The concept of wave field synthesis (WFS) was introduced by Berkhout in 1988 [1]. It enables the generation of sound fields with natural temporal and spatial properties within a volume or area bounded by arrays of loudspeakers. Applications are found in real time performances as well as in reproduction of multi-track recordings. A logic next step was the formulation of a new wave field analysis (WFA) concept by Berkhout et al. in 1997 [2], where sound fields in enclosures are recorded with arrays of microphones and analyzed with postprocessing techniques commonly used in acoustical imaging. This way, both the temporal and spatial properties of the sound field can be investigated and understood. WFS and WFA meet in auralization applications: sound fields measured (or simulated) along arrays of microphone positions can be generated by arrays of loudspeakers for perceptual evaluation. Array technology was a core issue in the successful IST 5th framework project CARROUSO [3], which strongly contributed to worldwide knowledge and application of the concept.

1. INTRODUCTION

In traditional sound enhancement and sound reproduction practice, individual (groups of) loudspeakers are used to generate a replica of the recorded sound. Using high-quality systems in the appropriate manner, the temporal properties of this replica may be correct. Spatially, however, its properties are fully determined by the interfering radiation patterns of the loudspeakers. In figure 1, this is illustrated for a monochromatic source signal reproduced by two omnidirectional loudspeakers. Only in a limited listener area, the 'sweet spot', the perceived spatial image is correct. If, for instance, the two loudspeakers enhance the signal of a primary point source positioned behind them, most listeners receive one of the loudspeaker signals earlier than the primary signal which leads to mislocalization: the first arriving wave front determines the direction from which the sound is heard.

Traditionally, when measuring the sound field in an auditorium in order to analyze its properties in a physical or perceptual way, impulse responses are recorded with microphones placed at a limited number of rather arbitrarily chosen 'representative' positions. Each measured response, and the acoustic parameters (predictors of perception) derived from it, are supposed to be valid for some area around the corresponding microphone position. In practice it appears that (1) it is often difficult to physically interpret a single impulse response, and that (2) significant differences between neighboring microphone positions are observed which are also difficult to explain. This is not surprising, since this way only local temporal information about the sound field is obtained. Information on spatial properties related to interference and diffraction can not be taken into account.

In order to overcome the above mentioned drawbacks in reproduction and analysis of sound fields, array technology should be applied based on the concepts of wave field synthesis (WFS) and wave field analysis (WFA) which will be explained in the following sections.

2. THE WAVE FIELD SYNTHESIS CONCEPT

In this section the underlying theory of WFS is summarized. For an extensive treatment of the theory and the mathematical formulation, the reader is referred to [4].

According to the Huygens principle, the propagation of a wave through a medium can be qualitatively described by adding the contributions of all secondary sources positioned along a wave front. This can be mathematically worked out to the Kirchhoff representation theorem. This implies that, when the wave field on the boundary surface $S$ of a closed, source-free volume $V$ is known in terms of pressure and normal particle velocity, the sound pressure at any point within that volume can
be determined. It appears that surface \( S \) can be interpreted as covered with a continuous distribution of secondary monopole sources driven by the local normal velocity generated by the primary sources, plus a distribution of dipole sources driven by the local primary pressure. Together, all these virtual secondary sources can be seen as generating a field within \( V \) which is identical to the field that the primary sources would have generated there - or, when \( S \) is a virtual boundary, really generates there. When the closed boundary is replaced by a (pseudo)infinite plane with the primary sources at only one adjacent half-space, the sound pressure at any point in the other half-space can be determined whether from the normal velocity distribution in that plane, or from the pressure distribution. (This can be mathematically described by the Rayleigh representation theorems.) In the first case, the plane can be interpreted as being covered with secondary monopole sources driven by the local primary normal velocity distribution, in the second case as covered with secondary dipole sources driven by the local primary pressure distribution.

