Stellingen

behorende bij het proefschrift

'On the design and psychophysical assessment of loudspeaker systems'

van

Ronald M. Aarts

1. Preferentie is een één-dimensionaal continuüm; derhalve is het niet zinvol om die meer-dimensionaal te schalen.

2. 'Gelijke appreciatie' wordt vaak verward met 'gelijkheid'.
[Dit proefschrift, Hoofdstuk 1.]

3. De Schroeder-toontjes hebben weliswaar een bijna optimale piekfactor ('crest factor'), doch zijn niet bruikbaar als een normaal verdeeld ruissignaal in psycho-akoestich onderzoek.
[M.R. Schroeder, IEEE Trans. on IT, Vol.16, p.85, 1970.]

4. Bij een elektrodynamische luidspreker, gemonteerd in een kleine kast, is het verkrijgen van een hoog rendement principieel strijdig met een goede basweergave.
[Dit proefschrift, Appendix A.]

5. Het gebruik van pulslengte-distributie-metingen als diagnostisch hulpmiddel bij de analyse van (magneto-)optische opslagsystemen is zeer nuttig.

6. Het genereren van zeer veel harmonische vervorming bij lage frequenties zonder dat daarbij veel intermodulatievervorming optreedt, komt de appreciatie van de basweergave ten goede.

7. Voor een digitale PID-regelaar verdient de parallelle topologie de voorkeur boven de gecascadeerde.

8. Door simultaan spectrale en temporele overgangen te detecteren en te combineren, is een betrouwbare automatische spraak/muziek-detector te construeren.

9. De A-weging als methode voor het schatten van het luidheidsniveau bij luidsprekerluisterproeven, is onbetrouwbaar.
[Dit proefschrift, Hoofdstuk 3.]

10. Voor vele sprekers geldt dat het produkt van volume en inhoud constant is.

11. De invloed van de interactie van geluid en beeld op de totale kwaliteitsbeleving wordt stelselmatig onderschat door beeldtechnici. De grootte van deze invloed kan men snel duidelijk maken door het geluid bij een TV-programma uit te schakelen.

12. De ontwikkelaars van (PC-)programmatuur zouden een significant groter deel van hun werkzaamheden behoren uit te voeren op een doorsnee huis-(tuin-en-keuken-)computer, in plaats van op een krachtige ontwikkelmachine.

13. Gezien de te bereiken luidheid zou een 'megafoon' beter 'hectofoon' kunnen heten.

14. De uitspraak 'zien is geloven' geldt sterk bij luidsprekerluisterproeven. In plaats van te kijken kan men echter beter luisteren.

15. Eerst was er geluid en daarna pas beeld.
[Genesis 1:3]

16. Het te pas en te onpas gebruiken van het woord 'perfect' duidt op niet-perfect taalgebruik.

17. Het is niet de mimetische begeerte, zoals Hobbes betoogt, maar lokale optimalisatie die de grondslag vormt van bijna alle ellende in de wereld.
[Thomas Hobbes, Leviathan 1651.]

18. Het zou voor de derde wereld een goede zaak zijn als de Wereldwinkel een 'wereldzaak' zou worden.

19. Het verdient aanbeveling om, net als bij golfwedstrijden, tijdens voetbalwedstrijden en andere rumoerige bijeenkomsten, bordjes met de tekst 'stilte/silence' te gaan gebruiken.

On the design and psychophysical assessment of loudspeaker systems

Ronald M. Aarts

The work described in this thesis has been carried out
at the PHILIPS RESEARCH LABORATORIES Eindhoven,
the Netherlands, as part of the Philips Research programme.

CIP-gegevens Koninklijke Bibliotheek, Den Haag
Aarts, R.M.
On the design and psychophysical assessment of
loudspeaker systems
Proefschrift Technische Universiteit Delft,-Met lit. opg.,
-Met samenvatting in het Nederlands.
ISBN 90-74445-18-7
Trefw.: luidsprekers, audio, geluidsreproduktie,
psycho-akoestiek

ON THE DESIGN AND PSYCHOPHYSICAL ASSESSMENT OF LOUDSPEAKER SYSTEMS

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus, Prof. ir. K.F. Wakker,
in het openbaar te verdedigen ten overstaan van een commissie,
door het College van Dekanen aangewezen,
op maandag 11 september 1995, te 13.30 uur
door

Ronaldus Maria AARTS

ingenieur elektrotechniek (HBO)
geboren te Amsterdam

Dit proefschrift is goedgekeurd door de promotor:

Prof.dr.ir. F.A. Bilsen

Samenstelling promotiecommissie:

Rector Magnificus, voorzitter
Prof.dr.ir. F.A. Bilsen, TU Delft, promotor
Prof.dr.ir. A.J. Berkhout, TU Delft
Prof.dr. A.J.M. Houtsma, IPO, Eindhoven
Prof.dr.ir. K.A. Schouhamer Immink, Philips Research Eindhoven
Prof.dr.ir. I.T. Young, TU Delft
dr.ir. M.M. Boone, TU Delft
dr.ir. W.F. Druyvesteyn, Philips Research Eindhoven.

Ter nagedachtenis aan mijn vader.
Voor mijn moeder, Doortje, Vincent & Wouter.

Sound, that Noble Accident of the Air...
Saggi... Accademia del Cimento
Waller translation

VERGILIUS ROMANUS - Pâtres gardant leurs troupeaux (Les Géorgiques, Livre III) - Première Moitié du V^e siècle - (H. 0,22; L. 0.225). Vatican, Bibliothèque Apostolique, Vat. Lat. 3867, Folio 44 verso.

Acknowledgements

This work has been performed at the Acoustics and Noise Control group of Philips Research Laboratories Eindhoven. Many people have contributed. In particular, I am indebted to the following persons:

Group leader dr. Erik Druyvesteyn, who in 1992 suggested writing this thesis and kept me enthusiastic, dr. Warner ten Kate who reviewed most of it and gave lots of constructive criticism.

Gillian Booles, Rachel Overton, Oliver Wright, Russell Richardson and Simon Webster, (former) students of the University of Salford, U.K., for performing experiments and giving many demonstrations.

Dr. Arie Kaizer, dr. Sjoerd Verduyn Lunel and former student Ronald Janmaat for valuable support.

My colleagues Robert Toonen Dekkers and Paul Boers for all kinds of support.

Ronald van Rijn for computer support, writing ps2d (a versatile plotting package) and for proof reading the manuscript from cover to cover.

The committee-members, in particular Prof.dr. A.J.M. Houtsma for thoroughly reviewing the manuscript.

Finally, I want to thank Prof.dr.ir. F.A. Bilsen, for his stimulating discussions and advice improving the coherence of the manuscript considerably.

Contents

Chapter 1

Introduction

1.1 Objectives and overview

Electroacoustics is as old and familiar as thunder and lightning[1], but the knowledge that is required to control such modes of energy conversions stems from the last two centuries. While the first patent for a moving-coil loudspeaker was filed in 1877, by Cuttriss and Redding [54], shortly after Bell's telephone invention [22], the real impetus to a commercial success was given by Rice and Kellog through their famous paper [161]. In defiance of the risk of yielding a somewhat kaleidoscopic view, this thesis aims to discuss the wide field of the design and assessment of loudspeaker[2] systems. In judging the quality of a loudspeaker, it is important to listen to it, either alongside other loudspeakers or by itself, using different parameters. Thereby, it is very important to ensure that all the systems being tested have the same loudness levels. In Chapter 2, two loudness measurement methods will be discussed and they are compared with simpler methods in Chapter 3, which will show that the most common method, the 'A-weighting' method, is not the best.

A particular sound power response of a loudspeaker may be more important than the traditional flat sound pressure level (SPL) design aim. Therefore, methods for measuring a loudspeaker's sound power response will be considered in Chapter 4. One driver alone cannot radiate sufficient acoustical power over a large frequency range. To overcome this problem, a

[1] A good introduction describing the position of (electro)acoustics in its proper historical setting is given by Hunt [105, 106] and others [15, 46, 169].

[2] 'Loudspeaker' is understood here in a wide sense. In general, a loudspeaker consists of one or more transducers, or drivers, a cabinet with one or more volumes and either active or passive electronic circuitry.

loudspeaker is usually built up using two or more drivers, each optimized for a restricted frequency range. An electronic network (crossover filter) is used to combine these drivers for the full audio band (usually 20 Hz–20 kHz). In Chapter 5, it will be discussed how these filters can be simulated in real-time, enabling the design engineer to select the best from a set of possible filters.

A method for investigating loudspeaker crossover filters in an analytical way will be presented in Chapter 6, and in Chapter 7 a new method for designing those filters will be presented, in which they are both used as crossover and equalizing filters, which enables extra design properties. As stated above, the sound power is to be considered very important, according to Staffeldt [183] even more important than the SPL response; the influence of the crossover filters on the sound power response will also be discussed in Chapter 4.

It is impossible to imagine sound reproduction today without stereophonic techniques, and it is to the credit of both the technology and human binaural hearing capabilities that a single pair of loudspeakers can evoke auditory perspectives so convincingly. It was in 1939, over 50 years after the first stereophonic demonstration at the Paris Opera [228], that the system was first taken into commercial use in Walt Disney's "Fantasia". Since then, a whole range of improvements to the system has been suggested, of which Blumlein's patent [29] is an early example. Various kinds of improvements for stereophonic sound reproduction will be discussed. The listening area for good stereophonic sound reproduction is limited, but can be enlarged using time-intensity trading, as will be described in Chapter 8.

For some stereophonic set-ups the necessary distance between the loudspeakers cannot be realized. Chapter 9 will discuss how this problem can be solved to some extend.

In some cases it is preferable to listen via headphones rather than via loudspeakers, but if one wishes to have the same perception as when listening via loudspeakers, specific signal processing is required, as will be discussed in Chapter 10.

1.2 The electrodynamic loudspeaker

An electrodynamic loudspeaker, of the kind depicted in Fig. 1.1 consists of a conical diaphragm, the cone, usually made of paper, being suspended by an outer suspension device, or rim, and an inner suspension device, or spider. The latter limits the maximum excursion of the cone at low frequencies so

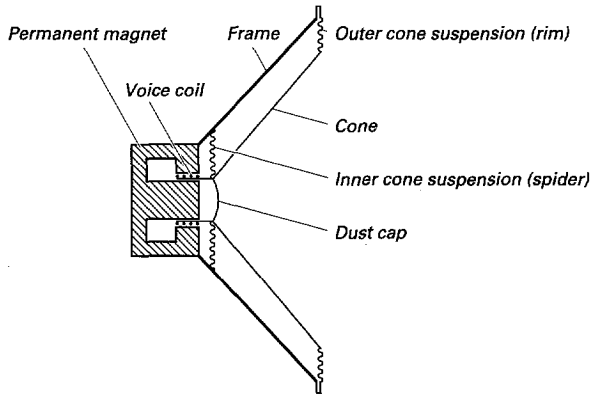

Figure 1.1: Cross-section of an electrodynamic cone loudspeaker.

that the voice coil remains inside the air gap of the permanent magnet, which can lead to non-linear distortion; see e.g. [113, 118, 120, 147, 194]. The voice coil is attached to the voice coil cylinder, generally made of paper or metal, which is glued to the inner edge of the cone. In most cases the spider is also attached to this edge. The voice coil is placed in the radial magnetic field of a permanent magnet and is fed with the signal current of the amplifier. For low frequencies the driver can be modeled in a relatively simple way, as it behaves as a rigid piston. This will be summarized in the next section, in which the electronic behavior of the driver will be described on the basis of a lumped-element model in which the mechanical and acoustical elements can be interpreted in terms of the well-known properties of their analogous electronic-network counterparts. At higher frequencies deviations from this model occur, as the driver's diaphragm is then no longer rigid. Both transverse and longitudinal waves then appear in the conical shell. These waves are coupled and together they determine the vibration pattern, which has a considerable effect on the sound radiation. Although this is an important issue, it will not be considered here (see e.g. [73, 113, 207]).

1.2.1 The lumped-element model

For low frequencies a loudspeaker can be modeled with the aid of some simple elements, allowing the formulation of some approximate analytical expressions for the loudspeaker sound radiation due to an electrical input current, or voltage, which proves to be quite satisfactory for frequencies below the cone break-up frequency. The model will not be derived here

Table 1.1: System parameters of the model of Fig. 1.2.

R_e	electrical resistance of the voice coil
L_e	inductance of the voice coil
I	voice coil current
U	voltage induced in the voice coil
B	flux density in the air gap
l	effective length of the voice coil wire
F	Lorentz force acting on the voice coil
V	velocity of the voice coil
k_t	total spring constant
m_t	total moving mass, without air load mass
R_m	mechanical damping
Z_{rad}	mechanical radiation impedance $= R_r + jX_r$

as that has been discussed extensively elsewhere; see e.g. [23, 113, 135, 146]. When use is made of the model shown in Fig. 1.2, which behaves as a simple second-order mass-spring system, then the following formula applies:

$$F = (R_m + j\omega m_t + \frac{k_t}{j\omega} + Z_{rad})V; \tag{1.1}$$

the system parameters are presented in Table 1.1. With the aid of Eq. 1.1

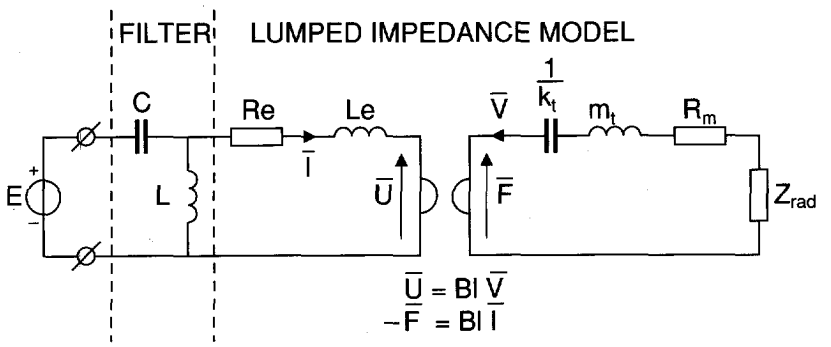

Figure 1.2: Lumped-element model of the impedance-type analogy of an electrodynamic loudspeaker preceded by an LC high-pass filter. The coupling between the electrical and mechanical parts is represented by a gyrator. The system parameters are given in Table 1.1.

and the properties of the gyrator shown in Fig. 1.2 the electrical impedance

of the loudspeaker, without[3] X_r, can be calculated as follows[4]

$$Z_{in} = R_e + j\omega L_e + \frac{(Bl)^2}{(R_m + R_r) + j\omega m_t + k_t/(j\omega)}. \tag{1.2}$$

Using the following relations

$$\begin{array}{llll} Q_m = \sqrt{k_t m_t}/R_m\ , & & Q_e = R_e\sqrt{k_t m_t}/(Bl)^2\ , \\ Q_r = \sqrt{k_t m_t}/R_r\ , & & \omega_0 = \sqrt{k_t/m_t}\ , \\ \nu = \omega/\omega_0 - \omega_0/\omega\ , & & \tau_e = L_e/R_e\ , \\ Q_{mr} = Q_m Q_r/(Q_m + Q_r)\ , & & \end{array} \tag{1.3}$$

we can write Z_{in} as

$$Z_{in} = R_e[1 + j\omega\tau_e + \frac{Q_{mr}/Q_e}{1 + jQ_{mr}\nu}]. \tag{1.4}$$

The time-averaged electrical power delivered to the driver is then

$$P_e = 0.5|I|^2\Re\{Z_{in}\} = 0.5|I|^2 R_e[1 + \frac{Q_{mr}/Q_e}{1 + Q_{mr}^2\nu^2}]. \tag{1.5}$$

The radiation impedance of a plane circular rigid piston[5] with a radius a in an infinite baffle can be derived as [143, p. 384]

$$Z_{rad} = \pi a^2\rho c[1 - 2J_1(2ka)/(2ka) + j2\mathbf{H}_1(2ka)/(2ka)], \tag{1.6}$$

where $\mathbf{H}_1$ is a Struve function [13, §12.1.7], J_1 is a Bessel function and k is the wavenumber ω/c. The real and imaginary parts of Z_{rad} are plotted in Fig. 1.3. Z_{rad} can be approximated as

$$Z_{rad} \approx \begin{cases} \pi a^2\rho c[(ka)^2/2 + j8ka/(3\pi)]\ , & \omega \ll \omega_t \\ \pi a^2\rho c[1 + j2/(\pi ka)]\ , & \omega \gg \omega_t\ , \end{cases} \tag{1.7}$$

where

$$\omega_t = 1.4c/a \tag{1.8}$$

[3] For $\omega \ll \omega_t$ (defined in Eq. 1.8) X_r/ω can be taken into account in m_t.

