AMBISONICS SYMPOSIUM 2009
June 25-27, Graz

CALCULATION OF NEAR-FIELD HEAD RELATED TRANSFER FUNCTIONS USING POINT SOURCE REPRESENTATIONS

Dylan Menzies

De Montfort University (rdmg@dmu.ac.uk)

Abstract:
We aim to improve the calculation of near-field head related transfer functions from far-field HRTFs using point source expansions into plane waves. We consider shifted expansions derived using the Fourier-Bessel representation and the Sommerfield-Weyl identity, which are also relevant to field reconstruction with speakers.

Key words: ambisonic binaural bessel sommerfield weyl point-source

1 INTRODUCTION

Far-field Head Related Transfer Functions (HRTFs) provide a well established tool for spatial auditory synthesis, but when applied directly are limited to the creation of distant images. The distance of sources within about 1.5m can be discriminated by listeners as a result of HRTF changes, independently of other factors such as environmental reflections.

Measuring HRTFs for near sources is both difficult, and dramatically increases the size of the complete data set. Near-field HRTFs can be estimated from far-field HRTFs by controlling the left and right angles and delays independently according to parallax to the ears, and by adjusting according to an analytic model for spherical head near-field HRTFs [1, 2, 3, 4]. However, informal listening shows that near-field HRTFs, especially the closest, are less convincing that the far-field HRTFs can be, when compared with the corresponding real sources, even when near-field listening is restricted by blocking the furthest ear. This implies that the unaccounted distance variation of HRTFs due to scattering sources such as pinna and shoulder, provide important location cues. To model the scattering accurately and generate near-field HRTFs we shall approximate point-source incident fields that are accurate on the most important scattering surfaces, using plane waves. The Sommerfield radiation condition then implies the scattered field, which includes the binaural signals, is accurate. In the following we review and expand some theoretical work previously presented [5], and consider an alternative incident field construction.

2 FOURIER BESSEL EXPANSION

In High Order Ambisonics (HOA), the field within a listening area is represented with a Fourier-Bessel expansion (FBE) [6]. The field is reproduced over a loudspeaker array, each speaker modeled as a point source or a plane wave source if the speakers are sufficiently distant. In connection with HRTF synthesis we consider full 3D HOA with plane waves speaker sources corresponding to far-field HRTFs. The FBE provides a route to generating a plane wave expansion (PWE).

Using HOA notation, the FBE expansion is given by

$$p(\boldsymbol{r},k) = \sum_m i^m j_m(kr) \sum_n B_{mn}(k) Y_{mn}(\hat{\boldsymbol{r}}) \qquad (1)$$

For a plane wave with source $S(k)$ the FBE coefficients are

$$B_{mn}(k) = S(k)\, Y_{mn}(\hat{\boldsymbol{r}}) \qquad (2)$$

From here on k dependence is omitted in most cases to simplify the presentation. For loudspeaker sources s_i the PWE is given by

$$p(\boldsymbol{r},k) = \sum_i s_i e^{-i\boldsymbol{k}_i \cdot \boldsymbol{r}} \qquad (3)$$

The expansions are related by

$$B_{mn} = \sum_i C_{imn} s_i \qquad (4)$$

where

$$C_{imn} = \sum_{m,n} Y_{mn}(-\hat{\boldsymbol{k}}_i) \qquad (5)$$

The inverse relation is obtained approximately with a regularized pseudo-inverse,

$$D_{imn} = (C^* C + \lambda I)^{-1} C^* \qquad (6)$$

where $C = C_{i\ mn}$ is taken as a two dimensional matrix. HRTF vector sets generally have large gaps in the downward direction where measurements cannot easily be made. This makes the simple source approach less attractive. The regularization helps to reduces ill-conditioning caused by the gap, and helps in other ways. Vectors are often patterned due to systematic measuring procedures, causing some degeneracy and loss of conditioning. Noise in the H_i makes them sensitive to large, ill-conditioned sets of s_i. Also higher unwanted harmonic components B_{mn} which are normally suppressed within the region of convergence can be come significant in the ill-conditioned case.

