. Amsterdam, The Netherlands: Elsevier Science, 1987.

[15] D. de Vries and M. Boone, "Wave field synthesis and analysis using array technology," presented at the IEEE Workshop Applications Signal Processing Audio Acoustics, New Paltz, NY, 1999.

[16] E. Hulsebos, D. de Vries, and E. Bourdillat, "Improved microphone array configurations for auralization of sound fields by wave field synthesis," presented at the 110th Audio Engineering Society (AES) Convention, Amsterdam, The Netherlands, Apr. 2001.

[17] A. J. Berkhout, D. de Vries, and P. Vogel, "Acoustic control by wave field synthesis," *J. Acoust. Soc. Amer.*, vol. 93, no. 5, pp. 2764–2778, May 1993.

[18] A. Berkhout, D. de Vries, and J. Sonke, "Array technology and wave field analysis in enclosures," *J. Acoust. Soc. Amer.*, vol. 102, pp. 2757–2770, 1997.

[19] J. Blauert, *Spatial Hearing*. Cambridge, MA: MIT Press, 2001.

[20] T. Ajdler, L. Sbaiz, and M. Vetterli, "The plenacoustic function on the circle with application to HRTF interpolation," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, Barcelona, Spain, May 2005.

[21] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *J. Acoust. Soc. Amer.*, vol. 65, pp. 943–950, 1979.

[22] E. G. Williams, *Fourier Acoustics*. New York: Academic, 1999.

[23] T. Ajdler, L. Sbaiz, and M. Vetterli, "The plenacoustic function and its sampling," Ecole Polytechnique Federale de Lausanne, Lausanne, Switzerland, 2005, Tech. Rep..

[24] M. Vetterli and J. Kovačević, *Wavelets and Subband Coding*, ser. Signal Processing. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[25] P. Vaidyanathan, *Multirate Systems and Filter Banks*. Englewood Cliffs, NJ: Prentice-Hall, 1992.

[26] S. Müller and P. Massarani, "Transfer function measurement with sweeps," *J. Audio Eng. Soc.*, vol. 49, pp. 443–471, 2001.

[27] G. Elko, E. Diethorn, and T. Gänsler, "Room impulse response variation due to thermal fluctuation and its impact on acoustic echo cancellation," in *Proc. Int. Workshop Acoustic Echo Noise Control (IWAENC)*, Kyoto, Japan, 2003, pp. 67–70.

[28] T. Theußl, T. Möller, and E. Gröller, "Optimal regular volume sampling," in *Proc. IEEE Visualization*, 2001, pp. 91–546.

[29] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 4th ed. New York: Academic, 1965.

[30] R. N. Bracewell, *The Fourier Transform and Its Applications*. New York: McGraw-Hill, 2000.

Thibaut Ajdler (S'03–M'06) received the B.Sc. and the M.Sc. degrees in electrical engineering from the Free University of Brussels (Vrije Universiteit Brussel), Belgium, in 2001. He is currently working towards the Ph.D. degree in communication systems at Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland.

From July to September 2000, he was a Research Assistant at the Massachusetts Institute of Technology (MIT), Cambridge. From October 2002 to February 2003, he was an intern at Intel Corporation, Santa Clara, CA, working on acoustic source localization using distributed sensor networks. His research interests include audio signal processing, acoustics and signal processing for communications.

Luciano Sbaiz (M'98) received the "Laurea in Ingegneria" degree in electronic engineering and the Ph.D. degree from the University of Padova, Padova, Italy, in 1993 and 1998, respectively.

Between 1998 and 1999, he was Postdoctoral Researcher at the Audiovisual Communications Laboratory at Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland, where he conducted research on the application of computer vision techniques to the creation of video special effects. In 1999, he joined Dartfish, Ltd., Fribourg, Switzerland, as Project Manager. Within the company, he developed video special effects for television broadcasting and sport analysis. In 2004, he took a position as Senior Researcher at the Audiovisual Communications Laboratory at EPFL, where he conducts research on signal processing. His activities are in the field of image and audio processing, superresolution techniques, and acoustics.

Martin Vetterli (S'86–M'86–SM'90–F'95) received the Engineering degree from Eidgenössische Technische Hochschule (ETH), Zurich, Switzerland, in 1981, the M.S. degree from Stanford University, Stanford, CA, in 1982, and the Ph.D. degree from Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland, in 1986.

In 1986, he joined Columbia University, New York, first with the Center for Telecommunications Research and then with the Department of Electrical Engineering, where he was an Associate Professor of Electrical Engineering. In 1993, he joined the University of California at Berkeley, were he was Full Professor until 1997. Since 1995, he has been a Professor at EPFL, where he headed the Communication Systems Division (1996–1997) and heads the Audiovisual Communications Laboratory. From 2001 to 2004, he directed the National Competence Center in Research on mobile information and communication systems. He has also been Vice-President for International Affairs at EPFL since October 2004. He has held visiting positions at ETH Zurich (ETHZ) (1990) and Stanford University (1998). His research interests are in the areas of applied mathematics, signal processing, and communications. He is the coauthor of a textbook on *Wavelets and Subband Coding* and of over 100 journal papers.