VIII International Conference on Recent Advances in Structural Dynamics
Southampton, UK, 14-16 July 2003

DEVELOPMENT OF AN EFFICIENT BINAURAL SIMULATION FOR THE ANALYSIS OF STRUCTURAL ACOUSTIC DATA

Marty E. Johnson and Aimee L. Lalime
Vibration and Acoustic Laboratories, Dept. of Mech. Eng, Virginia Tech
Blacksburg, Virginia, VA 24061-0238, USA

Ferdinand W. Grosveld
Lockheed Martin Engineering Sciences Company
NASA Langley Research Center
Hampton, VA 23681-2199, USA

Stephen A. Rizzi and Brenda M. Sullivan
Structural Acoustics Branch, NASA Langley Research Center
Hampton, VA 23681-2199, USA

ABSTRACT

Applying binaural simulation techniques to structural acoustic data can be very computationally intensive as the number of discrete noise sources can be very large. Typically, Head Related Transfer Functions (HRTFs) are used to individually filter the signals from each of the sources in the acoustic field. Therefore, creating a binaural simulation implies the use of potentially hundreds of real time filters. This paper details two methods of reducing the number of real-time computations required by: (i) using the singular value decomposition (SVD) to reduce the complexity of the HRTFs by breaking them into dominant singular values and vectors and (ii) by using equivalent source reduction (ESR) to reduce the number of sources to be analyzed in real-time by replacing sources on the scale of a structural wavelength with sources on the scale of an acoustic wavelength. The ESR and SVD reduction methods can be combined to provide an estimated computation time reduction of 99.4% for the structural acoustic data tested. In addition, preliminary tests have shown that there is a 97% correlation between the results of the combined reduction methods and the results found with the current binaural simulation techniques

INTRODUCTION

Virtual prototyping allows designers to view and analyze prototypes of automobiles, airplanes, buildings, etc., using three-dimensional virtual reality technology. This concept of virtual prototyping can be extended to include the acoustical properties of a prototype. For example, a three-dimensional audio-visual representation of an aircraft interior would allow designers and analysts to test, subjectively, the acoustic sound quality of the aircraft.

Paper No. 77

Unfortunately, the three-dimensional simulation (using binaural techniques) of structural-acoustic data can be difficult due to the large number of computations required in the analysis. The objective of this research has been to reduce the number of real-time calculations required in order to create a binaural representation of the acoustic field produced by a vibrating structure. Two reduction methods, and their combination, are investigated: Singular Value Decomposition (SVD) and Equivalent Source Reduction (ESR). SVD reduces the number of real-time computations and ESR performs pre-processing calculations that reduce the number of sources, and hence computations. The ESR also transforms the data into a format that allows the binaural signals to be calculated using the standard Head Related Transfer Functions or HRTFs.

Literature review. A great deal of research has been done on the modeling of HRTFs including principal components analysis [1,2], Karhunen-Loeve expansion [3,4], balanced model truncation [5], state-space analysis [6,7], representation as spherical harmonics [8,9], pole-zero approximation [10,11] and structural modeling [12,13]. Duda [14] presents a comprehensive summary of these different HRTF models, which have been researched for several reasons: to better understand the properties of HRTFs, to reduce the number of measurements required in order to create a full set of individualized HRTFs, and (as with the authors' research) to reduce the number and complexity of real-time calculations. In contrast, the research presented here uses the singular value decomposition (SVD) to model the HRTFs. SVD produces similar results to Principal Components Analysis and Karhunen-Loeve Expansion, but through a more direct computational method. Abel [15,16] introduced SVD as a computational reduction method for binaural simulation, but results regarding the accuracy and computational reduction of this method were not discussed. Equivalent source reduction (ESR) [17-19] is also used in order to simplify the binaural simulation of a distributed source.

This work combines the ESR and SVD methods and applies them to the binaural simulation of structural acoustic data. In addition, the reduction methods are implemented and results are compared with the computationally exhaustive method in order to determine the accuracy and computation time associated with the number of singular values.

THEORY

This section first describes a non-reduced or "exhaustive" method for creating a binaural simulation of a vibrating structure. Equivalent source reduction and singular value decomposition are then presented as ways to reduce the computation time involved in creating the binaural simulation of a vibro-acoustic source.

Calculating Binaural Signals. In order to couple together a model of the radiation of sound from a vibrating structure and the HRTFs, the authors have modeled the structure as an array of monopoles (acting as equivalent sources). Accordingly, this discussion focuses first on the radiation of sound from a monopole, then on sound radiation from a group of monopoles, and

finally on sound radiation from a vibrating structure. The use of monopoles is advantageous due to their simplicity, uniform directivity and ability to be used directly with HRTFs.

