PREDICTION AND MEASUREMENT OF ACOUSTIC CROSSTALK CANCELLATION ROBUSTNESS

Felipe Orduña-Bustamante

Centro de Instrumentos
Universidad Nacional Autónoma de México
México D.F., CP 04510, MEXICO
e-mail: felipe@aleph.cinstrum.unam.mx

Jose J. López, Alberto González[*]

Departamento de Comunicaciones
Universidad Politécnica de Valencia
Camino de Vera s/n, 46022, Valencia, Spain
e-mail: {jjlopez, agonzal}@dcom.upv.es

ABSTRACT

The condition number of the matrix of electro-acoustic head-related transfer functions (HRTF) in a two-channel sound reproduction system has been used as a measure of robustness of the Atal-Schroeder crosstalk canceler. A comparative study has been made using results produced by computer simulations and HRTFs measured in an anechoic chamber by means of a dummy head. It has been found that acoustic scattering by the head has a very important and beneficial influence on robustness, specially for large loudspeaker separations. For narrow loudspeaker separations of less than about 40 degrees it is found that crosstalk cancellation exhibits a large variation of alternating very low and very high robustness. Also, simulations and measurements have been made of the natural channel separation under the same conditions. Scattering by the head is seen to provide a good level of natural channel separation at high frequencies and large loudspeaker angles. At low frequencies or small loudspeaker angles natural channel separation is poor.

1. INTRODUCTION

The robustness of acoustic crosstalk cancellation systems and the levels of natural channel separation already present in the sound reproduction system (without dedicated signal processing) has been studied as a function of frequency and loudspeaker separation. The condition number of the matrix of electro-acoustic head- related transfer functions (HRTF) in a two-channel sound reproduction system has been used as a measure of robustness of the Atal-Schroeder crosstalk canceler for these systems [1]. Also, natural channel separation can be used as an indication of how cooperative is the sound reproduction system with respect to the application of digital signal processing for crosstalk cancellation. A comparative study has been made using results produced by computer simulations and head-related transfer functions measured in an anechoic chamber by means of an artificial head and torso simulator. Transfer functions have been simulated in two ways: one which uses the sound pressure signal at two points in free space, set apart without obstruction at the standard inter-aural distance [2], [3], [4], and another which includes acoustic scattering from a rigid sphere to model the presence of a human head [5]. The sound sources are represented by simple acoustic monopoles in both cases. The rigid sphere model has been kept as simple as possible in order to emphasize the importance of scattering by the head alone in the explanation of the results [6], [7]. An HRTF data set has been measured for this work at Universidad Politécnica de Valencia (Spain), this has an angular resolution of 1 degree in azimuth on the plane of the head. This set has been used to assess the robustness of crosstalk cancellation and the levels of natural channel separation for symmetrical arrangements of two loudspeakers with angular separations from 0 to 180 degrees (plus or minus 90 degrees) and frequencies up to 8 kHz. These two aspects investigated here are related between them and they are also related to other aspects of cross-talk cancellation and spatial 3D sound systems which have not been touched upon here, such as size and shape of the zones of good sound reproduction [8], [9], [10] (also known as equalization zones or sweet spots), subjective acceptability [11], [12] etc. Some of these aspects and their relationships are the subject of current research, and they are of prime interest in the design of transaural systems for the reproduction of spatial sound.

[*]Partially supported by CICYT Grants TIC98-0343/99-0444

2. CONDITION NUMBER FOR CROSS-TALK CANCELLATION

Crosstalk cancellation in a two-channel, two-loudspeaker sound reproduction system, aims to minimize the difference between input signals and signals reproduced at the corresponding ears. That is, it tries to make the following identity hold as closely as possible:

$$\mathbf{I} = \begin{bmatrix} C_{11}(\omega) & C_{12}(\omega) \\ C_{21}(\omega) & C_{22}(\omega) \end{bmatrix} \begin{bmatrix} H_{11}(\omega) & H_{12}(\omega) \\ H_{21}(\omega) & H_{22}(\omega) \end{bmatrix}, \quad (1)$$