As described by Berkhout et al. [4], this concept is a strong base for application in audio and acoustics technology. When a plane is covered with loudspeakers having monopole source (i.e., omnidirectional) characteristics, being driven with signals corresponding to the normal velocity distribution in that plane generated by a real or virtual source or sources in one half-space, a replica (in case of real sources) or a simulation (in case of virtual sources) is generated in the entire other half-space. The same holds for a plane covered with dipole-type ('figure-of-eight') loudspeakers driven with signals corresponding to the primary pressure distribution. It has been shown [4] that with the use, instead of planar arrays as prescribed by the theory above, of linear arrays of loudspeakers - which are much more appropriate for practical use from a visual point of view and with respect to the hardware and computational power required - good results can be obtained in a horizontal plane, e.g., the earplane of an audience. Also, it has been shown [5] that the concept holds for any type of loudspeaker, by adapting the driving signal operator to its properties.

This means that, instead of the spatially erroneous wave field shown in figure 1, now a spatially (and temporally) correct wave field of a point source positioned behind the array is synthesized by all loudspeakers together, as illustrated in figure 2.

In the next section, applications of wave field synthesis will be discussed.

3. APPLICATIONS OF WAVE FIELD SYNTHESIS

With traditional electro-acoustic systems, the sound field is spatially correct at only one or a few 'sweet spots': local solutions are obtained. Systems based on WFS, however, generate a wave field with natural properties in time and space in an extensive audience area, yielding a volume solution.

3.1. Sound enhancement in theatres

When the instantaneous positions of actors and singers in theatre performances are known, their direct sound fields at the position of a loudspeaker array addressing the audience area can be calculated - usually, omnidirectionality of the primary sources is assumed as an approximation - and used to drive the individual elements of the array. This way, replicas of the primary field are generated which can be amplified with full preservation of the original properties in time and space. In order to know the instantaneous source positions, directional microphones, focussed microphone arrays or source tracking systems (in case of 'on body'-miking) have to be used. Care should be taken that feedback between loudspeakers and microphones is avoided.

3.2. Reproduction of multi-channel recordings

When multi-track recordings of sources (voices, instruments) at known positions are available, they can be replayed with preservation of the original spatial properties by means of WFS: within the total area between a (e.g., rectangular) configuration of loudspeaker arrays, a wave field is created that matches the wave field on the recording location [6]. This is especially the case when the acoustics of the recording venue (reflections, reverberation) has been recorded separately and is regenerated as a limited number (e.g., 8) of plane waves [7].

It is also possible to position the sources on other positions than their original ones ("build your favorite acoustic virtual reality"), even within the listeners area between the loudspeaker arrays using the focussing principle as illustrated in figure 3. Also, the sound of moving sources can be reproduced in a natural way.

WFS-based reproduction can well be combined with visual information. Applications are found in cinemas, home theatres and virtual reality theatres. A real-time application in this

Figure 2. Monochromatic signal of a point source placed behind an array of loudspeakers, reproduced by that array. Note the correct spatial and temporal properties.

Figure 3. With WFS, not only the sound of sources outside the listeners area (a), but also within that area (b) can be reproduced.
context is the WFS-based reproduction of speech in teleconferencing systems, in order to improve the agreement between visual and acoustical perception [8]. An interesting development here is the introduction of multi-actuator panel (MAP) loudspeaker modules as a WFS reproduction medium, which can simultaneously be applied as projection screens [9].

There is an increasing interest for WFS application from the world of electronic music, since it is recognized that the concept enables to play the compositions in a spatially controlled way [10].

3.3. Auralization

As discussed in the next section, wave fields in halls should be physically analyzed when recorded along arrays of microphone positions. When these recordings are regenerated by arrays of loudspeakers, the acoustics of the hall is auralized, i.e., made audible for perceptual evaluation [11]. In commonly used auralization techniques, the (often binaural) sound field at one listener position is reproduced, to be perceived with headphones or a pair of near-field loudspeakers. This way, the spatial properties of the sound field are quite difficult to assess, since often the full acoustic image is localized within the listener's head. In the WFS approach of auralization, however, listeners can 'walk around' along or within the wave field generated by the arrays. In order to properly simulate the sound field in a hall it is necessary to include the first order ceiling reflections. For this purpose, a ceiling array should be added to the configuration in the horizontal plane.

Note that this auralization can also be done for wave fields, simulated by some modeling algorithm along an array of microphone positions. This way, the differences between measured and modeled data can be perceptually evaluated. When modeling is done properly, architects and consultants can acquire a realistic impression of the acoustic properties of a space under design, and of the acoustic effects of possible modifications.

