Cover: 'Synthesis of a spherical wave field by an increasing number of point sources'

Direct sound enhancement by wave field synthesis

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus,
Prof. dr. ir. J. Blaauwendraad,
in het openbaar te verdedigen
ten overstaan van een commissie,
door het College van Dekanen aangewezen,
op dinsdag 24 juni 1997 te 10:30 uur

door

Evert Walter START

natuurkundig ingenieur
geboren te Kerkrade

Dit proefschrift is goedgekeurd door de promotor:
Prof. dr. ir. A.J. Berkhout

Samenstelling promotiecommissie:

Rector Magnificus (voorzitter)	
Prof. dr. ir. A.J. Berkhout (promotor)	*Technische Universiteit Delft, Natuurkunde*
Prof. dr. ir. F.A. Bilsen	*Technische Universiteit Delft, Natuurkunde*
Prof. dr. ir. A.J.M. Houtsma	*Technische Universiteit Eindhoven, Natuurkunde*
Prof. dr. ir. G.L. Vermeir	*Katholieke Universiteit Leuven, Bouwkunde*
Prof. ir. L.C. Röling	*Technische Universiteit Delft, Bouwkunde*
Dr. ir. D. de Vries	*Technische Universiteit Delft, Natuurkunde*
Dr. ir. D.J. Verschuur	*Technische Universiteit Delft, Natuurkunde*

ISBN 90-9010708-8

Typesetting system: Framemaker 5.1
Printed in the Netherlands by: Beeld en Grafisch Centrum Technische Universiteit Delft.

Preface

In August 1992 I finished my M.Sc. project in the research group 'Perceptual Acoustics', supervised by professor Bilsen, at the faculty of Applied Physics in Delft. During my M.Sc. project I worked in the field of psycho-acoustics, where I was initiated in the principles of human auditory perception. The primary issue in psycho-acoustics is establishing the relationship between physical and perceptual characteristics of acoustical signals by performing subjective listening tests. An important part of the work is the generation of test the signals, in which digital signal processing plays an important role. This combination of subjective human auditory processing and the generation of signals which can be measured objectively, I have always found attractive.
In April 1992, professor Bilsen drew my attention to a Ph.D. project in the group of 'Seismics and Acoustics' supervised by professor Berkhout. The subject of the project, 'Development of a holographic array for direct sound enhancement', contained the ideal combination of subjective and objective acoustics. So, in September 1992 I started with my Ph.D. study.
Thanks to Peter Vogel, my predecessor on the project, I got familiar with the new research field very quickly. In the first year it became clear that in order to develop further the sound enhancement system Peter made, I had to go back to the 'roots' of wave field synthesis. This step back finally proved fruitful and resulted in an optimized and more flexible solution for direct sound enhancement by wave field synthesis. The results, which are hopefully very useful for other wave field synthesis applications, are reported in this thesis.

There are many people I worked with the last 5 years. First of all I would like to thank my promoter, professor Berkhout for supervising this thesis, his energy, stimulating ideas and fund-raising efforts.
Especially I wish to mention the great contribution of my co-promoter Diemer de Vries. I am very grateful for his conscientious reading of the manuscript and his comments. He always helped me getting things done and I certainly appreciate his frankness in many matters.

Also, I would like to thank my 'room-mates' and 'colleque-sound-controllers', Antwan, Edwin, Ivo, Jan-Jakob and Jaap. They all contributed to the pleasant working atmosphere in our section.
The contribution of several students to this thesis may not be forgotten. Thank you all: Vincent, Ben, Menno and last but not least Chiel, who helped me a lot with the evaluation of the DSE system and conducted the final listening tests.
It was always a pleasure to work with my old group 'Perceptual Acoustics'. Especially I am very grateful to Johan Raatgever for his advise and suggestions in the psycho-acoustical experiments.
Thanks to all the members of the staff, Leen and Edo for the computer support, Henry for the technical support and of course Riaz, who assisted me with the preparation of the cover of this thesis.
Finally, I want to thank my parents and Carlien for their love and support, and all my friends who helped me and make life fun. I love you all.

Contents

1

Introduction to wave field synthesis

In the last decades the application of electro-acoustical systems in enclosures, especially systems for speech enhancement, has assumed large proportions. Nowadays we shall hardly ever find a meeting room of medium or large size or auditorium which is not provided with a public address system for speech amplification.

The rapid development of micro-electronics, in particular the development of digital signal processors (DSP's), offered a whole new set of possibilities for acoustic control. These electroacoustic techniques allow a much simpler and more flexible variation of the acoustical conditions than the variation of acoustics by architectural means, like demountable orchestra shells, movable curtains, adjustable ceilings etc. The need for variable acoustics is given by the fact that many halls, largely for economical reasons, are frequently used for very different kinds of presentations, ranging from theatrical productions to various music performances.

However, the use of such systems does not mean that we can omit a careful acoustic design of the room. Without knowledge of the acoustical factors that influence e.g. speech intelligibility in a room, it is hardly possible to operate any acoustical system with optimal performance. Moreover, the experience acquired in the last decades with the application of electro-acoustical systems, extensively described by Ahnert (1993), shows that the generation of a natural sound field is not untroublesome. A well known problem is the danger of mis-localization with the use of public address systems. This means that listeners locate the sound sources at the loudspeaker positions instead of at their true positions on stage. A related problem is the difficulty of creating a natural sound image with conventional sound reproduction systems in e.g. cinemas, living rooms etc. A solution for these problems is given by the *wave field synthesis* concept, introduced by Berkhout (1988). Application of this principle in acoustics offers

unprecedented opportunities in sound control: direct sound enhancement, variable reverberation and sound reproduction.
In this thesis the design trajectory, including a physical and psychoacoustical evaluation, of a direct sound enhancement system based on wave field synthesis is described. The present work is a continuation and optimization of the work done by Vogel (1993).

1.1 The Huygens principle

First, a short historical overview will be given of the relevant theories which form the basis of the wave field synthesis concept.
According to the Huygens principle, formulated as early as 1690, each element of a wave front may be regarded as the center of a disturbance which gives rise to a secondary spherical wave front. Moreover he stated that the position of the full wave front at any later time is the envelope of all such elementary wave fronts, as shown in figure 1.1. A point source Ψ, positioned in a

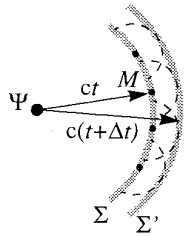


Figure 1.1: *Schematic representation of the Huygens principle.*

homogeneous medium, emits a spherical wavelet and Σ is the instantaneous position of the spherical wave front with radius ct at a time t, where c is the speed of sound. In the Huygens principle each point M of Σ is considered as a secondary source emitting a spherical elementary wave front. At time $t+\Delta t$ the wave front Σ' is the envelope of the secondary wave fronts with radius $c\Delta t$.
Fresnei, in 1818, supplemented this principle with the postulate that the secondary wave fronts mutually interfere. This combination of Huygens' construction with the principle of interference is called the Huygens-Fresnel principle.
With the Huygens principle in mind, it seems possible to calculate the wave field of the source Ψ by placing a distribution of secondary sources on the wave front Σ, each driven with a signal that is related to the local vibrations on Σ.
This idea was put on a mathematical basis by Kirchhoff in 1882, who showed that the Huygens-Fresnel principle may be regarded as an approximate form of a certain representation theorem, which states that at any receiver point within a source-free volume V, the sound pressure, due to sources outside V, can be calculated if both the sound pressure and its gradient are known on

the surface S enclosing V. This theorem also shows that each point on S contributes as a secondary dipole driven with the sound pressure at that point, and simultaneously as a secondary monopole, the strength of which is proportional to the gradient of the sound pressure (i.e. the normal component of the particle velocity).
If the surface S degenerates to a infinite plane surface, separating the receiver area from the source area, Kirchhoff's theorem can be transformed to the so-called Rayleigh theorems. In this case only knowledge of the sound pressure *or* the normal component of the particle velocity is required, yielding only a dipole *or* monopole distribution on the surface S.

1.2 Synthesis of sound fields

Based on Kirchhoff's and Rayleigh's representation theorems, Berkhout (1988) introduced the concept of *wave field synthesis* in acoustics. Berkhout proposed to *synthesize an acoustic wave field* by actually driving a real secondary monopole or dipole source distribution on a plane with the appropriate signals as defined by Rayleigh's theorem. In practice the sound pressure of a primary source Ψ, recorded by pressure microphones at a certain plane $z=z_S$, could be re-radiated by loudspeakers having dipole characteristics, as indicated by figure 1.2. In this way the

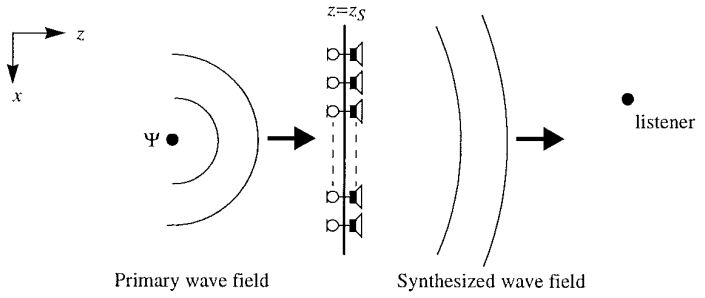


Figure 1.2: *Illustration of the basic principle of wave fields synthesis.*

wave field of the primary source is synthesized with full conservation of its temporal and spatial properties. The arrays of transducers used for recording and synthesis of the wave fronts can also be placed in planes with different coordinates z_M and z_S, as shown in figure 1.3. In this case the microphone signals recorded at the plane $z=z_M$ are extrapolated numerically with an operator $\boldsymbol{W}$ to the loudspeaker plane $z=z_S$.
When this process is carried out in real time, i.e., it is done simultaneously with the generation of the direct sound of the primary source, it means *direct sound enhancement*. The wave field on the plane $z=z_M$ could also be recorded first and reradiated at a later time and/or in a different space, in which case it means *sound reproduction*.

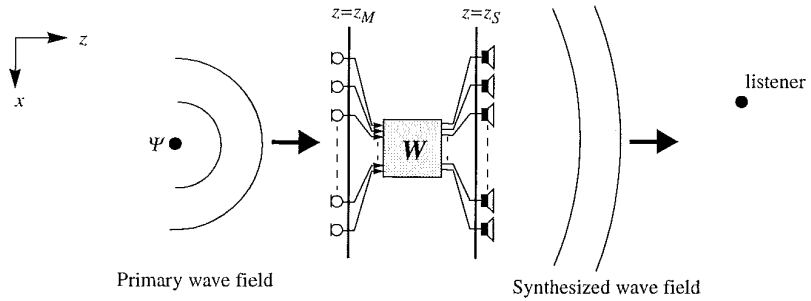


*Figure 1.3: Generalization of figure 1.2; wave field emission and recording in a different plane. Before emission, the wave field is extrapolated with an operator **W** from the microphone array to the loudspeaker array.*

Though the principle of wave field synthesis is quite simple, the practical realization is much more complicated, as will become evident in the next chapters.

1.3 Recording of source signals

In the basic processing schemes of figure 1.2 and 1.3 both the temporal and the spatial properties of the primary wave field are *recorded* with the microphone array. In a theater or concert hall situation, the primary wave field is the superposition of the wave fields of sound sources, e.g., actors, singers and/or instruments. To enable an individual processing and manipulation of each source signal, each source signal should be recorded separately. As a result, the source signal is decoupled from the spatial properties of the sound field.

When the positions of the primary sources are known and when we make assumptions about the directivity characteristics of the primary sources, the loudspeaker driving signals can also be *calculated*. Often, the sound sources will be represented by monopole sources, which means that they are assumed to have omni-directional characteristics. However, any desired directivity may be included. In addition, each recorded source signal can be assigned to a so-called *notional source* at the corresponding primary source position.

Basically three different strategies for the primary source position can be used:

- Use of a priori knowledge of the source positions combined with close-miking.
- Application of directional microphones (or microphone arrays) which each cover a certain part of the source area (remote-miking).
- Close-miking integrated with a suitable source tracking device.

Each of these techniques will be explored in this thesis.

1.4 The localization problem

An optimal public address system, if not used for special effects, should be implemented in such a way that its operation does not disagree with the usual natural acoustical perception. Therefore, attention must be paid to the following issues (Ahnert, 1993):

- The sound pressure level should be adjusted to the surroundings (background noise).
- The frequency response of the system should ideally be frequency independent in the frequency range of interest.
- Nonlinear distortions should remain below an acceptable level. The audibility of nonlinear distortions depends strongly on the frequency span of the original signal.
- The equipment should have a sufficiently large dynamic range and a low noise level.
- Acoustical feedback, which may lead to unwanted coloration and eventually to instability, must be reduced to acceptable proportions.
- The visual and acoustical localization of the sound sources should coincide.

The first five requirements all apply to the temporal properties of the amplified sound, and can be solved by common technical measures. The localization issue, however, which relates to the spatial properties of the emitted wave field, is not solved fundamentally in conventional public address systems.

To avoid false localization cues, many public address systems make use of the so- called 'Haas effect', 'law of the first wave front' or 'precedence effect'. Due to this psychoacoustic effect, the localization of a sound source is dominated by the first arriving wave front which is leading one or more wave fronts from other directions (see Haas, 1951; Gardner, 1968). In e.g. the Delta Stereophony System (DSS) delays are applied such that the loudspeaker signals arrive later at the listener than the actual source signal: the loudspeaker signal is perceived as pseudo-direct sound (Ahnert, 1986).

When the echo delay exceeds the *echo threshold*, the sound image breaks up in two separate parts. It has been shown that the echo threshold strongly depends on the type of stimulus employed. For speech of average speed the echo threshold is about 20 ms, while for a single click or brief noise burst the threshold decreases to 2-5 ms.

Even with the proper time delays (2-20 ms), the precedence effect only occurs if the leading wave fronts of the primary source and the wave fronts of the lagging loudspeakers are not too different with respect to orientation and loudness.

Other drawbacks of the precedence effect have been discussed by Blauert (1983) and Zurek (1987) and can be summarized as follows:

1. The presence of the lagging source is quite detectable. It results in a change of the loudness (which is not undesired in a public address system), spaciousness, and timbre of the perceived image.
2. Changes of the relative intensities of the primary sound and the indirect sound changes the echo threshold.
3. There is some influence of the lagging source on the apparent source position. For slow onset tones the precedence effect fails completely.

A fundamental approach to the localization problem can be accomplished by designing a direct sound enhancement system based on wave field synthesis. Unlike all existing methods, the wave field synthesis solution is a so-called volume solution that generates an accurate representation of the original wave field in the entire listening space.

1.5 The state of affairs at the start of the present research

As mentioned above, the present research is a continuation and optimization of the work done by Vogel (1993). Therefore, it is appropriate to give a short summary of his results.
In his thesis he argued that a high-quality electroacoustic system for (pseudo)direct sound and/or reflections should be based on the concept of wave field synthesis. The application of this concept leads, at least theoretically, to electro-acoustically generated sound fields, that have an arbitrarily close resemblance with natural sound fields.
Furthermore, it was made clear how the basic principle, described in section 1.2, could be adapted for practical application in auditoria:

- The impractical planar loudspeaker distribution (planar loudspeaker array) was replaced by a horizontal linear loudspeaker array. In this way the shape of the wave fronts remains unaffected in the horizontal plane through the linear array. A so-called 2½D operator was derived for driving a linear array consisting of monopole or dipole point sources.
- In order to reduce the number of microphones, the source area (i.e. the stage) was subdivided into sub-areas, each covered by a directive microphone. Hence, any signal recorded by a certain microphone can be assumed to be generated by a 'notional monopole source' in the center of the corresponding sub-area. Next, the driving signals for the loudspeaker array can be calculated.
- Due to the limited number of channels that can be processed by the hardware, the continuous and infinitely long linear array was sampled and truncated.

A prototype system was built, consisting of 48 loudspeakers. and 6 microphone inputs. The hardware consisted of 3 DSP's for the required signal processing, which comprised a complete matrix of weighted delay lines, connecting all microphone inputs to all loudspeaker outputs.
Listening tests in an anechoic room showed that the prototype system, with a spatial sampling distance of minimally 0.11 m, can generate a monopole source wave field in the horizontal plane with a well defined source localization.
The prototype system has also been tested in the auditorium of the Delft University of Technology using six microphones to address the stage area. The experiments showed that a 5 dB(A) level increase of the direct sound was possible, without any intolerable coloration due to acoustic feedback. Due to this experiment, a first wave field synthesis system has been installed in a theatre in Sweden (Malmø Stadsteater).

1.6 Objective of this thesis

By the research, summarized in section 1.5, new insights and practical experience with a wave field synthesis system had been gained. However, many topics remained to be explored deeper.
In contrast to the solid theoretical basis of the wave field synthesis concept, formulated by Kirchhoff and Rayleigh, the operation of a system adapted to the practical requirements in a theater, was insufficiently understood.
Therefore, a main objective of the present thesis is to enlarge the knowledge and expertise of a wave field synthesis system when applied in practical circumstances.
The design trajectory of a direct sound enhancement system based on wave field synthesis will be described, and a physical and psychoacoustical evaluation of the system will be given.
Much of the expertise obtained with the development of the direct sound enhancement system, (DSE system) can also be applied in other areas of sound control (sound reproduction and reverberation control).
The thesis is set-up in a way parallel to the design trajectory. It starts with acoustical principles of wave field synthesis, presented in Chapter 2.
After that, in Chapter 3, synthesis operators for line arrays in a 3D space are derived. Due to a new approach, besides operators for straight line arrays, also operators for bent line arrays can be derived.
In Chapter 4 the performance of continuous finite secondary source arrays is analyzed. Diffraction theory plays an important role.
The next Chapter 5 is devoted to the performance of discretized secondary source arrays. It is investigated under which conditions it is possible to achieve an accurate representation of the desired primary wave field. To improve the performance of the system, the technique of spatial bandwidth reduction applied to the synthesis of sound fields is introduced. It is shown that the proposed method of spatial bandwidth reduction can be combined elegantly with the application of bent arrays.
Chapter 6, which is an intermezzo, deals with the psychoacoustic aspects of synthesized wave fields. A preliminary investigation has been carried out, in order to determine the similarities and differences between a synthesized wave field and the original wave field emitted by a source. These measurements were done in an anechoic room.
Chapter 7 describes the implementation of the developed synthesis techniques in the DSE system. Two main elements of the DSE system are described: the recording part and the reproduction part.
Finally, in Chapter 8 the DSE system is evaluated, both physically and psychoacoustically, under several acoustic conditions.

2

Acoustical principles of wave field synthesis

In this chapter the theoretical principles of wave field synthesis are formulated. Most of the acoustical relations used can be found in many textbooks on acoustics like Pierce (1981) or Berkhout (1987). However, for the succeeding parts of this thesis it is useful to summarize the most important results. It will be demonstrated that a given wave field in an arbitrary closed volume can be exactly reconstructed by a suitable chosen source distribution. This phenomenon is closely related to the Huygens principle of wave field propagation. The quantification of the Huygens principle, as formulated in the Kirchhoff-Helmholtz theorem and the Rayleigh theorems, will be discussed and illustrated with some examples.

2.1 Fundamental wave theory

2.1.1 The wave equation in a sourceless medium

In this section the two fundamental equations of linear acoustical wave theory are presented. From these equations the wave equation will be derived. The medium considered is a homogeneous and isotropic fluid, that is inviscid and non-thermally conducting.
The law of conservation of mass states that the rate of change of mass inside an arbitrary volume equals the net mass flux entering the volume. Expressed in acoustic field quantities the linear approximation of the law of mass conservation for the fluid reads

$$\frac{1}{c^2}\frac{\partial}{\partial t}p(\mathbf{r}, t) + \rho\nabla \cdot \mathbf{v}(\mathbf{r}, t) = 0, \tag{2.1}$$

in which the scalar p and the vector $\mathbf{v}$ are the acoustic pressure and velocity as a function of position in scalar coordinates, $\mathbf{r}=(x,y,z)^T$, and t denotes time. The quantities c and ρ stand for the speed of sound and density respectively.
Secondly, the equation of motion gives the relation between the spatial variation of pressure and the corresponding particle acceleration. The linearized form of this momentum conservation law is given by

$$\rho\frac{\partial}{\partial t}\mathbf{v}(\mathbf{r}, t) + \nabla p(\mathbf{r}, t) = \mathbf{0}. \tag{2.2}$$

Combining the conservation expressions (2.1) and (2.2) gives the following scalar wave equation:

$$\nabla^2 p(\mathbf{r}, t) - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}p(\mathbf{r}, t) = 0. \tag{2.3}$$

An identical relation can be given for each of the Cartesian components of the particle velocity $\mathbf{v}$.
The acoustic pressure p and velocity $\mathbf{v}$ may be decomposed into monochromatic functions $e^{j\omega t}$, in which ω stands for the angular frequency, by applying the Fourier transform defined by

$$P(\omega) = \int_{-\infty}^{\infty} p(t)e^{-j\omega t}dt. \tag{2.4}$$

The inverse transform is given by

$$p(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} P(\omega)e^{j\omega t}d\omega. \tag{2.5}$$

Application of the Fourier transform (2.4) to the wave equation (2.3) results in

$$\nabla^2 P(\mathbf{r}, \omega) + k^2 P(\mathbf{r}, \omega) = 0, \tag{2.6}$$

the homogeneous Helmholtz equation, in which k represents the wavenumber given by ω/c.

2.1.2 The inhomogeneous wave equation

The fundamental relations presented in the previous section all apply to a sourceless medium. In order to have a general description of wave phenomena in presence of acoustical sources, the fundamental equations have to be extended. Two basic types of sources are distinguished: volume sources and force sources.
The first type introduces a fluctuating volume flow into the medium, for example a radially pulsating sphere. The equation of mass conservation can be rewritten for this situation in the following form:

$$\frac{1}{c^2}\frac{\partial}{\partial t}p(\mathbf{r}, t) + \rho\nabla\cdot\mathbf{v}(\mathbf{r}, t) = \rho q(\mathbf{r}, t), \tag{2.7}$$

in which q indicates the source strength density [1/s], and the product ρq represents the rate of mass introduction per unit volume.
The second source type acts as a fluctuating force per unit volume. In this case the law of conservation of momentum can be adapted to

$$\rho\frac{\partial}{\partial t}\mathbf{v}(\mathbf{r}, t) + \nabla p(\mathbf{r}, t) = \mathbf{f}(\mathbf{r}, t), \tag{2.8}$$

in which the vector $\mathbf{f}$ is the external force per unit volume.
Combination of these conservation equations (2.7) and (2.8) yields the inhomogeneous wave equation (in a homogeneous medium)

$$\nabla^2 p(\mathbf{r}, t) - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}p(\mathbf{r}, t) = \nabla\cdot\mathbf{f}(\mathbf{r}, t) - \rho\frac{\partial}{\partial t}q(\mathbf{r}, t). \tag{2.9}$$

The analysis of this inhomogeneous wave equation is often more convenient in frequency components. Application of the Fourier transform gives

$$\nabla^2 P(\mathbf{r}, \omega) + k^2 P(\mathbf{r}, \omega) = \nabla\cdot\mathbf{F}(\mathbf{r}, \omega) - j\omega\rho Q(\mathbf{r}, \omega), \tag{2.10}$$

which is called the inhomogeneous Helmholtz equation.
In order to solve the inhomogeneous Helmholtz equation (2.10) initial and boundary conditions are needed. For an infinite homogeneous medium Sommerfeld's radiation condition holds, which requires that merely outgoing waves are allowed. The outward propagation can be characterized by the two criteria (Bleistein, 1984)

$$\lim_{r\to\infty}\left[r\left(\frac{\partial}{\partial r}P(\mathbf{r}, \omega) + jkP(\mathbf{r}, \omega)\right)\right] = 0 \tag{2.11a}$$

$$P(\mathbf{r}, \omega) = O\left(\frac{1}{r}\right). \tag{2.11b}$$

Here r is the radial distance measured from the finite domain containing the acoustical sources to the observation position at $\mathbf{r}$. The first criterion characterizes the direction of propagation while the second criterion states that the solution must decay to zero at the rate of $1/r$ as $r\to\infty$.
A special kind of volume source is the monopole, that can be described as a volume source at $\mathbf{r}_0$ with a source strength distribution given by

$$Q_m(\mathbf{r}, \omega) = U(\omega)\delta(\mathbf{r} - \mathbf{r}_0), \tag{2.12}$$

where $U(\omega)$ represents the volume velocity of the source. The spatial delta function describes the point source behavior, i.e., a source strength that equals zero everywhere except at $\mathbf{r}_0$.
The solution of the inhomogeneous wave equation (2.9) for the free field boundary conditions (2.11a,b) with the monopole source term (2.12) is

$$P_m(\mathbf{r}|\mathbf{r}_0, \omega) = S(\omega)\frac{e^{-jk|\mathbf{r}-\mathbf{r}_0|}}{|\mathbf{r}-\mathbf{r}_0|}, \tag{2.13}$$

in which $S(\omega)$, given by

$$S(\omega) = j\omega\frac{\rho_0 U(\omega)}{4\pi} \tag{2.14}$$

represents the source spectrum. Note that the wave fronts are spherically symmetrical and propagate away from the source at $\mathbf{r}_0$ with propagation velocity c. An observer at position $\mathbf{r}$ receives the source information retarded by travel time $|\mathbf{r}-\mathbf{r}_0|/c$ and attenuated by a factor proportional to the distance.

The particle velocity $\mathbf{V}_m(\mathbf{r}|\mathbf{r}_0,\omega)$ of a monopole can be found by inserting the expression for the monopole pressure (2.13) in the momentum conservation equation (2.2)

$$\mathbf{V}_m(\mathbf{r}|\mathbf{r}_0, \omega) = \frac{S(\omega)}{\rho c}\frac{1+jk|\mathbf{r}-\mathbf{r}_0|}{jk|\mathbf{r}-\mathbf{r}_0|}\frac{e^{-jk|\mathbf{r}-\mathbf{r}_0|}}{|\mathbf{r}-\mathbf{r}_0|}\mathbf{i}_r \tag{2.15}$$

with $\mathbf{i}_r$ the unit vector pointing in the direction of $\mathbf{r}$-$\mathbf{r}_0$.

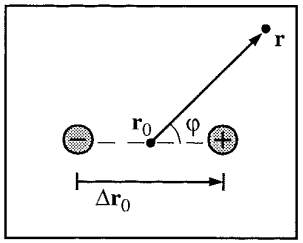


Figure 2.1: *Schematic representation of a dipole source.*

Another fundamental source is the dipole (see figure 2.1), consisting of two monopoles with opposite source strengths located at a short distance $|\Delta\mathbf{r}_0|$ ($|\Delta\mathbf{r}_0| \ll\lambda$) from each other. So the dipole field equals

$$P_d(\mathbf{r}|\mathbf{r}_0, \omega) = P_m(\mathbf{r}|\mathbf{r}_0 + \frac{\Delta\mathbf{r}_0}{2}, \omega) - P_m(\mathbf{r}|\mathbf{r}_0 - \frac{\Delta\mathbf{r}_0}{2}, \omega). \tag{2.16}$$

Expanding equation (2.16) in a Taylor series around $\mathbf{r}_0$ up to the first order (limit of small $\Delta\mathbf{r}_0$), yields the following dipole field

$$P_d(\mathbf{r}|\mathbf{r}_0, \omega) = \Delta\mathbf{r}_0 \cdot \nabla_0 P_m(\mathbf{r}|\mathbf{r}_0, \omega) = -\Delta\mathbf{r}_0 \cdot \nabla P_m(\mathbf{r}|\mathbf{r}_0, \omega) \tag{2.17}$$

Note that the gradient ∇_0 is taken with respect to the source coordinates. Evaluating eq. (2.17) by inserting the monopole wave field P_m yields

$$P_d(\mathbf{r}|\mathbf{r}_0, \omega) = S(\omega)|\Delta\mathbf{r}_0|\frac{1+jk|\mathbf{r}-\mathbf{r}_0|}{|\mathbf{r}-\mathbf{r}_0|}\cos\varphi\frac{e^{-jk|\mathbf{r}-\mathbf{r}_0|}}{|\mathbf{r}-\mathbf{r}_0|} \tag{2.18}$$

in which φ is the angle between the axis $\Delta\mathbf{r}_0$ of the dipole and the vector $\mathbf{r}$-$\mathbf{r}_0$. Taking the gradient on both sides of the inhomogeneous wave equation (2.10) with merely the monopole source term Q_m in the right hand side term, followed by the dot product with $-\Delta\mathbf{r}_0$ yields the dipole wave equation

$$\nabla^2 P_d(\mathbf{r}, \omega) + k^2 P_d(\mathbf{r}, \omega) = j\omega\rho_0\Delta\mathbf{r}_0 \cdot \nabla Q_m(\mathbf{r}, \omega) \tag{2.19}$$

By comparing the source term in eq. (2.19) to that in the inhomogeneous wave equation (2.10), inserting eq. (2.12) and using eq. (2.14), the dipole wave equation can be written as

$$\nabla^2 P_d(\mathbf{r}, \omega) + k^2 P_d(\mathbf{r}, \omega) = \nabla \cdot \mathbf{F}_d(\mathbf{r}, \omega) \tag{2.20}$$

in which

$$\mathbf{F}_d(\mathbf{r}, \omega) = 4\pi S(\omega)\delta(\mathbf{r}-\mathbf{r}_0)\Delta\mathbf{r}_0 \tag{2.21}$$

with $S(\omega)$ as the source spectrum. So a dipole is equivalent to a source of force, with the direction of the force along its axis.

2.2 The Kirchhoff-Helmholtz integral

In this section the Kirchhoff-Helmholtz integral will be derived, which forms the basis of the wave field synthesis concept. This integral is also called the representation theorem which gives a general solution for the acoustic wave field inside a given volume V. The geometrical situation for the following analysis is drawn in figure 2.2.

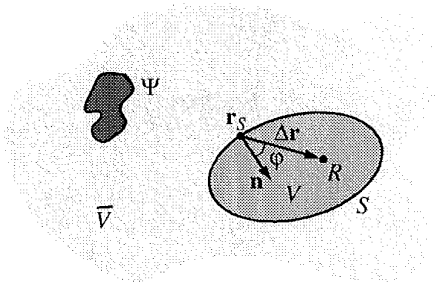


Figure 2.2: *Diagram of volume V enclosed by a surface S with a distribution of sources Ψ outside V.*

The volume V, enclosed by a surface S with inward pointing unit normal vector $\mathbf{n}$, lies inside an unbounded homogeneous medium. The primary source distribution Ψ is located inside volume $\bar{V}$, the spatial complement of V and S.

The wave field P inside V, due to the source distribution Ψ outside V, obeys the homogeneous wave equation

$$\nabla^2 P(\mathbf{r}, \omega) + k^2 P(\mathbf{r}, \omega) = 0 \tag{2.22}$$

Suppose that inside V a monopole source at point R is located with a frequency independent source spectrum $S(\omega)=1$. Inside V the pressure field G of this monopole given by

$$G(\mathbf{r}|\mathbf{r}_R, \omega) = \frac{e^{-jk|\mathbf{r}-\mathbf{r}_R|}}{|\mathbf{r}-\mathbf{r}_R|}, \tag{2.23}$$

which is the solution of the inhomogeneous wave equation

$$\nabla^2 G(\mathbf{r}|\mathbf{r}_R, \omega) + k^2 G(\mathbf{r}|\mathbf{r}_R, \omega) = -4\pi\delta(\mathbf{r}-\mathbf{r}_R) \tag{2.24}$$

under free field boundary conditions. Such a solution G of eq. (2.24) with an elementary source term is called a Green's function. The Green's function remains unchanged if the positions of the source and observer are reversed, as can be seen in eq. (2.23).

$$G(\mathbf{r}|\mathbf{r}_R, \omega) = G(\mathbf{r}_R|\mathbf{r}, \omega). \tag{2.25}$$

This reciprocity principle is an important property of the Green's function, and will be used in the following derivation.
Multiplication of eq. (2.22) by $G(\mathbf{r}|\mathbf{r}_R)$, eq. (2.24) by $P(\mathbf{r})$, and subtracting the results yields

$$G(\mathbf{r}|\mathbf{r}_R)\nabla^2 P(\mathbf{r}) - P(\mathbf{r})\nabla^2 G(\mathbf{r}|\mathbf{r}_R) = 4\pi\delta(\mathbf{r}-\mathbf{r}_R)P(\mathbf{r}). \tag{2.26}$$

From this point we omit the indication of the angular frequency ω for notational convenience. Integration of eq. (2.26) over volume V gives

$$\frac{1}{4\pi}\int_V [G(\mathbf{r}|\mathbf{r}_R)\nabla^2 P(\mathbf{r}) - P(\mathbf{r})\nabla^2 G(\mathbf{r}|\mathbf{r}_R)]dV = \begin{cases} P(\mathbf{r}_R) & \text{for } \mathbf{r}_R \in V \\ \frac{1}{2}P(\mathbf{r}_R) & \text{for } \mathbf{r}_R \text{ on } S \\ 0 & \text{for } \mathbf{r}_R \in \bar{V} \end{cases} \tag{2.27}$$

Application of the second theorem of Green with $\mathbf{n}$ the inward pointing normal vector

$$\int_V (G\nabla^2 P - P\nabla^2 G)dV = -\oint_S (G\nabla_S P - P\nabla_S G)\cdot\mathbf{n}dS \tag{2.28}$$

and the reciprocity relation (2.25) results in the so called Kirchhoff-Helmholtz integral

$$P(\mathbf{r}_R) = -\frac{1}{4\pi}\oint_S \left[G(\mathbf{r}_R|\mathbf{r}_S)\nabla_S P(\mathbf{r}_S) - P(\mathbf{r}_S)\nabla_S G(\mathbf{r}_R|\mathbf{r}_S) \right]\cdot\mathbf{n}dS \qquad \text{for } \mathbf{r}_R \in V, \tag{2.29}$$

in which $\mathbf{r}_S$ denotes a point on the closed surface S. Further evaluation of eq. (2.29) by substituting the Green's function of eq. (2.23) and using the equation of conservation of momentum (2.2) yields

$$P(\mathbf{r}_R) = \frac{1}{4\pi}\oint_S \left[j\omega\rho V_n(\mathbf{r}_S)\frac{e^{-jk\Delta r}}{\Delta r} \right]dS + \frac{1}{4\pi}\oint_S \left[P(\mathbf{r}_S)\frac{1+jk\Delta r}{\Delta r}\cos\varphi\frac{e^{-jk\Delta r}}{\Delta r} \right]dS \qquad \text{for } \mathbf{r}_R \in V \tag{2.30}$$

in which $\Delta r=|\Delta\mathbf{r}|=|\mathbf{r}_R-\mathbf{r}_S|$ is the distance from a secondary source point $\mathbf{r}_S$ at the surface S to the reconstruction point $\mathbf{r}_R$ inside V, and φ the angle between the vectors $\Delta\mathbf{r}$ and $\mathbf{n}$. The Kirchhoff-Helmholtz integral (2.30) states that any wave field P due to sources outside V can be reconstructed inside V by means of a continuous distribution of monopoles and dipoles at the surface S. The axes of the dipoles coincide with the inward pointing unit normal vector $\mathbf{n}$. The strength per unit area of the monopole source layer is proportional to the normal component of the particle velocity $V_n(\mathbf{r}_S)$ of the incident wave field, while the force strength per unit area of the dipole source layer is proportional to the pressure $P(\mathbf{r}_S)$ of the incident wave field. It should be realized, that the secondary source layer of monopoles and dipoles on S gives a zero contribution inside the primary source region $\bar{V}$.
So far the monopole and dipole sources enclose the reconstruction volume V. However it is also possible to turn the surface S 'inside out' in order to enclose the primary source distribution Ψ. The configuration for this situation is shown in figure 2.3. Again, the unit normal vector $\mathbf{n}$ is

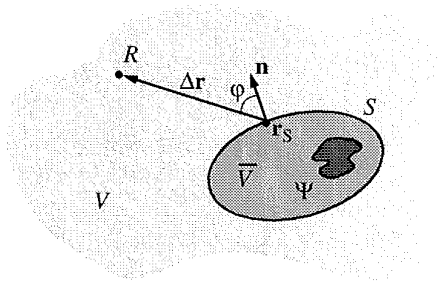


Figure 2.3: Diagram of volume V enclosed by a surface S with a distribution of sources Ψ inside $\bar{V}$.

pointing inward the reconstruction volume V. The Kirchhoff-Helmholtz integral for this situation is also given by eq. (2.30).

2.3 The Rayleigh integrals

In order to simplify the Kirchhoff-Helmholtz integral it should be noted that the choice of the Green's function is not unique. Since the Green's function G has to obey eq. (2.24) inside V

only, any convenient boundary condition on (parts of) S may be imposed. It would be interesting to find Green's functions such that either

$$\nabla_S G(\mathbf{r}_S|\mathbf{r}_R, \omega) \cdot \mathbf{n} = 0 \qquad \mathbf{r}_S \text{ on } S \tag{2.31}$$

or

$$G(\mathbf{r}_S|\mathbf{r}_R, \omega) = 0 \qquad \mathbf{r}_S \text{ on } S \tag{2.32}$$

is fulfilled. Condition (2.31) means that S acts as a perfectly reflecting rigid surface, which in general yields a complicated expression for the Green's function G. Under condition (2.32) S behaves like a perfectly reflecting free surface, which may yield a complicated expression for $\nabla G \cdot \mathbf{n}$. For the special case that S is a plane surface simple expressions for G and $\nabla G \cdot \mathbf{n}$ can be found. Consider the geometry of figure 2.4.

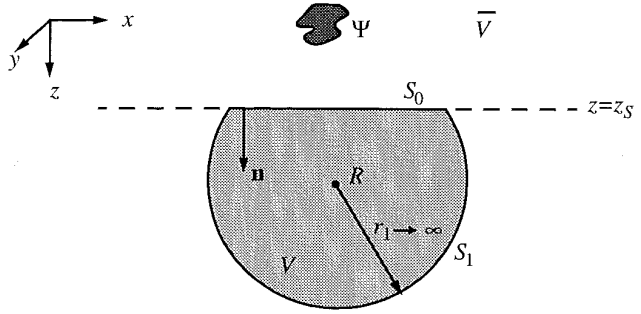


Figure 2.4: Diagram of volume V enclosed by a plane surface S_0 at $z=z_S$ and a spherical surface S_1 with a radius r_1. The distribution of sources Ψ lies in the upper half- space $\overline{V}$.

The reconstruction volume V is enclosed by a plane surface S_0 and a spherical surface S_1 with radius r_1. The sources of the acoustical pressure field P are situated in the upper half-space ($z<z_S$). Letting the radius r_1 tend to infinity and applying Sommerfeld's radiation condition (2.11a,b) to boundary S_1, results in a zero contribution of the Kirchhoff-Helmholtz over S_1 to the pressure at R (Wapenaar, 1989).
Now, the geometry of figure 2.4 can be modified into the configuration of figure 2.5. The primary source volume ($z<z_S$) and the reconstruction volume ($z>z_S$) are separated by an infinite secondary source plane at $z=z_S$.
Now, a Green's function has to be found for this geometry such that either condition (2.31) or condition (2.32) is fulfilled. Since the surface S_0 acts like a perfect reflector in both cases the Green's function may be interpreted, in volume V, as being the wave field of two monopole sources situated symmetrically with respect to S_0. In order to satisfy the rigid boundary condition (2.31) these two monopoles should have the same polarity. In this case the Kirchhoff-Helmholtz integral (2.29) may be replaced by

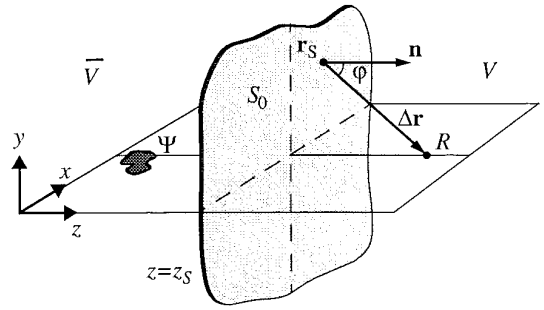


Figure 2.5: Diagram of a primary source distribution in the half space $z<z_S$ and an infinite secondary source plane S_0 at $z=z_S$.

$$P(\mathbf{r}_R) = -\frac{1}{4\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\Big[G_I(\mathbf{r}_R|\mathbf{r}_S)\nabla_S P(\mathbf{r}_S)\Big]\cdot\mathbf{n}\,dx_S dy_S \qquad \text{for } \mathbf{r}_R \in V \ . \tag{2.33}$$

The Green's function G_I for this situation equals twice the free space solution (2.23) for G at S_0. Using the equation of conservation of momentum (2.2) integral (2.33) can be written as

$$P(\mathbf{r}_R) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\Big[j\omega\rho V_z(\mathbf{r}_S)\frac{e^{-jk\Delta r}}{\Delta r}\Big]dx_S dy_S \qquad \text{for } \mathbf{r}_R \in V \tag{2.34}$$

which is called the Rayleigh I representation integral. It states that any wave field due to sources in the half space $\overline{V}$ ($z<z_S$) can be reconstructed in half space V ($z>z_S$) by means of a continuous distribution of monopoles at the surface S_0. The strength of each monopole is proportional to the normal component (positive z-direction) of the particle velocity $V_z(\mathbf{r}_S)$ of the incident wave field measured at the position of that monopole.
In order to satisfy the free boundary condition (2.32) the two monopoles should have opposite polarity. In this case the Kirchhoff-Helmholtz integral (2.29) may be replaced by

$$P(\mathbf{r}_R) = -\frac{1}{4\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\Big[P(\mathbf{r}_S)\nabla_S G_{II}(\mathbf{r}_R|\mathbf{r}_S)\Big]\cdot\mathbf{n}\,dx_S dy_S \qquad \text{for } \mathbf{r}_R \in V \tag{2.35}$$

The function $\nabla_S G_{II}$ for this situation equals twice the free space solution for $\nabla_S G$ at S_0. Integral (2.35) can be written as

$$P(\mathbf{r}_R) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\Big[P(\mathbf{r}_S)\frac{1+jk\Delta r}{\Delta r}\cos\varphi\frac{e^{-jk\Delta r}}{\Delta r}\Big]dx_S dy_S \qquad \text{for } \mathbf{r}_R \in V \tag{2.36}$$

which is called the Rayleigh II representation integral. It states that any wave field due to sources in the half $\bar{V}$ space ($z<z_S$) can be reconstructed in half space V ($z>z_S$) by means of a continuous distribution of dipole sources at the surface S_0, with their axes parallel to the normal vector **n**. The strength of each dipole is given by the pressure $P(\mathbf{r}_S)$ of the incident wave field measured at the position of that dipole.
In contrast to the Kirchhoff-Helmholtz integral the Rayleigh integrals do not yield a null value for points inside the domain of the primary sources. The wave field in the source domain may be interpreted as being the reflection of the primary source wave field at the plane surface S_0. The monopole distribution in the Rayleigh I integral acts like a perfect reflector with R=+1, while the dipole distribution in Rayleigh II acts like a reflector with R=-1.

2.4 2D wave field reconstruction

In section 2.2 it was shown that any wave field P can be reconstructed in a 3 dimensional sourceless domain by enclosing the source domain (figure 2.2) or the receiver domain (figure 2.3) with a continuous distribution of secondary sources, according to the Kirchhoff-Helmholtz integral. A degenerated case of such a closed surface is an infinite plane surface between the source and the receiver domain (figure 2.5), in which case the Kirchhoff-Helmholtz integral can be simplified, yielding the Rayleigh integrals. So, in order to reconstruct a primary wave field P in 3 dimensions planar distributions of secondary sources are required.
The 3D versions of the Kirchhoff and Rayleigh integrals over a (closed) surface can be transformed to integrals over a (closed) line in 2 dimensions, which will be shown in the next sections.

2.4.1 The 2D Kirchhoff-Helmholtz integral

Consider the special 3D geometry in figure 2.6. The reconstruction volume V is enclosed by a cylindrical surface S_0 perpendicular to the plane y=0 and two plane surfaces S_1 and S_{-1} at $y=y_1$ and $y=-y_1$ respectively.
The sources of wave field P are situated outside V in the plane y=0. The closed contour L is the intersection of surface S_0 and plane y=0. The contribution of the Kirchhoff-Helmholtz integral over S_1 and S_{-1} to the pressure at R vanishes if y_1 goes to infinity (Sommerfeld's radiation condition), yielding

$$P(\mathbf{r}_R) = \frac{1}{4\pi}\int_{S_0}\left[j\omega\rho V_n(\mathbf{r}_S)\frac{e^{-jk\Delta r}}{\Delta r}\right]dS_0 + \frac{1}{4\pi}\int_{S_0}\left[P(\mathbf{r}_S)\frac{1+jk\Delta r}{\Delta r}\cos\varphi\frac{e^{-jk\Delta r}}{\Delta r}\right]dS_0 \quad \text{for } \mathbf{r}_R \in V \qquad (2.37)$$

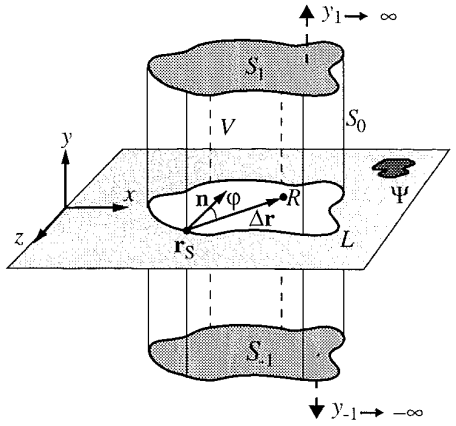


Figure 2.6: *Diagram of volume V enclosed by a cylindrical surface S_0 and two plane surfaces S_1 and S_{-1} at $y=y_1$ and $y=-y_1$ respectively. The distribution of sources Ψ lies outside V in the plane y=0.*

For the 2D version of the Kirchhoff-Helmholtz integral it must be assumed that P and V_n are independent of y. This means that the primary source distribution Ψ is replaced by a primary line source distribution. By integration over the y-axis, the secondary point sources on S_0 for the 3D case can be transformed into secondary line sources perpendicular to the xz-plane on L for the 2D case. The wave field of a secondary monopole line source is given by

$$P_m^{\text{line}} = -j\pi H_0^{(2)}(k\Delta r), \qquad (2.38a)$$

the wave field of a secondary dipole line source by

$$P_d^{\text{line}} = -jk\pi\cos\varphi H_1^{(2)}(k\Delta r) \qquad (2.38b)$$

The functions $H_0^{(2)}$ and $H_1^{(2)}$ represent the zeroth-order and first-order Hankel functions of the second kind respectively. Now the Kirchhoff-Helmholtz integral can be written as

$$P(\mathbf{r}_R) = \frac{1}{4\pi}\oint_L[j\omega\rho V_n(\mathbf{r}_L)\{-j\pi H_0^{(2)}(k\Delta r)\}]dL + \frac{1}{4\pi}\oint_L[P(\mathbf{r}_L)\{-jk\pi\cos\varphi H_1^{(2)}(k\Delta r)\}]dL \quad \text{for } \mathbf{r}_R \in D \qquad (2.39)$$

in which the surface D is the area enclosed by contour L, as depicted in the 2D configuration of figure 2.7.
Expressions (2.38a and b) can be approximated by an exponential function for $k\Delta r \gg 1$ (Abramowitz et al.,1965), giving

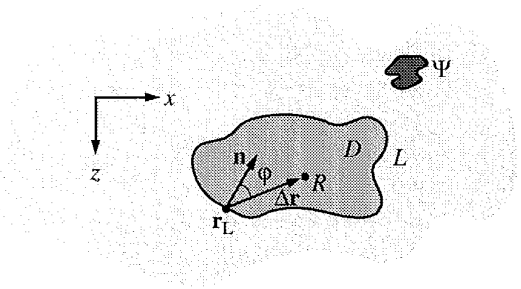


Figure 2.7: 2D diagram of surface D enclosed by contour L with inward pointing normal vector **n**. *The distribution of sources Ψ lies outside S.*

$$P_{\mathrm{m}}^{\mathrm{line}} = \sqrt{\frac{2\pi}{jk}}\,\frac{e^{-jk\Delta r}}{\sqrt{\Delta r}} \tag{2.40a}$$

$$P_{\mathrm{d}}^{\mathrm{line}} = \sqrt{2\pi jk}\,\cos\varphi\,\frac{e^{-jk\Delta r}}{\sqrt{\Delta r}} \tag{2.40b}$$

which are the far field expressions for the pressure field of a monopole and a dipole line source respectively. Using these approximations eq. (2.39) can be written as

$$P(\mathbf{r}_R) = \frac{1}{4\pi}\oint_L \left[\mathrm{j}\omega\rho V_{\mathrm{n}}(\mathbf{r}_L)\sqrt{\frac{2\pi}{jk}}\frac{e^{-jk\Delta r}}{\sqrt{\Delta r}}\right] dL \quad \text{for} \quad \mathbf{r}_R \in D \tag{2.41}$$
$$+ \frac{1}{4\pi}\oint_L \left[P(\mathbf{r}_L)\sqrt{2\pi jk}\,\cos\varphi\,\frac{e^{-jk\Delta r}}{\sqrt{\Delta r}}\right] dL$$

which is the far field approximation of the 2D Kirchhoff-Helmholtz integral.

2.4.2 The 2D Rayleigh integrals

The 2D versions of the Rayleigh integrals can be derived in the same way as the 2D Kirchhoff-Helmholtz integral. Consider again the 3D geometry in figure 2.5. The 2D Rayleigh I integral can be obtained by integration of eq. (2.34) over the y-axis, replacing the primary source by a primary line source such that the integrand V_{n} is independent of y, yielding

$$P(\mathbf{r}_R) = \frac{1}{2\pi}\int_{-\infty}^{\infty} [j\omega\rho V_{\mathrm{n}}(x_L)\{-j\pi H_0^{(2)}(k\Delta r)\}]dx_L \quad \text{for} \quad z_R > z_L \ . \tag{2.42}$$

The far field approximation ($k\Delta r \gg 1$) of eq. (2.42) reads

$$P(\mathbf{r}_R) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \left[\mathrm{j}\omega\rho V_{\mathrm{n}}(\mathbf{r}_L)\sqrt{\frac{2\pi}{jk}}\frac{e^{-jk\Delta r}}{\sqrt{\Delta r}}\right] dx_L \quad \text{for} \quad z_R > z_L \ . \tag{2.43}$$

Similarly, the Rayleigh II integral (2.36) can be transformed to a line integral by applying the same integration procedure, giving

$$P(\mathbf{r}_R) = \frac{1}{2\pi}\int_{-\infty}^{\infty} [P(x_L)\{-jk\pi\cos\varphi H_1^{(2)}(k\Delta r)\}]dx_L \quad \text{for} \quad z_R > z_L \ . \tag{2.44}$$

The far field approximation ($k\Delta r \gg 1$) of eq. (2.44) is given by

$$P(\mathbf{r}_R) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \left[P(\mathbf{r}_L)\sqrt{2\pi jk}\,\cos\varphi\,\frac{e^{-jk\Delta r}}{\sqrt{\Delta r}}\right] dx_L \quad \text{for} \quad z_R > z_L \ . \tag{2.45}$$

The configuration for the 2D Rayleigh integrals is depicted in figure 2.8, in which the infinite

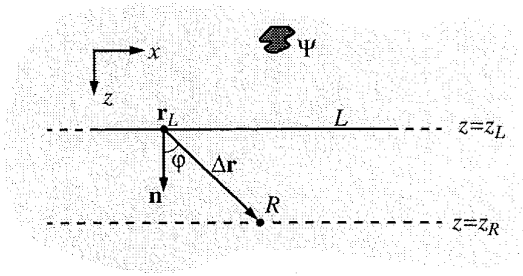


Figure 2.8: 2D diagram of a primary source distribution Ψ in the half plane $z<z_L$ and reconstruction point R at $z=z_R$ in the half plane $z_R>z_L$.

secondary source plane has S_0 been replaced by an infinite secondary source line L at $z=z_L$.

2.5 Examples of 2D wave field synthesis

In the previous sections the principles of wave field reconstruction have been presented. It has been pointed out that an arbitrary wave field P due to a source distribution Ψ can be reconstructed by means of a secondary source distribution consisting of monopoles and/or dipoles. However, the source distribution Ψ does not need to be physically present in order to determine the pressure P and the normal component of the velocity V_{n} at the secondary source positions. The required secondary source strengths can also be calculated by pretending that a certain desired notional source distribution is present. In this way the wave field of any notional source (distribution) can be synthesized using a proper secondary source distribution.

As an example the wave field of a monopole line source will be synthesized by two different secondary source distributions. In the first example the wave field will be synthesized by a monopole and dipole line source distribution on a contour line. In the second example the same wave field will be synthesized by means of a monopole line source distribution along an infinite straight line parallel to the x-axis.

2.5.1 Reconstruction of a monopole wave field with the 2D Kirchhoff-helmholtz integral

As an example consider the 2D situation depicted in figure 2.9. A notional monopole line

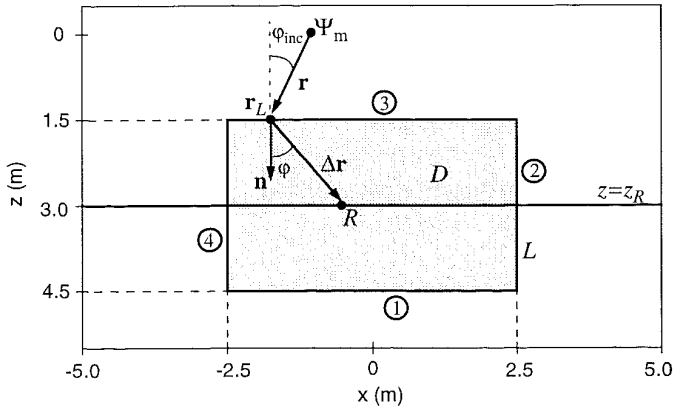


Figure 2.9: Diagram of the geometry of a notional monopole line source Ψ_m at $\mathbf{r}_\Psi$=(-1,0) and a secondary monopole and dipole line source distribution along the 4 sides of rectangle L, numbered counter-clockwise from 1 to 4. The primary and the synthesized wave field are recorded at the line $z=z_R$.

source at position $\mathbf{r}_\Psi$=(-1,0) emits a short pulse $s(t)$ starting at t=0. The primary wave field p, which is the wave field we want to synthesize, recorded at the line $z=z_R$ is shown in the x-t diagram of figure 2.10a.

Next, it is attempted to synthesize the same wave field p by a secondary monopole and dipole line source distribution along the 4 sides of the indicated rectangle, forming the contour L. According to the 2D version of the Kirchhoff-Helmholtz integral (2.39) the wave field p can be reconstructed in area D enclosed by contour L if the pressure P^{line} and normal velocity V_n^{line} of the notional monopole line source are known at contour L. In the frequency domain, the pressure P^{line} at L is given by eq. (2.38a)

$$P^{\text{line}}(r_L) = -j\pi S(\omega) H_0^{(2)}(kr) \tag{2.46}$$

while the normal component of the particle velocity V_n^{line} can be found by applying the equation of momentum conservation (2.2) to eq. (2.46)

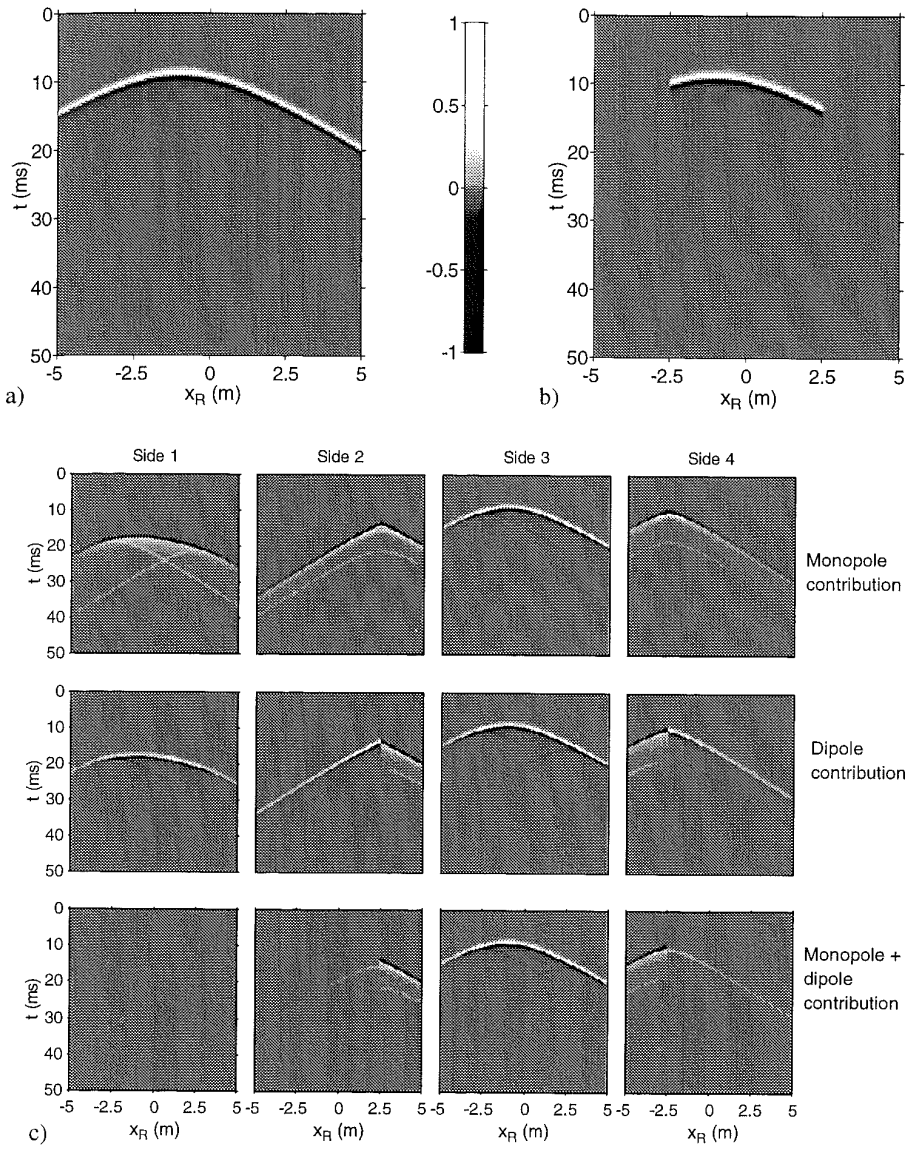


Figure 2.10: Synthesized wave field p. Registration along the line $z=z_R$ for receiver positions $-5.0<x_R<5.0$ m.
a) Primary wave field
b) Synthesized wave field
c) Individual contributions of the secondary monopole and dipole sources

$$V_n^{line}(r_L) = -\frac{\pi}{\rho c} S(\omega) \cos\varphi_{inc} H_1^{(2)}(kr) \qquad (2.47)$$

with $r=|\mathbf{r}|=|\mathbf{r}_L-\mathbf{r}_\Psi|$ the distance from the notional source to a secondary source at the line L, and φ_{inc} the angle between the vectors $\mathbf{r}$ and $\mathbf{n}$.

Figure 2.10b shows that indeed the synthesized wave field is in perfect agreement with the primary wave field for positions inside D ($|x_R|<2.5$ m.). Outside D ($|x_R|>2.5$ m.) the secondary sources give a zero contribution to the total wave field.

It is interesting to look at the individual contributions of the secondary monopoles and dipoles at the recording line for each side of the rectangle L. This is shown in figure 2.10c. Note that the main contribution emerges from side 3. The contribution of side 1is negligible because the wave field of the monopoles and the dipoles are almost in counter phase. Sides 2 and 4 merely give a contribution outside the area D in order to neutralize the wave field produced by side 3.

2.5.2 Reconstruction of a monopole wave field with the 2D Rayleigh I integral

As a second example consider the 2D situation of figure 2.11. The closed contour L of figure 2.9 has been replaced by an infinite long line parallel to the x-axis. The recording line $z=z_R$ and

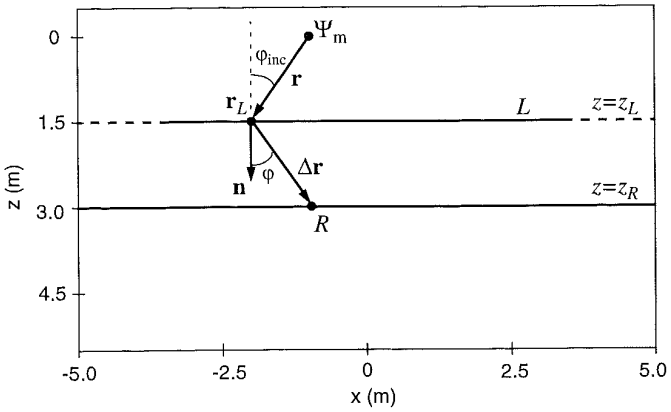


Figure 2.11: Diagram of the geometry of a notional monopole line source Ψ_m at $\mathbf{r}_\Psi=(-1,0)$ and a secondary monopole line source distribution at an infinite line L. The primary and the synthesized wave field are recorded at the line $z=z_R$.

the notional monopole line source retain their positions. In this case the wave field of the notional source is synthesized by a secondary monopole line source distribution at the line L. According to the 2D version of the Rayleigh I integral (2.42) the wave field p can be reconstructed in the area $z>z_L$. The primary wave field, which equals the primary wave field depicted in figure 2.10a, and the synthesized wave field are shown in figure 2.12a and b. Indeed the synthesized wavefield matches perfectly the original wave field for all x-positions.

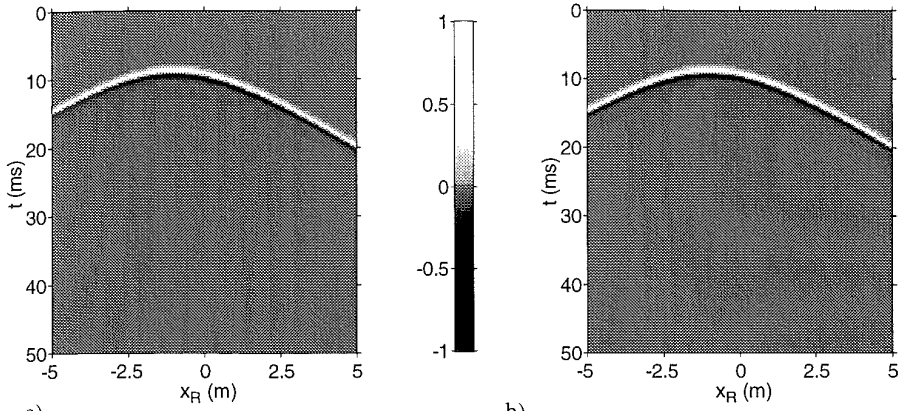


Figure 2.12: Wave field p at the receiver line $z=z_R$ for receiver positions $-5.0<x_R<5.0$ m.
a) Primary wave field
b) Synthesized wave field

3

Synthesis operators for line arrays in a 3D space

In the previous chapter the fundamental principles of wave field synthesis have been addressed. It has been demonstrated that the Kirchhoff-Helmholtz integral leads to the possibility of wave field synthesis. It has been shown that according to the 3D Rayleigh and Kirchhoff-Helmholtz integrals a planar distribution of secondary sources between the source area and the receiver area is required in order to synthesize an arbitrary 3D wave field. This would imply that the source area is visually hidden from the receiver area, which is absolutely unacceptable.

In this chapter we will derive synthesis operators which make it possible to synthesize a 3D wave field with a horizontal line array consisting of secondary *point* sources instead of vertical *line* sources, as in the 2D case. These so-called '2½D operators' will form the basis of the practical implementation of many wave field synthesis applications, such as direct sound enhancement, audio reproduction (Boone et al., 1995) and acoustical control systems (de Vries et al., 1994).

Vogel (1993) derived synthesis operators for straight line arrays. However, the method Vogel used is inappropriate for generalization towards bent line arrays, which have many advantages in practice. Therefore, in this chapter we follow a different approach which leads to synthesis operators for various array geometries.

Using a horizontal line array instead of a planar array, the wave fronts are modified in the direction perpendicular to the line array. It will be shown that in practical applications a horizontal line array will satisfy our requirements for direct sound enhancement as well for other applications.

In section 3.1 and 3.2 of this chapter the 2½D Rayleigh operators are derived, followed in section 3.3 by the derivation of the 2½D Kirchhoff-Helmholtz integral. Departing from this 2½D Kirchhoff-Helmholtz operator the 'generalized' Rayleigh operators will be derived in section

3.5, which make it possible to reconstruct a wave field with only secondary monopole or dipole sources at a bent line array. Finally, in section 3.6 the effect of a difference in height between the array and the receivers on the apparent source position will be addressed.

3.1 The 2½D Rayleigh I integral

Again consider the 3D Rayleigh configuration of figure 3.1. Both the notional monopole source

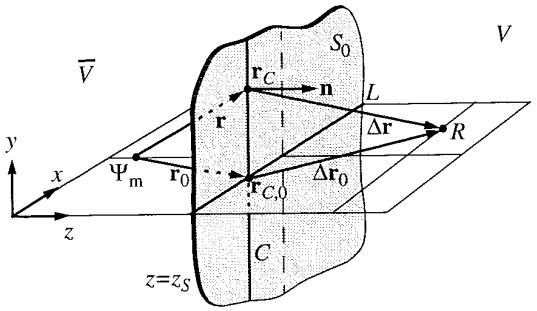


Figure 3.1: Diagram of a notional monopole source Ψ_m at position $\mathbf{r}=\mathbf{r}_\Psi$ in the plane $y=0$ and an infinite secondary source plane at $z=z_S$. The line L is the cross section of the planes S_0 and $y=0$. The contribution in R of a secondary source column C through $\mathbf{r}_{C,0}$ perpendicular to the plane $y=0$ can be approximated by a contribution of point $\mathbf{r}_{C,0}$ only.

Ψ_m at position $\mathbf{r}_\Psi$ and the receiver R at position $\mathbf{r}_R$ are positioned in the horizontal plane $y=0$. The wave field of the notional monopole source is synthesized by a secondary monopole source distribution on the plane S_0. If we look at the contribution of the secondary sources on the vertical line C trough $\mathbf{r}_{C,0}$, it should be realized that the points in the neighborhood of $\mathbf{r}_{C,0}$ give the largest contribution to the total wave field at point R. It will be shown that the contribution of the secondary source column C trough $\mathbf{r}_{C,0}$ can be approximated by the contribution of one secondary monopole source in point $\mathbf{r}_{C,0}$ only. This approximation can be applied to all secondary source columns in the plane S_0 yielding a secondary source distribution on the horizontal line L only.

By substitution of expression (2.15) for the particle velocity of a monopole, the Rayleigh I synthesis integral (2.34) over the surface S_0 can be written as an integration over all column contributions P_C, yielding

$$P(\mathbf{r}_R) = \frac{1}{2\pi}\int_{-\infty}^{\infty} P_C(\mathbf{r}_R)dx_C, \qquad \text{with} \tag{3.1a}$$

$$P_C(\mathbf{r}_R) = \int_{-\infty}^{\infty} S(\omega)\left(\frac{1+jkr}{r}\right)\cos\varphi_{inc}\frac{e^{-jkr}}{r}\frac{e^{-jk\Delta r}}{\Delta r}dy_C \tag{3.1b}$$

in which the coordinate x_C denotes the x-position of the secondary source column C and $r=|\mathbf{r}|=|\mathbf{r}_C-\mathbf{r}_\Psi|$, $\Delta r=|\Delta\mathbf{r}|=|\mathbf{r}_R-\mathbf{r}_C|$ with $\mathbf{r}_C=(x_C, y_C, z_S)^T$ a point on column C. The angle φ_{inc} is the angle between the vectors $\mathbf{r}$ and $\mathbf{n}$.