According to Fahy [20] the sound pressure at point R, $p(t)$, due to a uniformly vibrating sphere or a monopole in a free field is defined by Eq. (1):

$$p(t) = \left(\frac{\rho_0}{4\pi r}\right)\frac{\partial}{\partial t}\left[Q\left(t-\frac{r}{c}\right)\right] = \left(\frac{\rho_0}{4\pi r}\right)a\left(t-\frac{r}{c}\right) \tag{1}$$

where r is the radial distance from the monopole to the measurement point (R), c is the speed of sound in air, ρ_o is the density of air, Q is the volume velocity (Su), S is the surface area of the monopole ($4\pi r_o^2$), r_o is the radius of the sphere, and u is the normal velocity of the surface of the monopole. The rate of change of the volume velocity with respect to time can be described by a, the volume acceleration of the source. This relationship means that the sound at point R will have a magnitude that varies inversely with the distance r and a time delay equal to the time it takes for the sound to reach point R, which is r/c.

Having determined the pressure at point R due to a single monopole source, the pressure resulting from a group of monopoles can be calculated as a linear superposition of the pressure due to each source. Equation 2 shows the total pressure, $p(t)$, at point R resulting from N monopole sources:

$$p(t) = \sum_{n=1}^{N} p_n(t) = \sum_{n=1}^{N}\left(\frac{\rho_0}{4\pi r_n}\right)a_n\left(t-\frac{r_n}{c}\right) \tag{2}$$

where a_n is the volume acceleration of the n^{th} source and r_n represents the distance between the n^{th} source and the point R.

The above equation allows a monaural signal (i.e. sound at one point in the field representing the center of an observer's head) to be calculated, but it is now necessary to apply HRTFs in order to determine the binaural signals. HRTFs are functions of time (or frequency), angle, and distance. The angular position is defined by the elevation angle, ϕ, and the azimuth angle, θ, between the sound source and the center of the head (see Fig. 1). Because the authors use HRTFs [21] that were measured using sound sources that were each placed at a uniform (and relatively large) distance from the head (i.e. assumes plane waves), they are not considered a function of distance. The sound attenuation and delay due to distance are calculated through the radiation model (Eq. 2). Also, while each person has their own individual HRTFs, this research uses a generalized set that were measured by researchers at MIT using the Knowles Electronic Manikin for Acoustic Research (KEMAR) [21]. However, all of the work presented here can

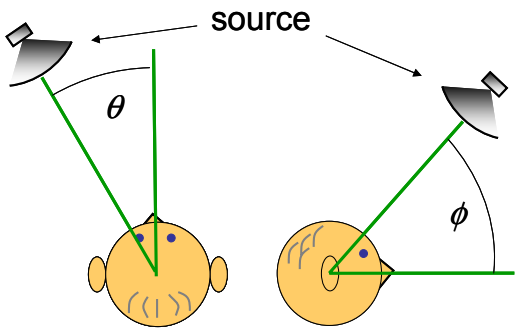

Figure 1: Angles relative to the center of the observer's head

easily be applied to any set of individualized HRTFs.

The binaural signals can be found by convolving the monaural signal with the appropriate Head Related Impulse Response (HRIR), which is the time domain version of the HRTF. However, since the HRIRs depend on direction, the monaural signal representing the vibrating structure (*p(t)* from Eq. 2) cannot be directly convolved with a single HRIR. Instead, the signal from each monopole must be convolved with its corresponding HRIR. For example the contribution, p_{Ln}, to the sound pressure at the left ear due to the n^{th} source can be calculated as:

$$p_{Ln}(t) = \int_0^\infty H_L(\theta_n, \phi_n, \tau) p_n(t-\tau) d\tau \tag{3}$$

where H_L is the HRIR for the left ear, and τ is a temporal variable. The pressure, p_n, resulting from the n^{th} monopole at angles θ_n, ϕ_n with respect to the head, represents the pressure at the center of the head *if the head were not present*. A similar equation can be written to describe the response at the right ear. The left and right signals contain all of the binaural information necessary to recreate a binaural simulation.

Since the binaural signals are calculated using a computer, the convolution of Eq. 3 is performed on discrete sampled signals rather than continuous ones.

$$p_{Ln}[i] = \sum_{j=0}^{M-1} H_L[\theta_n, \phi_n, j] p_n[i-j] \tag{4}$$

where M is the number of taps used in the HRIR finite impulse response (FIR) filter representation and i is the i^{th} discrete time step. This representation assumes that there will be a different HRIR filter for every angle. In reality a finite set of angles (spaced every 1 degree for example) are used and the filter closest to the true angle is used in the calculation.

The resulting left ear channel of the binaural signal, p_L, from the set of monopoles (representing the vibrating structure) is found by summing the pressure contributions from each monopole:

$$p_L[i] = \sum_{n=1}^{N} p_{Ln}[i] = \sum_{n=1}^{N} \sum_{j=0}^{M-1} H_L[\theta_n, \phi_n, j] p_n[i-j] \tag{5}$$

Note that this is a double summation requiring roughly NM multiply and add calculations every time step. Using this equation directly without modification will be termed *the exhaustive method*.

Equivalent Sources. In order to describe the vibration of a structure the measurement positions (or nodes) must be spaced at distances less than half of a structural wavelength (considerably less if possible). For thin panels, typically found in aircraft and cars, this can result in a huge number of points required to describe the vibration of the source. To overcome this, an equivalent source technique can be used to both reduce the number of sources or nodes (and hence computations), account for the non-planar geometry of the source and more generally convert the acoustic data into a format where fast real time binaural computations can be performed (i.e. make use of Eq. 5).