[4] Due to the eddy current losses in the voice coil, the voice coil does not behave as an ideal coil, but it can be modeled very well by means of $L_e = L_0(1 - j\alpha)$, where α is in the order of magnitude of 0.5.

[5] The radiation impedance of rigid cones and that of rigid domes is studied in e.g. [191, 192]. They appeared to be significantly different with respect to rigid pistons for $ka > 1$, revealing that Z_{rad} for convex domes is generally lower than that for pistons and higher than that for concave domes.

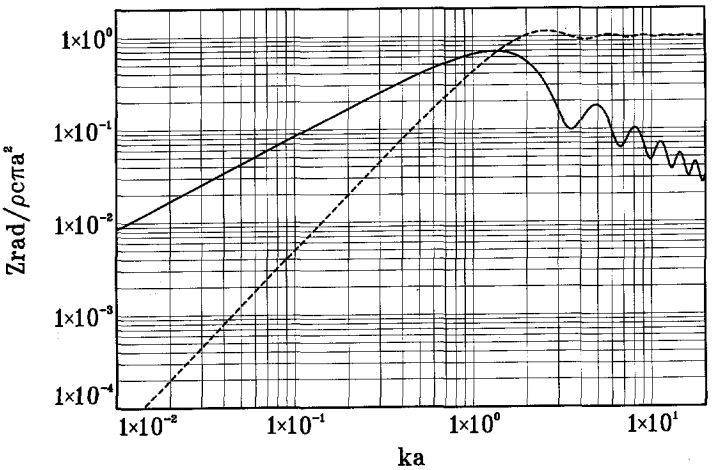

Figure 1.3: Real (dashed line) and imaginary (solid line) parts of the normalized radiation impedance of a rigid disk with a radius a in an infinite baffle.

is the transition frequency (-3 dB point). The real part of Z_{rad} is depicted in Fig. 1.4. The time-averaged acoustically radiated power can then be calculated as follows

$$P_a = 0.5|V|^2 \Re\{Z_{rad}\}, \tag{1.9}$$

and with the aid of Eqs. 1.1 - 1.9 as follows

$$P_a = \frac{0.5(Bl/(R_m + R_r))^2 I^2 R_r}{1 + Q_{mr}^2 \nu^2}, \tag{1.10}$$

as depicted in Fig. 1.4. The acoustic pressure in the far field at (r, θ) is

$$p(r,t) = jf\rho_0 V \pi a^2 / r \left[\frac{J_1(ka\sin\theta)}{ka\sin\theta}\right] e^{j\omega(t-r/c)}, \tag{1.11}$$

on the assumption of an axis of symmetry, where V is the velocity of the piston [23, 115]. Assuming a velocity profile as depicted in Fig. 1.4, we can calculate the magnitude of the on-axis response ($\theta = 0$). This is depicted in Fig. 1.4, which shows a 'flat' SPL for $\omega \gg \omega_0$. However, due to the term in square brackets in Eq. 1.11, the off-axis pressure response ($\theta \neq 0$) decreases with an increasing ka. This yields an upper frequency limit, together with the mechanical lower frequency limit ω_0, and is the reason why a practical loudspeaker system needs more than one driver (a multi-way system) to handle the whole frequency range. Using Eq. 1.11, we can calculate the

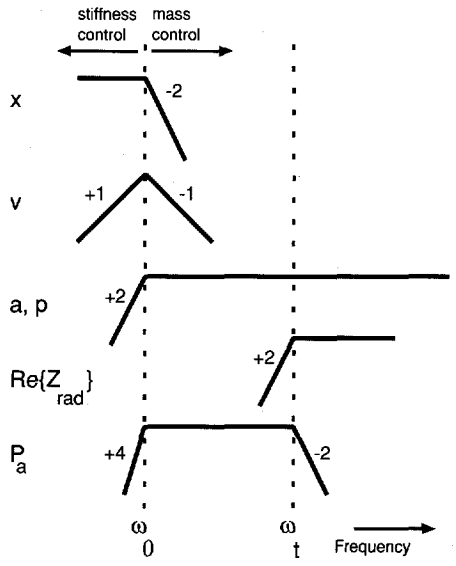

Figure 1.4: The displacement x, velocity v, acceleration a, on-axis pressure p, the real part of the radiation impedance $\Re\{Z_{rad}\}$ and the acoustical power P_a of a rigid-plane piston in an infinite baffle. The numbers denote the slopes of the curves.

directivity D as follows

$$D = \frac{(ka)^2}{1 - J_1(2ka)/ka}. \tag{1.12}$$

This is plotted in Fig. 1.5, which clearly shows the narrowing of the beam width for $ka > 1$. The efficiency can be calculated as follows

$$\eta(\nu) = P_a/P_e = [Q_e Q_r(\nu^2 + 1/Q_{mr}^2) + Q_r/Q_{mr}]^{-1}, \tag{1.13}$$

or for low frequencies[6] so that $Q_{mr} \approx Q_m$

$$\eta(\nu) = P_a/P_e \approx [Q_r \left\{ Q_e(\nu^2 + 1/Q_m^2) + 1/Q_m \right\}]^{-1}. \tag{1.14}$$

Some practical values for the efficiency and the lumped-element parameters of some low-frequency loudspeakers are given in Appendix A. While the directivity of a rigid piston increases for $ka > 1$, a non-rigid single driver can be used over a much wider frequency band, because the central portions

[6]It should be noted that Q_r depends on ω, using Eq. 1.3 and Eq. 1.7 for $\omega \ll \omega_t$, Q_r can be approximated as $Q_r \approx 2c\sqrt{k_t m_t}/(\pi a^4 \rho \omega^2)$.

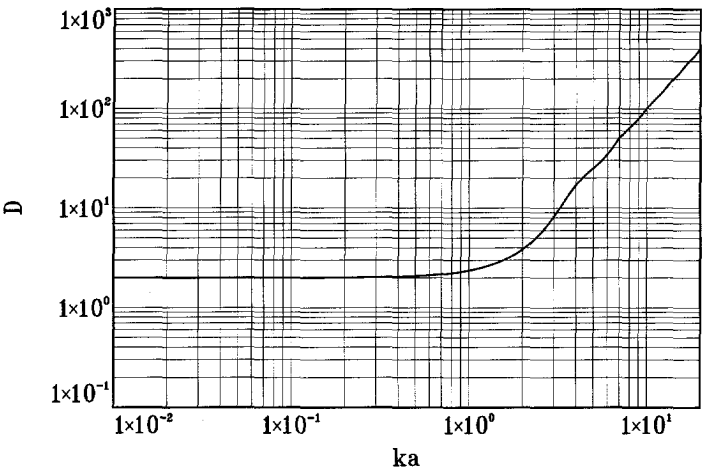

Figure 1.5: Directivity for a rigid-plane piston in an infinite baffle.

of a conical diaphragm are stiffer and more tightly coupled to the driving voice coil than the outer, flatter portions of the conical surface. It is normal behavior, therefore, for the outer zones of a conical diaphragm to become progressively decoupled with increasing frequency, leaving the stiffer central portion to carry the burden of sound radiation at higher frequencies. This behavior is symbolized by the model depicted in Fig. 1.6. However, it remains beneficial to control this crossing to a smaller a by means of electronic filtering (crossover filters) and by using separate drivers, each optimized for its own frequency band. The final crossing from one band to the other is very important for the total system behavior and will largely determine the sound radiation in its environment and, consequently, the sound assessment. The importance of crossover filters is depicted in Fig. 1.7.

1.3 Perception of loudspeakers

In many disciplines subjective assessment is more of an art than a science. However, subjective assessment is an important part of an acoustician's work, especially in such tasks assessing room acoustics or the design of loudspeakers. For good sound reproduction, a loudspeaker must not only meet the traditional design requirements, but must also be tuned to the desired performance, which implies subjective measurements. A general review of listening tests for loudspeakers is given in [197], a more thorough analysis of the variables being presented in [76, 199]. The relation-

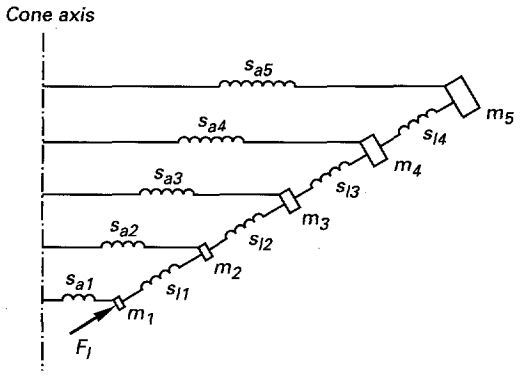

Figure 1.6: Lumped-element model of an electrodynamic loudspeaker (Fig. 4.3 from [73]). The cone is represented by masses m and springs with stiffness s_l in the longitudinal direction and s_a in the azimuthal direction. The values of m increase with their distance from the center, while the values of s_a decrease.

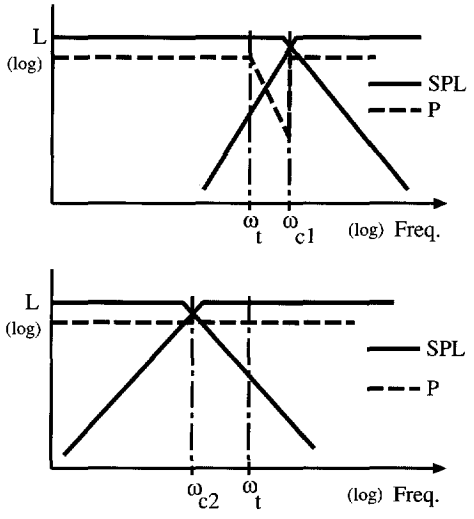

Figure 1.7: Influence on the total power response P, of the different crossover frequencies ω_{c1} and ω_{c2}. While the on-axis response for two pistons will add to a flat SPL, the total radiated power starts to decrease at frequency ω_t, the cut-off frequency of the largest piston, if $\omega_t < \omega_{c1}$, as depicted in the upper panel.

ship between objective and subjective loudspeaker measurements is studied in e.g. [35, 79–81, 183, 193, 200, 201]. In general, a loudspeaker affects the timbre of a sound, and since timbre is a multi-dimensional and subjective attribute (see e.g. [154, 155]) the ultimate test is to judge a loudspeaker's

quality, preferably by comparing it with that of other systems, by listening to it. When comparing a loudspeaker with other systems it is important that all the systems have the same loudness. It is known that, to a certain extent, the quality increases with increasing loudness; see e.g. [107]. For that reason attention will be paid to this topic in Chapters 2 and 3. In the following section some objective measures will be discussed, which can affect the perception of a loudspeaker's sound.

1.3.1 SPL and power response

The traditional objective measure is the on-axis free-field sound pressure measurement. For several reasons this is not an adequate measure for characterizing a loudspeaker.

First of all, a loudspeaker is, under normal conditions, not used in an anechoic environment. Sound radiated in directions other than that of the main-axis arrives at the ears via room boundary reflections. Therefore, acoustic power is an important parameter, as will be discussed in greater detail in Chapter 4. Shorter [179] suggested that, if a single quantity representing 'effective' response is to be found at all, it will lie somewhere between the axial and the mean spherical response. Staffeld [183] recognized that the power response shows a resemblance to listening-room responses. This was elaborated by Tannaka and Koshikawa [193], who defined a partial-space-power response. Bech [20] has shown that the room and the loudspeaker position will influence the absolute level of fidelity of the reproduced sound as well as relative differences between different loudspeakers or different positions.

Secondly, due to different polar patterns, the system pressure response can be equalized to another (hypothetical) system for one position only and will consequently be perceived differently at other listening positions, even due to (modest) head movements. This problem can be relaxed in the case of a multi-way loudspeaker system by equalizing drivers individually rather than the total system, which allows extra freedom. This 'per-driver' equalization can be integrated in the crossover filters, as will be discussed in Chapter 7.

Thirdly, due to the finite dimensions of the drivers and the finite distance between them, different path lengths to the ears occur. Since different frequency regions are operated by different drivers. Even worse, at the crossover frequency two drivers radiate the same SPL. This can produce spurious effects such as instability of the image. For a stereo set-up this is even more important, because a phantom image is less robust than a real

source.

Finally, allowance must be made for loudspeaker room interactions. Irregularities in the frequency response may be emphasized due to room resonances, giving the music a colored sound.

The frequency response generally has to cover the 20 Hz–20 kHz range, although most people cannot detect the effect of low-pass filtering speech or music at 16 kHz [145]. Even if they can, the low-pass filtering does not result in any degradation of sound quality [140]. It was found [116] that a group of subjects who were exposed to a restricted frequency range sound reproduction system for a long period of time, preferred that above the full frequency range system.

If the frequency response has peaks and dips then these can affect the reproduced sound in two ways.

First of all, the timbre of the reproduced sound may be altered, which causes coloration. Several psychoacoustic studies have been devoted to this topic [38, 90, 142, 203]. These studies indicate that, under ideal conditions, subjects can detect changes in spectral shape when the level in a given frequency region is increased or decreased by 1–2 dB, relative to the level in other frequency regions. The detection of changes in spectral shape is partly limited by the frequency selectivity of the auditory system. According to Moore and Glasberg [141] the equivalent rectangular bandwidth of the auditory filters (ERB) is

$$ERB = 24.7(4.37F + 1), \tag{1.15}$$

where F is expressed in kHz, but the ERB in Hz. At higher frequencies the auditory system is less sensitive due to the increasing ERB, since irregularities are smoothed by the auditory filters. Discriminating a change in spectral shape by simultaneous comparisons of the intensity level at different parts of the spectrum, known as profile analysis, can be an important contributor to detection performance [90].

Secondly, due to irregularities in the frequency response, loudness fluctuations may occur.

1.3.2 Phase response

Another aspect of loudspeaker performance is its phase response. There is a vast amount of literature on this topic [19, 24, 25, 27, 28, 41–43, 47, 60, 65, 71, 87, 97, 98, 124, 149, 155, 163]; for historical notes see [87, 155]. The statement that the ear is *phase deaf* is often ascribed to Helmholtz [101], but his formulation was actually more subtle. He used tuning-forks to

obtain a tone complex of eight tones and left the question as to whether the phase difference is of importance for other sounds unanswered. The linear phase rule was formulated by de Boer [60]

$$\Delta\varphi(f) = a + bf, \tag{1.16}$$

where φ is the phase of the component with respect to frequency, f. It says that the timbre of a sound does not change when the phases of the components are shifted by a constant amount or by amounts that are linearly dependent on frequency. Buunen et al. [43] showed that de Boer's phase rule can be explained by the mutual interaction of acoustic components and combination tones evoked in the auditory system. Craig and Jeffress [52] noted that this rule can be violated and Rosen [163] showed that it must be viewed in a more stringent context.

Other studies, e.g. [91, 97, 98, 189], have shown that one can even perceive a 180° phase shift, known as the Wood effect [52].

Another way of describing phase differences is by the group delay

$$\tau_g = -\frac{d\varphi(\omega)}{d\omega}, \tag{1.17}$$

or its deviation from a constant value, the group delay distortion

$$\Delta\tau_g = \tau_g - T, \tag{1.18}$$

where T is the frequency-independent delay in the system. Investigations by Blauert and Laws [27] showed that measured distortions can approach the magnitude of the threshold of perceptibility, but will in most cases be below this value. However, it was noted in [71] that group delay distortion at low frequencies can produce subtle, but clearly audible, changes in sound quality.

One may conclude that the audibility of phase distortion is very dependent on the stimuli, but in *real* music it is very difficult to detect.

1.3.3 Non-linear distortion

A loudspeaker shows minor non-linearities that produce typical distortion phenomena. A literature overview pertaining to the perception of non-linear distortion is given by Cabot [44]. Possible non-linearities may be found in various parts of the loudspeaker, the most prominent being

- the force factor Bl, which depends on the voice coil excursion

- the non-linear (suspension) stiffness k_t
- the limited excursion capability of the moving parts
- 'Doppler'-distortion: a weak modulation of the high frequencies due to finite cone excursions at a low frequency, which occurs when two tones at a low and a high frequency are presented simultaneously.