Next, a mathematical concept called the 'stationary phase approximation' is introduced (Bleistein, 1984), which states that an integral I of the form

$$I = \int_{-\infty}^{\infty} f(y_C)e^{j\phi(y_C)}dy_C \tag{3.2}$$

can, for large values of ϕ, be approximated by

$$I \approx \sqrt{\frac{2\pi}{|\phi''(y_{C,0})|}}f(y_{C,0})e^{j\phi(y_{C,0})+j\,\mathrm{sgn}[\phi''(y_{C,0})]\frac{\pi}{4}} \tag{3.3}$$

in which y_0 is a simple stationary point of the phase function $\phi(y)$, so that

$$\phi'(y_{C,0}) = 0, \qquad \text{while} \qquad \phi''(y_{C,0}) \neq 0. \tag{3.4}$$

We see that, with this approximation, the result of integral (3.2) does not depend on global properties of the integrand, but only on local properties, namely, f and ϕ and its derivatives near the stationary point $y_{C,0}$. It means that the stationary point and its immediate surroundings give the major contribution to the integral.

Expression of the integral I in eq. (3.2) in terms of eq. (3.1b) yields for the complex amplitude $f(y)$

$$f(y_C) = S(\omega)\left(\frac{1+jkr}{r^2\Delta r}\right)\cos\varphi_{inc} \tag{3.5}$$

which for $kr>>1$ (far field or high frequency approximation) can be written as

$$f(y_C) = S(\omega)\left(\frac{jk}{r\Delta r}\right)\cos\varphi_{inc} \tag{3.6}$$

The phase function ϕ and its derivatives, for this situation, are given by

$$\phi(y_C) = -k(r+\Delta r), \tag{3.7a}$$

$$\phi'(y_C) = -k\left(\frac{y_C}{r}+\frac{y_C}{\Delta r}\right), \tag{3.7b}$$

$$\phi''(y_C) = -k\left(\frac{(x_C-x_\Psi)^2+(z_S-z_\Psi)^2}{r^3}+\frac{(x_R-x_C)^2+(z_R-z_S)^2}{\Delta r^3}\right) \tag{3.7c}$$

The y-coordinate $y_{C,0}$ of the stationary point $\mathbf{r}_{C,0}$ can be found by substituting eq. (3.7b) into eq. (3.4) and solving for y_C, yielding: $y_{C,0}=0$. Consequently, the stationary point $\mathbf{r}_{C,0}$ of the secondary source column at x_C is found at the intersection point of that secondary source column with the horizontal plane $y=0$, yielding $\mathbf{r}_{C,0}=(x_C, 0, z_S)^T$. This result agrees with what we expected. Substitution of the value $y_{C,0}$ in eqs.(3.6), (3.7a) and (3.7c) gives

$$f(y_{C,0}) = S(\omega)\left(\frac{jk}{r_0 \Delta r_0}\right)\cos\varphi_{0,\,\mathrm{inc}} \tag{3.8a}$$

$$\phi(y_{C,0}) = -k(r_0 + \Delta r_0), \tag{3.8b}$$

$$\phi''(y_{C,0}) = -k\left(\frac{1}{r_0} + \frac{1}{\Delta r_0}\right) \tag{3.8c}$$

in which $r_0=|\mathbf{r}_0|=|\mathbf{r}_{C,0}-\mathbf{r}_\Psi|$, $\Delta r_0=|\Delta\mathbf{r}_0|=|\mathbf{r}_R-\mathbf{r}_{C,0}|$ and $\varphi_{\mathrm{inc},0}$ is the angle between the vectors $\mathbf{n}$ and $\mathbf{r}_{C,0}$.

Application of the stationary phase approximation (3.3) to the integral representation of P_C given by eq. (3.1b) and substitution of eqs. (3.8a-c) yields for the contribution of a secondary source column C at position x_C to the synthesized wave field

$$P_C(\mathbf{r}_R) = S(\omega)\sqrt{2\pi jk}\sqrt{\frac{\Delta r_0}{r_0 + \Delta r_0}}\cos\varphi_{\mathrm{inc},\,0}\frac{e^{-jkr_0}}{\sqrt{r_0}}\frac{e^{-jk\Delta r_0}}{\Delta r_0} \tag{3.9}$$

Now, the total synthesis integral can be built up by integrating the stationary phase contributions along the line L. Omitting the subscripts of Δr and r, which all lie in the plane $y=0$, we can write the following synthesis integral:

$$P(\mathbf{r}_R) = \int_{-\infty}^{\infty}\sqrt{\frac{jk}{2\pi}}S(\omega)g(\mathbf{r}_R, \mathbf{r}_L)\cos\varphi_{\mathrm{inc}}\frac{e^{-jkr}}{\sqrt{r}}\frac{e^{-jk\Delta r}}{\Delta r}dx_L. \tag{3.10}$$

The amplitude factor g is given by

$$g(\mathbf{r}_R, \mathbf{r}_L) = \sqrt{\frac{\Delta r}{r + \Delta r}}, \tag{3.11}$$

where $r=|\mathbf{r}|=|\mathbf{r}_L-\mathbf{r}_\Psi|$, $\Delta r=|\Delta\mathbf{r}|=|\mathbf{r}_R-\mathbf{r}_L|$ with $\mathbf{r}_L=(x_L,0,z_L)$ a point on the line L. The angle φ_{inc} is the angle between the vectors $\mathbf{r}$ and $\mathbf{n}$.

The 3D configuration of figure 3.1 is now transformed to the configuration depicted in figure 3.2, since eq. (3.9) shows that the contribution to the total wave field of a secondary source column C through the line L can be replaced by the contribution of one single secondary monopole source at $\mathbf{r}_L$.

Because the amplitude factor g is a function of the secondary source-to-receiver distance Δr, the correct driving signal for each secondary monopole depends on the receiver position $\mathbf{r}_R$, which is impossible to realize.

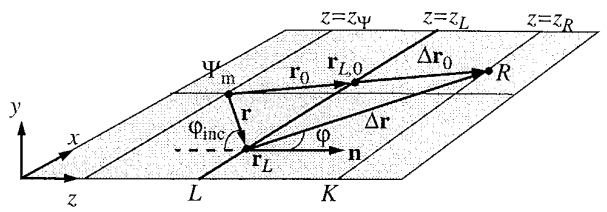


Figure 3.2: *Diagram of the geometry of a notional monopole source Ψ_m and a secondary monopole point source distribution at an infinite line L. The secondary point source at $\mathbf{r}_{L,0}$ gives the largest contribution to the wave field in point R at the line K.*

In order to find a practical function g, that is valid for more than one receiver position, we consider a receiver line K through point R parallel at the line L, as shown in figure 3.2. Next, we argue that the contribution to the total wave field at a certain receiver position is highly dominated by the contribution of the stationary phase point in the x-direction, on the line L. The stationary phase point $\mathbf{r}_{L,0}=(x_{L,0}, 0, z_L)^T$ on the line L for a certain receiver position is defined by

$$\phi'(x_{L,0}) = 0, \qquad \text{while} \qquad \phi''(x_{L,0}) \neq 0. \tag{3.12}$$

The phase function $\phi(x_L)$ of the integrand in eq.(3.10) and its derivatives are determined by

$$\phi(x_L) = -k(r + \Delta r), \tag{3.13a}$$

$$\phi'(x_L) = -k\left(\frac{x_L - x_\Psi}{r} - \frac{x_R - x_L}{\Delta r}\right), \tag{3.13b}$$

$$\phi''(x_L) = -k\left(\frac{(z_L - z_\Psi)^2}{r^3} + \frac{(z_R - z_L)^2}{\Delta r^3}\right) \tag{3.13c}$$

Substitution of eq. (3.13b) in eq. (3.12) yields the following implicit expression for the x-coordinate $x_{L,0}$ of the stationary point:

$$\frac{x_{L,0} - x_\Psi}{x_R - x_{L,0}} = \frac{r_0}{\Delta r_0} = \frac{z_L - z_\Psi}{z_R - z_\Psi}, \tag{3.14}$$

where $r_0=|\mathbf{r}_0|=|\mathbf{r}_{L,0}-\mathbf{r}_\Psi|$, $\Delta r_0=|\Delta\mathbf{r}_0|=|\mathbf{r}_R-\mathbf{r}_{L,0}|$.

Note that each receiver position on K defines a unique stationary point $\mathbf{r}_{L,0}$ on L, which is found at the intersection point of the line from the notional source Ψ_m to receiver R with the line L. And, vice versa: each secondary source at position $\mathbf{r}_L$ defines a unique 'stationary' receiver position $\mathbf{r}_{R,0}$ on K, which is determined by the intersection point of the line through $\mathbf{r}_\Psi$ and $\mathbf{r}_L$ and the receiver line K. In this way, we can express the 'stationary' receiver position as a function of the secondary source position $\mathbf{r}_L$.

When we approximate the secondary source-to-receiver distance Δr in the expression (3.11) with the 'stationary' secondary source-to-receiver distance $\Delta r_0(\mathbf{r}_L)$, given by

$$\Delta r_0(\mathbf{r}_L) = |\mathbf{r}_{R,0}(\mathbf{r}_L) - \mathbf{r}_L|, \tag{3.15}$$

the amplitude factor g can be written as a function of the secondary source position only, yielding

$$g_0(\mathbf{r}_L) = \sqrt{\frac{\Delta r_0(\mathbf{r}_L)}{r_0(\mathbf{r}_L) + \Delta r_0(\mathbf{r}_L)}} = \sqrt{\frac{z_R - z_L}{z_R - z_\Psi}} = \sqrt{\frac{(z_R - z_L)/(z_L - z_\Psi)}{1 + (z_R - z_L)/(z_L - z_\Psi)}} \tag{3.16}$$

with $\mathbf{r}_{R,0} \in K$.
For a certain receiver line K parallel to the secondary source line L, g_0 is a constant.
The synthesis integral of eq. (3.10) can now be written as

$$P(\mathbf{r}_R) = g_0 \int_{-\infty}^{\infty} \sqrt{\frac{jk}{2\pi}} S(\omega) \cos\varphi_{inc} \frac{e^{-jkr}}{\sqrt{r}} \frac{e^{-jk\Delta r}}{\Delta r} dx_L. \tag{3.17}$$

This is called the 2½D Rayleigh I integral, which is identical to the 2½D monopole operator, derived by Vogel (1993).
In the case that the receivers are positioned at different depth levels z_R, an average depth $\bar{z}_R$ has to be chosen, introducing only an amplitude error in the synthesized wave field for receivers on the lines $z_R \neq \bar{z}_R$. In figure 3.3 the dependence of the amplitude function g_0 on the ratio between

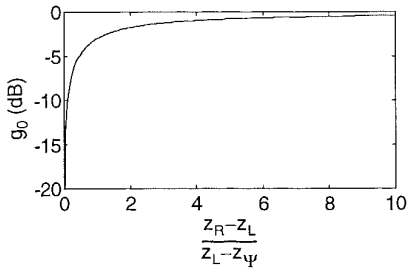


Figure 3.3: *Amplitude factor g_0 as a function of the ratio between the array-to-receiver distance and the notional source-to-array distance*

the array-to-receiver distance $(z_R\text{-}z_L)$ and the notional source-to-array distance $(z_L\text{-}z_\Psi)$ is shown. With this diagram the amplitude error for receivers on the lines $z_R \neq \bar{z}_R$ can be calculated. Since the amplitude function g_0 is calculated for an average depth $\bar{z}_R$, the amplitude for the deeper positioned receivers will be too low, while for the receivers closer to the array the amplitude will be to high.
For receiver positions far away from the secondary sources on line L in comparison with the distance between the notional source Ψ_m and the line L, the amplitude factor g_0 can be approximated by

$$g_0 \approx 1 \qquad \text{for} \qquad \frac{z_R - z_L}{z_L - z_\Psi} \gg 1, \tag{3.18}$$

and becomes independent of the depth position of the receivers.
Introducing the far field approximation of the normal component of the particle velocity V_z^{line} of a monopole line source at $\mathbf{r}_\Psi$,

$$V_z^{line}(x_L) = \frac{1}{\rho c} \sqrt{\frac{2\pi}{jk}} S(\omega) \cos\varphi_{inc} \frac{e^{-jkr}}{\sqrt{r}} \tag{3.19}$$

the 2½D Rayleigh I integral (3.17) can be rewritten as

$$P(\mathbf{r}_R) = \frac{1}{2\pi} \int_{-\infty}^{\infty} j\omega\rho V_z^{line}(x_L) \frac{e^{-jk\Delta r}}{\Delta r} dx_L \qquad \text{for} \qquad \frac{z_R - z_L}{z_L - z_\Psi} \gg 1 \tag{3.20}$$

while $kr \gg 1$. The source strength for each secondary monopole source is proportional to the normal component of the particle velocity V_z^{line} of a monopole line source at $\mathbf{r}_\Psi$.
For receiver positions close to the secondary sources on line L in comparison with the distance between the notional source Ψ_m and the line L, the amplitude factor g of eq. (3.11) can be approximated by

$$g(\mathbf{r}_R, \mathbf{r}_L) \approx \sqrt{\frac{\Delta r}{r}} \qquad \text{for} \qquad \frac{z_R - z_L}{z_L - z_\Psi} \ll 1. \tag{3.21}$$

Introducing the far field approximation of the normal component of the particle velocity V_z of a monopole point source at $\mathbf{r}_\Psi$,

$$V_z(x_L) = \frac{S(\omega)}{\rho c} \cos\varphi_{inc} \frac{e^{-jkr}}{r} \tag{3.22}$$

the 2½D Rayleigh I integral (3.17) can be written as

$$P(\mathbf{r}_R) = \frac{1}{2\pi} \int_{-\infty}^{\infty} j\omega\rho V_z(x_L) \sqrt{\frac{2\pi}{jk}} \frac{e^{-jk\Delta r}}{\sqrt{\Delta r}} dx_L \qquad \text{for} \qquad \frac{z_R - z_L}{z_L - z_\Psi} \ll 1 \tag{3.23}$$

while $k\Delta r \gg 1$. In this case it can be stated that for receiver positions close to the line L in comparison with the distance r, the wave field of a notional monopole source can be synthesized in the plane $y=0$, using secondary monopole line sources. The source strength for each monopole line source is proportional to the normal component of the particle velocity V_z of the notional monopole point source.

3.2 The 2½D Rayleigh II integral

Taking the 3D Rayleigh II integral (2.36) as a starting point, in which expression (2.13) for the pressure of the notional monopole source is inserted, and following the same procedure as in the previous section for the same configuration as considered before, the 2½D Rayleigh II integral can be found

$$P(\mathbf{r}_R) = g_0 \int_{-\infty}^{\infty} \sqrt{\frac{1}{2\pi jk}} S(\omega) \frac{e^{-jkr}}{\sqrt{r}} jk\cos\varphi \frac{e^{-jk\Delta r}}{\Delta r} dx_L \tag{3.24}$$

The amplitude factor g_0 is given by eq. (3.16).
For receiver positions far away from the secondary sources on line L in comparison with the distance between the notional source Ψ_m and the line L the amplitude factor g_0 is given by eq. (3.18), yielding

$$P(\mathbf{r}_R) = \frac{1}{2\pi} \int_{-\infty}^{\infty} P^{\text{line}}(x_L) jk\cos\varphi \frac{e^{-jk\Delta r}}{\Delta r} dx_L \qquad \text{for} \qquad \frac{z_R - z_L}{z_L - z_\Psi} \gg 1 \tag{3.25}$$

while $kr \gg 1$. The strengths of the secondary dipole point sources are given by P^{line}; the pressure of a notional monopole line source at $\mathbf{r}_\Psi$, given by eq. (2.40a).
For receiver positions close to the secondary sources on line L in comparison with the distance between the notional source Ψ_m and L, the amplitude factor g for the stationary point is given by eq. (3.21), yielding for the 2½D Rayleigh II integral

$$P(\mathbf{r}_R) = \frac{1}{2\pi} \int_{-\infty}^{\infty} P(x_L) \sqrt{2\pi jk}\cos\varphi \frac{e^{-jk\Delta r}}{\sqrt{\Delta r}} dx_L \qquad \text{for} \qquad \frac{z_R - z_L}{z_L - z_\Psi} \ll 1 \tag{3.26}$$

while $k\Delta r \gg 1$. In this case it can be stated that for receiver positions close to the line L in comparison with the distance r, the wave field of a notional monopole source can be synthesized in the plane $y=0$, using secondary dipole line sources (eq. (2.40b)). The source strength for each dipole line source is given by the pressure P of the notional monopole point source.

3.3 The 2½D Kirchhoff-Helmholtz integral

In order to derive the 2½D Kirchhoff-Helmholtz equation consider the special 3D Kirchhoff-Helmholtz geometry of figure 3.4.
Both the notional source Ψ_m at position $\mathbf{r}_\Psi$ and the receiver R at position $\mathbf{r}_R$ are positioned in the horizontal plane $y=0$. The wave field of the notional monopole source is synthesized by a secondary monopole and dipole point source distribution on the surface S_0.

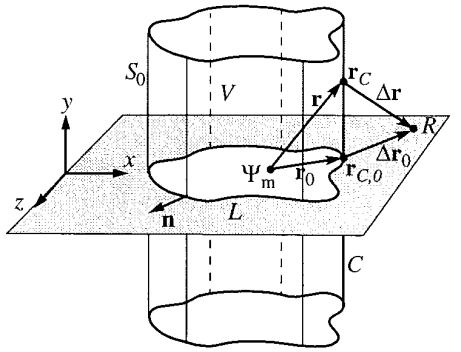


Figure 3.4: Diagram of a notional monopole source Ψ_m at position $\mathbf{r}=\mathbf{r}_\Psi$ in the plane $y=0$ and a infinite long cylindrical secondary source surface S_0. The contour L is the cross section of the surface S_0 and $y=0$. The contribution in R of a secondary source column C trough $\mathbf{r}_{C,0}$ perpendicular to the plane $y=0$ can be approximated by a contribution of point $\mathbf{r}_{C,0}$ only.

Making the same approximations as in the previous sections, the contribution of each secondary monopole and dipole source column may be replaced by the contribution of one secondary monopole and one dipole point source on the contour L. Consequently, the 3D configuration of figure 3.4 is transformed to the configuration depicted in figure 3.5.
The 2½D Kirchhoff-Helmholtz integral can be written as

$$P(\mathbf{r}_R) = \frac{1}{4\pi} \oint_{S_0} \left[j\omega\rho_0 V_n^{\text{line}}(\mathbf{r}_L) g(\mathbf{r}_R, \mathbf{r}_L) \frac{e^{-jk\Delta r}}{\Delta r} \right] dL$$
$$+ \frac{1}{4\pi} \oint_{S_0} \left[P^{\text{line}}(\mathbf{r}_L) g(\mathbf{r}_R, \mathbf{r}_L) jk\cos\varphi \frac{e^{-jk\Delta r}}{\Delta r} \right] dL \qquad \text{for } \mathbf{r}_R \in D \tag{3.27}$$

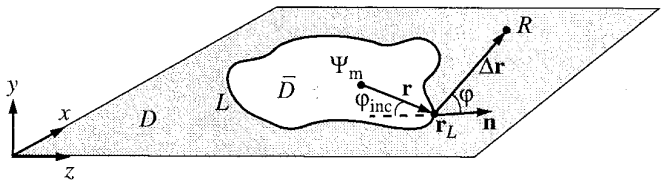


Figure 3.5: 2½D Kirchhoff-Helmholtz configuration with a notional monopole source Ψ_m enclosed by a secondary monopole and dipole point source distribution on the contour L.

in which D is the receiver area situated outside the contour L. The strengths of the secondary dipole and monopole point sources are given by the pressure P^{line} and the particle velocity V_n^{line} of a monopole line source, given by eqs. (2.40a) and (3.19). The amplitude factor g is given by eq. (3.11). In order to find a practical function g that is valid for more than one receiver position, consider a receiver contour K through point R around contour L, as shown in figure 3.6. This

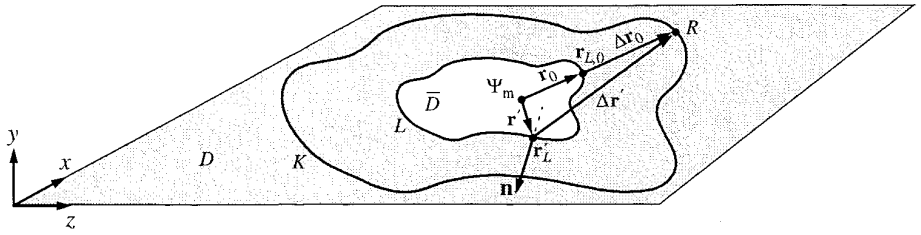


Figure 3.6: 2½D Kirchhoff-Helmholtz configuration with a notional monopole source Ψ_m enclosed by a secondary monopole and dipole point source distribution on the contour L. The secondary source position $\mathbf{r}_{L,0}$ and $\mathbf{r}'_L$ are 'stationary phase' points.

receiver contour K and the secondary source contour L must be chosen such that the line piece connecting the notional source Ψ_m and an arbitrary receiver point R on contour K crosses both the contour L and the contour K in one point only. The secondary source at the intersection point $\mathbf{r}_{L,0}$ on contour L is a 'stationary phase' point along the contour L for this notional source-receiver combination. However, note that in such a configuration multiple 'stationary phase' points can exist for a certain receiver R in which the phase, given by

$$\phi(x_L, z_L) = -k(r + \Delta r), \tag{3.28}$$

reaches a local minimum. For instance, position $\mathbf{r}'_L$; at this point the phase ϕ also reaches a local minimum, since $\varphi=\pi-\varphi_{\text{inc}}$, with φ_{inc} the angle between $\Delta\mathbf{r}'$ and the normal $\mathbf{n}$, and φ the angle between $\mathbf{r}'$ and the normal $\mathbf{n}$. Position $\mathbf{r}_{L,0}$ is the global 'minimum phase' point which gives the main total contribution at the receiver R. So, by this 'minimum phase' condition, each secondary source at position $\mathbf{r}_L$ defines a unique 'minimum phase' receiver point $\mathbf{r}_{R,0}$ on K, which is determined by the intersection point of the line through $\mathbf{r}_\Psi$ and $\mathbf{r}_L$ and the receiver contour K, with $\varphi=\varphi_{\text{inc}}$.

When we approximate the secondary source-to-receiver distance Δr in the expression (3.11) for the amplitude factor g with the 'minimum phase' secondary source-to-receiver distance $\Delta r_0(\mathbf{r}_L)$, given by

$$\Delta r_0(\mathbf{r}_L) = |\mathbf{r}_{R,0}(\mathbf{r}_L) - \mathbf{r}_L|, \tag{3.29}$$

the amplitude factor g can be written as a function of the secondary source position only, yielding

$$g_0(\mathbf{r}_L) = \sqrt{\frac{\Delta r_0(\mathbf{r}_L)}{r_0(\mathbf{r}_L) + \Delta r_0(\mathbf{r}_L)}} \qquad \text{for} \quad \mathbf{r}_R \in K\,. \tag{3.30}$$

In a similar way, it is also possible to derive the 2½D Kirchhoff-Helmholtz integral for a configuration in which the receiver area D is enclosed by secondary sources. However, a difficulty arises when defining the amplitude factor g_0 for such a configuration. In this case it is impossible to assign a unique receiver position inside the contour L to each secondary source. The best way to work around this problem is to insert a fixed receiver position $\mathbf{r}_R$ in eq. (3.11).

3.4 Wave field synthesis example with the 2½D Kirchhoff-Helmholtz integral

Consider the situation of a notional monopole source Ψ_m at position $\mathbf{r}_\Psi$, emitting a bandlimited impulse. The wave field is synthesized by applying the 2½D Kirchhoff-Helmholtz operator (3.27), using a distribution of secondary monopole and dipole point sources along the contour L.

Figure 3.7 specifies the geometry of the configuration in the plane y=0. The notional monopole source is positioned at $\mathbf{r}_\Psi$=(-1,0,-4), the contour L is defined by the angular points (-2.5,0,0), (2.5,0,0), (5,0,-5) and (-5,0,-5). The source signal has a bandwidth of 0-1 kHz.

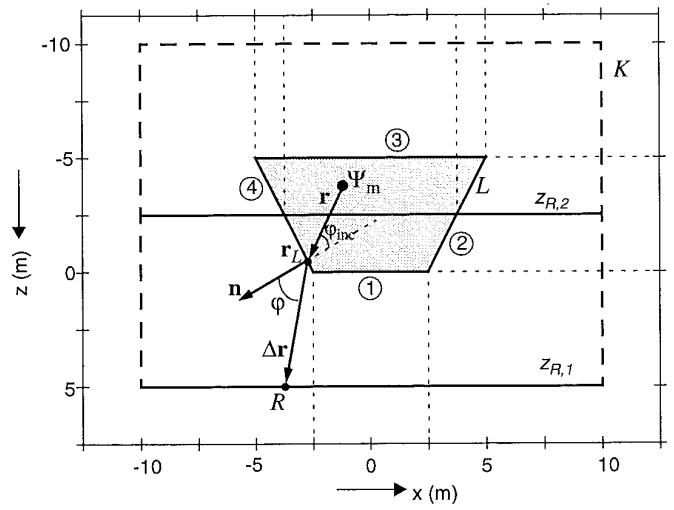


Figure 3.7: Diagram of the geometry of a notional monopole point source Ψ_m at $\mathbf{r}_\Psi$=(-1,0,-4) and a secondary monopole and dipole point source distribution along the 4 sides of contour L, numbered counter clockwise from 1 to 4. The primary and the synthesized wave field are recorded at the line $z_{R,1}$=5m and at $z_{R,2}$=-2.5m for -10<x<10m.

In order to assign a unique 'minimum phase' receiver position to each secondary source in the way described in the previous section, consider a receiver contour K around contour L, as shown in figure 3.7. For this configuration the correct amplitude factor g_0 can be calculated for each secondary source according to eq. (3.30).

Firstly, the primary and synthesized wave fields are registered at the lower part of the rectangle K, at the line $z_{R,1}$=5 (-10<x<10m).

In figure 3.8 the primary (a) and the synthesized (b) wave fields at the receiver line $z_{R,1}$ are shown. Figure 3.8a and b show that the synthesized wave field is in good agreement with the primary wave field. Besides the correct wave front a weak 'shadow' wave front can be observed in figure 3.8b. The origin of this shadow wave front can be recovered by looking at the at the individual contributions of the secondary monopoles and dipoles for each side of the contour L. This is shown in figure 3.8c. It is demonstrated that side 1 gives the main contribution for centrally positioned receivers. For the receivers at the line $z_{R,1}$ the contributions of the secondary monopoles and dipoles at side 1 are in phase. Sides 2 and 4 mainly contribute to the wave field at the more laterally positioned receivers. The contribution of side 3 is negligible because the wave field of the monopoles and dipoles are in counterphase. However, since the 2½D Kirchhoff-Helmholtz integral is a far field and/or high-frequency approximation, a small residue remains. Notice that due to the finite length of the sides of the contour L, diffraction effects become visible in the wave field contributions of each side. However, each diffraction 'tail' is compensated by a diffraction tail produced by a neighboring contour side, which results in a diffraction-free synthesized wave field.

The primary and secondary wave field are also registered at the line $z_{R,2}$=-2.5 for -10<x_R<10 m, which crosses the contour L. This registration line is not a part of the receiver contour K, and, as a consequence, the amplitude factor g_0 for the secondary sources is not correct for this configuration. This can be solved by choosing a fixed receiver position $\mathbf{r}_R$=(0,0,-2.5), that can be inserted at eq. (3.11).

In figure 3.9 the primary (a) and the synthesized (b) wave fields are shown. Figure 3.9b shows that, as expected, the contribution of the secondary sources at the receiver line, inside L (-3.75<x<3.75 m), is negligible. Outside L, the wave field is in good agreement with the primary field (figure 3.9a). Near the contour L, the synthesized wave field differs somewhat from the desired response which is not surprising because the 2½D operator is a far field and/or high-frequency approximation.

Inspection of the individual contributions of the secondary monopoles and dipoles for each side of the contour L, shown in figure 3.9c, reveals that the monopole and dipole contributions of side 1 cancel each other, giving a negligible total contribution. The same holds for side 3. Side 2 only contributes to receiver positions for which x>3.75 m. Inside and at the left-hand side of contour L the monopole and dipole contributions are in counterphase. The opposite side (side 4) only contributes to receiver positions for which x<-3.75 m.

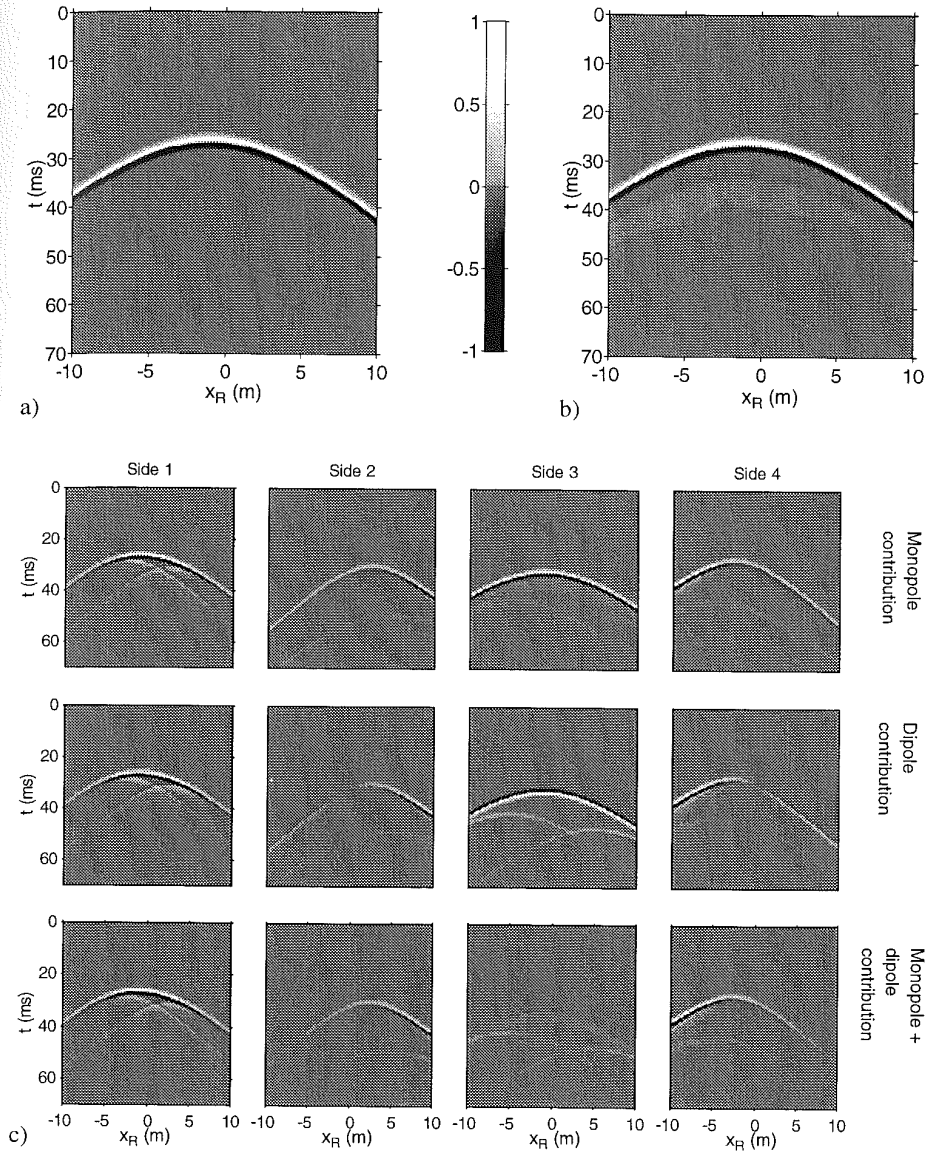


Figure 3.8: *Wave field p at the receiver line $z_{R,1}$=5m for -10<x_R<10 m.*
a) Primary wave field
b) Synthesized wave field
c) Individual contributions of the secondary monopole and dipole sources

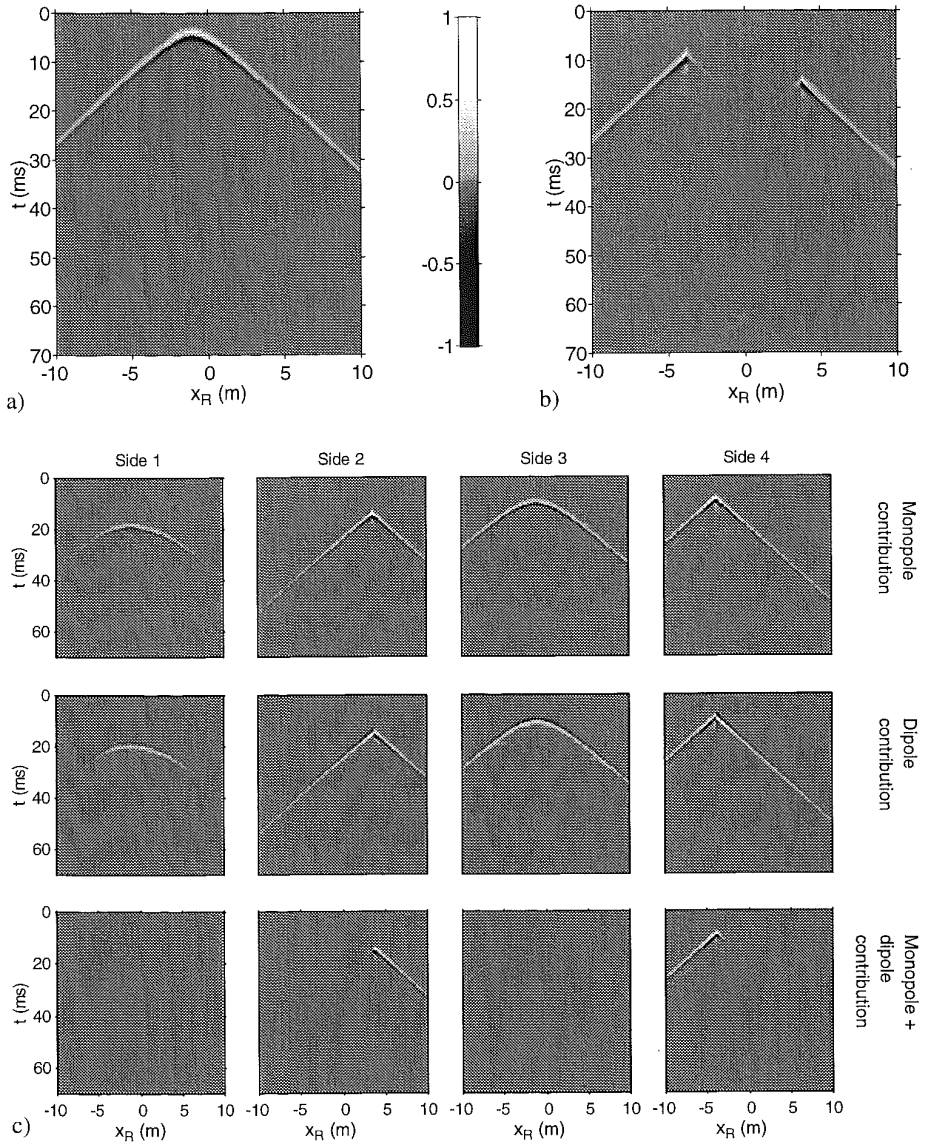


Figure 3.9: *Wave field p at the receiver line $z_{R,2}$ for $-10<x_R<10$ m.*
a) Primary wave field
b) Synthesized wave field
c) Individual contributions of the secondary monopole and dipole sources

The findings of the preceding example can be generalized as follows:

- Inside contour L the synthesized wave field is negligible and would ideally be zero.
- Outside contour L the contributions of the secondary monopoles and dipoles interfere in a constructive manner, and are almost equal.
- Only the parts of the contour L that are positioned between the notional source and a certain receiver position contribute to the wave field at that receiver position. It is equivalent to state that mainly the global 'minimum phase' points contribute.

3.5 The generalized 2½D Rayleigh integrals

In order to construct a more efficient synthesis operator, we will break up the complete Kirchhoff-Helmholtz integral, using the findings of the previous section.
With concern to the receivers at the line $z_{R,1}$ for $-10<x_R<10$m (see figure 3.7) the contribution of side 3 may be omitted. It may be expected that by driving simply the monopoles or the dipoles at sides 1, 2 and 4, practically the same wave field can be synthesized at the line $z_{R,1}=5$ as with the complete secondary source distribution. In order to have a correct amplitude the strengths of the monopoles or dipoles must be doubled. In fact the secondary monopoles are driven with the Rayleigh I operator, while the secondary dipoles are driven with the Rayleigh II operator. In figure 3.10a the result of the synthesis is shown using only secondary monopoles. Figure 3.10b represents the synthesis using only the secondary dipoles. In both cases the synthesized wave field matches the primary wave field in figure 3.8a rather accurately. However, note that for both situations the diffraction artifacts, indicated by the arrows marked 1a and 1b,

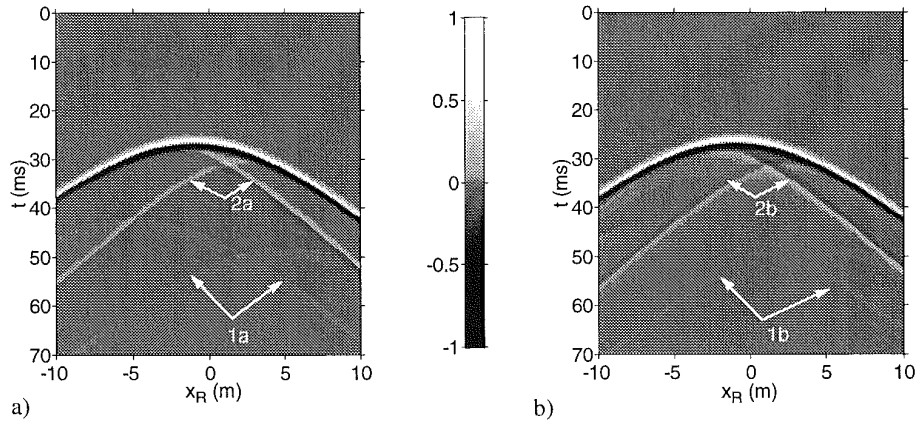


Figure 3.10: *Wave field p at the receiver line $z=z_{R,2}$ for lateral receiver positions $-10<x_R<10$ m.*
a) Synthesized wave field using the secondary monopoles of side 1,2 and 4 only
b) Synthesized wave field using the secondary dipoles of side 1,2 and 4 only.

due to the ends of the array are not compensated anymore by the contribution of side 3. Also note that diffraction artifacts become visible, indicated by the arrows marked 2a and 2b, emerging from the bents in the array at $\mathbf{r}_L$=(−2.5,0,0) and $\mathbf{r}_L$=(2.5,0,0). Addition of both wave fields 3.10a and b would display a cancellation of the artifacts 2a and 2b, while the artifacts 1a and 1b would add up.

Based on the above considerations, we postulate a generalized 2½D Rayleigh I operator for a bent line array L, as depicted in figure 3.11:

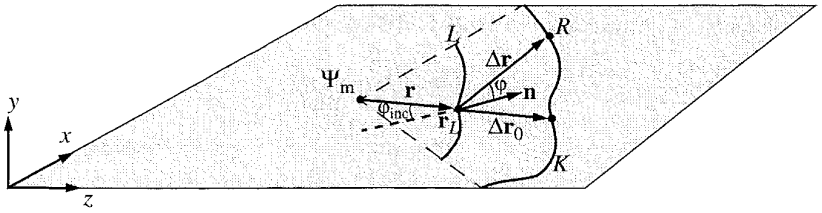


Figure 3.11: Diagram of a notional monopole source Ψ_m and a secondary monopole source distribution on a bent array L. The operator can be optimized for the receivers at curve K.

$$P(\mathbf{r}_R) = \frac{1}{2\pi}\int_L g_0(\mathbf{r}_L) j\omega\rho_0 V_n^{\text{line}}(\mathbf{r}_L)\frac{e^{-jk\Delta r}}{\Delta r}dL, \qquad (3.31)$$

with $g_0(\mathbf{r}_L)$ defined by eqs. (3.29)-(3.30). The receiver curve K and the secondary source curve L must be chosen such that the line through the notional source Ψ_m and an arbitrary receiver point R on curve K crosses both the curve L and the K in one point only. This condition can be fulfilled by choosing a suitable array and receiver curve geometry.

The generalized 2½D Rayleigh II operator for a bent line array L is given by

$$P(\mathbf{r}_R) = \frac{1}{2\pi}\int_L g_0(\mathbf{r}_L) P^{\text{line}}(\mathbf{r}_L) jk\cos\varphi\frac{e^{-jk\Delta r}}{\Delta r}dL, \qquad (3.32)$$

in which case the secondary monopole sources are replaced by dipole sources.

3.6 Synthesized wave fields in the vertical plane

Until now we assumed that the notional source, the array and the receivers were positioned in the same horizontal plane. In this section the effect of differences in height will be investigated. As shown in the previous sections, the wave field of a notional monopole source Ψ_m can be synthesized correctly in a plane through the notional source and the horizontal line array L. This is illustrated in figure 3.12a, in which the wave fronts in the horizontal plane are shown. How-

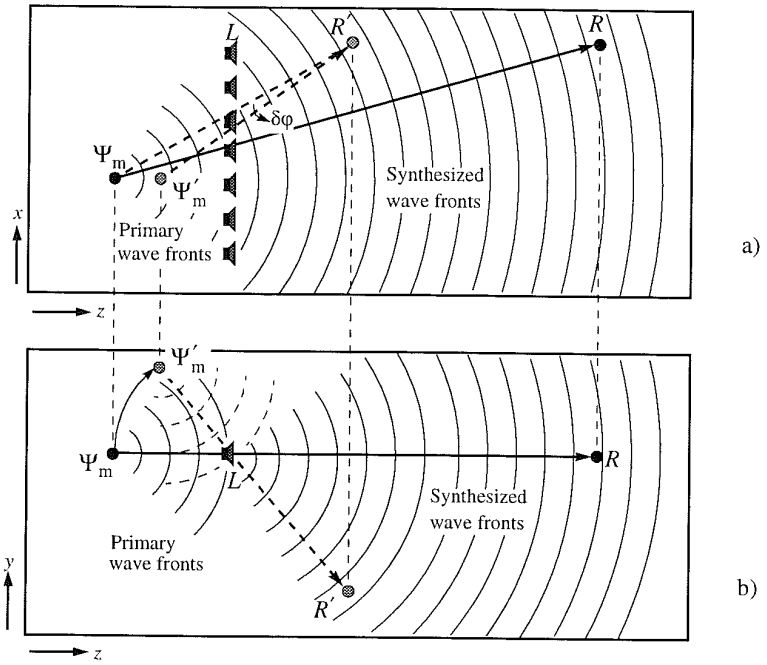


Figure 3.12: The apparent source position is different for receivers at unequal heights. For receiver R, which is positioned at the same height as the array L, the apparent source is positioned correctly with respect to azimuth and elevation. For receiver R′ at a lower position the apparent source is rotated around the array L, yielding an elevation of the source and a small azimuthal deviation δφ.

a) Cross-section in the horizontal plane.

b) Cross-section in the vertical plane.

ever, in the vertical plane perpendicular to the array, the wave fronts become circular and originate from the array instead of the notional source, as shown in figure 3.12b. Since the receiver R is positioned at the same height as the array L, the apparent source is positioned correctly with respect to azimuth and elevation.

For a second receiver R' at a lower position, the situation becomes more complicated. Since, inherent to 2½D wave field synthesis, the notional source is positioned in the same (tilted) plane as the array and the receiver, the notional source Ψ'_m is rotated around the array L for receiver R'. This has two effects; first, the apparent source is elevated, and secondly, a small azimuthal deviation $\delta\varphi$ is introduced for receiver R'. Note that for a certain fixed elevation of the array, as seen from a receiver, this azimuthal deviation can be corrected for, by adjusting the z-coordinate of the notional source. However, generally in a hall, the elevation of the array is not

the same for all receivers (listeners), and a small azimuthal deviation will exist for those receivers which experience an array elevation below or above the averaged elevation.
In many configurations the difference between the averaged and the actual array elevation is very small, which means that the resulting azimuthal deviation may be neglected.
The investigation to the vertical localization aspects of synthesized sound fields will be postponed until chapter 6.

3.7 Conclusions and discussion

In this chapter the 2½D wave field synthesis operators have been derived according to a new method. These operators offer new possibilities in wave field synthesis applications by making it possible to synthesize a 3D wave field in a plane through a (bent) line array consisting of only monopole or dipole sources.
Depending on the source-to-array distance with respect to the array-to-receiver distance, two different approximations to the full 3D synthesis operators have been made:

1. For the situation that the ratio between the source-to-array distance and the array-to-receiver distance is small, the wave field of a notional *point* source can be synthesized using a *horizontal line* array consisting of secondary monopole or dipole *point* sources. The driving signals of the secondary *monopole point* sources are proportional to the normal component of the particle velocity of a *vertical line* source at the same position as the notional point source. For a *horizontal line* array consisting of secondary *dipole point* sources the driving signals are given by the pressure of a *vertical line* source.
   These versions of the 2½D operators are well suited for direct sound enhancement applications and for situations where the wave field of early reflections is synthesized.
2. For the situation that the ratio between the source-to-array distance and the array-to-receiver distance is large, the wave field of a notional *point* source can be synthesized using a *horizontal line* array consisting of *vertical* secondary monopole or dipole *line* sources. The driving signals of the secondary *monopole line* sources are proportional to the normal component of the particle velocity of a *point* source at the same position as the notional point source. For a *horizontal line* array consisting of secondary *dipole line* sources the driving signals are given by the pressure of a *point* source.
   These versions of the 2½D operators are well suited for the generation of late reflections and reverberation since in those situations the sources are positioned at relatively large distances from the array.

For a configuration which does not fulfill one of these conditions, the wave field can be synthesized using configuration 1. The amplitude of the synthesized wave field can be optimized for a receiver line at a certain averaged depth, which means that for receivers at other depths only an amplitude error is introduced.

4

Finite secondary source distributions

In the previous chapter the foundation has been laid for 3D wave field synthesis with continuous (bent) line arrays, by deriving the (generalized) 2½D Rayleigh integrals. With the practical implementation of a system based on these wave field synthesis techniques some limitations have to be accepted. The main limitations originate from the limited number of channels that can be processed by a certain hardware configuration. Therefore, in practice a discretized and finite version of a continuous distribution of secondary sources is used.
In this chapter we will investigate the performance of finite secondary source arrays. The discretization of the secondary source distribution will be addressed in the next chapter.
To enable an objective comparison between the synthesized and the primary wave field, two physical performance measures will be defined.

4.1 Physical performance measures

4.1.1 The error level

The performance of a wave field synthesis system can be measured in the space-time domain by the synthesis error e, which is given by

$$e(\mathbf{r}_R, t) = p(\mathbf{r}_R, t) - p_{\mathrm{pr}}(\mathbf{r}_R, t), \tag{4.1}$$

where $p_{\mathrm{pr}}(\mathbf{r}_R,t)$ is the 'desired' primary wave field and $p(\mathbf{r}_R,t)$ the synthesized wave field at the receiver position $\mathbf{r}_R$.
Fourier transformation to the space-frequency domain yields the spectral error E

$$E(\mathbf{r}_R, f) = P(\mathbf{r}_R, f) - P_{\mathrm{pr}}(\mathbf{r}_R, f). \tag{4.2}$$

Note that $P_{\mathrm{pr}}(\mathbf{r}_R,f)$ and $P(\mathbf{r}_R,f)$ are complex spectra, yielding a complex spectral error signal $E(\mathbf{r}_R,f)$.
It is convenient to express the spectral error E as an error level L_E, according to

$$L_E(\mathbf{r}_R, f) = 10\log\left\{\frac{|E(\mathbf{r}_R, f)|^2}{|P_{\mathrm{pr}}(\mathbf{r}_R, f)|^2}\right\} \quad [\mathrm{dB}], \tag{4.3}$$

with $P_{\mathrm{pr}}(\mathbf{r}_R, f) \neq 0$ on the domain of interest. The error level L_E gives the sound pressure level of the error E relatively to the sound pressure level of the primary wave field P_{pr}. Ideally, L_E=$-\infty$ dB for all receivers and for all frequencies of interest, which means that the synthesized wave field is both temporally and spatially correct.

4.1.2 The pressure ratio level

A second performance measure of a wave field synthesis system is called the synthesized-to-primary pressure ratio R_P, which is defined in the space-frequency domain by the ratio between the pressure of the synthesized wave field $P(\mathbf{r}_R,f)$ and the pressure of the primary wave field $P_{\mathrm{pr}}(\mathbf{r}_R,f)$

$$R_P(\mathbf{r}_R, f) = \frac{P(\mathbf{r}_R, f)}{P_{\mathrm{pr}}(\mathbf{r}_R, f)}, \tag{4.4}$$

with $|P_{\mathrm{pr}}(\mathbf{r}_R, f)| \neq 0$ on the domain of interest. The pressure ratio R_P, gives the complex amplitude of each frequency component in the synthesized wave field relative to the complex amplitude of the primary wave field. Ideally, R_P=1 for all receivers and for all frequencies of interest, in which case the synthesized wave field is identical to primary wave field.
When we are only interested in the amplitude level of the pressure ratio, we can write

$$L_R(\mathbf{r}_R, f) = 10\log\{|R_P(\mathbf{r}_R, f)|^2\} \quad [\mathrm{dB}] \tag{4.5}$$

The pressure ratio level L_R gives the sound pressure level of the synthesized wave field P relative to the sound pressure level of the primary wave field P_{pr}. In the case that L_T=0dB, the amplitude spectrum of the synthesized wave field is identical to the amplitude spectrum of the primary wave field, which does not mean that the wave field is synthesized correctly with respect to phase aspects. With the pressure ratio level L_R we can analyze the temporal amplitude spectrum of the synthesized wave field at the receiver positions while the phase properties are disregarded.

4.2 Diffraction effects in synthesized wave fields

In this section the 2½D Rayleigh operators for finite secondary source arrays will be formulated, and the effects of the truncation of the secondary source array will be described.

4.2.1 Truncation of the 2½D Rayleigh integrals

In chapter 3 the 2½D Rayleigh I and II integrals were derived in the space-frequency domain.
2½D Rayleigh I:

$$P(x_R) = \sqrt{\frac{z_R - z_L}{z_R - z_\Psi}} \int_{-\infty}^{\infty} \sqrt{\frac{jk}{2\pi}} S(\omega) \cos\varphi_{\mathrm{inc}} \frac{e^{-jkr}}{\sqrt{r}} \frac{e^{-jk\Delta r}}{\Delta r} dx_L. \tag{4.6}$$

2½D Rayleigh II:

$$P(x_R) = \sqrt{\frac{z_R - z_L}{z_R - z_\Psi}} \int_{-\infty}^{\infty} \sqrt{\frac{1}{2\pi jk}} S(\omega) \frac{e^{-jkr}}{\sqrt{r}} jk \cos\varphi \frac{e^{-jk\Delta r}}{\Delta r} dx_L. \tag{4.7}$$

Since z_L and z_R are fixed, both integrals can be written as

$$P(x_R) = \int_{-\infty}^{\infty} Q(x_L) W(x_R - x_L) dx_L, \tag{4.8}$$

where $Q(x_L)$ is the driving signal for the secondary source at $\mathbf{r}_L$=(x_L,0,z_L) on the line z=z_L, and $W(x_R$-$x_L)$ is the response of that secondary source at the receiver position $\mathbf{r}_R$=(x_R,0,z_R) on the line z=z_R. $Q(x_L)$ is determined by the field of a vertical line source through the primary source position, measured at the position of the secondary source. Eq.(4.8) represents a convolution along the x-axis, which can symbolically be written as

$$P(x_R) = Q(x_R) * W(x_R). \tag{4.9}$$

For the 2½D Rayleigh I case, where the secondary sources are monopoles, Q and W are determined respectively by the monopole driving signal Q_{m} and the secondary monopole response W_{m}, given by

$$Q(x_L) = Q_{\mathrm{m}}(x_L) = \sqrt{\frac{z_R - z_L}{z_R - z_\Psi}} \sqrt{\frac{jk}{2\pi}} S(\omega) \cos\varphi_{\mathrm{inc}} \frac{e^{-jkr}}{\sqrt{r}} \tag{4.10a}$$

and

$$W(x_R - x_L) = W_{\mathrm{m}}(x_R - x_L) = \frac{e^{-jk\Delta r}}{\Delta r}. \tag{4.10b}$$

For the 2½D Rayleigh II case, where the secondary sources are dipoles, Q and W are determined respectively by the dipole driving signal Q_d and the secondary dipole response W_d, given by

$$Q(x_L) = Q_d(x_L) = \sqrt{\frac{z_R - z_L}{z_R - z_\Psi}}\sqrt{\frac{1}{2\pi jk}}S(\omega)\frac{e^{-jkr}}{\sqrt{r}} \tag{4.11a}$$

and

$$W(x_R - x_L) = W_d(x_R - x_L) = jk\cos\varphi\frac{e^{-jk\Delta r}}{\Delta r}. \tag{4.11b}$$

In the remainder of this section we will only consider the 2½D Rayleigh I integral. Therefore, Q will be referred to as the monopole driving function Q_m and W will be referred to as the secondary monopole response W_m.

In practical applications we always have to deal with finite secondary source arrays. Considering the case that the secondary sources are positioned on the line $z=z_L$ at a finite interval $[x_A,x_B]$, the convolution integral (4.8) can be rewritten as

$$P(x_R) = \int_{x_A}^{x_B} Q(x_L)W(x_R - x_L)dx_L, \tag{4.12}$$

In fact we multiplied the driving signal $Q(x_L)$ with a spatial window $T(x_L)$, with $T(x_L)=1$ inside the interval $[x_A, x_B]$ and $T(x_L)=0$ outside the interval $[x_A, x_B]$.

For an arbitrary window function T, eq.(4.12) can be generalized to

$$P(x_R) = \int_{-\infty}^{\infty} [T(x_L)Q(x_L)]W(x_R - x_L)dx_L, \tag{4.13}$$

Writing this result in symbolic convolution notation, yields

$$P(x_R) = [T(x_R)Q(x_R)] * W(x_R) = Q^T(x_R) * W(x_R), \tag{4.14}$$

where Q^T is the windowed driving signal.

4.2.2 Wave field synthesis with an infinite linear array

Before we investigate the performance of a truncated array, consider the 'reference' configuration shown in figure 4.1. The wave field of a notional monopole source Ψ_m at position $\mathbf{r}_\Psi$=(-1,0) is synthesized by applying the 2½D Rayleigh I integral (4.6). The secondary monopole point sources are continuously distributed along the line L of infinite length (z_L=2m). The notional source emits a bandlimited impulse (0-1.7 kHz), as shown in figure 4.2a and b. The synthesized wave field p and the primary wave field p_{pr} are recorded at the line $z=z_R$=5m for -6.4<x_R<6.4m. The x-t diagram of the primary wave field p_{pr} is shown in figure 4.2c.

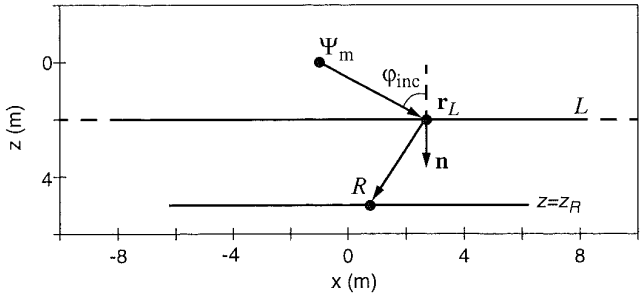


Figure 4.1: *Diagram of the geometry of a notional monopole source Ψ_m at $\mathbf{r}_\Psi$=(-1,0) and a infinite secondary monopole point source distribution along line L. The primary and the synthesized wave field are recorded at the line $z=z_R$=5m for -6.4<x_R<6.4m.*

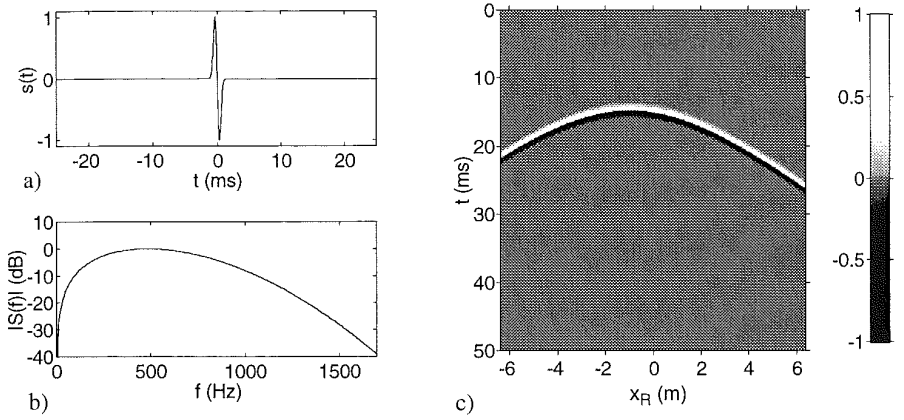


Figure 4.2: *Primary wave field p_{pr} emitted by the notional monopole source.*
a) source signature s(t)
b) amplitude spectrum |S(f)| of the source signal, bandwidth 0-1.7 kHz.
c) x-t diagram

The synthesized wave field P is shown in the x-t diagram of figure 4.3a. The synthesis error $e(x,t)$ is depicted in figure 4.3b. As expected, the synthesized wave field matches well the primary wave field.

Figure 4.4a shows the x-f diagram of the error level L_E for -6.4<x_R<6.4m and 0<f<1.7kHz. Except for very low frequencies (f<100 Hz), the error level L_E is very low and does not exceed -20 dB. The error in the wave field for the low frequencies is not surprising since in the derivation of the 2½D operators a high-frequency and/or a far field assumption has been made.

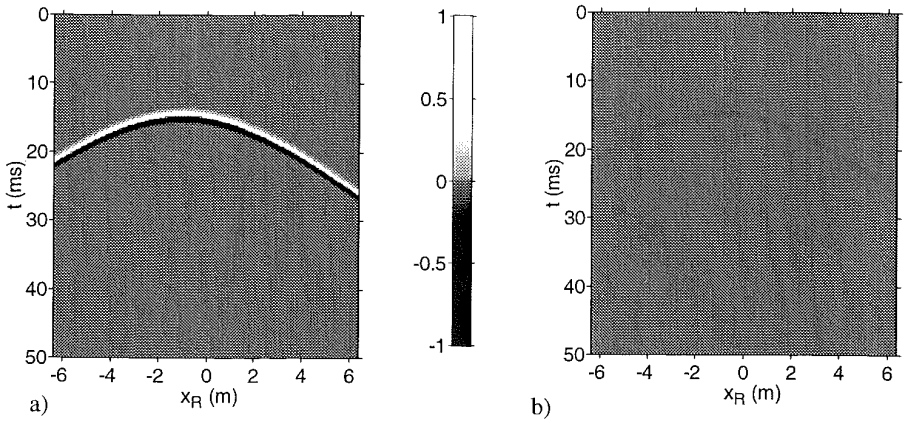


Figure 4.3: *Wave field synthesis with a infinite monopole source array. Registration along the line $z=z_R$ for $-6.4<x_R<6.4$m. Temporal bandwidth: 0-1.7 kHz*
a) x-t diagram of the synthesized wave field p
b) x-t diagram of the synthesis error e.

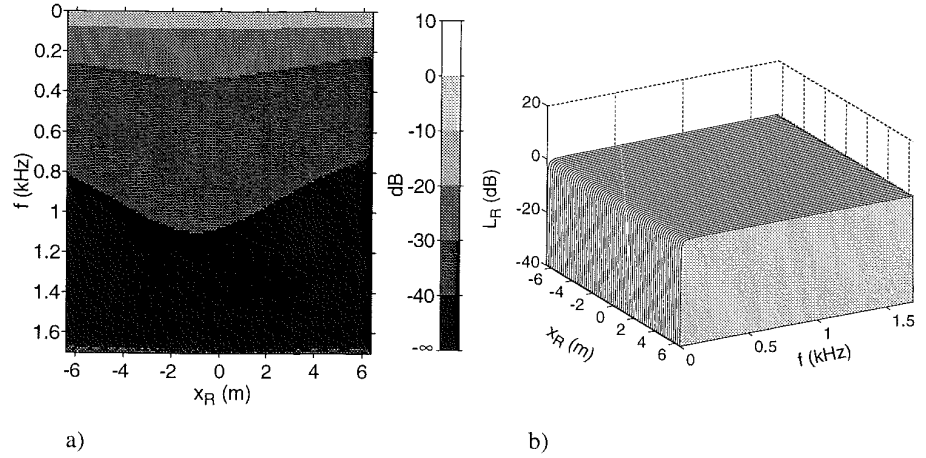


Figure 4.4: *Performance measures in the x-f domain. Wave field synthesized with a infinite monopole source array. Registration along the line $z=z_R$ for $-6.4<x_R<6.4$m. Temporal bandwidth: 0-1.7 kHz.*
a) x-f diagram of the error level L_E
b) x-f diagram of the pressure ratio level L_R

Figure 4.4b shows the *x-f* diagram of the pressure ratio level L_R. Note that the pressure ratio is completely flat at a level of L_R=0 dB, which means that the temporal spectral contents of the synthesized wave field is identical to the contents of the primary wave field. The amplitude for very low frequencies is too small, for reasons which are mentioned above.

4.2.3 Example of wave field synthesis with a truncated array

Now, consider the configuration shown in figure 4.5 in the horizontal plane y=0. The infinite

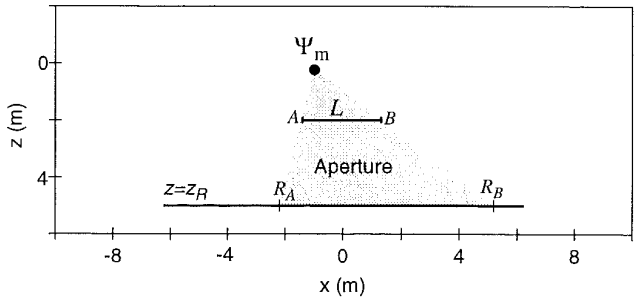


Figure 4.5: *Diagram of the geometry of a notional monopole source Ψ_m at $\mathbf{r}_\Psi=(-1,0)$ and a continuous array L of finite length. The end-points A and B are positioned at $x_A=-1.5$m and $x_B=1.5$m. The synthesized wave field is recorded at the line $z=z_R=5$m for $-6.4<x<6.4$m.*

array is replaced by a linear truncated secondary monopole source array at z_L=2m. The end-points *A* and *B* of the array *L* are positioned at x_A=-1.5m and x_B=1.5m. For this situation the windowed 2½D Rayleigh I integral (4.13) is applied, using a rectangular window *T*.

Due to the limited aperture of the array as seen from the notional source (indicated by the grey triangle), the receiver line cannot be covered completely by the array. Only for receivers positioned between R_A at $\mathbf{r}_{R,A}=(-2.25,5)^T$ and R_B at $\mathbf{r}_{R,B}=(5.25,5)^T$, which have a stationary point on the array and consequently lie within the aperture, reliable results can be expected.

Figure 4.6a shows the *x-t* diagram of the synthesized wave field $p(x,t)$. The synthesis error $e(x,t)$ is depicted in figure 4.6b.

Within the aperture, the synthesized wave field consists of a correct main wave front plus two additional wave fronts, originating from the edges *A* and *B* of the array. For the individual traces, the time delays between the main wavelet and the diffraction wavelets depend on the registration position. At receiver position R_A the main wavelet and the diffraction wavelet originating from the left end-point *A* coincide. The same occurs at R_B for the main wavelet and the diffraction wavelet originating from the right end-point *B*. Outside the aperture only a diffracted wave field exists.

The *x-f* diagram of the error level L_E of the synthesized wave field for $-6.4<x_R<6.4$m and $0<f<1.7$kHz is shown in figure 4.7a. Even for receiver positions inside the aperture, the error level is high. The error level decreases for increasing frequency. Figure 4.7b shows the pressure

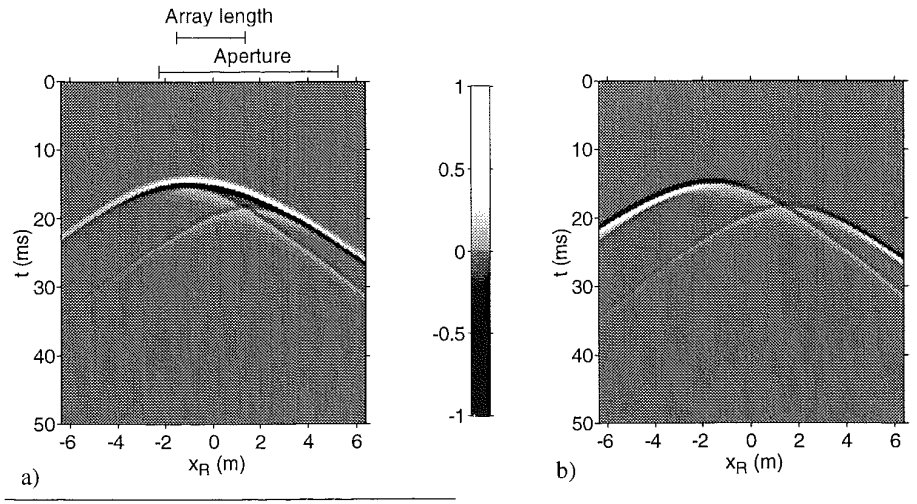


Figure 4.6: *Wave field synthesis with a truncated monopole source array $-1.5 \le x_L \le 1.5$m.*
a) x-t diagram of the synthesized wave field p
b) x-t diagram of the synthesis error e.

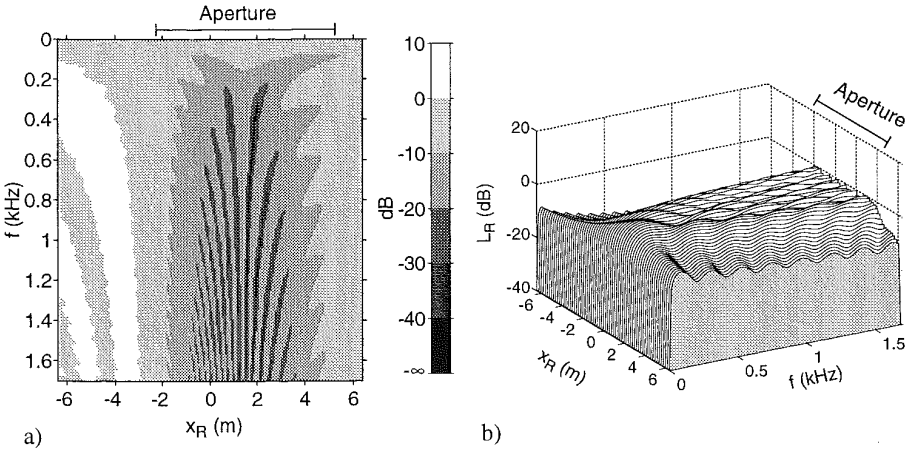


Figure 4.7: *Performance measures in the x-f domain. Wave field synthesized with a truncated continuous monopole source array. Registration along the line $z=z_R$ for $-6.4<x_R<6.4$m. Temporal bandwidth: 0-1.7 kHz*
a) x-f diagram of the error level L_E
b) x-f diagram of the pressure ratio level L_T

ratio level L_R. Inside the aperture the effect of truncation is observable as a ripple superimposed on the desired pressure ratio level L_R of figure 4.4b. For receiver positions at increasing distances from the aperture and for increasing frequency, the pressure ratio decreases.