The PWE is now given by

$$s_i = \sum_{m,n} D_{imn} B_{mn} \tag{7}$$

The signal at the ear is the sum of the HRTFs H_i applied to the corresponding plane waves,

$$H(\hat{\boldsymbol{r}})S = \sum_{i,m,n} H_i D_{imn}(S\, Y_{mn}(\hat{\boldsymbol{r}})) \tag{8}$$

From which the HRTF for the direction $\hat{\boldsymbol{r}}$ is

$$H(\hat{\boldsymbol{r}}) = \sum_{i,m,n} D_{imn} Y_{mn}(\hat{\boldsymbol{r}})\, H_i \tag{9}$$

For a near source the FBE has additional terms $F_m(kr)$, known as distance functions [6]. The HRTF is modified as follows, with the optimal order of calculation made explicit,

$$B_{mn} = SF_m(kr)Y_{mn}(\hat{\boldsymbol{r}}) \tag{10}$$

$$H(\boldsymbol{r}, k) = \sum_{i,m,n} D_{imn} F_m(kr) Y_{mn}(\hat{\boldsymbol{r}})\, H_i(k) \tag{11}$$

$$H(\boldsymbol{r}, k) = \sum_{m} F_m(kr) H_m(\hat{\boldsymbol{r}}, k) \tag{12}$$

$$H_m(\hat{\boldsymbol{r}}, k) = \sum_{n} Y_{mn}(\hat{\boldsymbol{r}})\, H_{mn}(k) \tag{13}$$

$$H_{mn}(k) = \sum_{i} D_{imn} H_i(k) \tag{14}$$

The FBE point source can be made to approximate a real source as closely as desired in the the valid region, including the evanescent component, although the cost in order is unbounded. This may be counter-intuitive because the reconstruction is using only non-evanescent components.

The FBE provides a natural way to create a centred valid expansion region. However in its direct application as described above the region radius is limited to the distance from the centre to the source. This can be overcome by creating an FBE of the source about another centre then phase shifting the derived PWE so that it is relative to the desired centre. Equivalently, the FBE can translated directly [7]. In this way the limiting region that can be extended to a source is a half space, although at the cost of increased order. For HRTF synthesis this amounts to phase shifting the HRTFs, or delaying in the time domain,

$$H_i \rightarrow H_i e^{i\hat{\boldsymbol{k}}_i \cdot \boldsymbol{r}_t} \tag{15}$$

where $\boldsymbol{r}_t$ is the translation from the HRTF centre to the FBE centre.

As well as relaxing restrictions on the area of valid reconstruction, changing the FBE centre can be used to focus on important scattering regions such as the ear.

We have assumed plane wave reconstruction, and have seen this makes HRTF translation via PWEs very simple. If the base HRTF set is significantly near-field, a far-field set could be pre-calculated by including distance filters for the loudspeakers rather than the synthesized field point.

$$H(\hat{\boldsymbol{r}}, k) = \sum_{i,m,n} \frac{D_{imn}}{F_m(kr)} Y_{mn}(\hat{\boldsymbol{r}})\, H_i(k) \tag{16}$$

where r is the radius to the sources used to measure $H_i(k)$. This method might also have signal to noise advantages over far-field measurement.

3 FOCUSED SOURCES AND THE SOMMERFIELD-WEYL IDENTITY

Wavefield synthesis, WFS, [8] is a two dimensional speaker reproduction methodology, for which there is an established technique for generating virtual sources either in front or behind the speakers, by summing over plane waves evenly along a semi-circle:

$$p(\boldsymbol{r}) = \frac{1}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} e^{i\boldsymbol{k}(\theta)\cdot\boldsymbol{r}} d\theta = \frac{1}{\pi}\int_{-k}^{k} \frac{1}{k_y} e^{ik_x x + ik_y y} dk_x \tag{17}$$

The field approximates an outgoing line source in one half space and a converging line source in the other, hence it is sometimes called a focused source. Clearly such a field only approximates a line source, even in the continuous limit, because the centre point has finite pressure, however at distance it converges.