It is important to note that all of the equivalent source calculations are done in a pre-processing stage, with the end result of these calculations being a reduced set of input monopole sources strengths acting as if in a free-field. The assumption is that the vibration of the structure has been measured or calculated at a sufficiently large number of positions on the surface and is known *a priori*.

Figure 2 shows how the ESR method is accomplished by creating an *evaluation surface,* S_E, that surrounds the vibrating surface, S_0. S_E is evenly covered by an array of Q evaluation points where Q is large enough to accurately describe the velocity of the evaluation surface and prevent ill conditioning [17]. If S_E is placed even a small distance from the vibrating surface S_0 then it will not be affected by the non-radiating or subsonic components on S_0 since the air gap acts as a natural low pass wavenumber filter (i.e. oscillation with wavenumbers larger than the wavenumber in air, decay exponentially with distance from S).

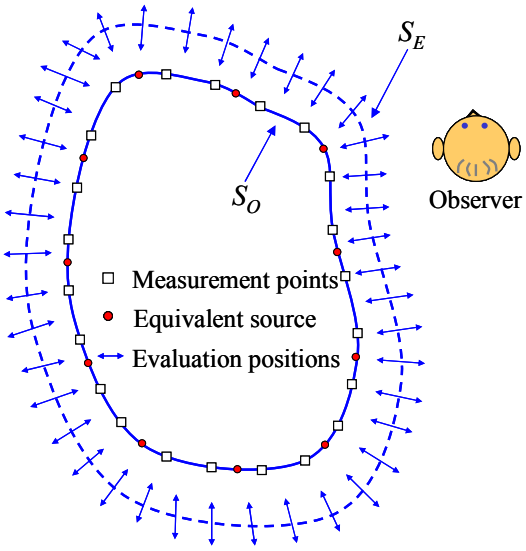

Figure 2: A reduced number of equivalent sources used to represent the radiation from a vibrating structure S_0

In this analysis the problem will be formulated in the frequency domain (for simplicity) and then converted into the time domain for real time implementation. This is possible because the analysis is all performed in a preprocessing stage. At a single frequency ω the complex particle velocity $\mathbf{v}(\omega)$, normal to the surface S_E, due to the original distributed source can be calculated using boundary element and/or finite element techniques, or in the case of a planar surface in a baffle using the Rayleigh integral. $\mathbf{v}(\omega)$ can also be calculated directly from the N length vector of measured (or calculated) accelerations $\mathbf{a}(\omega)$ using a transfer matrix approach.

$$\mathbf{v}(\omega) = \mathbf{T}(\omega)\mathbf{a}(\omega) \tag{6}$$

where $\mathbf{T}$ is an Q by N matrix of complex transfer functions between the Q measurement positions on the structure and the normal velocity at each evaluation position. Similarly, the normal particle velocity on the evaluation surface due to a reduced set of N_E equivalent sources can be calculated as,

$$\mathbf{v}_E(\omega) = \mathbf{T}_\mathbf{E}(\omega)\mathbf{a}_E(\omega) \tag{7}$$

where $\mathbf{T_E}$ is an Q by N_E matrix of complex transfer functions relating the normal particle velocity at each evaluation point to the acceleration of each equivalent monopole source and is calculated using the equations for radiation from a monopole acting in a free-field [17]. As long as the set of equivalent sources, of acceleration $\mathbf{a_E}(\omega)$, is driven such that it creates a particle velocity $\mathbf{v_E}(\omega)$ normal to the surface S_E that is the same (or very similar) to the velocity $\mathbf{v}(\omega)$**,** then the acoustic field outside the surface S_E will be the same as the original field. This can be deduced from the Kirchoff-Helmholtz equation [17, 22] assuming that there are no other sources acting in the field. The Kirchoff-Helmholtz equation requires that both the velocity and pressure at a surface be specified in order to calculate the acoustic field.

However, these quantities are not independent and "the pressure field is everywhere uniquely determined by a specified distribution of surface velocity" (from Fahy [22]). So as long as the velocity distribution on the surface S_E is correctly re-created the pressure field at the surface will also be correctly re-created. The equivalent source method circumvents the direct use of the Kirchoff-Helmholtz equation by matching the conditions at the boundary and the free-field Green's function (Eq. 1) can be directly used with the equivalent source strengths to calculate the pressure anywhere in the acoustic field outside the surface S_E.