where the matrix with elements $C_{ij}(\omega)$ represents the electroacoustic response of the sound reproduction system, $H_{ij}(\omega)$ are the signal processing filters and $\mathbf{I}$ is the identity matrix. Disregarding the problem of practical realization, the signal processing filters which lead to an exact solution of the cross-talk cancellation problem are those corresponding to the inverse of the electroacoustic response matrix:

$$\begin{bmatrix} H_{11}(\omega) & H_{12}(\omega) \\ H_{21}(\omega) & H_{22}(\omega) \end{bmatrix} = \frac{1}{D(\omega)} \begin{bmatrix} C_{11}(\omega) & -C_{12}(\omega) \\ -C_{21}(\omega) & C_{22}(\omega) \end{bmatrix}, \quad (2)$$

where $D(\omega) = C_{11}(\omega)C_{22}(\omega) - C_{12}(\omega)C_{21}(\omega)$.

A relevant question is that of robustness of the cross-talk cancellation problem. This can be characterized by the condition number of the electroacoustic response matrix, in relation with the problem of solving a system of linear equations with this matrix. In the case of sound reproduction systems with two listening points (ears) and two loudspeakers, the electroacoustic response matrix is square. This might not be the case, and for greater generality, a pseudo-inverse approach must be considered where the matrix to be inverted (or solved for the unknown filter response functions) is either $\mathbf{C}^H\mathbf{C}$ or $\mathbf{C}\mathbf{C}^H$ depending on the dimensions and rank of the matrix $\mathbf{C}$ (subscript H denotes hermitian transpose). Thus, the condition number can be calculated as the ratio of the maximum and minimum singular values, as follows:

$$\kappa(\mathbf{C}) = \sigma_{max}/\sigma_{min}; \quad (3)$$

and, in linear algebra parlance, the condition number is a measure of the sensitivity of the solution vector against perturbations in the matrix of the system or in the right-hand side vector. Large condition numbers characterize ill-posed problems, which lack robustness in the sense that their solutions are highly sensitive to these perturbations.

2.1. Condition number results

Figure 1 shows contours of the reciprocal condition number for cross-talk cancellation at two points in free space separated the standard interaural separation $2a = 0.175$ m. Regions of low robustness (shown in white) are found at very low frequencies, also for very small loudspeaker separations, and for particular combinations of loudspeaker angle and frequency (of which several families exist) which make the cross-talk path exactly proportional to the direct path, which, in turn, render a singular cross-talk matrix. For instance, in this free space (non-scattering) case, and at a loudspeaker separation of 90 degrees, the frequencies of low robustness appear at frequencies for which the interaural separation $2a = 0.175$ m is a multiple of half a wavelength $n = \lambda/2$. These are frequencies which are multiples of $c/4a \approx 980$ Hz. This expectation corresponds very well with the results shown in Figure 1. And in fact, is a condition which remains approximately valid (for these extremely wide loudspeaker angles and at low frequencies) even when a rigid sphere or a head simulator is present, compare with Figures 2 and 3. In general, the similarities with this non-scattering case can be extrapolated approximately to the other cases for all loudspeaker angles at low frequencies, and for all frequencies at small loudspeaker angles.

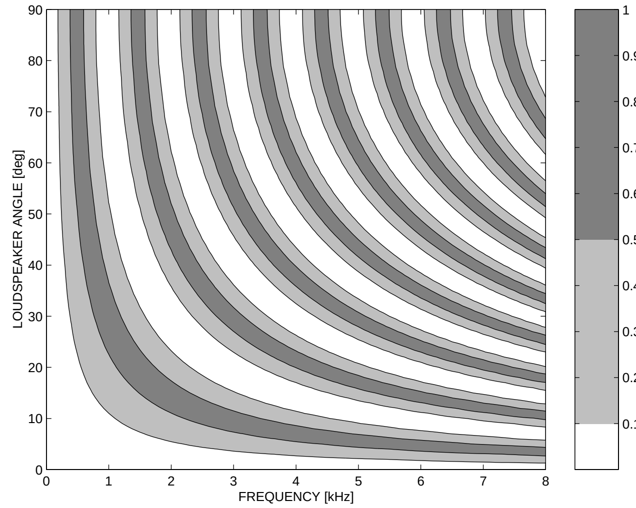

Fig. 1. Reciprocal condition number for cross-talk cancellation at two points in free space, separated the standard interaural distance $2a = 0.175$ m.