There is a natural extension of the focused source expansion, FSE, to three dimensions by integrating plane waves over a hemisphere. If such an FSE converges it could have several advantages over the FBE for near-field synthesis. The 2D FSE similarly could have uses for 2D only HRTF synthesis. An FSE of any order has a wide valid aperture from source, whereas an FBE must have high order. An FSE only uses plane wave vectors over a hemisphere, so reliance on HRTF directions which have not been measured can be avoided. HRTFs from an FSE require no additional distance filtering and can be generated directly by quadrature of shifted measured HRTF, using the simple-source formulation [9].

There are natural limitations to using FSEs. The near-field of the FSE is not accurate, however if the scattering surfaces are in the good FSE region, the synthesized HRTF should still be good at that frequency. For the lowest frequencies, the HRTF can be approximated with conventional, more direct techniques, as there is little angular variation.

The FSE is centred around the source, so more distant sources will require higher orders of expansion to reach the scattering surfaces. At some point other techniques will be preferable, perhaps using the conventional methods.

To better understand the theoretical basis for the focused source and its convergence in three dimensions, we now look at the Sommerfield-Weyl identity (18), [11]. A derivation via the Fourier Transform of the point-source is given in the Appendix.

$$\frac{1}{r}e^{ikr} = \frac{i}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{k_z}e^{ik_x x + ik_y y + ik_z |z|}\,\mathrm{d}k_x\,\mathrm{d}k_y \tag{18}$$

$k_z = \sqrt{k^2 - {k_x}^2 - {k_y}^2}$ is the positive real or imaginary root. This is not a plane wave expansion as it is discontinuous at $z = 0$. We can use this to write an expression for an ideal focused source over all of space, with outwards radiation on the positive z side,

$$\frac{1}{r}e^{-\bar{z}ikr} = -\bar{z}\frac{i}{2\pi}\iint_{-\infty}^{\infty}\frac{1}{k_z}e^{-\bar{z}(ik_x x + ik_y y) + ik_z |z|}\,\mathrm{d}k_x\,\mathrm{d}k_y \tag{19}$$

where $\bar{z}$ is the sign of z. The integral can be divided into two parts. For k_z real the integral identifies with the 3D FSE, with plane waves integrated uniformly over the hemisphere. The remaining part with k_z imaginary generates two half spaces of evanescent waves, which are not smoothly joined. In the positive z region, this part is the error in 3D FSE compared with a point-source. The error along the z-axis is given by

$$\frac{i}{2\pi}\int_{k}^{\infty}\frac{1}{k_z}e^{ik_z z}2\pi k_{xy}dk_{xy} \tag{20}$$

where $k_{xy} = \sqrt{{k_x}^2 + {k_y}^2}$. Changing the integration variable to k_z leads to an error of $1/z$. This is alarming because this is the same magnitude as the point-source for all z. It is is not a short range effect that would be associated with a real evanescent wave. For the 2D FSE a similar analysis, where the positive y axis bisects the outward radiating region, leads to an error

$$\frac{1}{\pi}\int_{0}^{\infty}\frac{1}{k_x}e^{ik_y y}dk_y \quad . \tag{21}$$

Since $k_x \to k$ as $k_y \to 0$, the asympototic error is α/y, the same decay rate as the 3D case. The 2D FSE error decays faster than the 2D source, which converges as $y^{-1/2}$, and so the 2D FSE is convergent along the y axis. This is consistent with successful use of focused line sources.