In order to find the acceleration $\mathbf{a_E}(\omega)$ that minimizes the difference between the true velocity $\mathbf{v}$ and the simulated velocity $\mathbf{v_E}(\omega)$ an error minimization method, typically used in active control [17], is employed.

$$\mathbf{a_E}(\omega) = \left[\mathbf{T_E^H}(\omega)\mathbf{T_E}(\omega)\right]^{-1}\mathbf{T_E^H}(\omega)\mathbf{T}(\omega)\mathbf{a}(\omega) = \mathbf{C}(\omega)\mathbf{a}(\omega) \qquad (8)$$

This will minimize the difference (in a least squares sense) between $\mathbf{v}(\omega)$ and $\mathbf{v_E}(\omega)$ [17]. Therefore, the N_E by N matrix $\mathbf{C}(\omega)$ is a matrix that converts the N original accelerations into N_E accelerations that most accurately represent the radiation from the surface. This process can be conducted over a range of frequencies to build up a N_E by N matrix of filters in the frequency domain. The conversion matrix $\mathbf{C}(\omega)$ can then be transformed into the time domain using an inverse Fourier transform (to create $\mathbf{C}(t)$). Since this conversion is done in a pre-processing stage the time domain filters contained in $\mathbf{C}(t)$ are not constrained to be causal. The N time domain acceleration signals can then be filtered through $\mathbf{C}(t)$ to produce N_E equivalent accelerations that can be directly used in Eqs. 2 and 5. It should be noted that once this is done the equivalent set of monopoles act as if *in a free-field*.

If the radiation is into an enclosed space the equivalent sources can still be used to represent the radiation from the structure [17] and will take into account the resonant behavior of the enclosure (or equivalently compensate for all of the image sources). The equivalent monopole sources will still act as if in a free-field and Eqs. 2 and 5 can still be directly used.

Singular Value Decomposition. In addition to the ESR method, the authors have also investigated singular value decomposition [23,15,16] as a reduction method. By breaking down the matrix of HRIRs (minimum phase version) into three separate matrices, the number of convolutions can be greatly reduced. Singular value decomposition separates a matrix (in this case a matrix of HRIRs) into three separate matrices:

$$HRIR = [U][\Sigma][\mathrm{V}]^T \qquad (9)$$

where *HRIRs* is an m by n matrix of the HRIRs for *every* elevation and azimuth angle (in one degree steps), $\mathbf{U}$ is an m by m unitary matrix of left singular vectors, $\mathbf{\Sigma}$ is an m by n diagonal matrix of singular values, and $\mathbf{V}$ is an n by n unitary matrix containing right singular vectors. While singular value decomposition can also be performed in the frequency domain (i.e., on the HRTFs), the following analysis is in the time domain.

The SVD of the HRIR matrix (without time delays) is shown in Fig. 3. Notice that the left singular vector matrix, $\mathbf{U}$, contains angular information for each singular value σ_k. The matrix of singular values is simply a diagonal matrix of singular values that diminish in value as k becomes larger. The time information for each singular value is contained in the right

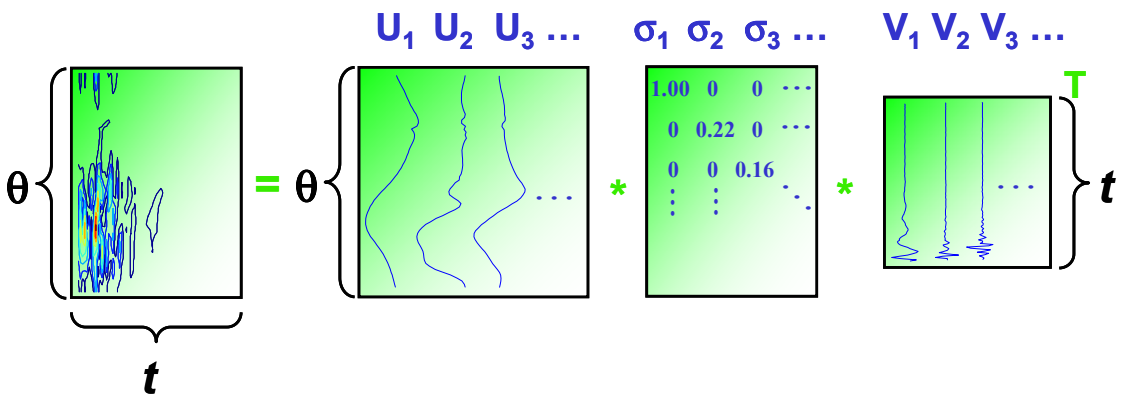

Figure 3: Graphical representation of the SVD of a minimum phase HRIR matrix. Left singular vectors are "mode shapes" and the right singular vectors are time domain FIR filters.

singular vector matrix. When these three matrices are multiplied together, the matrix of HRIRs (without time delays) will be the result. The advantage of SVD is that an approximate HRIR matrix can be obtained using only a few singular values (K is small), which is important when a large number of sources (N) are analyzed. Equation 5 can be approximated using K singular values/vectors as,

$$p_L[i] \approx \sum_{k=1}^{K} \sum_{j=0}^{M-1} V_k[j] s_k[i-j] \qquad (10)$$

where

$$s_k[i] = \sum_{n=1}^{N} \sigma_k U_k[\theta_n, \phi_n] p_n[i]$$

where V_k is the k^{th} right singular vector or filter, U_k is the k^{th} left singular vector or angular dependency. s_k is a time domain signal due to all of the sources driving the k^{th} mode or singular value. Although this may appear to be adding complexity to the calculation it actually represents a significant decrease in computation because the filtering process (summation over j) is performed on only K signals instead of N signals. Each time step now only requires the computation to calculate s (KN multiply and adds) plus the computations in the filtering stage (KM multiply and adds). This represents a significant reduction in computation if K is smaller than N (typically M=128). It should be noted that the right singular vectors are independent of any motion of the listener (i.e. filters are fixed).