4.2.4 Reduction of the truncation artifacts by tapering

In the previous example we saw that due to truncation of the secondary source distribution, diffraction artifacts occur in the synthesized wave field. The occurrence of diffraction artifacts causes a disturbance of the spatial information in the synthesized wave field and, additionally, a distortion of the temporal information, possibly causing unwanted coloration of the perceived sound. Therefore, attention must be paid to the suppression of diffraction artifacts.
One approach to suppressing these artifacts is tapering the driving signal Q near the end-points, instead of truncating abruptly the driving signal. This means that a spatial window T is applied to the driving signal Q, such that a decreasing weight is given to the secondary sources near the edges of the array. The tapering function should be chosen such that the amplitude reaches zero for the secondary sources at x_A and x_B. Next, consider again the configuration of figure 4.5. For this situation the driving signal Q is tapered by a one-sided cosine function over the first and last 20% of the array length, with a flat middle part, as shown in figure 4.8. Figure 4.9a shows

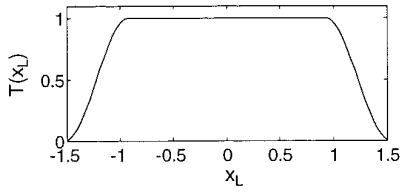


Figure 4.8: *Tapering function T*

the x-t diagram of the synthesized wave field $p(x,t)$. Note that the amplitude of the main wave front is still correct while the diffraction wave fronts are weaker and smeared out in time compared to figure 4.6a. The synthesis error $e(x,t)$ is depicted in figure 4.9b. The x-f diagram of the error level L_E of the synthesized wave field for $-6.4<x_R<6.4$m and $0<f<1.7$kHz is shown in figure 4.10a. Inside the aperture the error is decreased substantially compared to figure 4.7a. However, for low frequencies the error level remains rather high. Figure 4.10b shows the pressure ratio level L_R. The pressure ratio level is flattened inside the aperture. For receiver positions outside the aperture, the amplitude has been decreased.
We can conclude that tapering is a very simple and easily implemented method to reduce truncation artifacts. It may be clear that the shape of the tapering function is always a trade-off between the effective size of the aperture, i.e. the size of the listening area, and the reduction of truncation effects.

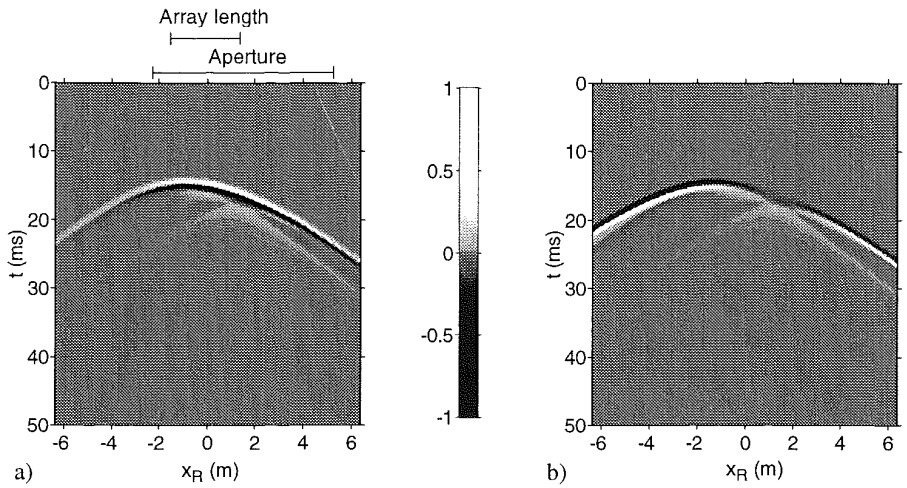


Figure 4.9: *Wave field synthesis with a tapered monopole source array $-1.5 \le x_L \le 1.5$ m.*
a) x-t diagram of the synthesized wave field p
b) x-t diagram of the synthesis error e.

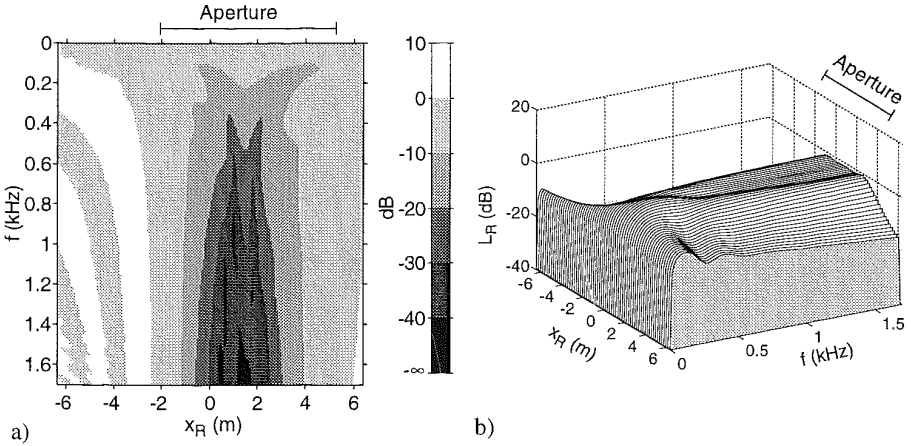


Figure 4.10: *Performance measures in the x-f domain. Wave field synthesized with a tapered continuous monopole source array. Registration along the line $z=z_R$ for $-6.4<x_R<6.4$. Temporal bandwidth: 0-1.7 kHz*
a) x-f diagram of the error level L_E
b) x-f diagram of the transfer function level L_T

4.3 Mathematical description of the truncation artifacts

Generally, in practical applications of wave field synthesis the tapering method is an adequate way to reduce truncation effects. However, in order to gain more fundamental insight, it is interesting to approach the diffraction phenomenon more theoretically. Finite aperture effects may be compared to diffractions that occur when a wave is incident on a barrier with an opening of finite extend. Hence, conventional diffraction theory can be used to model truncation effects. Consider the situation in which the wave field of a notional source Ψ is synthesized with a finite array L. Using optical terms, the receiver area can be divided into an illuminated zone $\mathcal{A}$ and two dark zones $\mathcal{D}_1$ and $\mathcal{D}_2$, as shown in figure 4.11. The points A and B mark the begin- and

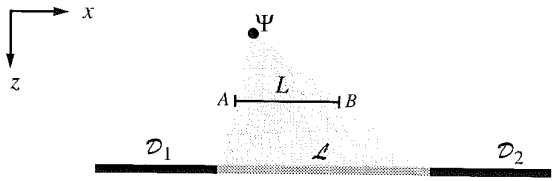


Figure 4.11: *Diagram of the geometry of a notional monopole source Ψ and a secondary source array L of finite length, illuminating the receiver area. The receiver area can be divided into an illuminated zone $\mathcal{A}$ (indicated with a light gray line), and two dark zones $\mathcal{D}_1$ and $\mathcal{D}_2$ (indicated with black lines). The points a and b mark the begin- and end-points of the array.*

end-points of the array. Due to diffraction, the transition from light to dark and vice versa is not very sharp, which is clear from the example in section 4.2.
For notational convenience the 2½D Rayleigh I integral will be rewritten as

$$I = \int_{-\infty}^{\infty} f(x_L) e^{j\phi(x_L)} dx_L \tag{4.15}$$

where $f(x_L)$ and $\phi(x_L)$ represent the amplitude and phase, respectively, of the integrand in eq. (4.6):

$$f(x_L) = \sqrt{\frac{z_R - z_L}{z_R - z_\Psi}} \sqrt{\frac{jk}{2\pi}} S(\omega) \cos\varphi_{inc} \frac{1}{\sqrt{r}\Delta r} \tag{4.16}$$

and

$$\phi(x_L) = -k(r + \Delta r) \tag{4.17}$$

Note that all the variables influencing f and ϕ are suppressed, except the integration variable x_L. In our situation the infinite integration interval is reduced to a finite interval $[x_A, x_B]$, yielding

$$\langle I\rangle = \int_{x_A}^{x_B} f(x_L)\mathrm{e}^{j\phi(x_L)}dx_L \tag{4.18}$$

in which x_A and x_B are the x-coordinates of the end-points of the array. In order to derive expressions for the diffraction artifacts, we express the approximated result <I> as the sum of a diffraction-free result I_0 plus two diffraction terms, as follows

$$\langle I\rangle = I_0 + I_A + I_B \tag{4.19}$$

Depending on the position of the receiver R, these terms are defined by

$$\left.\begin{aligned} I_0 &= 0 \\ I_A &= \int_{x_A}^{\infty} f(x_L)\mathrm{e}^{j\phi(x_L)}dx_L \\ I_B &= -\int_{x_B}^{\infty} f(x_L)\mathrm{e}^{j\phi(x_L)}dx_L \end{aligned}\right\} \quad \text{for} \quad R\in\mathcal{D}_1 \tag{4.20a}$$

$$\left.\begin{aligned} I_0 &= \int_{-\infty}^{\infty} f(x_L)\mathrm{e}^{j\phi(x_L)}dx_L \\ I_A &= -\int_{-\infty}^{x_A} f(x_L)\mathrm{e}^{j\phi(x_L)}dx_L \\ I_B &= -\int_{x_B}^{\infty} f(x_L)\mathrm{e}^{j\phi(x_L)}dx_L \end{aligned}\right\} \quad \text{for} \quad R\in\mathcal{L} \tag{4.20b}$$

$$\left.\begin{aligned} I_0 &= 0 \\ I_A &= -\int_{-\infty}^{x_A} f(x_L)\mathrm{e}^{j\phi(x_L)}dx_L \\ I_B &= \int_{-\infty}^{x_B} f(x_L)\mathrm{e}^{j\phi(x_L)}dx_L \end{aligned}\right\} \quad \text{for} \quad R\in\mathcal{D}_2 . \tag{4.20c}$$

Note that the integral I_0 in eq.(4.20b) is identical to the 2½D Rayleigh integral (4.15). Inside the illuminated area, the integral I_0 may be replaced by the wave field of the primary source. Now that we have set up the integral formulation for the diffraction effects, we will try to approximate the diffraction contributions I_A and I_B separately, since an exact solution does not exist. Note that, in our situation, the exact expression for the diffraction-free wave field I_0 can be subtracted from the synthesized wave field <I>, yielding the sum of the left and right diffraction contributions, according to

$$(I_A + I_B) = \langle I\rangle - I_0 \tag{4.21}$$

In the literature several solutions can be found for this 'classical' diffraction problem. In the next sections we will deal with two methods.

4.3.1 Asymptotic expansion of the diffraction terms

Assuming that the receiver R lies far away from a light-dark transition, the diffraction terms I_A and I_B may be approximated, for high frequencies, by an asymptotic expansion of the diffraction integrals. According to Bleistein (1984), integrals of the form

$$I_{L,U} = \int_{L}^{U} f(x)\mathrm{e}^{j\phi(x)}dx \tag{4.22}$$

can be expanded to order N, under the assumption that

$$\phi'(x)\neq 0 \qquad L\le x\le U, \tag{4.23}$$

by a repeated process of integration by parts. The N^{th} order asymptotic expansion $\tilde{I}_{L,U}$ reads

$$\tilde{I}_{L,U} \approx \sum_{n=0}^{N-1} \frac{(-1)^n \mathrm{e}^{j\phi(x)}}{(j)^{n+1}} \left[\frac{1}{\phi'(x)}\frac{d}{dx}\right]^n \left[\frac{f(x)}{\phi'(x)}\right]\Bigg|_L^U . \tag{4.24}$$

For the physical understanding, it is illustrative to apply a first order expansion to the diffraction integrals I_A and I_B of eqs.(4.20a-c), yielding

$$\tilde{I}_A \approx -\frac{f(x_A)}{j\phi'(x_A)}\mathrm{e}^{j\phi(x_A)} \tag{4.25a}$$

$$\tilde{I}_B \approx \frac{f(x_B)}{j\phi'(x_B)}\mathrm{e}^{j\phi(x_B)} . \tag{4.25b}$$

in which $\tilde{I}_A$ and $\tilde{I}_B$ denote the approximated results. The derivative of the phase at the end-points $\phi'(x_i)$ is given by

$$\phi'(x_i) = -k\left(\frac{x_i - x_\Psi}{r_i} - \frac{x_R - x_i}{\Delta r_i}\right) = -k(\sin\varphi_{\text{inc},i} - \sin\varphi_i) \qquad i = A, B . \tag{4.26}$$

Looking at the expressions (4.25a,b) it can be noted that the diffraction terms I_A and I_B are approximated by the integrand $f(x_i)\exp(j\phi(x_i))$ at the end-points x_i $(i=A,B)$ multiplied by a correction factor $\pm 1/j\phi'(x_i)$.
Inserting the positions of the end-points in expressions (4.16) and (4.26), and using eqs.(4.25a,b) gives

$$\tilde{I}_A \approx \sqrt{\frac{z_R - z_L}{z_R - z_\Psi}} \sqrt{\frac{1}{2\pi jk}} S(\omega) \cos\varphi_{\text{inc},A} \frac{e^{-jkr_A}}{\sqrt{r_A}} \frac{1}{(\sin\varphi_{\text{inc},A} - \sin\varphi_A)} \frac{e^{-jk\Delta r_A}}{\Delta r_A} \tag{4.27a}$$

$$\tilde{I}_B \approx -\sqrt{\frac{z_R - z_L}{z_R - z_\Psi}} \sqrt{\frac{1}{2\pi jk}} S(\omega) \cos\varphi_{\text{inc},B} \frac{e^{-jkr_B}}{\sqrt{r_B}} \frac{1}{(\sin\varphi_{\text{inc},B} - \sin\varphi_B)} \frac{e^{-jk\Delta r_B}}{\Delta r_B} \tag{4.27b}$$

in which $\Delta r_A = |\Delta \mathbf{r}_A|$ and $\Delta r_B = |\Delta \mathbf{r}_B|$ are the distances between the end-points A and B of the array and the receiver R, while $r_A = |\mathbf{r}_A|$ and $r_B = |\mathbf{r}_B|$ are the distances from the notional source Ψ to the end-points A and B. The angles $\varphi_{\text{inc},A}$, φ_A, $\varphi_{\text{inc},B}$ and φ_B are indicated in figure 4.12. The sign

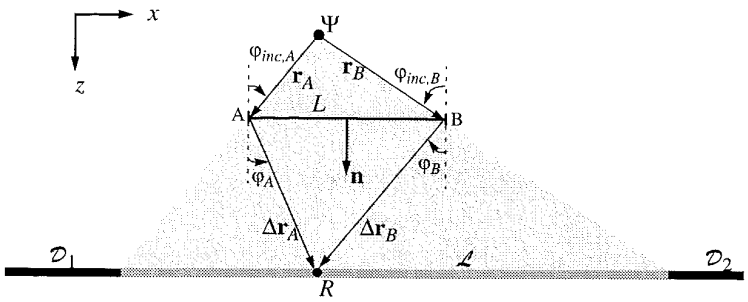


Figure 4.12: Diagram of the geometry of a notional monopole source Ψ and a secondary source array L of finite length, illuminating the receiver area. The diffraction effects can be approximated by a contribution of the end-points only, multiplied with a correction factor

of the angles is defined such that angle between a vector and the normal **n**, is taken positively, when that vector is turned counterclockwise with respect to the normal **n**.
Eqs.(4.27a,b) show that, as we saw in figure 4.6, the diffraction contributions can be represented by scaled point sources at the ends of the array (x-coordinates: x_A and x_B) with a certain directivity characteristic. The delay time τ_i $(i=A,B)$ between the diffraction-free signal and the diffraction contributions inside the aperture can be written as:

$$\tau_i = \frac{r_i + \Delta r_i - \hat{r}}{c} \qquad i = A, B, \tag{4.28}$$

in which $\hat{r}$ is the distance between the notional source and the receiver.

From eqs. (4.25a,b) it is clear that this approximation fails when the stationary point x_0, defined by

$$\phi'(x_0) = 0 \qquad \text{with} \qquad \phi''(x_0) \neq 0 \tag{4.29}$$

lies on or in the neighborhood of x_A or x_B. For this situation $\varphi_{\text{inc},A} = \varphi_A$ or $\varphi_{\text{inc},B} = \varphi_B$, yielding an unbounded amplitude in eqs. (4.27a,b). This will be illustrated by the following example in which the diffraction artifacts are estimated by the method described above. The results for two frequencies will be compared to the exact solution.
Again, consider the configuration of figure 4.5. The synthesized wave field <I>, the diffraction-free wave field I_0, and the exact diffraction term (I_A+I_B), given by eq.(4.21), are plotted in figure 4.13a for f=150 Hz. Figure 4.13b shows the results for f=1500 Hz.

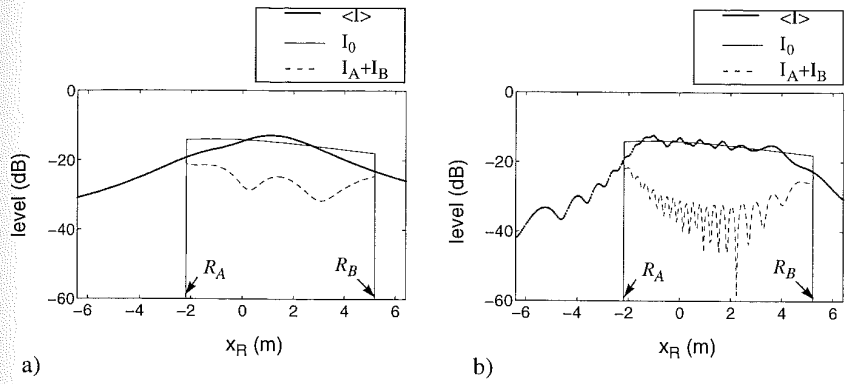


Figure 4.13: Results of the exact calculation of the diffraction contribution. The graphs give the level as a function of the receiver position, for two different frequencies.
a) <I>, I_0 and (I_A+I_B) for f=150 Hz.
b) <I>, I_0 and (I_A+I_B) for f=1500 Hz.

Next, the diffraction contributions $\tilde{I}_A$ and $\tilde{I}_B$ are calculated according to eqs. (4.27a,b) and displayed in figure 4.14a. For comparison, the approximated sum $(\tilde{I}_A + \tilde{I}_B)$ and the exact diffraction contribution $(I_A + I_B)$, are shown in figure 4.14b. In figure 4.14c-d the results are shown for f=1500 Hz.
Looking at these figures it can be observed that indeed near the light-dark transitions, indicated with R_A and R_B, the amplitude of the approximated diffraction terms $\tilde{I}_A$ and $\tilde{I}_B$ becomes unbounded. Moving away from these points the approximation becomes more accurate. It can be noted that for high frequencies the diffraction is weaker than for low frequencies. Furthermore the diffraction artifact estimation is accurate in a larger area for high frequencies than for low frequencies, which was expected because the asymptotic expansion method is a high-frequency approximation.

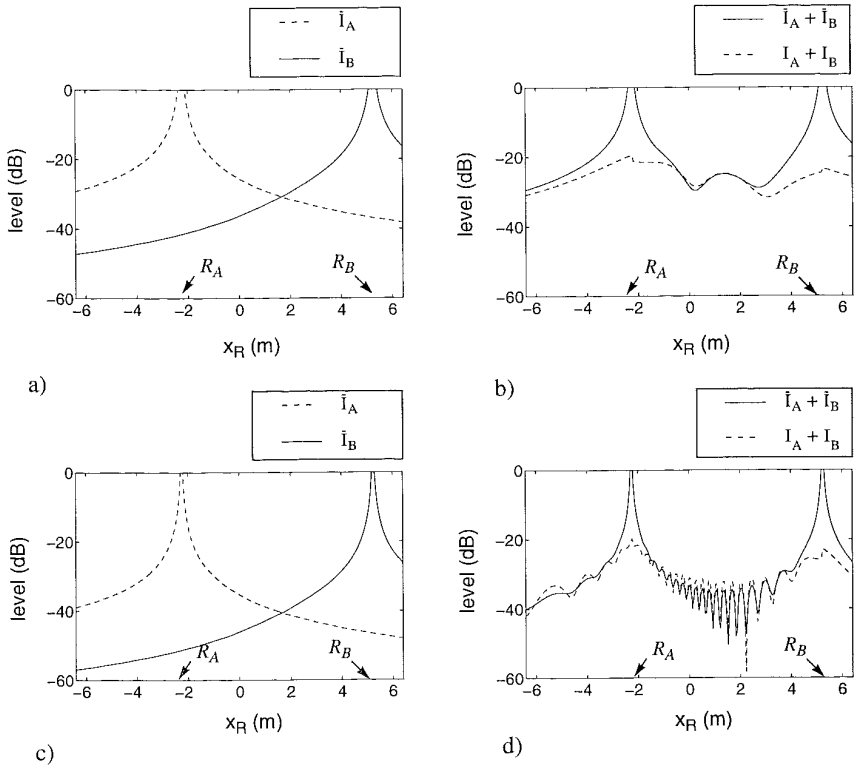


Figure 4.14: Results of the first order asymptotic expansion for the diffraction artifacts compared to the exact results. The graphs give the level as a function of the receiver position, for two different frequencies.
a) $\check{I}_A$ and $\check{I}_B$ for f=150 Hz.
b) (I_A+I_B) and $(\check{I}_A+\check{I}_B)$ for f=150 Hz.
c) $\check{I}_A$ and $\check{I}_B$ for f=1500 Hz.
d) (I_A+I_B) and $(\check{I}_A+\check{I}_B)$ for f=1500 Hz.

4.3.2 Fresnel approximation of the diffraction terms

In order to get a description of the diffraction effects which is also valid in the neighborhood of a light-dark transition, we apply a so-called 'Fresnel' approximation to the integrals I_A and I_B in eq.(4.20a-c). In this Fresnel approximation the phase $\phi(x_L)$ in the diffraction terms I_A and I_B is expanded around $x_L=x_i$ $(i=A,B)$ upto the second order in a Taylor series, according to

$$\phi(x_L) \approx \phi(x_i) + (x_L - x_i)\phi'(x_i) + \tfrac{1}{2}(x_L - x_i)^2\phi''(x_i) \qquad i = A, B. \tag{4.30a}$$

or

$$\phi(x_L) \approx \phi(x_i) + \mu_i(\xi^2 - \xi_i^2) \qquad i = A, B \tag{4.30b}$$

where

$$\mu_i = \operatorname{sgn}\phi''(x_i) \qquad i = A, B, \tag{4.30c}$$

$$\xi = \sqrt{\tfrac{1}{2}|\phi''(x_i)|}\left\{x_L - x_i + \frac{\phi'(x_i)}{\phi''(x_i)}\right\} \qquad i = A, B \tag{4.30d}$$

and

$$\xi_i = \xi_{x_L = x_i} = \frac{\mu_i\phi'(x_i)}{\sqrt{2|\phi''(x_i)|}} \qquad i = A, B. \tag{4.30e}$$

the first derivative of the phase at the end-points $\phi'(x_i)$ is given by eq.(4.26), while the second derivative $\phi''(x_i)$ at the end-points reads

$$\phi''(x_i) = -k\left(\frac{(z_L - z_\Psi)^2}{r_i^3} + \frac{(z_R - z_L)^2}{\Delta r_i^3}\right) \qquad i = A, B. \tag{4.31}$$

Note that $\operatorname{sgn}\phi''(x_i) < 0$ and consequently $\mu_i = -1$. Substitution of eq.(4.30b) into the integrals I_A and I_B in eqs.(4.20a-c) using

$$dx_L = \frac{1}{\sqrt{\tfrac{1}{2}|\phi''(x_i)|}}d\xi \qquad i = A, B. \tag{4.32}$$

yields

$$\check{I}_A \approx K(x_A)f(x_A)e^{j\phi(x_A)} \tag{4.33a}$$

$$\check{I}_B \approx -K(x_B)f(x_B)e^{j\phi(x_B)} \tag{4.33b}$$

where

$$K(x_i) = \operatorname{sgn}(\xi_i)\sqrt{\frac{\pi}{|\phi''(x_i)|}}e^{j\xi_i^2}\left[\frac{1}{2} - C_1(|\xi_i|) - j\left\{\frac{1}{2} - S_1(|\xi_i|)\right\}\right], \tag{4.34}$$

C_1 and S_1 being the Fresnel integrals (Abramowitz, 1965). When the Fresnel integrals are plotted along the axes in the complex plane with ξ_i as a parameter, the 'spiral of Cornu' is obtained, as shown in figure 4.15. In this figure the behavior of C_1 and S_1 can easily be determined.
Note that eqs.(4.33a,b) have the same form as eqs.(4.25a,b): the diffraction terms I_A and I_B are approximated by the integrand $f(x_i)\exp(j\phi(x_i))$ at the end-points x_i $(i=A,B)$ multiplied by a cor-

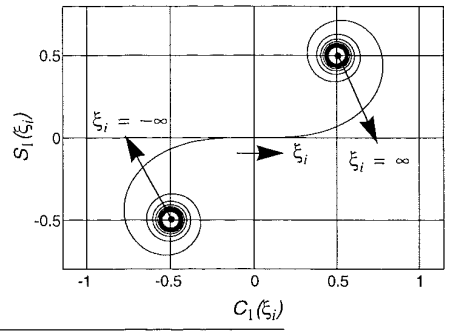


Figure 4.15: *Spiral of Cornu, given by the complex function $C_1(\xi_i)+jS_1(\xi_i)$ with ξ_i as a parameter.*

rection factor $\pm K(x_i)$. Although all variables are suppressed here except x_i, remind that the amplitude $f(x_i)$, the phase $\phi(x_i)$ and their derivatives are also functions of the receiver position.
In order to determine the accuracy of this Fresnel approximation, we will consider again the configuration of figure 4.5. The diffraction contributions $\tilde{I}_A$ and $\tilde{I}_B$ are calculated according to eqs. (4.33a,b) and displayed in figure 4.16a. For comparison, the approximated sum $(\tilde{I}_A + \tilde{I}_B)$ and the exact diffraction contribution $(I_A + I_B)$, are shown in figure 4.16b. In figure 4.16c-d the results are shown for f=1500 Hz.
Note that with the Fresnel approximation the diffraction artifacts can be calculated rather accurately for low frequencies. For high frequencies the calculated and exact diffraction effects almost match perfectly.
Evaluating eqs. (4.33a,b) at the light-dark transitions R_A and R_B reveals that the amplitude of the diffracted wave fields I_A and I_B is half the amplitude of the diffraction-free wave field I_0. Just inside the aperture, near R_A and R_B, the diffracted wave fields are in counter-phase with I_0, outside the aperture the phases of the diffracted wave fields are reversed. The result is that inside the aperture, near R_A and R_B, the amplitude of the synthesized wave field <I> is 6 dB reduced in comparison to the diffraction-free wave field I_0. This can be verified in figure 4.13, were the exact results were plotted. Also note that the amplitude in the synthesized wave field outside the aperture decreases more rapidly for high frequencies than for low frequencies.
The broadband wave field <I>, synthesized with the finite array, was shown in the x-t diagram of figure 4.6. Figure 4.17a shows the approximated diffraction wave fronts $\tilde{I}_A + \tilde{I}_B$, calculated with the Fresnel approximation. By subtracting this result from the synthesized wave field <I>, the wave field of figure 4.17b is obtained. The diffraction wavelets are almost completely vanished and the amplitude within the aperture is restored.

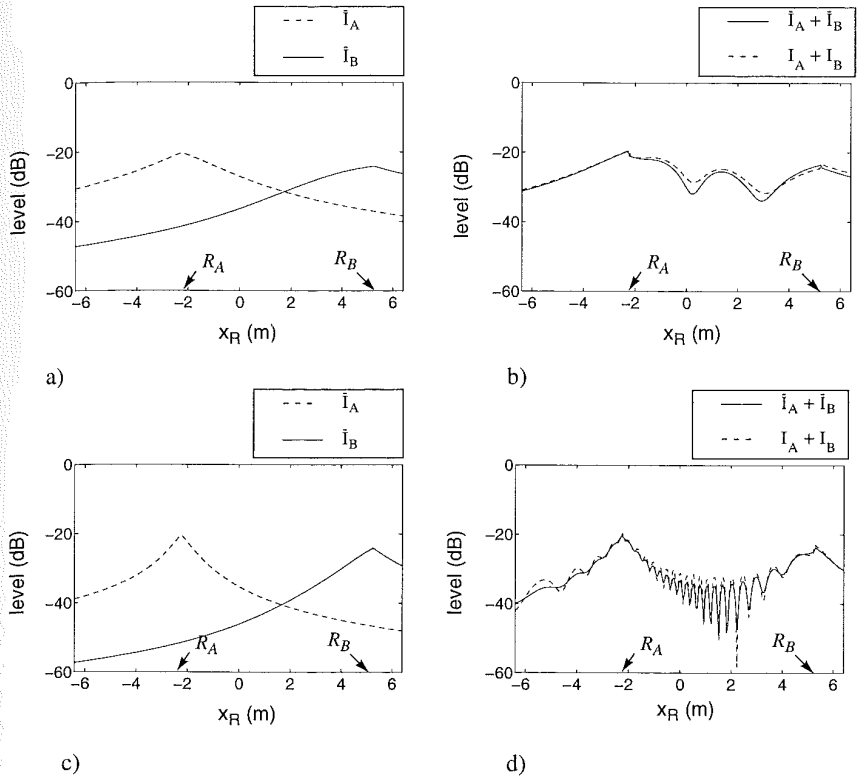


Figure 4.16: *Results of the Fresnel approximation for the diffraction artifacts compared to the exact results. The graphs give the level as a function of the receiver position, for two different frequencies.*
a) $\tilde{I}_A$ and $\tilde{I}_B$ for f=150 Hz.
b) (I_A+I_B) and $(\tilde{I}_A + \tilde{I}_B)$ for f=150 Hz.
c) $\tilde{I}_A$ and $\tilde{I}_B$ for f=1500 Hz.
d) (I_A+I_B) and $(\tilde{I}_A + \tilde{I}_B)$ for f=1500 Hz.

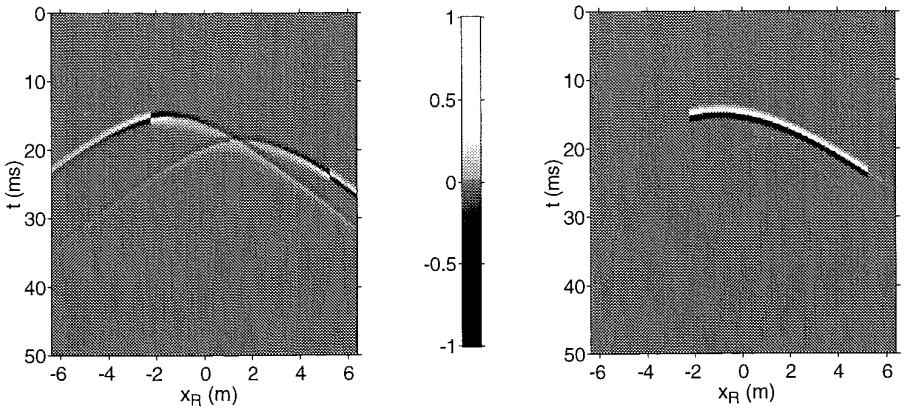


Figure 4.17: Diffraction artifacts due to the finite size of a monopole source array -1.5 ≤ x_L ≤1.5 m.
a) x-t diagram of the diffraction artifacts $\check{I}_A + \check{I}_B$ calculated by the Fresnel approximation
b) x-t diagram of the synthesized wave field p corrected for the diffraction artifacts ($\langle I \rangle - \check{I}_A - \check{I}_B$).

5

Discretization of the synthesis operators

In this chapter we will explore the performance of discretized secondary source arrays. We will investigate under which conditions it is possible to achieve an accurate representation of the primary wave field. We will start with the analysis of discretized linear arrays (straight line arrays). It will be shown that a reduction of the spatial bandwidth of the wave field improves the performance of a discrete linear array. Therefore, an important aspect of this chapter will be the development of a spatial bandwidth reduction technique for wave field synthesis applications. This technique will put new demands on the directivity characteristics of the secondary sources (loudspeakers). Further, it will be demonstrated that the proposed spatial bandwidth reduction technique can be elegantly combined with the application of bent line arrays.

5.1 Wave field synthesis in the wave number domain

A very important tool in the spatial analysis of wave fields is the spatial Fourier transform. Also, for a full understanding of the effects of spatial sampling, the spatial Fourier transform is indispensable. Therefore, first a short review of the definitions and interpretations of the spatial Fourier transform are given, which can also be found in many textbooks on multidimensional digital signal processing (e.g. Dudgeon, 1984).

5.1.1 The spatial Fourier transform and plane wave decomposition

In section 2.1 the temporal forward and inverse Fourier transform were introduced by eqs. (2.4) and (2.5). In addition to the temporal Fourier transform, also a spatial Fourier transformation

can be defined. A pressure recording $P(x,z_0,\omega)$ at a certain line $z=z_0$ in the plane $y=0$ along the x-axis can be spatially Fourier transformed to a signal $\tilde{P}(k_x, z_0, \omega)$ according to

$$\tilde{P}(k_x, z_0, \omega) = \int_{-\infty}^{\infty} P(x, z_0, \omega) e^{jk_x x} dx \tag{5.1}$$

in which k_x is the variable in the Fourier domain. The inverse is given by

$$P(x, z_0, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{P}(k_x, z_0, \omega) e^{-jk_x x} dk_x \tag{5.2}$$

The physical importance of eq.(5.1) and (5.2) and the physical meaning of the Fourier variable k_x will be made clear in the following example.
Consider a two-dimensional impulsive plane wave in the x-z domain, traveling under an angle θ with the x-axis with propagation velocity c. A snapshot in the x-z domain of the plane wave at a certain time t_0 is shown in figure 5.1a. This plane wave can be described mathematically in the space-time domain by

$$p(x, z, t) = s\left(t - \frac{x}{c_x} - \frac{z}{c_z}\right), \tag{5.3}$$

in which $s(t)$ is the source signature and $c_x=c/\sin(\theta)$ and $c_z=c/\sin(\theta)$ are the phase velocities along the x-axis and the z-axis respectively. By making a x-t registration of this wave field along

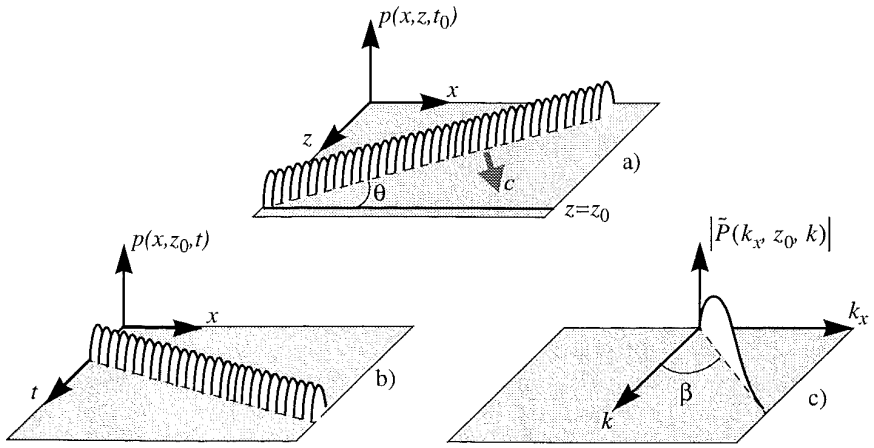


Figure 5.1: *An impulsive traveling plane wave making an angle θ with the x-axis.*
a) Snapshot in the x-z domain of the pressure p at a certain time t_0
b) x-t diagram of the pressure registration along the line $z=z_0$ parallel to the x-axis.
c) k_x-k diagram of the pressure registration the line $z=z_0$ parallel to the x-axis. The amplitude $|\tilde{P}|$ of the spatial spectrum is zero everywhere, except on the line $k_x = k\sin(\theta) = k\tan(\beta)$

the line $z=z_0$ parallel to the x-axis the x-t diagram of figure 5.1b is obtained. Applying a Fourier transform from time to frequency yields

$$P(x, z_0, \omega) = S(\omega) e^{-j\omega x/c_x} e^{-j\omega z_0/c_z}, \tag{5.4}$$

where $S(\omega)$ is the source spectrum. Note that, for a certain value of ω, the phase velocity c_x and the apparent wavelength $\lambda_x=\lambda/\sin(\theta)$ along the x-axis are dependent of the propagation angle θ, which is illustrated in figure 5.2.

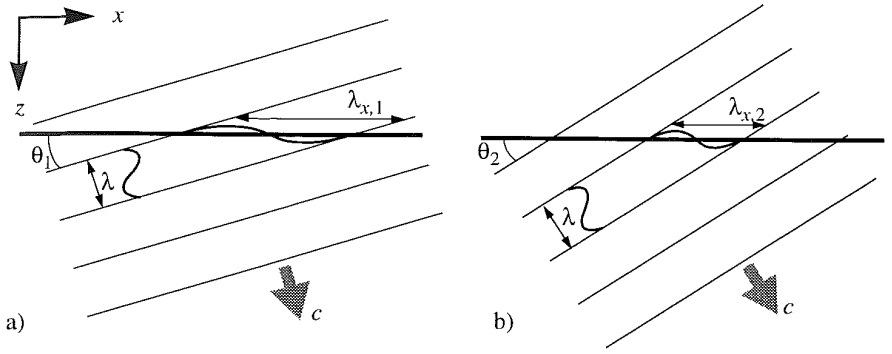


Figure 5.2: *x-z domain representation of a monochromatic plane wave with angular frequency ω and wavelength λ.*
a) propagation angle θ_1, apparent wavelength $\lambda_{x,1}=\lambda/\sin(\theta_1)$ along the x-axis.
a) propagation angle θ_2, apparent wavelength $\lambda_{x,2}=\lambda/\sin(\theta_2)$ along the x-axis.

Now we define the spatial frequency $k_x=2\pi/\lambda_x$, which is the spatial equivalent of the temporal (angular) frequency $\omega=2\pi/T$, in which T is the period. The spatial frequency k_x and the temporal frequency ω are related by $k_x=\omega/c_x$.
To proof that the Fourier variable k_x and the spatial frequency k_x are identical, the spatial Fourier transform (5.1) is applied to eq. 5.4, yielding

$$\tilde{P}(k_x, z_0, \omega) = 2\pi S(\omega) e^{-j\omega z_0/c_z} \delta\left(k_x - \frac{\omega}{c_x}\right). \tag{5.5}$$

Since the data in any domain should represent a a non-zero pressure distribution, the delta function must have a non-zero value. Hence, the Fourier variable k_x equals ω/c_x which is identical to the spatial frequency k_x. Usually the temporal frequency ω is replaced by the wave number $k=\omega/c$, such that can be written

$$k_x = \frac{\omega}{c_x} = \frac{\omega}{c}\sin(\theta) = k\sin(\theta) \tag{5.6}$$

The k_x-k diagram of the double Fourier transform $\tilde{P}$ of the plane wave is shown in figure 5.1c. The amplitude $|\tilde{P}|$ of the spatial spectrum $\tilde{P}$ is zero everywhere, except on the line $k_x = k\tan(\beta)$. The angle β in the k_x-k domain is related to the angle θ in the x-z domain by

$$\tan(\beta) = \frac{k_x}{k} = \sin(\theta) \tag{5.7}$$

Using these results, eq.(5.2) can be interpreted as the composition along the x-axis of the wave field $P(x,z_0,\omega)$ in terms of monochromatic plane waves $\tilde{P}(k_x, z_0, \omega)e^{-jk_x x}$. Hence, eq. (5.1) may be interpreted as a decomposition of the wave field $P(x,z_0,\omega)$ into monochromatic plane waves.

Since, as a Fourier variable, k_x can assume all values in the interval $(-\infty,\infty)$, the wave field $p(x,z_0,\omega)$ is decomposed in both travelling waves ($|k_x|<k$) and evanescent waves ($|k_x|>k$). The effect of evanescent waves can be neglected in the far field of the acoustical sources.

5.1.2 Transformation of the 2½D Rayleigh integrals to the wave number domain

In chapter 3 the 2½D Rayleigh I and II integrals were derived in the space-frequency domain, and can be written as

$$P(x_R) = \int_{-\infty}^{\infty} Q(x_L)W(x_R - x_L)dx_L \tag{5.8}$$

or, in convolution notation:

$$P(x_R) = Q(x_R) * W(x_R). \tag{5.9}$$

Q is the driving signal for the secondary source array and W is the secondary source response, defined by eqs.(4.10a,b) and (4.11a,b) for the Rayleigh I and II integral respectively.

The driving signal Q and the secondary source response W can be Fourier transformed with respect to x_R, so that the convolution of Q and W may be replaced by multiplication of the Fourier transforms $\tilde{Q}$ and $\tilde{W}$:

$$\tilde{P}(k_x) = \tilde{Q}(k_x)\tilde{W}(k_x), \tag{5.10}$$

in which $\tilde{P}$ is the spatial Fourier transform of the synthesized wave field P.

In order to find analytical expressions for the spatial Fourier transforms $\tilde{Q}$ and $\tilde{W}$, we introduce the following Fourier transform pair (Wapenaar, 19xx):

$$F_\alpha(x) = \frac{e^{-jkr}}{r^\alpha} \qquad \text{for } kr>>1 \tag{5.11a}$$

and

$$\tilde{F}_\alpha(k_x) = \frac{\sqrt{2\pi}}{k^{\alpha-1}} \frac{k_z^{\alpha-3/2}}{|\Delta z|^{\alpha-1/2}} e^{-j(k_z|\Delta z| + \pi/4)}, \tag{5.11b}$$

in which $r = \sqrt{x^2+\Delta z^2}$, $k_z = \sqrt{k^2-k_x^2}$ for $k_x^2 \le k^2$ and $k_z = -j\sqrt{k_x^2-k^2}$ for $k_x^2 \ge k^2$.

Using this Fourier transform pair, the spatial Fourier transforms of the monopole driving function Q and the secondary monopole response W are given by

$$\tilde{Q}(k_x) = \sqrt{\frac{z_R - z_L}{z_R - z_\Psi}} S(\omega) e^{-jk_z(z_L - z_\Psi)}, \tag{5.12a}$$

$$\tilde{W}(k_x) = \sqrt{\frac{2\pi}{jk_z(z_R - z_L)}} e^{-jk_z(z_R - z_L)}. \tag{5.12b}$$

Insertion of (5.12a,b) in eq.(5.10) yields

$$\tilde{P}(k_x) = \sqrt{\frac{2\pi}{jk_z(z_R - z_\Psi)}} e^{-jk_z(z_R - z_\Psi)}. \tag{5.13}$$

With eqs.(5.11a,b) it can be verified that this is exactly the spatial Fourier transform of the notional monopole wave field registered along the line $z=z_R$ in the plane $y=0$, which we expected.

5.1.3 Wave field synthesis with a continuous linear array

As an example consider the situation shown in figure 5.3, which is identical to the configuration

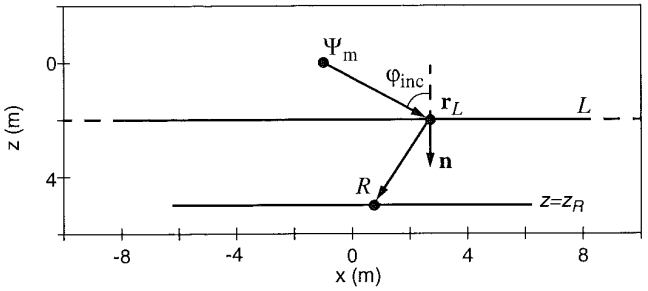


Figure 5.3: Diagram of the geometry of a notional monopole source Ψ_m at $\mathbf{r}_\Psi$=(-1,0) and an infinite continuous secondary monopole point source distribution along line L. The primary and the synthesized wave field are recorded at the line $z=z_R$=5m for -6.4<x_R<6.4m.

of figure 4.1 in the previous chapter. The wave field of a notional monopole source Ψ_m at position $\mathbf{r}_\Psi$ =(-1,0) is synthesized by applying the 2½D Rayleigh I integral (4.6). The secondary monopole point sources are continuously distributed along the line L of infinite length (z_L=2m). The notional source emits a bandlimited impulse (0-1.7 kHz), as shown in figure 5.4a and b. The synthesized wave field p and the primary wave field p_{pr} are recorded at the line $z=z_R$=5m for -6.4<x_R<6.4m. The x-t diagram of the primary wave field p_{pr} is shown in figure 5.4c.

The k_x-k diagrams of the secondary monopole response $\tilde{W}$ at the line $z=z_R$ and the driving signal $\tilde{Q}$ at the line L are shown in figure 5.5a and b. Only the amplitude is depicted, indicated

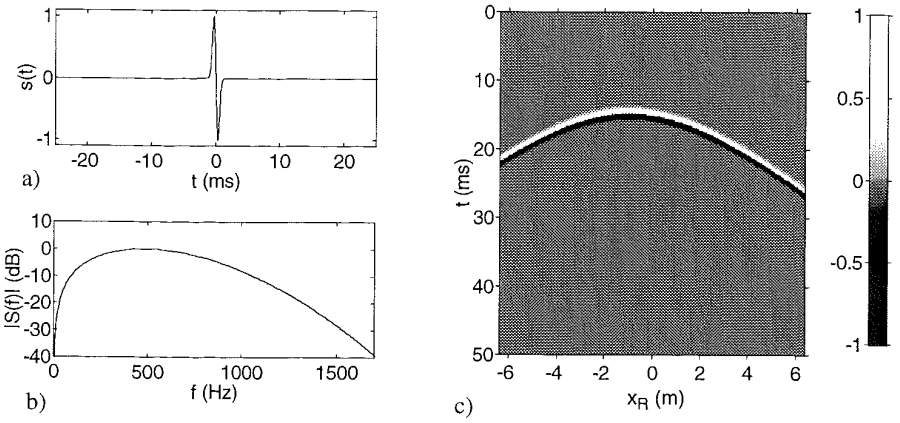


Figure 5.4: *Primary wave field p_{pr} emitted by the notional monopole source.*
a) source signature s(t)
b) amplitude spectrum of the source signature, bandwidth 0-1.7 kHz.
c) x-t diagram

with the greyness-level on a scale from 0 to 1 as shown below the graphs. The k_x-k diagram of the synthesized wave field $\tilde{P}$, which is the multiplication of $\tilde{W}$ and $\tilde{Q}$, is shown in figure 5.5c. Applying a double inverse Fourier transform to the synthesized wave field $\tilde{P}$ results in the x-t diagram of figure 5.6a. The synthesis error $e(x,t)$ is depicted in figure 5.6b. As expected, the synthesized wave field matches well the primary wave field.

From figure 5.5a and b it is clear that the main part of the wave energy is contained in the area $|k_x|<k$, i.e. the effect of evanescent waves can be neglected. Using eq. 5.6, it follows that the secondary monopole response $\tilde{W}$ and the driving signal $\tilde{Q}$ contain plane wave components with propagation angles $-90°<\theta<90°$.

For the x-f diagram of the error level L_E and the x-f diagram of the pressure ratio level L_R we refer to figure 4.4a and b.

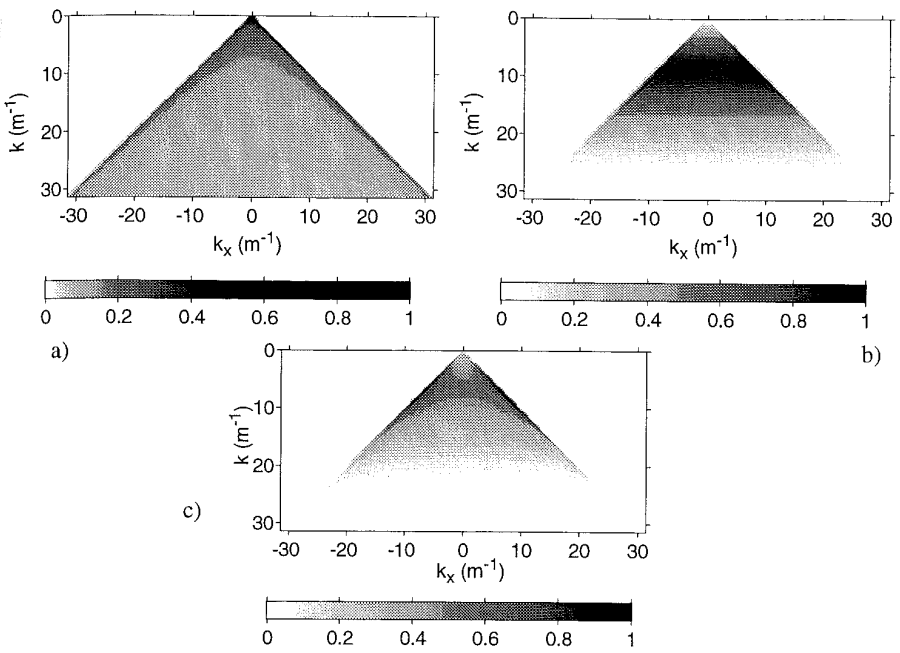


Figure 5.5: *Wave field synthesis in the spatial frequency domain with an infinite continuous monopole source array.*
a) k_x-k diagram of the secondary monopole response $\tilde{W}$
b) k_x-k diagram of the driving signal $\tilde{Q}$
c) k_x-k diagram of the synthesized wave field $\tilde{P}$

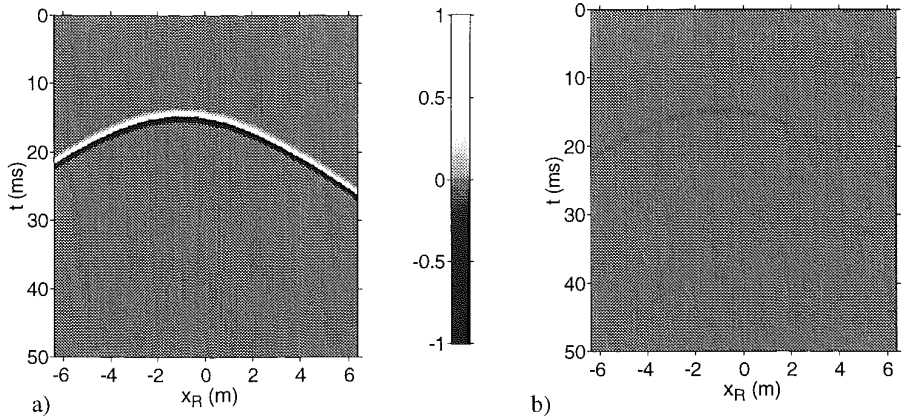


Figure 5.6: Wave field synthesis with an infinite continuous monopole source array. Registration along the line $z=z_R$ for $-6.4<x_R<6.4$. Temporal bandwidth: 0-1.7 kHz.
a) x-t diagram of the synthesized wave field p
b) x-t diagram of the synthesis error e.

5.2 Discretization of the 2½D Rayleigh integrals

With the practical implementation of a system based on wave field synthesis techniques, some adaptations must be made. The continuous secondary source distribution has to be replaced by a discrete array of loudspeakers with monopole or dipole directivity characteristics.
In this section we will explore under which conditions it is possible to synthesize a wave field with a discrete secondary source distribution.

5.2.1 Spatial sampling

The discrete version of a wave field $P(x,\omega)$, that is recorded or calculated at a certain line $z=z_0$ in the plane $y=0$, can be written as

$$P_\Delta(x,\omega) = P(x,\omega)\Delta x \sum_{n=-\infty}^{\infty} \delta(x - n\Delta x) \tag{5.14}$$

This discretization is allowed if the spatial anti-aliasing condition is fulfilled. Spatial aliasing does not occur for a spatially bandlimited signal $P(x,\omega)$ having a symmetrical spatial amplitude spectrum $\tilde{P}(k_x,\omega)$ with $\tilde{P}(k_x,\omega) = 0$ for $|k_x|>k_{x,\max}$ if

$$k_{x,\max} \le k_{x,N} = \frac{\pi}{\Delta x} \tag{5.15}$$

in which $k_{x,N}$ is the spatial Nyquist frequency. Since the spatial bandwidth is related to the temporal bandwidth via eq. (5.6), the anti-aliasing condition (5.15) can also be expressed in terms of the maximum temporal frequency $f_{\max}=k_{\max}c/2\pi$. For wave fields containing plane wave components in the range $-\theta_{\max} \le \theta \le \theta_{\max}$, the anti aliasing condition can be rewritten as

$$f_{\max} \le f_{al} = \frac{c}{2\Delta x \sin(\theta_{\max})} \tag{5.16}$$

where f_{al} is called the spatial aliasing frequency. Note that the spatial aliasing frequency depends on the maximum angle $\theta_{\max}$. This important fact will be further explored in section 5.3.
It can be easily derived that for asymmetrical spatial amplitude spectra, the anti-aliasing condition must be replaced by

$$f_{\max} \le f_{al} = \frac{c}{\Delta x[\sin(\theta_{\max}) - \sin(\theta_{\min})]}, \tag{5.17}$$

in which $\theta_{\max}$ and $\theta_{\min}$ indicate the maximum and minimum angle of incidence of the plane wave components of the wave field P respectively.
The discretized version of the convolution integral (5.8) reads

$$P(x_R) = \sum_{n=-\infty}^{\infty} Q_\Delta(x_L) W(x_R - x_L) \tag{5.18}$$

with

$$Q_\Delta(x_L) = Q(x_L)\Delta x \sum_{n=-\infty}^{\infty} \delta(x_L - n\Delta x). \tag{5.19}$$

Note that the driving signal $Q(x_L)$ has been sampled yielding $Q_\Delta(x_L)$, while $W(x_R\text{-}x_L)$, the response at the continuous receiver line $z=z_R$ of a secondary source at x_L, remains a continuous function.

Applying the spatial Fourier transform to the convolution sum of eq.(5.18) yields

$$\tilde{P}(k_x) = \tilde{Q}_p(k_x)\tilde{W}(k_x), \tag{5.20}$$

where $\tilde{Q}_p$ is the periodic spectrum of the sampled driving signal $Q_\Delta(x_L)$.

In figure 5.7 the two steps in the discrete wave field synthesis process are presented schematically in the k_x-k domain. In the first step, representing the sampling process, the spatial Fourier transform $\tilde{Q}$ of the continuous driving signal is convolved with the spatial sampling operator $\tilde{\Delta}_p$ given by

$$\tilde{\Delta}_p(k_x, \omega) = \sum_{n=-\infty}^{\infty} \delta\left(k_x - n\frac{2\pi}{\Delta x}\right) \tag{5.21}$$

yielding the periodic driving signal $\tilde{Q}_p$. In the second step this periodic driving function $\tilde{Q}_p$ is multiplied with the spatial secondary source response $\tilde{W}$, which will be called the synthesis process. In the synthesis process the periodic repetitions of the driving signal $\tilde{Q}_p$ are suppressed by the spatial secondary source response $\tilde{W}$.

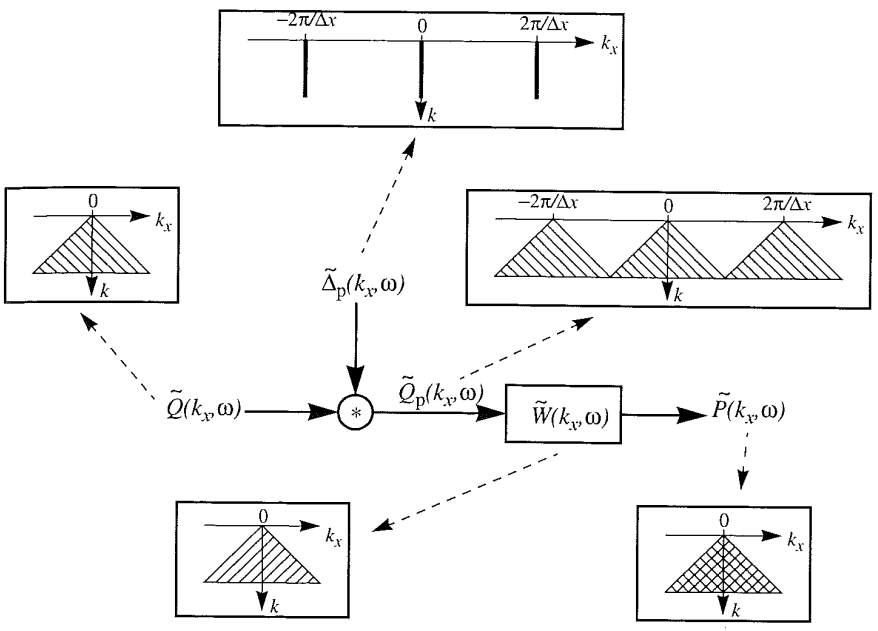


Figure 5.7: Schematical representation of the discrete wave field synthesis process in the k_x-k domain. In the first step, representing the sampling process, the spatially Fourier transformed continuous driving function $\tilde{Q}$ is convolved with the spatial sampling operator $\tilde{\Delta}_p$ yielding the periodic driving function $\tilde{Q}_p$. In the second step this periodic driving function $\tilde{Q}_p$ is multiplied with the spatial monopole response $\tilde{W}$, which will be called the synthesis process.

5.2.2 Wave field synthesis with a discrete linear array

Next consider the situation of figure 5.8. The continuous secondary monopole source distribution of figure 5.3 has been discretized with sampling interval Δx=0.1 m. For this situation, the spatial Nyquist frequency is given by $k_{x,N}=\pi/0.1=31.4$ m^{-1}. For a maximum propagation angle $\theta_{max}=90°$ of plane wave components in the driving signal $\tilde{Q}$, and a source signal containing no energy above f_{max}=1.7 kHz, it can be verified that the anti-aliasing condition (5.16) is fulfilled. Figure 5.9a shows the secondary monopole response $\tilde{W}$ at the line $z=z_R$. One period of the periodic k_x-k diagram of the discrete driving signal $\tilde{Q}_p$ at the line L is shown in figure 5.9b. The k_x-k diagram of the synthesized wave field $\tilde{P}$, which is the multiplication of $\tilde{W}$ and $\tilde{Q}_p$, is depicted in figure 5.9c. Since the anti-aliasing condition is fulfilled, the synthesized wave field $\tilde{P}$, and consequently the inverse transform p, are exactly identical to the results obtained with the continuous array, as presented in the previous section (see figure 5.6).

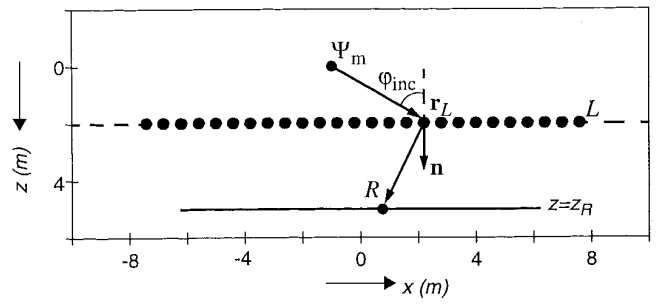


Figure 5.8: *Diagram of the geometry of a notional monopole source Ψ_m at $\mathbf{r}_\Psi$=(-1,0) and a discrete secondary monopole point source distribution along line L with sampling distance Δx=0.1 m. The primary and the synthesized wave field are recorded at the line $z=z_R$=5m for -6.4<x_R<6.4m.*

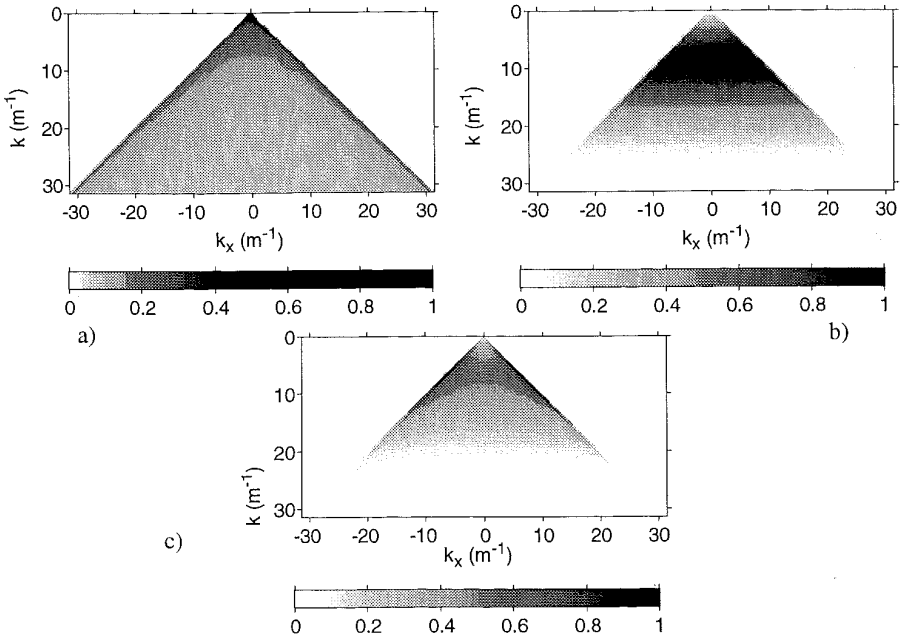


Figure 5.9: *Wave field synthesis in the spatial frequency domain with a discrete monopole source array (Δx=0.1 m).*
a) k_x-k diagram of the secondary monopole response $\tilde{W}$
b) k_x-k diagram of the driving signal $\tilde{Q}_p$ (1 period)
c) k_x-k diagram of the synthesized wave field $\tilde{P}$

5.2.3 Wave field synthesis with an undersampled linear array

Next consider the same configuration, but now with the driving signal Q sampled at half the sample rate of the previous example (Δx=0.2 m). For this situation, the spatial Nyquist frequency $k_{x,N}$ equals $\pi/0.2$=15.7 m^{-1}, and the anti-aliasing condition is not fulfilled anymore. Figure 5.10a shows the secondary monopole response $\tilde{W}$ at the line $z=z_R$. Two periods of the

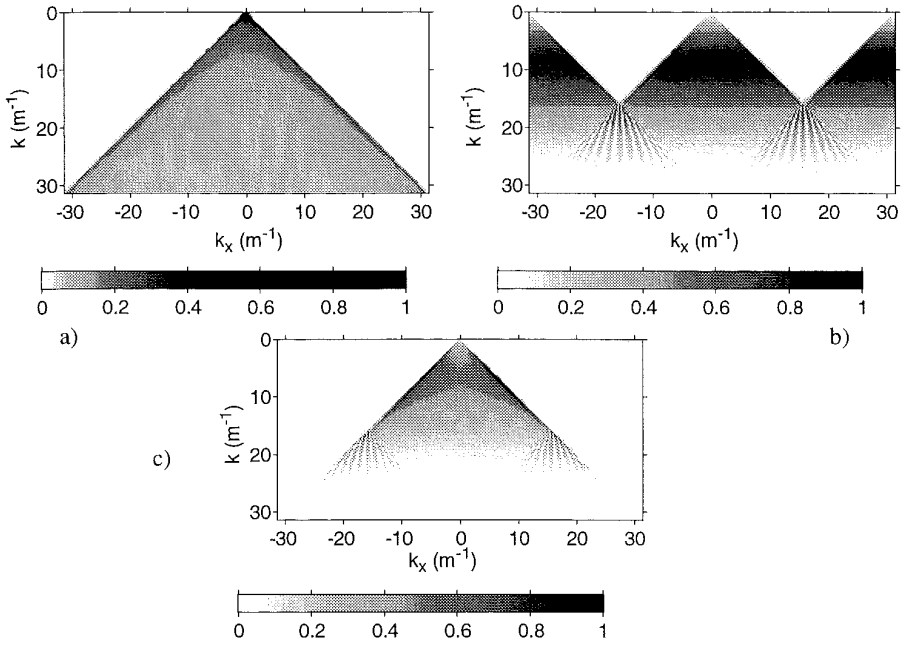


Figure 5.10: *Wave field synthesis in spatial frequency domain with an undersampled monopole array (Δx=0.2 m)*
a) k_x-k diagram of the secondary monopole response $\tilde{W}$
b) k_x-k diagram of the driving signal $\tilde{Q}_p$ (2 periods)
c) k_x-k diagram of the synthesized wave field $\tilde{P}$

periodic k_x-k diagram of the discrete driving signal $\tilde{Q}_p$ at the line L are shown in figure 5.10b, in which it is clearly visible that the repetitions of the spatial spectrum overlap each other, which represents the spatial aliasing phenomenon. The k_x-k diagram of the synthesized wave field $\tilde{P}$, which is the multiplication of $\tilde{W}$ and $\tilde{Q}_p$, is depicted in figure 5.10c.
Figure 5.11 shows the x-t diagram of the double inverse transform $p(x,t)$. The desired wave front is clearly visible, but due to spatial aliasing the pressure behind the main wave front is not equal to zero: the contributions of the secondary sources do not cancel each other behind the

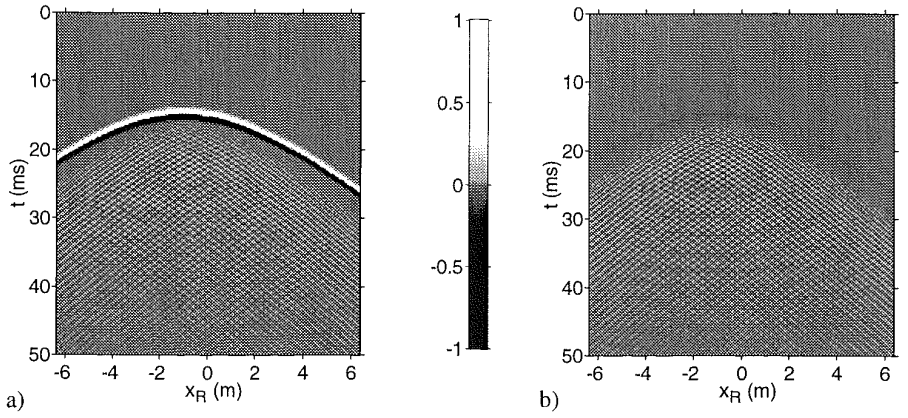


Figure 5.11: Wave field synthesis with an undersampled monopole source array (Δx=0.2 m). Registration along the line $z=z_R$ for -6.4<x_R<6.4m. Temporal bandwidth: 0-1.7 kHz.
a) x-t diagram of the synthesized wave field p
b) x-t diagram of the synthesis error e.

main wave front for $f>f_{al}$, which may lead to worse localization performance of the listeners. The synthesis error $e(x,t)$ is depicted in figure 5.11b.
The x-f diagram of the error level L_E of the synthesized wave field for -6.4<x_R<6.4m and 0<f<1.7kHz is shown in figure 5.12a. For frequencies $f<f_{al}$=850 Hz, the error level is very low,

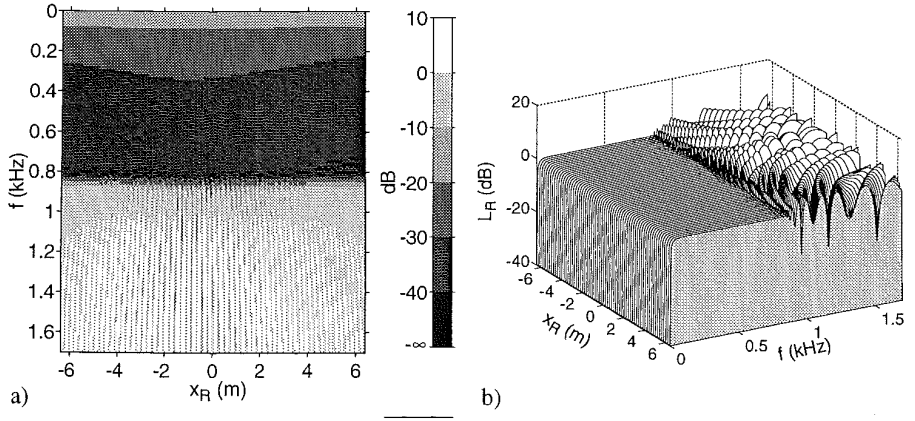


Figure 5.12: Performance measures in the x-f domain. Wave field synthesized with an undersampled monopole source array (Δx=0.2 m). Registration along the line $z=z_R$ for -6.4<x_R<6.4m. Temporal bandwidth: 0-1.7 kHz.
a) x-f diagram of the error level L_E
b) x-f diagram of the pressure ratio level L_R

which means that for those frequencies the wave field is synthesized correctly. However, for $f>f_{al}$=850 Hz, the error level increases rapidly and fluctuates around L_E=0dB, i.e., the phase and amplitude of the pressure $P(x_R,f)$ for $f>f_{al}$=850 Hz are incorrect.
Figure 5.12b shows the pressure ratio level L_R. For frequencies $f>f_{al}$ =850 Hz the spectrum is irregular, which may lead to coloration of the perceived sound.