So far, many attempts have been made to reduce these non-linearities by electrical means; see e.g. [114, 117, 119]. At high drive levels these non-linearities can cause audible distortion that may be

- harmonic distortion. This is usually specified in terms of the extent to which frequency components are generated (and thus present at the output) that are not present at the input. Distortion is usually presented as a percentage of the total harmonic distortion of the loudspeaker with respect to its input, being a single sinusoid. This measure allows for an estimate of its audibility. If the distortion is concentrated at lower harmonics it will not be easily detectable. However, if the harmonics introduced are not easily masked, then the distortion should be kept below about 0.1% [140].
- intermodulation distortion. When two tones with frequencies f_1 and f_2 are applied to the loudspeaker, frequencies $mf_1 \pm nf_2$ are generated due to non-linearities. In general, intermodulation distortion is somewhat better audible than harmonic distortion, but values of less than 0.5% are unlikely to be detected.
- Doppler distortion. This is a type of distortion that occurs because low frequencies with high cone excursions simultaneously modulate high frequencies, both radiated by the same loudspeaker. This type of distortion is studied in e.g. [21, 40, 137, 209, 212]. Although it is considered to be of theoretical interest only, it may be of importance for modern loudspeaker systems. If a small cabinet with a high output is requested, then a long-stroke woofer is necessary. The distortion factor, defined as the square root of the ratio of the sideband power to the total power is given by [21],

$$d.f. = \sqrt{1 - [J_0(2\pi A_1/\lambda_2)]^2} \tag{1.19}$$

where λ_2, the wavelength of the modulated frequency, is approximately proportional to A_1, the amplitude of the cone excursion. For high cone excursions one may then expect audible distortion.

1.4 Subjective quality assessment

When a subject is asked to rank loudspeakers, using either an implicit or an explicit method, from poor to good, on a bipolar one-dimensional scale, it is hypothesized that the following rule will be used. The appreciation is the weighted sum of several attributes

$$Ap_j = \sum_i^n W_{ij} A_{ij}, \tag{1.20}$$

where W_{ij} is the weighting for the i^{th} attribute A_{ij} of the j^{th} subject. The attributes may be clarity, fullness, spaciousness, etc. Although the A_{ij} are, in general, elements of a multi-dimensional continuum, Ap_j is by definition a *one*-dimensional outcome. This quality assessment can be performed for example via paired comparisons [58], in which all pairs, or a subset of pairs, as discussed in Appendix B, are presented to the subjects. One element of each pair is judged to be the best. A frequency matrix is then constructed from the outcomes of the individual pairs. With the aid of scaling techniques, e.g. Thurnstone-Torgerson scaling [92, 196, 204], a *one*-dimensional scale can be constructed. However, because of variance in judgements, due to such factors as

- differences in the physical background noise
- differences in the physiological background noise (e.g. heart beat, breathing)
- differences in the psychological background noise (e.g. alertness, concentration)
- changes in the weights of Eq. 1.20,

it may be impossible to scale the preference matrix one-dimensionally. However, this does not justify the use of a multi-dimensional scaling technique, such as that discussed in Appendix C.2. But, if the subjects were to be asked to assess the similarity of the loudspeakers, e.g. by using the method of triadic comparisons discussed in Appendix C, then it would be possible to scale the resulting similarity matrix with the aid of a multi-dimensional scaling technique, since the timbre of the loudspeakers would differ, this being a multi-dimensional aspect of sound. This may be illustrated by the following example. Suppose there are ten stimuli, each having four attributes (arbitrary units), as shown in Table 1.2. For the sake of

convenience, W_{ij} are assumed to be unity. If the subject uses Eq. 1.20 to determine the total appreciation Ap, the result will be as shown in Table 1.2 and Fig. 1.8. However, if the subject is asked to assess the similarity of the

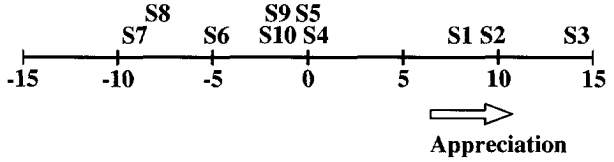

Figure 1.8: Appreciation of the stimuli given in Table 1.2.

Table 1.2: Four attributes belonging to ten stimuli and their total appreciations.

	A_4	A_3	A_2	A_1	Ap
S1	2.4	2.0	1.7	1.9	8.0
S2	-0.7	3.1	5.5	1.8	9.7
S3	0.2	4.3	6.8	2.8	14.1
S4	-2.1	0.4	2.2	-0.1	0.4
S5	0.0	0.0	0.0	0.0	0.0
S6	-0.4	-1.4	-2.0	-1.0	-4.8
S7	0.2	-2.8	-4.7	-1.8	-9.1
S8	-8.5	-1.0	4.1	-2.5	-7.9
S9	5.0	-1.3	-5.5	0.2	-1.6
S10	-2.9	0.0	1.9	-0.6	-1.6

stimuli he will probably use the following 'rule': the distance D_{kl} between two stimuli S_k and S_l is the absolute difference between their corresponding attributes

$$D_{kl} = \{\sum_{i}^{n} W_i |A_i(S_k) - A_i(S_l)|^p\}^{1/p} \tag{1.21}$$

where $A_i(S_k)$ denotes the i^{th} attribute of the k^{th} stimulus, p is Minkowski's parameter (in this example $p = 2$ is used) and W_i is the i^{th} weighting term. With the aid of Eq. 1.21 and the data from Table 1.2 a distance matrix can then be constructed, which can be used as input for a multi-dimensional scaling program. The output of the program is shown in Fig. 1.9. Stimuli S_9 and S_{10} have the same appreciations ($Ap = -1.6$), but are very dissimilar, as can be seen in Figs. 1.8 and 1.9.

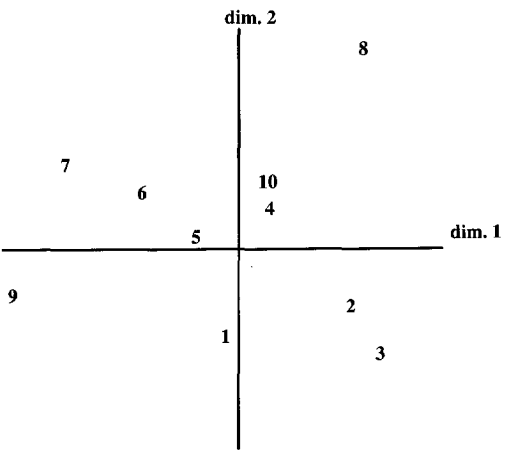

Figure 1.9: Perceptual configuration; see Table 1.2.

Chapter 2

Calculation of loudness for loudspeaker listening tests [1]

A computer program is used to calculate loudness levels in order to determine the reproduction levels of loudspeakers during listening tests. The calculation results are compared with the results of listening tests. Pink noise and music, reproduced by various loudspeakers, are used as the stimuli. Subjective listening tests show that the A-weighting, which is currently the defacto standard, can result in misleading conclusions. The use of the ISO 532B loudness model appears to be the best technique among the unweighted, A-weighted and ISO 532A methods, for adjustment of the inter-loudness levels of loudspeakers during listening tests.

2.1 Introduction

The purpose of this chapter is twofold. First, after an introduction to loudness and the history of scaling of loudness, a brief description is given of the calculation of loudness of stationary signals, written according to ISO standards [229]. A comparison is made between the program results, the traditional A-weighted sum using several noise spectra, and the subjective loudness ratings known from the literature.
Second, an application is discussed concerning the objective measurement of reproduction levels of loudspeakers during listening tests. It is considered to be very important that during listening tests the reproduction levels of the loudspeakers being tested are the same, otherwise the test results can

[1] Parts of this chapter have been published in [5].

be seriously biased. To validate the calculation results, they are compared with several subjective ratings obtained from listening tests, using a varied repertoire and a variety of loudspeakers. From this the conclusion is drawn that the A-weighting method is in general too simple, and may lead to misleading loudspeaker quality assessments.

2.1.1 Definitions

The following definitions are used [235].

Loudness: That attribute of auditory sensation in terms of which sounds may be ordered on a scale extending from soft to loud. Loudness is expressed in sone.

Loudness Level: Of a given sound, the sound pressure level of a reference sound, consisting of a sinusoidal plane progressive wave of frequency 1 kHz coming from directly in front of the listener, which is judged by otologically normal persons to be equally loud to the given sound. Loudness level is expressed in phon.

Critical bandwidth: The widest frequency band within which the loudness of a band of continuously distributed random noise of constant band sound pressure level is independent of its bandwidth.

2.1.2 Scaling of loudness

Several experimenters have made contributions to the scaling of loudness. The earliest published work seems to be that credited to Richardson and Ross [162], who required an observer to rate one of two tones of different intensities which he heard as a certain multiple or fraction of the other. Over the past five decades various methods of evaluating loudness of complex sounds from objective spectrum analysis have been proposed. The earliest attempt to use algebraic models in psychophysical measurement is probably that of Fletcher and Munson [72]. However, there is still much interest in this subject, and nowadays two procedures are used for calculating loudness levels; both are standardized by the ISO. The first is based on a method developed by Stevens, hereafter referred to as method 532A and the second one, 532B, by Zwicker [224, 226]. A-weighting is a widely used method, traditionally applied in sound level meters to measure the loudness of signals or to determine the annoyance of noise. It is based on an early 40 phon contour and is a rough approximation of a frequency weighting of

the human auditory system. However, considerable differences are ascertained between subjective loudness ratings and the A-weighted measurements. The loudness of noise increases when the dBA level is kept constant and the bandwidth of the noise is increased. As depicted in Fig. 2.1, with

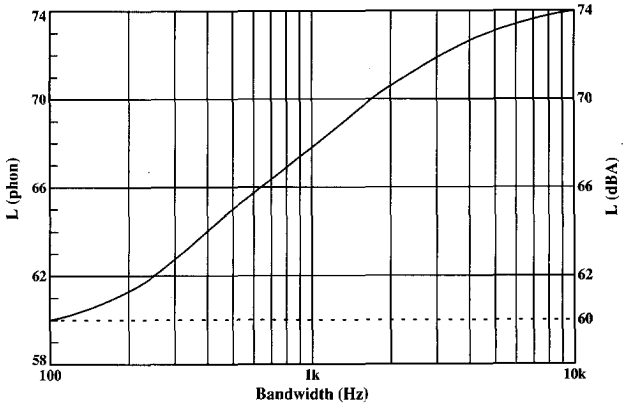

Figure 2.1: Increasing loudness (solid curve) vs. bandwidth of white noise while the dBA Level (dashed line) was kept constant.

increasing bandwidth the loudness has increased from 60 to 74 phon (solid curve), while the dBA level (dashed line) was kept constant. The effect has been studied by Brittain [34], and is a striking example that for wide-band signals the A-weighted method is generally too simple.

2.2 Methods of loudness calculation

A computer program has been written according to both Stevens' and Zwicker's method as specified in the ISO 532 standard: 'Acoustics - method for calculating loudness level' [229]. The input for both programs is a file containing SPL values measured in one third-octave bands. The essential parts of both methods is briefly discussed in the following sections.

2.2.1 The sone scale

The loudness level is expressed in phon. However, loudness values expressed on this scale do not immediately suggest the actual magnitude of the sensation. Therefore the sone scale, which is the numerical assignment of the strength of a sound, has been established. It has been obtained by subjective magnitude estimation by observers with normal hearing.

As a result of numerous experiments [166], the following equation has evolved to calculate the loudness S of a 1 kHz tone in sone:

$$S = 0.01(p - p_0)^{0.6} \tag{2.1}$$

where $p_0 = 45$ μPa approximates the effective threshold and p is the sound pressure in μPa. For values $p \gg p_0$, equation 2.1 can be approximated by the well-known equation

$$S = 2^{(P-40)/10} \tag{2.2}$$

or

$$P = 40 + 10 \log_2 S \tag{2.3}$$

where P is the loudness in phon, leading to the ISO 532 definition: 'One sone is the loudness of a sound whose loudness level is 40 phon'. It is important to discriminate between loudness levels that have been computed and those obtained by subjective measurements. In the following, the computed methods will be labeled by the suffix (OD) for Stevens' method, and by (GD) or (GF) for Zwicker's method, where O stands for measurements being made in octave bands, G for group (Frequenzgruppen or critical bands). D stands for diffuse and F for free-field, depending on the measurement environment of the sound source. Method 532A (Stevens) assumes that the measurements are taken in a diffuse sound field, while Method 532B (Zwicker) allows diffuse and free-field conditions to be set as a parameter in the program.

2.2.2 Stevens' method: ISO 532A

The method is equal to the Mark VI version as described in [187]. However, Stevens refined the method resulting in the Mark VII version [188] which is not standardized. In the following the 532A method is discussed briefly. The SPL of each one-third octave is converted into a loudness index using a table based on subjective measurements. The total loudness in sone(OD) S_t is calculated by means of the equation

$$S_t = S_m + F(\sum S_i - S_m) \tag{2.4}$$

where S_m is the greatest of the loudness indices and ΣS_i is the sum of the loudness indices of all the bands. For one-third octave bands the value of F is 0.15, 0.2 for one-half octave bands and 0.3 for octave bands. The total loudness may be converted into loudness level in phon(OD) using equation 2.3.

2.2.3 Zwicker's method: ISO 532B

An early version is described in [224] and later it has been refined, see e.g. [150, 225, 227]. In the following the essential steps in the procedure are discussed. The spectrum measured in one-third octaves is converted into frequency bands having a bandwidth approximating the critical bands of the human ear. This warping of the frequency axis into the critical band rate, with units of Bark, is depicted in Fig. 2.2. The one-third octave bands

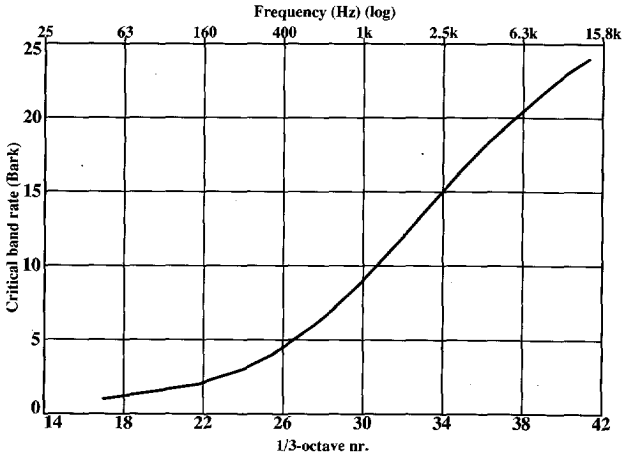

Figure 2.2: The relationship between frequency, one-third octave number and critical band rate.

up to 90 Hz are assembled as the first critical band, the next three one-third octaves (90–180 Hz) as the second band and the bands ranging from 180 Hz up to 280 Hz as the third critical band, leading to the critical band levels L_{c1}, L_{c2} and L_{c3}. The rule of combination may be understood from the example.

$$L_{c2} = 10\log(10^{L_{100}/10} + 10^{L_{125}/10} + 10^{L_{160}/10}) \tag{2.5}$$

where L_{125} etc., is the measured one-third octave band SPL for the band with center frequency 125 Hz, and L_{c2} is the level in the second critical band. The relation between the frequency f_i and the one-third octave number i is given by

$$f_i = 10^{i/10}. \tag{2.6}$$

Each critical band is subdivided into bands of 0.1 Bark wide. The SPL in each critical band is converted, by means of a table, into a loudness index for each of its sub-bands. In order to incorporate masking effects, contributions

are also made to higher bands. As an example the transformation is plotted in Fig. 2.3 for a 1 kHz tone at 90 dB SPL. Clearly, there is a 'tail' after

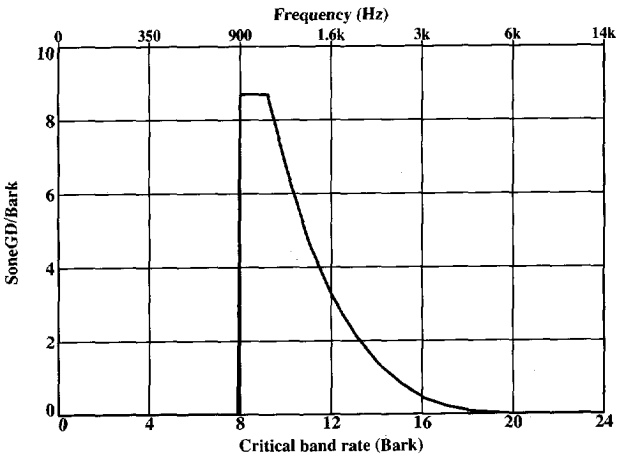

Figure 2.3: The transformation from SPL into loudness for a 1 kHz tone at 90 dB.

the actual excitation. When there are more tones in the spectrum it is possible that there is an overlap between the various tails. However, only the highest loudness index in each 0.1 Bark band makes a contribution to the overall result, and as a consequence the weaker tones are masked. The total loudness is finally calculated by integrating the loudness indices over all the sub-bands, resulting in the loudness in sone(GF) or sone(GD).