4. WAVE FIELD ANALYSIS

Berkhout et al. [2] have shown that multi-channel recording or calculation of impulse responses in an enclosed space along an array of microphone positions introduces a new concept of wave field analysis (WFA), yielding much insight in the temporal and spatial structure of the wavefield. An example is given in figure 4, showing the impulse responses measured in a 50-seat, rectangular lecture room at Delft University, along an array of microphone positions with 0.05m interspacing, over the full width of the hall at a distance of 5m from an omnidirectional source placed on the usual lecturer position centrally at the front side of the room. The vertical axis represents the traveltime coordinate t which equals zero when the pulse leaves the source. The horizontal axis gives the lateral microphone position x, the so-called offset, r, the center of the array which in this case coincides with the center of the room.

Already without any further processing the dataset clearly shows the wave character of the sound field. In spite of the complex structure of the field due to interference and diffraction, with array-based WFA many reflection and diffraction events can be easily discriminated since - other than when displaying individual, isolated impulse responses - now the spatial correlation between neighboring responses is revealed. By taking the hall geometry into account, the origins of many reflected or diffracted wave fronts can be identified, as indicated in figure 4. By applying a spatial Fourier transform or a Radon transform to the dataset, the wave field is decomposed into plane wave components [2] which enables further study of properties as diffusivity, lateral energy content, etc as a function of time.

Since the sound pressure was measured with an omnidirectional microphone, the dataset of figure 4 allows no discrimination in the elevation plane around the microphone array: wave fronts from front, back, above and below are all projected in the same offset-traveltime plane. When, however, not only the sound pressure is recorded, but also the three components of the particle velocity vector, on each array position a directional microphone can be simulated by post-processing [12]. This simulated microphone can be rotated around the microphone array under each azimuthal angle with the array between -90 and +90 degrees, such that wave components incident on the array under different azimuthal as well as elevation angles can now be discriminated.

Using wave field extrapolation techniques as developed for seismic extrapolation purposes [13], from a multi-channel recording along one array the responses at the position of any other array in the hall can be estimated. In principle, by combining the techniques of wave field extrapolation and directional microphone synthesis described above, one array measurement gives ample 3D information on the acoustics of the hall. Taking aspects of spatial resolution into account, it becomes clear that, in addition to recording or calculation along an array over the width of the hall, data acquisition along an array with front-to-back orientation and a vertical array improves this resolution.

Multi-channel array measurements have been carried out in several auditoria and concert halls, among which the Amsterdam Concertgebouw ('shoebox') and "De Doelen" in Rotterdam (hexagonal plane- as well as cross-section). The data
allow a physical comparison between the wave fields in such geometrically so different spaces. Besides, since the measured responses can be auralized with loudspeaker arrays as discussed in subsection 3.3, also perceptual comparison is now possible.

One of the results of the physical analysis is that the traditional room acoustical parameters as Clarity Index, Early Decay Time, Lateral Fraction, etc show such significant fluctuations on a small spatial scale - e.g., within the 12 microphone array positions at one and the same seat - that their relevance to predict perceptual cues is quite doubtful [14]. Apparently, these parameters are sensitive to local interference where the human perception is not. New, spatially more stable versions of the parameters should be defined based on perceptual evaluation of the measured datasets.

Measurements along a linear or cross array of microphone positions are relatively time consuming. Hulsebos et al. [15] showed that WFA can be done much faster and more efficient using a circular array of positions: a microphone is mounted on a rod and rotates slowly on a turntable; see figure 5. This way, for each source position a data set is acquired in about 15 minutes.

Figure 5. Impulse response measurements in the Amsterdam Concertgebouw along a circular array of microphone positions.

5. STATE OF THE ART

In the 1990s, more and more acousticians and audio engineers became aware of the unique possibilities of array technology for spatial sound control. In 2000, ten universities, research institutes and industries cooperated successfully in the CARROUSO project on the real time realization of multi-channel recording, encoding, transmission in MPEG-4 format, decoding and reproduction by WFS of acoustic scenes. The technology for such process is now available. As a spin-off, WFS systems are now commercially available and applied in cinemas and teleconferencing systems. Special recordings for WFS reproduction are made using microphone array technology. Composers of computer music are designing their own loudspeaker array configurations dedicated to the WFS effects they have in mind. Research is going on to further improve the sound quality of MAP loudspeaker systems.

6. REFERENCES