5.3 Reduction of spatial aliasing

In the previous section it was shown under which conditions the 2½D Rayleigh integrals may be discretized. The measure of performance is determined by the required temporal bandwidth and the sampling distance of the secondary source array. Since, for a travelling wave the maximum spatial frequency $k_{x,max}$ never exceeds the maximum wavenumber k_{max}, in theory it is always possible to fulfill the spatial anti-aliasing condition (5.16). The secondary point sources can be placed arbitrarily close to each other.
In practice however, loudspeaker arrays will be used, of which the individual elements have a finite size. Consequently, the individual elements cannot be placed arbitrarily close to each other anymore, yielding a lower limit for the sampling distance Δx. Besides this physical restriction on the sampling distance, there may be also technical and economical restrictions, e.g., a limited number of channels that can be processed by a certain hardware configuration.
According to the results of the previous section, undersampling of the secondary source array results in a distortion of the spatial and temporal properties of the synthesized wave field above the spatial aliasing frequency. Therefore, in this section a method will be presented, which suppresses spatial aliasing, given a certain sampling distance Δx.

5.3.1 Spatial bandwidth reduction in wave field synthesis applications

Eq.(5.16) shows that a reduction of the maximum propagation angle θ_{max} will give an increase of the spatial aliasing frequency f_{al}. This means that, given a certain sampling distance Δx, the spatial bandwidth can be reduced in favor of the temporal bandwidth. This property will be further explored and implemented in the wave field synthesis process.
In the previous section the two steps in the discrete wave field synthesis process were presented: the sampling step and the synthesis step. If the sampling distance is too large to avoid spatial aliasing, given the spatial bandwidth of the driving function Q, we must reduce the spatial bandwidth of the driving function Q before it is spatially sampled. This step may be considered as spatial anti-aliasing filtering.
The discrete wave field synthesis process can now be divided into four steps:

I. Spatial anti-aliasing filtering
II. Spatial sampling
III. Spatial synthesis
IV. Spatial reconstruction filtering

Each of these steps will be discussed more extensively and visualized schematically in the k_x-k domain.

I. Since, inherent to the notional source concept, the position of the notional source is known a priori, an analytical expression in the continuous x-z domain is available for the driving signal Q (see eq.(4.10a) or (4.11a)). By applying the spatial Fourier transform, the k_x-k domain representation $\tilde{Q}$ can be obtained, as shown in figure 5.13a. Suppose that, for a given temporal bandwidth (0-k_{max}) of the notional source signal, the sampling interval Δx is too large to avoid spatial aliasing. Next, a spatial low-pass filter $\tilde{H}$ with spatial cutoff frequency $k_x^{co} = \pi/\Delta x$ is applied, which is schematically depicted in figure 5.13b. The spatially filtered driving signal $\tilde{Q}^H$, which is the multiplication of $\tilde{Q}$ and $\tilde{H}$, is shown in figure 5.13c. The spatial bandwidth of the driving signal is now reduced; no aliasing will occur after sampling.

II. The spatially filtered driving function Q^H, which is the spatial inverse Fourier transform of $\tilde{Q}^H$, can be sampled at intervals Δx, the distance between the secondary sources. In the k_x-k domain, sampling can be represented by a convolution with the spatial sampling operator $\tilde{\Delta}_p$, given by eq. (5.21). The k_x-k diagram of $\tilde{\Delta}_p$ is shown in figure 5.13d.
The spatially bandlimited and sampled driving signal can be represented in the k_x-k domain by the periodic driving signal $\tilde{Q}_p^H$, as shown in figure 5.13e. The repetitions do not overlap each other, i.e. spatial aliasing is avoided.

III. The spatial synthesis can be described by the multiplication of the spatially bandlimited driving signal $\tilde{Q}_p^H$ with the secondary source response $\tilde{W}$, which is shown in figure 5.13f. The synthesized wave field $\tilde{P}$ is shown in figure 5.13g.

IV. In order to remove the grating lobes (the repetitions in the k_x-k domain), the synthesized wave field $\tilde{P}$ is filtered with a spatial reconstruction filter $\tilde{H}_r$, which has the same shape as the anti-aliasing filter $\tilde{H}$. The spatially bandlimited synthesized wave field $\tilde{P}^H$ is shown in figure 5.13h.

5.3.2 Spatial bandwidth reduction in wave field synthesis: an idealized example

The spatial bandwidth reduction technique, developed in the previous subsection, will be applied to the configuration that was investigated in the previous example. The individual effects of each step in the processing scheme of figure 5.13 will be extensively analyzed.
Consider again the configuration discussed in section 5.2.3. The notional monopole source emits a bandlimited impulse (0-1.7 kHz), as shown in figure 5.4a and b. In figure 5.4c the x-t diagram of the primary wave field p_{pr} is depicted. The sampling distance Δx, i.e. the distance between the individual secondary monopole sources, equals 0.2 m. For this situation, the spatial Nyquist frequency $k_{x,N}$ equals $\pi/0.2 = 15.7$ m^{-1}. The k_x-k diagram of the continuous driving function Q is shown in figure 5.14a, which is identical to figure 5.5b. It is evident that the anti-aliasing condition cannot be fulfilled, since $k_{x,max} = k_{max}\sin(90°) = 31.4$ m^{-1}, twice the value of the spatial Nyquist frequency $k_{x,N}$.

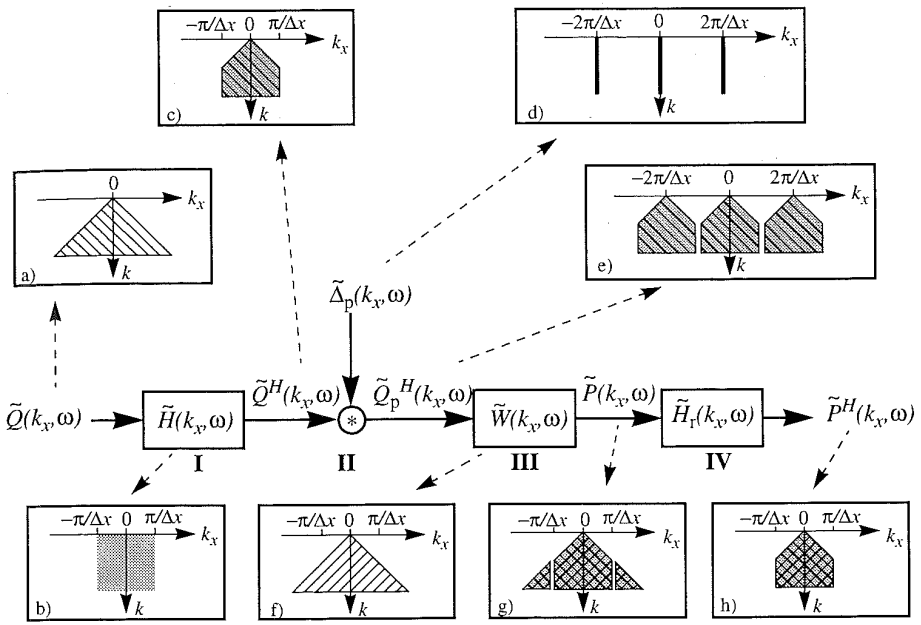


Figure 5.13: *General processing scheme in the k_x-k domain for discrete wave field synthesis applications, consisting of four steps: spatial anti-aliasing filtering (**I**), spatial sampling (**II**), spatial synthesis (**III**) and spatial reconstruction filtering (**IV**). The eight blocks give the k_x-k domain representations of the:*

a) continuous driving function $\tilde{Q}$
b) spatial low-pass filter $\tilde{H}$
c) spatially bandlimited driving signal $\tilde{Q}^H$
d) sampling operator $\tilde{\Delta}_p$
e) spatially bandlimited and sampled driving signal $\tilde{Q}_p^H$
f) secondary source response $\tilde{W}$
g) synthesized wave field $\tilde{P}$
h) spatially bandlimited synthesized wave field $\tilde{P}^H$

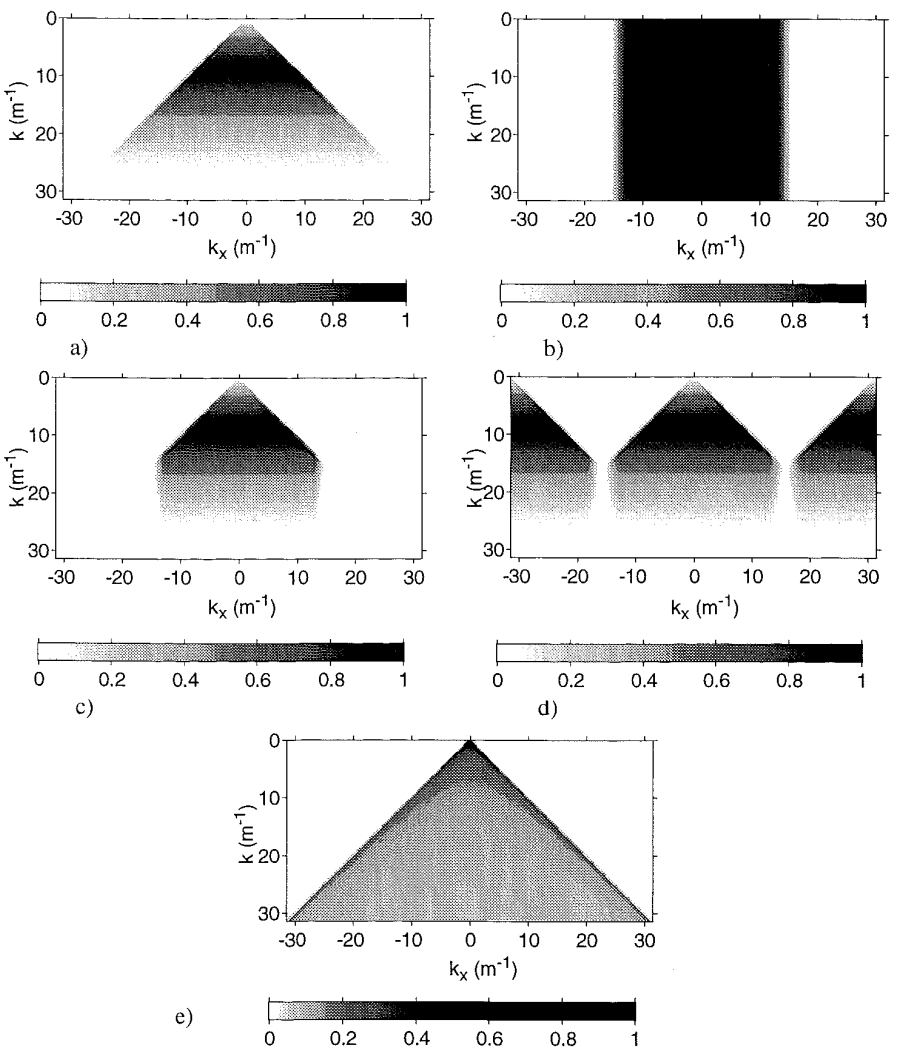


Figure 5.14: The bandwidth reduction process applied to wave field synthesis. In the pictures the k_x-k diagrams of the following function are shown:
a) continuous driving function $\tilde{Q}$
b) spatial low-pass filter $\tilde{H}$
c) spatially bandlimited driving signal $\tilde{Q}^H$
d) spatially bandlimited and sampled driving signal $\tilde{Q}_p^H$
e) secondary source response $\tilde{W}$

According to the processing scheme of figure 5.13 a spatial low-pass filter $\tilde{H}$ will be applied to the driving signal $\tilde{Q}$ with a cutoff frequency $k_x^{co}=\pi/\Delta x=15.7\ m^{-1}$, as shown in figure 5.14b. This filter can be interpreted as a frequency-independent spatial filter. The filtered driving signal $\tilde{Q}^H$ is presented in figure 5.14c. Note that for low frequencies the spatial bandwidth of the filtered driving signal $\tilde{Q}^H$ is unaffected. For the highest frequency k_{max} the maximum spatial frequency $k_{x,max}=k_{max}/2$, yielding a spatial bandwidth determined by $-30°<\theta<30°$.
The periodic spectrum $\tilde{Q}_p^H$ of the sampled version of the filtered driving signal $\tilde{Q}^H$ is shown in figure 5.14d. Due to the anti-aliasing filter $\tilde{H}$ the repetitions in the k_x-k domain of the periodic spectrum $\tilde{Q}_p^H$ do not overlap each other; spatial aliasing is avoided.
In the spatial synthesis process, the periodic driving signal $\tilde{Q}_p^H$ is multiplied with the secondary source response $\tilde{W}$ (figure 5.14e). In this way the spatial spectrum of the synthesized wave field $\tilde{P}$ is obtained (figure 5.15a). From this diagram it is clear that omitting the reconstruction filter

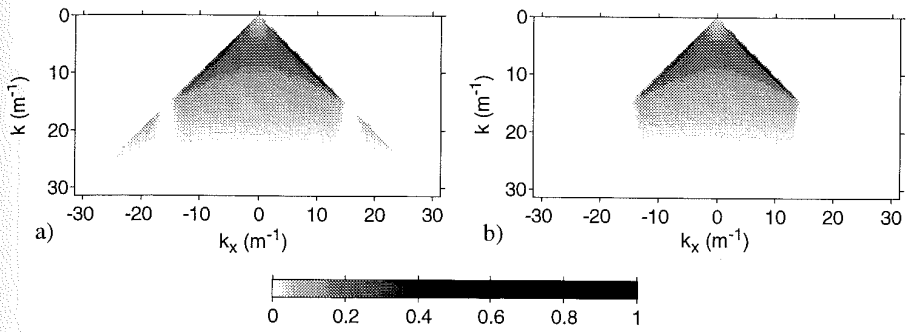


Figure 5.15: Influence of the reconstruction filter on the synthesized wave field in the k_x-k domain.
a) k_x-k diagram of the synthesized wave field $\tilde{P}$ without applying a spatial reconstruction filter
b) k_x-k diagram of the synthesized wave field $\tilde{P}^H$ using a spatial reconstruction filter

gives rise to false spatial high-frequency contributions in the synthesized wave field.
To remove the non-evanescent grating lobes in the spatial spectrum, the synthesized wave field $\tilde{P}$ must be filtered with the reconstruction filter $\tilde{H}_r$, which should have the same shape as the anti-aliasing filter of figure 5.14b. After applying the spatial reconstruction filter, the spatially bandlimited synthesized wave field $\tilde{P}^H$ is obtained (figure 5.15b).
Inverse Fourier transformation of the non-reconstructed synthesized wave field $\tilde{P}$ gives the x-t diagram of the synthesized wave field p, which is shown in figure 5.16a. The synthesis error $e(x,t)$ is depicted in figure 5.16b. The unwanted spatial contributions can be observed in the synthesized wave field as extra high-frequency wave fronts behind the main wave front.
Inverse Fourier transformation of the spatially reconstructed synthesized wave field $\tilde{P}^H$ gives the x-t diagram of the spatially bandlimited synthesized wave field p^H, which is shown in figure 5.16c. The synthesis error $e(x,t)$ is shown in figure 5.16d.

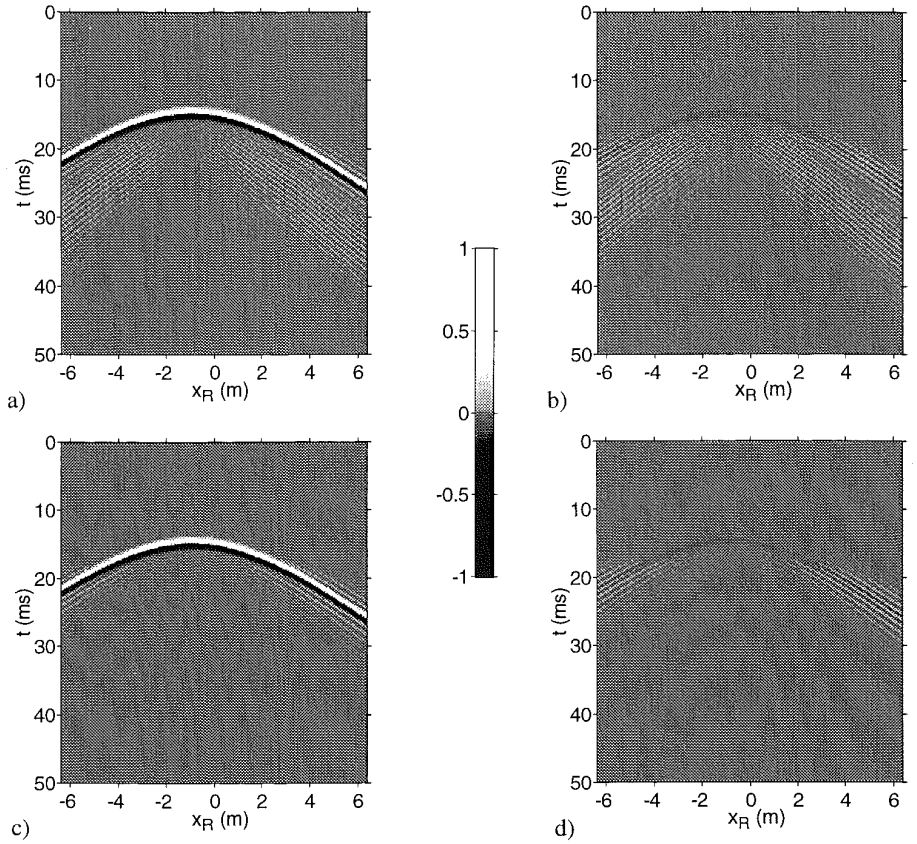


Figure 5.16: *Influence of the reconstruction filter on the synthesized wave field in the x-t domain.*
a) x-t diagram of the synthesized wave field p, without applying a reconstruction filter
b) x-t diagram of the synthesis error e, without applying a reconstruction filter
c) x-t diagram of the synthesized wave field p^H, using a spatially bandlimited reconstruction filter
b) x-t diagram of the synthesis error e, using a spatially bandlimited reconstruction filter

In figure 5.17a and b the error level L_E of the non-reconstructed synthesized wave field p and the error level of the spatially reconstructed wave field p^H are compared. The error level L_E in figure 5.17b shows a significant improvement of the synthesized wave field for f>850Hz in a region around x_R=-1m, compared to the error level of the previous example (figure 5.12a). For the situation in which the reconstruction filter is omitted, the improvement is smaller, however not negligible. The explanation for the fairly accurate results without a reconstruction filter will be postponed until section 5.5.3.

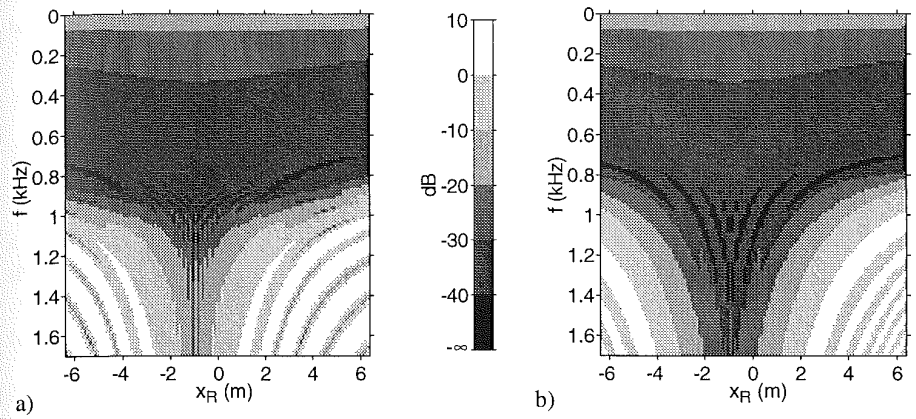


Figure 5.17: *Influence of the reconstruction filter on the error level L_E. Registration along the line $z=z_R$ for -6.4m<x_R<6.4m. Temporal bandwidth: 0-1.7 kHz.*
a) x-f diagram of the error level L_E without applying a spatial reconstruction filter.
b) x-f diagram of the error level L_E using a spatially bandlimited reconstruction filter.

In figure 5.18a and b a comparison is made of the pressure ratio level L_R for both situations. Obviously, the grating lobes in the spatial spectrum $\tilde{P}$ of the synthesized wave field cause a distortion of the temporal spectrum for high frequencies in the region around x_R=-1m, whereas the pressure ratio level L_R of the spatially reconstructed wave field p^H shows a better behavior in this area. In figure 5.18a and b it can also be observed that reconstruction filtering reduces

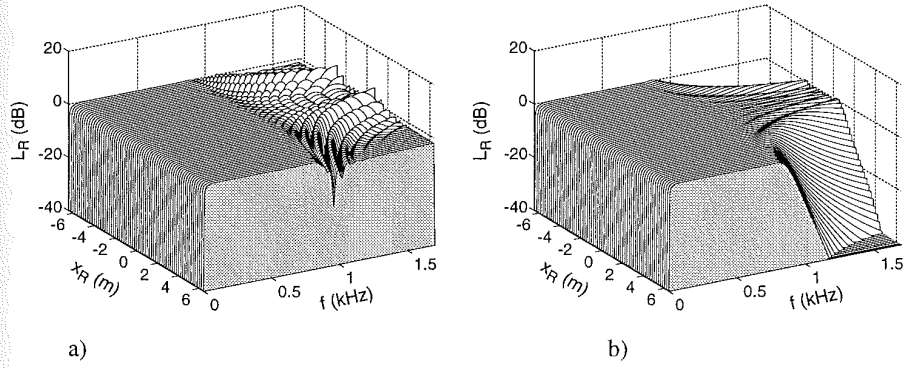


Figure 5.18: *Influence of the reconstruction filter on the pressure ratio level L_R. Registration along the line $z=z_R$ for -6.4<x_R<6.4. Temporal bandwidth: 0-1.7 kHz.*
a) x-f diagram of the error level L_E without applying a spatial reconstruction filter.
b) x-f diagram of the error level L_E using a spatially bandlimited reconstruction filter.

the high-frequency contents of the synthesized sound field at the outer receiver positions. This effect will be explained in the next section where spatial filtering will be related to directional behavior of the notional source and the secondary sources.

5.4 Physical interpretation of the spatial filtering process

In the previous sections it was shown that spatial bandwidth reduction can be applied successfully in wave field synthesis applications in order to suppress spatial aliasing. In this section a physical interpretation of the spatial bandwidth reduction process in wave field synthesis will be given in the space domain, leading to practical methods for implementation.
First, a mathematical description of the total wave field synthesis process, including spatial bandwidth reduction, will be given in the k_x-k domain.
The spatially bandlimited synthesized wave field $\tilde{P}^H$ can be written as

$$\tilde{P}^H(k_x, \omega) = \tilde{H}_r(k_x, \omega) \cdot \tilde{W}(k_x, \omega) \cdot \tilde{Q}_p^H(k_x, \omega), \tag{5.22}$$

in which $\tilde{H}_r$ and $\tilde{W}$ are respectively k_x-ω domain representations of the spatial reconstruction filter and the secondary source response. $\tilde{Q}_p^H$ is the k_x-ω domain representation of the spatially sampled and bandlimited driving signal, determined by

$$\tilde{Q}_p^H(k_x, \omega) = \tilde{\Delta}_p(k_x, \omega) * \tilde{Q}^H(k_x, \omega) \tag{5.23a}$$

$$\tilde{Q}^H(k_x, \omega) = \tilde{H}(k_x, \omega) \cdot \tilde{Q}(k_x, \omega), \tag{5.23b}$$

where $\tilde{H}$ is the spatial anti-aliasing filter.
Transforming expressions (5.22) and (5.23a-b) to the space-frequency domain yields

$$P^H(x_R) = H_r(x_R) * W(x_R) * Q_\Delta^H(x_R), \tag{5.24}$$

with

$$Q_\Delta^H(x_L) = \Delta_p(x_L) \cdot Q^H(x_L), \tag{5.25a}$$

$$Q^H(x_L) = H(x_L) * Q(x_L) \tag{5.25b}$$

For notational convenience the angular frequency ω is omitted. Δ_p is the spatial sampling operator in the space domain, defined by

$$\Delta_p(x_L) = \Delta x \sum_{n=-\infty}^{\infty} \delta(x_L - n\Delta x) \tag{5.26}$$

The driving function Q can be interpreted as the wave field of the notional monopole line source Ψ_m at $\mathbf{r}_\Psi=(x_\Psi,z_\Psi)^T$ at the secondary source at $\mathbf{r}_L=(x_L,z_L)^T$ (see eq.(4.10a) or (4.11a)). The secondary source response W is the response of the secondary source at the receiver position $\mathbf{r}_R=(x_R,z_R)^T$ (see eq.(4.10b) or (4.11b)).

5.4.1 Physical interpretation of the spatial anti-aliasing filter

The driving signal Q can be written as a convolution along the x-axis of a notional line source distribution Ψ_m and a notional line source response Q_σ:

$$Q(x_L) = \Psi_m(x_\sigma) * Q_\sigma(x_L|x_\sigma). \tag{5.27}$$

In the case of a single notional line source, the notional monopole line source distribution Ψ_m is given by

$$\Psi_m(x_\sigma) = \delta(x_\sigma - x_\Psi), \tag{5.28}$$

while the notional line source response Q_σ is given by
for the 2½D Rayleigh I case:

$$Q_\sigma(x_L|x_\sigma) = \sqrt{\frac{z_R - z_L}{z_R - z_\Psi}} \sqrt{\frac{jk}{2\pi}} S(\omega) \cos\varphi_{inc} \frac{e^{-jk|\mathbf{r}_L - \mathbf{r}_\sigma|}}{\sqrt{|\mathbf{r}_L - \mathbf{r}_\sigma|}} \tag{5.29}$$

for the 2½D Rayleigh II case:

$$Q_\sigma(x_L|x_\sigma) = \sqrt{\frac{z_R - z_L}{z_R - z_\Psi}} \sqrt{\frac{1}{2\pi jk}} S(\omega) \frac{e^{-jk|\mathbf{r}_L - \mathbf{r}_\sigma|}}{\sqrt{|\mathbf{r}_L - \mathbf{r}_\sigma|}} \tag{5.30}$$

where $\mathbf{r}_\sigma=(x_\sigma,z_\Psi)^T$. Insertion eq.(5.27) in eq.(5.25b) yields

$$\begin{aligned} Q^H(x_L) &= H(x_L) * [\Psi_m(x_\sigma) * Q_\sigma(x_L|x_\sigma)] \\ &= [H(x_\sigma) * \Psi_m(x_\sigma)] * Q_\sigma(x_L|x_\sigma), \\ &= \Psi_m^H(x_\sigma) * Q_\sigma(x_L|x_\sigma) \end{aligned} \tag{5.31}$$

in which Ψ_m^H is the spatially filtered notional monopole source distribution, given by

$$\Psi_m^H(x_\sigma) = H(x_\sigma - x_\Psi) \tag{5.32}$$

This result can be interpreted as follows: convolution of the driving signal Q with a spatial filter H is equivalent to convolving the notional line source Ψ_m with that spatial filter H. This means that, as a result of the spatial anti-aliasing filter H, the notional monopole line source Ψ_m is replaced by a notional monopole line source distribution Ψ_m^H along the line $z=z_\Psi$. This is illustrated in figure 5.19.
In general, the length $D(\omega)$ of the spatial filter H in the x-ω domain is inversely proportional to the width $B(\omega)$ of the passband ($B(\omega)=2\pi/\Delta x$, see figure 5.13b) of $\tilde{H}$ in the k_x-ω domain. If the distance $|z_L - z_\Psi|$ between the notional monopole source distribution Ψ_m^H along the line $z=z_\Psi$ and the secondary source distribution along the line $z=z_L$ obeys

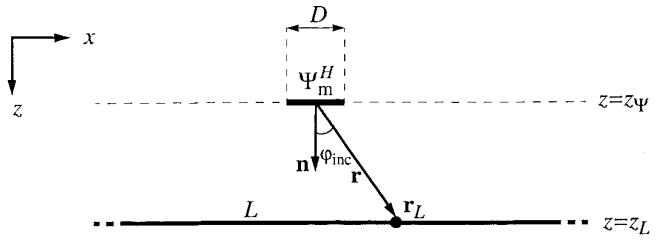


Figure 5.19: As a result of the spatial anti-aliasing filter H, the notional monopole line source Ψ_m is replaced by a notional monopole line source distribution Ψ_m^H along the line $z=z_\Psi$.

$$|z_L - z_\Psi| \gg \frac{D^2}{2\lambda_{min}}, \tag{5.33}$$

the Fraunhofer condition is fulfilled. In that case, the notional monopole source distribution Ψ_m^H may be regarded as a notional point source with a directivity pattern $G(\varphi_{inc},\omega)$. Using the fact that $\theta=\varphi_{inc}$ in the Fraunhofer domain, the directivity pattern G is given by

$$G(\varphi_{inc}, \omega) = \tilde{H}(k_x, \omega), \tag{5.34}$$

in which $k_x=k\sin\varphi_{inc}$. Applying a frequency-independent spatial filter $\tilde{H}$, yields a frequency-dependent directivity pattern G.
The spatially bandlimited driving signal Q^H, given by eq.(5.25b), can be rewritten as

$$Q^H(x_L, \omega) = G(\varphi_{inc}, \omega) Q(\varphi_{inc}, r, \omega). \tag{5.35}$$

Note that this result is only valid under the Fraunhofer condition of eq.(5.33).

5.4.2 Physical interpretation of the spatial reconstruction filter

Subsequent to the spatial sampling process, described by eq.(5.25a), the wave field is synthesized, yielding the synthesized wave field P, which must be spatially filtered by a reconstruction filter H_r, described by eq.(5.24).
The spatial reconstruction filtering process can be described by a continuous convolution of the synthesized wave field along the receiver line with the spatial filter H_r. Since the synthesized wave field can not be processed afterwards in practice, the spatial reconstruction filter must be implemented in a different way. Therefore, the spatially bandlimited wave field P^H is written as the convolution of a spatially bandlimited secondary source response W^H and the bandlimited and sampled driving signal Q_Δ^H, according to

$$P^H(x_R) = W^H(x_R) * Q_\Delta^H(x_R), \tag{5.36}$$

where W^H is given by

$$W^H(x_R) = H_r(x_R) * W(x_R) \tag{5.37}$$

The secondary source response W can be written as a convolution along the x-axis of a source distribution Ψ_s and a secondary point source response W_τ:

$$W(x_R) = \Psi_s(x_\tau) * W_\tau(x_R|x_\tau). \tag{5.38}$$

The secondary source distribution Ψ_s is given by

$$\Psi_s(x_\tau) = \delta(x_\tau - x_L), \tag{5.39}$$

while the secondary point source response W_τ is given by
for the 2½D Rayleigh I case:

$$W_\tau(x_R|x_\tau) = \frac{e^{-jk|\mathbf{r}_R - \mathbf{r}_\tau|}}{|\mathbf{r}_R - \mathbf{r}_\tau|} \tag{5.40}$$

for the 2½D Rayleigh II case:

$$W_\tau(x_R|x_\tau) = jk\cos\varphi \frac{e^{-jk|\mathbf{r}_R - \mathbf{r}_\tau|}}{|\mathbf{r}_R - \mathbf{r}_\tau|} \tag{5.41}$$

where $\mathbf{r}_\tau=(x_\tau,z_L)^T$. Inserting eq.(5.38) in eq.(5.37) yields

$$\begin{aligned} W^H(x_R) &= H_r(x_R) * [\Psi_s(x_\tau) * W_\tau(x_R|x_\tau)] \\ &= [H_r(x_\tau) * \Psi_s(x_\tau)] * W_\tau(x_R|x_\tau), \\ &= \Psi_s^H(x_\tau) * W_\tau(x_R|x_\tau) \end{aligned} \tag{5.42}$$

in which Ψ_s^H is the spatially filtered secondary point source, given by

$$\Psi_s^H(x_\tau) = H(x_\tau - x_L) \tag{5.43}$$

This result can be interpreted as follows: convolution of the synthesized wave field is equivalent to convolving the secondary source response W with a spatial filter H_r, which is equivalent to convolving each secondary point source Ψ_s with that spatial filter H_r. This means that, as a result of the spatial reconstruction filter H_r, the discrete set of secondary point sources Ψ_s is replaced by a set of (possibly overlapping) secondary sub-array elements Ψ_s^H with length D_r along the line $z=z_L$ (figure 5.20).
In the Fraunhofer domain, each secondary source element Ψ_s^H may be regarded as a secondary point source (monopole or dipole) with a extra directivity pattern $G_r(\varphi,\omega)$. The directivity pattern G_r is determined by

$$G_r(\varphi, \omega) = \tilde{H}_r(k_x, \omega), \tag{5.44}$$

with $\theta=\varphi$. The spatially bandlimited secondary source response W^H, given by eq.(5.37), can be rewritten as

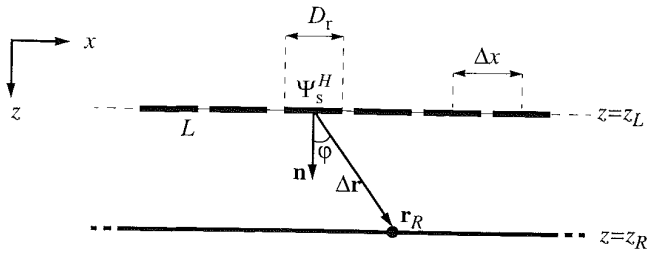


Figure 5.20: As a result of the spatial reconstruction filter H_r, the discrete set of secondary point sources Ψ_s is replaced by a set of (overlapping) secondary sub-array elements Ψ_s^H

$$W^H(x_R, \omega) = G_r(\varphi, \omega) W(\varphi, \Delta r, \omega) \tag{5.45}$$

Note that this result is only valid under the Fraunhofer condition.

5.5 Practical aspects of spatial bandwidth reduction

In the previous sections the theoretical framework was set up for spatial bandwidth reduction in wave field synthesis applications.
In this section we will study the performance of finite and discrete secondary source arrays consisting of directional loudspeakers. It will be investigated how aperture limitation can be incorporated in the spatial bandwidth reduction concept.

5.5.1 Aperture limitation versus notional source directivity

In the previous chapter it was shown that a limitation of the aperture of a secondary source array causes truncation effects in the synthesized wave field. These truncation effects were described as diffractions at the boundaries of the aperture. Now, we will continue the analysis of aperture limitations from a different viewpoint.
Consider the configuration of figure 5.21. The aperture of the array, as seen from the notional monopole source Ψ_m is determined by φ_{min} and φ_{max}. As shown in chapter 4, the effect of aperture limitation can be taken into account by multiplying the driving signal Q with a window function T, yielding the following synthesis integral

$$P(x_R) = \int_{-\infty}^{\infty} [T(x_L)Q(x_L)]W(x_R - x_L)dx_L, \tag{5.46}$$

Writing this result in a symbolic convolution notation yields

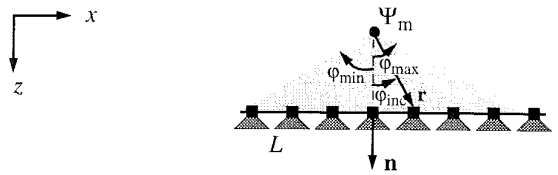


Figure 5.21: Configuration of a notional monopole source Ψ and a finite loudspeaker array L.

$$P(x_R) = [T(x_R)Q(x_R)] * W(x_R) = Q^T(x_R) * W(x_R), \tag{5.47}$$

where Q^T is the windowed driving signal.
The configuration of figure 5.21 is equivalent to the configuration of figure 5.22, in which the

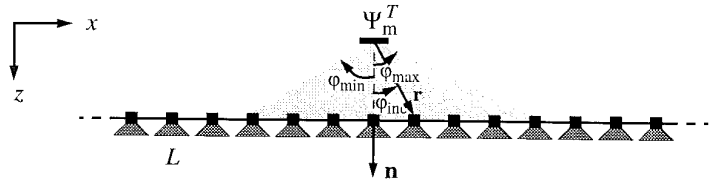


Figure 5.22: Configuration of a directional notional source Ψ_m^T and an infinite loudspeaker array L. This configuration is equivalent to the configuration of figure 5.21.

finite loudspeaker array is replaced by an infinite array, and the notional monopole source Ψ_m is replaced by a directional source Ψ_m^T, having frequency independent directivity characteristics G:

$$G(\varphi_{inc}, \omega) = T(\varphi_{inc}) \tag{5.48}$$

with $\varphi_{inc}=\text{asin}[(x_L-x_\Psi)/r]$. So, alternatively, the windowed driving signal Q^T may be written as

$$Q^T(x_L, \omega) = T(\varphi_{inc})Q(\varphi_{inc}, r, \omega). \tag{5.49}$$

Note that this expression has the same form as eq.(5.35), in which the Fraunhofer approximation of the spatial filtered driving signal Q^H was given. The correspondence between eq.(5.35) and eq.(5.49) will be further established in the next subsection, where the spatial bandwidth of the windowed driving signal Q^T will be determined.

5.5.2 Aperture limitation versus the spatial bandwidth of the driving signal

Applying the spatial Fourier transformation to eq.(5.47), gives the following result:

$$\tilde{P}(k_x) = [\tilde{T}(k_x) * \tilde{Q}(k_x)]\tilde{W}(k_x) = \tilde{Q}^T(k_x)\tilde{W}(k_x), \tag{5.50}$$

in which $\tilde{T}$ and $\tilde{Q}^T$ are respectively the spatial Fourier transforms of the window function T and windowed driving signal Q^T.

Generally, the influence of the aperture size on the spatial spectrum $\tilde{Q}^T$ of the windowed driving signal Q^T cannot be expressed in an exact analytical form. Therefore, first a rather intuitive description will be given of the consequences of aperture limitation. Later, we will investigate under which conditions aperture limitation can be described as spatial filtering of the driving signal.

The spherical wave fronts emitted by the notional source Ψ_m can be regarded locally as plane wave fronts, making an angle φ_{inc} with the array L parallel to the x-axis, which is shown in fig-

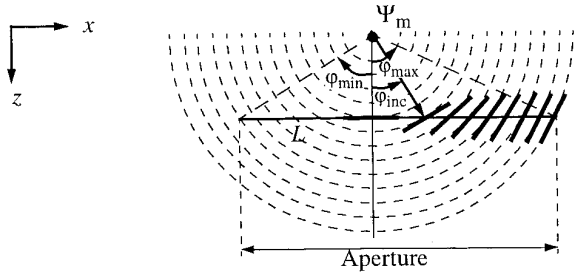


Figure 5.23: The spherical wave fronts emitted by the source Ψ can be regarded locally as plane wave fronts, making an angle φ_{inc} with the line L

ure 5.23. For a given array aperture the angle φ_{inc} lies in the interval $[\varphi_{min},\varphi_{max}]$. Intuitively, we can feel that a limitation of the array aperture will reduce the spatial bandwidth of the driving signal Q, i.e. a limitation of the range of angles θ in the spatial spectrum. It is important to note that the angle φ_{inc} in the x-z domain is *not* identical to the propagation angle θ of a plane wave component in the k_x-k domain. However, it can be shown that for each plane wave component in the spatial spectrum, the main contribution comes from one point on the line L, i.e. the point where the plane wave is tangent to the wave front (Wapenaar, 1992). So, as a crude approximation we may state that $\theta \approx \varphi_{inc}$. This means that a limitation of the aperture in the x-z domain ($\varphi_{min} \le \varphi_{inc} \le \varphi_{max}$) gives a frequency-dependent spatial bandwidth reduction in the k_x-k domain ($k\sin\varphi_{min} \le k_x \le k\sin\varphi_{max}$). The conditions, under which this relationship is valid, will be analyzed in the following.

In the previous sub-section we saw that applying a window T to the driving signal Q is equivalent to assigning a directivity function $G(\varphi_{inc}) = T(\varphi_{inc})$ to the notional source Ψ_m^T.

In the Fraunhofer area, the directional notional source may be interpreted as a notional monopole source distribution Ψ_m^T with a frequency dependent length $D(\omega)$. The notional source distribution Ψ_m^T is the convolution of a certain spatial filter $H(x_\sigma)$ with the notional monopole source distribution $\Psi_m(x_\sigma)$ of eq.(5.28) along the line $z=z_\Psi$, yielding

$$\Psi_m^T(x_\sigma) = H(x_\sigma - x_\Psi). \tag{5.51}$$

The spatial filter $H(x_\sigma)$ is the inverse spatial Fourier transform of $\tilde{H}(k_x)$, which is related to the directivity function $T(\varphi_{inc})$ in the Fraunhofer area by

$$\tilde{H}(k_x, \omega) = T(\varphi_{inc}), \tag{5.52}$$

in which the angle $\theta=\varphi_{inc}$. Reversing the argumentation given in section 5.4.1, it can be stated that: convolution of the notional monopole source Ψ_m with a spatial filter H is equivalent to convolving the driving signal Q with that spatial filter H.

$$Q^T(x_L) = Q^H(x_L) = H(x_L) * Q(x_L) \tag{5.53}$$

The spatial spectrum of the windowed driving signal Q^T can be written as

$$\tilde{Q}^T(k_x, \omega) = T(\varphi_{inc}) \cdot \tilde{Q}(k_x, \omega). \tag{5.54}$$

This result is in agreement with which we found intuitively before. However, note that eq.(5.54) is only valid under Fraunhofer conditions, which will be more precisely specified now.

Generally, the length $D(\omega)$ of the spatial filter H is inversely proportional to the width $B(\omega)$ of the passband of $\tilde{H}$ in the k_x-ω domain, yielding

$$D(\omega) = \frac{C}{B(\omega)} = \frac{C}{2k\sin|\varphi_{max}|} \tag{5.55}$$

where C is a proportionality constant, which is determined by the shape of the window T. It is assumed that $|\varphi_{min}|=|\varphi_{max}|$. Inserting eq.(5.55) into eq.(5.33), yields the Fraunhofer condition

$$|z_L - z_\Psi| \gg \frac{C^2\lambda}{32\pi^2\sin^2|\varphi_{max}|}, \tag{5.56}$$

showing that for increasing wavelengths λ (decreasing frequencies f) the Fraunhofer condition is fulfilled more difficultly.

As an example, a tapered rectangular window T is chosen, with $|\varphi_{max}|=|\varphi_{min}|=30°$. Figure 5.24a shows the window T as a function of $\sin(\varphi_{inc})$. In the Fraunhofer domain, this function may be

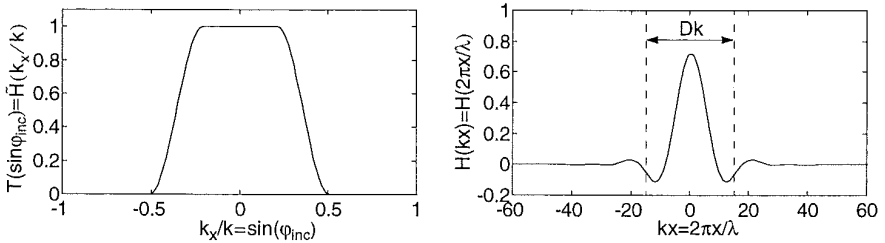


Figure 5.24: In the Fraunhofer domain the tapered rectangular window $T(\sin\varphi_{inc})$ may be interpreted as a spatial filter $\tilde{H}(k_x/k)$.
a) Spatial filter $\tilde{H}(k_x/k)$
b) Spatial inverse Fourier transform H(kx)

interpreted as a spatial filter $\tilde{H}(k_x/k)$, which is a function of the normalized spatial frequency k_x/k. The spatial inverse Fourier transform $H(kx)=H(2\pi x/\lambda)$, which is a function of a relative distance kx, is shown in figure 5.24b. The relative length Dk of the spatial filter H is found rather arbitrarily and equals: $Dk \approx 35$. Rewriting the Fraunhofer condition (5.33) into

$$|z_L - z_\Psi| \gg \frac{(Dk)^2 \lambda}{8\pi^2}, \tag{5.57}$$

and inserting Dk in eq.(5.57) yields

$$|z_L - z_\Psi| \gg 16\lambda, \tag{5.58}$$

for this example. Comparing eq.(5.56) and eq.(5.58) it can be found that $C \approx 15$ for this window shape. Condition (5.58) puts a severe demand on the notional source-to-array distance, which will not be met in practice for low frequencies.

5.5.3 Aperture limitation versus receiver directivity and reconstruction filtering

A limitation of the aperture of the loudspeaker array, as seen from the notional source, can also be interpreted as an aperture limitation, as seen from the receiver positions. This is illustrated in figure 5.25. The aperture, determined by φ_{min} and φ_{max}, as seen from one receiver, which is

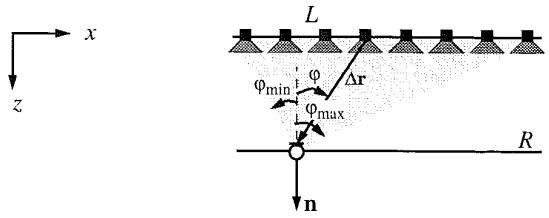


Figure 5.25: Configuration of a finite loudspeaker array L and a receiver line R, consisting of omni-directional microphones.

represented by an omni-directional microphone, is indicated by the grey triangle. This configuration is equivalent to the configuration of figure 5.26, in which the finite loudspeaker array is replaced by an infinite array, and the omni-directional receiver is replaced by a directional receiver.

In the Fraunhofer area, each directional microphone may be interpreted as an omni-directional microphone distribution, with a frequency dependent length $D_r(\omega)$. The Fraunhofer condition is fulfilled if

$$|z_R - z_L| \gg \frac{D_r(\omega)^2}{2\lambda_{min}}, \tag{5.59}$$

where $|z_R - z_L|$ is the distance between the loudspeaker array L and the receiver line R.

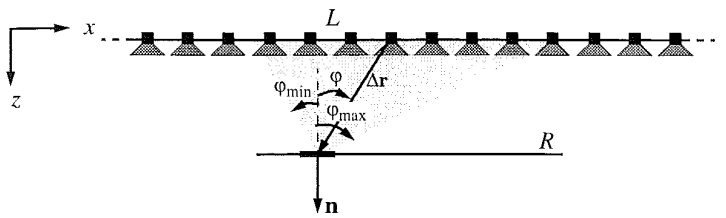


Figure 5.26: Configuration of a finite loudspeaker array L and a receiver line R, consisting of directional microphones. This configuration is equivalent to the configuration of figure 5.25.

This result means that by windowing the array, the continuous distribution of omni-directional receivers is replaced by a continuous distribution of (overlapping) microphone sub-arrays with a position- dependent directivity pattern.

Virtually, the microphone sub-array distribution can be described by the result of a position-dependent convolution of the omni-directional microphone distribution, along the line R. This is precisely the effect of a reconstruction filter; in this case, a position-dependent reconstruction filter.

In conclusion we may state that, for receiver positions sufficiently far away from a windowed linear loudspeaker array, no reconstruction filter has to be applied, which was already demonstrated at the end of section 5.3.2. Due to the aperture limitation of the array, as seen from the receiver ($\varphi_{min} \le \varphi \le \varphi_{max}$), a frequency- and position-dependent spatial reconstruction filter is realized. Note that only for centrally positioned receivers with respect to the loudspeaker array, a fairly accurate reconstruction filter is realized. More lateral receivers are doing without a proper reconstruction filter, and consequently suffer from false grating lobes in the synthesized wave field.

5.5.4 Practical approach to spatial bandwidth reduction: possibilities and limitations

In order to implement a spatial anti-aliasing filter, the continuous driving signal Q must be convolved with an anti-aliasing filter H. It was shown in section 5.4.1 that spatial anti-aliasing filtering may be regarded as replacing the notional point source Ψ_m by a notional point source distribution Ψ_m^H. Applying a frequency-independent spatial low-pass filter $\tilde{H}$, results in a frequency-dependent directivity pattern G of the notional source in the Fraunhofer area.

Consequently, the temporal bandwidth of the driving signal is dependent of the position of the loudspeakers. This results in a position-dependent temporal bandwidth at the receiver line, as was shown by the pressure ratio level L_R in figure 5.18a.

In section 5.5.2 an alternative spatial filtering method was presented, viz, aperture limitation of the loudspeaker array. This method yields a frequency independent directivity pattern of the notional source in the Fraunhofer domain, and can be easily implemented by windowing the driving signals, according to

$$Q^H(x_L, \omega) = Q^T(x_L, \omega) = T(\varphi_{\text{inc}})Q(\varphi_{\text{inc}}, r, \omega)\,. \tag{5.60}$$

The spatial reconstruction filtering process can be described, on the one side, by a continuous convolution of the synthesized wave field along the receiver line with the spatial filter H_r, and, on the other hand, by a convolution of the discrete secondary source distribution. Since the synthesized wave field can not be processed afterwards, the spatial reconstruction filter must be implemented by a convolution of the discrete secondary source distribution. This means that the discrete set of secondary point sources Ψ_s must be replaced by a set of (overlapping) secondary sub-array elements Ψ_s^H with length D_r along the line $z=z_L$.

In the Fraunhofer area, each secondary sub-array may be replaced by one secondary source (monopole or dipole) with an extra directivity pattern G_r, yielding

$$W^H(x_R, \omega) = G_r(\varphi, \omega)W(\varphi, \Delta r, \omega) \tag{5.61}$$

So, in general, to implement a spatial reconstruction filter in a practical wave field synthesis system, loudspeaker arrays must be used consisting of directional loudspeakers. For the special case that the receivers are centrally and remotely positioned from the windowed loudspeaker array, we saw that a reconstruction filter can be omitted.

The directional properties of a loudspeaker (diameter D_r), can be modeled by a rigid piston located in an infinitely large baffle. Virtually, a rigid piston can be regarded as the convolution of a monopole source with a distribution H_r. In the Fraunhofer area, the directivity characteristics G_r are given by the spatial Fourier transform $\tilde{H}_r$ of H_r:

$$G_r(\varphi, \omega) = \tilde{H}_r(k_x, \omega)\,. \tag{5.62}$$

Since the spatial filter H_r is frequency-independent (i.e. a frequency-independent displacement of the piston), the directivity pattern G_r is frequency-dependent.

However, in general, the directivity of a loudspeaker is limited and cannot be controlled effectively. Higher directivity and better control can be obtained by applying (overlapping) finely sampled subarrays of (directional) loudspeakers. The sampling distance Δx_r of the subarrays should be small compared to the sampling distance Δx of the driving signal Q^H, otherwise the grating lobes in the spatial spectrum of the synthesized wave field cannot be removed.

Note, however, that the limited number of loudspeaker channels was the initial constraint in this chapter, which led to the concept of spatial filtering. Fortunately, the application of finely sampled subarrays can be implemented by simple technical means, and is more economical than sampling the driving signal at a higher rate. This problem will be further addressed in chapter 7, in which the design aspects of the DSE system are presented.

Finally, an overview is given of the various interpretations of spatial filtering, described in the last sections (figure 5.27).

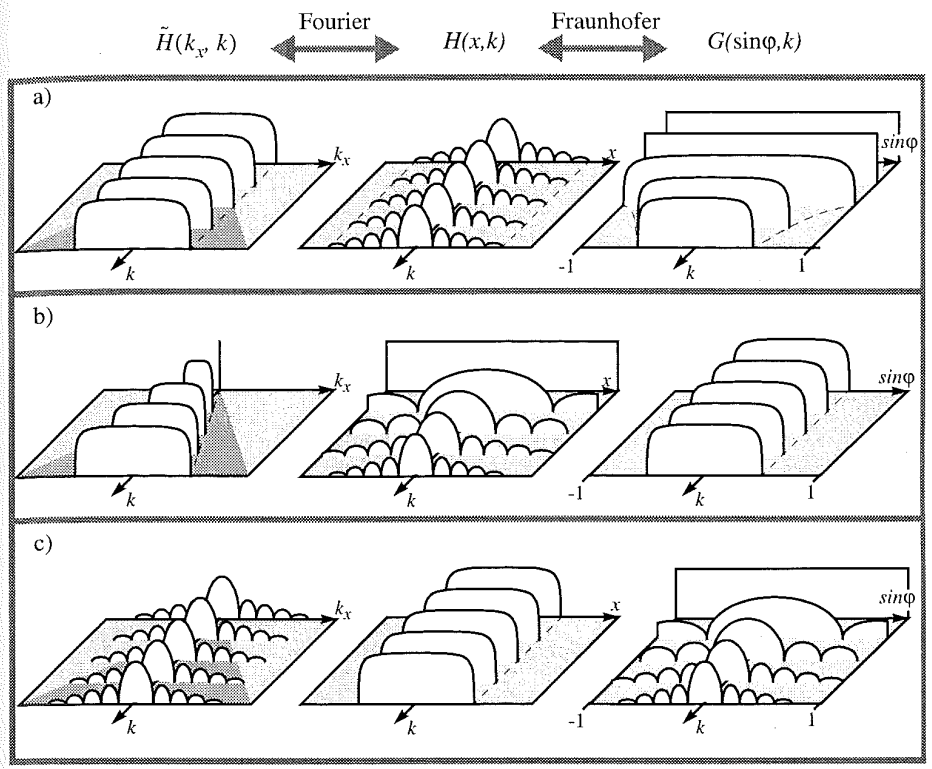


Figure 5.27: *Overview of the various interpretations of spatial filtering.*

a) Applying a frequency independent spatial filter $\tilde{H}$ to the response of a (notional or secondary) point source may be interpreted as replacing the point source by a frequency independent point source distribution H, with a frequency dependent directivity pattern G in the Fraunhofer area.

b) Windowing the driving signals is equivalent to replacing the notional source by a frequency independent directivity pattern G. In the Fraunhofer area, the directional notional source may be regarded as a frequency dependent point source distribution H, with a frequency dependent spatial Fourier transform $\tilde{H}$.

c) Loudspeakers of finite size may be represented by a rigid piston, which can be regarded as applying a frequency independent spatial filter H with spatial Fourier transform $\tilde{H}$, to the secondary monopole response. In the Fraunhofer area, the loudspeaker has a frequency dependent directivity pattern G.

5.5.5 Practical approach to spatial bandwidth reduction: example with a linear array

In the next example, we will study the proposed practical method of spatial bandwidth reduction. The spatial bandwidth of the driving signal will be reduced by windowing the loudspeaker array. Two situations will be compared: 1) With a reconstruction filter by applying finely sampled loudspeaker subarrays, and: 2) Without reconstruction filtering, which means that only monopole-type loudspeakers are used.

Consider the configuration of figure 5.28. The discrete linear array L at z_L=7m, consisting of

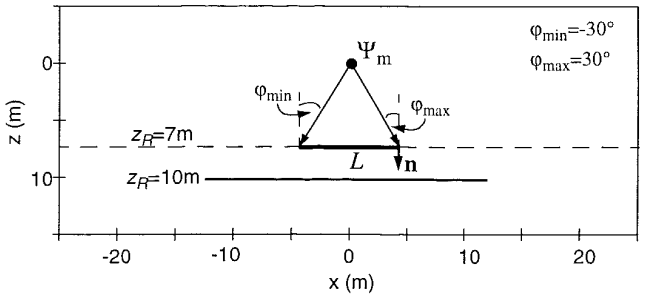


Figure 5.28: Diagram of the geometry of a notional monopole source Ψ_m at $\mathbf{r}_\Psi$=(0,0) and a linear discrete array of 39 loudspeaker elements along the line L (Δx=0.2 m). The primary and the synthesized wave field are recorded at the line $z=z_R$=10 for -12.8<x_R<12.8.

39 elements (subarrays or single loudspeakers), is sampled at intervals Δx=0.2 m. In order to demonstrate the effect of the spatial reconstruction filter, the geometry has been chosen such that the aperture of the array as seen from the notional source is small, while the aperture as seen from the receivers is large. The notional source at $\mathbf{r}_\Psi=(0,0,0)^T$ emits a bandlimited impulse (0-3.4 kHz), as shown in figure 5.29a and b. The synthesized wave field p and the primary wave field p_{pr} are recorded at the line $z=z_R$=10m for -12.8<x_R<12.8m. The x-t diagram of the primary wave field p_{pr} is shown in figure 5.29c.

The aperture of the array is reduced to $|\varphi_{max}|=|\varphi_{min}|=30°$, and the driving signal Q_Δ is tapered with a window function T. Since, in the Fraunhofer approximation, $\theta_{max} \approx \varphi_{max}$, the spatial aliasing frequency is given by $f_{al}=c/(2\Delta x \sin(\theta_{max}))$=1.7 kHz. Note that spatial aliasing is not completely avoided since f_{max}=3.4 kHz.

For situation 1, each loudspeaker sub-array consists of 9 monopole point sources at intervals $\Delta x_r=\Delta x/2$=0.1 m. The reconstruction filter H_r, which is a discretized and windowed version of an ideal reconstruction filter is given by

$$H_r(x_L) = \sin\left(\frac{\pi}{\Delta x}x_L\right) \Big/ \left(\frac{2\pi}{\Delta x}x_L\right) \qquad (5.63)$$

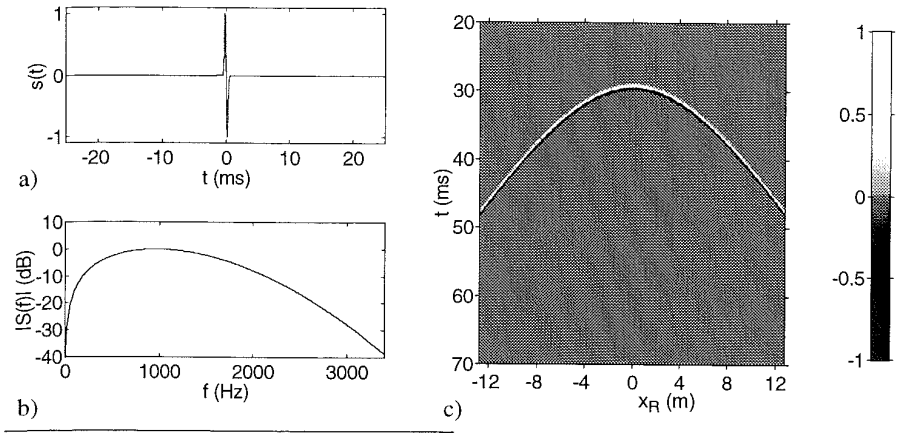


Figure 5.29: Primary wave field p_{pr} emitted by the notional monopole source.
a) source signature s(t)
b) amplitude spectrum $|S(\omega)|$ of the source signature, bandwidth 0-3.4 kHz.
c) x-t diagram

In figure 5.30a the ideal reconstruction filter (solid line) is plotted together with the discretized and windowed version (open circles, numbered 1-9). Note that only 5 of the 9 weights are unequal to zero.

The ideal reconstruction filter has a low-pass characteristic in the k_x-k domain with spatial cut-off frequency $k_x^{co}=\pi/\Delta x$, as shown in figure 5.30b. The spatial spectrum $\tilde{H}_r$ of the discretized and windowed version H_r is plotted in figure 5.30c, showing that the discrete reconstruction filter has a periodic spatial spectrum with period $2\pi/\Delta x_r$.

The k_x-k diagram of the discrete sub-array response $\tilde{W}^H$ at the line $z=z_R$ is shown in figure 5.31a, which is the multiplication of the periodic reconstruction filter $\tilde{H}_r$ and the monopole source response $\tilde{W}$, depicted in figure 5.31b. The k_x-k diagram of the sampled and tapered driving signal $\tilde{Q}_p^T$ at the line L is plotted in figure 5.31c, in which the lines $k_x=k\sin(\varphi_{min})$ and $k_x=k\sin(\varphi_{max})$ are indicated. In the Fraunhofer area, the energy is expected to be located between these two lines, which is true up to the spatial aliasing frequency f_{al}=1.7 kHz (or, up to wave number: k_{al}=31.4 m^{-1}).

Figure 5.31d shows the k_x-k diagram of the spatially band limited synthesized wave field $\tilde{P}^H$, which is the multiplication of $\tilde{W}^H$ and $\tilde{Q}_p^T$, while the synthesized wave field $\tilde{P}$ for situation 2, which is the multiplication of $\tilde{W}$ and $\tilde{Q}_p^T$, is shown in 5.31e. Notice that due to the reconstruction filtering the grating lobes in the spatial spectrum are reduced in figure 5.31d.

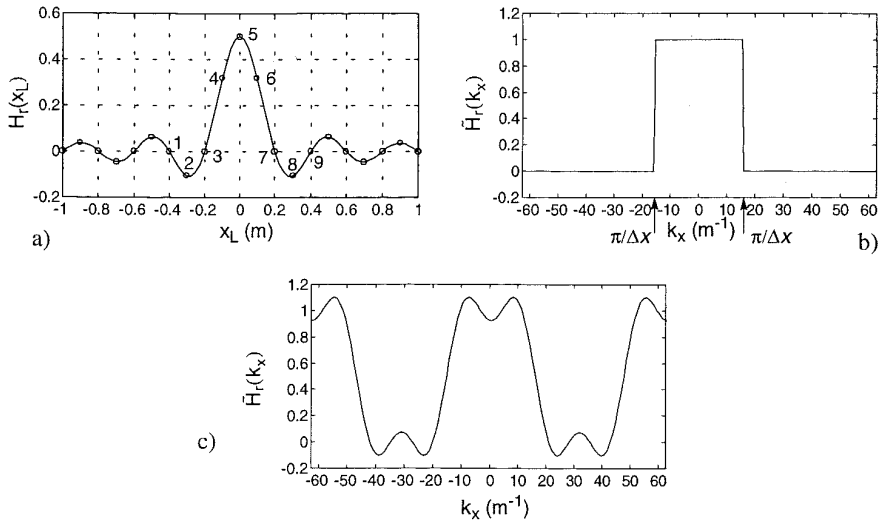


Figure 5.30: *The reconstruction filter H_r as a discretized and windowed version of an ideal reconstruction filter*
a) Ideal reconstruction filter H_r (solid line), discretized and windowed version of H_r (circles 1-9)
b) Ideal spatial frequency response $\tilde{H}_r$
c) Spatial spectrum $\tilde{H}_r$ of the discretized and windowed version

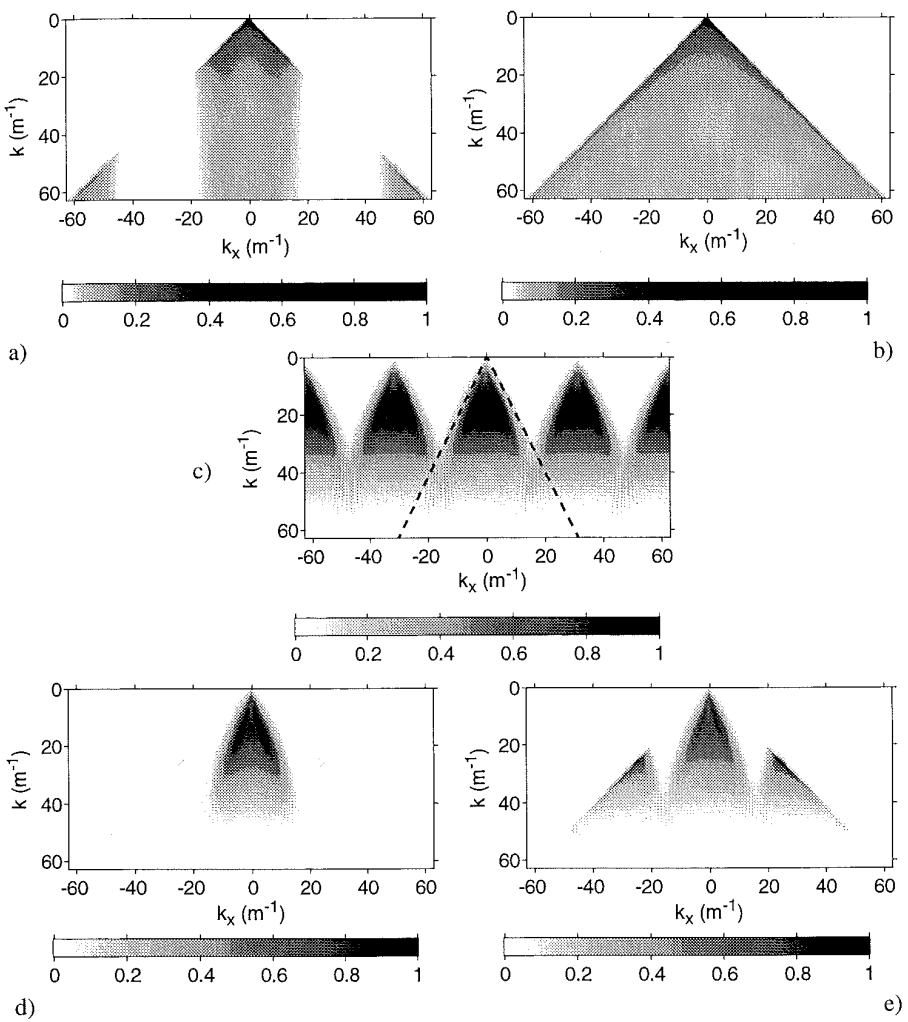


Figure 5.31: *Wave field synthesis in the spatial frequency domain with a linear discrete loudspeaker array, consisting of 39 elements.*
a) k_x-k diagram of the discrete sub-array response $\tilde{W}^H$
b) k_x-k diagram of the discrete sub-array response $\tilde{W}$
c) k_x-k diagram of the windowed and sampled driving signal $\tilde{Q}_p^T$
d) k_x-k diagram of the synthesized wave field $\tilde{P}^H$ using secondary subarrays
d) k_x-k diagram of the synthesized wave field $\tilde{P}$ without using secondary subarrays

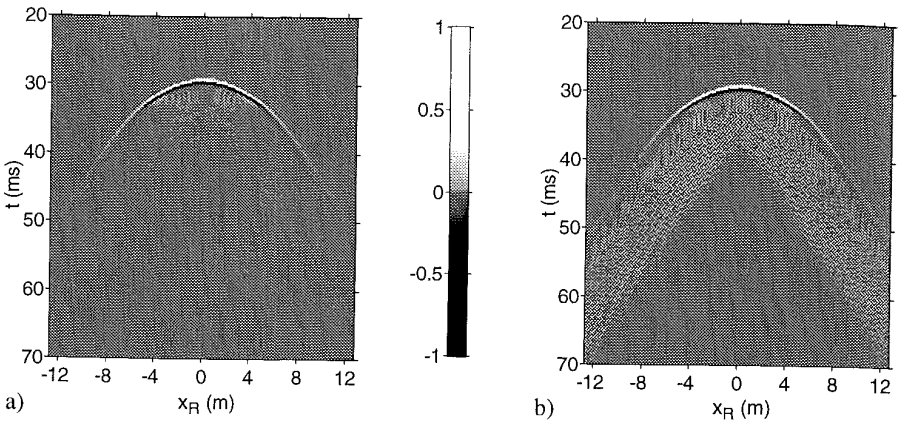


Figure 5.32: Influence of the spatial reconstruction filter on the synthesized wave field. A discrete linear loudspeaker array (Δx=0.2 m), consisting of 39 elements has been applied. Registration along the line $z=z_R$ for $-12.8<x_R<12.8$m. Temporal bandwidth: 0-3.4 kHz.
a) x-t diagram of the synthesized wave field p with application of secondary subarrays
b) x-t diagram of the synthesized wave field p with application of secondary monopole loudspeakers

Applying the double inverse Fourier transform to the synthesized wave fields $\tilde{P}^H$ and $\tilde{P}$ results in the x-t diagrams of figure 5.32a and b respectively. Notice that in figure 5.32a outside the aperture the wave field is suppressed as a result of the spatial filtering, while in figure 5.32b the high-frequency aliasing tails (see the grating lobes in figure 5.31e) are still visible.

The effect of the spatial reconstruction filter can also be observed in the x-f diagram of the error level L_E which is shown in figure 5.33a and b for both situations. For the first situation (figure 5.33a) the error level is rather low within the aperture upto 1.7 kHz. For higher frequencies the error level increases. For low frequencies (f<100 Hz) the error level is rather high, due to diffraction. For the second situation (figure 5.33b) the result is accurate up to ±1.1 kHz; at that frequency grating lobes become visible in the synthesized wave field (see figure 5.31e).