2.3 Loudness program

To gain insight into the methods used, some standard noise spectra will be used and comparison is made between program results, the traditional A-weighted sum and the subjective loudness ratings known from the literature. The results of this are summarized in Table 2.1 (on page 26). The table has the following vertical entries:

- Type of noise spectrum.
- Noise level in the one-third octave band of 400 Hz.
- Total SPL of the spectrum, i.e. the value obtained by simply adding the sound pressure in all the third-octaves.
- Total A-weighted level of the noise spectrum, i.e. the value obtained by using the A-weighting formula [1].

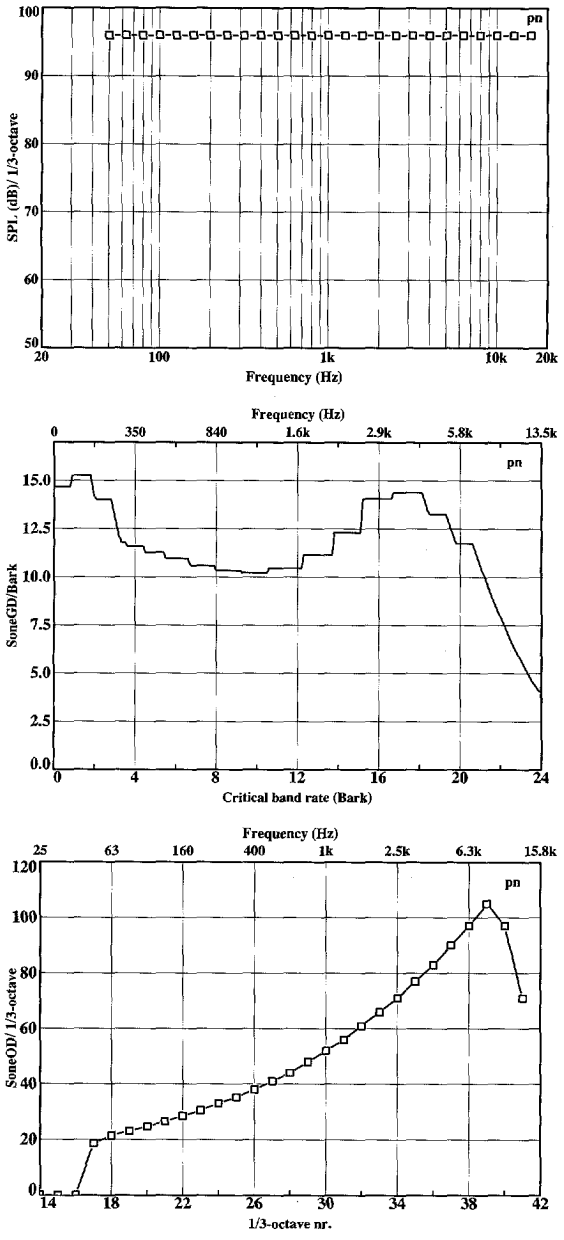

Figure 2.4: Upper panel: (a) Pink noise spectrum at 96 dB SPL/one-third octave. (b) Pink noise loudness spectrum according to 532B sone(GD)/Bark. Lower panel: (c) Pink noise loudness spectrum according to 532A sone(OD)/one-third octave.

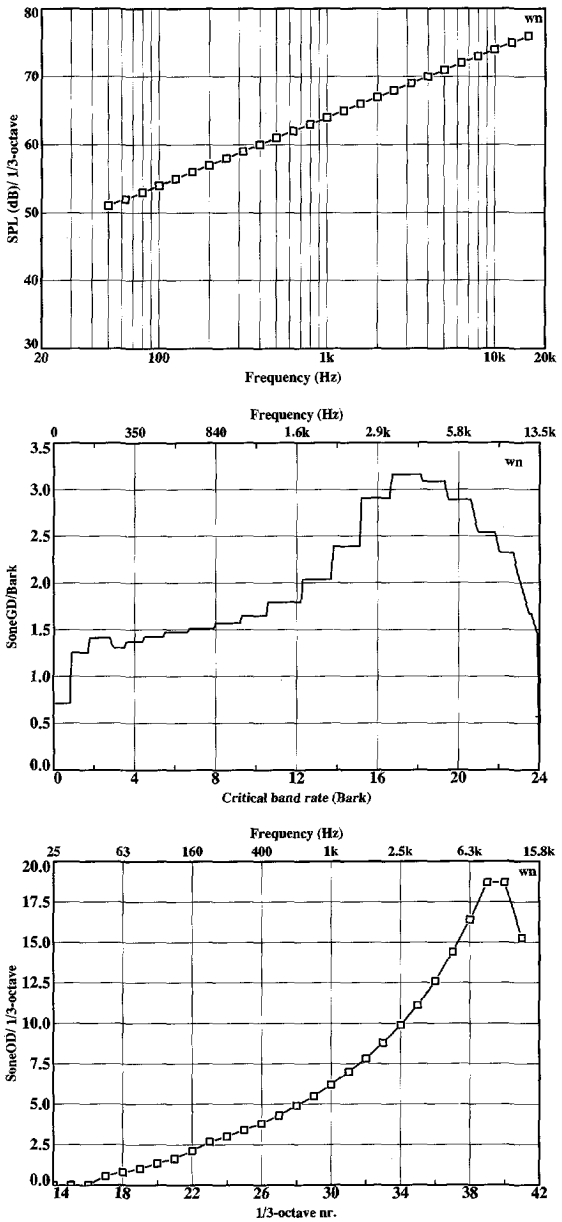

Figure 2.5: Upper panel: (a) White noise spectrum at 81.2 dBA level. (b) White noise loudness spectrum according to 532B sone(GD)/Bark. Lower panel: (c) White noise loudness spectrum according to 532A sone(OD)/one-third octave.

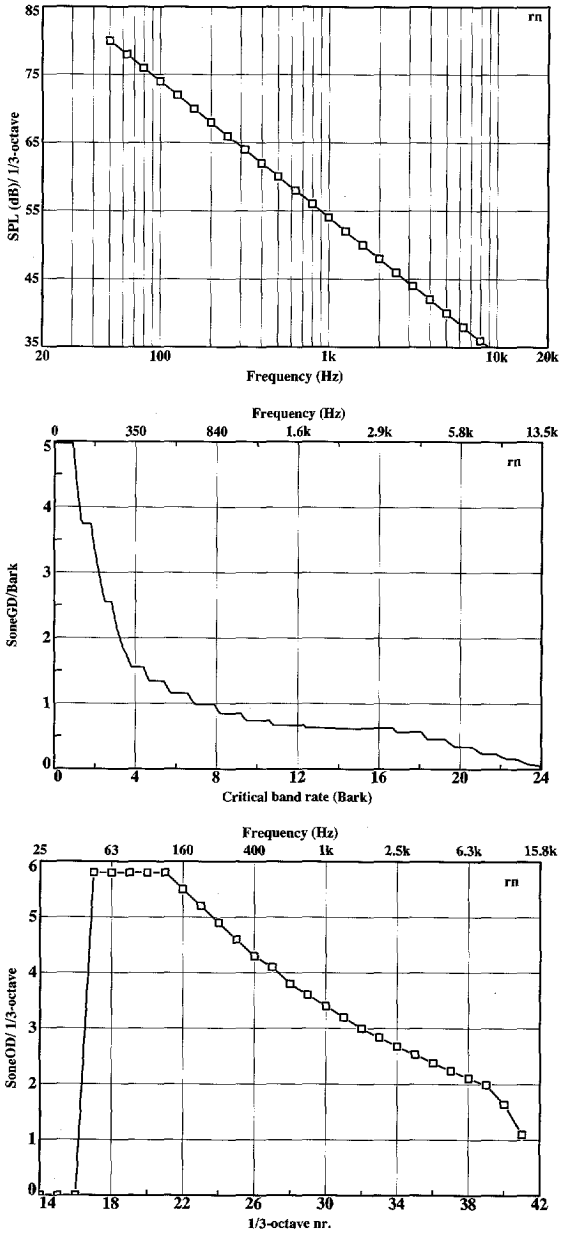

Figure 2.6: Upper panel: (a) Red noise spectrum at 67.4 dBA level. (b) Red noise loudness spectrum according to 532B sone(GD)/Bark. Lower panel: (c) Red noise loudness spectrum according to 532A sone(OD)/(one-third octave).

Table 2.1: Comparison of loudness measures.

signal	400 Hz Level dB	SPL dB	Weighted Level dBA	Judged level dB	Stevens OD phon	Zwicker GD phon	Fig.
M18	65.5	74.0	69.8	81.5	81.0	87.7	-
K19	59.5	74.0	73.5	85.0	85.0	90.7	-
M4	59.0	74.0	71.4	83.5	83.2	88.4	-
K11	55.0	74.0	70.7	84.0	83.0	89.4	-
Sine1k	-	90.0	90.0	-	90.4	91.0	2.3
Pink	96.0	110.2	107.9	-	121.8	122.1	2.4
White	60.0	82.9	81.2	-	94.3	96.9	2.5
Red	62.0	84.3	67.4	-	81.2	86.4	2.6

- Judged loudness level of the noise spectrum obtained by comparing the noise sources with a narrow band noise at 1 kHz (these values are taken from Table III of [108]).
- Loudness in phon(OD) (method 532A) calculated by assuming diffuse field measurement conditions.
- Loudness in phon(GD) (method 532B) calculated by assuming diffuse field measurement conditions.
- Figure numbers referring to the corresponding spectrum.

The table contains the following types of noise:

- M18, K19, M4, and K11 are spectra measured from engines by Lübcke et al. [131]. They recorded 19 different types of noise for subjective loudness ratings. These experiments were repeated and extended by Jahn [108]. The noise sources were all band limited from 90 Hz to 11 kHz and scaled to a noise power of 74 dB.
- Sine1k (Fig. 2.3) is a pure tone of a frequency of 1 kHz at 90 dB SPL.
- Pink noise (Fig. 2.4) for several power levels.
- White noise (Fig. 2.5) for several power levels.
- Red noise (Fig. 2.6). Red noise falls at a rate of 6 dB/oct.

Parts (a) of Fig. 2.4–2.6 represent the spectra of the various noise sources all measured with a one-third octave analyzer in a diffuse sound field. They serve as the input for the actual loudness calculations. The corresponding parts (b) of these figures represent the loudness versus the critical band rate according to method 532B assuming a diffuse sound field, parts (c) show the loudness according to method 532A.

2.3.1 Comparison of ISO 532A and 532B

At first glance there is quite a difference when the Figs. 2.4(b)–2.6(b) are compared with their counterparts, Figs. 2.4(c)–2.6(c), but one has to bear in mind that the frequency axes are different (as discussed in Section 2.2.3). Another apparent difference is the low-frequency gain for 532B, which may be due to the addition of the first three groups of one-third octaves into critical bands, as discussed in Section 2.2.3. Finally, there remain discrepancies due to the different approaches of the two methods. Zwicker's method is elegant because of its compatibility with the accepted models of the human ear, but Stevens' method is based on a heuristic approach and fits better to the subjective ratings. The tendency for the Zwicker procedure to give values systematically larger than Stevens' has also been noted by others. This is not a drawback however, since, as discussed in Section 2.4.2, we were interested in relative loudness levels only.

2.4 Calculation of loudness

As is necessary for loudness calculations, the responses of the loudspeakers have to be determined first. Six loudspeakers were used for this purpose. The loudspeakers, hereafter labelled LS1–LS6, were of different brands and covered a wide price and quality range. The frequency responses of the loudspeakers were measured both in an anechoic chamber and in the listening room where the listening tests were performed. The results of the free-field measurements are shown in Fig. 2.7. Clearly, each loudspeaker exhibits a distinct frequency response, and consequently they sound very different. The loudspeaker measurements in the listening room were made with a real-time one-third octave analyzer and a microphone pointed towards the ceiling, while the loudspeaker being tested was reproducing pink noise. The microphone was placed in the middle of the listening area at ear height and the loudspeakers were placed next to each other. The results are depicted in Fig. 2.8. From these measurement results the weighted and unweighted sound pressure levels and the loudness levels were calculated. However,

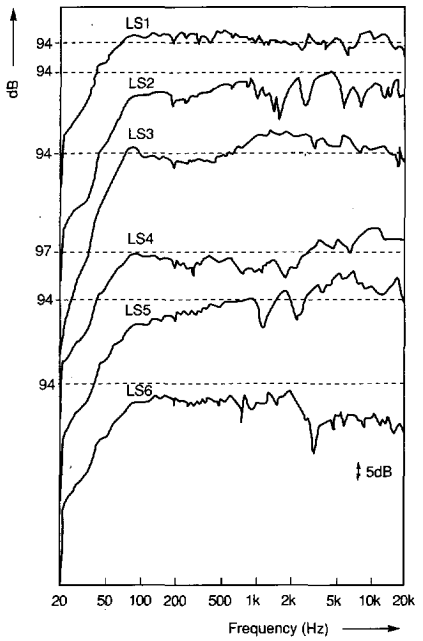

Figure 2.7: Free field on-axis response of the six loudspeakers (3 V across the terminals).

both direct sound as well as reflections due to the room were measured by the microphone and thus contributing to the measured SPL. The loudness models ignored non-simultaneous masking [139] of the human ear, such as backward masking (also known as pre-stimulatory masking) and forward masking (also known as post-stimulatory masking). The calculated loudness versus the critical band rate, using the Zwicker procedure, is given in Fig. 2.9, showing the prediction of the model as what one would perceive it. As can be seen, there is quite a difference among the various loudspeakers; for example, LS5 exhibits a poor low-frequency response.

2.4.1 Subjective loudness measurements

The perceived sound quality of a loudspeaker and its relation to the various physical properties of the loudspeaker have been a subject for discussion and research for a long time, see e.g. [79, 193, 197, 199–201]. An important parameter during listening tests is the setting of the sound level, both for the different programs and for the relative levels between the different loudspeakers. The latter is especially important as it is well known that a higher reproduction level, or loudness level, of a loudspeaker can lead to

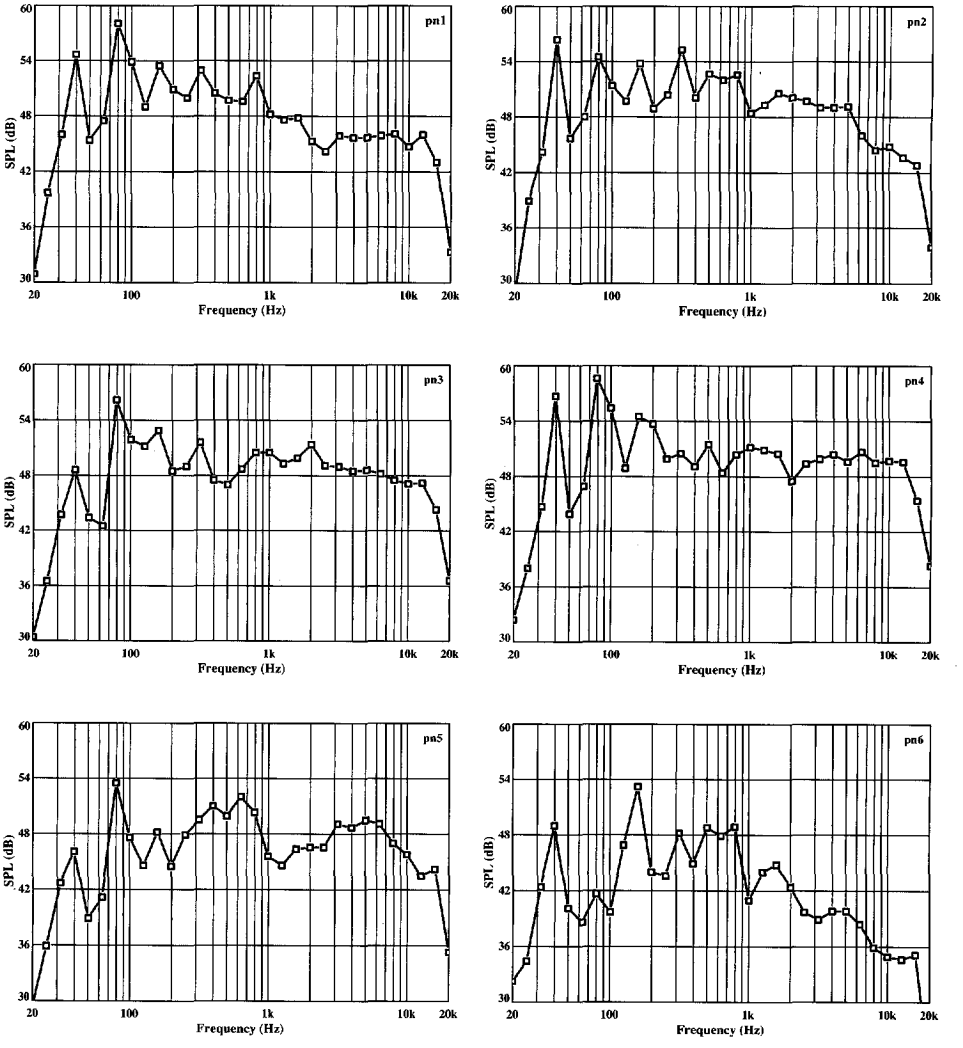

Figure 2.8: Response of the six loudspeakers reproducing pink noise, measured in one-third octaves in the listening room.