To understand the 3D FSE error over the full half space, we calculate it over a plane through the z axis. This has been carried out directly using vector sets, and also by calculating the FBE about the point source, with the FBE symmetry axis matching the z axis. This produces a compact set of harmonic coefficients, and provides a faster and more accurate way to evaluate the FSE field. The FBE representation can be used to realize a 3D FSE in HOA, in a similar way to the 2D FSE used in [10]. Shifting of the FSE centre relative to the HOA centre can be achieved using multipole

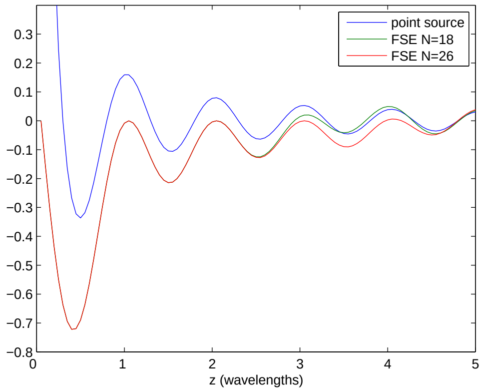

Figure 1: Re(p(z)) for the point source, and FBE FSEs.

re-expansion[1].

The FBE coefficients B_{mn} for a plane wave with vector $-\hat{\boldsymbol{k}}$, using HOA conventions, are $Y_{mn}(-\hat{\boldsymbol{k}})$. Referring to (19) the coefficients for a 3D FSE are then

$$\frac{-i}{2\pi}\int_{\boldsymbol{k}\cdot\hat{\boldsymbol{z}}>0}Y_{mn}(-\hat{\boldsymbol{k}})d\Omega \tag{22}$$

Integrating using spherical coordinates, coefficients for $n \neq 0$ vanish due to the azimuthal term vanishing. Changing the integration variable to $z = \cos(\theta)$ and using the Legendre polynomial identity $P_m(-z) = (-1)^m P_m$, the coefficients in m become, for all k,

$$B_{mn} = -i(-1)^m\sqrt{2m+1}\int_{0}^{1}P_m(z)dz \tag{23}$$

The integral is 1 for $m = 0$, 0 for even $m > 0$, and odd m have the closed form [12]

$$(-1)^{(m-1)/2}\frac{m!!}{m(m+1)(m-1)!!} \tag{24}$$

Fig **??** shows the error of the 3D FSE relative to the point source at different orders N. Fig **??** shows the real component along the z-axis. We can see the error already calculated along the z-axis for higher N.

It is striking that the most suitable overall error profile could be between $N = 6$ and $N = 10$, and good results can be achieved at even lower orders. High resolution PWEs are then formed by sampling only this bandlimited FBE, and not a higher order FBE, as might be expected.

4 SUMMARY

The FBE can be used to generate accurate half spaces of point sources, but at high cost. The FSE can produce half spaces of point sources with limited accuracy more efficiently, and it appears accuracy can be improved by using a

[1] The most efficient method of multipole re-expanslion may actually be via a PWE, which is readily shifted and converted back to a FBE