The somewhat higher error level in figure 5.33a for frequencies up to 1.1 kHz can be explained by realizing that the amplitude $|\tilde{H}_r|$ of the reconstruction filter in figure 5.30c is not exactly equal to unity inside the passband.

Figure 5.34a and b show the pressure ratio level L_R for both situations. Comparison of figure 5.34a and b reveals that by applying the reconstruction filter the amplitude spectrum is almost flat up to 1.7 kHz (inside the aperture), while for situation 2 the spectrum becomes irregular above 1.1 kHz.

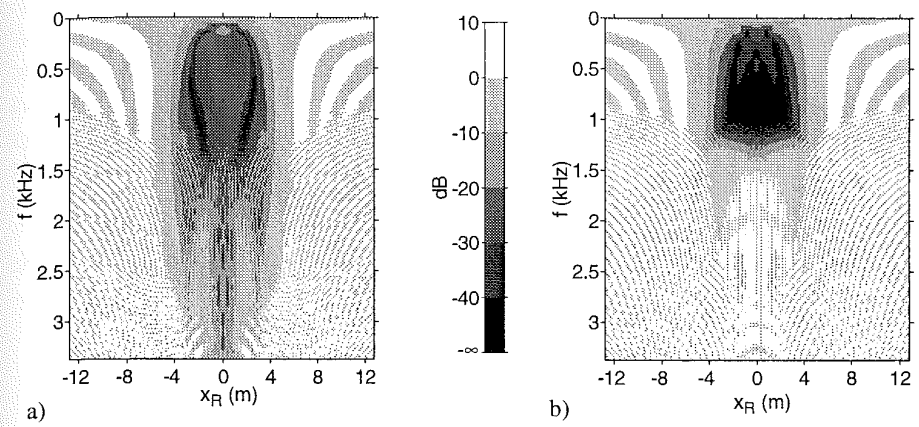


Figure 5.33: Influence of the spatial reconstruction filter on the error level. Wave field synthesized with a linear discrete loudspeaker array (Δx=0.2m). Registration along the line $z=z_R$ for $-12.8<x_R<12.8$m. Temporal bandwidth: 0-3.4 kHz.
a) x-f diagram of the error level L_E with application of secondary subarrays
b) x-f diagram of the error level L_E with application of secondary monopole loudspeakers

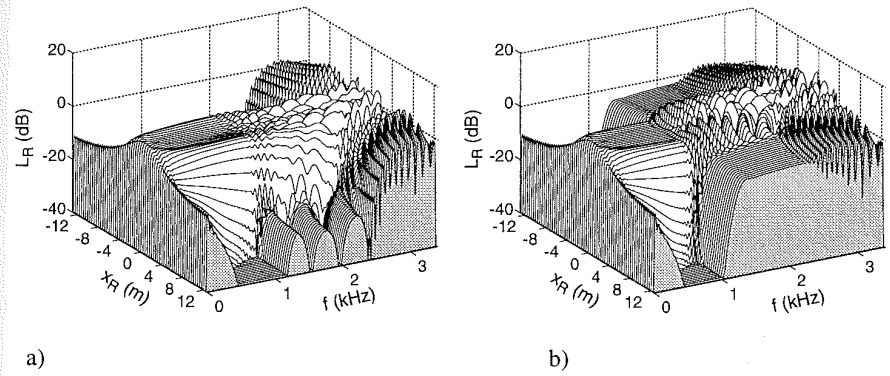


Figure 5.34: Influence of the spatial reconstruction filter on the pressure ratio level. Wave field synthesized with a linear discrete loudspeaker array (Δx=0.2m). Registration along the line $z=z_R$ for $-12.8<x_R<12.8$m. Temporal bandwidth: 0-3.4 kHz.
a) x-f diagram of the pressure ratio level L_R with application of secondary subarrays
b) x-f diagram of the pressure ratio level L_R with application of secondary monopole loudspeakers

5.6 Bent loudspeaker arrays

In the previous sections, we considered the performance of discrete linear loudspeaker arrays. In order to suppress spatial aliasing effects, a spatial bandwidth reduction technique has been developed. This technique puts new demands on the directivity characteristics of the secondary sources (loudspeakers). Further, it was shown that the array length must be decreased in order to reduce the spatial bandwidth of the driving signal. As a result, the synthesized wave field is only correct in a limited area, of which the size is dependent of the aperture of the array.
In this section we will focus on the application of bent arrays. It will be demonstrated that the spatial bandwidth reduction technique, developed for linear arrays, can be successfully integrated in the application of bent arrays.

5.6.1 Spatial bandwidth reduction by array-shaping

In section 5.5.2 it was shown that the spherical wave fronts incident on the array, emitted by a notional source, can be interpreted locally as plane wave fronts (figure 5.23). Further, it was demonstrated that the spatial bandwidth of the driving signal Q for a finite array depends on maximum angle of incidence. Consider the configuration of figure 5.35, in which we replaced

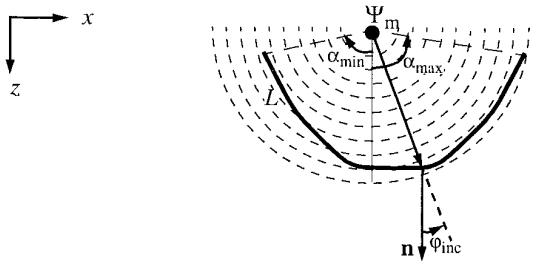


Figure 5.35: The maximum angle of incidence $\varphi_{max}=max\{\varphi_{inc}\}$ is small, while the aperture of the array $\alpha_{max}-\alpha_{min}$ is large.

the linear array by a bent array L. It is striking that the maximum angle of incidence $\varphi_{max}=\max\{|\varphi_{inc}|\}$ is small, while the aperture of the array $\alpha_{max}-\alpha_{min}$ is large.
Recalling the spatial anti-aliasing condition

$$f_{max} \leq f_{al} = \frac{c}{2\Delta x \sin(\theta_{max})}, \qquad (5.64)$$

and assuming $\theta_{max}=\varphi_{max}$, it may be expected that the spatial aliasing frequency f_{al} increases by using a bent array.
Virtually, a bent array may be considered as a special case of a piece-wise linear array; i.e. the limit for continually decreasing linear array elements. For each of these linear pieces, the spatial bandwidth reduction method holds, which was described in the previous section.

In order to implement a spatial reconstruction filter for the bent array configuration, directional loudspeakers must be used. Ideally, the directivity pattern $G_r(\varphi)$ of the loudspeakers must be chosen such that

$$G_r(\varphi) = 0, \qquad \text{for} \qquad |\varphi| > \max\{\varphi_{inc}\}. \qquad (5.65)$$

In this way the grating lobes in the spatial spectrum of the wave field, synthesized by a certain array element, are filtered. Application of loudspeakers which are too directional, will give rise to defects in the total synthesized wave field, since the wave fields, synthesized by each array element, are not connected properly (see figure 5.36).

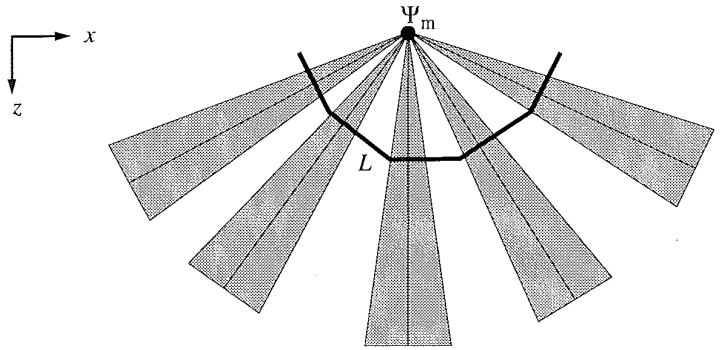


Figure 5.36: Effect of applying too directional loudspeakers. The wave fields, synthesized by each array element are not connected properly.

As for the configuration with a linear array, the directivity characteristics of the reconstruction sources can be improved by applying finely sampled subarrays.

5.6.2 Practical approach to spatial bandwidth reduction: example with a bent array

In this final example, we will study the proposed method of spatial bandwidth reduction by array-shaping. The spatial bandwidth of the driving signal will be reduced by limiting the maximum angle of incidence $\max\{\varphi_{inc}\}$ by applying a bent loudspeaker array. Also for this configuration two situations will be compared: 1) With a reconstruction filter by applying finely sampled loudspeaker subarrays, and: 2) Without reconstruction filtering, which means that only monopole-type loudspeakers are used.
Consider the configuration of figure 5.37, in which the linear loudspeaker array of figure 5.28, is replaced by a bent array L, consisting of 88 elements (subarrays or single loudspeakers). The notional source position and the recording line z_R are kept the same as in the previous example. The sampling interval ΔL=20 cm. The aperture of the array is given by α_{max} and α_{min}, while the maximum angle of incidence $\max\{\varphi_{inc}\}$=25°, as indicated in figure 5.37. The driving signal Q_Δ is tapered at the ends of the bent array.

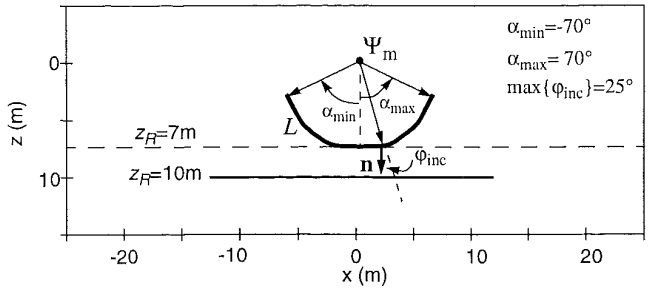


Figure 5.37: *Diagram of the geometry of a notional monopole source Ψ_m at $\mathbf{r}_\Psi$=(0,0) and a discrete bent loudspeaker array consisting of 88 elements along the line L (ΔL=0.2 m). The primary and the synthesized wave field are recorded at the line $z=z_R$=10m for -12.8<x_R<12.8m.*

For the situation in which a reconstruction filter is applied, the loudspeaker subarrays are identical to the subarrays used in the previous example. The reconstruction filter H_r is shown in figure 5.30a. The *x-t* diagram of the primary field is shown in figure 5.29c. The synthesized wave field *p* for situation 1 is shown in the *x-t* diagram of figure 5.38a, while the result for situation 2 is shown in figure 5.38b.

In contrast to the wave field in figure 5.32, also for lateral receiver positions the wave field is synthesized correctly. Notice that the aliasing tails in figure 5.38a are substantially reduced compared to the second situation in figure 5.38b. The effect of the spatial reconstruction filter

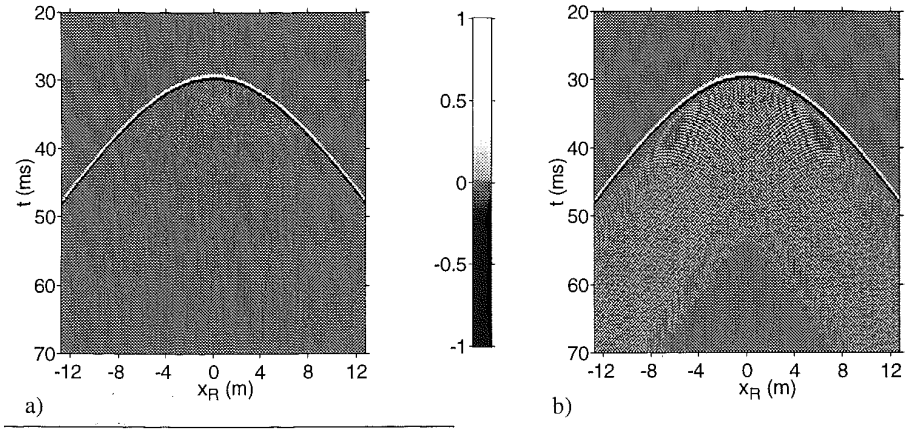


Figure 5.38: *Influence of the spatial reconstruction filter on the synthesized wave field. A discrete bent loudspeaker array (ΔL=0.2m), consisting of 88 elements has been applied. Registration along the line $z=z_R$ for -12.8<x_R<12.8m. Temporal bandwidth: 0-3.4 kHz.*

a) x-t diagram of the synthesized wave field p with application of secondary subarrays

b) x-t diagram of the synthesized wave field p with application of secondary monopole loudspeakers

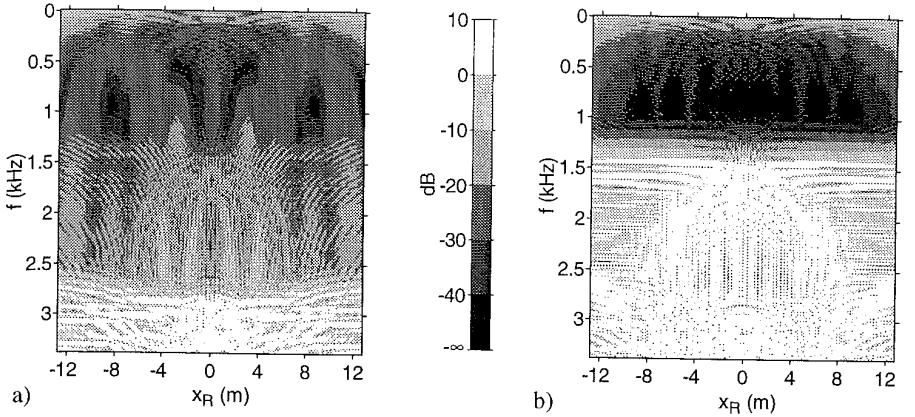


Figure 5.39: *Influence of the spatial reconstruction filter on the error level. Wave field synthesized with a discrete bent loudspeaker array (ΔL=0.2m). Registration along the line $z=z_R$ for -12.8<x_R<12.8m. Temporal bandwidth: 0-3.4 kHz.*

a) x-f diagram of the error level L_E with application of secondary subarrays

b) x-f diagram of the error level L_E with application of secondary monopole loudspeakers

can also be observed in the *x-f* diagrams of the error level L_E, which is shown in figure 5.39a and b for both situations.

Figure 5.40a shows the pressure ratio level L_R for the wave field synthesized using loudspeaker subarrays. Comparison of figure 5.40a with 5.40b, in which the result is shown for monopole-

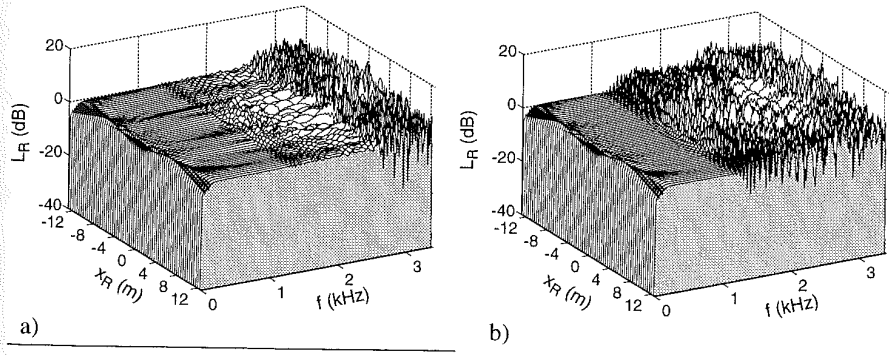


Figure 5.40: *Influence of the spatial reconstruction filter on the pressure ratio level. Wave field synthesized with a discrete bent loudspeaker array (ΔL=0.2m). Registration along the line $z=z_R$ for -12.8<x_R<12.8m. Temporal bandwidth: 0-3.4 kHz.*

a) x-f diagram of the pressure ratio level L_R with application of secondary subarrays

b) x-f diagram of the pressure ratio level L_R without application of secondary monopole loudspeakers

type loudspeakers, reveals that the amplitude spectrum has been flattened by the reconstruction filtering. Apart from a small ripple, the pressure ratio level in figure 5.40a is almost flat up to 1.7kHz.

5.7 Conclusions and discussion

In this chapter the performance of discretized secondary source arrays has been explored. We saw that by fulfilling the spatial Nyquist criterion, wave fields can be synthesized perfectly with a sampled secondary source array.
However, due to technical and economical reasons, the number of channels, that can be processed by a certain hardware configuration and the number of loudspeakers that can be driven, may be limited. Therefore in practical applications, we often have to deal with undersampled loudspeaker arrays, which leads to spatial aliasing. As a result of spatial aliasing, the temporal and spatial properties of the synthesized wave fields are distorted. This may lead to coloration of the perceived sound and possibly to worse localization performance.
To suppress spatial aliasing, an approach has been proposed for reducing the spatial bandwidth of synthesized wave fields in favor of the temporal bandwidth. The discrete wave field synthesis process can be divided into four steps:

I. Spatial anti-aliasing filtering
II. Spatial sampling
III. Spatial synthesis
IV. Spatial reconstruction filtering

It was shown that this rather theoretical approach could be translated very well into a practical method to reduce the spatial bandwidth. On the one hand, the relationship was established between spatial anti-aliasing filtering and aperture limitation of the array, and, on the other hand, spatial reconstruction filtering was related to the application of directional loudspeakers.
Inherent to the concept of spatial bandwidth reduction for linear arrays is the decrease of the size of the receiver area, in which the sound field is synthesized correctly. A solution for this problem was found by applying bent loudspeaker arrays. By shaping the loudspeaker array the beneficial effect of spatial bandwidth reduction can be combined with a large aperture of the array, yielding a large receiver area.

6

Psycho-acoustical aspects of synthesized sound fields

In the previous chapters an extensive physical analysis was conducted on the performance of finite and discretized loudspeaker arrays. The transition from the continuous to the discrete domain introduced a restriction on the maximum temporal frequency that can be synthesized correctly. With the use of spatial bandwidth reduction techniques, this upper limit can be increased, but it does not change the fact that beyond the spatial aliasing frequency, the synthesized wave field differs from the 'ideal' primary wave field both spatially and temporally. Diffraction effects due to the finiteness of the loudspeaker array only have a small effect on low frequencies.
Besides physical performance measures, the subjective appreciation of synthesized sound fields is of great importance. Synthesized sound fields can be judged and compared to 'real' sound fields by a number of perceptual (subjective) characteristics. Since this study forms the basis of several wave field synthesis applications under various acoustical conditions, it is appropriate to start a systematic investigation of these perceptual aspects under free field conditions.
First, in this chapter we will investigate the relation between physical performance measures and subjective criteria for synthesized sound fields (section 6.1). Secondly, the vertical localization aspects of horizontal line arrays are considered (section 6.2). Finally, a preliminary investigation is conducted to the perceptual aspects of spatial aliasing and diffraction, based on listening tests with human subjects (sections 6.3-6.7).

6.1 Perceptual criteria and physical measures

In this section the main subjective criteria for direct sound will be summarized, and, if possible, be related to physical measures.

6.1.1 Loudness

The most fundamental property of a sound field is the sound pressure. In relation to loudness perception, its strength is expressed as an effective sound pressure:

$$p_{\mathrm{rms},T}(t) = \sqrt{\frac{1}{T}\int_{t-T}^{t} p^2(\tau)\mathrm{d}\tau}, \tag{6.1}$$

where T is the length of the integration interval. Usually, the strength is given in terms of a sound pressure level, defined by

$$L_p = 20\log\left(\frac{p_{\mathrm{rms}}}{p_0}\right) \qquad [\mathrm{dB\ SPL}], \tag{6.2}$$

with p_0=20 μPa. This reference value corresponds with the hearing threshold for normal human hearing at 1000 Hz.
Since the sensitivity of the ear depends on frequency in a complicated way, the sound pressure level does not correspond to the subjective loudness sensation. In order to account for this frequency-dependence of the ear, the measured signals can be filtered. The most often applied filter is the A-filter, which is the approximate inverse of the 40 phone curve (Boone et al., 1991). The level that is measured with this filter is called the sound level expressed in dB(A).

6.1.2 Localization of sound sources

The ability of human listeners to identify the direction of a sound source is called (acoustical) localization. The localizability, which is the localization resolution, is dependent of many factors such as: the source signal, source position, acoustics of the room. When a sound is presented from a loudspeaker, the perceived sound is generally described as coming from outside the head. In contrast, when a sound is presented with a headphone, the sound image is perceived within the head. The differential placement of sound images along the imaginary line between the ears is called lateralization of the sound images.
The physical cues which are available to the auditory system, depend on whether the sound source is positioned in the horizontal plane or in the vertical (median) plane. In the horizontal plane both interaural time differences (ITD) and interaural level differences (ILD) are available. In the vertical plane no ITD's and ILD's are present; only spectral cues can be used.
In the processing of ITD's, low frequencies (below 1600 Hz) play a dominant role. However, it has been shown by many authors (e.g. McFadden, 1976; Raatgever, 1980) that, also at high

frequencies, lateralization is possible on the basis of interaural time differences between the time-envelopes of the signals. ILD's mainly occur for high frequencies (above 2kHz), as a result of the shielding of the head.
Localization can be measured subjectively by performing a source-identification experiment, in which the relation between the physical direction of a sound source and the perceived (subjective) source direction is established. To quantify the localization performance Hartmann (1983) introduced some statistical error measures, which we will adopt. These error measures provide information about localization bias as well as localization variability.

In a source identification experiment the following quantities are given:

Number of sources:	L,
Number of trials per source:	M,
Number of trials per run:	$N=M\,L$,
Angular position of source for trial i:	S_i,
Angular response for trial i:	R_i,
Error for trial i:	$e_i=R_i - S_i$.

Using:

$$\Delta_i(k) = \begin{cases} 1 & i = k \\ 0 & i \neq k \end{cases}, \tag{6.3}$$

where k (=1, 2,...,L) is the source number, the statistical quantities for a given source k are defined by:

Rms error: $$D(k) = \sqrt{\frac{1}{M}\sum_{i=1}^{N}\Delta_i(k)e_i^2} \tag{6.4a}$$

Error: $$E(k) = \frac{1}{M}\sum_{i=1}^{N}\Delta_i(k)e_i \tag{6.4b}$$

Mean response: $$\bar{R}(k) = \frac{1}{M}\sum_{i=1}^{N}\Delta_i(k)R_i \tag{6.4c}$$

Standard deviation: $$s(k) = \sqrt{\frac{1}{M}\sum_{i=1}^{N}\Delta_i(k)[R_i-\bar{R}(k)]^2}. \tag{6.4d}$$

The rms error $D(k)$ is the rms average of the discrepancy between the azimuth of a given source k and the subject's responses to that source. The error $E(k)$ is the average deviation of the mean responses from the actual source, in which the sign of the deviation is taken into account. The

standard deviation *s(k)* is the standard deviation of the subject's responses; it differs from *D(k)* in that it is a measure of the rms discrepancy between the subject's individual responses to a particular source and his mean response to that source (not the actual source position).
The overall quantities for a given subject, averaged across the sources, are:

Run rms error:
$$\bar{D} = \sqrt{\frac{1}{L}\sum_{k=1}^{L} D^2(k)} \tag{6.5a}$$

Run error:
$$\bar{E} = \frac{1}{L}\sum_{k=1}^{L} E(k) \tag{6.5b}$$

Run standard deviation:
$$\bar{s} = \sqrt{\frac{1}{L}\sum_{k=1}^{L} s^2(k)}\,. \tag{6.5c}$$

The run rms error $\bar{D}$ may be regarded as the most meaningful single number to describe localization performance since it incorporates both localization bias and variability. Note that $\bar{D} \geq \bar{s}$.
Quantities averaged over subjects will be indicated by $\langle ... \rangle$.
The measure $\bar{s}$ quantifies a subject's consistency or resolution in localizing sound sources. A more direct measure of spatial resolution is the minimum audible angle (MAA), introduced by Mills (1958). In an MAA experiment the subject hears a sound from a reference source at a certain angle θ. He then hears a sound from a second source which is either to the left or to the right of the reference, by an angle $\pm\delta\theta$. The subject must declare whether the source is to the left or to the right of the reference. The MAA $\Delta\theta$ is defined as that value of $\delta\theta$ where the subject's responses are 75% correct.
This method cannot reveal any localization biases, which can be measured with a source identification experiment. If it is assumed that there is no bias ($\bar{D} = \bar{s}$) it can be shown (Hartmann, 1983) that under certain conditions the MAA $\Delta\theta \approx \bar{s}$.

6.1.3 Auditory distance perception

Besides the identification of the direction of sound sources, a second aspect is involved in sound localization; auditory distance perception. Compared to the study of localization in the horizontal and vertical planes, relatively little attention has been paid to distance perception in literature. The factors which influence the judgement and discrimination of auditory distance are thoroughly discussed in Blauert (1983) and Nielsen (1993). At least three cues for auditory distance judgement have been identified:

- Sound pressure level (the greater the SPL, the shorter is the judged distance)
- The amount of reverberation (the greater the ratio of direct to reverberant energy in the received signal, the shorter is the judged distance)
- Spectral shape of the perceived signal (the greater the high-frequency content of the stimulus, the shorter is the perceived distance)

In reverberant rooms the perceived distance of a source primarily depends of the direct-to reverberant ratio and is almost independent of the sound pressure level. In more dry environments the sound pressure level is the most important cue.
Finally, Wagenaars (1990) found that distance perception is better for broadband signals than for narrowband signals, as is also the case for horizontal localization.

6.1.4 Spaciousness

Spaciousness is usually loosely defined as the sense of being enveloped by the sound or as an increase in the apparent width of the source. In the past, many authors did not differentiate between listeners envelopment (LEV) and apparent source width (ASW). At present, it is widely agreed upon by acousticians that in normal concert halls ASW emerges from a decrease in correlation of the left- and right-ear signal, resulting mainly from early lateral reflections. Reflections arriving within 80 ms after the direct sound are generally considered as early reflections. Recent subjective experiments show that LEV is related to the level, direction of arrival, and temporal distribution of late arriving reflections (Bradley, 1994).
Since in this thesis we are only concerned with the synthesis of the direct sound, LEV is excluded from this research. In this study, the term spaciousness will be used for the subjective broadening of the source or, equivalently, the apparent source width. Most room acoustical measures for spaciousness are either based on the interaural degree of coherence, e.g., the interaural correlation coefficient *IACC* (Keet, 1968; Ando, 1985; Hidaka et al., 1991) or on the ratio of lateral sound energy to frontal or total energy arriving at the listener, e.g. the lateral energy fraction *LF* (Barron et al., 1981). Spaciousness also depends on the frequency content of the signal.
The *IACC* is defined as the maximum absolute value of the normalized interaural cross-correlation function, in the delay range of $|\tau|<1$ ms, in formula:

$$IACC = \max_{\tau} |\rho_{lr}(\tau)| \qquad |\tau| \leq 1\,\text{ms} \tag{6.6a}$$

in which the interaural cross-correlation function ρ_{lr} coefficient is defined as

$$\rho_{lr}(\tau) = \frac{E[s_l(t)s_r(t+\tau)]}{\sqrt{E[s_l^2(t)]E[s_r^2(t)]}}, \tag{6.6b}$$

where $s_l(t)$ is the left-ear signal and $s_r(t)$ is the right-ear signal. According to Ando, *IACC* can be measured either with a dummy head with microphones at the ear opening or with a human head, by setting small microphones at the entrances of the ear canals. If the two sounds are identical, *IACC*=1, and spaciousness is likely to be small, for uncorrelated signals *IACC*=0, in which case spaciousness is large.

A rather new measure of spaciousness is the central modulation coefficient (*CMC*) developed by Potter et al. (1995), which is based on a binaural interaction model: the Central Spectrum model (Bilsen, 1977; Raatgever and Bilsen, 1986).
In the Central Spectrum model the processing of interaural time or phase differences is described by linear system theory. The internal power distribution $P(f_c,\tau_i)$ can be calculated from the left- and right-ear signals:

$$P(f_c,\tau_i) = \int_{-\infty}^{\infty} |C(f_c,f)| \left\{ |S_l(f)|^2 + |S_r(f)|^2 + 2Re[S_{lr}(f)e^{j2\pi f\tau_i}] \right\} df, \tag{6.7}$$

where $C(f_c,f)$ denotes the transfer function of the peripheral auditory filter with center frequency f_c; $|S_l(f)|^2$ and $|S_r(f)|^2$ are the power density of the left- and right-ear signal respectively and $S_{lr}(f)$ represents the cross-power spectral density. The bandwidth and the shape of the auditory filters will be described in the next section.
By weighting this power distribution $P(f_c,\tau_i)$ in frequency, the non-ideal operation of the binaural interaction is incorporated. This weighting function is based on lateralization data of Raatgever (1980):

$$w(f_c) = \begin{cases} e^{-(f_c/600-1)^2} & f_c \geq 600\text{Hz} \\ e^{-(f_c/300-2)^2} & f_c < 600\text{Hz} \end{cases} \tag{6.8}$$

By integrating the frequency-weighted power distribution $P(f_c,\tau_i)$ over frequency, and by normalizing, the lateralization function $F_{lat}(\tau_i)$ is obtained

$$F_{lat}(\tau_i) = \frac{\int_{-\infty}^{\infty} w(f)P(f,\tau_i)df}{\int_{-\infty}^{\infty} w(f)df}, \tag{6.9a}$$

in which the subscript of f is omitted. This lateralization function may be regarded as a frequency weighted internal cross-correlation function. Next, the *CMC* is defined as the modulation depth in the lateralization function $F_{lat}(\tau_i)$ according to

$$CMC = \frac{\max[F_{lat}(\tau_i)] - \overline{\min}[F_{lat}(\tau_i)]}{C_n}, \tag{6.9b}$$

where $\overline{\min}$ denotes the mean of the minima on both sides of the maximum in the lateralization function $F_{lat}(\tau_i)$. The *CMC* is also able to predict the spaciousness for artificial stimulus conditions like anti-phase signals, while the *IACC* fails to predict the spaciousness for these signals. The constant C_n is a normalization constant, which we choose such that the *CMC* lies in the range of 0 to1, as the *IACC*.

6.1.5 Coloration

The notion of coloration, as we will use it, comprises both timbral changes and changes in pitch of a sound. The color of a signal is a quality of the sound, that may be changed by the surroundings in which the sound is produced or by the apparatus that (re)produces the sound. Coloration can be best perceived when the colored signal is compared to the original signal. The latter is not always available, and therefore coloration is often judged by comparing a signal to an 'internal reference' (Salomons, 1995).
The character of the coloration varies with the length of the impulse response, describing the transmission from source to listener. For short impulse responses (< 20 ms), which fall within the integration time of the ear, coloration can be mainly described as a spectral phenomenon. For longer impulse responses, also temporal characteristics of the sound are involved; e.g. infrapitch (Warren at al. (1980)). Since, in this study, we are mainly interested in the response of the wave field synthesis system, of which the length of the impulse response seldom exceeds 20 ms, the discussion is restricted to spectral coloration.

Coloration in rooms is mainly the result of reflections, which are repetitions of the direct sound. Also in synthesized wave fields repetitions of the main wave front, resulting from diffraction or spatial aliasing, may color the synthesized sound.
Adding to a signal, especially white noise, one delayed version gives rise to a tonal sensation. The pitch, which is associated with this sensation is called repetition pitch. For noise with one repetition (cosine noise), the pitch depends on the delay time τ and the phase difference $\Delta\phi$ between the direct sound and the repetition. For phase differences $\Delta\phi \neq 0°$, the pitch is ambiguous. Along the same lines as in Bilsen (1985), the (possibly ambiguous) pitch f_0 of cosine noise with delay τ and phase difference $\Delta\phi$ is found to be

$$f_0 = \frac{1}{\tau} \frac{n}{n \pm \frac{\Delta\phi}{2\pi}} \qquad (\Delta\phi \text{ in radians}) \tag{6.10}$$

in which the number n of the dominant harmonic must be inserted. Taking into account the experimentally found dominance about the fourth harmonic (n=4) with respect to pitch perception, the ambiguous pitch for e.g. $\Delta\phi = \pi$ is given by the values $0.88/\tau$ and $1.14/\tau$.
When more than one repetition is present, the pitch sensation will be determined by the relative strength and the delay-time ratios of the repetitions.

It is obvious to relate (spectral) coloration with a change in the temporal power spectrum of a sound. However, in order to relate the perception of coloration with the physical power spectrum, we need to know more about the frequency analysis by the ear.
The peripheral frequency analysis by the cochlea can be represented by a set of parallel bandpass filters with increasing center frequency; the 'auditory filters' or 'critical band filters'. These auditory filters evidently have a large influence on the perception of spectral coloration. Therefore, in an attempt to find objective criteria for coloration, the physical spectra, which can

be calculated or measured, should be analyzed in the same way as in the ear, i.e., by filtering with the auditory filters. Patterson (1986) proposed the following shape of the auditory filter:

$$C(f,f_c) = \left(1+\frac{4|f-f_c|}{W(f_c)}\right)\exp\left(-\frac{4|f-f_c|}{W(f_c)}\right) \tag{6.11}$$

where f is the frequency, f_c the center frequency and $W(f_c)$ the equivalent rectangular bandwidth. The equivalent rectangular bandwidth is defined as the bandwidth of a rectangular filter with unit amplitude, having the same area as the auditory filter:

$$\int_{-\infty}^{\infty} C(f,f_c)df = W(f_c)\,. \tag{6.12}$$

According to Patterson, the relation between center frequency and bandwidth is given by:

$$W(f_c) = 6.23\ 10^{-6} f_c^2 + 93.4\ 10^{-3} f_c + 28.5\,. \tag{6.13}$$

In figure 6.1a the critical bandwidth $W(f_c)$ as a function of center frequency f_c is shown. Examples of the auditory filters at different center frequencies are given in figure 6.1b.

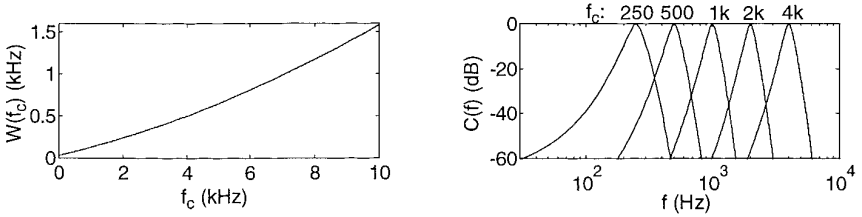


Figure 6.1: *Auditory filters*
a) Equivalent rectangular bandwidth of an auditory filter as function of the center frequency
b) Auditory filter at 5 different center frequencies (f_c=250, 500, 1k, 2k, 4k Hz)

Filtering of the incoming power spectrum $|S(f)|^2$ yields the internal power spectrum $|S_c(f_c)|^2$, described by:

$$|S_c(f_c)|^2 = \frac{\int_{-\infty}^{\infty} C(f,f_c)|S(f)|^2 df}{\int_{-\infty}^{\infty} C(f,f_c)df}\,. \tag{6.14}$$

In order to compare the internal power spectrum of an arbitrary signal with a white signal, the internal spectrum is normalized with the internal spectrum of that white signal.

Based on the internal spectrum, Salomons (1995) adopts an objective criterion, the A_0-criterion, for (spectral) coloration by stating that: "Coloration is perceptible if the maximum modulation depth of the internal spectrum (i.e. the level difference between the global maximum and minimum) exceeds a certain threshold A_0." Supported with data obtained with correlated and uncorrelated harmonic cosine noise stimuli, she estimates A_0=1.5±0.2 dB. However, based on data obtained with diotic hall signals, she found A_0=2.5±0.5dB.
Although successful for predicting coloration thresholds of broadband rippled noise stimuli, this method is unable to predict thresholds for purely local changes in the spectrum. This is supported by several studies, of which some will be reviewed now.
Green (1987) reports measurements on the ability to detect a change in a narrow frequency region. He showed that the detectability is influenced by the frequency at which the intensity change is produced. The midfrequency region (500-2000 Hz) appears to be the most sensitive. Changes in spectral shape near 7 kHz are more difficult to detect than the same type of change near 1 kHz by about 12 dB.
In earlier experiments Green (1983) found that: "the detection of a change in spectral shape, or profile analysis, appears to be mediated by comparisons across widely separated frequency 'channels' rather than by local comparisons among adjacent frequency regions." The estimated relative bandwidth of these frequency channels (16-32%) is somewhat larger than the accepted values of the critical bandwidth (about 15%), for which no satisfying explanation can be found (Bernstein, 1987).
In conclusion it can be stated that the detection of coloration is not limited to peripheral processes. Some aspects of profile analysis suggest a recombination of information from different frequency channels (or critical bands) at a higher central level. Nonetheless, the perception of coloration starts with a frequency analysis by the peripheral auditory filters. And, though the A_0-criterion cannot give exact coloration threshold predictions for all kinds of stimuli, it can serve as a qualitative measure of the strength of coloration, especially for signals with repetitions.

6.2 Line arrays: the sound localization aspect

In many direct sound enhancement applications the loudspeaker array will not be positioned in the same plane as the primary source and the listeners. As a result, the synthesized wave fronts differ from the primary wave fronts in the plane perpendicular to the array, as shown in section 3.6. In the vertical plane, the synthesized wave fronts become circular and originate from the array (figure 6.2a).
In the horizontal plane of the listener the synthesized wave fronts are almost identical to the wave fronts of the primary source (figure 6.2b). This means that the azimuthal direction of the amplified sound coincides with the direction of the primary source. This is an extremely important property, since in a theatre or concert hall the horizontal localization of sounds is much more important than the vertical localization.

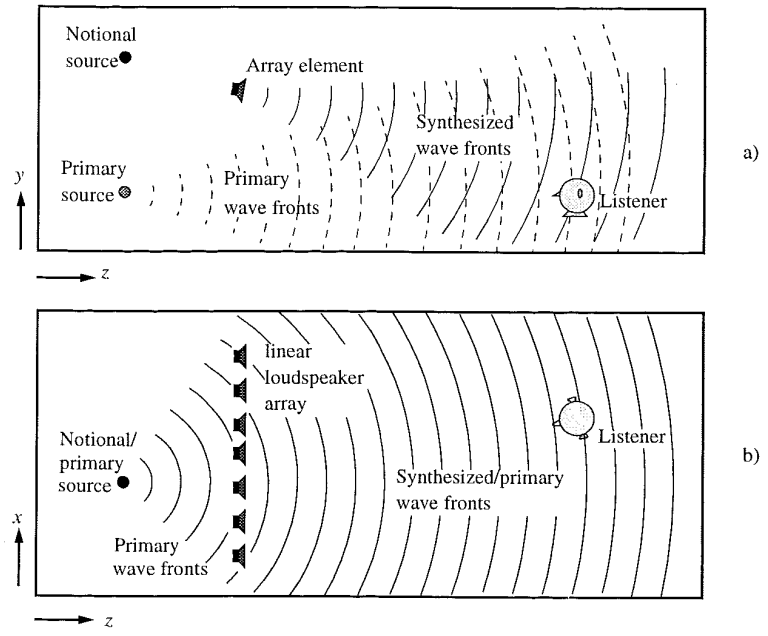


Figure 6.2: *Using a linear array, the synthesized wave fronts become circular in a plane perpendicular to the array and originate from the array. In many situations this will lead to an elevated notional source (a). In the horizontal plane of the listener the synthesized wave fronts are very accurate (b).*

In the vertical plane, i.e., in the median plane of the listener, there are two sound sources; a real source which is visible for the listener and a notional source, invisible for the listener.
In this context, it is noteworthy to quote what Hartmann (1983) found: "the azimuthal localization of impulsive tones significantly improves when the ceiling of a reverberant room is lowered". When the ceiling was lowered, its reflection, which has the same azimuthal angle as the direct sound, reached a listener much earlier than the side wall reflections, which have differing azimuths. Hartmann hypothesized that this gave the ceiling reflection an increased importance and accounted for the enhancement of localization accuracy. He also interprets this result as indicating a limitation of the precedence effect as it operates in rooms. Hartmann states that his results are most economically explained if one assumes that reflections from floor and ceiling reinforce the sense of localization, while reflections from the side provide only confusion.
In our situation, the wave field, synthesized by a lifted horizontal array, can also be regarded as a low ceiling reflection, i.e. a reflection with the same azimuthal angle as the direct sound. So, when the arrival time of the synthesized wave field is adjusted such that it succeeds the direct

sound, but precedes the first side wall reflection, the listener's sense of horizontal location can even be reinforced.
With a configuration like that of figure 6.2a the danger exists, that the apparent source position shifts upwards, when the intensity difference between the synthesized wave field and the direct sound is too large. Fortunately, there is some evidence that the precedence effect also works in the median plane, and, therefore, that the precedence effect does not require interaural differences for its operation (Rakerd, 1992). Although the precedence effect does not apply to the vertical plane in the same way as it does to the horizontal plane (Somerville, 1966), the elevation of the sound image can be reduced by delaying the synthesized wave field with respect to the direct sound by 2-20 ms. For clicks this is demonstrated in figure 6.3, which shows the

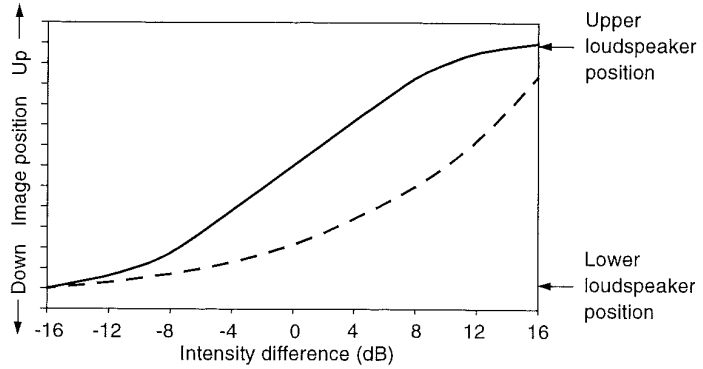


Figure 6.3: *Effect of the intensity difference between the upper and lower loudspeaker on the vertical image position. Solid line: speakers in phase. Dashed line: 20 ms delay on the upper loudspeaker. (Adapted from Somerville, 1966)*

effect of the intensity difference between the upper and lower loudspeaker on the vertical image position for the case that the loudspeakers are in phase and the case that the upper loudspeaker is 20 ms delayed.
The discussion on the vertical localization of synthesized sound fields will be resumed in chapter 8, where the localization in the median plane is explored further by performing measurements in a practical configuration with realistic sounds, such as speech and music.

6.3 Preliminary psycho-acoustical experiments

6.3.1 Description of the prototype system

At the start of the present research a wave field synthesis prototype system was available at the laboratory. This prototype system, designed by Vogel (1993), consisted of 48 loudspeakers having a width of 0.11 m, thus creating an array length of about 5.4 m when placed closely to each other. A complete overview of the setup of the prototype system can be found in Vogel's thesis.

Some important specifications will be summarized here:

- Sampling distance: Δx=0.113 m.
- Spatial aliasing frequency: f_{al}=1.5 kHz
- Number of secondary sources: 48
- Array length: L=48Δx=5.424 m.
- Type of secondary sources: 2½'' loudspeaker built-in into a box of $11 \times 11 \times 20$ cm
- Driving signal: 2½D Rayleigh I operator
- Temporal sampling frequency: f_s=16 kHz

Vogel conducted some objective and subjective tests with the prototype loudspeaker system. The objective test results will not be shown here, but the subjective tests, consisting of several source identification experiments, will be re-analyzed and presented in a different form in section 6.4.

6.3.2 Objectives of the present perception experiments

The main objective of the subjective listening tests is to get insight into the main differences between 'real' and synthesized sound fields. Since spatial aliasing and diffraction occur in synthesized wave fields, the subjective performance may be degraded. The listening tests were performed with the prototype system, described in the previous sub-section. The results of the subjective evaluation with this prototype system will be taken into account in the design of the direct sound enhancement system (DSE-system), which will be given in the next chapter.

The preliminary subjective listening tests can be discerned into localization experiments, spaciousness experiments and coloration experiments.

6.4 Localization experiments

The purpose of the source identification experiment (done by Vogel) was to determine the ability of subjects to localize a synthesized notional source. In particular, Vogel was interested in the effect of spatial aliasing on the localization performance.

6.4.1 Method

A monopole wave field was synthesized in an anechoic room. The notional source position was chosen randomly from a pre-defined set of 20 possible locations. Three subjects participated in this experiment. The subject's task was to indicate the perceived location of the source by giving the number of the loudspeaker at the intersection of the line from the simulated source to the subject (figure 6.4). The sources were positioned at a line 2 m behind the array. The listener

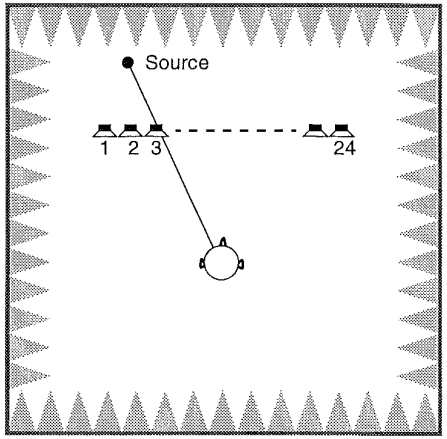


Figure 6.4: During the source identification experiment, the subjects were asked to indicate the notional source position by giving the number of the loudspeaker at the intersection of the line from source to listener. The loudspeaker array was placed in an anechoic room.

was seated 3 m in front of the array and was free to turn his head. Each experimental run consisted of 60 trials, 3 presentations of 20 different sources. No feedback was given during the run.

Several source signals have been used during the experiments:

1. Speech; recording of a novel fragment, read by a male voice.
2. Band limited pink noise; bandwidth 100-1500 Hz.
3. Broadband pink noise; bandwidth 100-8000 Hz.

6.4.2 Results and discussion

Data from the individual runs, obtained by Vogel, were re-analyzed, and the following statistics are calculated (see section 6.1.2): (1) Rms error D, which indicates the overall accuracy of localization. (2) Signed error E, which indicates any systematic response biases to the left (negative E) or right (positive E). (3) Response standard deviation s, which indicates the variability of subject's choices.

The results of each subject are averaged across all source positions, according to eqs.(6.5a-c). The standard deviation of the run values have been determined from a sample of 20 sources. Finally, the results of all subjects have been pooled, yielding a sample based on 3 subjects and 20 sources. The sample mean and the standard deviation of the mean have been calculated. The results are given in figure 6.5 for all three stimuli.

It can be concluded from these figures that the overall localization accuracy averaged across all subjects $\langle \overline{D} \rangle$ for all stimuli is very good (less than 2°). The localization accuracy for the broadband noise stimulus ($\langle \overline{D} \rangle$ =2.0°±0.2°) is somewhat worse than for the other stimuli ($\langle \overline{D} \rangle$ =1.7°±0.1°). This is, however, not a significant difference. Apparently, the effect of spatial aliasing above 1.5 kHz does not degrade localization performance for the broadband noise stimulus. Our broadband noise result ($\langle \overline{D} \rangle$ =2.0° (0.2)) does not significantly differ from that of Hartmann (1983), who found an rms error, averaged across subjects of $\langle \overline{D} \rangle$ =2.3° (0.6) in a broadband noise localization experiment in an absorbing room.

Note that for all subjects and all stimuli, the localization accuracy $\overline{D}$ is larger than the response variability $\bar{s}$ which can be explained by the fact that the localization accuracy $\overline{D}$ comprises both the response variability $\bar{s}$ and the response bias $\overline{E}$.

6.5 Localization resolution experiments

In addition to the localization experiments performed by Vogel, which were re-analyzed in the previous sub-section, we are interested in the localization resolution or localizability of 'real' and synthesized sources. Therefore, a minimum audible angle (MAA) experiment has been conducted. The results of this MAA experiment will be compared to the results of the localization experiment.

6.5.1 Method

Synthesized and real sound fields have been registered with a KEMAR dummy head in an anechoic room. For the registration of the synthesized sound fields the dummy head is placed centrally, 3.75 m in front of the loudspeaker array. The source signal is broadband white noise (100-8000Hz). The azimuths of the notional sources were arbitrarily chosen in the range θ=(-10°, 10°), such that stimulus pairs could be made with $\delta\theta$ in the range 0°-2.5° with steps of 0.1°. Each stimulus pair with a certain $\delta\theta$ could be synthesized from three different main directions.

For the recording of the real sources, the complete loudspeaker array was removed except for one loudspeaker, which was used as a real source. Next, real source positions were chosen similarly as for the notional sources. During the recordings, the position of the dummy head was not changed.

In the experiments a two-interval, two-alternative forced-choice paradigm (2-AFC) was used. The recorded stimuli were presented to the subjects at a level of 50 dB SL through headphones.

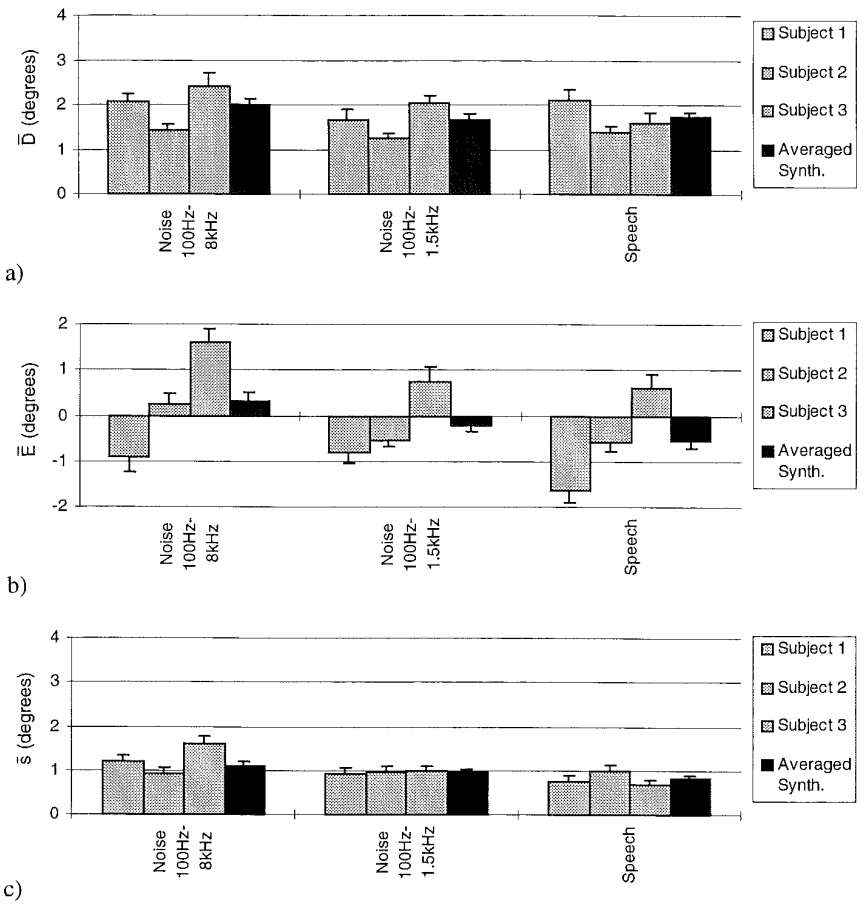


Figure 6.5: *Results of the localization experiment (derived from Vogel's data). The error bars represent the standard deviation of the run values.*
a) Run rms error $\overline{D}$
b) Run error $\overline{E}$
c) Run standard deviation $\bar{s}$

Since the dummy head used in the experiments were provided with microphones at the position of the eardrums, an inverse filter has been applied to correct for the ear canal transmission.
Each trial consisted of a reference stimulus, presented for 1000 ms, followed by 300 ms silence, and a second stimulus with a duration of 1000 ms. The subjects must declare whether the second source is to the left or to the right of the reference. Azimuthal changes in the range 0°-2.5° were presented in random order. In order to avoid recognition of a stimulus pair by the subjects, each pair with a certain $\delta\theta$, was randomly chosen from the three possible pairs with the same $\delta\theta$.
Each experimental run consisted of 100 trials, 10 trials for each azimuthal change $\delta\theta$. Three subjects participated, each making three runs, so each azimuthal change is presented 30 times.
The experiment is repeated for:
1. Real sources, bandwidth 100-1500 Hz.
2. Real sources, bandwidth 100-8000 Hz.
3. Synthesized sources, bandwidth 100-1500 Hz.
4. Synthesized sources, bandwidth 100-8000 Hz.

The minimum audible angle is determined by calculating the likelihood $L_{\delta\theta}$ for each azimuthal change $\delta\theta$, yielding a likelihood distribution. The mean of the likelihood distribution is an estimate of the MAA. The variance of the likelihood distribution is a measure of the accuracy of the MAA.

6.5.2 Results and discussion

The MAA's for each subject and each stimulus condition, are given table 6.1. The standard deviation is added between brackets. It is striking that the averaged MAA's for real and synthesized sources are almost identical. For both situations, the MAA decreases for increasing bandwidth. A similar dependence of localization accuracy as a function of bandwidth, could not be found in the localization results of figure 6.5.

Table 6.1: *The minimum audible angle (MAA) under different stimulus conditions for three subjects. The averaged values across subjects are indicated. The standard deviation is added between brackets.*

	MAA (deg)			
	Real		Synthesized	
Subject	100-1500 Hz	100-8000 Hz	100-1500 Hz	100-8000 Hz
1	1.02 (0.40)	0.79 (0.07)	0.89 (0.23)	0.72 (0.21)
2	0.90 (0.15)	0.87 (0.16)	0.99 (0.07)	0.82 (0.11)
3	1.53 (0.21)	0.72 (0.21)	1.41 (0.23)	0.76 (0.22)
Averaged	1.15 (0.16)	0.79 (0.09)	1.10 (0.11)	0.77 (0.11)

The expected correlation between the MAA and the response standard deviation $\langle\bar{s}\rangle$ is confirmed on the basis of the present results; the averaged MAA for synthesized sources for band limited and broadband noise, are respectively 1.10° and 0.77°, while the response standard deviations $\langle\bar{s}\rangle$ are 1.10° and 0.97°.
MAA thresholds under anechoic conditions have also been reported for clicks by Perrot (1990), and for white noise bursts by Saberi (1991). They report a mean MAA of approximately 1.0° in the horizontal plane, which is in good agreement with the present results.
Based on the findings in the previous chapter, it was expected that the localization accuracy (D) and localization resolution (MAA), would become worse due to spatial aliasing. Clearly, the results do not answer this expectation.
In order to give an explanation for this result, one could hypothesize that: (1) the low frequencies (<1.5 kHz) dominate localization performance, and consequently mainly ITD cues are used for horizontal localization. (2) High-frequency localization (>1.5 kHz) is dominated by the correct first arriving high-frequency wave front, which is leading the 'aliasing tails'. Based on our results, we cannot distinguish between these two hypotheses.
Support for the first hypothesis can be found in a study by Wightman and Kistler (1992). In their study they were able to independently assess the weights of ITD and ILD cues in horizontal localization. By digital manipulation of the signals, stimuli were created that had ITD cues associated with one location in space and ILD cues associated with a different position; subjects were required to indicate the location of these stimuli. The results were straightforward: For broadband noise, subjects' responses were governed completely by the ITD cue. For high-pass filtered noises, on the other hand, responses were governed by the ILD cue. Consequently, they concluded that in the horizontal plane, low-frequency ITD's are the dominant cues for localization of broadband sounds, while ILD's are the dominant cues for localizing high-frequency sounds. Apparently, the ITD's in the envelopes of high-frequency signals do not contribute significantly to localization performance.
Since, in the processing of ITD's, low frequencies (below 1600 Hz) play a dominant role, we may expect that the spatial aliasing frequency (f_{al}=1.5 kHz) of the prototype system, is a lower limit. Increasing the spatial sampling distance Δx, yielding a decrease of the spatial aliasing frequency, would result in a worse localization resolution.
Therefore a second MAA experiment was done, in which the sampling distance was doubled (Δx=0.227 m), yielding a spatial aliasing frequency of f_{al}=750 Hz. The length of the array was held constant, so the loudspeaker array consisted of 24 elements in this experiment. For two subjects (1 and 3), the MAA was measured for synthesized sources with the same bandwidths as in the first experiment. The results are shown in table 6.2, and indeed indicate that the localization resolution decreases when we lower the spatial sampling rate.

Table 6.2: The minimum audible angle (MAA) for two subjects. The sampling distance was doubled in comparison to the first experiment. The averaged values across subjects are indicated. The standard deviation is added between brackets.

	MAA (deg)	
	Synthesized	
Subject	100-1500 Hz	100-8000 Hz
1	1.17 (0.19)	1.24 (0.31)
3	1.84 (0.25)	1.89 (0.33)
Averaged	1.51 (0.16)	1.57 (0.23)

6.6 Spaciousness experiments

A noteworthy effect that we discovered in informal listening tests was that the apparent source width of a synthesized noise source depends on frequency. Especially, when listening to moving synthesized sources, we noticed that the source consisted of a rather broad high-frequency image and a more compact low frequency image. When we replaced the noise source by a speech source, this effect vanished almost completely. To investigate these effects more systematically, we determined the subjective spaciousness of real and synthesized wave fields.

6.6.1 Method

Six noise stimuli of synthesized and real sources were recorded with a KEMAR dummy head in an anechoic room. Afterwards, the signals were filtered, yielding bandwidths of 100-1500 Hz and 100-8000 Hz, for both the real and synthesized signals. The *IACC* (eq.6.6a,b) of the six signals was measured for the four stimulus conditions (table 6.3). For the low-frequency condition (100-1500 Hz), the differences in the *IACC* of real and synthesized sound fields are very small. For the broadband condition (100-8000 Hz) the differences are much larger.
In a magnitude estimation experiment the influence of the *IACC* on the spaciousness has been investigated. The observers had to indicate the relative width of a stimulus with respect to a reference stimulus (source 2) which had a fixed value of 10 on the scale.
The spaciousness of real and synthesized sources was determined in two separate experiments; first, for the low-frequency condition, and, secondly, for broadband condition. The reason for this is that the subjects had great difficulty with estimating the width of a broadband stimulus with respect to the width of a band limited stimulus, and vice versa.
All stimuli were presented at a level of 50 dB SL through headphones. Four subjects participated in the experiment.

Table 6.3: The interaural cross-correlation coefficient ρ_{lr} (IACC) for real and synthesized wave fields at different bandwidths.

	IACC			
	Real		Synthesized	
Source	100-1500 Hz	100-8000 Hz	100-1500 Hz	100-8000 Hz
1	0.85	0.85	0.80	0.64
2	0.86	0.86	0.81	0.61
3	0.86	0.85	0.81	0.61
4	0.88	0.84	0.78	0.60
5	0.87	0.84	0.83	0.73
6	0.85	0.84	0.79	0.75

6.6.2 Results and discussion

The results of the magnitude estimation for the bandlimited stimulus condition, averaged across subjects, are given in figure 6.6a. The errors are estimated by the standard deviation among the

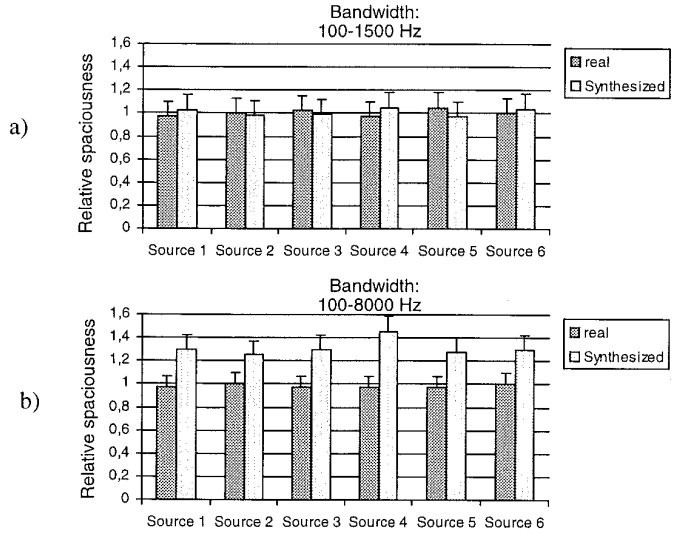


Figure 6.6: Spaciousness of real and synthesized sound sources for different bandwidth conditions
a) Bandwidth: 100-1500 Hz (frequency range below the spatial aliasing frequency)
b) Bandwidth: 100-8000 Hz (entire frequency range, including spatial aliasing)

subject's responses. From these data, it is clear that the differences in spaciousness between real and synthesized sources, for the low-frequency condition, are very small, which corresponds with the small differences in the *IACC*.
The results of the broadband stimulus condition are shown in figure 6.6b. These results indicate that the perceived source width is larger for synthesized sources than for real sources, which could be predicted on the basis of the *IACC* values.

6.7 Experiments on coloration due to diffraction

In listening tests with synthesized noise sources it is noticed that, when the listener moves his head, the coloration of the sound changes rapidly with the position of his head. From the results of the previous chapters it can be found that both diffraction effects and spatial aliasing effects give a distortion of the spectrum of the perceived sound. It is expected that aliasing effects mainly contribute to the coloration, since they produce a more severe distortion of the spectrum than the truncation effects, as will be shown in the next section.
To assess the strengths of the coloration due to diffraction and spatial aliasing independently, we investigated the audibility of diffraction in the absence of spatial aliasing. Therefore, in a threshold experiment, the array length is determined at which coloration is just (in)audible. The results of this experiment will be compared to the A_0-criterion, given in section 6.1.
The effect of diffraction can be described by the addition of two diffraction terms to a diffraction-free signal (see eq.(4.19)). In a first order approximation the diffraction contributions can be described by eqs.(4.27a,b). Note that the phase difference between the diffraction-free signal and the diffraction contributions has a fixed value of 135°, while the attenuation of the diffraction signal is frequency-dependent. The delay times τ_A and τ_B between the diffraction-free signal and the diffraction contributions are given by eq.(4.28).
It is expected that, especially with noise as the input signal, the delayed diffraction contributions give rise to a tonal sensation. In our situation, the ambiguous pitch of the diffracted sound field is determined by delay times τ_A and τ_B of both diffraction contributions; inserting $\Delta\phi$=135° (=3π/4 radians) in eq.(6.10) with n=4, the values $0.91/\tau_A$, $1.10/\tau_A$, $0.91/\tau_B$ and $1.10/\tau_B$ are found.

6.7.1 Method

Since it is impossible to create spatial aliasing-free sound fields with the prototype system, nor it is possible to create a diffraction-free sound field, which is needed as a reference, we will simulate the effect of diffraction. In chapter 4 analytical expressions (see eqs.(4.27a,b)) were found for the diffraction terms in eq.(4.19). Using these expressions it is possible to create digitally generated signals that have a very close resemblance with the desired signals.
We proceed as follows. A recording is made of a noise source (bandwidth: 60-8000 Hz) and is stored on a hard disk. This signal represents the diffraction-free signal, which can serve in the

experiment as the reference stimulus. Next, the signal is digitally processed with a filter approximating the ideal $\sqrt{1/2\pi jk}$ characteristic, according to eqs.(4.27a,b), and is also stored on the hard disk. This signal represents the diffraction signal, which can be added to the reference stimulus after applying a delay and a scaling factor, according to eq.(4.27a,b). By adding a second diffraction signal, diffraction from both ends of the array can be simulated.
The source is simulated at a central position, 10 m behind the array. The receiver position is simulated 10 m in front of the array at two different locations: at a central position and at a position 1/4 of the array length right from the center. At the central receiver position the diffraction contributions are in phase for all frequencies, while for the more lateral receiver position the phase relationship is frequency-dependent.
The signals created in this way, are diotically presented, which means that a possible beneficial effect of binaural decoloration is not taken into account. Therefore, the stimuli may be regarded as worst case stimuli with respect to the audibility of coloration.
In the experiments a four-interval, two-alternative forced-choice paradigm is used. Stimuli were presented in pairs; one pair comprises the reference and the reference plus two diffraction signals, while the other pair comprises two reference stimuli. The stimuli are presented to the subjects at equal sensation level (50 dB SL) through headphones. The duration of a stimulus is 600 ms and the time between the stimuli is 300 ms. The subjects must declare whether the deviating stimulus is presented in the first pair or in the second pair.
Fifteen different array lengths in the range 5-11 m have been chosen in random order. Each experimental run consists of 150 trials, 10 trials for each array length. Two subjects have participated, each making three runs, so for each array length stimuli have been presented 30 times. Next, the percentage correct answers of the total number of trials has been calculated for a each array length. A maximum likelihood estimator was used to determine the threshold.

6.7.2 Results and discussion

It was noted during the experiments by both subjects that the dominant cue was a change in the low-frequency contents of the signal.
The diffraction thresholds are given in table 6.4, for both subjects and each receiver position. The standard deviation is added between brackets. The thresholds for the central receiver position are a bit larger than the thresholds for the lateral receiver position. However, the difference is not significant at a high confidence level.
The subjective results may be compared to the A_0-criterion, given in section 6.1. Therefore, the spectra of the stimuli at the averaged threshold level and the spectra convolved with the auditory filters are shown in figure 6.7. Figures 6.7a and b give the incoming spectrum $S(f)$ and the internal spectrum $S_c(f_c)$ at the central receiver position respectively. Figures 6.7c and d show the results for the lateral receiver position. Since diffraction is stronger at low frequencies, and consequently the highest modulation depth is found at low frequencies, only the frequency region between 60 and 1500 Hz is shown. The global maximum and minimum of the internal spectrum in the figures 6.3b and d are taken to obtain the threshold A. The maximum modula-

Table 6.4: Diffraction thresholds expressed in the length of the array. The receiver is simulated at two positions: a central position and a lateral position. The averaged values across subjects are indicated. The standard deviation is added between brackets.

Subject	Diffraction threshold (m)	
	Receiver at a central position	Receiver 1/4 of the array length to the right from the center
1	9.11 (0.43)	8.14 (0.30)
2	7.80 (0.19)	7.61 (0.25)
Averaged	8.46 (0.23)	7.88 (0.20)

tion depth equals A=3.3±0.1 dB for the central receiver position, and A=2.1±0.1 dB for the lateral receiver position. Both values of A well agree well with the expected threshold A_0=2.5±0.5 dB, based on diotic hall signals (see section 6.1).

Note that the results are given for one particular configuration. Extension of the results for various configurations would require the collection of more data. However, using the objective A_0-criterion for coloration with $A_0 \approx 2.5$ dB, the diffraction threshold can be predicted for other configurations.

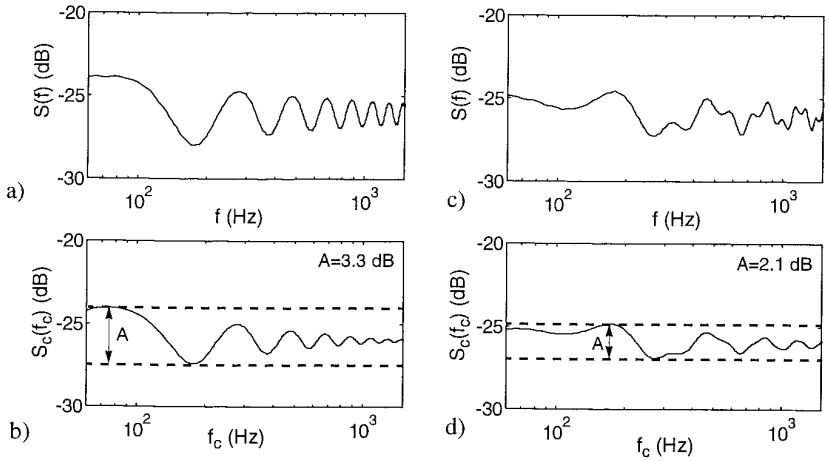


Figure 6.7: Spectra of the stimuli at the diffraction threshold
a) Spectrum S(f) at the central receiver position
b) Internal spectrum $S_c(f_c)$ at the central receiver position
c) Spectrum S(f) at the lateral receiver position
d) Internal spectrum $S_c(f_c)$ at the lateral receiver position

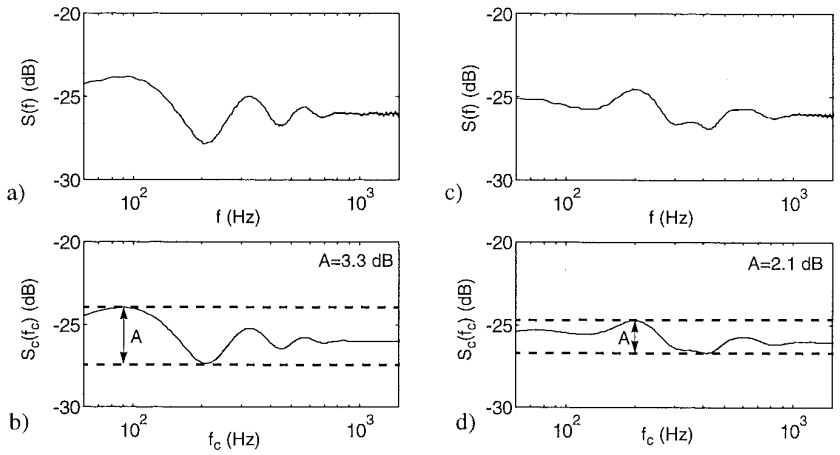


Figure 6.8: Spectra of the signals for the case that the loudspeaker array is tapered.
a) Spectrum S(f) at the central receiver position
b) Internal spectrum $S_c(f_c)$ at the central receiver position
c) Spectrum S(f) at the lateral receiver position
d) Internal spectrum $S_c(f_c)$ at the lateral receiver position

Next, we calculate the internal spectrum for the case that the driving signals are tapered. The configuration remains the same. The results are plotted in figure 6.8. Figures 6.8a and b give the incoming spectrum $S(f)$ and the internal spectrum $S_c(f_c)$ at the central receiver position respectively. Figures 6.8c and d show the results for the lateral receiver position. Note that, although the modulation in the spectra has decreased for higher frequencies, the maximum modulation depth, which is found at low frequencies, has not changed, in comparison with the previous results. This could be expected in view of the results of section 4.2.4. In that section it was shown that tapering yields a significant improvement for high frequencies, while no or very little improvement was found at low frequencies.