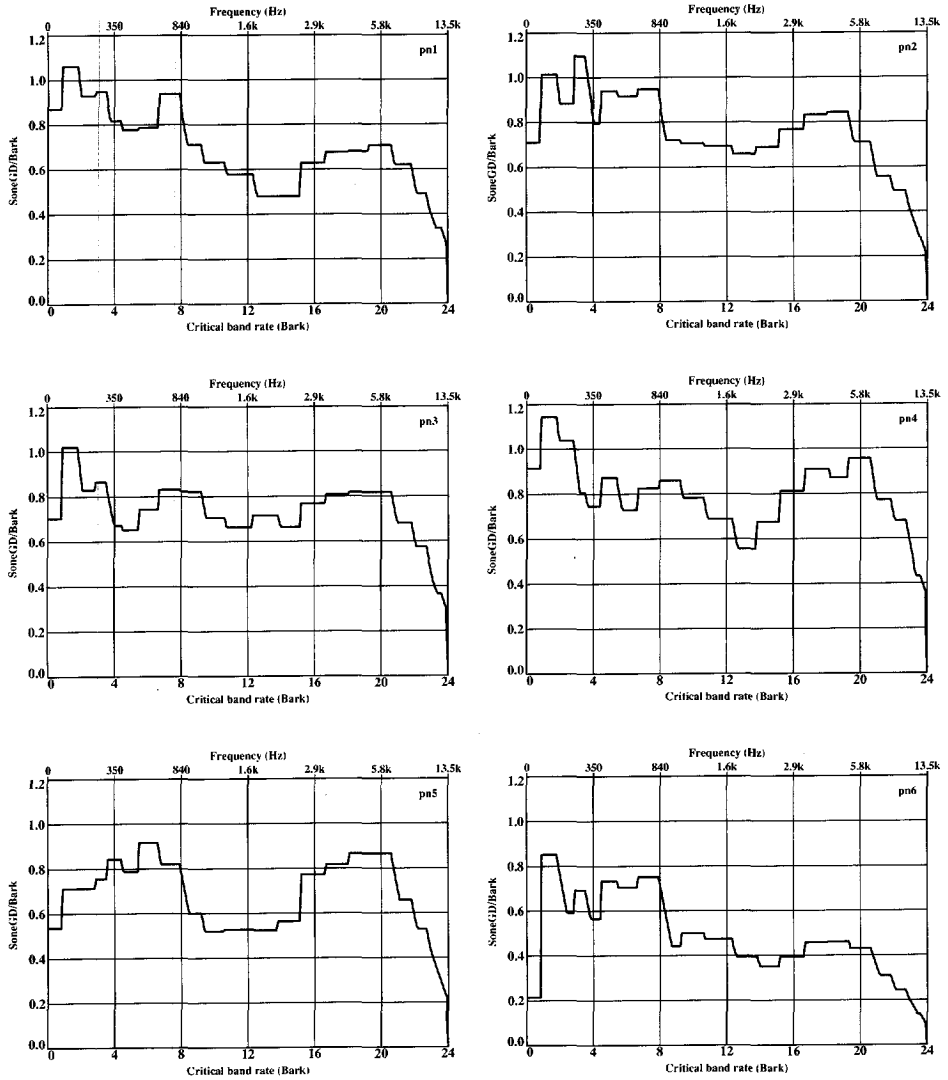

Figure 2.9: Loudness of the loudspeakers in the listening room according to 532B (sone(GD)/Bark).

a higher appreciation score than that of another one of the same quality, or even the same loudspeaker. The importance of equal loudness levels of the sounds being compared is shown by a striking investigation of Illényi and Korpássy [107]. They found that the rank order of the loudspeakers according to the subjective quality judgements was in good agreement with

the rank order obtained by the corresponding calculated loudness.

2.4.2 First experiment

Aim The aim of this experiment was to investigate the capability of subjects to match the loudness of noise sounds, each one having a different timbre. The aim was also to inquire into the variability amongst the subjects for such a task.

Method Usually, loudness is determined directly by using a pure tone at 1 kHz, and is obtained at other frequencies indirectly by means of loudness matching. However, with this experiment only the relative loudness of several loudspeakers is considered, which eases the subjects' task considerably.

Listening Conditions The subjects were seated in front of the loudspeakers, at a distance of 3.5m (see Fig. 2.10). The listening room was a

Figure 2.10: Experimental set-up.

soundproof room of 8.35 m length, 4.50 m width, and 2.62 m height and fulfilled the requirements of IEC 268-13 [234]. The room was arranged and equipped as a normal living room with chairs and furniture. Diffusion of the sound field was enhanced by a glass window, bookcases and framed pictures. The room has a reverberation time of about 500 ms at 125 Hz, which gradually decreases to about 300 ms at 4 kHz.

Technical Equipment We used six different loudspeakers, including the standard (LS1). An ostensible 7th loudspeaker (LS7) was used, but this was physically the same as LS1. The loudspeakers were not seen by the subjects, due to an acoustically transparent but visually opaque screen. The loudspeakers were connected to a switching facility, which contained a set of high-quality relays, remotely controlled by the subject. Variable attenuators were placed in the signal path from the CD player to the power amplifier. Each loudspeaker could be attenuated by the experimenter, by adjusting the knob corresponding to the loudspeaker that was playing. To eliminate possible cues, the knobs could not be seen by the subjects.

Subjects There were five male subjects involved ranging in age from 22 to 46 years, some of whom were naive listeners and some highly experienced. All subjects were recently tested for normal hearing and they could all be considered as otologically normal persons.

Procedure The stimuli were presented by reproducing pink noise via six different loudspeakers, LS1–LS6. The subjects could compare the loudspeakers LS2–LS7 with the reference one (LS1) as often as they desired. The reference loudspeaker was used as an anchor or standard, its volume setting remaining constant during all the tests, resulting in an SPL of 60 dBA for pink noise. The other loudspeakers were to be matched by the subjects so that they perceived an equal loudness level in comparison with the standard. The subjects gave a signal to the experimenter to lower or raise the volume of the loudspeaker being tested. The test was accomplished within two runs for each subject. In the first run the subject balanced the loudspeakers LS2–LS7 to the reference one, which took approximately 10 minutes. The result of the test was the volume setting of each attenuator. After a short break of about two minutes, the second run began following the same procedure. Both runs used the same pink noise source for the programs, however the programs hereafter referred to as PN1 and PN2 respectively, were considered as a factor to study the ability of the subjects to reproduce the same loudness level as during the first run.

Results

Analysis of variance (ANOVA) [82, 221, 234] was performed to analyze the results of the listening tests. ANOVA essentially means that the total variance in the data is split up into different components due to the different sources of variation. These sources can be the program, the loudspeakers,

Table 2.2: Means of 5 subjects (values in dB), for the first experiment.

	PN1	PN2	Avg.
LS 2	-0.05	-0.10	-0.07
LS 3	-0.40	-0.50	-0.45
LS 4	-1.00	-0.80	-0.90
LS 5	1.15	0.35	0.75
LS 6	4.70	5.05	4.88
LS 7	-0.05	0.10	0.02
Avg.	0.72	0.68	0.70

Table 2.3: Analysis of Variance Table, for the first experiment.

Source	SS	df	F	Fc
LSS	223.72	5	43.39	6.6
Program	0.03	1	0.03	1.8
Interaction	2.07	5	0.40	6.6
Residual	49.50	48		
Total	275.31	59		

the subjects or other causes or the interaction between them. The statistical tests make it possible to decide whether the differences between the sources of variation are real, with a certain probability, or whether they are caused by other (random) errors. The six loudspeakers and the program were considered as factors, being the possible sources of variation. The subjects were not considered as a factor, however their variation was taken into account also. Formally this is called a 6X2 design with repeated measures on the same elements. All the responses of the subjects are in dB, relative to the absolute level of LS1. The individual responses are averaged and listed in the Table of Means in Table 2.2. The results of the ANOVA computations are recorded in Table 2.3. The ratio of the resulting variance quantities will possess, under the 'ordinary' normality assumptions, an F distribution [221]. When the F values exceed a certain critical F value (Fc), one may conclude that the corresponding factor is of statistical significance. The value of Fc depends on the value of p, which gives the probability that the variance is caused by change and not by the studied factor. For the computation of Fc, p was chosen to be equal to 0.0001. The entries SS are

the sums of squares and df are the number of degrees of freedom for the particular factors. It appears from the table that the loudspeakers are the only statistically significant factor (p <0.0001). Furthermore, it was very surprising, at least to the author, that most subjects were very good at reproducible loudness balancing. They could adjust the relative loudness of a loudspeaker, say LS4 to that of LS1, and reproduced that level after a retention period of, say, 10–15 minutes with a good accuracy. One has to bear in mind that the timbres of all the loudspeakers were very different.

2.4.3 Second experiment

The aim of this experiment was to investigate the capability of subjects to match the loudness of several loudspeakers for different programs, consisting of pink noise and music excerpts. A further aim was to study the influence of the program, or the existence of a possible interaction between program and loudspeakers. The same method as in the first experiment was used, but the program was extended by including music and the set of subjects was increased.

The program material consisted of excerpts from Compact Discs, listed in Appendix D. The subjects were given as much time as was required. In this experiment 7 male and 3 female subjects participated, ranging from 21 to 46 years of age, all of whom were tested for normal hearing (< 20 dB hearing loss, ISO 389).

Results

The responses of all subjects were measured in decibels relative to the absolute level of LS1. The individual responses were averaged and recorded in Table 2.4, while the ANOVA results are in Table 2.5. It appears from the table that the loudspeakers are the only statistically significant factor ($p < 0.0001$). The influence of the program alone, as well as the interaction between program and loudspeakers are not statistically significant. There remains a substantial residual variance however, due in addition to normal errors such as inadvertence of the subjects, to the lack of a full consensus amongst the subjects. Another reason for the residual variance is that not every subject receives the same signal due to the poor directivity of some loudspeakers, their positioning and the different ear position of the several subjects.

Table 2.4: Table of Means of 10 subjects (values in dB), for the second experiment.

	PN	Ferry	Rabbit	Jazz	Clas.	Avg.
LS 2	-0.90	-1.40	-0.85	-1.62	-1.17	-1.19
LS 3	-0.57	-0.63	-0.35	0.13	0.32	-0.22
LS 4	-1.40	-1.23	-1.35	-1.10	-1.70	-1.36
LS 5	1.17	1.17	0.85	1.00	0.70	0.98
LS 6	5.75	5.02	5.09	5.32	5.60	5.36
LS 7	-0.13	-0.15	-0.10	0.42	-0.35	-0.06
Avg.	0.65	0.46	0.55	0.69	0.57	0.58

Table 2.5: Analysis of Variance Table second experiment.

Source	SS	df	F	Fc
LSS	1545.54	5	256.58	5.1
Program	1.97	4	0.41	5.8
Interaction	20.58	20	0.85	2.6
Residual	325.27	270		
Total	1893.36	299		

Table 2.6: Computed loudness values.

	SPL dB	SPL dBA	ISO 532A phon(OD)	ISO 532B phon(GD)	subj. dB
LS 1	65.0	59.6	75.3	80.7	-
LS 2	-0.4	-2.0	-1.0	-1.4	-0.9
LS 3	0.7	-1.8	-1.1	-0.9	-0.6
LS 4	-1.3	-2.8	-2.5	-2.1	-1.4
LS 5	2.3	-0.7	-0.1	0.2	1.2
LS 6	5.3	4.2	5.7	5.0	5.8
LS 7	-	-	-	-	-0.1

Table 2.7: Difference between subjective and objective measurements.

	SPL dB	SPL dBA	ISO 532A dB(OD)	ISO 532B dB(GD)
LS 2	0.53	-1.14	-0.05	-0.46
LS 3	1.26	-1.15	-0.53	-0.33
LS 4	0.13	-1.36	-1.13	-0.69
LS 5	1.11	-1.90	-1.32	-1.02
LS 6	-0.50	-1.58	-0.08	-0.80
HT2	14.94	54.32	16.58	5.94
α	$< 10^{-3}$	$< 10^{-4}$	10^{-3}	~ 0.06

2.5 Comparison of computed and measured loudness

The computed loudnesses for the loudspeakers mentioned in Section 2.4 are recorded in Table 2.6. Loudnesses (LS2–LS7) are relative to the loudness level of LS1. The results of the second experiment (the average over subjects and pink noise as program) are gathered in the same table, as the last column. The error of an objective measure is considered to be the difference between the loudness levels of the subjective and an objective method respectively, these values are in Table 2.7. The entry HT2 is Hotelling's (squared) generalized Student ratio, or briefly, Hotelling's T^2 [221].

Clearly, the table shows that the dBA values can differ considerably from the subjective measurements. The statistical significance of the data

presented in Table 2.7 can be tested against the following zero hypothesis 'The differences between the various methods (dB, dBA, ISO-A and ISO-B) and the subjective ratings are due to chance variations only'. Using the T^2 values from Table 2.7 one can not reject the zero hypothesis for ISO 532B, the hypothesis is rejected for the other three methods. The entry α in Table 2.7 denotes the level of significance of the T^2 test, it is the probability of making the decision to reject the zero hypothesis when in fact it is true. One may conclude that the ISO 532B method provides results similar to those of the subjects; the three other methods are not consistent with the subjective ratings.

2.6 Conclusions

The best technique for adjusting the interloudness levels of loudspeakers during listening tests is using the ISO 532B method. It provides similar results as the population of 10 subjects. The A-weighting method is not recommended for accurate loudness balancing. The results of the computed loudness level of pink noise reproduced by several loudspeakers agree very well with the average loudness level adjusted by several subjects. The average loudness for a varied repertoire is very similar to the computed loudness for pink noise as the program. There is a little variability among the subjects. However, analysis of variance shows that the only significant factor is the difference between the loudspeakers. The loudness levels are hardly influenced by the program choice.

Chapter 3

A Comparison of some loudness measures [1]

For the purpose of loudspeaker listening tests, simple weighting methods and ISO loudness models are compared with listener-adjusted loudness levels. For a loudness level of about 80 phon, the B-weighting appeared to be the best method while the commonly used A-weighting is unreliable.

3.1 Introduction

In listening tests, opinions that are formed about sound quality and stereo imaging are influenced by many factors in addition to the one that may be of specific interest, see [202] for a brief overview. One of the sources of variability is loudness. The loudness balancing of loudspeakers during listening tests on loudspeakers is considered to be very important. Among a variety of literature [77–79, 107, 193, 197, 199–202] it was recently noted by Gabrielsson [77, 78] that an increase in sound level will increase the perceived fullness, and spaciousness, and will give a better clarity and fidelity. In Chapter 2, calculation of the loudness due to loudspeaker sound reproduction was discussed. Some standardized loudness calculations were compared with the traditional method relying on A-weighted sound levels and with subjective loudness measurements obtained through listening tests. One of the conclusions of that chapter was that the A-weighted sound level method was not recommended for accurate loudness balancing and that the loudness differences between the loudspeakers are hardly influenced by

[1] Parts of this chapter have been published in [6].

the program choice.