Figure 4 shows the magnitude of the normalized singular values versus singular value number for four different frequency ranges (normalized such that the first singular value is equal to one). Notice that as the low-pass filter is applied at successively lower frequencies, the number of important singular values also decreases. It is found that only three singular values are necessary to create accurate results up to 2500 Hz.

RESULTS

In order to verify the effectiveness of the reduction methods, the authors chose to analyze a vibrating plate. This plate is set in a baffle and radiates sound into the free field. This structure and configuration was chosen for two reasons: (1) the plate is a relatively simple structure to analyze, allowing us to focus on the binaural processes rather than structural effects; and (2) a plate in this configuration is essentially the same as a plate secured in a baffle, radiating sound into an anechoic chamber, which can be arranged in a typical acoustic laboratory.

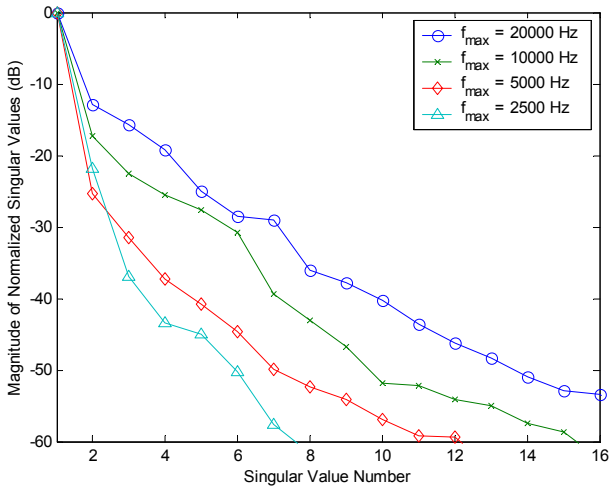

Figure 4: Magnitude of singular values of the HRIR matrix covering four different frequency ranges.

Description of Measurements. A 1.415 m x 1.415 m x 4.9 mm aluminum panel that was mounted in the transmission loss window of the SALT facility at NASA Langley Research Center and subjected to mechanical excitation. Using a scanning laser vibrometer, accelerations were collected by Grosveld [24] over a 23 x 23 grid (so $N = 529$). Also included in the data set are the transfer functions from the input force: to microphones in the anechoic chamber, and to two microphones in the left and right ears of NASA's KEMAR dummy head taken for a range of head positions and orientations. This data was used to verify the exhaustive method and was discussed in detail in [24, 25]. The authors also used this data to compare the reduction methods with the exhaustive method, which will be discussed in the following sections.

Results with the Exhaustive Method. With the plate measurements collected by Grosveld, the exhaustive method (Eq. 5) was used to find the binaural signals for the head and plate orientation shown in Fig. 5. This was achieved using a numerical form of the Rayleigh integral where the volume accelerations used in Eq. 2 were simply the measured accelerations times an elemental area. If the accelerations used are taken as the impulse responses from the input force to the acceleration at the measurement positions, then the output of the exhaustive method is the impulse response from the input force to the left ear. This calculated impulse response was compared with the measured [24] impulse response for this same head orientation. The two were found to match satisfactorily [25].

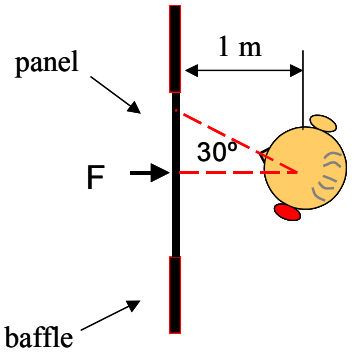

Figure 5: Head and plate orientation used in experiments

ESR. An evenly spaced grid of eight by eight equivalent sources (i.e., $N_E = 64$) and an evenly spaced grid of four by four equivalent sources (i.e., $N_E = 16$) were used to represent the radiation from the plate. The evaluation surface was a flattened hemisphere where the radius in the plane of the plate was 1.2m, and the radius normal to the plate was 0.6m. This surface was covered by 884 evenly spaced evaluation locations (see Fig. 6). The matricies **T** and $\mathbf{T_E}$ were calculated using the gradient of the pressure calculated from Eq. 2 [17].