6.8 Coloration due to spatial aliasing

To investigate the coloration due to spatial aliasing, we will consider again the configuration of the previous section; the notional source is positioned centrally, 10 m behind the array. The wave field of the white noise source, with a bandwidth of 60 Hz-10 kHz, is synthesized and registered at a line parallel to the array, 10 m in front of the array. The array consists of 34 monopole loudspeakers at distances of Δx=0.25 m, yielding an array length of 8.5 m, such that coloration due to diffraction is just inaudible. The maximum angle of incidence φ_{max}=23°, yielding a spatial aliasing frequency of f_{al}=1.74 kHz. Figure 6.9a shows the x-t diagram of the

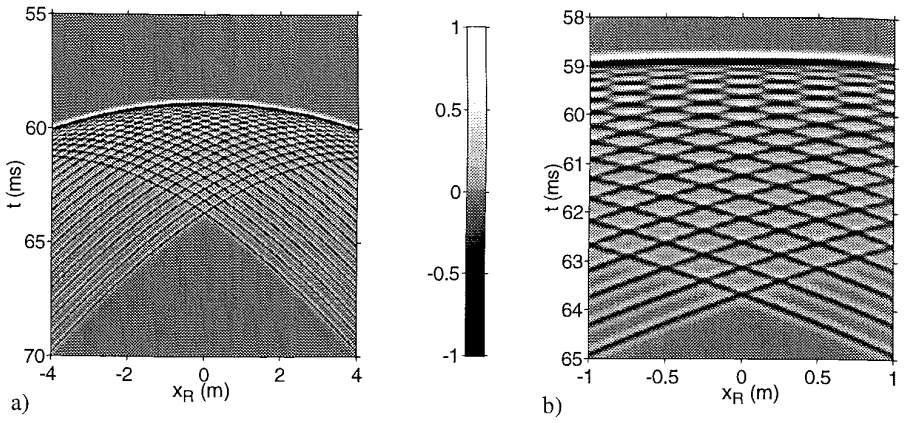


Figure 6.9: *Wave field synthesized with a discrete monopole source array (Δx=0.25 m). Temporal bandwidth: 60Hz-10kHz. In the aliased pressure distribution following the main wave front periodic structures are visible.*
a) x-t diagram of the synthesized wave field p (-4<x_R<4m)
b) x-t diagram of the synthesized wave field p (-1<x_R<1m).

synthesized wave field p for -4<x_R<4 m. By enlarging the center part of the diagram, as shown in figure 6.9b, periodic structures become clearly visible. Directly behind the main wave front, at equal distances in the x-direction, zero-pressure areas are visible. The distances between these 'nodes' are equal to the sampling interval of the secondary source, viz. 0.25 m.

Figure 6.10 shows the physical spectra and the internal spectra at three receiver positions. The receiver positions are chosen such that x_R=0, $\Delta x/4$ and $\Delta x/2$ respectively. The first receiver at x_R=0 is positioned at the perpendicular bisector of the line connecting the two center loudspeakers of the array.

It can be noticed that due to the auditory filters, the fine-structure in the spectra is smoothened. However, even for very small receiver displacements, the internal spectrum S_c changes very strongly for high frequencies; a distinct spectral edge is visible at different locations in each internal spectrum (indicated with an arrow). This can explain the effect that when subjects move their heads slightly (half a sampling distance), while listening to broadband noise, a distinct change of color is perceived. Due to reciprocity reasons, the same effect occurs when the source is moving and the listeners are at a fixed position.

Further it can be noticed that the averaged amplitude of the frequencies above the spatial aliasing frequency is too high. This can be explained by recalling that the individual contributions of the loudspeakers do not cancel each other behind the main wave front for those frequencies (see e.g. figure 5.11a). By applying a filter to the source signal, the amplitude of the high frequencies can be corrected.

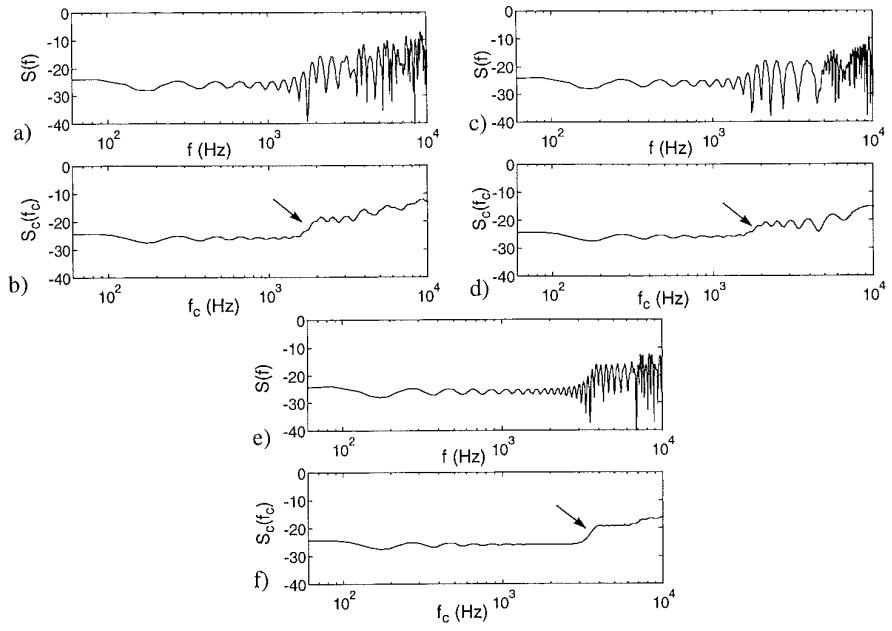


Figure 6.10: *Spectra at three receiver positions. Bandwidth: 60 Hz-10 kHz. In all three internal spectra a distinct edge is visible (indicated with an arrow).*
a,b) Physical spectrum S(f) and internal spectrum $S_c(f_c)$ at x_R=0
c,d) Physical spectrum S(f) and internal spectrum $S_c(f_c)$ at x_R=$\Delta x/4$
e,f) Physical spectrum S(f) and internal spectrum $S_c(f_c)$ at x_R=$\Delta x/2$

6.9 Conclusions and discussion

A preliminary investigation on the perceptual aspects of synthesized sound fields has been conducted under free field conditions. The results can be summarized as follows:

1. The localization performance with synthesized sound fields (spatial aliasing frequency: 1.5 kHz) is as good as with real sound fields.
2. The localization resolution with synthesized sound fields is very good and equals the resolution in real sound fields, provided that the spatial aliasing frequency is above 1.5 kHz. It is very likely that the excellent localization performance is caused by the dominance of ITD cues for broadband signals.
3. The imperfectness of the synthesized sound field above the spatial aliasing frequency gives rise to an increase of the apparent source width (ASW). The source broadening can be predicted on the basis of the interaural correlation coefficient (*IACC*). Large values of the

IACC predict a small source width, while small values of the *IACC* predict a large source width.

4. The coloration effects due to diffraction at the ends of the loudspeaker array are very small; the diffraction coloration threshold is approached for realistic array lengths. It has been shown that the coloration threshold can be predicted by the A_0-criterion applied to the human internal spectrum.
5. The coloration due to spatial aliasing changes rapidly as a function of the receiver position and is very distinct, especially when listening to white noise. Likely, the distinct change of color is produced by the movement of a spectral edge in the internal spectrum.

These findings will be incorporated in the design of the Direct Sound Enhancement system (DSE), which will be presented in the next chapter. In addition, a method will be presented which makes it possible to reduce the coloration effects due to spatial aliasing.

7

Design of the DSE system

In this chapter the implementation in the Direct Sound Enhancement system (DSE-system) will be described of the wave field synthesis techniques, which were discussed in the previous chapters. Although this chapter is mainly devoted to the design of the sound reproduction part of the DSE system, also strategies for the sound pick-up and methods for tracking of the sound sources will be mentioned.
In section 7.1 several methods for sound pick-up at stage will be investigated. Depending on the application of the DSE system in a specific room or hall, different kinds of sound recording techniques can be used. Next, in section 7.2, several loudspeaker array geometries will be considered for direct sound enhancement. In section 7.3 the synthesis operators are transformed to the time domain, in which form they can easily be implemented in a real-time processing system. Finally, section 7.4 gives a description of the DSE prototype system.

7.1 Sound recording at stage

7.1.1 Basic principles

As mentioned in the introduction of this thesis, the primary sound fields should be recorded at a certain plane $z=z_M$, extrapolated, and re-radiated at a plane $z=z_S$ (see figure 1.2 and 1.3). In this way both the temporal and the spatial properties are controlled, and an accurate representation of the desired primary wave field can be obtained in theory.
Even if the planar arrays are replaced by horizontal linear arrays, a major disadvantage of this method is the large amount of microphones needed for the sound pick-up. Another disadvantage is the huge computational power needed for the real-time extrapolation of the microphone array signals to the loudspeaker array.

In a theater or concert hall situation, the primary wave field is the superposition of the wave fields of several sound sources, e.g., actors, singers and/or instruments. To make possible an individual processing and manipulation of each source signal, each source signal should be recorded separately. As a result, the source signal is uncoupled from the spatial properties of the sound, which are disregarded. Usually, the sound sources are represented by *notional* monopole sources, which were introduced in section 1.3.
In many circumstances, the wave fields will be synthesized such that the notional sources coincide with the actual primary sources, however this is not obligatory. Since the source signals are known separately, the notional sources can be positioned anywhere, just by calculating the appropriate driving signals for the loudspeaker array. The primary source may even be absent, in which case the wave fields of *virtual* sources are synthesized.
It may be clear that in this way a very flexible processing of the sound fields can be achieved and many possibilities are offered for various artistic performances and virtual reality applications.
The complexity of the recording system depends on whether the sources are statically positioned on stage or are dynamically moving on the stage. For moving sources, a system is required to determine the position of each source, and to identify the sources.

7.1.2 Statically positioned sound sources

In a number of direct sound enhancement applications the sound sources are positioned at fixed locations. For instance, a lecturer in an auditorium behind a microphone-provided lectern, a music ensemble playing on stage, several people on stage having a forum discussion, etc.
In these cases the notional source positions are known a priori and can be programmed once-only in the DSE system. The source signal of each source is recorded by one microphone, preferably positioned closely to the source.
As an example consider the configuration of figure 7.1. The fixed positions of K sources are given by $\mathbf{r}_{\Psi,k}$ (k=1,...K). The wave field of each static source Ψ_k is recorded with a microphone M_k. When the microphones are placed closely to the sources, such that mainly the direct sound is recorded, the recorded pressure p_k is given by

$$p_k(t) = \frac{s_k\left(t-\frac{d_k}{c}\right)}{d_k}, \tag{7.1}$$

where d_k is the distance between the k^{th} source and its corresponding microphone M_k. The pressure p_k is a scaled and delayed version of the source signature $s_k(t)$. In the frequency domain the source spectrum $S_k(\omega)$, which is needed for the calculation of the driving signal Q_k, is given by

$$S_k(\omega) = d_k P_k(\omega) e^{jkd_k}. \tag{7.2}$$

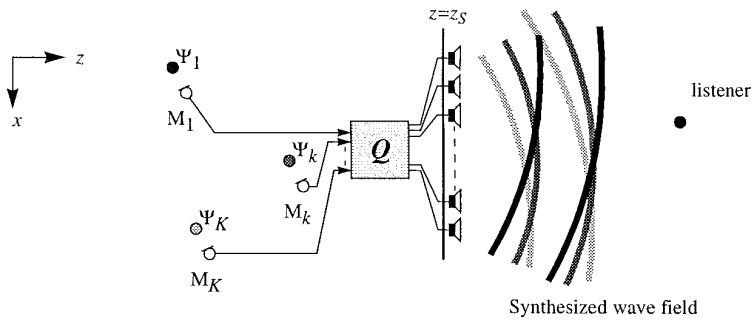


Figure 7.1: *The sound field of each static source Ψ_k (k=1,...K) is recorded with a microphone M_k. When the position of each source is known, the driving signal Q for the loudspeakers can be calculated.*

The calculation of the driving signal Q_k depends on the array configuration; e.g., for a linear monopole loudspeaker array, the 2½D Rayleigh I driving signal (4.10a) is used. The total driving signal Q for each loudspeaker at position $\mathbf{r}_L$ is the superposition of all driving signals Q_k, according to

$$Q(\mathbf{r}_L, \omega) = \sum_{k=1}^{K} Q_k(\mathbf{r}_L, \omega). \tag{7.3}$$

7.1.3 Moving sound sources

A more complicated situation arises when the sources are moving, which is the case in many theater performances, e.g., drama, musical, opera etc. In these situations basically two different recording techniques can be applied: remote miking and close miking.

I. Remote miking.

In this technique the sound is picked-up at a relative large distance from the source. Since, generally, in room acoustics the sound sources are positioned at a well-defined stage area, the stage can be divided into sub-areas, each covered by a directional microphone (see figure 7.2a). Hence, any signal recorded by a certain microphone can be assumed to be generated by a notional monopole source in the center of the corresponding sub-area of which the position is known (figure 7.2b). Next, the driving signals for the loudspeaker array can be calculated in the same way as in the previous section.
To avoid the notional source jumping from one sub-area the another when the source is moving around the stage, some overlap must be permitted of adjacent microphone beams.
Since the rms localization error $\langle\overline{D}\rangle$, as reported by Hartmann (1983), amounts 3.22° (0.75) in a reflecting room for noise band stimuli, the spatial resolution, obtained with the technique described above, is far beyond the desired resolution in practical applications. For example,

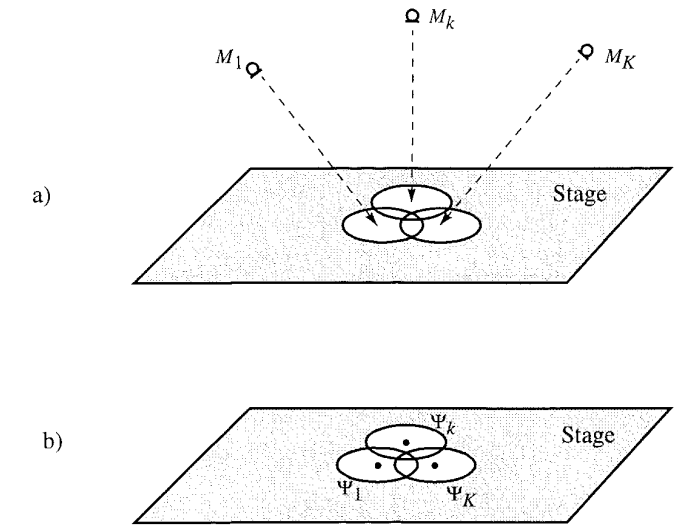


Figure 7.2: *By illuminating the stage with a number of highly directional microphones M_k (k=1,...K), the stage can be divided into sub-areas (a). Each signal recorded by a certain microphone M_k can be assumed to be generated by a notional monopole source Ψ_k in the center of the corresponding sub-area of which the position is known (b).*

when the listeners are seated at an average distance of 20 m from the stage, sub-areas should be created with a diameter of 1.13 m in order to meet the localization accuracy of the listeners. Since the microphones must be placed at large height from the stage floor for visibility reasons, this requirement is not feasible with standard directional microphones.

A solution for this problem is given by Van der Wal et al. (1996) by proposing the application of logarithmically spaced constant directivity microphone line arrays. In this way, with a minimum number of microphones, an adjustable frequency-independent directivity pattern can be achieved. Since the perception of lateral movements of sources on stage are more important than front-to-back movements, in many applications it is sufficient to apply one horizontal microphone line array. With one device it is possible to realize any desired selectivity by processing the microphone signals such that the array is steered simultaneously in several directions. Consequently, the stage is divided into 'strips', distributed along the left-right direction, as shown in figure 7.3a. Hence, any signal recorded in a certain steering direction can be assumed to be generated by a notional monopole source in the center of the corresponding strip of which the position is known (figure 7.2b).

The length of the microphone array is dependent on the desired resolution and the temporal bandwidth which is needed. When the distance between the microphone array and the stage is

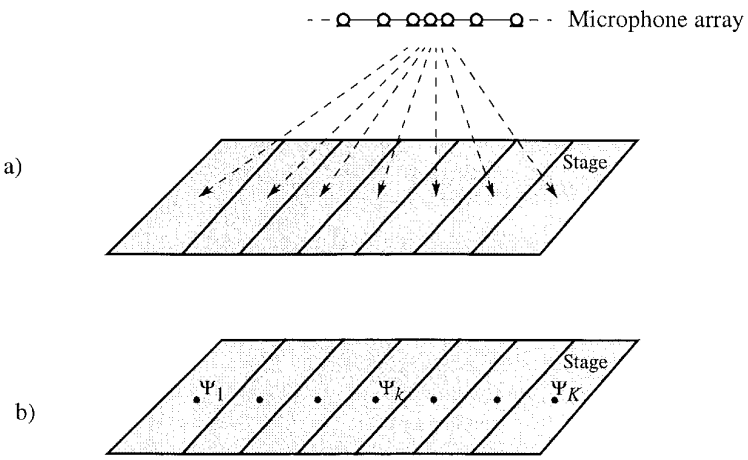


Figure 7.3: *With the application of logarithmically spaced microphone array the stage can be divided into strips (a). Any signal recorded in a certain lobe can be assumed to be generated by a notional monopole source Ψ_k in the center of the corresponding strip of which the position is known (b).*

too small with respect to the length of the array, that is, the Fraunhofer condition is not fulfilled, the array can be focussed on the center of each strip to compensate for this effect.

When, in addition to left-right information, also front-to-back information is required, the microphone line array must be extended to a planar array or must be supplemented with other line arrays, positioned at different depths. The individual microphones should have a directivity pattern, such that sufficient resolution in the front-to-back direction is obtained.

With the practical application of a remote miking system, one must be aware of acoustical feedback, yielding unwanted coloration of the reproduced sound. The stability of the total system, sound recording and reproduction, depends on several factors such as the transducer directivities, relative positioning of the microphones and the loudspeakers, and the geometry of the room and the stage. To increase the gain before instability (GBI) many modern reverberation systems use time-varying electronic filters, in which basic schemes such as phase modulation or delay-time modulation are applied (Svensson, 1995). Although rather successful for the improvement of the stability in multi-channel reverberation systems, it is very difficult to apply these techniques to increase the GBI in a direct sound enhancement system. The main drawback of time-varying filters applied to the direct sound is given by the fact that the resulting frequency modulation is extremely well audible.

In order to reduce acoustic feedback effectively, and to achieve the utmost flexibility for moving sources in several sceneries, a close miking technique is preferred.

II. Close miking.
In this technique the sound is picked-up very close to the source, preferably with wireless microphones attached to the head or body. In this case no information is obtained about the position of the sources and consequently the recording system must be extended with a so-called 'source tracking' system. Such a tracking system should be capable to determine the location and identity of each moving object, at least in 2 dimensions, viz, the horizontal plane.

7.1.4 Tracking systems

In many applications there is a need for accurately measuring the positions of objects. For example, in virtual reality applications a tracking system is needed to provide real time adjustments to views that are projected to virtual reality users, and also in robotics information about the position is crucial.
Current tracking devices, which are commercially available and may be applicable in a theater situation, are based on electromagnetic, acoustic, or optical technology. An overview of each of these approaches follows, together with a discussion of their advantages and disadvantages.

- *Optical tracking devices.*
  Basically, the setup is as follows: One or more cameras are mounted on the ceiling or a fixed frame and a few LEDs are placed at the object. The projections of the LEDs on the image of the camera contain enough information to uniquely identify the position and orientation of the object. To track several objects simultaneously, the system must be completed with a pattern recognition algorithm, which should be able to identify a certain LED-structure on each object. Generally, this is a very complicated task and very troublesome.
  Optical trackers in general have high update rates, and sufficiently short lags. However, they suffer from the line of sight problem, in that any obstacle between sensor and source seriously degrades the tracker's performance. Ambient light and infrared radiation also adversely affect optical tracker performance. As a result, the environment must be carefully designed to eliminate as much as possible these causes of uncertainty; a demand which is difficult to meet in a theater.
- *Electromagnetic tracking devices.*
  These devices function by measuring the strength of the magnetic fields generated by three electro-magnets, oriented perpendicular to one another. These electro-magnets are embedded in a small unit that is attached to the object. By sequentially activating each electro-magnet, and measuring the magnetic fields at a receiver at a reference position, it is possible to determine the position and orientation of the sending unit. These tracking units may experience interference while operating in the vicinity of devices that produce magnetic fields, as well as metal objects that disrupt magnetic fields. Another disadvantage of these tracking devices is that the working volume tends to be rather small.
- *Acoustic tracking devices.*
  These devices use ultrasonic sound waves for measuring the position and orientation of the target object. There are two ways of doing this: so-called time-of-flight tracking and phase-coherence tracking.
  Time-of-flight tracking works by measuring the amount of time that it takes for sound emitted by transmitters on the target to reach sensors located at fixed positions in the environment. By triangulation the position of the object can be calculated.
  Time-of-flight trackers typically suffer from a large latency, brought about by the low speed of sound in air. Of course, another problem is that the speed of sound in air is affected by environmental factors as temperature, barometric pressure and humidity.
  Phase coherence tracking works by measuring the difference in phase between sound waves emitted by a transmitter on the target and those emitted by a transmitter at some reference point. As long as the distance travelled by the target is less than one wavelength between updates, the system can calculate the position of the target.
  Since phase coherence tracking devices work by periodic updates of position, rather than by measuring absolute position at each time step, they are subject to error accumulation over time. As with optical devices, ultrasonic acoustical devices suffer from the line of sight problem.
  To get round the line of sight problem, a tracking method has been proposed which uses the source signal itself to determine the source position. By correlating the signal of the (wireless) microphone, which is attached to the source, with the signals recorded by microphones at fixed positions in space, the time-of-flight of the source signal to each microphone can be calculated. The performance of such a tracking system needs to be explored further. Especially, attention must be paid to the latency of the system, which will depend on the spectral contents of the signal and the chosen correlation length.

It can be concluded that the 'ideal' tracking system does not exist yet. Each method has its own advantages and disadvantages. Researchers and industry are actively seeking the robust tracking device with a large working volume, high accuracy and resolution, very short lag time and high update rate. Until such a device is found, the close miking technique in combination with an automatic notional source positioning, which is preferred for the DSE system, cannot be realized.

7.2 Array geometries for direct sound enhancement

The application of the 2½D Kirchhoff-Helmholtz and the generalized 2½D Rayleigh operators, which were derived in chapter 3, offers a powerful and flexible solution for many problems in sound control. The shape of the loudspeaker array can be adapted to any room or hall geometry. In this section several loudspeaker array geometries will be considered which can be used for direct sound amplification.

I. Linear loudspeaker arrays:

The linear loudspeaker array is a basic geometry which is applicable in many halls. Especially in halls with a rectangular (shoebox) shape, a linear loudspeaker array is very convenient for direct sound enhancement. Generally, the configuration in such rooms can be represented by the geometry in figure 7.4, in which the synthesized wave field for a source Ψ at stage is indi-

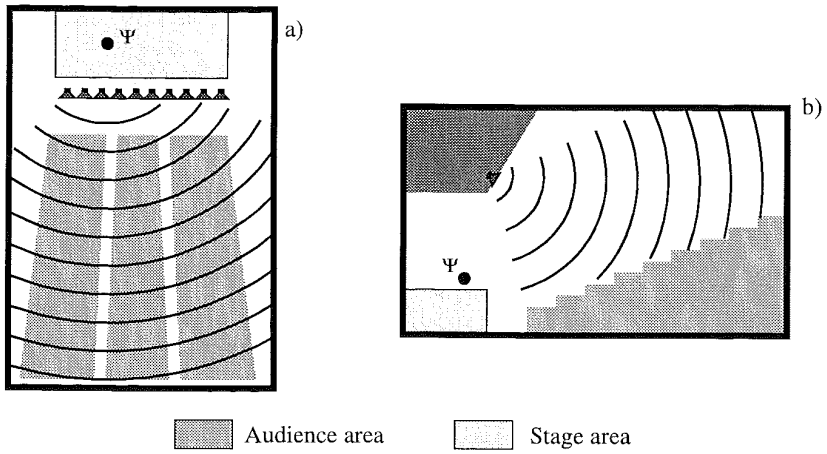


Figure 7.4: Schematic diagram of a linear array in a rectangular-shaped hall.
a) Top-view
b) Side-view

cated. The driving signal for the linear array, consisting of omni-directional loudspeakers, is a discretized version of the 2½D Rayleigh driving function Q, given by eq.(4.10a). In practice, the omni-directional loudspeakers may be replaced by loudspeakers with an arbitrary directivity pattern G. In theory it is possible to correct the driving signal for the directional behavior of the loudspeakers (de Vries, 1996). However, as shown in section 5.5.4, directional behavior is often desirable, since in this way a spatial reconstruction filter can be practically implemented. Moreover, application of directional loudspeakers reduces the risk of coloration due to acoustic feedback to the microphones.

II. Convex loudspeaker arrays:

An architectural form, which is often applied in auditoria and some modern concert halls, is the fan-shape. This shape is characterized by a wide divergent seating area, as shown in figure 7.5.

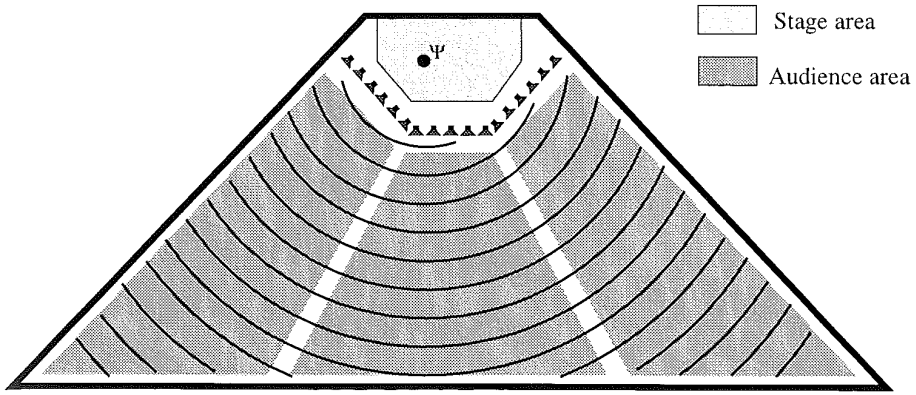


Figure 7.5: Schematic diagram of a convex loudspeaker array in a fan-shaped hall.

For this geometry, the application of a convex bent loudspeaker array is the proper solution. The driving signals are determined by the discretized version of the generalized 2½D Rayleigh operator.
A convex loudspeaker array has the additional advantage that the aperture as seen from a source Ψ at stage is very large, while the maximum angle of incidence $\max\{\varphi_{inc}\}$ on the array remains small (section 5.6.2), thus reducing spatial aliasing.

III. Concave loudspeaker arrays:

Finally, a third shape will be considered; the concave loudspeaker array. Basically, a concave array can be applied in two different forms; the array is 'embracing' the stage or 'embracing' the audience. Both situations are illustrated in figure 7.6. The configuration in figure 7.6a is not very efficient for direct sound enhancement, since for a source Ψ at stage only the central front part of the loudspeaker array will be used. However, for the generation of lateral reflections (represented by image sources), the array can be successfully applied, as indicated in figure 7.6a for one image source Ψ'. In this way a reflection pattern can be generated, which is related to any desired virtual hall. Like in the configuration with a convex array, the driving signals are determined by the discretized version of the generalized 2½D Rayleigh operator.
The configuration figure 7.6b is a special case for which no solution is obtained by the generalized 2½D Rayleigh integrals. Boone et al. (1996) derived a synthesis operator for the situation in which the notional source is positioned in front of the loudspeaker array by combining the wave field synthesis method with the principle of focussing. Notice, however, that the synthesized wave fronts radiate unidirectionally, i.e., in the figure from the array downward. In the area between the array and the notional source the wave field converges to the focus-point (i.e. the notional source Ψ). After having crossed the focus-point, the wave field becomes divergent

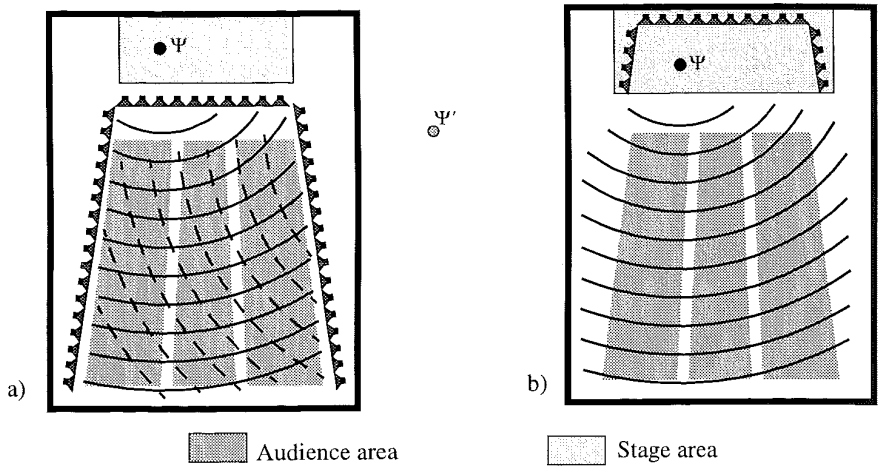


Figure 7.6: Schematic diagram of concave loudspeaker arrays in a rectangular-shaped hall.
a) The array is 'embracing' the audience
b) The array is 'embracing' the stage

and is perceived by listeners in the audience area as originating from the notional source position.

Using the focussed-source-technique new possibilities are obtained for sound reproduction. However, practical application of this technique for real-time sound amplification has one major disadvantage: since the sound of the real source is picked-up from the same position on which the synthesized wave field is focussed, a serious danger exists for instability of the system as a result of acoustic feedback. It is, however, a challenge to solve this problem with further research.

7.3 Wave field synthesis in the time domain

The wave field synthesis operators derived in chapter 3 are formulated in the space-frequency domain. The discretized versions of the operators were given in chapter 5. For the configuration of a discrete array L driven with N independent loudspeaker signals, we have to deal with the following driving signal Q_n as a function of the loudspeaker number n:

$$Q_n(\omega) = g_0(\mathbf{r}_{L,n})\sqrt{\frac{jk}{2\pi}}S(\omega)\cos\varphi_{\text{inc},n}\frac{e^{-jkr_n}}{\sqrt{r_n}}, \tag{7.4}$$

where, for the 2½D Rayleigh I case, the amplitude factor g_0 is given by

$$g_0(\mathbf{r}_{L,n}) = \sqrt{\frac{z_R - z_L}{z_R - z_\Psi}}, \tag{7.5}$$

and for the generalized 2½D Rayleigh I situation by

$$g_0(\mathbf{r}_{L,n}) = \sqrt{\frac{\Delta r_0(\mathbf{r}_{L,n})}{r_0(\mathbf{r}_{L,n}) + \Delta r_0(\mathbf{r}_{L,n})}}, \tag{7.6}$$

in which the 'minimum phase' secondary source-to-receiver distance $\Delta r_0(\mathbf{r}_n)$ is given by

$$\Delta r_0(\mathbf{r}_{L,n}) = \left|\mathbf{r}_{R,0}(\mathbf{r}_{L,n}) - \mathbf{r}_{L,n}\right|. \tag{7.7}$$

The distance r_n between the notional source Ψ and the n^{th} loudspeaker at $\mathbf{r}_{L,n}$ equals $|\mathbf{r}_n| = |\mathbf{r}_{L,n} - \mathbf{r}_\Psi|$ with $\mathbf{r}_\Psi$ the notional source position, given by $\mathbf{r}_\Psi = (x_\Psi, 0, z_\Psi)$. The angle $\varphi_{\text{inc},n}$ is the angle between the vector $\mathbf{r}_n$ and the normal $\mathbf{n}$ on the array.

Using $k = \omega/c$, the phase in the exponent of eq.(7.4) can be written as

$$kr_n = \omega\frac{r_n}{c} = \omega\tau_n. \tag{7.8}$$

Inverse Fourier transformation of the driving signal $Q_n(\omega)$ to the space-time domain yields the driving signal $q_n(t)$ as a function of the loudspeaker number n and time t, according to

$$q_n(t) = a_n[h(t) * s(t) * \delta(t - \tau_n)], \tag{7.9}$$

where $s(t)$ is the notional source signature, a_n the weighting factor for the n^{th} loudspeaker given by

$$a_n = \frac{g_0(\mathbf{r}_n)\cos\varphi_{\text{inc},n}}{\sqrt{r_n}}, \tag{7.10}$$

and τ_n the delay-time for the n^{th} loudspeaker given by r_n/c. The impulse response $h(t)$ is the inverse Fourier transform of the filter $\sqrt{jk/2\pi}$.

For multiple notional sources Ψ_k $(k = 1,\ldots,K)$ eq.(7.9) can be generalized to

$$q_n(t) = \sum_{k=1}^{K} q_{n,k}(t), \tag{7.11}$$

with

$$q_{n,k}(t) = a_{n,k}[h(t) * s_k(t) * \delta(t - \tau_{n,k})], \tag{7.12}$$

and

$$a_{n,k} = \frac{g_0(\mathbf{r}_{n,k})\cos\varphi_{\text{inc},n,k}}{\sqrt{r_{n,k}}}. \tag{7.13}$$

The total wave field synthesis operator for N independent loudspeaker signals and K notional sources can be described by the processing diagram of figure 7.7.

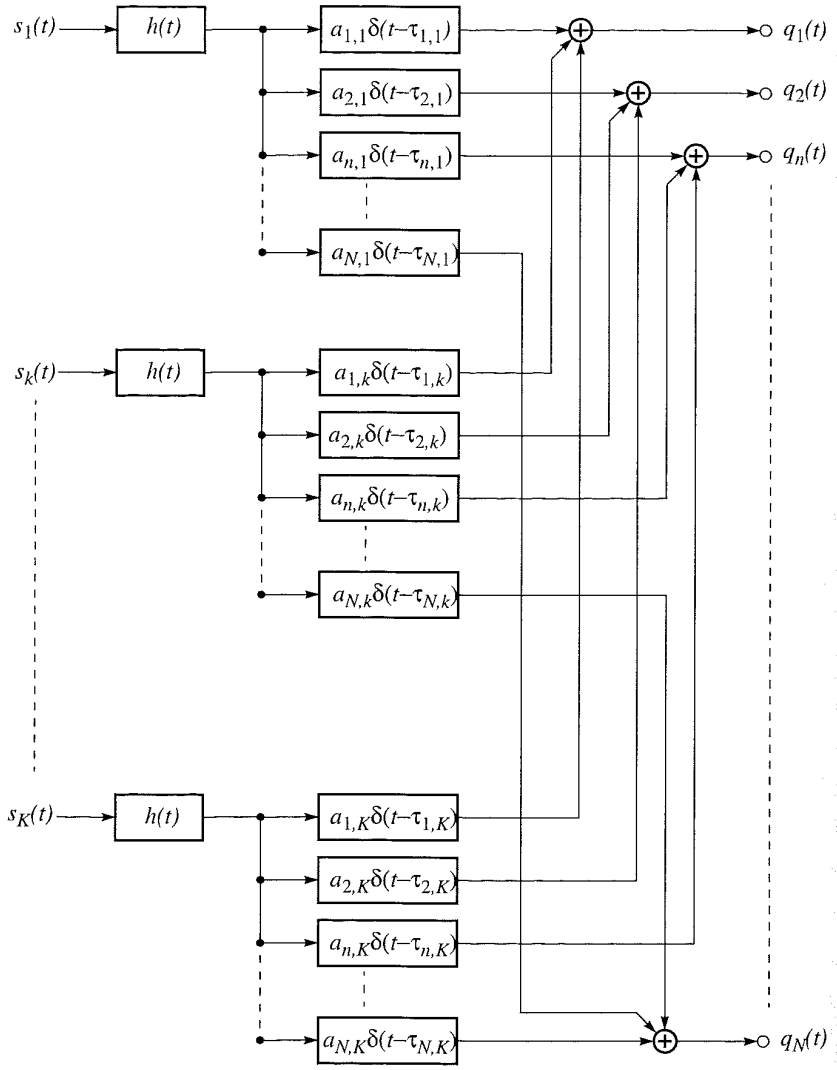


Figure 7.7: *General processing scheme for K notional sources and N independent loudspeaker signals. Each notional source signal is filtered with h(t). Next, the n^{th} driving signal $q_n(t)$ is calculated by summation of K weighted and delayed notional source signals.*

It is shown that the wave field synthesis operator can be described in the time domain simply by weighting factors and delay times. In this form the operator is efficiently applicable in a real time processing system. Ideally, the filter *h(t)* should be approximated by a FIR filter. In practice the filter *h(t)* can be efficiently combined with the correction filter for the non-ideal loudspeaker frequency response and also with the averaged amplitude correction for frequencies above the spatial aliasing frequency (see section 6.8).

7.4 The DSE proto-type system

Taking into account the results of the preliminary perception experiments, presented in the previous chapter, the practical design aspects of the direct sound wave field synthesis system will be summarized and a technical description of the hardware configuration will be given.
The implementation of the spatial bandwidth reduction technique, introduced in chapter 5, will be described. In addition, a method will be proposed to reduce the unwanted audible effects of spatial aliasing, i.e., the coloration due to undersampling of the loudspeaker array.

7.4.1 General set-up of the DSE system

The main goal of the DSE prototype system was to measure its performance under several acoustic conditions, e.g., anechoic conditions, auditorium acoustics and concert hall acoustics. Therefore a system was designed that could be easily mounted and dismounted in several rooms and halls. The individual parts of the system should have dimensions such that they could be transported in a van.
Based on the geometry of the rooms of interest and based on the findings in chapter 6, the following design criteria were compiled:

- The length of the loudspeaker array should be large (24 m), in order to meet the geometrical requirement in all halls.
- Preferably, the spatial aliasing frequency should be higher than 1.5 kHz.
- The temporal bandwidth should at least meet the requirements for speech and singing.
- It should be possible to investigate the effect of reconstruction filtering.
- A sufficiently high pressure level in all rooms should be obtainable.
- A large number (>8) of notional sources must be processed.
- Software driven digital signal processing should be applied for flexibility.
- The produced sound should have a high-quality: low distortion and a flat frequency response.

In order to meet these requirements, a custom made system was provided by a Dutch firm (Duran Audio BV).
The DSE proto-type system is provided with 96 output channels which are driving, via a spatial reconstruction filter, a number of 192 loudspeakers. The sampling distance Δx equals 0.25 m,

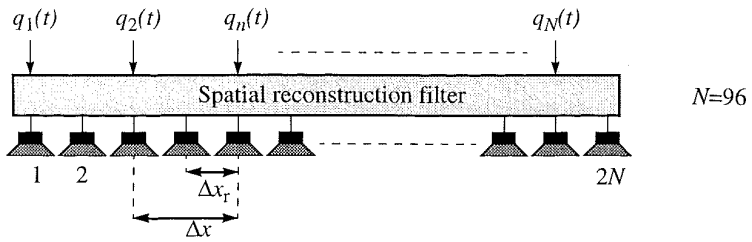


Figure 7.8: Schematic diagram of the loudspeaker configuration. The sampling distance Δx equals 0.25 m, while the loudspeakers are placed at a distance of Δx_r=0.125 m. The 96 output channels are distributed over 192 loudspeakers.

and the loudspeakers were placed at a distance Δx_r=0.125 m of each other, as shown in figure 7.8. The implementation of the spatial reconstruction filter will be described in the next section.
The loudspeakers have been mounted in 8 aluminum bars with a length of 3m and a cross section of 60 × 120 mm. In each bar 24 loudspeakers are mounted, thus creating a maximum array length of 24m.
Each loudspeaker is provided with an amplifier to reserve the possibility of driving each loudspeaker separately, yielding an array of 96 loudspeakers with a total length of 12m.
The digital hardware of the DSE system is formed by 12 floating point DSP processor boards (Motorola DSP96002). The processor boards are plugged into a PC (Intel 486DX2), which is coupled to a custom made 16-input and 96-output A/D-D/A convertor unit (see figure 7.9).

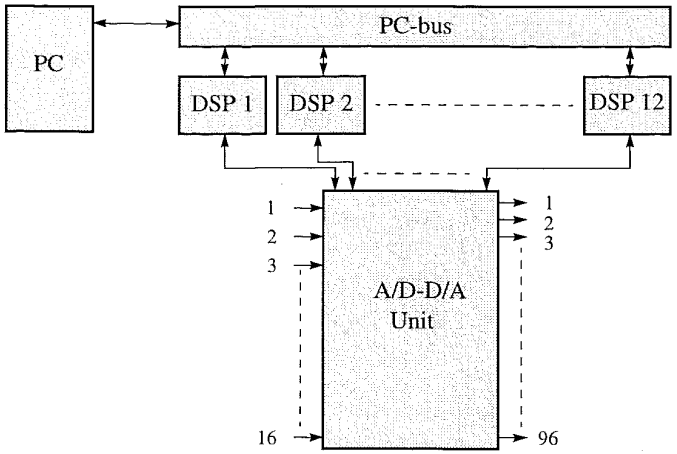


Figure 7.9: Schematic diagram of the digital hardware of the DSE proto-type system

Each DSP can access all A/D inputs, while the outputs can be divided in variable block sizes over the DSPs. Communication between the PC and the DSPs is possible over the PC-bus. Adaptation of parameters during program execution of the DSPs is accomplished by Dual-Port RAM, which is a memory area shared by the PC and the DSPs.
The 18-bit input channels and the 16-bit output channels are sampled at 32 kHz. The input filtering and loudspeaker response equalization is realized by a set of equalizers.
The loudspeaker arrays are supported by stands which can be adjusted in height, as sketched in figure 7.10. By means of 7 hinge-joints the 8 array elements could be coupled such that 4 to 6

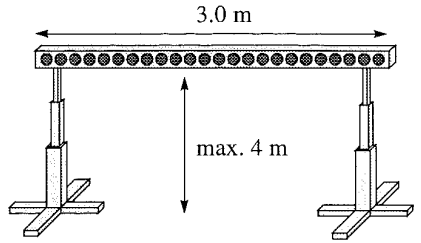


Figure 7.10: The array elements are supported by stands which can be adjusted in height.

stands could support the whole construction. By bending the array elements with respect to each other, various array geometries could be realized. Each bend must be supported by one stand.

7.4.2 Implementation of spatial bandwidth reduction in the DSE system

The spatial reconstruction filtering is implemented in the DSE-system by forming subarrays, each consisting of 3 loudspeakers at a distance Δx_r=0.125 m. By switching-off the even numbered loudspeakers, the reconstruction filter can be disabled, which is illustrated in figure 7.11.
The characteristics of the reconstruction filter are determined both by the weighting factors b and the directivity pattern of the individual loudspeakers. First the effect of the sub-array itself will be presented. After that the total directivity pattern, determined by sub-array plus loudspeaker directivity, will be determined.
Next, we follow the same approach as in the example of section 5.5.5. The weighting factors b are samples (open circles b_{-2} to b_2) of an ideal reconstruction filter (solid line) plotted in figure 7.12a. The weighting factors are determined by eq.(5.63). Note that b_{-2} and b_2 are equal to zero, for which reason they are omitted in the diagram of figure 7.11. The spatial amplitude spectrum $|\tilde{H}_r|$ of the reconstruction filter H_r is plotted in figure 7.12b.
Given the sampling distance Δx of 0.25 m and assuming a reduction of the maximum angle of incidence $\max\{\varphi_{inc}\}$ to 30° in the ideal situation, a spatial aliasing frequency f_{al}=1360 Hz is obtained. For wave numbers above $k=2\pi/\Delta x$ =25.1 m^{-1} (expressed as a temporal frequency:

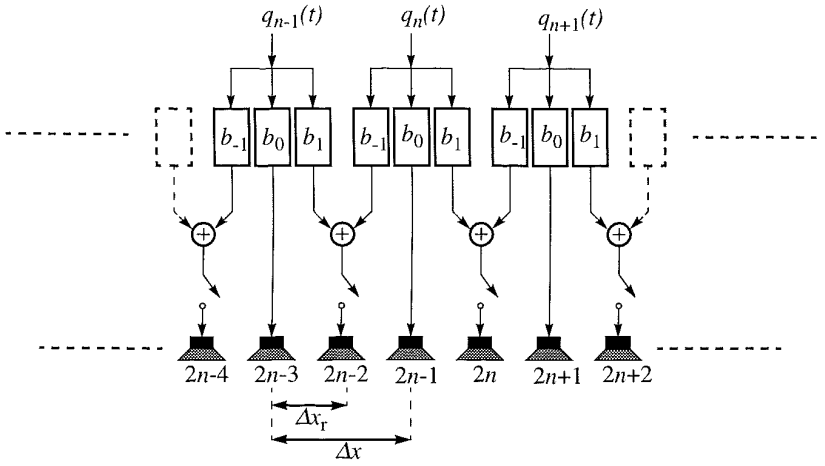


Figure 7.11: Schematic diagram of the implementation of the spatial reconstruction filter. By switching-off the even numbered loudspeakers, the reconstruction filter can be disabled.

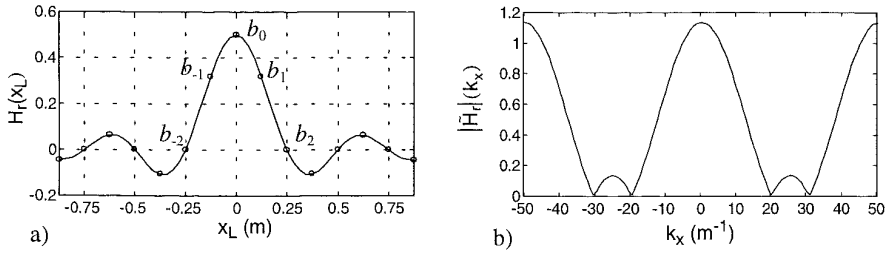


Figure 7.12: The spatial reconstruction filter H_r of the DSE system (b_0=0.5, b_1=b_{-1}=1/π=0.32 and b_2=b_{-2}=0)
a) Ideal reconstruction filter H_r (solid line), discretized and windowed version of H_r (open circles b_{-2} to b_2)
b) Spatial amplitude spectrum $|\tilde{H}_r|$ of the discretized and windowed version of H_r

f=1360 Hz) non-evanescent grating lobes become visible in the spatial spectrum $\tilde{H}_r$. For increasing frequency, more and more grating lobes become non-evanescent. However, due to the directivity pattern of the loudspeakers in the DSE system, which are given in appendix A, these non-evanescent grating lobes are attenuated, as shown in figure 7.13. Figure 7.13a shows the spatial amplitude spectrum $\tilde{H}_r$ for f=1360 Hz, in which the grating lobes are just visible, and figure 7.13b for the doubled frequency f=2720 Hz.

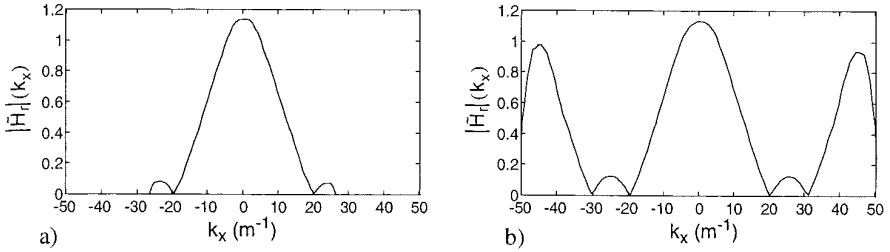


Figure 7.13: Combined effect of sub-arrays and directivity of the loudspeakers
a) Spatial amplitude spectrum $|\tilde{H}_r|$ for f=1360 Hz
b) Spatial amplitude spectrum $|\tilde{H}_r|$ for f=2720 Hz

In some array configurations the upper limit for f_{al} will not be reached and consequently the spatial aliasing frequency is far below the requirement of f_{al}>1.5 kHz given in section 7.4.1. It may be clear that the set-up of the DSE prototype system is a compromise between the desired length of the array, number of independent channels that can be processed and the spatial aliasing frequency.

7.4.3 Processing of high frequencies

In section 6.8 an explanation was found for the experimental fact that, when a listener moves his head slightly (half a sampling distance) while listening to a synthesized broadband noise source, a distinct change of color is perceived. Namely, in the x-t diagram of the synthesized wave field p the 'aliased part' displays a regular pattern with period Δx, which is also found in the physical and internal power spectra. Therefore, a method based on randomizing the arrival times of the loudspeaker signals, is proposed to reduce the unwanted coloration effects.
Initially, it was assumed that by positioning the loudspeakers at irregular intervals the coloration threshold could be increased. However, calculation of internal spectrum induced by the wave field synthesized with an irregularly sampled loudspeaker array, showed that the effect was insignificant. This negative result was contributed to the limited possibility of rearranging the loudspeaker positions due to their finite dimensions, yielding very small variations in the arrival time of the loudspeaker signals.
Therefore, the method has been adapted by applying externally random delays to the loudspeaker signals. In this way, the range of delays is not limited by any physical boundary. The processing scheme of figure 7.7 has been adapted in the following way (see figure 7.14). The source signal of the k^{th} notional source is convolved with a filter $h(t)$, which includes the $\sqrt{jk/2\pi}$-filtering, the loudspeaker response correction filter and a correction filter to obtain an averaged flat response for the frequencies above the spatial aliasing frequency. Next, in order to keep intact the low-frequency performance (i.e. the response below the spatial aliasing frequency), the driving signals are led through a cross-over filter with cut-off frequency f_{co} which

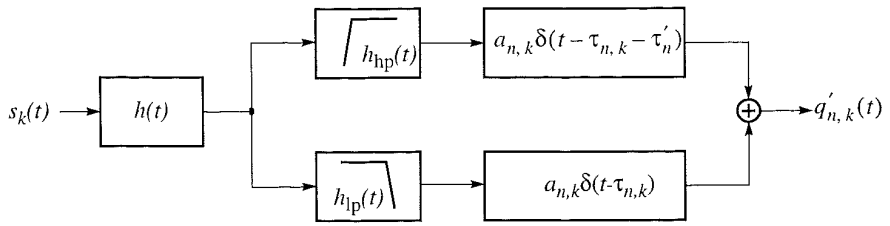


Figure 7.14: *Block diagram of the configuration for separate processing of low and high frequencies. The cut-off frequency of the complementary filters is equal to the spatial aliasing frequency. Both the low-pass and high-pass filtered signals are processed identically as in figure 7.7. However, the high-pass filtered signal is also delayed with a randomly chosen delay time τ'_n.*

equals the spatial aliasing frequency f_{al}. Both the low-pass and high-pass filtered signals are processed by applying the correct weighting factor $a_{n,k}$ and delay-time $\tau_{n,k}$. The high-pass filtered signal is additionally delayed with a randomly chosen delay-time τ'_n. To obtain the driving signal $q'_{n,k}$ for the n^{th} loudspeaker (sub-array), the low-pass and high-pass signals are combined again.

The low-pass filter h_{lp} and the high-pass filter h_{hp} are complementary; the frequency responses H_{lp} and H_{hp} are related by:

$$|H_{lp}(f)|^2 + |H_{hp}(f)|^2 = 1 \qquad (7.14)$$

The delay-time τ'_n is randomly chosen from a set of N delay-times, which have non-integer ratios. The range of delay-times will be determined experimentally, however, the maximum delay-time is limited by the occurrence of temporal side-effects for large delay-times.

For one special configuration, we will investigate by simulation the effect of randomizing the high-frequency delays. Consider the configuration of figure 7.15. The discrete linear array L, consisting of 24 overlapping loudspeaker sub-arrays, is sampled at intervals Δx=0.25 m. The sub-array weighting factors are given in figure 7.12. The notional source at $\mathbf{r}_\Psi=(0,0,0)^T$ emits a broadband impulse (0-10 kHz), as shown in figure 7.16a and b. To simulate accurately the directivity characteristics of the DSE-loudspeakers, the measured directivity responses (see Appendix A) were used in the calculations. The angular responses of the loudspeaker were deconvolved with the loudspeaker response on its main axis.

The synthesized wave field p and the primary wave field p_{pr} are recorded at the line $z=z_R$=8m for -3<x_R<3m. The x-t diagram of the primary wave field p_{pr} is shown in figure 7.16c.

The aperture of the array is reduced to max{φ_{inc}}=30°, and the driving signal is tapered at the ends of the array. The spatial aliasing frequency is given by f_{al}=1360 Hz.

Two situations are considered:

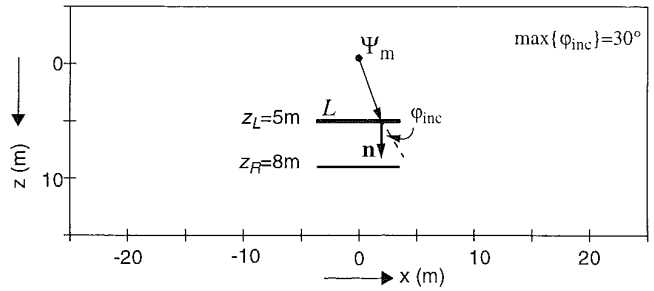


Figure 7.15: *Diagram of the geometry of a notional monopole source Ψ_m at $\mathbf{r}_\Psi$=(0,0) and a linear loudspeaker array consisting of 24 overlapping sub-array elements along the line L. The primary and the synthesized wave field are recorded at the line $z=z_R$=8m for -3<x_R<3m.*

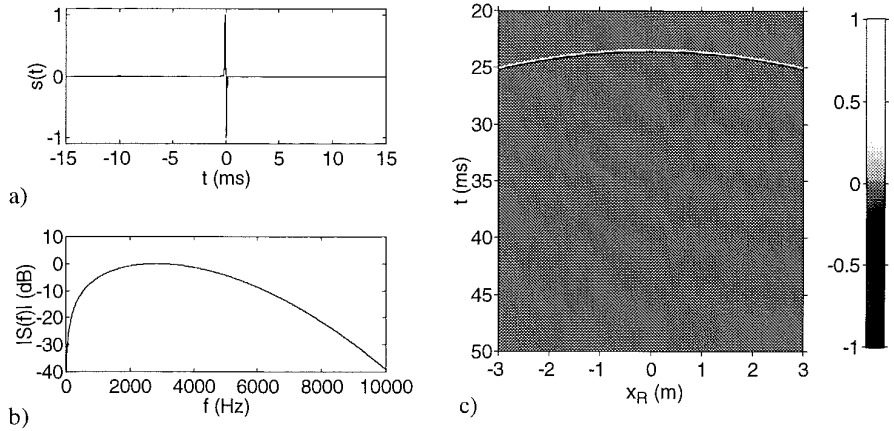


Figure 7.16: *Primary wave field p_{pr} emitted by the notional monopole source.*
a) source signal s(t)
b) amplitude spectrum |S(ω)| of the source signal, bandwidth 0-10 kHz.
c) x-t diagram

1.) No randomization: τ'_n=0 ms, for all n.

2.) Randomization of the high-frequency delays τ'_n

The amplitude responses of the complementary low-pass and high-pass filter are shown in figure 7.17.

The synthesized wave fields for situation 1 and 2 are shown in figure 7.18a and b respectively. Note that in situation 1 the arrival times of the individual sub-array contributions are regularly distributed over a time interval of about 12 ms, while for situation 2 the arrival times are randomly distributed over a larger time interval (about 25 ms). Moreover, it is clearly visible that

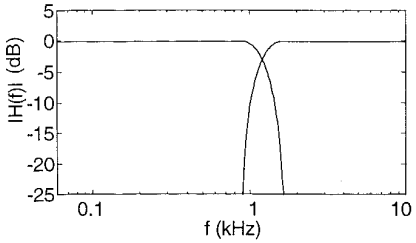


Figure 7.17: *Amplitude response of the cross-over filter*

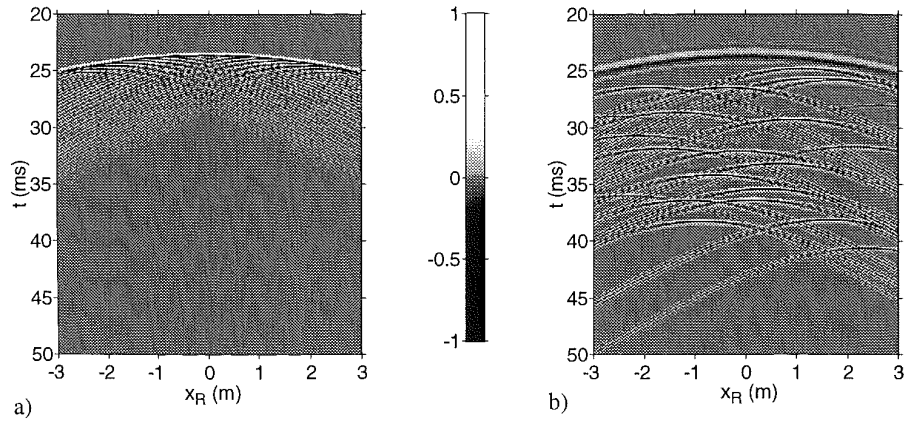


Figure 7.18: *Wave field synthesis with a finite discrete loudspeaker array (Δx=0.25 m), consisting of 24 sub-array elements. Registration along the line $z=z_R$ for $-3<x_R<3$m. Temporal bandwidth: 0-10 kHz.*
a) x-t diagram of the synthesized wave field p without randomizing the high-frequency delays
b) x-t diagram of the synthesized wave field p with randomizing the high-frequency delays.

the main wave front in situation 1 is composed of all frequencies (0-10kHz), while for situation 2 the main front merely consists of low frequencies (0-1.36 kHz). By randomizing the high-frequency delays, the high-frequency contributions are separated in time from the low-frequency main wave front.

The spectral ratio L_R has been calculated on the basis of the internal spectra of the primary and synthesized wave field. Figure 7.19 a and b show the spectral ratio L_R for $-1<x_R<1$m and $0<f<10$ kHz for situation 1 and 2 respectively. It has been demonstrated that, as a result of randomizing the high-frequency delays, the internal spectrum of the synthesized wave field is flattened. Even more important is the decrease of the spectral differences in the internal spectrum for neighboring receiver positions.

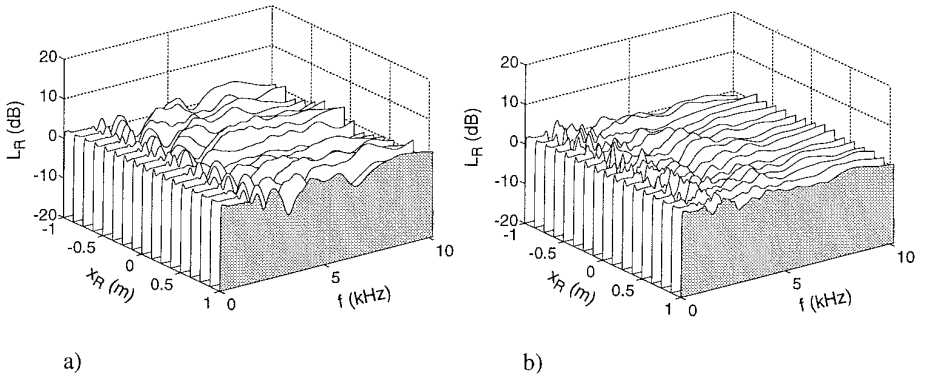


Figure 7.19: *Spectral ratio based on the internal spectra, $-1<x_R<1$m. Temporal bandwidth: 0-10 kHz.*
a) x-f diagram of the pressure ratio level L_R without randomizing the high-frequency delays
b) x-f diagram of the pressure ratio level L_R with randomizing the high-frequency delays.

Based on these simulations, it is expected that randomization of the high-frequency delays yields a reduction of the audibility of coloration (differences) in synthesized wave fields. The performance of the DSE system will be evaluated under several room acoustic conditions; anechoic room, auditorium and concert hall acoustics, of which the results will be given in chapter 8.

8

Evaluation of the DSE system

To complete the design of the direct sound enhancement system based on wave field synthesis, the proto-type DSE system has been evaluated in different rooms and halls. Physical as well as perceptual experiments have been conducted to test the performance of the system. A full-scale investigation had never been conducted before into the physical and perceptual aspects of synthesized wave fields under various acoustic conditions.
The physical experiments have been done in such a way that real and synthesized wave fields of a point source were spatially measured and compared. In the psycho-acoustical experiments the main subjective aspects of real and synthesized sound fields, as summarized in chapter 6, have been judged by several observers.
The measurements have been conducted in four different halls. The first three halls have been chosen such that the performance of the proto-type DSE system is tested under various acoustic conditions, ranging from a dead to a reverberant acoustic situation. The fourth hall was only used to test the physical performance of a special array geometry; a concave loudspeaker array. The halls included in this investigation are shown in table 8.1, in which for each hall, the volume V, the number of seats N_S, the seated area S_S and the reverberation time T_{60} are given.

8.1 Physical measurements

To obtain information about the spatial properties of the synthesized wave field, the impulse response of the notional source should be measured along a microphone array. For the practical realization of this measurement technique we measured sequentially the impulse response with one omni-directional microphone of which the position was changed equidistantly along the registration line after each measurement.
In the next four sub-sections the room geometry, loudspeaker array configuration and the registration positions will be given for each hall. The results of the measurements on real and syn-

Table 8.1: *Halls included in the evaluation experiments with volume V, number of seats N_S, seated area S_S and reverberation time T_{60}.*

Name	$V\,(m^3)$	N_S	$S_S\,(m^2)$	$T_{60}\,(s)$
Anechoic room, laboratory of acoustics of Delft University	512 ($8\times8\times8$m)	-	-	0
Auditorium, Delft University of Technology	±8,000	977	610	1.1
Concert hall "De Doelen", Rotterdam, The Netherlands	27,070	2,230	1,867	2.3
Congress room, Jaarbeurs Utrecht, The Netherlands	1,350	±200	300	±1.0

thesized sound fields will be presented. Since the array geometry of the DSE system could be adapted rather easily, the shape of the array was chosen such that, within the technical boundaries of the bearing construction of the array and within the possibilities of the architectural shape of the hall, an optimal result was obtained for all situations.

8.1.1 Measurements on synthesized sound fields in the anechoic room

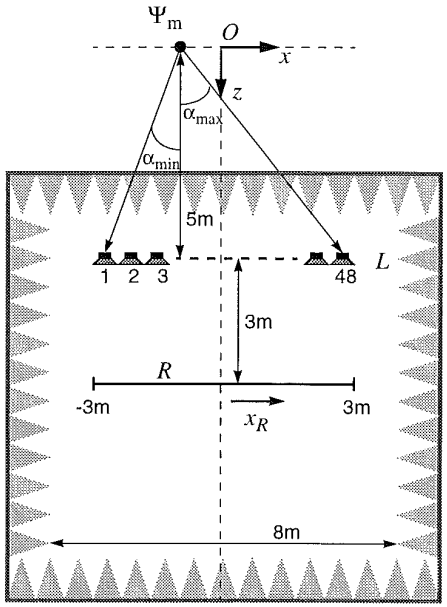
The first experiments were carried out in the anechoic room which has inner dimensions of $8\times8\times8$m. The configuration is shown in figure 8.1. The origin of the co-ordinate system is indicated with *O*. The downward direction is taken as the positive *z*-direction, and the positive *x*-direction to the right. The notional monopole source Ψ_m is positioned at $\mathbf{r}_\Psi$=(-1,0). The loudspeaker array at z_L=5m consists of 48 elements at intervals of 0.125 m. However, the sampling distance Δx equals 0.25 m. Due to the limited aperture of the array as seen from the notional source, given by α_{min}=21.8° and α_{max}=38.7°, the spatial aliasing frequency f_{al}, which is determined by eq.(5.17), equals 1.4 kHz.

In the anechoic room three situations were considered:

1. No reconstruction filtering; only the odd-numbered loudspeakers (1,3,5,...47) are used.
2. With reconstruction filtering; 24 overlapping sub-arrays, each consisting of 3 loudspeakers at a distance of 0.125 m, as described in section 7.4.2.
3. Reconstruction filtering combined with random high-frequency delays, as explained in section 7.4.3. The cross-over frequency was set at 1.2 kHz.

It is expected that the maximum frequency that is synthesized correctly amounts about 1.4 kHz and 1.0 kHz for the first and second situation respectively.

The broadband impulse response (0-16kHz) of the notional source is registered at 121 microphone positions with 0.05m intervals on the line *R* at z_R=8m parallel to the array. To visualize the measured wave field, the impulse responses have been processed with a low-pass filter (0-5kHz), yielding a compact source signal. Since the wave field of a real source in an anechoic room is known a priori, a measurement of the real wave field is omitted.



α_{min}=-21.8°
α_{max}= 38.7°

Figure 8.1: *Measuring configuration in the anechoic room. The notional source Ψ_m is positioned at $\mathbf{r}_\Psi$=(-1,0), 5m behind the array. The loudspeaker array consists of 48 loudspeakers at intervals of 0.125 m (Δx=0.25 m). The wave field was registered at 121 positions with 0.05 m intervals on the line z_R=8m, 3m in front of the array.*

The measurement results of the synthesized wave fields are shown in figure 8.2a-c for the three situations, presented in *x-t* diagrams. For comparison the simulated synthesized wave field for situation 1 is shown in figure 8.2d. By comparing figure 8.2a and d it can be observed that the measured and simulated synthesized sound field match very well. Also notice that due to the reconstruction filtering the aliasing tails have been reduced. Further, in figure 8.2c the high-frequency contributions have been randomly rearranged to reduce the audibility of coloration due to spatial aliasing. The results of the perceptual experiments on coloration will be presented in section 8.5.

The temporal spectra (80 Hz-16 kHz) for seven receiver positions (intervals of 1m) are shown in figure 8.3a-c for situation 1 to 3 respectively. These spectra indeed show that, as a result of the spatial reconstruction filter, in situation 2 the synthesized wave field is correct up to a frequency of about 1.4 kHz, while for situation 1 this frequency is approximately 1.0 kHz.