The recommended measures (ISO 532) correlated well with the subjective ratings of the various subjects. However, many consider these methods to be cumbersome and too complicated for everyday use. While the ISO methods are intended to work over the whole dynamic range of the auditory system, the practical use for loudspeaker listening tests is only, say, 70–80 phon. This relaxes the problem of loudness measurements, and makes simpler methods feasible. Therefore in this chapter the comparison of loudness measures will be extended to include other simple ones such as the B-, C- and D-weighting functions. It should be noted that the A-, B-, C- and D-weighting curves are only intended for rank ordering of noises according to loudness and not for measuring absolute loudness, while the ISO 532 methods are based on psychoacoustical data of human ears and can be used to measure absolute loudness. In the following sections it will be discussed why the A-weighting is not recommended, and a better alternative will be examined.

3.1.1 Equal loudness

In the equal-loudness-level contours for pure tones, plotted in Fig. 3.1, two psychoacoustic phenomena may be observed. Firstly, the contours are heavily frequency dependent and, secondly the curves are level dependent. The latter can be illustrated by Fig. 3.2. It shows that the normalized differences between the 80 phon curve and the 20-, 40-, 60- and 100-phon curves are increasing for decreasing frequency, below 200 Hz. The shapes of equal-loudness contours have been used in the design of sound level meters, which attempt to give an approximate measure of the loudness of complex sounds. Such meters contain weighting networks, so that the meter does not simply sum the power at all frequencies but, instead, weights the power at each frequency according to the shape of the equal-loudness contours. At low sound levels low-frequency components contribute little to the total loudness of a complex sound, so A-weighting is used to reduce the contribution of low frequencies to the final meter reading. At high levels all frequencies contribute more or less equally to the loudness sensation, so that a more nearly linear weighting characteristic, the C-network, is used [139]. B-weighting is used for intermediate levels, while D-weighting is used for very high levels, such as aircraft noise. The A-weighting, which is traditionally used for general purposes, is supposed to be an approximation of the 40-phon contour. This level is much too low for loudspeaker listening tests. When the 80-phon contour is used to obtain a weighting function by

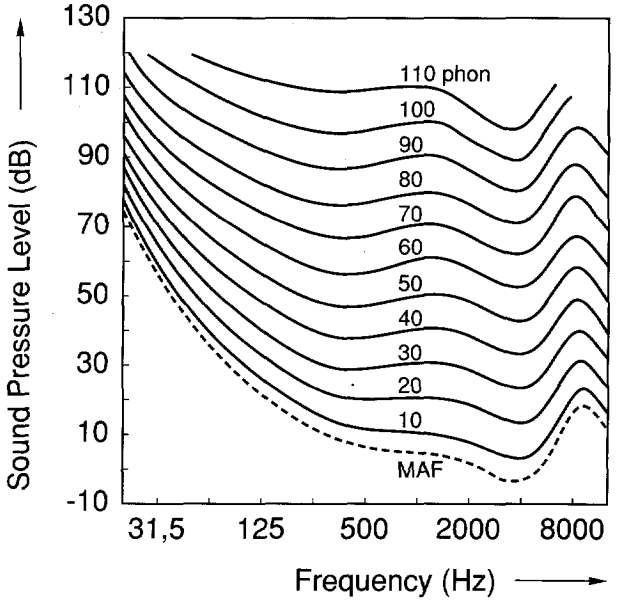

Figure 3.1: Normal equal-loudness level contours for pure tones (binaural free-field listening, frontal incidence), from [235] Fig. 1.

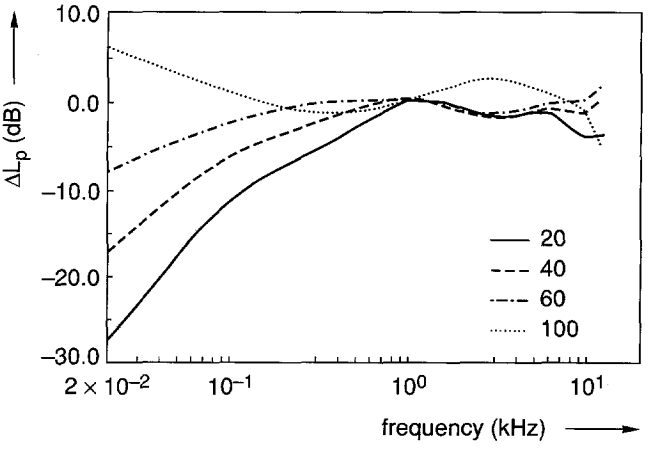

Figure 3.2: Differences between the 80 phon curve and the 20-, 40-, 60- and 100-phon curves, respectively. The difference curves have been normalized to 0 dB at 1 kHz.

normalizing it to 0 dB at 1 kHz, the curve labeled 80-phon weighting will result, as shown in Fig. 3.3. As a reference, A-weighting and B-weighting are also plotted in Fig. 3.3. It appears however, that at low frequencies

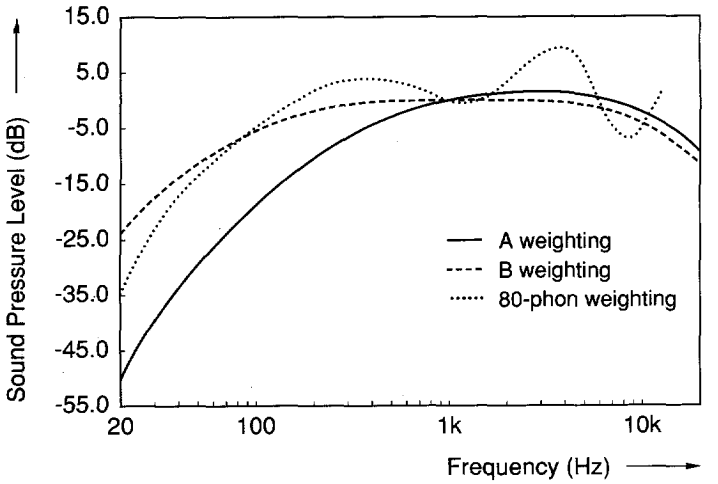

Figure 3.3: A-, B- and 80-phon (free-field) weighting functions.

the A-weighting is too strong while the B-weighting is a reasonable approximation of the 80-phon weighting curve. Another way to demonstrate the weakness of a simple weighting in general and the A-weighting in particular is the following. When a subject listens to a pure tone at 200 Hz or 2 kHz, each with the same SPL of 60 dB, each tone will give about the same loudness (see Fig. 3.4). The A-weighted value of the 200 Hz tone does not reflect the perceived strength, however. If the subject listens to the two tones simultaneously, with a frequency separation of more than a critical band, the perceived loudness will increase by about 10 phon(GD) with respect to a single tone. (The suffix GD is an acronym for Group and Diffuse field, see [229].) The A-weighted level will remain the same as for the 2 kHz tone presented alone. This is due to the too rigorous weighting at low frequencies (at higher levels), and because the addition of signals has different effects in psychoacoustics than for electrical signals. The addition rules for loudness are incorporated in the more advanced loudness measurements, as discussed in Chapter 2. However, if loudspeakers under tests are similar, then there is no serious objection to a simpler weighting. For reference purposes, the A-, B-, C-, and D-weighting functions are plotted together in Fig. 3.5. In Appendix E a computer procedure to compute these functions is presented. In the following sections these global comparisons will be tested against listener-adjusted loudness levels.

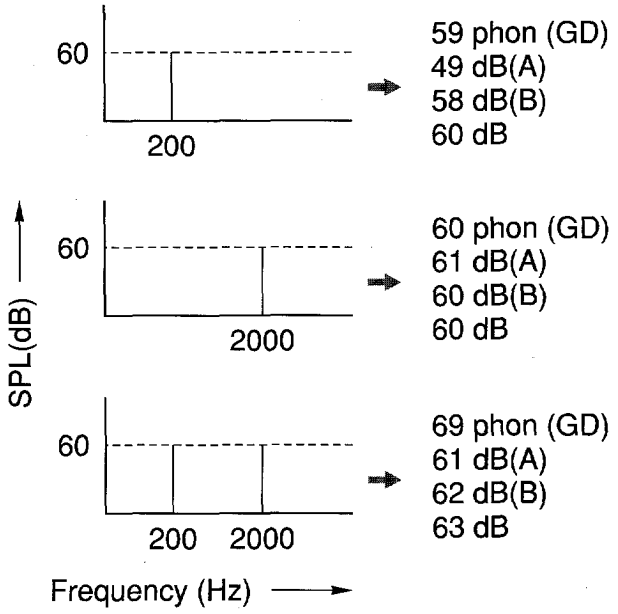

Figure 3.4: Levels of tones at 200 Hz (upper), 2 kHz (middle) and the two tones simultaneously (lower figure).

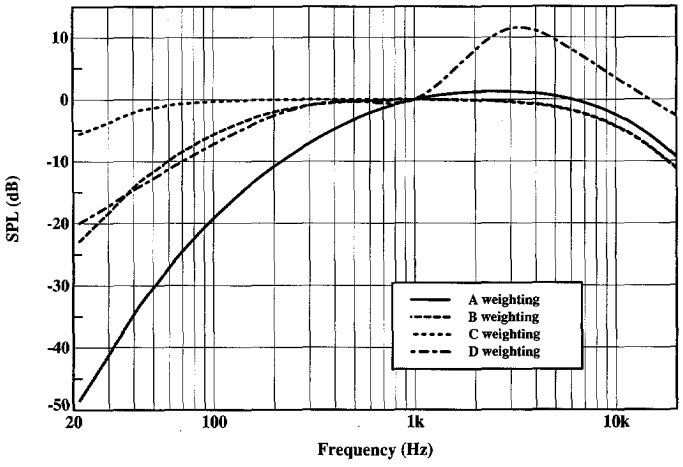

Figure 3.5: A-, B-, C- and D-weighting functions.

3.1.2 Subjective loudness measurements

To test the usefulness of the objective loudness measures, these values will be compared with listener-adjusted loudness levels. These subjective values

were obtained by an experiment, which will be summarized here; the details are to be found in [5]. Ten subjects, one at a time, listened to six different loudspeakers (LS1–LS6), including the standard, or reference (LS1), at a distance of 3.5m. The loudspeakers were of different brands and covered wide ranges of price and quality. They exhibited very dissimilar frequency responses and different efficiencies, especially at low frequencies. The listening room was a soundproof room arranged and equipped as a normal living room. The loudspeakers could not be seen by the subjects, due to an acoustically transparent but visually opaque screen. They were connected to a switching facility, which contained a set of high-quality relays, remotely controlled by the subject. Variable attenuators were placed in the signal path from the CD player to the power amplifier. Each loudspeaker could be attenuated by the experimenter by adjusting the knob corresponding to the loudspeaker that was playing. The stimuli were presented by reproducing pink noise via the six different loudspeakers, LS1–LS6. The subjects could compare the loudspeakers LS2–LS6 to the reference one (LS1) as often as they desired. The loudspeakers LS2–LS6 were to be matched by the subjects so that they perceived an equal loudness level in comparison with the standard. The subjects gave a signal to the experimenter to lower or raise the volume of the loudspeaker under test. When the subject was satisfied with all loudness levels (which took approximately 10 min.), the level of each attenuator was stored, this was done once per subject. These values were averaged (over the ten subjects) and hereafter referred to as L_{subj}. The reference loudspeaker was used as an anchor or standard. Its volume setting remained constant during all the tests, resulting in a loudness level of 80 phon(GD) for pink noise. See Table 3.1 for a comparison with other measures.

Table 3.1: Comparison of some loudness measures for a pink noise source with an SPL of 48 dB in each one-third-octave in the range of 20 Hz–20 kHz.

SPL	Sound Level				532A	532B
dB	dB(A)	dB(B)	dB(C)	dB(D)	phon(OD)	phon(GD)
62.91	59.93	60.31	61.69	67.04	75.71	80.33

Table 3.2: Difference between objective and subjective measurements.

	SPL	Sound Level				532A	532B
	dB	dB(A)	dB(B)	dB(C)	dB(D)	dB(OD)	dB(GD)
LS2	0.530	-1.140	-0.300	0.490	-1.690	-0.050	-0.460
LS3	1.235	-1.175	0.405	1.235	-1.845	-0.555	-0.355
LS4	0.130	-1.360	0.080	0.290	-1.910	-1.130	-0.690
LS5	1.135	-1.875	-0.185	1.165	-2.785	-1.295	-0.995
LS6	-0.450	-1.530	-1.580	-0.670	-0.690	-0.030	-0.750
HT2	13.62	48.22	4.16	14.58	74.35	18.90	6.54
α	$< 10^{-5}$	$< 10^{-15}$	~ 0.7	$< 10^{-6}$	$< 10^{-15}$	$< 10^{-10}$	~ 0.1

3.1.3 Results

To compare the results of the various methods tested the error was calculated as

$$\Delta_{mj} = (L_{m1} - L_{mj}) - L_{subj} \tag{3.1}$$

where L_{m1} is the sound level of the reference loudspeaker (LS1) using method m, L_{mj} is the sound level of the jth loudspeaker using method m, and L_{subj} the averaged relative level adjusted by the subjects for the jth loudspeaker. The results of the listening test and using Eq. 3.1 are summarized in Table 3.2 (the first column is the unweighted SPL). The entry HT2 is Hotelling's T^2 [221], given as

$$T^2{}_m = \delta_m^T \mathbf{C}^{-1} \delta_m \tag{3.2}$$

where δ_m is the vector of differences of method m (the columns of Table 3.2), and $\mathbf{C}$ the covariance matrix of the subjects' ratings. A large value of T^2 indicates a large deviation from the subjects' ratings. Clearly, Table 3.2 shows that D-weighting is not applicable. The second worst method is the A-weighted sound level. The simple B-weighting is surprisingly the best in this test. The statistical significance of the data presented in Table 3.2 can be tested against the following zero hypothesis: 'The differences between the various methods (unweighted; A-, B-, C- and D-weighted; ISO-A and ISO-B) and the subjective ratings are due to random variations only'. Using the T^2 values from Table 3.2, one cannot reject the zero hypothesis for B-weighting and ISO 532B. The hypothesis is rejected for the other five methods. The entry α in Table 3.2 denotes the level of significance of the T^2 test, which is the probability of making the decision to reject the zero hypothesis when in fact it is true (type I error). The power of this test

cannot be calculated explicitly. However, it can be shown that the power of the present test is much higher than the power of a one-dimensional test and is sufficient to reject some methods. One may conclude that the B-weighting and ISO 532B methods provide results similar to those of the subjective assessments. The five other methods are not consistent with the subjective ratings. It should be noted that the ISO 532 method is intended for general absolute loudness measures applicable for various levels and sound sources; while in this present case only relative loudness measures of comparable sound sources are of interest.

3.2 Conclusions

Experimental evaluations were made of seven measurement techniques to identify those that would be useful for the adjustment of loudness levels of loudspeakers for listening tests, at a level of 80 phon. The most satisfactory results were obtained by the use of a B-weighted measure of sound level. This provided results similar to those derived from subjective adjustments by a population of 10 subjects. The elaborate ISO 532B method also gave good results. The A-weighted measure yielded poor results and therefore is not recommended for accurate loudness balancing.

Chapter 4

Loudspeaker sound power vs. sound pressure level

The listening-room response of loudspeakers is determined largely by the power response. Attention is therefore paid to room-correction values of reverberation rooms. The room-correction values of two reverberation rooms have been determined. This correction is given to a high degree of accuracy by a simple expression, which includes the influence of changes in atmospheric conditions such as humidity, temperature and barometric pressure. Using these room-correction values, the total power response of a loudspeaker has been determined and revealed to be an important parameter for loudspeaker quality assessment.

4.1 Introduction

In order to compare the results of the loudspeaker power measurements performed at the Philips Research Laboratories in Eindhoven ('PRLE') and the loudspeaker factory in Dendermonde ('PSS'), we have (re-)calculated the room-correction values of PSS' reverberation room. In PRLE's reverberation room the sound power level, L_W, of a Reference Sound Source was measured. The sound pressure level of the same source was measured in PSS' reverberation room. Via a rather simple relationship between those two values, the room-corrections can easily be calculated. To obtain an impression of the expected accuracy of the measurements, some calculations were performed to account for changes in atmospheric conditions.