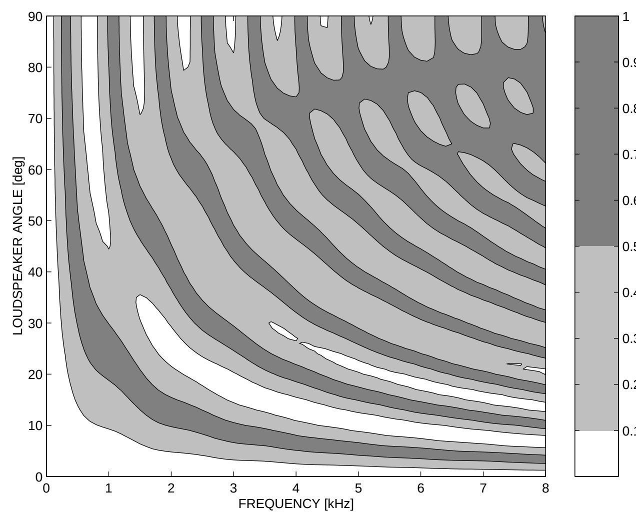

Fig. 2. Reciprocal condition number for cross-talk cancellation at two points on a rigid sphere of diameter $2a = 0.175$ m.

Figure 2 shows contours of the reciprocal condition num-

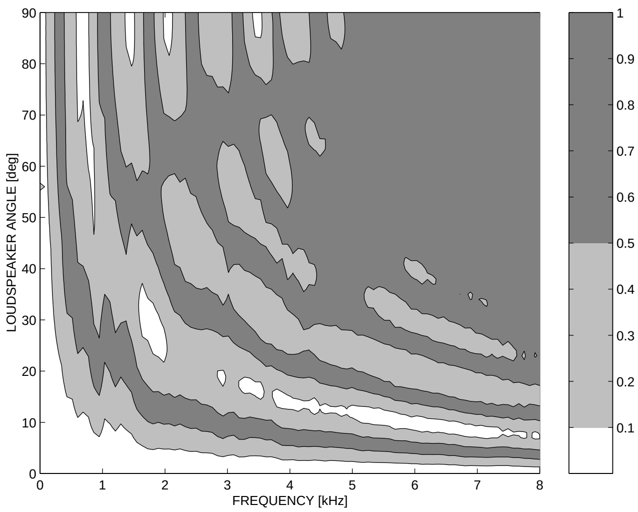

Fig. 3. Reciprocal condition number for cross-talk cancellation with the Bruel & Kjaer Head and Torso simulator measured at 1 degree intervals.

ber for cross-talk cancellation at two opposite points on a rigid sphere of diameter $2a = 0.175$ m. The most notable effect is a general increase in robustness at high frequencies and wide loudspeaker separations. Regions of low robustness persist at low frequencies and small loudspeaker separations. Figure 3 shows contours of the reciprocal condition number for cross-talk cancellation with the Bruel& Kjaer Head and Torso Simulator measured at 1 degree intervals. These corresponds acceptably well with the predictions from the rigid sphere calculations. The rigid sphere model is shown to provide a remarkably good agreement with these experimental results. The main difference being a somewhat higher increase in robustness at high frequencies and wide loudspeaker separations. Among other reasons, this difference appears because of the presence of pinna simulators in the experiments which are absent in the scattering calculations.

3. NATURAL CHANNEL SEPARATION

The aim of cross-talk cancellation is, of course, to make the matrix product in (1) as close as possible to the identity matrix. This is not normally achieved exactly in practice. One possible measure of the goodness of this approximation is *channel separation* defined as follows:

$$S_i(\omega) = \frac{|C_{ii}H_{ii} + C_{ij}H_{ji}|^2}{|C_{ii}H_{ij} + C_{ij}H_{jj}|^2}, \quad i,j \in \{1,2\}, \quad i \neq j; \quad (4)$$

which measures the transfer of energy from the direct (desired) signal relative to the cross-talk (unwanted) signal for each ear.