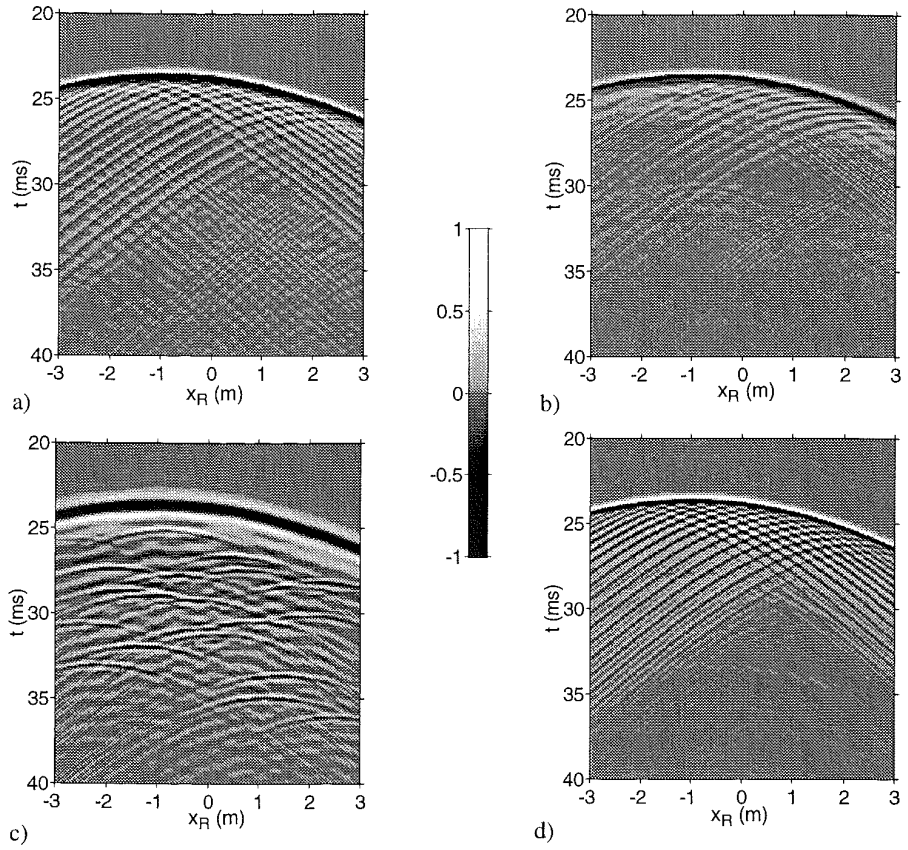


Figure 8.2: *Measured and simulated synthesized wave fields in the anechoic room. Registration along the line R for -3<x_R<3m. Temporal bandwidth: 0-5 kHz*

a) x-t diagram of the measured synthesized wave field; situation 1

b) x-t diagram of the measured synthesized wave field with reconstruction filtering; situation 2

c) x-t diagram of the measured synthesized wave field with reconstruction filtering and random high-frequency delays; situation 3

d) x-t diagram of the simulated synthesized wave field; comparable to situation 1

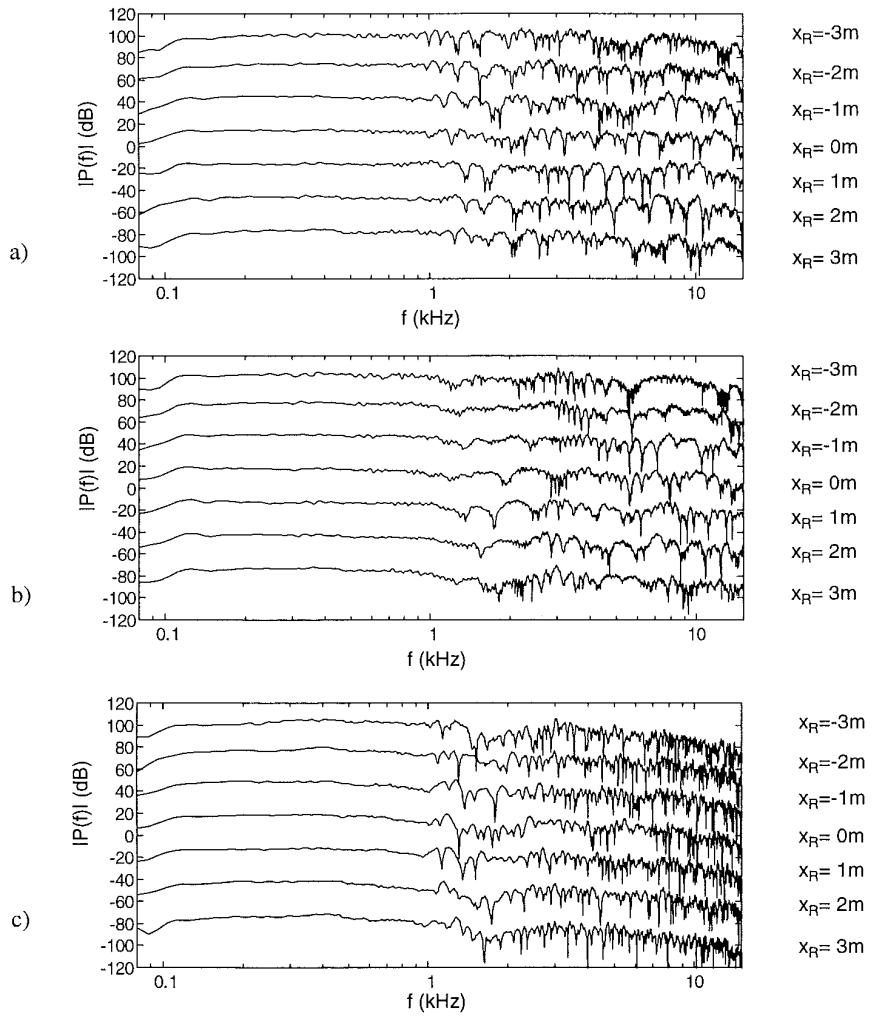


Figure 8.3: *Measured spectra at seven receiver positions on the line R (intervals of 1m). Spectra are shifted 30 dB. Bandwidth: 80 Hz-16 kHz.*

a) Spectrum of the measured synthesized wave field; situation 1

b) Spectrum of the measured synthesized wave field with reconstruction filtering; situation 2

c) Spectrum of the measured synthesized wave field with reconstruction filtering and random high-frequency delays; situation 3

8.1.2 Measurements on synthesized sound fields in the auditorium

In this section the results of the measurement in the auditorium are presented. The ground plan, which is the horizontal projection of the hall, is shown in figure 8.4. The seated parts have a rather steep inclination and occupy an arch around the optional stage area which can also be used as seated area. Given this geometry, it is advantageous to apply a convex loudspeaker array. The advantage of this type of array is that the loudspeakers are placed relatively close to the audience, and a large aperture as seen from a source at stage (indicated with the dashed lines) is achieved. For listening positions outside the aperture no correct sound field can be synthesized, however, at some of these positions listeners would also have a line-of-sight problem.

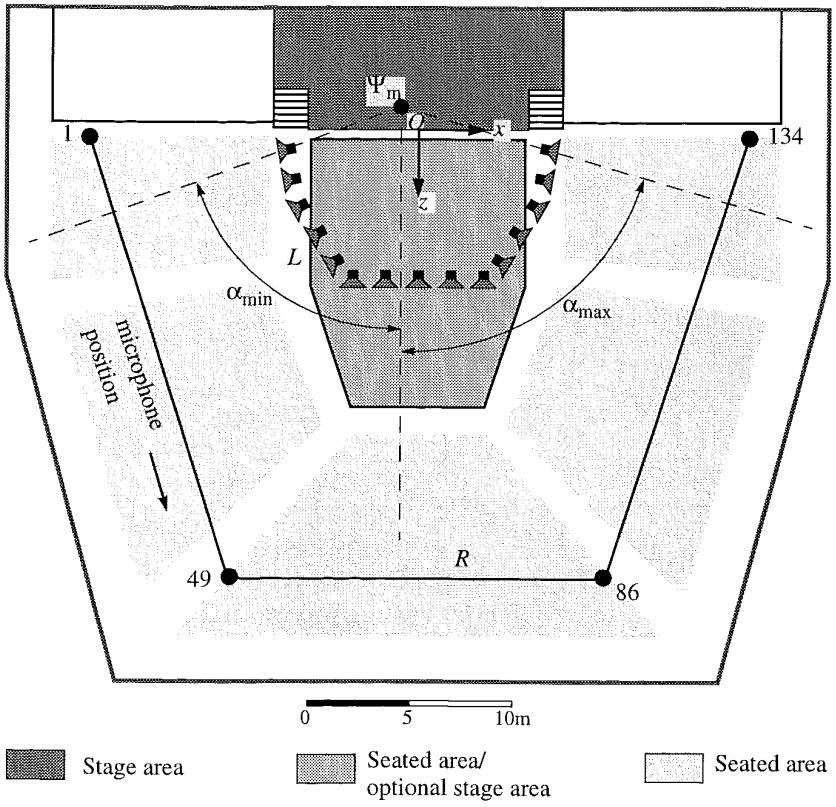


Figure 8.4: *Ground plan of the auditorium of Delft University of Technology. The measurement configuration can be described as follows: The notional monopole source Ψ_m is positioned at $\mathbf{r}_\Psi=(x_\Psi,y_\Psi,z_\Psi)=(-1,4,-1)$. The wave field is synthesized using an array of 96 overlapping loudspeaker sub-arrays at the bent line L ($\Delta L=0.25$ m.), consisting of 192 loudspeakers. The wave field is registered at 134 positions on the bent line R.*

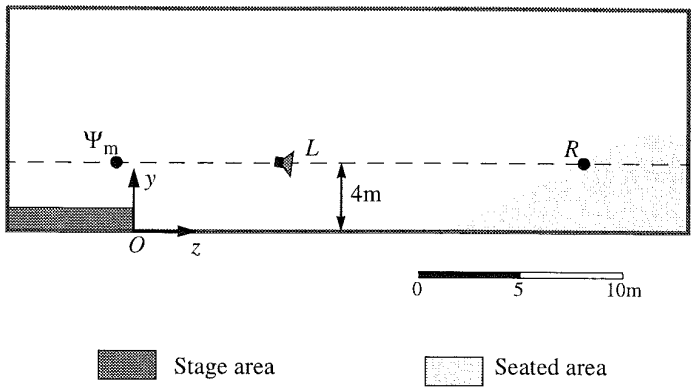


Figure 8.5: *Cross-section of the measurement configuration in figure 8.4. The notional monopole source Ψ_m is positioned at $\mathbf{r}_\Psi=(x_\Psi,y_\Psi,z_\Psi)=(-1,4,-1)$.*

In figure 8.4 also a description of the measurement configuration in the horizontal plane is given. The geometry in the vertical plane x=0 is shown in figure 8.5. The origin O of the coordinate system is defined at the floor level at the center of the edge of the stage. The downward direction in figure 8.4 is taken as the positive z-direction, and the positive x-direction is taken to the right. The positive y-direction is chosen upwards as shown in figure 8.5. The bent loudspeaker array L lies in the horizontal plane y=4m and consists of 96 overlapping sub-arrays, each consisting of three loudspeakers at intervals of 0.125 m. The intermediate distance ΔL between the subarrays equals 0.25 m.

Inherent to 2½D wave field synthesis, the notional monopole source Ψ_m lies in the same horizontal plane at which the loudspeaker array is positioned; in this configuration the notional source is found $\mathbf{r}_\Psi$=(-1,4,-1). For this source-array geometry the maximum angle of incidence $\max\{\varphi_{inc}\}$ equals 34.3°, yielding a spatial aliasing frequency of about 1.2 kHz. The aperture of the array, as seen from the notional source, is given by α_{min} and α_{max} equaling -71.1° and 75.3° respectively.

The broadband impulse response (0-16kHz) of the notional source is registered on the line R at 134 microphone positions with 0.5 m intervals indicated in figure 8.4 and 8.5. The bandlimited wave field (0-5 kHz) is shown in figure 8.6a. In addition, the wave field of a real omni-directional source at the same x-z location as the notional source, however at height y_Ψ=2 m, has been recorded at the line R. The measured real sound field is shown in figure 8.6b. Notice that the first arriving wave front (i.e. the direct sound) is almost identical in both situations. Also notice that, as expected, the synthesized direct sound is too weak for receivers outside the aperture; that is for microphone positions 1-10 and 126-134.

Further, it can be observed that the reflection patterns of the synthesized and real wave field differ. Especially, the reflections from the stage floor and the back side wall of the stage, which

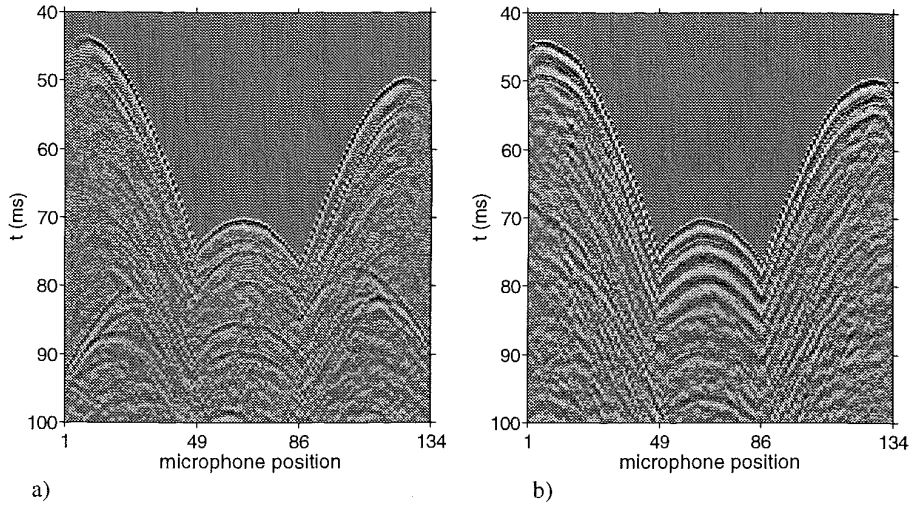


Figure 8.6: *Comparison of the measured synthesized and real wave fields in the auditorium. Registration along the bent line R. Temporal bandwidth: 0-5 kHz*
a) Measured synthesized wave field
b) Measured real wave field

are present in the reflection pattern of the real source, are lacking in the wave field of the notional source. The reason for this is twofold: 1) It is physically impossible to reconstruct the sound field in the primary source area (behind the array) correctly, and 2) in each vertical plane, perpendicular to the array, the wave fronts radiate circularly with the array as the acoustic center.

8.1.3 Measurements on synthesized sound fields in the concert hall

In this section the results of the measurement in the concert hall "De Doelen" are presented. The ground plan and a description of the measurement configuration in the horizontal plane are shown in figure 8.7.
The seated parts have a small inclination. Given this hall geometry, a large part of the audience can be covered by applying a linear loudspeaker array.

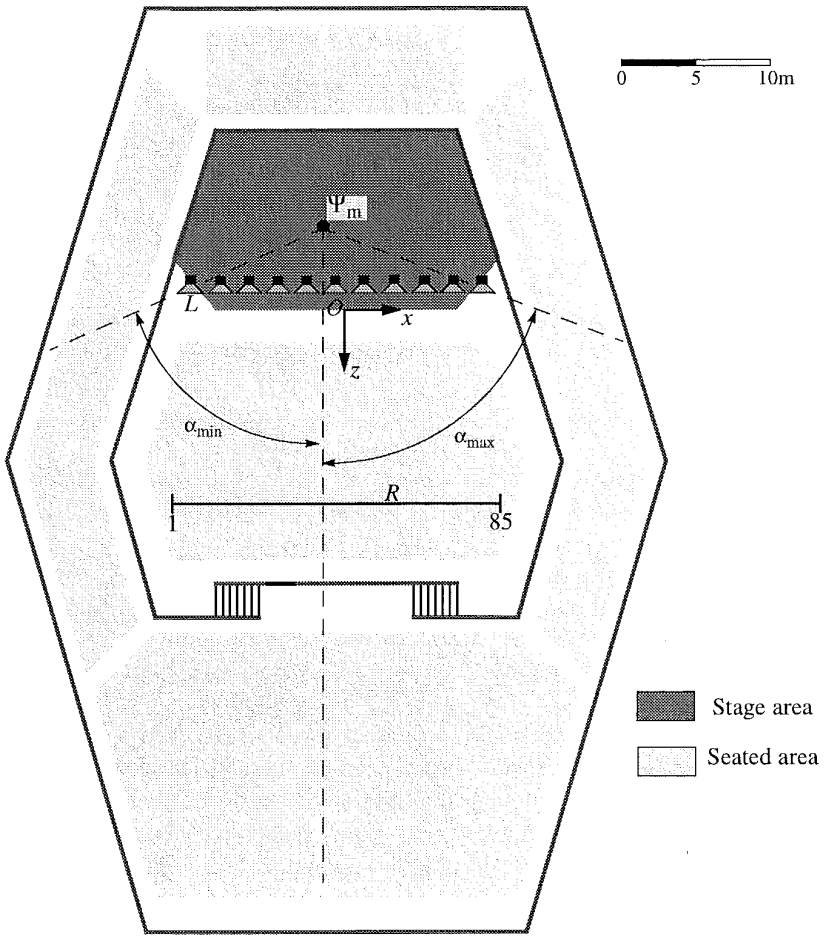


Figure 8.7: *Ground plan of the concert hall "De Doelen", Rotterdam, the Netherlands. The measurement configuration can be described as follows: The notional monopole source Ψ_m is positioned at $r_\Psi=(x_\Psi,y_\Psi,z_\Psi)=(-1,7,-5)$. The wave field is synthesized using a linear array of 72 loudspeakers at the line L ($\Delta x=0.25$ m.). The wave field is registered at 85 positions on the line R.*

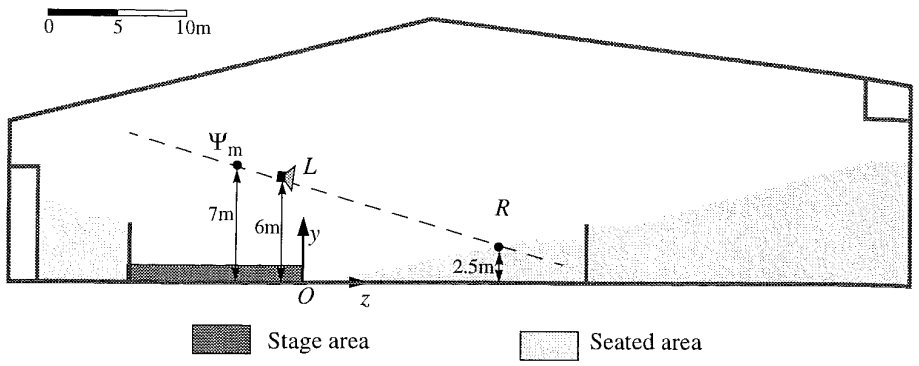


Figure 8.8: *Cross-section of the measurement configuration in figure 8.7. The notional monopole source Ψ_m is positioned at $\mathbf{r}_\Psi=(x_\Psi, y_\Psi, z_\Psi)=(-1,7,-5)$.The linear loudspeaker array, consisting of 72 loudspeakers at intervals of 0.25 m, is positioned at the horizontal line L given by z_L=-1m and height y_L=6m.*

The geometry in the vertical plane x=0 is shown in figure 8.8. The origin O of the co-ordinate system is defined at the floor level at the center of the edge of the stage. The linear loudspeaker array, consisting of 72 loudspeakers at intervals of 0.25 m, is positioned at the horizontal line L given by z_L=-1m and height y_L=6m.

The notional monopole source Ψ_m is positioned at a x-z location given by (x_Ψ, z_Ψ)=(-1,-5). For this source-array geometry the maximum angle of incidence $\max\{\varphi_{inc}\}$ is determined by the aperture, which is given by α_{min} and α_{max} equaling -63.4° and 68.2° respectively, yielding a spatial aliasing frequency of about 750 Hz.

The broadband impulse response (0-16kHz) of the notional source is registered at 85 microphone positions with 0.275 m intervals on the line R parallel to the array at z_R=13.4m and height y_R=2.5 m, as indicated in figure 8.7 and 8.8. Note that in this configuration the array and the receiver line R lie in a tilted plane. It can be verified that consequently the y-position of the notional source equals 7.0 m.

The bandlimited wave field (0-5 kHz) is shown in figure 8.9a. In addition, the wave field of a real omni-directional source at the same x-z location as the notional source, however at height y_Ψ=2 m, has been recorded at the line R. The measured real sound field is shown in figure 8.9b. The first arriving wave front (i.e. the direct sound) is almost identical in both situations. As explained in the previous section, the reflection patterns of the synthesized and real wave field differ.

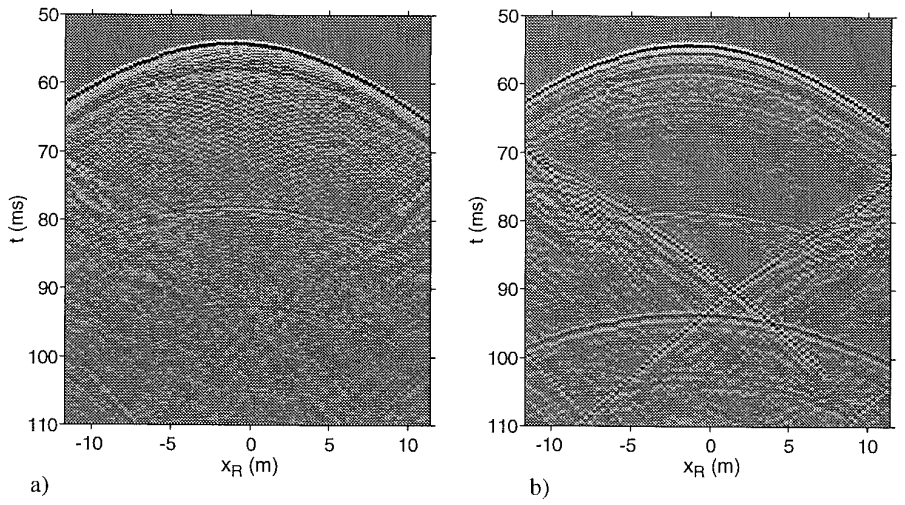


Figure 8.9: *Comparison of measured synthesized and real wave fields in De Doelen. Registration along the line R. Temporal bandwidth: 0-5 kHz*
a) x-t diagram of the measured synthesized wave field
b) x-t diagram of the measured real wave field

8.1.4 Measurements on synthesized sound fields in a congress room

Finally the results of the measurement in a congress room in De Jaarbeurs are presented. In this room a special array geometry was tested; a concave loudspeaker array. The configuration is shown in figure 8.10. The origin O of the co-ordinate system is defined at the floor level at the center of the front side of the room. The concave bent loudspeaker array, consisting of 96 loudspeakers at intervals of 0.25 m, is positioned at the horizontal bent line L with height y_L=2m.
Using this array, various source configurations were tested. Here, we will only show the results of the measurement on the wave field for one special configuration; viz, a notional monopole source Ψ_m in front of the array at (x_Ψ,y_Ψ,z_Ψ)=(3,2,8). To generate the wave field of this focussed notional source, only the loudspeakers at x_L>0 m are used. For this geometry the spatial aliasing frequency is about 960 Hz.
First, the broadband impulse response (0-16kHz) of the notional source is registered at 49 microphone positions with 0.25 m intervals on the line R_1 parallel to the array at z_R=11 m and height y_R=2 m, as indicated in figure 8.10.
The synthesized bandlimited wave field (0-5 kHz) is shown in figure 8.11a. For comparison, the wave field of a real omni-directional source at the same position as the notional source has

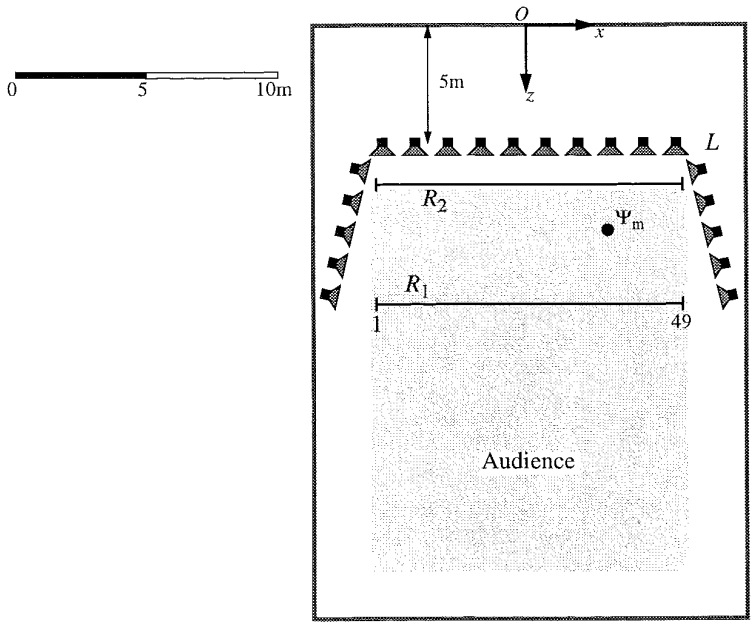


Figure 8.10: Measurement configuration in a congress room in de Jaarbeurs Utrecht, the Netherlands. The notional monopole source Ψ_m is positioned at $\mathbf{r}_\Psi=(x_\Psi,y_\Psi,z_\Psi)=(3,2,8)$. The wave field is synthesized using a concave array of 96 loudspeakers at the line L (Δl=0.25 m.). The wave field is registered at 49 positions on the line R_1 and also on the line R_2.

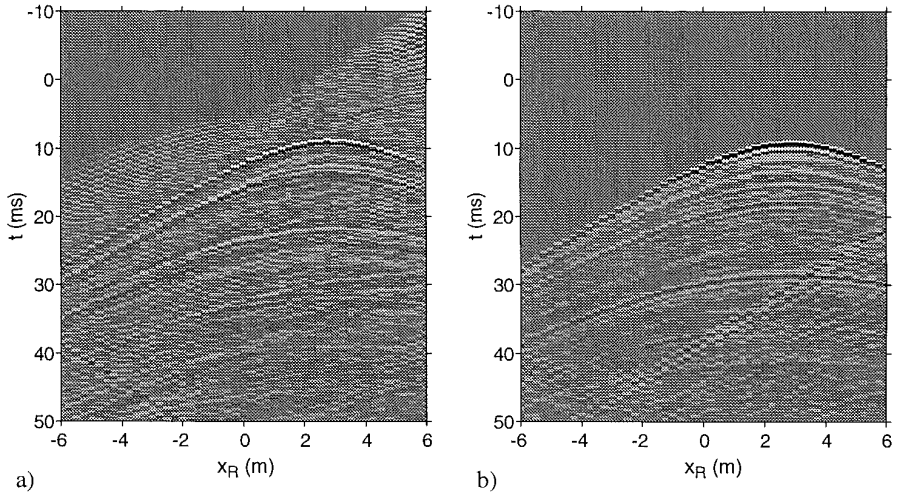


Figure 8.11: Comparison of measured synthesized and real wave field in a congress room. Registration along the line R_1 at z_R=11m. Temporal bandwidth: 0-5 kHz
a) x-t diagram of the measured synthesized wave field.
b) x-t diagram of the measured real wave field.

been recorded at the line R_1. The measured real sound field is shown in figure 8.11b. Notice in figure 8.11a that the desired wave front is clearly visible in the synthesized wave field. However, due to spatial aliasing the pressure before the arrival of the main wave front is unequal to zero. In contrast to the situation with a notional source behind the array, the aliasing tails are now leading the desired main wave front. Also notice that the reflection patterns of the synthesized and real wave field differ, which is not surprising since the synthesized wave field is only correct in the diverging part of wave field in the horizontal plane.
It is interesting to take a closer look at the synthesized wave field between the array and the notional source. Therefore, the wave field is also measured along the line R_2 at z_R=6.5 m. The *x-t* diagram is shown in figure 8.12. Since the registration is done in the converging part of the synthesized wave field, the main wave front in the spatial impulse response has a mirror-reversed hyperbola shape in comparison to the shape of the main wave front in figure 8.11a.

8.2 Localization experiments

In this section the results are presented of the source-identification experiments. These experiments were designed to determine whether horizontal localization performance of listeners differs for real and synthesized wave fields under various acoustic conditions. The influence of the

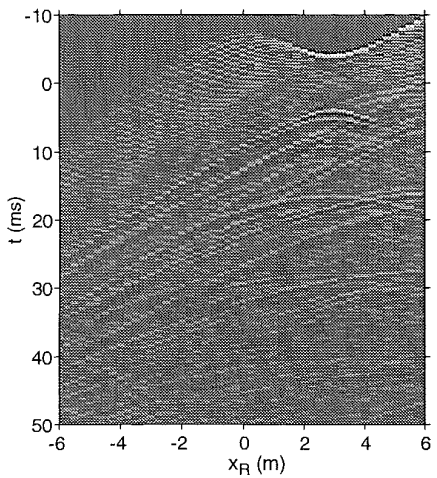


Figure 8.12: *x-t diagram of the measured synthesized wave field in a congress room. Temporal bandwidth: 0-5 kHz. Registration along a line between the array and the focussed source; z_R=6.5m*

nature of the source signal has also been taken into account by conducting the experiments for several source signals. The results have been analyzed with respect to localization accuracy, localization bias and random variability in localization.

The horizontal localization experiments have been conducted in the first three rooms, which were mentioned in the introduction of this chapter. In each room basically five different stimuli were used:

1. lowpass-filtered noise (80Hz-1.2kHz)
2. highpass-filtered noise (1.2kHz-16kHz)
3. broadband noise (80Hz-16kHz)
4. speech (male voice)
5. music (a fragment of Godár: Concerto Grosso für 12 streicher und cembalo)

The low-frequency and high-frequency noise stimuli have been chosen such that the spectral regions below and above the spatial aliasing frequency are tested separately.

For the synthesized sound fields three additional stimuli were used:

6. broadband noise (80Hz-16kHz) with random high-frequency delays
7. speech with random high-frequency delays
8. music with random high-frequency delays

The range of additional random high-frequency delays is given by: $0<\tau_n'<15$ ms. These stimuli are used to test whether the source localization is affected by the randomization.

The loudspeaker array configurations, which were described in the previous sections, were used for the localization experiments in each room.

In the second hall, the auditorium, additional experiments have been performed. In that situation localization (direction) as well as auditory distance perception and vertical localization have been investigated.

8.2.1 Anechoic room

I. Method

A monopole wave field was synthesized in the anechoic room. The notional source position was chosen randomly from a pre-defined set of 13 possible locations. The set of notional source positions was distributed uniformly at a line 5 m behind the array in the interval $-5m<x_\Psi<5m$. Forty eight numbered labels were attached to the loudspeaker array with a spacing of 0.125 m (figure 8.13). The subject's task was to indicate the perceived location of the source by giving

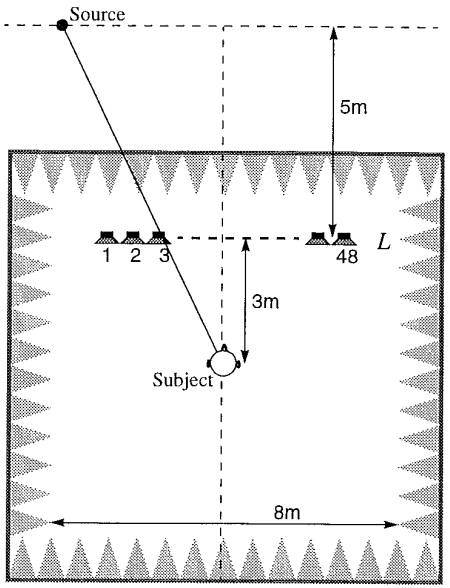


Figure 8.13: *During the source identification experiment in the anechoic room the subjects were asked to indicate the notional source position by giving the number of the loudspeaker at the intersection of the line from source to listener.*

the label number at the intersection of the line from the simulated source to the subject.

Since quite some experience and results have been obtained in the anechoic room in the past (see section 6.4), only two subjects participated in this experiment. The listener was seated 3 m in front of the array and was free to turn his head. Each experimental run consisted of 65 trials, 5 presentations of 13 different sources. No feedback was given during the run.

During the real source identification experiments at each stimulus presentation one loudspeaker of the array was used as a real source. The listener was seated 4.6m in front of the array.

II. Results

The statistical quantities *D, E* and *s*, as defined in section 6.1.2, were calculated. The run rms localization error $\langle \bar{D} \rangle$, the run error $\langle \bar{E} \rangle$ and the run standard deviation $\langle \bar{s} \rangle$ averaged across subjects are shown in figure 8.14a-c for real and synthesized sources. The individual results of the two subjects are given in appendix B.

These results show that the localization accuracy of low-frequency noise stimuli is almost identical for synthesized ($\langle \bar{D} \rangle$ =2.5°) and real sound fields ($\langle \bar{D} \rangle$ =2.2°). As expected, localization performance is seriously degraded for high-frequency noise stimuli (above spatial aliasing). Both the rms error $\langle \bar{D} \rangle$ and the standard deviation $\langle \bar{s} \rangle$ are significantly larger for the high-frequency synthesized sound field than for the high-frequency real sound field.

With respect to the broadband stimuli 3-5 it can be noted that the rms localization error is very small for synthesized sound fields. The rms error $\langle \bar{D} \rangle$ is only about 0.8-1.2° larger for synthesized sound fields in comparison with real sound fields.

Due to the randomization of the high-frequencies in the synthesized wave field (stimuli 6-9) the averaged rms error $\langle \bar{D} \rangle$ across subjects has increased in comparison with stimuli 3-5. However, looking at the individual results of the subjects in appendix B (figure B.1a), it must be noted that one of the two subjects was able to maintain a small localization error while for the other subject the localization error has increased considerably. During the source identification experiments the subjects reported that randomization of the high-frequency contributions 'diffuses' the high-frequency image of the source, while the low-frequency image remains rather compact. Likely, one of the subjects was 'distracted' by the broad high-frequency image, which does not contain useful localization cues.

The averaged run error $\langle \bar{E} \rangle$ is rather small; $|\langle \bar{E} \rangle| < 1.5°$ for the stimuli 1-5, and $|\langle \bar{E} \rangle| < 2.2°$ for the stimuli 6-9. The values of the averaged run standard deviations $\langle \bar{s} \rangle$, which are plotted in figure 8.14c, show that, except for the high-frequency noise stimulus, the subjects were very consistent in determining the source position.

8.2.2 Auditorium: horizontal localization

I. Method

In the Auditorium the notional source position was chosen randomly from a pre-defined set of 33 possible locations on stage. The set of notional source positions was distributed uniformly along three rows at different depth levels, 2m from each other. In the lateral direction the sources were positioned in the interval of $-4m < x_\Psi < 4m$. At each row 41 numbered labels were placed ranging from $-5m < x_\Psi < 5m$ with a spacing of 0.25 m, as shown in figure 8.15. The subject's task was to indicate the perceived location of the source by giving the row number (depth A, B or C) and the label number within that row (lateral position 1-41). Seven subjects participated in this experiment at the same time. The listeners were seated at the positions indicated

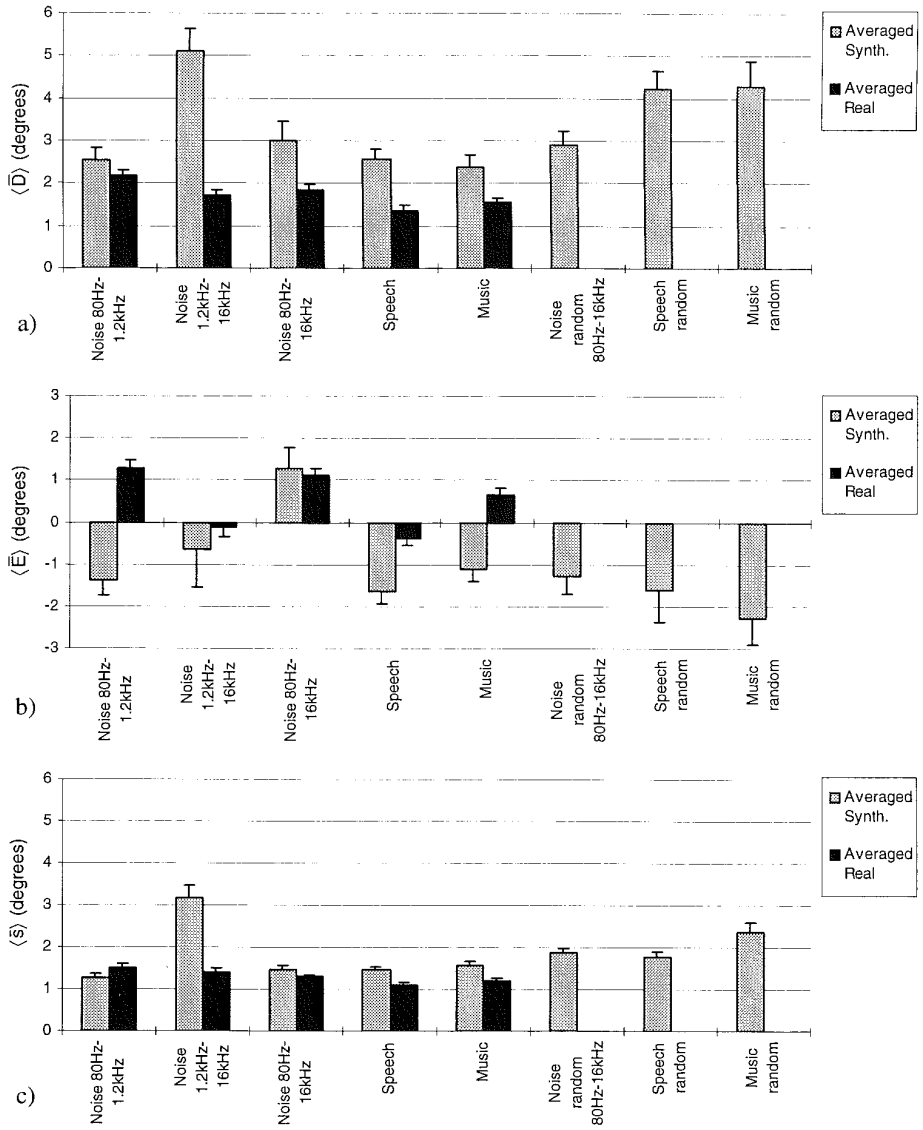


Figure 8.14: *Results of the localization experiments for synthesized and real sound fields in the anechoic room under several stimulus conditions.*

a) Run rms error $\langle \bar{D} \rangle$ averaged across subjects

b) Run error $\langle \bar{E} \rangle$ averaged across subjects

c) Run standard deviation $\langle \bar{s} \rangle$ averaged across subjects

in figure 8.15. Each experimental run consisted of 165 trials, 5 presentations of 33 different sources. No feedback was given during the run.
During the real source identification experiments the loudspeakers were placed on stage in three rows. At each stimulus presentation one loudspeaker was used as the real source.

II. Results

The statistical quantities D, E and s were calculated. The run rms localization error $\langle \bar{D} \rangle$, the run error $\langle \bar{E} \rangle$ and the run standard deviation $\langle \bar{s} \rangle$ averaged across subjects are shown in figure 8.16a-c for real and synthesized sources. The rms localization error $\langle \bar{D} \rangle$ shows the same tendency as in the anechoic room. For both synthesized and real sound fields, the localization accuracy is somewhat worse than in the anechoic room.
Note that for the synthesized sound field the averaged run error $\langle \bar{E} \rangle$ is negative for all stimuli, which means that the localization judgements are biased in the left direction; averaged across

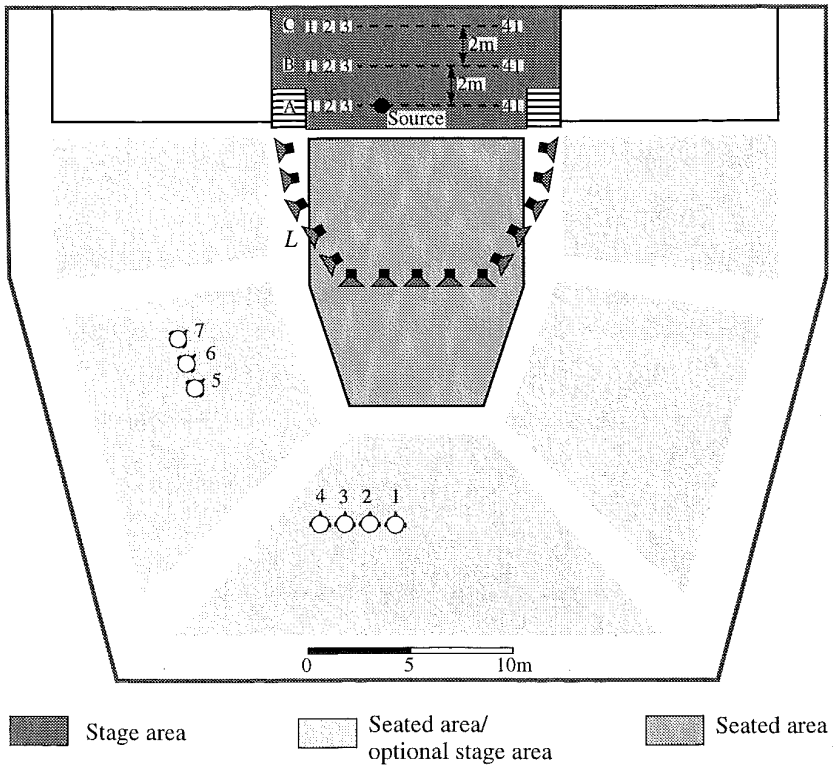


Figure 8.15: During the source identification experiment in the auditorium the subjects were asked to indicate the notional source position by giving the row number (A,B or C) and the label number within that row.

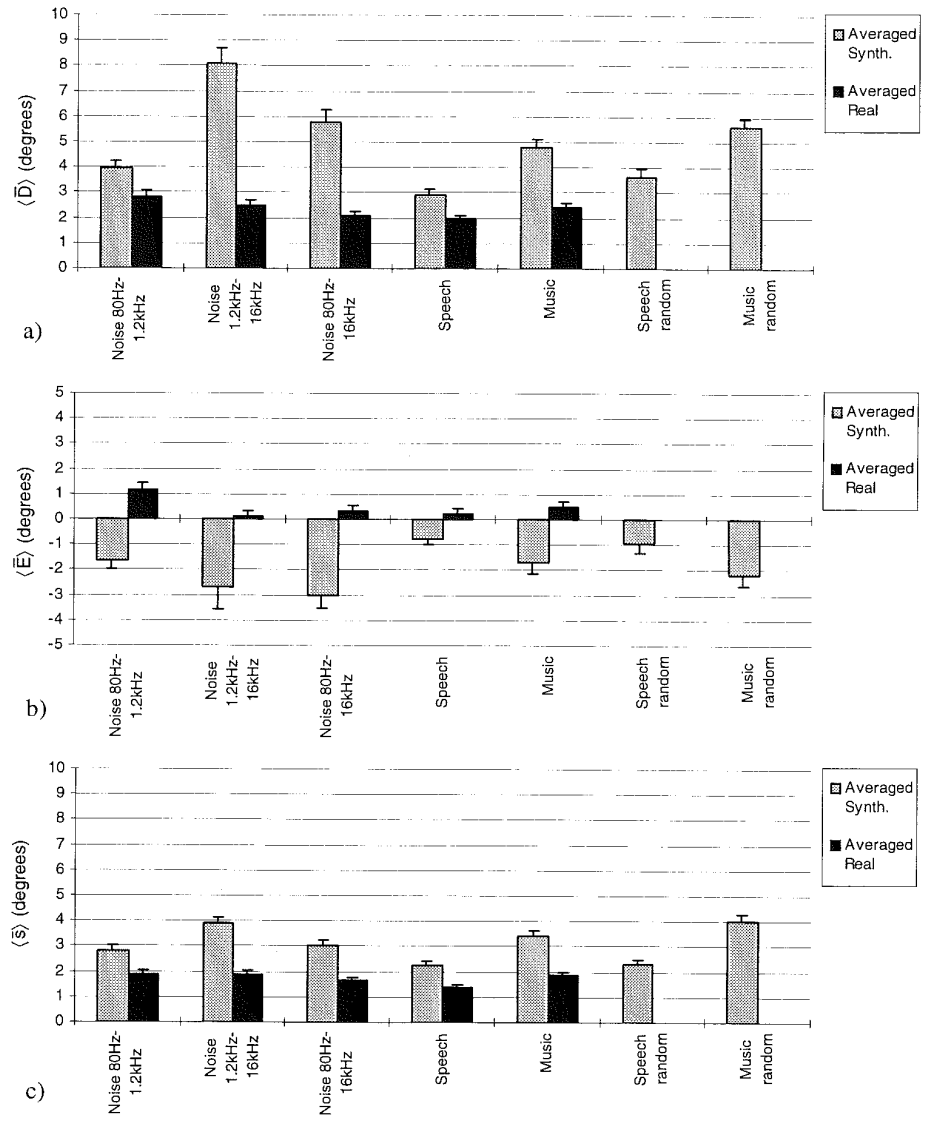


Figure 8.16: Results of the localization experiments for synthesized and real sound fields in the auditorium under several stimulus conditions.
a) Run rms error $\langle \bar{D} \rangle$ averaged across subjects
b) Run error $\langle \bar{E} \rangle$ averaged across subjects
c) Run standard deviation $\langle \bar{s} \rangle$ averaged across subjects

subjects and stimuli the run error $\langle E \rangle$ equals -1.9°. Note that the averaged run error $\langle E \rangle$ is smallest for the stimuli which contain relatively little high frequencies (speech and low-frequency noise). Note that the smaller bias for speech and low-frequency noise carries over into the results for the averaged rms error $\langle \overline{D} \rangle$.
The negative bias occurred for almost every subject, as can be found in appendix B. The bias may have an acoustical origin, since the listeners were all seated in the left half of the hall, which gives rise to a relative strong left side wall reflection. However, a systematic bias to the left has not been observed for real sources in the auditorium, so, the origin of the bias remains uncertain.
For those azimuthal angles at which the subjects could choose from 3 depths levels (A,B or C), an analysis has been made of auditory distance perception. A score-matrix can be calculated which gives the percentage of the total number of presentations that are perceived at a certain depth level for each actual depth level. The score-matrices are given in table 8.2 and 8.3 for real and synthesized sources respectively. With respect to distance perception the results in table 8.2 and 8.3 show that, for both real and synthesized sound fields, judgement of the perceived distance by subjects is possible. However, in both situations the distance resolution is very poor. By adding the diagonal elements of table 8.2 and 8.3, it is found that for real sources 49% and for synthesized sources 40% of all presentations are perceived at the correct depth level, which is above chance (33.3%).

Table 8.2: *Score matrix for the distance perception of real sources. Values are averaged across all stimuli conditions and all subjects. Total number of presentations is set 100%.*

		Perceived row		
		A	B	C
Actual row	A	15.8	12.4	2.2
	B	8.5	16.2	10.6
	C	3.2	13.9	17.2

Table 8.3: *Score matrix for the distance perception of synthesized sources. Values are averaged across all stimuli conditions and all subjects. Total number of presentations is set 100%.*

		Perceived row		
		A	B	C
Actual row	A	12.2	13.2	6.0
	B	9.9	15.0	9.4
	C	8.0	13.6	12.7

8.2.3 Auditorium: vertical localization

In this section we will resume the discussion about the vertical localization of an elevated notional source in combination with a primary source e.g. at stage. In section 6.2 it was found that applying a delay to the synthesized wave field with respect to the primary wave field, will reduce the upward shifting of the perceived sound image. To verify the validity of this method, we performed a listening test under realistic conditions.

I. Method

Consider the configuration in figure 8.17, in which the array configuration in the auditorium is

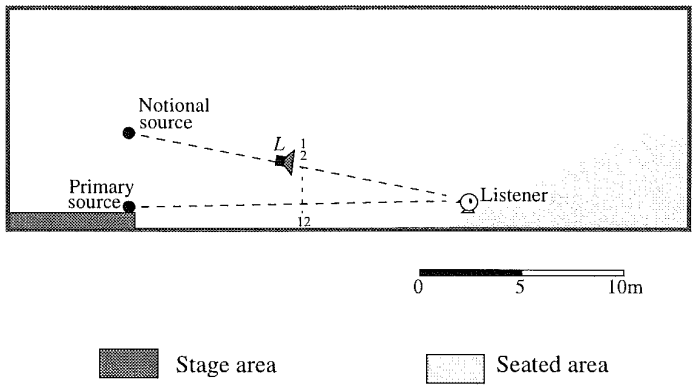


Figure 8.17: *Cross-section of the measurement configuration for the vertical localization experiments. The notional monopole source Ψ_m is positioned at $\mathbf{r}_\Psi=(x_\Psi,y_\Psi,z_\Psi)=(-1,4,-1)$. The notional monopole signal has been delayed by 5ms compared with the primary source.*

depicted, as described in section 8.1.2. In this situation the sound of a primary source at stage (one loudspeaker) is amplified by a loudspeaker array L which emits an accurate representation of the primary sound field in the horizontal plane with a delay of 5 ms. The listener is seated almost at ground level, which means that the apparent notional source has an elevation of 14.3°. The primary source at stage has an elevation of -1.2° as seen from the listener. Both the primary and the notional source are positioned in the median plane of the listener. A vertical strip with 12 numbered labels was placed in front of the listener near the array. The labels covered elevations from -3.4° to 15.8°.
Broadband noise stimuli were presented at 7 different gain levels. The gain level was varied randomly by changing the output level of the loudspeaker array in 6-dB steps. This resulted in the following gain levels: 0, 0.9, 2.8, 6.5, 11.2, 16.7 and 22.6 dB. A gain of 0dB means that only the primary source was on. Each gain level was presented six times, yielding a total of 42 presentations. The subject's task was to indicate the vertical position of the perceived source by

giving the number of the label with the corresponding source elevation. To avoid that the sound pressure level (gain) provides a direct cue for source elevation to the listeners, namely, the higher the SPL the higher the elevation, we randomly varied the overall level on each presentation. The level was selected from a 25-dB range. Four subjects participated in this experiment. Subject 1 and 2 were actively involved in building and testing the DSE system in this hall and were more familiar with the stimuli than subject 3 and 4 . The experiment was repeated for speech and music stimuli.

II.Results and discussion

The responses of the four subjects are given in figure 8.18 for three stimuli. In the graphs the number of times (max. 6) that the source is perceived at a certain elevation is plotted as a function of the gain level.

Clearly, the subjects can be divided into two groups. The first group, subject 1 and 2, is able to localize the sound source at rather distinct positions in the median plane. It can be observed that for low gains (0-5 dB) the sound image is perceived at the height of the primary source. Increasing the gain to 5-17 dB yields an ambiguous vertical source position; the lower image is perceived at the primary source position and an upper image at the notional source position. Further increase of the gain removes the uncertainty in some cases and the source is localized at the notional source position.

The second group, subjects 3 and 4, has great difficultly with localizing the sound source in the median plane. Even for the 0-dB gain setting, which means that only the primary source was on, the perceived sound image was elevated in many cases.

The results for the second group may be explained by the fact that they were less familiar with the stimuli. It has been shown by Blauert (1983) that familiarity with the sound has an enormous influence on the "localization blur" (which may be compared with the MAA) in the median plane. For example, the localization blur is about 17° for continuous speech by an unfamiliar person, and about 9° for continuous speech by a familiar person.

Though we did not test experimentally the effect of applying a zero-delay to the notional source signal, we expect on the basis of evidence in the literature that delaying the notional source signal has a positive effect on the vertical image position. One major discrepancy is found in the results of Somerville (see section 6.2, figure 6.3) and ours. Somerville reports that the subjects perceived one fused image that slowly shifts upwards for increasing gain, while our results indicate that some subjects perceive two separated images instead of one fused image.

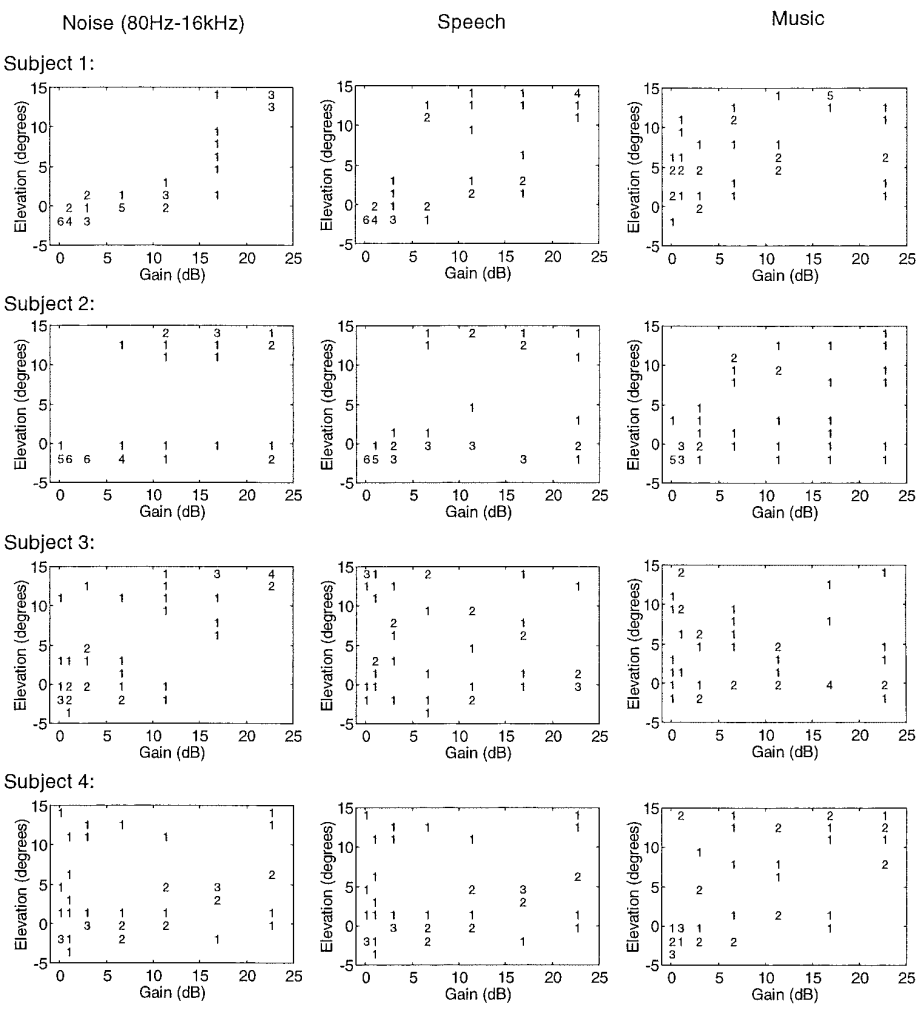


Figure 8.18: *Results of the vertical localization experiments in which 4 subjects participated. The number of times that the sound image is perceived at a certain elevation is plotted for 7 gain settings of the DSE system. The gain level was varied randomly, each gain level was presented 6 times, while the overall level was randomized. The 7 gain levels are: 0, 0.9, 2.8, 6.5, 11.2, 16.7, 22.6 dB.*

8.2.4 Concert hall

I. Method

In "De Doelen" the notional source position was chosen randomly from a pre-defined set of 17 possible locations on stage. The set of notional source positions was distributed uniformly at a line 4 m behind the array in the interval -8m<x_Ψ<8m. At this line 73 numbered labels were placed ranging from -9m<x_Ψ<9m with a spacing of 0.25 m, as shown in figure 8.19. The subject's task was to indicate the perceived location of the source by giving the label number. Six subjects participated in this experiment at the same time. The listeners were seated at the positions indicated in figure 8.19. Each experimental run consisted of 85 trials, 5 presentations of 17 different sources. No feedback was given during the run.

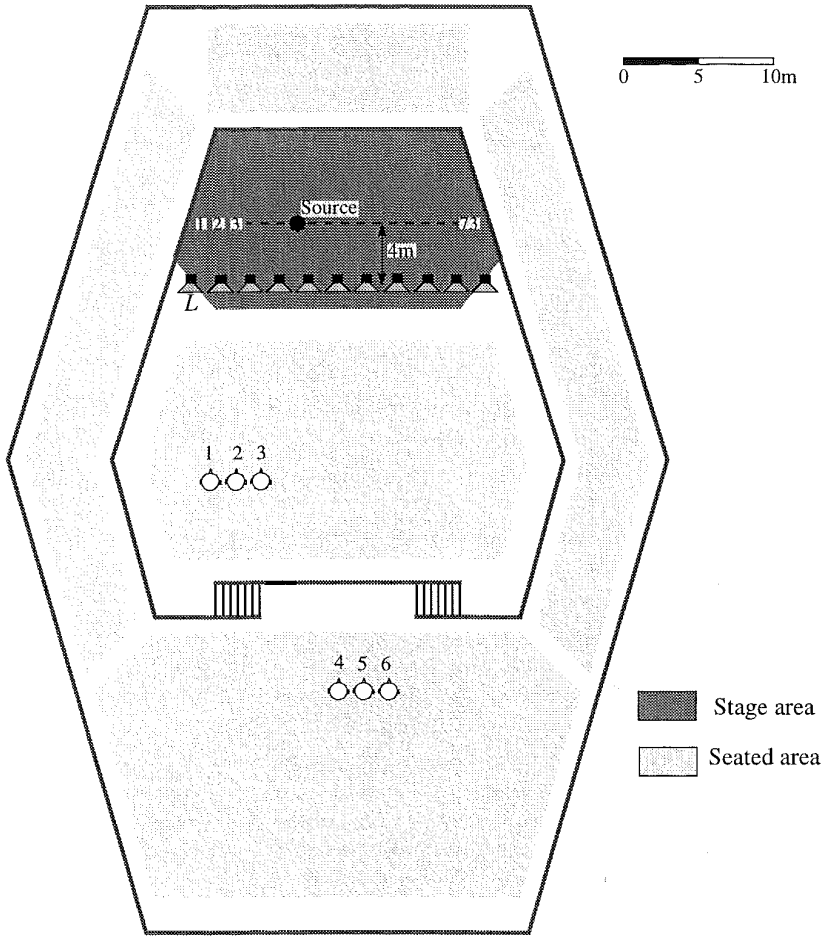


Figure 8.19: During the source identification experiment in the concert hall the subjects were asked to indicate the notional source position by giving the label number at the perceived source position at stage.

During the real source identification experiments the loudspeakers were placed on stage at the line 4 m behind the array. At each stimulus presentation one loudspeaker was used as a real source.

II. Results

The statistical quantities *D, E* and *s* were calculated. The run rms localization error $\langle \bar{D} \rangle$, the run error $\langle \bar{E} \rangle$ and the run standard deviation $\langle \bar{s} \rangle$ averaged across subjects are shown in figure 8.20a-c for real and synthesized sources. These results show that the rms localization error $\langle \bar{D} \rangle$ for synthesized sound fields is larger than in the previous halls. The run rms localization error $\langle \bar{D} \rangle$ for speech signals amounts 4.1°, while in the anechoic room and the auditorium $\langle \bar{D} \rangle$ equals 2.6° and 2.9° respectively. The increase of the rms localization error $\langle \bar{D} \rangle$ for the stimuli with random high-frequency delays, which occurred in the anechoic room and the auditorium, cannot be found in this hall.

8.2.5 Discussion of the localization experiments

In the following we will digress upon the localization experiments presented in the previous sections. First a few general observations will be summarized concerning the horizontal localization in real sound fields:

- Localization performance for broadband stimuli seems independent of the acoustic condition (reverberation time).
- Localization accuracy is best for speech stimuli in all configurations which were considered.

These results are in agreement with what Hartmann (1983) found; localization of impulsive sounds with strong attack transients, is independent of the room reverberation time. Since the speech stimulus we used contained the strongest attacks, the good localization result is not surprising.

Regarding the synthesized sound fields the following observations can be made:

- For synthesized sound fields, the localization accuracy is best for speech stimuli. In the anechoic room and the auditorium (spatial aliasing frequency of about 1.2 kHz) the averaged run rms error $\langle \bar{D} \rangle$ for synthesized speech (±2.7°) only deviates slightly from the results for a real speech source (±1.7°).
- The averaged run rms error $\langle \bar{D} \rangle$ for the low-frequency noise stimulus (below spatial aliasing) under anechoic conditions is almost identical for real and synthesized sources. In the auditorium and the concert hall the result for real sources is somewhat better than for synthesized sources.
- In all rooms the averaged run rms error $\langle \bar{D} \rangle$ for the high-frequency noise stimuli (above spatial aliasing) is markedly larger for synthesized sources than for real sources, which was

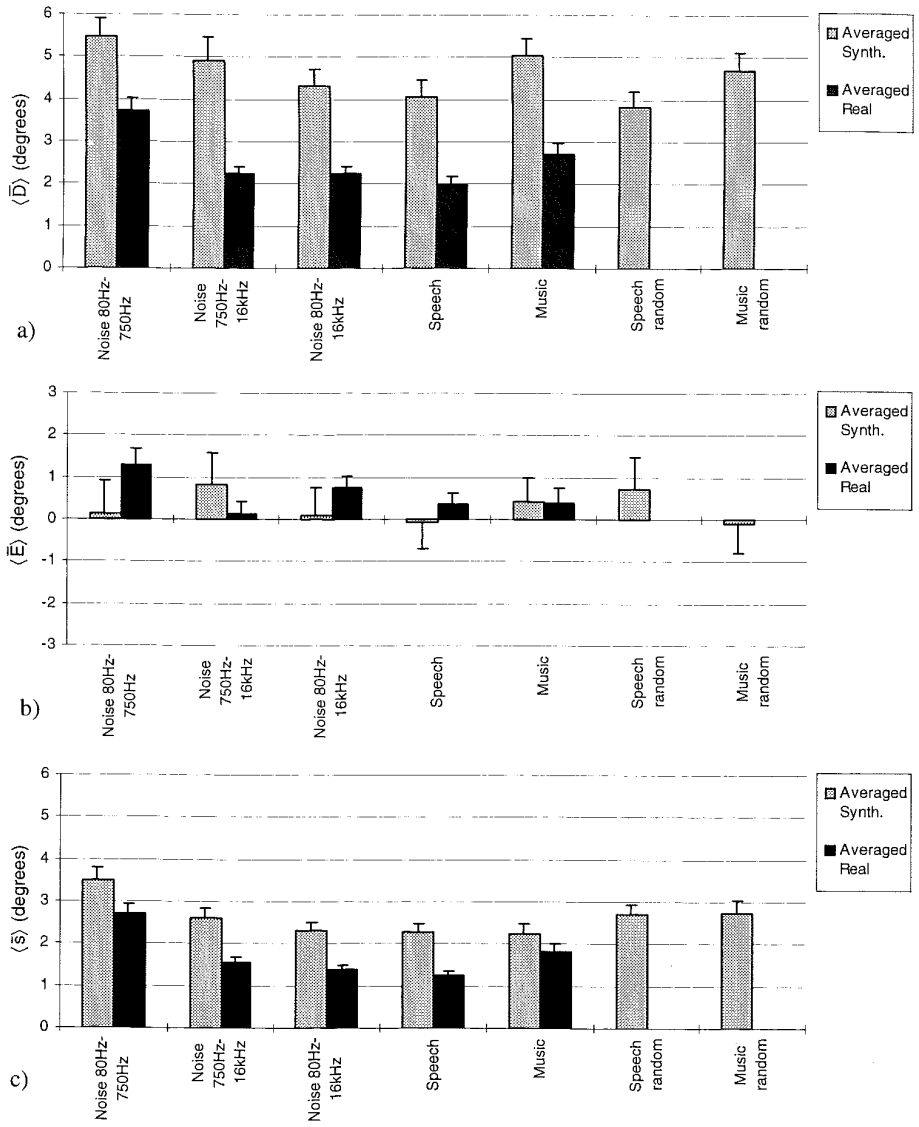


Figure 8.20: Results of the localization experiments for synthesized and real sound fields in the concert hall under several stimulus conditions.
a) Run rms error $\langle \bar{D} \rangle$ averaged across subjects
b) Run error $\langle \bar{E} \rangle$ averaged across subjects
c) Run standard deviation $\langle \bar{s} \rangle$ averaged across subjects

expected.

- In the concert hall the localization performance for synthesized sound fields is worse than in the other rooms, while the localization accuracy for real sound fields in this hall does not deviate much from the results in the other rooms.

It is supposed that, since in the anechoic room and the auditorium almost the complete ITD-frequency area is covered (spatial aliasing frequency of 1.2kHz), localization results are almost as good as for real wave fields. The worse localization performance in concert hall for synthesized sources in comparison with real sources is likely due to the fact that the synthesized wave field was only correct up to a frequency of about 750 Hz (see section 8.1.3), which means that a large part of the frequency area in which ITD cues are important is not synthesized correctly.

With respect to the localization experiments in the vertical plane, we can conclude the following:

- From evidence in the literature it is expected that delaying the notional source signal reduces the risk that the apparent source image shifts upwards. Our results indeed indicate that for gain levels up to about 5 dB the sound image remains positioned at the primary source. At higher gain levels the sound image breaks-up in two parts, a lower image at the primary source and a upper image at the notional source. It is expected that the visual cue will be decisive in those cases.
- Some listeners displayed a very poor localization accuracy in the median plane. This is contributed to the fact that they were less familiar with the stimuli, which is of great importance for vertical localization accuracy.

8.3 Scaling of perceptual attributes of sounds by paired comparison

In order to judge and to compare synthesized sound fields in different halls with each other and with real sound fields, perceptual scaling methods can be very valuable. Several psycho-physical scaling methods are available to scale a set of stimuli on the basis of a certain perceptual attribute (e.g. spaciousness or coloration). One of them we already encountered in section 6.6.1, where the direct magnitude estimation procedure was applied to scale a set of stimuli with respect to spaciousness.

An alternative method is the paired comparison procedure. In this procedure, a randomly chosen pair of stimuli from a given set is presented to the subject. The subject is asked to judge the dissimilarities between pairs of signals using only one perceptual criterion (cue). An advantage of the paired comparison method is that each stimulus is compared directly to all other stimuli of the set instead to one reference stimulus.

In the first experiment which will be presented in the next section, the subject's task is to indicate which of the two stimuli has the broadest image (the most spacious). Additionally, it can also be asked to indicate the measure of dissimilarity by giving one of five categories:

- equal (0)
- small difference (±1)
- moderate difference (±2)
- reasonable difference (±3)
- large difference (±4)

The answers can be arranged in a so-called dissimilarity matrix with elements $a_{j,k}$. If stimulus j is compared to stimulus k, and j is judged more spacious than k, the value is positive, otherwise $a_{j,k}$ is negative. If the subject is consistent in his (or her) answers, the values $a_{j,k}$ will be anti-symmetric: $a_{j,k}=-a_{k,j}$. Once all possible pairs have been judged, a perceptual scale for the stimuli can be constructed (Thurstone, 1927), ranging from zero, for the least spacious signals, to one, for the most spacious signals.

8.4 Source width experiments

In the previous section the localization for real and synthesized sound fields was investigated in several rooms. In a source identification experiment the subjects were asked to indicate the position of the center of the perceived sound image. Besides the position of the sound image, the width of the sound image is also an important quality (perceptual attribute) of the sound.
As shown in section 6.6, due to spatial aliasing the apparent source width or spaciousness (see section 6.1.4) of notional sources is larger than for real sources. Besides a 'system-induced' source broadening, reflections in a hall also increase the spaciousness of sound sources (Potter 1993), which can be regarded as 'hall-induced' spaciousness. Additionally we want to investigate the effect of randomizing the high-frequency delays, as explained in section 7.4.3, on spaciousness. It is expected that the coloration due to spatial aliasing will be reduced (see section 8.5), while the apparent source width of the synthesized sounds will increase.
In order to determine the relative spaciousness of real sources, notional sources and notional sources with random high-frequency delays, a comparative perceptual study has been made of the source width for several stimuli:

1. Broadband noise (80Hz-16kHz)
2. Speech
3. Music

The influence of the hall in which the sound is reproduced is investigated by conducting the experiments in several halls; the anechoic room, the auditorium and the concert hall.

8.4.1 Method

The spaciousness of real sources, notional sources and notional sources with random high-frequency delays, which were recorded in the three halls, was determined in a paired-comparison experiment. The total of 9 stimuli has been evaluated, yielding 81 possible pairs. Since the stimuli are not compared with themselves, a total of 72 paired comparisons per session is obtained.