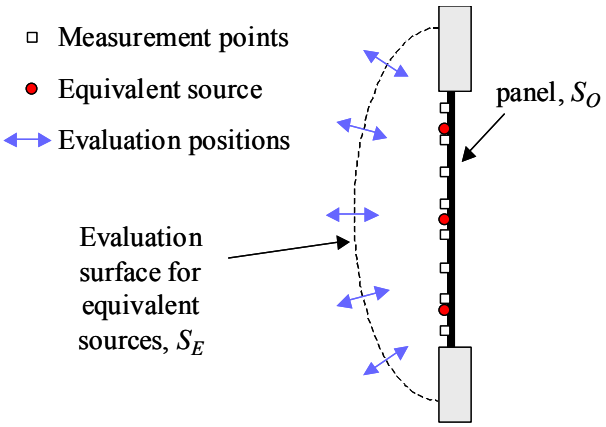

Figure 6: Evaluation surface used to calculate equivalent source strengths for the baffled panel

Figure 7 shows the frequency domain comparison of the pressure at a listener's left ear due to an impulsive point force driving the plate when 529 (exhaustive), 64, and 16 sources were used in the calculation. For this and all the following comparisons, the head and plate are configured according to Fig. 5. The decrease in accuracy above 1000 Hz (64 source case) is because the spacing of the equivalent sources becomes less than half an acoustic wavelength above 980 Hz. At times when only lower frequencies are important, a fewer number of equivalent sources can be used. For example, when only 16 sources are used, the ESR method is accurate up to 490 Hz. If frequencies up to 800Hz are of interest then the ESR method (with 64 sources) yields an excellent correlation of 98.4% in the time domain.

SVD. Figure 8 shows the beginning of the impulse response from the input force to the left ear as it is calculated with the exhaustive (529 sources) and SVD methods (only three singular values used). Since the collected data is good only up to 2000 Hz (due to the measurement spacing [25]), a low-pass filter at 1600 Hz was applied. The correlation between the SVD and exhaustive methods was 97.7%, while using only three singular values. From this result the authors concluded that only three singular values are necessary to make SVD an accurate reduction method for the frequency range of interest.

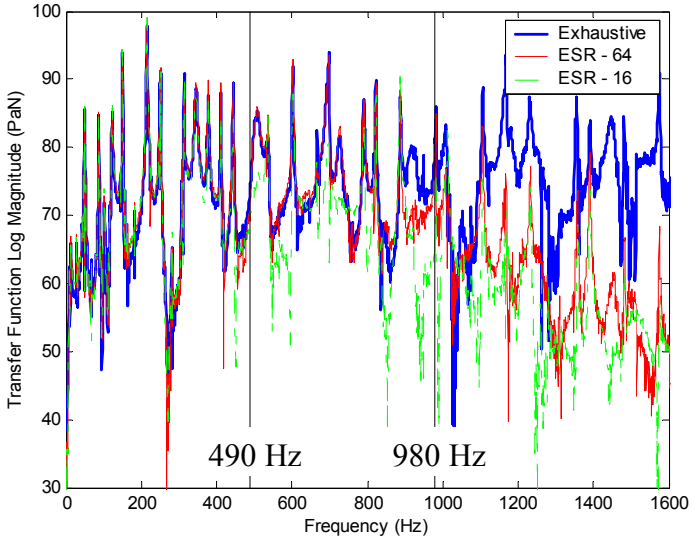

Figure 7: Frequency domain version of the sound at the left ear due to a force impulse when 529, 64 and 16 sources are used.

Combined ESR and SVD. Since the ESR method works to reduce the number of input sources while the SVD method changes how HRIRs are applied, the two methods can be used simultaneously to yield an even greater reduction in computation time. Figure 9 shows a time domain comparison of the combined reduction methods and the exhaustive method. In this case, three singular values and 64 equivalent sources were used. Because the ESR method for 64 sources is accurate only up to 980 Hz, a low-pass filter at 800 Hz was applied. A correlation of 97% was found between the combined reduction method and the exhaustive method.

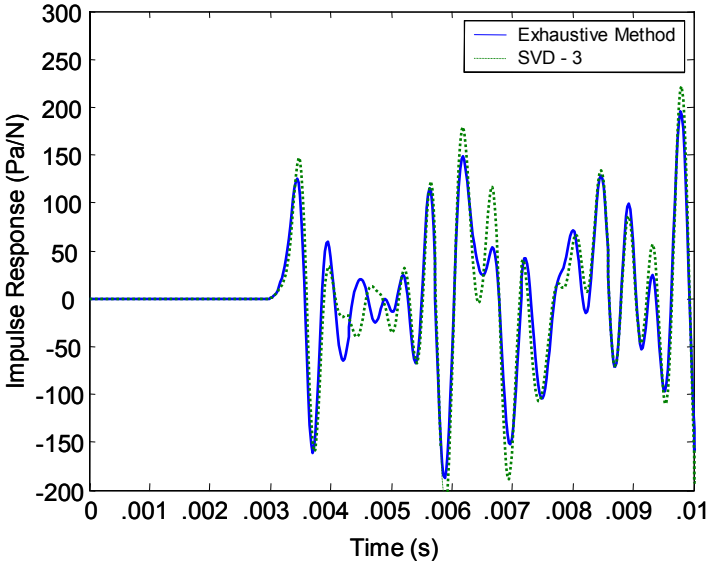

Figure 8: Beginning of the impulse response (low-pass at 1600Hz) between the input force and the sound at the left ear calculated using the full HRIR matrix and using only the first 3 singular values