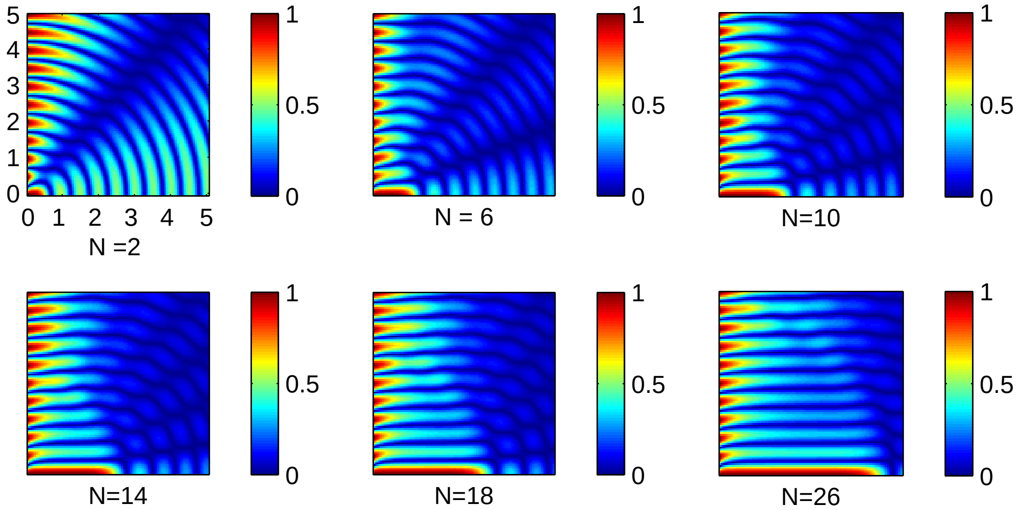

Figure 2: Relative error for a 3D FSE at different orders. The horizontal axes are z in units of wavelength.

limited FBE order approximation, with higher FBE orders set to zero. This may be valuable for near-field binaural synthesis, and 3D HOA over speakers.

5 APPENDIX

The following derivation of the Sommerfield-Weyl identity is adapted from [13].

A 3D point source is described by the Helmholtz equation,

$$\left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} + k_0^2\right] \varphi(x,y,z) = -\delta(x)\,\delta(y)\,\delta(z)\,, \tag{25}$$

Where $k_0 c = \omega$, the frequency of φ. In spherical coordinates the solution is

$$(2) \qquad \varphi(r) = \frac{1}{4\pi r}\, e^{ik_0 r}\,. \tag{26}$$

Negative time convention is used here, as it simplifies the form of the final result. We can also try and solve by expressing as a Fourier transform first. Assuming the Fourier transform exists, then the inverse Fourier transform has the form

$$\varphi(x,y,z) = \frac{1}{8\pi^3}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \widehat{\varphi}(\mathbf{k})\, e^{i\,\mathbf{k}\cdot\mathbf{x}}\, \mathrm{d}\mathbf{k} \tag{27}$$

The previous equations together with

$$\delta(x)\,\delta(y)\,\delta(z) = \frac{1}{8\pi^3}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{i\mathbf{k}\cdot\mathbf{x}}\, \mathrm{d}\mathbf{k} \tag{28}$$

give

$$\frac{1}{8\pi^3}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left[-k_x^2 - k_y^2 - k_z^2 + k_0^2\right]\, \widehat{\varphi}(\mathbf{k})\, e^{i\,\mathbf{k}\cdot\mathbf{x}}\, \mathrm{d}\mathbf{k} = -\frac{1}{8\pi^3}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{i\mathbf{k}\cdot\mathbf{x}}\, \mathrm{d}\mathbf{k} \tag{29}$$

Since the above equation holds for all values of $\mathbf{x}$ the Fourier components must agree, i.e.,

$$\left[-k_x^2 - k_y^2 - k_z^2 + k_0^2\right]\, \widehat{\varphi}(\mathbf{k}) = -1 \tag{30}$$

Or,

$$\widehat{\varphi}(\mathbf{k}) = -\frac{1}{k_0^2 - \mathbf{k}\cdot\mathbf{k}}\,. \tag{31}$$

The inverse Fourier transform is now

$$\varphi(x,y,z) = -\frac{1}{8\pi^3}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{k_0^2 - \mathbf{k}\cdot\mathbf{k}}\, e^{i\,\mathbf{k}\cdot\mathbf{x}}\, \mathrm{d}\mathbf{k} \tag{32}$$

Although the integral is well defined, it does not immediately express the field in terms of plane waves, as would be the case for a sourceless field. In order to find an expression that relates more directly to wave decompositions, we evaluate the integral over k_z first, without loss of generality. The poles are at

$$k_0^2 - \mathbf{k}\cdot\mathbf{k} = 0 \qquad \Longrightarrow \qquad k_z = \pm\sqrt{k_0^2 - k_x^2 - k_y^2}\,. \tag{33}$$

For $z > 0$ the integral is exponentially decreasing when $\mathrm{Im}(k_z) \to \infty$. Therefore, the integral over k_z can be split into the sum of an integral along the real line + an integral over an arc of a circle of radius infinity = sum of the residues at each of the poles in the upper arc.