The noise, speech and music stimuli have been recorded binaurally using a KEMAR dummy head. The positions of the dummy head in each hall are given by:

1. Anechoic room: at the same position as the subject in figure 8.13
2. Auditorium: at the position of subject 2 in figure 8.15
3. Concert hall: at the position of subject 5 in figure 8.19

The noise stimuli were presented in pairs of 2.5-s signal duration, the speech stimuli had a duration of 3.0 s and the music stimuli 4.0 s. Between the stimuli a silent interval of 500 ms was applied. The subject had to indicate which stimulus was the most spacious, and additionally, he had to indicate the measure of difference. The signals were presented over Etymotic Research ER-4B earphones at a level of 50 dB SL. An inverse filter has been applied to correct for the ear canal transmission. Four experienced, normal hearing subjects participated. For each of the three stimuli the experiment was repeated 3 times for each observer. The standard deviation has been calculated from this data set.
In order to compare the subjects' results with the physical measures for spaciousness, the inter-aural cross-correlation coefficient (*IACC*) and the central modulation coefficient (*CMC*) have been calculated for all signals.

8.4.2 Results

The perceived relative spaciousness of the sources has been calculated according to the method described in section 8.3. The results are shown in figure 8.21 for noise, speech and music. In addition, the measured values 1-*IACC* and 1-*CMC* are also given.

8.4.3 Discussion

The small standard deviations show that the subjects were very consistent in their answers. Although the subjective measure of spaciousness can only be expressed on a relative scale which may vary among subjects, the graphs do show that qualitatively the perceived spaciousness is very similar for all subjects. For all stimuli the relative spaciousness is smallest for the anechoic room, increases in the auditorium and is highest in the concert hall. Except for the concert hall, the spaciousness of notional sources is slightly larger than for real sources.
In the anechoic room and the auditorium, the application of random high-frequency delays increases the subjective spaciousness of the notional source. This effect is not found in the concert hall. Probably, since the spaciousness in the concert hall is already very large for real and notional sources, randomization of the high-frequency delays does not further increase the spaciousness.
It is striking that the spaciousness of the real sources in the concert hall is larger than for the notional sources. A possible explanation for this effect is the fact that the notional source is positioned higher than the real source. Especially in the concert hall this has a large effect on the strength of the lateral reflections, as shown in the *x-t* recordings of the real and the notional

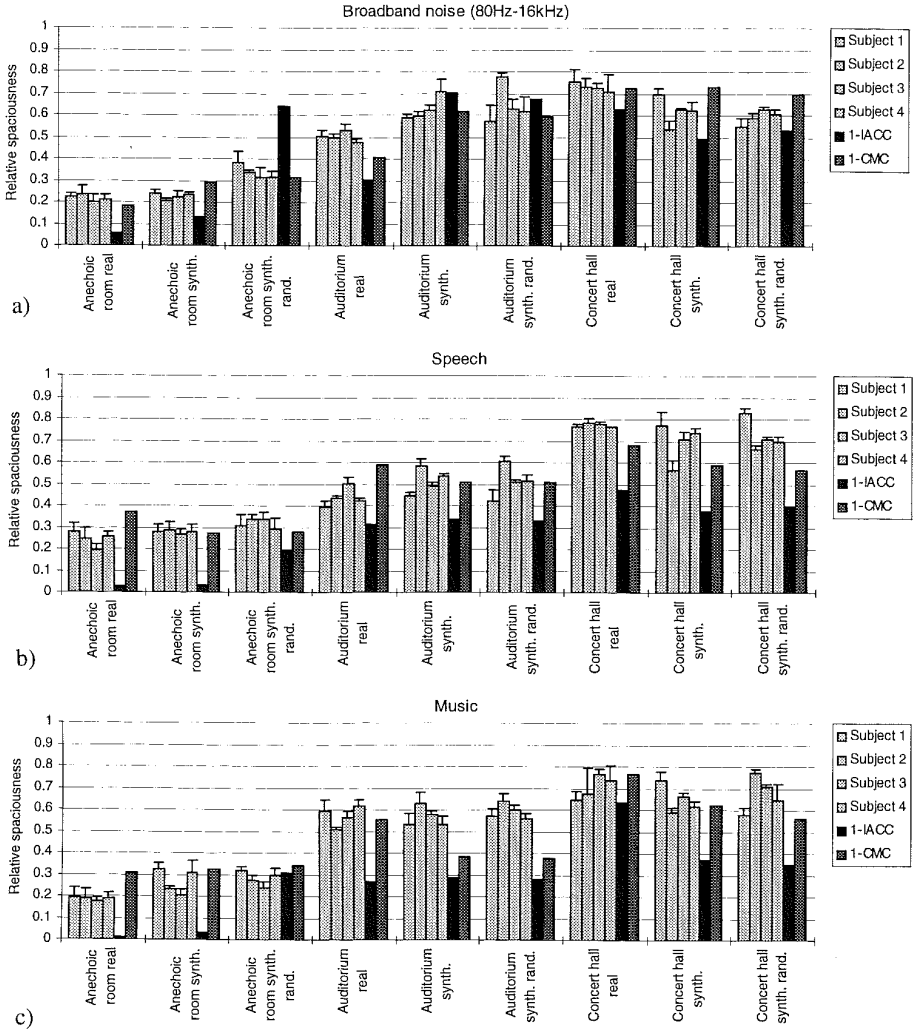


Figure 8.21: *Perceived relative spaciousness by 4 subjects measured in paired-comparison experiments for several stimuli and different halls. The physical measures for spaciousness 1-IACC and 1-CMC are also indicated. (Error bars indicate the standard error)*

a) Noise (80Hz-16kHz)

b) Speech

c) Music

source in the concert hall as shown in figure 8.9a and b. As mentioned before, the interaural correlation decreases as a result of lateral reflections.

The similarity between the physical measure 1-*CMC* and the perceived spaciousness is remarkable. The results show that the value 1-*IACC* is a less accurate indicator of spaciousness for our signals.

8.5 Experiments on coloration

In this section the coloration aspects of real and synthesized sound fields are considered. As shown in section 6.8, the coloration of synthesized sound fields due to spatial aliasing changes rapidly as a function of the position of the listener and/or the source. It has been found that, especially in an anechoic room, listeners are very sensitive for these 'color changes'. It is expected that under reverberant conditions the audibility of the 'system-induced' color changes will decrease due to the decoloration of the synthesized sound field by hall reflections. On the other hand, the hall will also introduce coloration of the synthesized sound field, however, it is experienced that this 'hall-induced' coloration is less source/receiver-dependent.

In section 7.4.3 a method was proposed to reduce the audibility of the unwanted coloration effects in synthesized wave fields by randomizing the high-frequency ($f>f_{al}$) delays of the loudspeaker signals. The objective of this experiment is twofold. Firstly, we want to investigate the effect of the acoustics of the hall and the effect of the randomization on coloration. Therefore, in three halls noise stimuli of broadband real sources, notional sources and notional sources with random high-frequency delays were compared. Secondly, we want to verify that coloration is mainly caused by spatial aliasing, which would imply that the coloration of low-frequency stimuli (below spatial aliasing) is negligible. To test the second hypothesis, low-frequency real sources and low-frequency notional sources are also included in the experiment.

8.5.1 Method

In the first part of the experiment the measure of change of coloration due to a moving source emitting stationary broadband noise was determined in a paired comparison experiment. In order to construct the signal of a moving source, the noise signal (80Hz-16kHz) of a source at different positions was binaurally recorded using a KEMAR dummy head. The azimuthal direction of the source with respect to the dummy head was changed with steps of 0.2° in a range of 2°. This range has been chosen such that the azimuthal change lies in the order of magnitude of the localization accuracy, however large enough to assure that the source is displaced several loudspeaker intervals. In this way the moving notional source signal contains time segments of which the temporal amplitude spectrum varies extremely (see section 7.4.3).

Nine series of signals have been constructed; for each hall a moving real source, notional source and a moving notional source with random high-frequency delays.

In the second part of the experiment the measure of change of coloration of a moving low-frequency ($f<f_{al}$) real and a notional source was compared with the change of coloration of the two most extreme stimulus conditions of the first part of the experiment; the real and the notional broadband noise sources in the anechoic room. So, eight series of signals have been compared: for each hall a moving low-frequency real and notional source plus the two 'reference' stimuli. The signal series were presented in pairs of 2.2-s signal duration, separated by a 1.0-s silence interval. Each series consisted of 11 segments of equal duration. During each run, consisting of 72 pairs in the first experiment and 56 pairs in the second experiment, the subject had to indicate which stimulus contained the most pronounced color differences, and additionally, he had to indicate the measure of difference in one of the five categories which were mentioned in section 8.3. The signals were presented over Etymotic Research ER-4B earphones at a level of 50 dB SL with a correction filter for the ear canal transmission. Four normal hearing subjects participated of which each completed three runs for each of the two parts of the experiment. The standard deviation has been calculated from this data set.

8.5.2 Results

The perceived relative coloration differences of the signal series in the two experiments have been scaled according to the method described in section 8.3. In order to display the results of the two experiments in one diagram on the same scale, the scale values of the second experiment have been re-scaled by matching the scale values of the two reference stimuli with the scale values obtained in the first experiment. The results are shown in figure 8.22 for both the broadband and low-pass noise stimuli in all rooms.

8.5.3 Discussion

The results of all subjects well agree with each other. The small standard deviations of the scale values indicate that the subjects were consistent in their answers. The results show that the coloration differences are smallest for the moving real source in the anechoic room, while the differences are largest for the moving notional source in the anechoic room. This was expected since in the anechoic room the decoloration of the sound due to reflections is absent.
The coloration differences for broadband notional sources in the auditorium are significantly smaller than for the broadband notional source in the anechoic room. Even for the broadband noise signals in the concert hall (f_{al}=750 Hz) the coloration has decreased in comparison to the anechoic room (f_{al}=1.2 kHz). This means that, as expected, the system-induced coloration is decolorized by reflections in a hall. Furthermore, the results indicate that, as we hypothesized, coloration differences are reduced by randomizing the high-frequency delays.
Since the scale values of the low-frequency real and notional noise sources are almost identical, it is shown that coloration in synthesized wave fields is mainly a result of spatial aliasing.

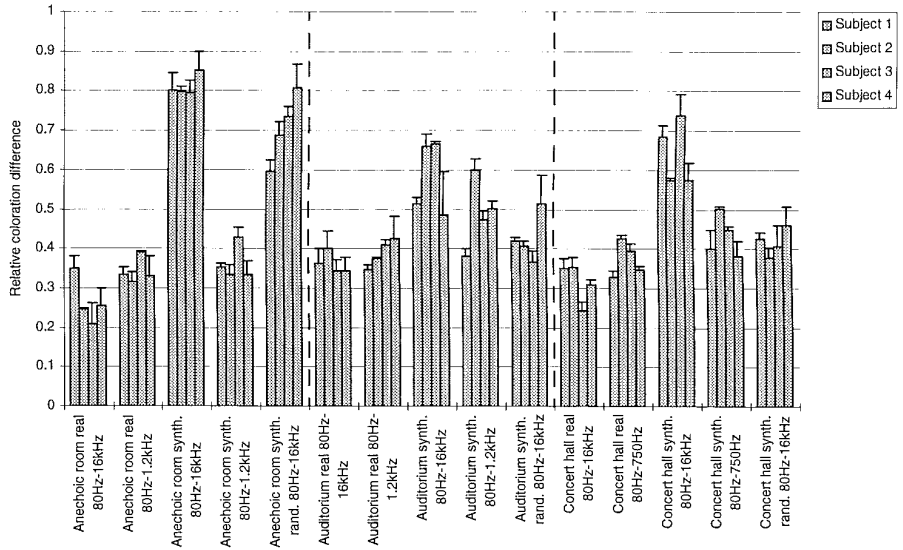


Figure 8.22: *Perceived relative coloration differences of 4 subjects measured in paired-comparison experiments for broadband noise (80Hz-16kHz) and low-pass noise (80Hz-f_{al}) in several halls. (Error bars indicate the standard deviation)*

8.6 Conclusions and discussion

The DSE system has been evaluated in several halls under various acoustic conditions. Physical as well as perceptual experiments have been conducted to test the performance of the system. In the physical experiments real and synthesized sound fields of a point source have been spatially measured and compared. The results show that by application of wave field synthesis an accurate representation of the direct real sound field can be created in a large listening area. The reflection pattern of the synthesized wave field in a hall differs from the reflection pattern of a real source. This can be explained by realizing that: 1) It is physically impossible to reconstruct the sound field in the primary source area (behind the array) correctly, and 2) in each vertical plane, perpendicular to the array, the wave fronts radiate circularly with the array as the acoustic center.
In the psycho-acoustical experiments the main subjective aspects of real and synthesized sound fields have been judged by several observers. Source identification experiments revealed that the localization of notional sound sources by listeners at various positions in a hall is nearly as good as for real sources.

Subjective scaling experiments have shown that, although the apparent width of a synthesized source is larger than the width of a real source, source broadening due to spatial aliasing is completely overruled by the source broadening due to the acoustics of a hall.
In coloration experiments, the change of coloration which is induced by a moving notional noise source has been compared to the change of coloration induced by a real noise source. By using noise signals the listeners are most sensitive for spectral changes in the signal, which may lead to coloration of the perceived noise. Results show that the coloration differences are larger for moving notional sources than for moving real sources. It has also been demonstrated that coloration is affected positively by reflections in a hall; reverberant halls tend to decolorize the notional source signal compared with an anechoic room. It has been experienced in practice that for less coloration-sensitive signals like speech coloration is hardly noticeable.
With respect to more 'coloration-sensitive' signals, like noise or music, randomization of the high-frequency contributions of the loudspeakers has proven to be a useful compromise between reduction of the audibility of coloration on the one hand and localization performance on the other hand.

Appendix A

Characteristics of the DSE loudspeakers

A.1 Frequency response of the DSE loudspeakers

The loudspeakers applied in the DSE proto-type system have a diaphragm with a diameter of 0.104 m and are mounted in aluminum bars with a length of 3 m and a cross section of 0.060×0.120 m. In each bar 24 loudspeakers are mounted. The compartment of each loudspeaker was separated from its neighbors by sound absorbing material, in order to reduce acoustic cross-talk.
The on-axis frequency response of the loudspeaker, mounted in center of the aluminum bar, is shown in figure A.1.

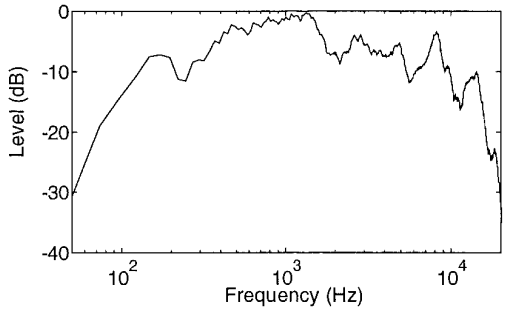


Figure A.1: *Frequency response of the DSE proto-type system loudspeakers for 50 Hz-20 kHz.*

A.2 Directivity patterns of the DSE loudspeakers

The broadband (0-25 kHz) complex frequency response of the DSE loudspeakers was measured as a fuction of the azimuthal angle. Azimuths varied in the range 0°-360° in steps of 5°. The measured directivity responses are used in simulations of the DSE system. The amplitude part of the directivity pattern of the DSE proto-type system loudspeakers is given in figure A.2 for several frequencies.

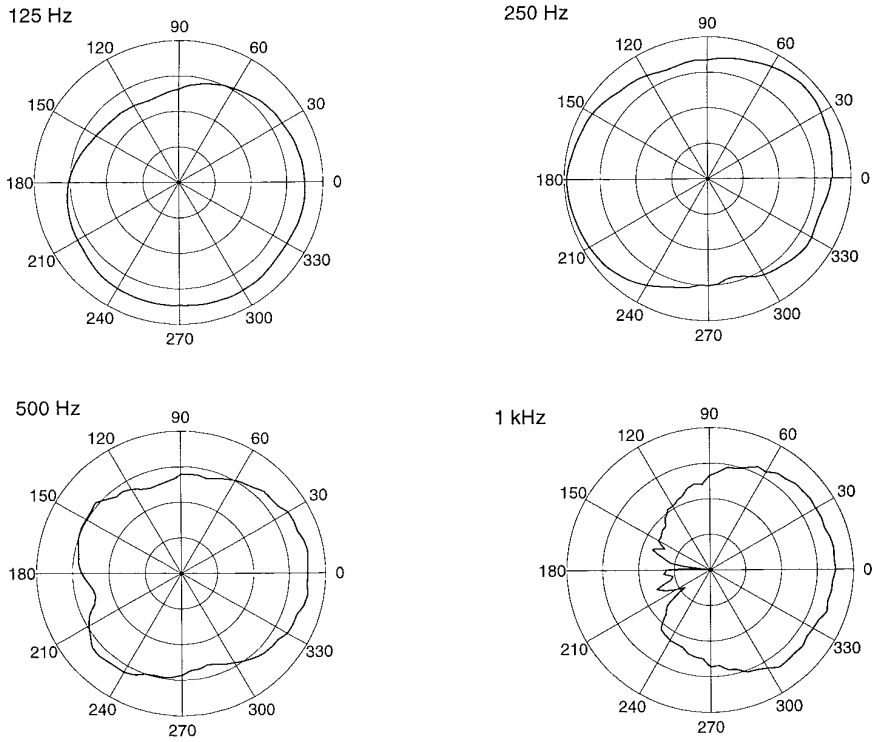


Figure A.2: *Directivity pattern of the DSE proto-type system loudspeakers at several frequencies. The radial direction is divided in 10 dB steps.*

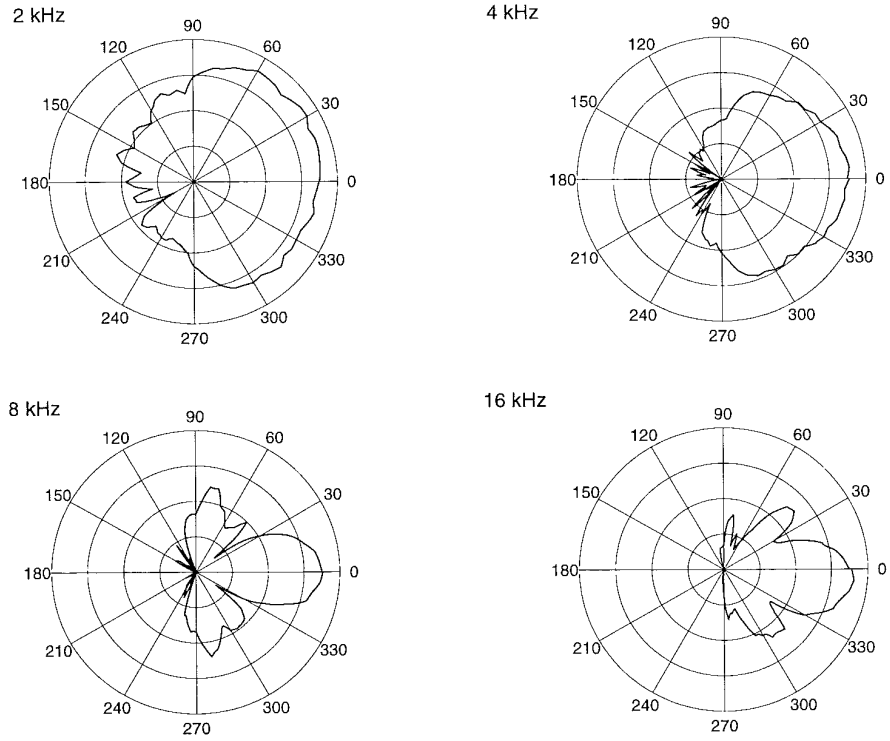


Figure A.2 Cont.: *Directivity pattern of the DSE proto-type system loudspeakers at several frequencies. The radial direction is divided in 10 dB steps.*

Appendix B

Localization results

In this appendix the individual results of the localization experiments are presented. Measurement were done in the anechoic room, auditorium and the concert hall.

B.1 Anechoic room

Figure B.1-B.3 show the run rms localization error $\overline{D}$, the run error $\overline{E}$ and the run standard deviation $\bar{s}$ for synthesized and real sound fields in the anechoic room.

B.2 Auditorium

Figure B.4-B.6 show the run rms localization error $\overline{D}$, the run error $\overline{E}$ and the run standard deviation $\bar{s}$ for synthesized and real sound fields in the auditorium.

B.3 Concert hall

Figure B.7-B.9 show the run rms localization error $\overline{D}$, the run error $\overline{E}$ and the run standard deviation $\bar{s}$ for synthesized and real sound fields in the concert hall.

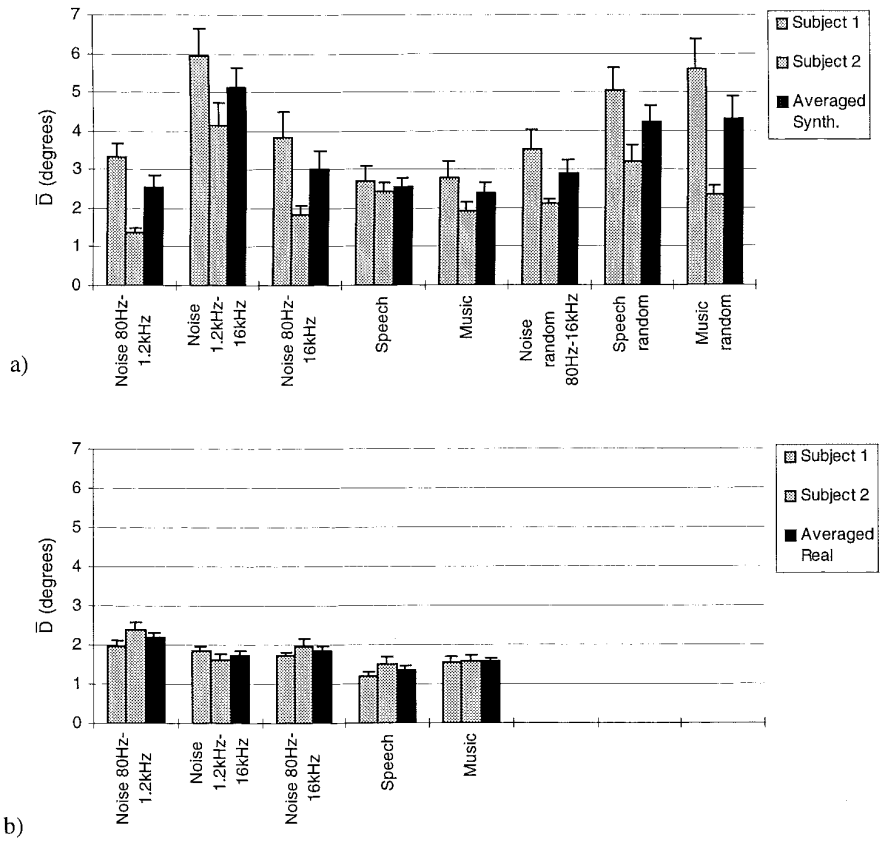


Figure B.1: *Individual and averaged values of the run rms error $\bar{D}$ calculated from localization data in the anechoic room under several stimulus conditions.*
a) Synthesized wave field (notional sources)
b) Real wave field (real sources) Run error $\bar{E}$

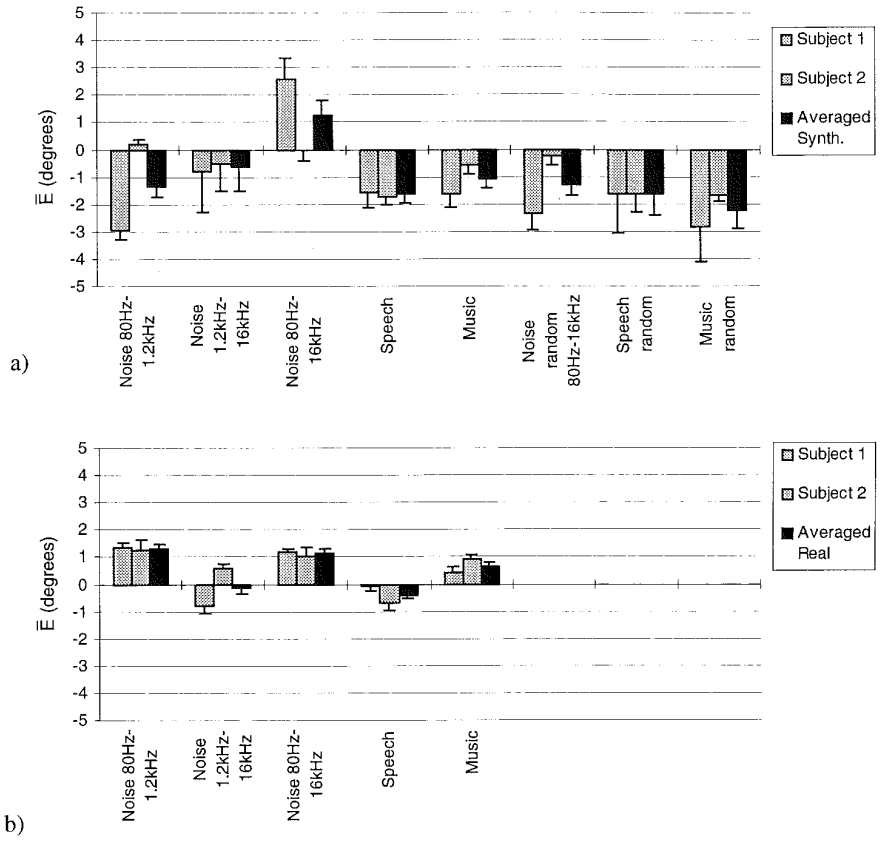


Figure B.2: *Individual and averaged values of the run error $\bar{E}$ calculated from localization data in the anechoic room under several stimulus conditions.*
a) Synthesized wave field (notional sources)
b) Real wave field (real sources)

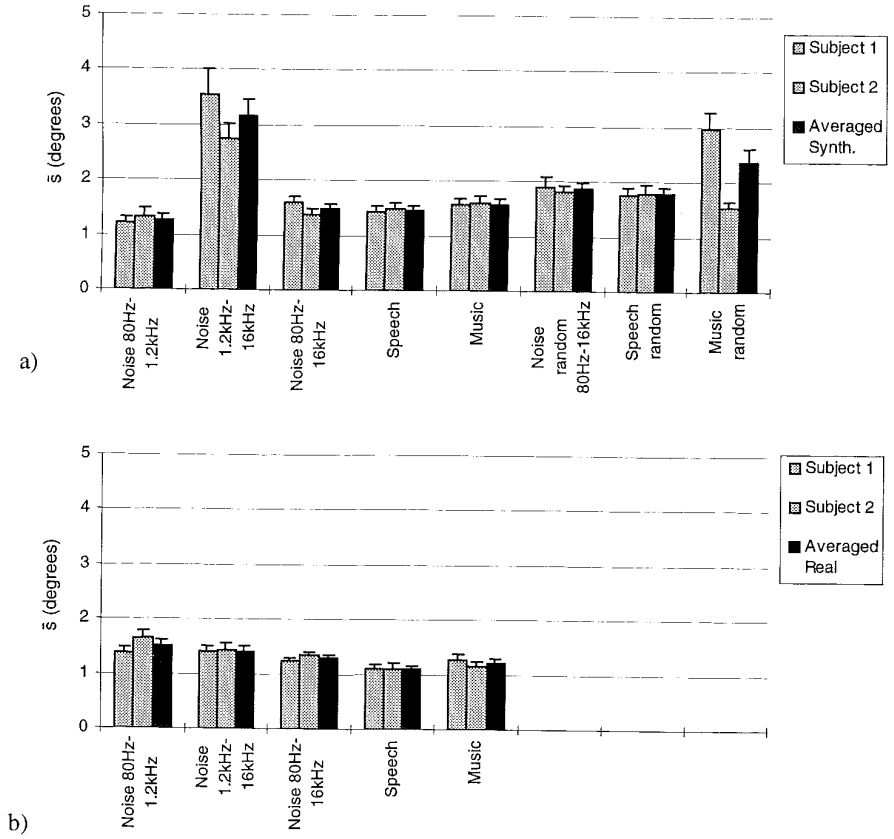


Figure B.3: *Individual and averaged values of the run standard deviation $\bar{s}$ calculated from localization data in the anechoic room under several stimulus conditions.*
a) Synthesized wave field (notional sources)
b) Real wave field (real sources)

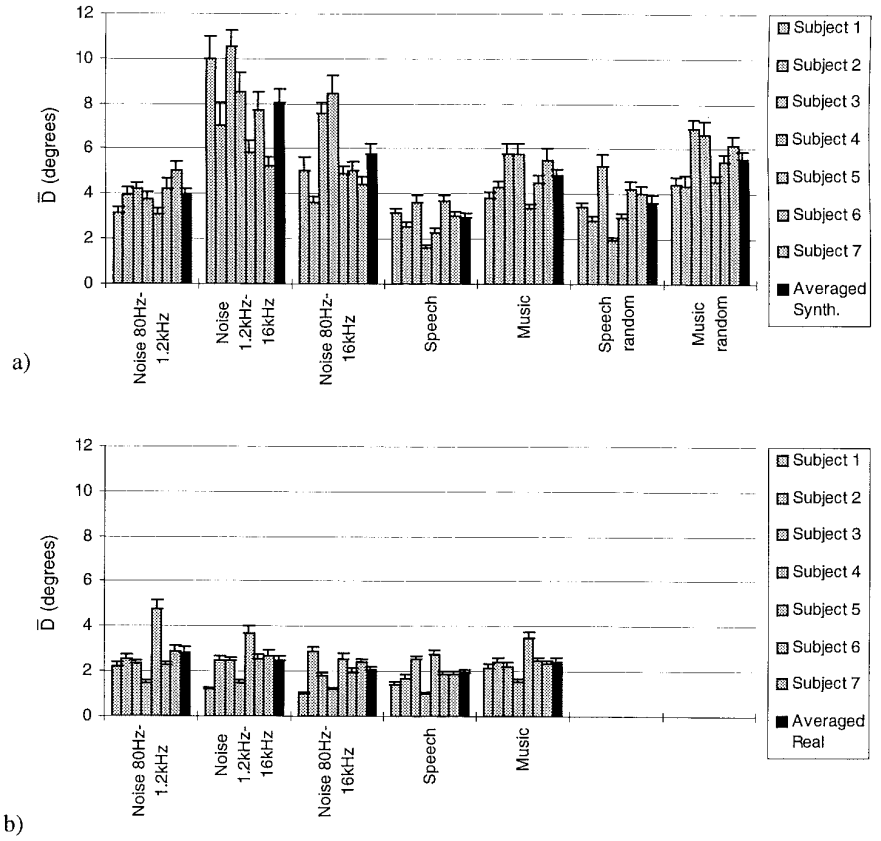


Figure B.4: *Individual and averaged values of the run rms error $\bar{D}$ calculated from localization data in the auditorium under several stimulus conditions.*
a) Synthesized wave field (notional sources)
b) Real wave field (real sources)

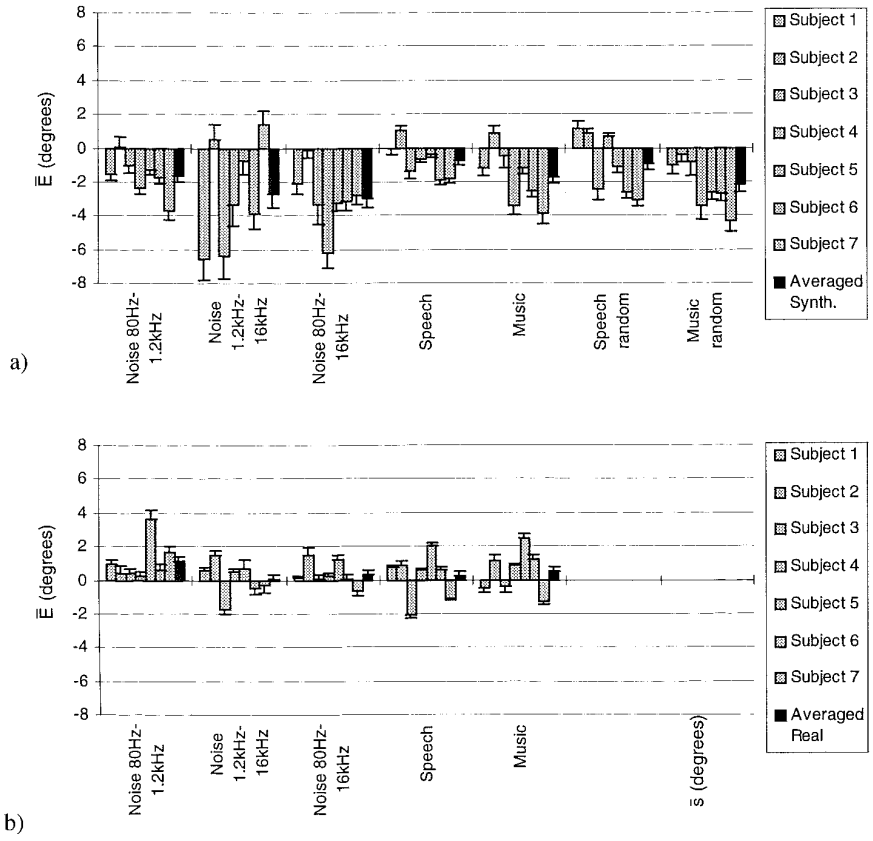


Figure B.5: *Individual and averaged values of the run error $\bar{E}$ calculated from localization data in the auditorium under several stimulus conditions.*
a) Synthesized wave field (notional sources)
b) Real wave field (real sources)

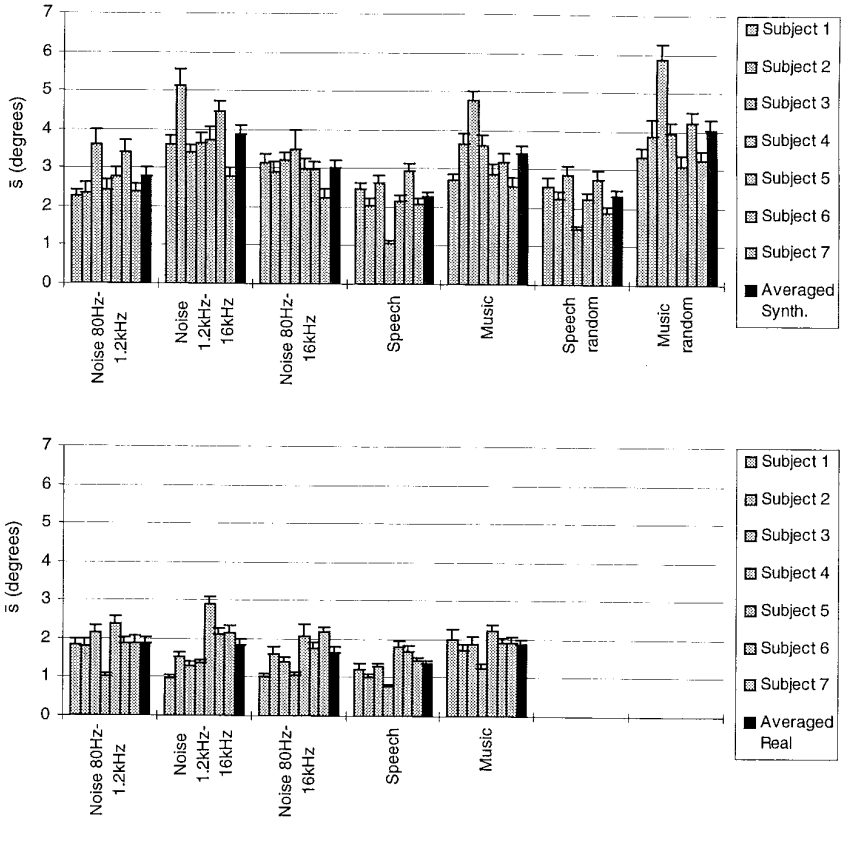


Figure B.6: *Individual and averaged values of the run standard deviation $\bar{s}$ calculated from localization data in the auditorium under several stimulus conditions.*
a) Synthesized wave field (notional sources)
b) Real wave field (real sources)

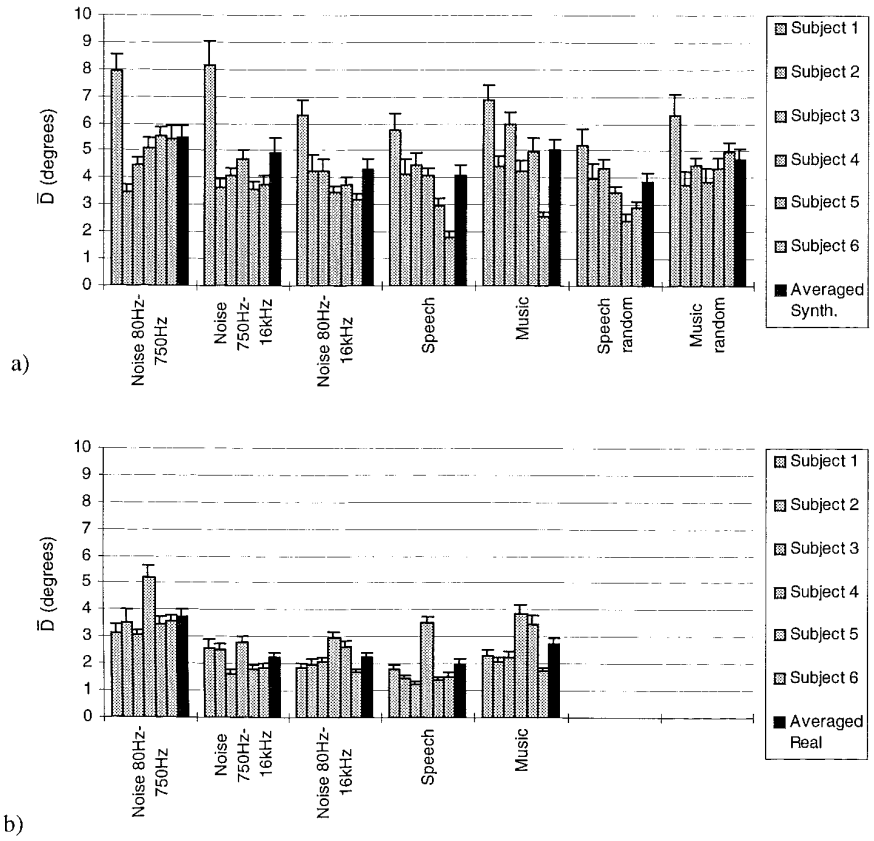


Figure B.7: *Individual and averaged values of the run rms error $\bar{D}$ calculated from localization data in the concert hall under several stimulus conditions.*
a) Synthesized wave field (notional sources)
b) Real wave field (real sources)

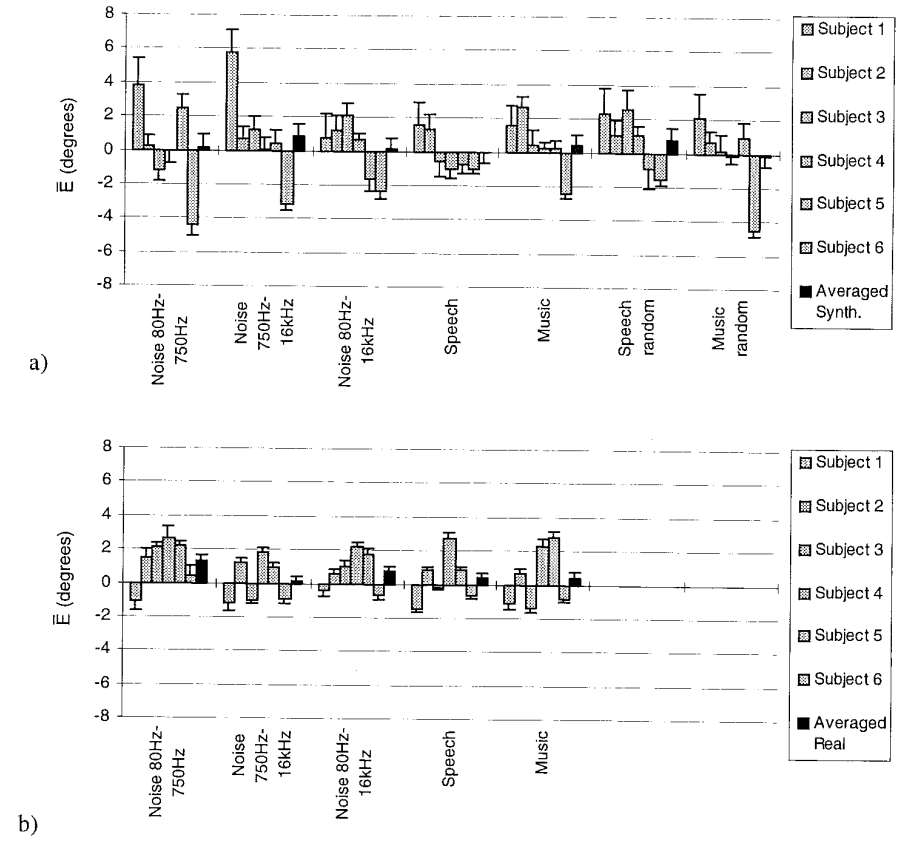


Figure B.8: *Individual and averaged values of the run error $\bar{E}$ calculated from localization data in the concert hall under several stimulus conditions.*
a) Synthesized wave field (notional sources)
b) Real wave field (real sources)

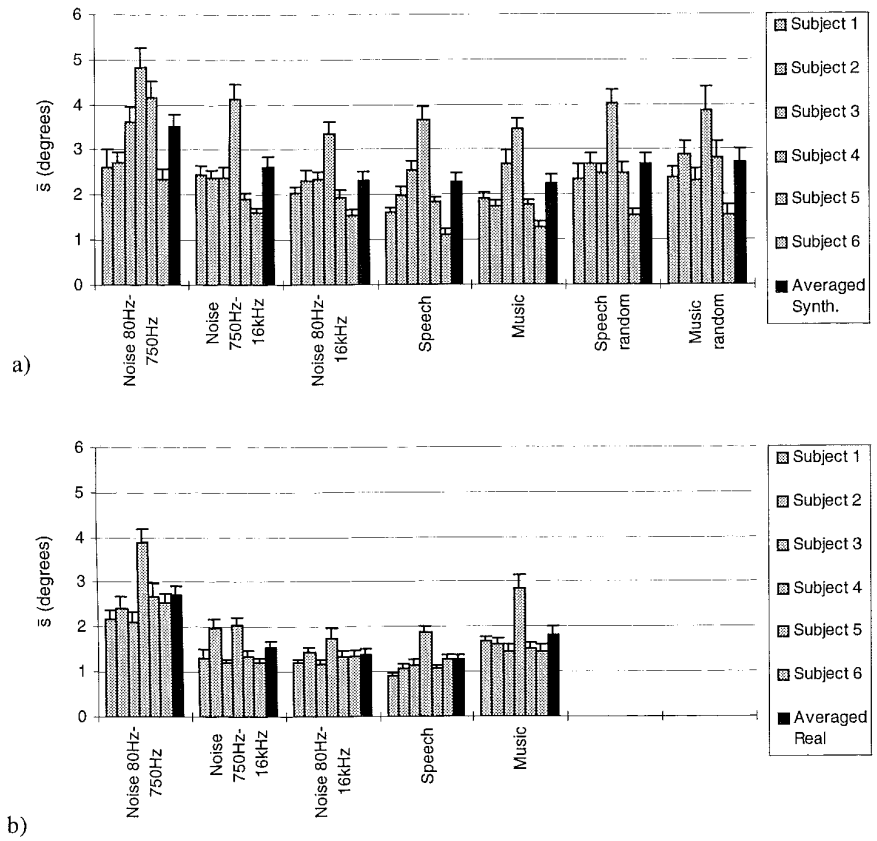


Figure B.9: *Individual and averaged values of the run standard deviation $\bar{s}$ calculated from localization data in the concert hall under several stimulus conditions.*
a) Synthesized wave field (notional sources)
b) Real wave field (real sources)

References

Abramowitz, M. and Stegun, I.A. (1965), Handbook of mathematical functions, Dover Publications Inc., New York

Ahnert, W. (1986), The complex simulation of acoustical sound fields by the delta streophony system (DSS), paper presented at the 81th Convention of the Audio Eng. Soc. (Abstracts), **34**, 1035, preprint 2418

Ahnert, W. and Steffen, F. (1993), Beschallungstechnik, Grundlagen und Praxis, S. Hirzel Verlag Stuttgart-Leipzig

Ando, Y. (1985), Concert hall acoustics, Springer Verlag, Berlin

Barron, M. and Marshall H.A. (1981), Spatial impression due to early reflections in concert halls, J, Sound and Vib. **77**(2), 211-232

Beranek, L.L. (1992), Concert hall acoustics—1992[a)], J. Acoust. Soc. Am. **92**, 1-39

Berkhout, A.J. (1987), Applied seismic wave theory, Elsevier, Amsterdam

Berkhout, A.J. (1988), A holographic approach to acoustic control, J. Audio Soc. **36** (12), pp. 977-995

Bernstein, L.R. and Green D.M. (1987), The profile-analysis bandwidth, J. Acoust. Soc. Am. **81**, 1888-1895

Bilsen, F.A. (1977), Pitch of noise signals: Evidence for a "central spectrum", J. Acoust. Soc. Am. **61**, 150-161

Bilsen, F.A. (1985), Akoestische perceptie, lecture notes

Blauert, J. (1983), Spatial hearing: The psychophysics of human sound localization, trans. J.S. Allen, Cambridge, MA: MIT

Bleistein, N. (1984), Mathematical methods for wave phenomena, Academic Press Inc., Orlando

Boone, M.M., de Vries D. and Berkhout, A.J. (1991), Geluidbeheersing/ Sound Control, lecture notes

Boone, M.M., Verheijen, E.N.G. and van Tol, P.F. (1995), Spatial sound-field reproduction by wave field synthesis, J. Audio Eng. Soc. **43** (12)

Boone, M.M., Verheijen, E.N.G. and Jansen, G. (1996), Virtual reality by sound reproduction based on wave field synthesis, paper presented at the 100th Convention of the Audio Eng. Soc.

Bradley, J.S. (1994), The influence of late arriving energy on spatial impression, J. Acoust. Soc. Am. **97**, 2263-2273

Dudgeon, D.E. and Mersereau, R.M. (1984), Multidimensional digital signal processing, Prentice/Hall International, Inc.

Gardner, M.B. (1968), Historical background of the Haas and/or precedence effect, J. Acoust. Soc. Am. **43**, 1243-1248

Green, D.M., Kidd, G. and Picardi, M.C. (1983), Successive versus simultaneous comparison in auditory intensity discrimination, J. Acoust. Soc. Am. **73**, 639-643

Green, D.M., Onsan, Z.A. and Forrest, T.G. (1987), Frequency effects in profile analysis and detecting complex spectral changes, J. Acoust. Soc. Am. **81**, 692-699

Haas, H. (1951), Einfluß eines Einfachechos auf die Hörsamkeit von Sprache, Acustica **1**, pp. 49-58

Hartmann, W.M. (1983), Localization of sound in rooms, J. Acoust. Soc. Am. **74**, 1380-1391

Hidaka, T., Okano, T., and Beranek, L. (1991), Studies of Inter-Aural Cross Correlation (IACC) and its relation to subjective evaluation of the acoustics of concert halls, 122nd Meeting of Acoust. Soc. Am., Houston, Texas, November

Huygens, C. (1690), Traite de la lumiere; ou sont expliquees les causes de ce qui luy arrive dans la reflexion et dans la refraction et particulierement dans l'etrange refraction du cristal d'Islande; avec un discours de la cause de la pesanteur, Van der Aa, P. , Leiden

Keet, W. de V. (1968), The influence of early lateral reflections on spatial impression, 6th International Congress on Acoustics, Tokyo, August

McFadden, D., Pasanen, E.G. (1976), Lateralization at high frequencies based on interaural time differences, J. Acoust. Soc. Am. **69**, 1112-1118

Mills, A.W. (1958), On the minimum audible angle, J. Acoust. Soc. Am. **30**, 237-246

Nielsen, S.H. (1993), Auditory distance perception in different rooms, J. Audio Eng. Soc. **41** (10)

Patterson, R.D. and Moore, B.C.J. (1986), Auditory filters and excitation patterns as representations of frequency resolution, Frequency selectivity in hearing, ed. B.C.J. Moore, Academic Press, London

Perrot, D.R. and Saberi, K. (1990), Minimum audible angle thresholds for sources varying in both elevation and azimuth, J. Acoust. Soc. Am. **87**, 1728-1737

Pierce, A.D. (1981), Acoustics: An Introduction to its Physical Principles and Applications, McGraw-Hill

Potter, J.M. (1993), On the binaural modelling of spaciousness in room acoustics, Ph.D. thesis, Delft University of Technology

Potter, J.M. (1995), Measures for spaciousness in room acoustics based on a binaural strategy, Acta Acust. 3, 429-443

Raatgever, J. (1980), On the binaural processing of stimuli with different interaural phase relations, Ph.D. thesis, Delft University of Technology, Dutch Eff. Bureau, Pijnacker

Raatgever, J. (1980), Lateralization on the basis of the signal envelope, in Proc. 10^{th} ICA, Sydney 1980

Raatgever, J. and Bilsen, F.A. (1986), A central spectrum theory of binaural processing. Evidence from dichotic pitch, J. Acoust. Soc. Am. **80**, 429-441

Rakerd, B. (1992), Precedence effect with and without interaural differences-Sound localization in three planes, paper presented at the 124th meeting of the Acoust. Soc. of Am.

Salomons, A.M. (1995), Coloration and binaural decoloration of sound due to reflections, Ph.D. thesis, Delft University of Technology, Eburon P&L

Saberi, K., Dostal, L., Sadralodabai, T. and Perrot, D.R. (1991), Minimum audible angles for horizontal, vertical, and oblique orientations: Lateral and dorsal planes, Acustica **75**, 57-61

Somerville, T., Gilford, C.L.S., Spring, N.F., Negus, R.D.M. (1966), Recent work on the effects of reflectors in concert halls and music studios, J. Sound Vib. **3** (2), 127-134

Svensson, P.U. (1995), Computer simulations of periodically time-varying filters for acoustic feedback control, J. Audio Eng. Soc. **43** (9)

Thurstone, L.L. (1927), Law of comparative judgement (case V), Psychological review **34**, 273-286

Van der Wal, M., Start, E.W. and de Vries, D. (1996), Design of logaritmically spaced constant-directivity transducer arrays, J. Audio Eng. Soc. **44** (6)

Vogel, P. (1993), Application of wave field synthesis in room acoustics, Ph.D. thesis, Delft University of Technology, Gebotekst, Zoetermeer

de Vries, D., Reijnen, A.J. and Schonewille, M.A. (1994), The wave field synthesis concept applied to generation of reflections and reverberation, paper presented at the 96th Convention of the Audio Eng. Soc.

Wagenaars, W.M., (1990), Localization of sound in a room with reflecting walls, J. Audio Eng. Soc. **38** (3)

Wapenaar, C.P.A., Berkhout, A.J. (1989), Elastic wave field extapolation, Elsevier, Amsterdam

Wapenaar, C.P.A. (1992), The stationary phase method and its application in seismics and acoustics, Lecture notes

Warren, R.M., Bradford, J.A. (jr.) and Wrightson, J.M. (1980), Infrapitch echo, J. Acoust. Soc. Am. **68**, 1301-1305

Wightman, F.L. and Kistler, D.J. (1992), The dominant role of low-frequency interaural time differences in sound localization, J. Acoust. Soc. Am. **91**, 1648-1661

Zurek, P. M. (1987), The precedence effect, In W. A. Yost & G. Gourevitch (Eds.), Directional hearing (pp. 85-105), New York: Springer-Verlag

Summary

Direct sound enhancement by wave field synthesis

In this thesis the design of a direct sound enhancement (DSE) system based on the principle of *wave field synthesis* is described. Wave field synthesis is the technique of combining the wave fields of a set of individual point sources (loudspeakers) to generate a wave field with predefined properties. Generally, the wave field of a virtual monopole point source, a so-called notional monopole source, will be synthesized. This technique is closely related to the Huygens principle.
Since the wave field synthesis method will also be applied in other areas of electro-acoustics, like reverberation control and sound reproduction, the main objectives of this thesis are to create a solid theoretical basis for wave field synthesis applications, to enlarge the expertise of a practical wave field synthesis system and to reduce the artifacts which arise when a system is adapted to practical and technical demands.

The acoustical principle of wave field synthesis lies in the Kirchhoff-Helmholtz theorem which describes the reconstruction of a primary wave field inside a closed surface by using a continuous distribution of secondary monopoles and dipoles on that closed surface. Simplification of this theorem yields the Rayleigh theorems, which state that the wave field can also be reconstructed in a half-space by means of only monopoles or dipoles on an infinite plane surface.

Since planar distributions of secondary sources are unacceptable in practical applications, the first step in the design of the DSE system is to replace them by horizontal line distributions of secondary sources, so-called line arrays. In order to synthesize a wave field correctly in a certain horizontal plane in a 3D space using only line arrays, the 2½D Rayleigh operators and the 2½D Kirchhoff-Helmholtz operator have been derived. Adapting the 2½D Kirchhoff-Helmholtz operator led to the generalized 2½D Rayleigh operators. These operators offer new pos-

sibilities in wave field synthesis applications by making it possible to synthesize a wave field using a bent line array consisting of only monopole or dipole sources.

In practical applications the infinite and continuous secondary source array must be replaced by a truncated and discretized loudspeaker array.
It has been shown that the effect of truncation of the array can be described as a diffraction phenomenon, and consequently, diffraction theory can be used to model the truncation effects. Truncation effects can be reduced very well, for the mid and high frequencies, by tapering the driving signals of the array.
The performance of discretized arrays has been analyzed extensively. The minimum spatial sampling distance of the loudspeakers that can be achieved in a practical application imposes a restriction on the maximum temporal frequency that can be processed correctly. For frequencies above this maximum frequency spatial aliasing occurs. Therefore a method has been introduced to reduce the spatial bandwidth of the synthesized wave field in favor of the temporal frequency contents. This method involves four steps: spatial anti-aliasing filtering, spatial sampling, spatial synthesis and spatial reconstruction filtering.
It has been shown that the first step, spatial anti-aliasing filtering of the driving signal, results in an aperture reduction of the linear array as seen from the notional source. After the second step, spatial sampling of the loudspeaker array, the wave field is synthesized using monopole or dipole louspeakers (step 3) and spatially reconstructed by applying a spatial reconstruction filter on the synthesized wave field in the listening area (step 4). This last step can only be realized in theory since it is impossible in practice to filter the sound field afterwards. However, it has been shown that step 3 and 4 can be combined by synthesizing the wave field using directional loudspeakers, which represent monopole or dipole point sources with an additional directivity pattern. In this way the theoretical method of spatial bandwidth reduction has been adapted to the practical situation.
Further it has been demonstrated that the technique of spatial bandwidth reduction, which has been developed for linear arrays, can be successfully integrated in the application of bent line arrays. By application of convex line arrays the beneficial effect of spatial bandwidth reduction is combined with a large array aperture, yielding a high spatial aliasing frequency and a large receiver area.

Besides the physical behavior of synthesized sound fields, the subjective appreciation is of great importance. Therefore a preliminary research has been conducted to the perceptual aspects of synthesized sound fields in an anechoic room. Synthesized wave fields must be evaluated at least with respect to acoustic source localization, spaciousness (apparent source width) and coloration. Each of these perceptual criteria can be quantified using psycho-acoustic measurement methods. Moreover, it is possible to relate the subjective results with objective physical measures. The outcome of the preliminary perceptual investigations can be summarized as follows:

1. Localization accuracy in synthesized sound fields is equal to the accuracy in real sound fields, provided that the sound field is synthesized correctly up to a frequency of at least 1.5 kHz.
2. Due to spatial aliasing the spaciousness of synthesized sound fields is larger than for real sound fields.
3. Coloration artifacts due to diffraction at the ends of the array are negligible for practically realizable loudspeaker arrays.
4. Distortion of the temporal spectrum due to spatial aliasing results in a source and receiver position-dependent coloration of synthesized signals.

Based on the wave field synthesis principle a full-scale Direct Sound Enhancement (DSE) prototype system has been designed and built. The developed spatial bandwidth reduction technique has been applied in the prototype system. The prototype system consists of a (bent) array of 192 loudspeakers, which can be adapted to the geometry of any hall. In the prototype system technical, economical and perceptual criteria have been combined to arrive at an optimal solution. In order to reduce the audibility of coloration due to spatial aliasing, an alternative method for the processing of high frequencies has been introduced, which randomizes the arrival times of the high-frequency contributions of the loudspeakers.
Although the research is mainly devoted to the reproduction part of the system, several strategies for the sound pick up and methods for tracking of sound sources have been proposed.

The DSE system has been evaluated in several halls under various acoustic conditions. Physical as well as perceptual experiments have been conducted to test the performance of the system. In the physical experiments real and synthesized sound fields of a point source have been spatially measured and compared. The results show that by application of wave field synthesis an accurate replica of the real sound field can be created in a large listening area.
In the psycho-acoustical experiments the main subjective aspects of real and synthesized sound fields have been judged by several observers. Source identification experiments revealed that the localization of notional sound sources by listeners at various positions in a hall is very good. Subjective scaling experiments have shown that, although the apparent width of a notional source is larger than the width of a real source, source broadening due to spatial aliasing is small compared to the source broadening due to the acoustics of a hall.
By comparison of moving notional sources with moving real sources it has been found that spatial aliasing may lead to coloration of the perceived sound. However, it has been demonstrated that coloration is reduced by reflections in a hall; reverberant halls tend to decolorize the notional source signal compared with an anechoic room. It has been experienced in several halls that for speech signals coloration is hardly noticeable. For more coloration-sensitive signals randomization of the high-frequency contributions has proven to be a useful compromise

between reduction of the audibility of coloration on the one hand and localization performance on the other hand.
Finally we conclude this summary by stating that application of wave field synthesis in direct sound enhancement makes it possible to amplify the sound of a certain source in a natural way. Unlike all existing methods, the wave field synthesis approach solves fundamentally the problem of mis-localization of amplified sound sources by listeners in a room by synthesizing the wave field of a notional source at the same position as the real source. In this way an accurate replica of the original wave field is obtained in the entire listening space.

Samenvatting

Direct-geluidsversterking met behulp van golfveldsynthese

In dit proefschrift wordt het ontwerpproces beschreven van een versterkingssysteem voor direct geluid (Eng: Direct Sound Enhancement system, ofwel DSE-system) dat gebaseerd is op het principe van *golfveldsynthese*. Golfveldsynthese is een techniek waarmee het mogelijk is om de golfvelden van individuele puntbronnen zodanig te combineren dat er één golfveld ontstaat met bepaalde eigenschappen. In het algemeen zal het golfveld van een virtuele monopoolbron gegenereerd worden. Deze techniek is nauw verwant aan het principe van Huygens.
Aangezien de golfveldsynthesemethode ook toegepast zal worden op andere gebieden van de elektro-akoestiek, zoals het beheersen van nagalm en geluidsreproductie, kunnen de belangrijkste doelen van dit proefschrift als volgt gesteld worden: het vormen van een theoretische basis voor het toepassen van golfveldsynthese, het vergroten van de ervaring met golfveldsynthesesystemen en het reduceren van artefacten die naar voren komen bij een system dat aangepast is aan eisen die de techniek en de praktijk stellen.

De oorsprong van golfveldsynthese ligt in het Kirchhoff-Helmholtz theorema. Dit theorema beschrijft hoe een primair veld gereconstrueerd kan worden binnen een gesloten oppervlak door middel van een continue verdeling van secundaire monopolen en dipolen op dat gesloten oppervlak. De Rayleigh-theorema's, die volgen uit een vereenvoudiging van dit theorema, geven aan dat het golfveld ook reconstrueerd kan worden door alleen monopolen of dipolen te plaatsen op een oneindig plat vlak.

Aangezien vlakverdelingen van secundaire bronnen onacceptabel zijn in praktische toepassingen, moeten deze vervangen worden door lijnverdelingen van secundaire bronnen, zogenaamde lijn-arrays. De volgende stap is het afleiden van de 2½D Rayleigh- en Kirchhoff-Helmholtz-operatoren, waarmee op een correcte manier, slechts gebruikmakend van lijn-

arrays, een golfveld gesynthetiseerd kan worden in een horizontaal vlak in een 3D ruimte. Door het aanpassen van de Kirchhoff-Helmholtz-operator zijn de gegeneraliseerde 2½D Rayleigh-operatoren verkregen. Met behulp van deze operatoren is het nu mogelijk om golfvelden te synthetiseren met gekromde lijn-arrays die slechts bestaan uit monopolen of dipolen.

In de praktijk moeten de oneindig lange en continue arrays van secundaire bronnen vervangen worden door afgekapte en gediscretiseerde luidsprekerarrays. Het is aangetoond dat de effecten tengevolge van het afkappen van het array beschreven kunnen worden als een diffactiefenomeen, en dientengevolge kan diffractietheorie gebruikt worden om deze afkapeffecten te modelleren. Voor de midden en lage frequenties kunnen de diffractie-effecten effectief bestreden worden door middel van het wegen van de stuursignalen van het array.
Het gedrag van gediscretiseerde arrays is uitgebreid onderzocht. De minimale spatiële bemonsteringsafstand tussen de luidsprekers die bereikt kan worden bij een praktische toepassing legt een beperking op aan de maximale temporele frequentie die correct weergegeven kan worden. Boven deze maximale frequentie treedt spatiële aliasing op. Om die reden is er een methode geïntroduceerd die de spatiële bandbreedte van het gesynthetiseerde golfveld beperkt ten gunste van de temporele bandbreedte. Deze methode bestaat uit vier stappen: spatiële anti-aliasing-filtering, spatiële bemonstering, spatiële synthese and spatiële reconstructie-filtering.
De eerste stap, het toepassen van een spatieel anti-aliasing-filter op de stuursignalen, resulteert in een apertuurbegrenzing van het lineaire array gezien vanuit de virtuele bron. Na de tweede stap, het spatieel bemonsteren van het luidspreker-array, wordt het golfveld gesynthetiseerd met behulp van monopool- of dipoolluidsprekers (stap 3) en wordt spatieel reconstrueerd door het toepassen van een spatieel reconstructiefilter op het gesynthetiseerde golfveld in het luistergebied (stap 4). Deze laatste stap is alleen realiseerbaar in theorie, aangezien het achteraf filteren van het geluidsveld in de praktijk onmogelijk is. Het is echter aangetoond dat stap 3 en 4 gecombineerd kunnen worden door het golfveld te synthetiseren met behulp van richtende luidsprekers, die gerepresenteerd kunnen worden als monopool- of dipool luidsprekers met een extra richtkarakteristiek. Op deze manier hebben we de theoretische aanpak van spatiële bandbreedteberking aangepast aan de praktische mogelijkheden.
Verder is aangetoond dat de spatiële bandbreedtebeperkingstechniek, die ontwikkeld is voor lineaire arrays, met succes geïntegreerd kan worden in de toepassing van gekromde lijn-arrays. Door het toepassen van convexe lijn-arrays kan het gunstige effect van spatiële bandbreedtebeperking gecombineerd worden met een grote array-apertuur, hetgeen een hoge spatiële aliasing-frequentie en een groot luistergebied tot gevolg heeft.

Naast het fysische gedrag van gesynthetiseerde geluidsvelden is een subjectieve beoordeling van groot belang. Om die reden is er in een reflectie-vrije ruimte verkennend onderzoek gedaan naar de perceptieve aspecten van gesynthetiseerde geluidsvelden. Het bleek dat gesynthetiseerde geluidsvelden op zijn minst geëvalueerd moesten worden ten aanzien van

akoestische lokalisatie, ruimtelijkheid (schijnbare bronverbreding) en kleuring. Elk van deze perceptieve criteria kan gekwantificeerd worden met behulp van psycho-fysische meetmethoden. Bovendien is mogelijk om de subjectieve resultaten te relateren aan objectieve fysische maten. De resultaten van het verkennende perceptieve onderzoek kunnen als volgt samengevat worden:

1. De lokalisatienauwkeurigheid in gesynthetiseerde geluidsvelden is gelijk aan de nauwkeurigheid in echte velden, mits het geluidsveld correct gesynthetiseerd is tot een frequentie van tenmiste 1,5 kHz.
2. Ten gevolge van spatiële aliasing is de ruimtelijkheid van gesynthetiseerde geluidsvelden groter dan voor echte geluidsvelden.
3. Kleuringsartefacten door diffractie aan de uiteinden van het array zijn verwaarloosbaar voor praktisch realiseerbare luidspreker-arrays.
4. Vervorming van het temporele spectrum ten gevolge van spatiële aliasing heeft een bron- en ontvangerpositie-afhankelijke kleuring tot gevolg.

Gebaseerd op het principe van golfveldsynthese is er een prototype van het direct geluidsversterkingssysteem (DSE systeem) ontworpen en gebouwd. In dit prototype is de ontwikkelde spatiële bandbreedtebeperkingstechniek toegepast. Het prototype-systeem bestaat uit een (gekromd) array van 192 luidsprekers, dat aangepast kan worden aan de geometrie van elke zaal. In het prototype zijn technische, economische en perceptieve criteria gecombineerd om tot een optimale oplossing te komen. Om de hoorbaarheid van kleuring ten gevolge van spatiële aliasing te verminderen, is er een alternatieve methode voor de verwerking van hoge frequenties onderzocht. Volgens deze methode worden de aankomsttijden van de hoogfrequente luidsprekerbijdragen gerandomiseerd.
Alhoewel dit onderzoek hoofdzakelijk gewijd is aan het weergave-deel van het systeem, zijn er ook verschillende technieken voor de geluidsregistratie en het volgen van geluidsbronnen voorgesteld.

Het DSE systeem is geëvalueerd in verschillende zalen onder verscheidene akoestische omstandigheden. Zowel fysische als perceptieve experimenten zijn uitgevoerd om het gedrag van het systeem te testen. In de fysische experimenten zijn echte en gesynthetiseerde geluidsvelden afkomstig van een puntbron spatieel gemeten en vergeleken. De resultaten tonen aan dat door middel van golfveldsynthese een nauwkeurige replica van het echte golfveld gegenereerd kan worden in een groot luistergebied.
In de psycho-akoestische experimenten zijn de belangrijkste subjectieve aspecten van echte en gesynthetiseerde golfvelden beoordeeld door verscheidene proefpersonen. Bronidentificatie-experimenten tonen aan dat het lokaliseren van virtuele geluidsbronnen zeer goed gaat op verschillende luisterposities in de zaal.

Subjectieve schalingsexperimenten hebben aangetoond dat, alhoewel de schijnbare breedte van een virtuele bron groter is dan die van een echte bron, bronverbreding door spatiële aliasing klein is in vergelijking met de bronverbreding door de akoestiek van de zaal.
Door middel van vergelijken van bewegende virtuele bronnen met bewegende echte bronnen is gevonden dat spatiële aliasing kan leiden tot kleuring van het waargenomen geluid. Het is echter aangetoond dat de kleuring vermindert door reflecties in een zaal; in vergelijking met reflectie-vrije ruimten ontkleuren galmende zalen het virtuele bronsignaal. Ervaring met verschillende zalen leert dat kleuring van spraaksignalen nauwelijks hoorbaar is. Met meer kleur-gevoelige signalen is gebleken dat de randomisatie van de hoogfrequente bijdragen van de luidsprekers een bruikbaar compromis is tussen de reductie van de hoorbaarheid van kleuring enerzijds en localisatie anderzijds.
Tenslotte besluiten we deze samevatting met te stellen dat door toepassing van golfveldsynthese in direct-geluidsversterking het mogelijk is om geluid van een bron te versterken op een natuurlijke manier. In tegenstelling tot de bestaande methoden lost golfveldsynthese het probleem van het verkeerd lokaliseren van versterkte geluidsbronnen op een fundamentele manier op door het golfveld van een virtuele bron te synthetiseren op dezelfde plaats als de echte bron. Op deze manier wordt er een nauwkeurige replica van het originele golfveld verkregen in het gehele luistergebied.