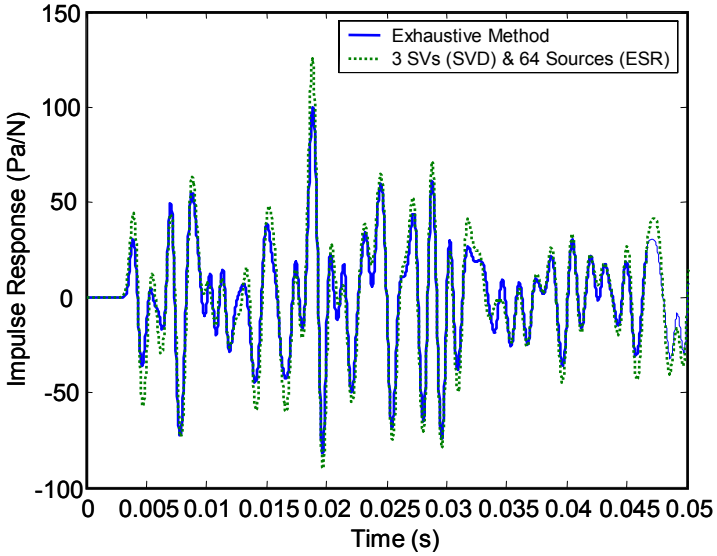

Figure 9: Beginning of the impulse response (low-pass at 800Hz) between the input force and the sound at the left ear calculated using the exhaustive method (full HRIR matrix and 529 sources) and using only 64 equivalent sources and the first 3 singular values

Computation comparisons. Since the radiation model process (Eq. 2) is a scalar multiplication (1/r), the number of computations is modeled to be N multiplications. Because convolution is required in the process of filtering the radiated pressure with the HRTFs (Eq. 5), the number of computations required will be MN multiplications plus MN additions (where M=128 is the number of samples in the HRIRs from MIT). Similarly, the additions required to sum the signal are counted as N summations. In this model, addition and multiplication are estimated to require approximately the same amount of computation time. Adding up the number of summations and multiplications required by the exhaustive method yields a total of $C_{EX}=(2M+2)N$ computations for each ear. This same equation is accurate for the ESR method as well. The only difference is found in the number

of sources (N). Notice from Table 1 that there is an obvious reduction in computation time achieved using the ESR method, but there is a tradeoff involving the frequency range of accuracy.

For the SVD method the computations required by the radiation model, though, are the same as with the ESR method: N multiplications. Multiplying the left singular vector, U_k, requires N multiplications for each of the K singular values, which is KN multiplications. By the same reasoning, the summing process requires MN summations. Since the convolution occurs after the summing process in the SVD method, only MK multiplications and MK additions are required. The scalar multiplication of the singular values is not included in the number of computations because it can be multiplied by the right singular vectors in pre-processing. The total number of computations is $C_{SVD}=N + 2KN + (2M+1)K$. Table 1 offers a comparison of the computation time associated with each method.

Table 1: Number of computations required for each ear for different simulations (M = 128)

Method	K	N	fmax	# comps
Exhaustive	N/A	529	N/A	**136482**
ESR	N/A	64	980	**16512**
ESR	N/A	16	490	**4128**
SVD	7	529	N/A	**9734**
SVD	3	529	N/A	**4474**
ESR & SVD	7	64	980	**2759**
ESR & SVD	7	16	490	**1911**
ESR & SVD	3	64	980	**1219**
ESR & SVD	3	16	490	**883**

This reduction in computation time is important in order to create a real-time binaural simulation of structural acoustic data. In a real-time virtual environment, the user's head orientation is measured by a head-tracking unit and read by the computer. Typical update rates for head-tracking units are 100 Hz, corresponding to one update every 10 ms. Until the computer receives another head orientation from the head-tracking unit, the left and right ear signals are calculated based upon the current head orientation. However, in audio systems the audio is typically updated at a sample rate of 44100 samples/second. This means that in order to create a structural acoustics simulation using the exhaustive method, the computer must make an estimated 136,482 calculations per ear in about 23 μs. With the combined reduction methods, only 883 calculations are required per ear in the same amount of time. This is an estimated computation time reduction of 99.4%, making real-time binaural analysis of structural acoustic data much more feasible.

CONCLUSIONS

The binaural simulation of structural acoustic data is a promising technology that is currently hampered by the extensive computations that are required. This paper has shown that singular value decomposition and equivalent source reduction can be employed simultaneously in the binaural simulation and, for this example, can reduce the number of real-time computations by 99.4%. In addition, these reduction methods are used while maintaining an accuracy of 97%, over the frequency range of interest, when compared with the exhaustive method. This work thus forms the foundation for implementation in a real-time filtering system. The only limiting factor with either method is that the ESR method has a frequency range of accuracy that depends upon the spacing of the equivalent sources.

Auditory tests will be required to determine if these reduction methods produce satisfactory psychoacoustic results. Also the application of ESR and SVD methods to enclosures and sources of complex geometries needs to be validated. Another important step in this research would be to use individualized HRTFs instead of generalized ones. Using individualized HRTFs would minimize HRTF error, allowing listeners in auditory tests to distinguish small errors in the reduction methods.

ACKNOWLEDGEMENTS

The authors would like to gratefully acknowledge the financial support provided by NASA under the Langley Research Center research cooperative agreement no. NCC1-01029.