An important characteristic of a sound reproduction system is its *natural channel separation*, that is, the amount of channel separation already present in the system even without specially designed signal processing filters (when $\mathbf{H}$ is the identity matrix). Natural channel separation can be defined as follows:

$$S_i(\omega) = \frac{|C_{ii}|^2}{|C_{ij}|^2}, \quad i,j \in \{1,2\}, \quad i \neq j. \quad (5)$$

In general, sound reproduction systems with high levels of natural channel separation are expected to produce better results when signal processing is incorporated in them for improvements.

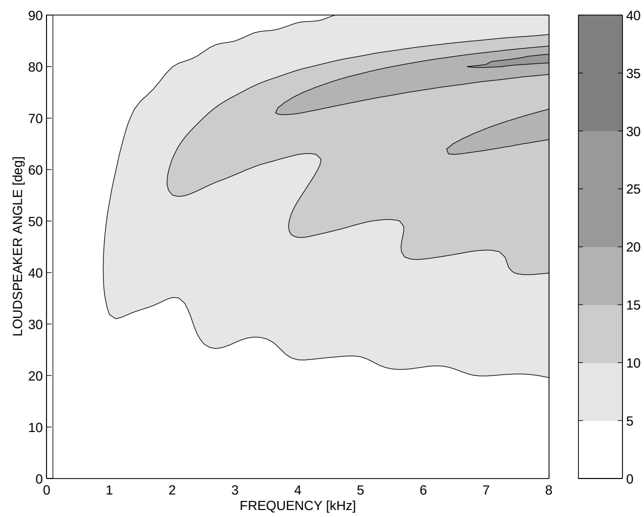

Fig. 4. Natural channel separation at two opposite points on a rigid sphere of diameter $2a = 0.175$ m.

3.1. Channel separation results

Figure 4 shows contours of natural channel separation at two opposite points on a rigid sphere of diameter $2a = 0.175$ m. A region of poor channel separation (less than 5 dB) appears at low frequencies or small loudspeaker separations. Channel separation increases progressively as frequency and loudspeaker separation increase. However, this increase is not monotonic. Note, for example, the appearance of bands of alternating high and low channel separation at high frequencies and very wide loudspeaker separations. These are characteristic of the angular dependence of scattering of sound by a rigid sphere (as shown, for instance, in Figure 8.4, page 420, in [5]).

Figure 5 shows contours of natural channel separation at the ears of the Bruel& Kjaer Head and Torso Simulator measured at 1 degree intervals. The general behaviour resembles that obtained with the rigid sphere model. Higher levels of channel separation are obtained at high frequencies

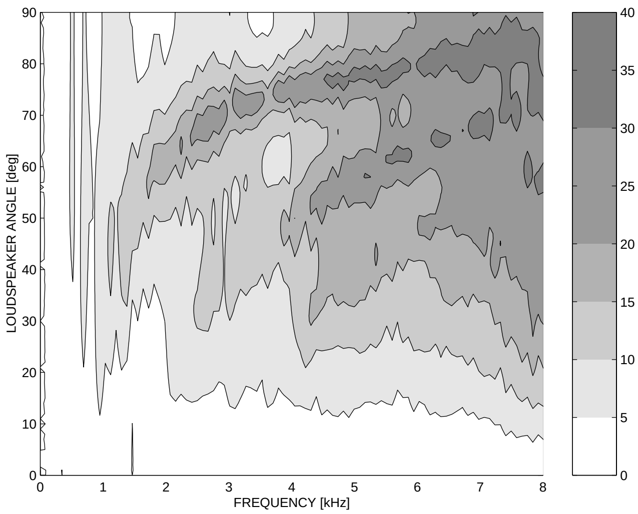

Fig. 5. Natural channel separation at the ears of the Bruel & Kjaer Head and Torso simulator.

and large loudspeaker angles. Also, bands of alternating low and high channel separation are hinted, which are similar to those predicted by the rigid sphere model.