4.2 Relation between sound power and sound pressure

The method is based on the assumption that the spatial average of the sound pressure squared, p^2, in the reverberant portion of the sound field, is proportional to the radiated power W. The proportionality constant depends on the room properties according to the equation [152]

$$W = \frac{p_{rms}^2 (6 \ln 10) V}{\rho c^2 T_{tot}}, \tag{4.1}$$

where V is the room volume, ρ the density of the air, c the sound velocity and T_{tot} the reverberation time (T_{60}). Owing to standing waves, p^2 varies with the position. In the low-frequency region the number of excited modes is too small to satisfy the assumption of random incidence and the spatial fluctuation of the sound pressure can be extremely large for narrow band sources. At a certain frequency the modes may be regarded as a smoothed-out continuum, this frequency (Schroeder cutoff frequency [170]) is identified as

$$f_{sch} = \sqrt{T_{60} c^3 / (4V \ln 10)}. \tag{4.2}$$

It is convenient to express Eq. 4.1 in terms of sound power level, L_w, and sound pressure level, L_p, and after substitution of

$$\rho = 1.29 \frac{273}{T_a} \frac{B}{B_0}, \tag{4.3}$$

$$c = 331.4 \sqrt{T_a / 273} \tag{4.4}$$

in Eq. 4.1 the following holds

$$L_w = L_p - 10 \log (T_{tot}/T_0) + 10 \log (V/V_0) - 10 \log (B/B_0) - 14 \tag{4.5}$$

where $V_0 = 1$ m^3, $T_0 = 1$ s, $B_0 = 1000$ mBar, B is the atmospheric pressure in mBar and T_a is the ambient temperature in Kelvin.

4.2.1 Reverberation time

There are a number of loss mechanisms that reduce the energy of sound waves when they are reflected from walls as well as during propagation through the air, all of which influence the reverberation time. Some are of practical relevance and are influenced by changes in atmospheric conditions.

Losses due to room boundaries

At low frequencies, most of the sound energy losses occur at the boundaries of the room, and the corresponding reverberation time is given by Sabine:

$$T_1 = \frac{4V \ln 10^6}{cA_t}, \tag{4.6}$$

where A_t, is the total equivalent absorption of the room boundaries.

Losses due to attenuation of sound in the air

The losses caused by dissipation during sound propagation through the air in the room also affect the reverberation time, particularly at high frequencies. If the sound pressure in a plane wave is attenuated exponentially while it travels a distance x, the pressure can be written as:

$$p(x) = p_0 e^{-\mu_a x} \tag{4.7}$$

where μ_a, is the dissipation coefficient, expressed in Neper/m. It is convenient to write the intensity J as:

$$J(x = ct) = J_0 10^{-2\mu_a ct \log e} \tag{4.8}$$

and it follows that the limiting case of the reverberation time, where all the losses occur during propagation and none during reflection, is given by:

$$T_2 = \frac{3}{c\mu_a \log e}. \tag{4.9}$$

The following approximation formula for the dissipation coefficient is used:

$$\mu_a = \frac{\alpha f^\beta}{RH} \tag{4.10}$$

where RH is the percentage relative humidity, f is the frequency in kHz, α and β are constants. Cremer and Müller [53] report: $\alpha = 0.0085$ and $\beta = 2.0$, Kinsler et al. [115] report: $\alpha = 0.0275$ and $\beta = 1.7$. When the measured values are fitted to Eq. 4.14 the following values are found: for PRLE's reverberation room: $\alpha = 0.0161$ and $\beta = 1.70$; for PSS' reverberation room: $\alpha = 0.0226$ and $\beta = 1.45$. They show a fair agreement with the values found by [53] and [115]. The approximation for μ_a, as given by Eq. 4.10, is given for 20 C°. For values at different temperatures, μ_a may be decreased by 4% for each degree below 20 C° .

Losses due to viscosity and heat conduction

The losses due to viscosity of the air are [53]

$$\mu_\nu = \frac{2\eta\omega^2}{3\rho c^3} \,. \tag{4.11}$$

Substitution of some typical values for air: $\eta = 1.8\,10^{-5}$ and $\rho c = 415$ lead to

$$\mu_\nu = \frac{10^{-6}}{\lambda^2} \,. \tag{4.12}$$

This value is very small in comparison with Eq. 4.10, even at the highest audio frequency, so it can be neglected for the calculation of the reverberation time. Losses due to heat conduction of the air are approximately one third of the viscous losses [115] and so can be neglected as well.

4.2.2 Total reverberation time

The total reverberation time is given by:

$$1/T_{tot} = 1/T_1 + 1/T_2 \tag{4.13}$$

or

$$T_{tot} = \frac{T_1}{1 + \frac{0.14\alpha f^\beta T_1 c}{RH}} \,. \tag{4.14}$$

Fig. 4.1 depicts the reverberation time of PRLE's reverberation room, calculated with Eq. 4.14. In the same figure, values for the reverberation given by [110] are marked. Eq. 4.14 is a very simple one, and for very low frequencies a more thorough calculation is necessary [37, 125, 208]. It fits from 250 Hz to 20 kHz with the measured reverberation time within ca. 0.5 dB. This degree of accuracy makes Eq. 4.14 suitable for the sound power calculation with Eq. 4.5. Eq. 4.14 shows the influence of changes in the humidity of the air on the reverberation time. Fig. 4.1 depicts the reverberation time of PRLE's reverberation room for three values of relative humidity, for RH= 30%, 45% and 65%, respectively.

Another cause of changes in the reverberation time is the behavior of the room itself at different temperatures. The eigen frequencies of the modes are inversely proportional to the square-root of the absolute temperature.

Problems arise in this respect when the frequency response shows steep slopes. Measurements should be made at the same room temperature as that at which the correction function is determined.

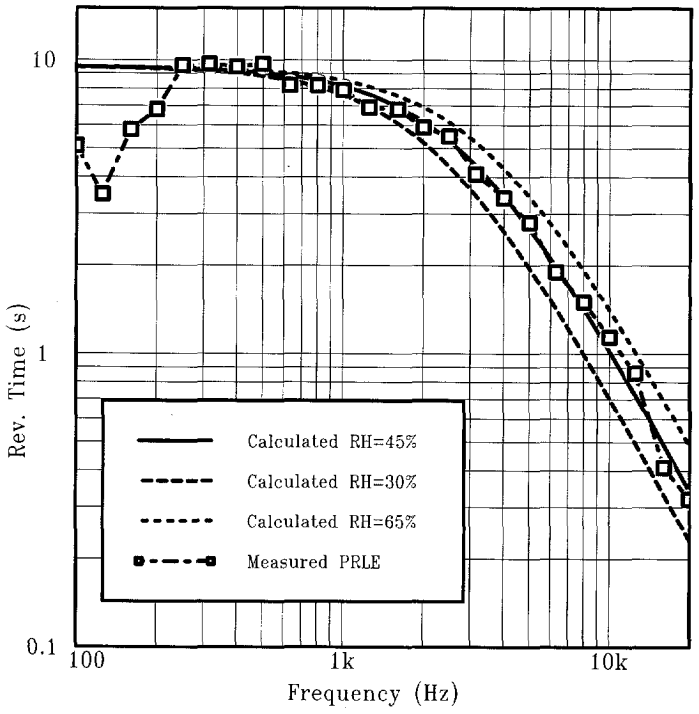

Figure 4.1: Reverberation time of PRLE's reverberation room for $T_1 = 9.5$ s, $\alpha = 0.0161$, $\beta = 1.70$.

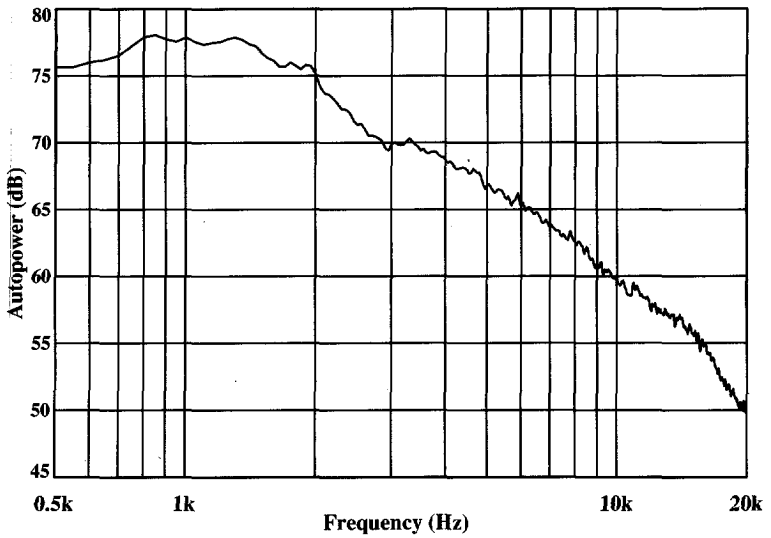

Figure 4.2: Power response of a B&K 4204 reference sound source measured in PRLE's reverberation room.

4.2.3 Determination of the room correction

To determine the room-correction values, R_c of a (reverberation) room, the use of a known sound source may be convenient, therefore we write Eq. 4.1 as

$$R_c = L_W - L_p. \tag{4.15}$$

In PRLE's reverberation room, the L_W of a reference sound source (type B&K 4204) is determined and plotted in Fig. 4.2. With Eq. 4.15, the room corrections of PSS's reverberation room are calculated, and with Eq. 4.5 the reverberation time is calculated and plotted in Fig. 4.3. The reverberation time calculated with Eq. 4.14 is plotted in the same figure. As can be seen, there is good agreement between the measured and calculated values.

4.3 Applications of sound power measurements

The radiated sound power of a loudspeaker system is an important performance indicator. Staffeldt [183] compared subjective loudspeaker data, acquired by listening tests, and objective data. The listening-room response is determined largely by the power response. Citing Staffeldt: 'Comparison of the anechoic sound pressure responses with the subjective data emphasizes once more the fact that anechoic response does not say much about

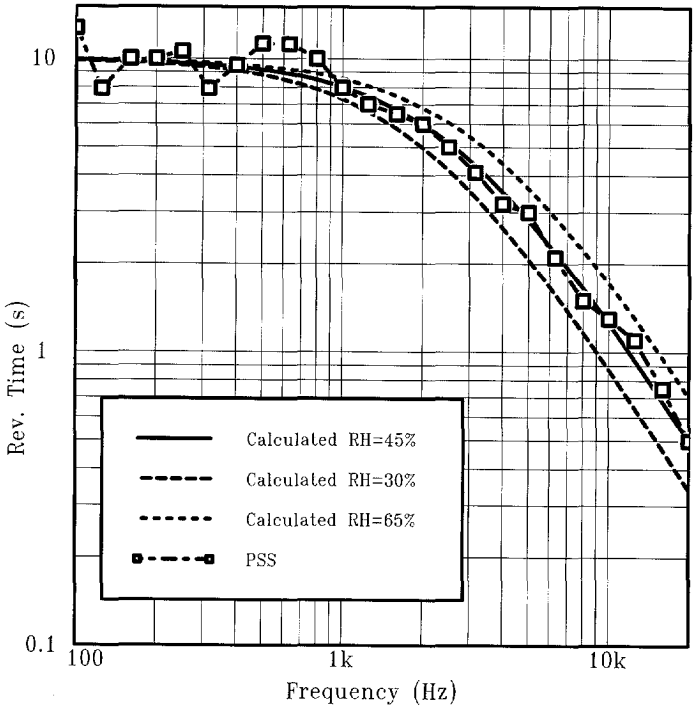

Figure 4.3: Reverberation time of PSS' reverberation room for $T_1 = 10$ s, $\alpha = 0.0226$, $\beta = 1.45$.

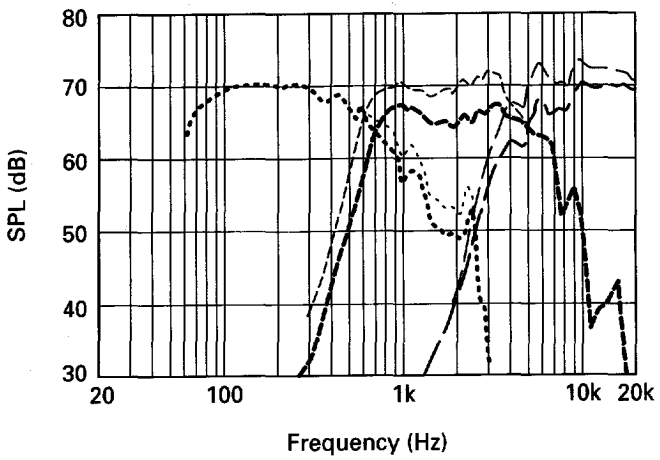

Figure 4.4: SPL of three loudspeakers driven by two different crossover filters, b and w. The three thick curves, b, (woofer, squawker and tweeter) are together ranked as the best, the thin curves, w, as the worst.

the perceived quality of a quality loudspeaker'.

4.3.1 The sound power of loudspeakers with different crossover filters

To combine several drivers in one loudspeaker system, an electronic network (crossover filter) is applied. In Chapter 5, it is discussed how these filters can be simulated. Using this simulation, a listening test was conducted by listening to one pair of three-way loudspeakers equipped with six different crossover filters. Two of this set are plotted in Fig. 4.4 and the total system response is plotted in Fig. 4.5.

Using Rayleigh's power summation formula [159],

$$\sum A^2 + 2\sum A_1 A_2 \cos(\epsilon_1 - \epsilon_2)\sin(kD)/(kD) \qquad (4.16)$$

the second summation being for every pair of sources of strength A, of which A_1 and A_2 are specimens, D denotes the distance between the specimen pair and $\epsilon_1 - \epsilon_2$ their phase difference. The average total power is calculated and plotted in Fig. 4.6. While the on-axis pressure between $\approx 3.5 - 5.5$ kHz is the same for both filters, see Fig. 4.5, the sound power differs in this region by 5 dB (Fig. 4.6), which clearly shows the influence of crossover filters. A listening test was conducted comparing six different sets of crossover filters, where the task of the subjects was to rank the six sets in order of preference.

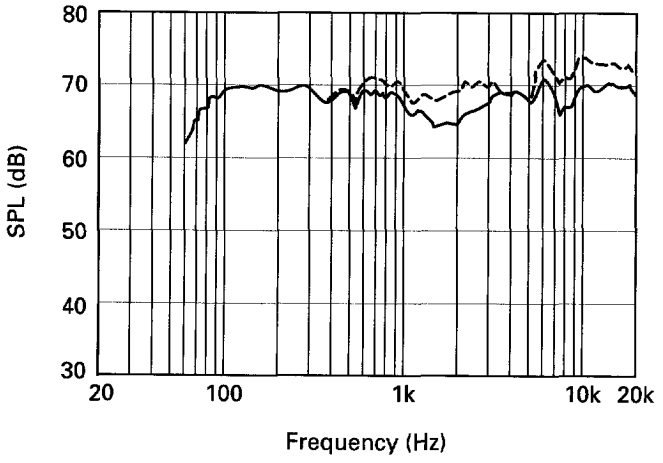

Figure 4.5: Total sound pressure level (SPL) of one loudspeaker system driven by two different crossover filters, b and w. The solid curve, b, was ranked as the best one, the dashed, w, as the worst.

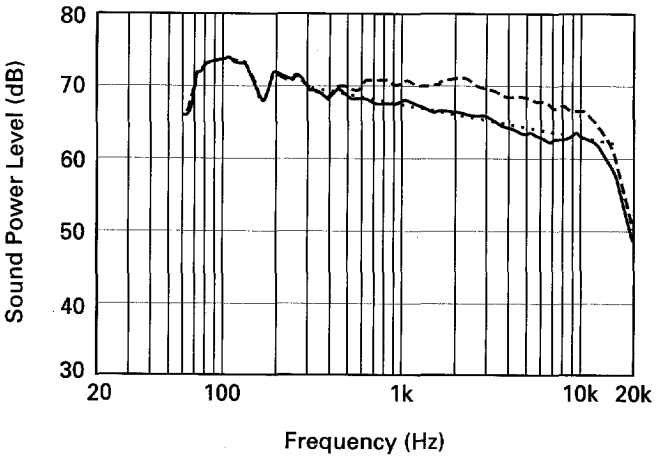

Figure 4.6: Total sound power level of one loudspeaker system driven by two different crossover filters b and w. The solid curve, b, was ranked as the best one, the dashed, w, as the worst. The dotted line is the target for the optimization procedure.

This listening test revealed that there was a remarkable consensus between the ten subjects; for this test: 'tastes *do not* differ'. The set marked *b* was chosen unanimously as the best and the one marked *w* the worst. Set *b* was optimized by a filter optimization program in order to minimize the deviation of the total power curve from a straight line, as indicated in Fig. 4.6. Irregularities in the power response indicate a spatially erratic sound radiation, which has to be avoided.

4.4 Conclusions

The room correction of PSS' reverberation room is determined by comparison of the sound pressure of a sound source with a known sound power output. This has proved to be the best method, since all variant parameters are included at once. It has been shown that the room correction can be calculated reasonably accurately with a single approximation formula. In this case, a number of parameters has to be determined. Combining these two methods leads to the conclusion that some parameters, such as losses due to room boundaries, can be determined once and that other parameters, such as temperature, air pressure and humidity, can be determined ad hoc, after which the room correction including those effects can be calculated easily.