REFERENCES

1. Martens, W.L. (**1987**), "Principal components analysis and resynthesis of spectral cues to perceived direction," *Proceedings of the International Computer Music Conference*, edited by J. Beauchamp (International Computer Music Association, San Francisco, CA), pp. 274-281.
2. Kistler, D.J. and F.L. Wightman (**1992**), "A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction," *J. Acoust. Soc. Am.*, **91**, pp. 1637-1647.
3. Chen, J., B.D. Van Veen, and K.E. Hecox (**1995**), "A spatial feature extraction and regularization model for the head related transfer function," *J. Acoust. Soc. Am.*, **97**, pp. 439-452.
4. Wu, Z., F.H.Y. Chan, F.K. Lam, and J.C.K. Chan (**1997**), "A time domain binaural model based on spatial feature extraction for the head-related transfer function," *J. Acoust. Soc. Am.*, **102**, pp. 2211-2218.
5. Mackenzie, J., J. Huopaniemi, V. Valimaki, and I. Kale (**1997**), "Low-order modeling of head-related transfer functions using balanced model truncation," *IEEE Signal Processing Letters*, **4**, pp. 39-41.
6. Georgiou, P.G., and C. Kyriakakis (**1999**), "Modeling of head related transfer functions for immersive audio using a state-space approach," *Conference Record of the Thirty-Third Asilomar Conference on Signals, Systems & Computers,* **1**, pp. 720-724.
7. Georgiou, P.G., A. Mouchtaris, S.I. Roumeliotis, and C. Kyriakakis (**2000**), "Immersive sound rendering using laser-based tracking," *Proceedings of the 109th Audio Engineering Society (AES) Convention*, preprint No. 5227.
8. Evans, M.J., J.A.S. Angus, and A.I. Tew (**1998)**, "Analyzing head-related transfer function measurements using surface spherical harmonics," *J. Acoust. Soc. Am.*, **104**, pp. 2400-2411.
9. Nelson, P.A., and Y. Kahana (**2001**), "Spherical harmonics, singular-value decomposition and the head-related transfer function," *J. Sound and Vib.*, **239**, pp. 607-637.
10. Jenison, R.L. (**1995**), "A spherical basis function neural network for pole-zero modeling of head-related transfer functions," *IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics*.

11. Runkle, P.R., M.A. Blommer, and G.H. Wakefield (**1995**), "A comparison of head related transfer function interpolation methods," *IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics*.
12. Brown, C.P., and R.O. Duda (**1998**), "A structural model for binaural sound synthesis," *IEEE Trans. on Speech and Audio Processing*, **6**, pp. 476-488.
13. Brown, C.P., and R.O. Duda (**1997**), "An efficient HRTF model for 3-D Sound," *1997 IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics*.
14. Duda, R.O. (**1994**), "Modeling head related transfer functions," *Proceedings of the 27th Asilomar Conference on Signals, Systems and Computers*, pp. 457-461.
15. Abel, J.S., and S.H. Foster (**1997**), "Method and apparatus for efficient presentation of high-quality three-dimensional audio," US Patent No. 5,596,644.
16. Abel, J.S., and S.H. Foster (**1998**), "Method and apparatus for efficient presentation of high-quality three-dimensional audio including ambient effects," US Patent No. 5,802,180.
17. Johnson, M.E., S.J. Elliott, K-H Baek, and J. Garcia-Bonito (**1998**), "An equivalent source technique for calculating the sound field inside an enclosure containing scattering objects," *J. Acoust. Soc. Am.*, **104**, pp. 1221-1231.
18. Koopmann, G.H., L. Song, and J.B. Fahnline (**1989**), "A method for computing acoustic fields based on the principle of wave superposition," *J. Acoust. Soc. Am.*, **86**, pp. 2433-2438.
19. Ochmann, M. (**1995**), "The source simulation technique for acoustic radiation problems," *Acustica*, **81**, pp. 512-527.
20. Fahy, F. (**2001**), *Foundations of engineering acoustics*, New York: Academic Press, pp. 110.
21. Gardner, W.G., and K.D. Martin (**1995**), "HRTF measurements of a KEMAR," *J. Acoust. Soc. Am.*, **97**, pp. 3907-3908, http://web.media.mit.edu/~kdm/hrtf.html.
22. Fahy, F. (**1985**), Sound and structural vibration: radiation, transmission and response, New York: Academic Press Inc. pp 115-116
23. Kung, S.Y. (**1978**), "A new identification and model reduction algorithm via singular value decompositions," *Conference Record of the Twelfth Asilomar Conference on Circuits, Systems and Computers*, pp. 705-714.
24. Grosveld, F.W. (**2001**), "Binaural simulation experiments in the NASA Langley structural acoustics loads and transmission facility," NASA/CR-2001-211-255, http://techreports.larc.nasa.gov/ltrs/PDF/2001/cr/NASA-2001-cr211255.pdf, Hampton: NASA Langley Research Center.
25. Lalime, A.L. and M.E. Johnson (**2002**), "Development of an efficient binaural simulation for the analysis of structural acoustics data," NASA/CR-2002-211753, Hampton: NASA Langley Research Center.