4. CONCLUSIONS

Sound reproduction systems using two closely separated loudspeakers must be carefully designed with respect to loudspeaker separation in order to avoid an ill-conditioned system response. A band of very high robustness over a wide band at high frequencies can be found at an angular separation around 10 degrees (left and right loudspeakers at plus and minus 5 degrees, respectively). However, the robustness of cross-talk cancellation with this arrangement reduces very quickly for frequencies below 2 kHz and at all frequencies when the loudspeaker separation deviates from the optimal value of around 10 degrees. Additionally, natural channel separation is shown to be typically low, specially at low frequencies. Sound reproduction systems using two loudspeakers separated by more than about 40 degrees (loudspeakers at more than plus and minus 20 degrees) tend to be naturally robust. This occurs mainly as a consequence of acoustic scattering by the head. At this loudspeaker separations, and for frequencies higher than about 500 Hz (corresponding to acoustic wavelengths smaller than the average radius of the human head), acoustic scattering has the beneficial effect of marking perceptually important differences in the signals reproduced at the left and right ears. This is explained by the fact that when the loudspeaker arrangement is sufficiently wide, the left ear lies in the illuminated field of the left loudspeaker, but in the shadow field of the right loudspeaker, etc. This condition in itself, already produces a good level of natural channel separation which is very beneficial because it requires less effort from the cross-talk cancellation filters.

5. REFERENCES

[1] B. S. Atal and M. R. Schroeder, *Apparent sound source translator*. US Patent 3,236,949, 1962.

[2] Darren B. Ward and Gary W. Elko, “Effect of loudspeaker position on the robustness of coustic crosstalk cancellation,” *IEEE Signal Processing Letters*, Vol. 6, no. 5, pp. 106-108, 1999.

[3] Darren B. Ward and Gary W. Elko, “A new robust system for 3-D audio using loudspeakers,” In Proceedings of ICASSP 2000, Istambul, Turkey, June 2000.

[4] Jun Yang and Woon-Seng Gan, “Speaker placement for robust virtual audio display system,” *Electronics Letters*, Vol. 36, no. 7, pp. 683-685, 2000.

[5] Philip M. Morse and K. Uno Ingard, *Theoretical Acoustics*. Princeton University Press, 1968.

[6] G. Neu, E. Mommertz, and A. Schmitz,“Investigations on true directional sound reproduction by playing head-referred recordings over two loudspeakers: Part I (in German),” *Acustica*, Vol. 76, pp. 183-192, 1992.

[7] G. Urbach, E. Mommertz, and A. Schmitz, “Investigations on the directional scattering of sound reflections from the playback of head-referred recordings over two loudspeakers: Part II (in German),” *Acustica*, Vol. 77, pp. 153-161, 1992.

[8] Philip A. Nelson, Felipe Orduña-Bustamante, and Hareo Hamada, “Inverse filter design and equalisation zones in multichannel sound reproduction,” *IEEE Transactions on Speech and Audio*, Vol. 3, no. 3, pp. 185-192, 1995.

[9] Alberto González, Jose J. López and Felipe Orduña-Bustamante, “Measurement of cross-talk cancellation and equalization zones in 3-d sound reproduction under real listening conditions,” In Proceedings of the 16th International Conference on Spatial Sound Reproduction, Rovaniemi, Finland, April 1999. Audio Engineering Society.

[10] Alberto González, Jose J. López and Felipe Orduña-Bustamante, “Equalization zones for cross-talk cancellation as a function of loudspeaker position and room acoustics,” In Proceedings of the International Symposium on Active Control of Sound and Vibration ACTIVE 99, pp. 649-660, Fort Lauderdale, Florida, USA, December 1999.

[11] Philip A. Nelson, Felipe Orduña-Bustamante, and Hareo Hamada, “Multichannel signal processing techniques in the reproduction of sound,” *Journal of the Audio Engineering Society*, Vol. 44, no. 11, pp. 973-989, 1996.

[12] Philip A. Nelson, Felipe Orduña-Bustamante, and Hareo Hamada, “Experiments on a system for the synthesis of virtual acoustic sources,” *Journal of the Audio Engineering Society*, Vol. 44, no. 11, pp. 990-1007, 1996.