Using the Residue theorem we can show that

$$\varphi(x,y,z) = \frac{i}{8\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{k_{zp}}\,e^{ik_x x+ik_y y+ik_{zp}z}\;\mathrm{d}k_x\;\mathrm{d}k_y \tag{34}$$

where k_{zp} is the value of $k_z$at the poles, i.e.,

$$k_{zp} := \pm\sqrt{k_0^2 - k_x^2 - k_y^2}\,. \tag{35}$$

When $z < 0$, the semicircular contour in the lower half plane must be taken, and picks up the residue at $-k_{zp}$. The result for all z can therefore be written as

$$\varphi(x,y,z) = \frac{i}{8\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{k_{zp}}\,e^{ik_x x+ik_y y+ik_{zp}|z|}\;\mathrm{d}k_x\;\mathrm{d}k_y\;. \tag{36}$$

This describes the sum of two half-spaces of travelling (k_{zp} real) and evanescent planewaves (k_{zp} imaginary) which meet non-smoothly.

Finally we can write the Sommerfield-Weyl identity,

$$\frac{1}{r}\,e^{ik_0 r} = \frac{i}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{k_z}\,e^{ik_x x+ik_y y+ik_z|z|}\;\mathrm{d}k_x\;\mathrm{d}k_y \tag{37}$$

REFERENCES

[1] G. F. Kuhn, "Model for the interaural time differences in the azimuthal plane," The Journal of the Acoustical Society of America, vol. 62, pp. 157-167, 1977.

[2] R. O. Duda and W. L. Martens, "Range dependence of the response of a spherical head model," The Journal of the Acoustical Society of America, vol. 104, pp. 3048-3058, 1998.

[3] D. S. Brungart, "Auditory localization of nearby sources. Head-related transfer functions," Journal of the Acoustical Society of America, vol. 106, pp. 1465-1479, 1999.

[4] A. Kan, "Distance Variation Function for Simulation of Near-Field Virtual Auditory Space," Proceedings of Acoustics, Speech and Signal Processing, ICASSP Toulouse 2006 .

[5] D. Menzies and M. Al-Akaidi, ÒNearÞeld binaural synthesis and ambisonics,Ó Journal of the Acoustical Society of America, 2007.

[6] J. Daniel. Spatial Sound Encoding Including Near Field Effect : Introducing Distance Coding Filters and a Viable, New Ambisonic Format. in AES 23rd International Conference. 2003.

[7] D. Menzies. Ambisonic synthesis of complex sources. JAES, 55(10):864Ð876, Oct. 2007.

[8] E.N.G. Verheijen. Sound reproduction by wave Þeld synthesis. PhD thesis, Delf tUniversity of Technology, 1997.

[9] M.A. Poletti. Three-dimensional surround sound systems based on spherical harmon- ics. Journal of the Audio Engineering Society (AES), 53(11):1004Ð1025, Nov. 2005

[10] J. Ahrens and S. Spors, "Focusing of virtual sound sources in higher order Ambisonics," Proceedings of the AES 124th International Convention, Amsterdam, 2008.

[11] H. Weyl. Ausbreitung electromagnetischer wellen uber einem ebenen leiter. ” Annalen der Physik”, 60:481–500, 1919.

[12] Byerly, W. E. "Zonal Harmonics." Ch. 5 in An Elementary Treatise on Fourier’s Series, and Spherical, Cylindrical, and Ellipsoidal Harmonics, with Applications to Problems in Mathematical Physics. New York: Dover, pp. 144-194, 1959.

[13] `http://en.wikiversity.org/wiki/Waves_in_com`