Using these correction values, the sound power of a loudspeaker can be calculated, and this appeared to be very important for the quality assessment of the system. Listening tests revealed that the loudspeaker system with a power response gradually decreasing with frequency and exhibiting no coarse irregularities yields the best sound assessment.

Chapter 5

Crossover filter simulation [1]

A method is presented for the evaluation of crossover filters in a multi-way loudspeaker system. A number of analog filters, designed by 'hand' or with CAD tools, are candidates for use in a loudspeaker system. These different filters are real-time simulated by a digital audio signal processor, which permits flexible and accurate switching between different crossover filters as needed in a listening test. The transformation of the analog filter to its digital equivalent and the implementation in a digital signal processor are discussed.

5.1 Introduction

Using a software package previously made for the numerical optimization of a multi-way loudspeaker system crossover filter [64], a set of component values is given for the crossover filter that satisfies the designer's demands and constraints to the greatest extent possible. In practice, however, many sets are found, e.g. because of different specifications of the optimization targets, different filter topologies or different initial values. These various sets can be differentiated by examining their numerical quality (standard deviation etc.) or by comparing their sensitivities to component value tolerances. It is our experience that after such a differentiation, several almost optimum solutions (sets of component values) remain, and that such solutions with almost identical numerical quality can sound quite different. Therefore we need a final check to select the best sounding solution. Obviously, such a final check is a listening test. In such a listening test it is desirable for the designer to switch between different crossover filter

[1] Parts of this chapter have been published in [11].

candidates, without changing the loudspeakers or the location of the loudspeaker system. Meticulous attention to the acoustical, psychological, and experimental variables is required to achieve subjective ratings that are reliable [199]. This can be achieved if a device is available that emulates the crossover filter, that is, it performs the same task as the analog filter. We will discuss passive crossover filters only but the same techniques can be used for active systems.
For a good comparison between the candidates it is essential to do the listening test very accurately. The above-mentioned reasons justify, in our opinion, the need for an accurate and flexible crossover switching device. This device enables the crossover filters to be changed without side-effects, that is, the same loudspeakers are used and the location of system and listener remains unchanged. This device can be built with analog crossover filters and a high-quality switching facility. However, since, this approach is quite cumbersome, we have implemented such a device with a digital audio signal processor [210].

Filtering digitally has some definite advantages:

1. It provides great flexibility.
2. It can switch fast and accurately between many different crossover filters.
3. All functions can be software-controlled.
4. It can be conveniently coupled with a computer-driven listening test.

A disadvantage is the need for rather complex hardware and software tools, which will be described in the following.

5.2 Simulation of a crossover filter

The simulation of a crossover filter by means of a digital signal processor means that it provides complex-valued transfer functions equal to those between the electric input of the crossover filter and the loudspeaker terminals, that is, the transfer function of the filter loaded with the complex-valued input impedance of the loudspeaker (Fig. 5.1). The digital filter is implemented with a Digital Signal Processor (the ASP described in [210]). To realize the digital filters we need a transformation from the analog domain (the s- or jω-plane) to the digital one (the z-plane), where s is introduced by means of a Laplace-transformation, and z by means of the

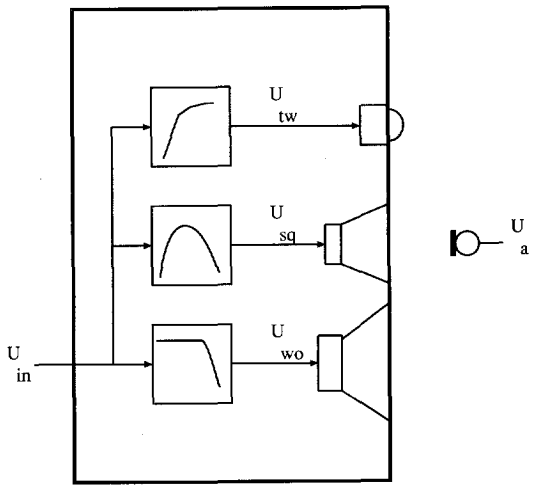

Figure 5.1: A three-way loudspeaker system with (analog) crossover filters.

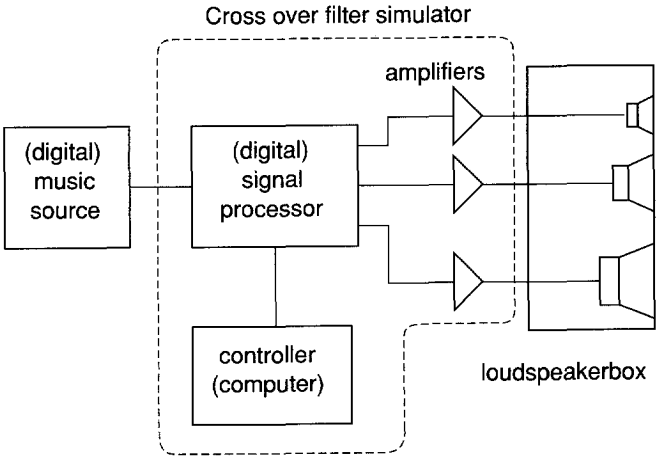

Figure 5.2: Experimental set-up of three drivers and their filters.

z-transformation [157]. When the 'analog' transfer function is given analytically, standard techniques are available for mapping from the analog to the digital one. In the case in question these cannot be used because an analytical description of the transfer function is not available. This transfer function is given at discrete frequency points only, because the input impedance of the loudspeaker has been determined by measurement at a finite number of frequencies. This necessitates the use of optimization or curve-fitting techniques for mapping the analog filter to its digital equivalent. The transfer function of the digital filter is written as the ratio of two z-polynomials [157] and we have to find the coefficients of these polynomials to simulate the specified complex-valued transfer function as good as possible.

5.2.1 Selecting the type of the digital filter

The choice to be made is whether the denominator of the 'digital' transfer function of $H(z)$ equals unity (a FIR or finite impulse response filter) or is given by a polynomial (an IIR or infinite impulse response filter). Despite the advantages of a FIR filter, its unconditional stability, the relatively easy design and its lack of limit cycles, we opted for an IIR approach. We did so because an appropriate FIR filter would have to be rather long in order to match the desired accuracy at low frequencies. A disadvantage of the IIR filter approach is the difficulty of determining the coefficients and the orders of the polynomials. When determining the coefficients of the polynomials we have to account for

1. a correction for the spectral shaping of the digital-to-analog converter [157].

2. the influence of the finite word length of the coefficients (that is, the number of bits used for each coefficient).

3. the calculated filter must be stable, which requires that the roots of the denominator polynomial be located inside the unit circle in the (complex) z-plane.

The next section describes the estimation of the order of the polynomials and the estimation of the initial values for the polynomial coefficients which serve as an input for an optimization procedure. These initial values are very important for a successful optimization, which determines the coefficients for the filter polynomials.

5.3 Numerical estimation of the digital filter

The desired transfer function is not given analytically but known at discrete frequency points only. The value of the z-domain transfer function depends in a nonlinear way on its (unknown) polynomial coefficients. This makes it hard to determine the coefficients of the z-domain polynomials directly. Consequently we make use of a numerical optimization technique. Such an optimization requires an initial value estimate. The choice of the initial values can be critical because the problem is nonlinear [222]. For a nonlinear optimization, convergence to a solution depends on the initial values of the coefficients, and it is therefore very important to supply good initial values for the coefficients.
In the following we derive an estimate for the analytical crossover transfer function, which is used to estimate the orders of the z-domain polynomials and an initial value for their coefficients.

5.3.1 Calculation of an initial value for the z-domain polynomials

The crossover filter optimization program [64] yields the transfer function of the filter (loaded with the loudspeaker) at discrete frequency points and a set of filter component values. This information, together with the filter topology, is used to determine an analytical s-domain estimate of the transfer function. This is used to find an initial value for the z-domain transfer function.
The loudspeaker input impedance greatly influences the transfer function of the analog filter and cannot be ignored. An analytical approximation of this input impedance can be derived from the lumped parameter circuit of the loudspeaker, (see Section 1.2.1)

$$Z_{in} = R_e + j\omega L_e + \frac{(Bl)^2}{(R_m + R_r) + j\omega m_t + k_t/(j\omega)}, \tag{5.1}$$

where R_e denotes the DC resistance of the voice coil, ω the angular frequency, L_e the voice coil self-inductance, Bl the force factor, R_m the mechanical damping, R_r the real part of the mechanical radiation impedance, m_t the total moving mass, and k_t the total spring constant. The lumped parameter values can be estimated roughly from the knowledge of whether the loudspeaker is a low-, mid- or high-frequency loudspeaker or, preferably, they can be estimated more precisely from a measured input impedance [144]. The topology and the component values of the crossover filter are

known and we can calculate the transfer function loaded with the loudspeaker in analytical form (see Chapter 6). This analytical representation for the s-domain transfer function is transformed to the z-domain using the bilinear transform; that is, the complex frequency variable s is replaced by

$$s = 2f_s \frac{(1 - z^{-1})}{(1 + z^{-1})}, \tag{5.2}$$

where f_s denotes the sampling frequency. The value of a high order s- or z-domain polynomial is in general very sensitive to a truncation of its coefficients. Also the application of the bilinear transformation to such a polynomial yields large expressions which are not easy to handle. Therefore the polynomials are split into a cascade of first- and second-order sections, which reduce the sensitivity to coefficient truncation considerably. The splitting into lower order sections is performed as follows. The poles and zeros of the s-domain transfer function are calculated. These poles and zeros are either real-valued or complex-conjugate pairs. A real pole and a real zero can be combined to form a first-order section

$$\frac{s - n_1}{s - p_1}, \tag{5.3}$$

where n_1 denotes the zero and p_1 the pole. Application of the bilinear transform Eq. 5.2 to Eq. 5.3 yields

$$\frac{(2f_s - n_1) - (2f_s + n_1)z^{-1}}{(2f_s - p_1) - (2f_s + p_1)z^{-1}}. \tag{5.4}$$

Complex-conjugate or real-valued poles and zeros can be combined to form a second-order section

$$\frac{(s - n_2)(s - n_3)}{(s - p_2)(s - p_3)} \tag{5.5}$$

and application of the bilinear transform yields the second-order z-domain section

$$\frac{a_0 + a_1 z^{-1} + a_2 z^{-2}}{b_0 + b_1 z^{-1} + b_2 z^{-2}}. \tag{5.6}$$

The parameters are given by

$$\begin{aligned}
a_0 &= 4f_s^2 - (n_2 + n_3)2f_s + n_2 n_3 \\
a_1 &= -8f_s^2 + 2n_2 n_3 \\
a_2 &= 4f_s^2 + (n_2 + n_3)2f_s + n_2 n_3 \\
b_0 &= 4f_s^2 - (p_2 + p_3)2f_s + p_2 p_3 \\
b_1 &= -8f_s^2 + 2p_2 p_3 \\
b_2 &= 4f_s^2 + (p_2 + p_3)2f_s + p_2 p_3.
\end{aligned} \tag{5.7}$$

When a (complex) pole and zero are very close in the z-plane and their joint influence is of only minor importance in the total transfer function, that is, a deviation with respect to the 'analog' transfer function within one half of a dB, they are both removed. This reduces the order of the polynomials. The product of the transfer functions of all remaining first- and second-order sections is the z-domain function $H(z)$, which serves as an initial value for the optimization. Having found estimates for both the order and coefficients of $H(z)$ we can proceed with the numerical optimization.

5.3.2 Numerical optimization of the z-domain polynomials

In the previous section we established the order of the z-domain polynomials and the initial value for their coefficients. These are required for the numerical optimization in order to obtain a good set of coefficients, that is, a set that ensures the complex values of the z-domain function $H(z)$ to approach those of the specified function values as closely as possible.
The optimization procedure searches for coefficients in order to minimize the difference, expressed as a least-squares sum, between the specified target function (the analog filter response) and the function to be optimized. The kernel of the actual optimization program is a NAG library [232] routine. It searches for an unconstrained minimum of a least-squares sum, which serves as a measure of fit for a nonlinear optimization. It is important both to provide a reasonable set of initial coefficient values and to match the optimization process to the crossover simulation problem. It is our experience that this greatly improves the numerical optimization. The specific measures are listed below.

1. The transfer function $H(z)$ is expressed as a cascade of second-order sections (each is the ratio of two second-order polynomials) instead of a ratio of two high order polynomials. In this case the function value of $H(z)$ is much less sensitive to a coefficient truncation, which greatly improves the optimization process. This technique is also important in a practical implementation of the filter when the coefficients have to be rounded to a finite wordlength, but even with the much higher wordlength of a mainframe computer this is important too.

2. The frequency points used in the calculation are located equidistantly on a logarithmic frequency scale.

3. The terms in the least-squares sum, the objective function, are weighted with a frequency-dependent weighting function in order to weaken

the influence of points whose amplitude is much lower than in the pass-band. The weighting function equals the magnitude of the specified 'analog' transfer function expressed in dB.

4. All sections are normalized with respect to gain and an overall gain factor or attenuation is defined, which reduces the number of variables to be optimized.

5. The z-domain function $H(z)$ is multiplied by a time delay function. The set of optimization variables is extended with the delay. This is done to deal with possible differences in delay between the analog and the digital filter.

6. An unstable section is rejected by applying a penalty factor to the objective function, that is, the least-squares sum is multiplied with a large number.

5.4 Implementation

Numbers in a digital signal processor are represented by binary words with a limited number of bits. As a consequence, errors are introduced which can seriously affect a practical implementation.
A first error to be mentioned is the rounding of the coefficients from the mainframe computer precision to signal processor precision. The effect is that the transfer functions before and after rounding can differ considerably; in the worst case the filter may even become unstable.
A second error source is the finite precision of the internal calculations of a signal processor, the results of the calculations being rounded. This rounding too, is a nonlinear operation and care must be taken to use the optimal sequence of the calculations.

5.4.1 Poles and zeros distribution

The signal processor used [210] has the following precision

- 12 bits for the coefficients.
- 24 bits for the rounded or truncated results of an arithmetic operation.
- 24 bits for data storage used as delay.
- 40 bits for the results of the accumulator.

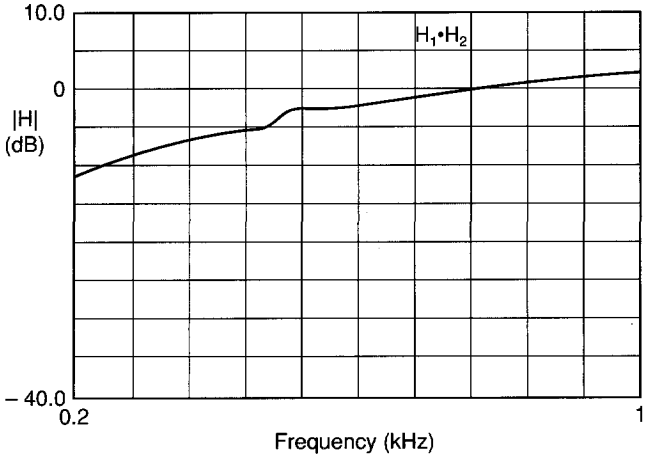

Figure 5.3: Amplitude response of two cascaded second-order sections H_1 and H_2.

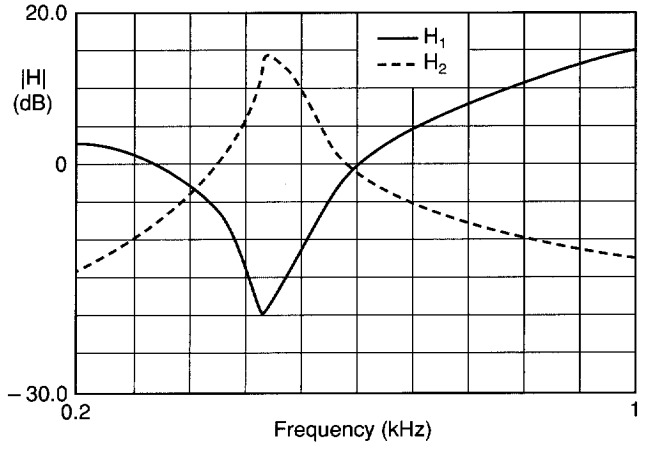

Figure 5.4: Amplitude response of a second-order section H_1 (solid curve) and H_2 (dashed curve).

The results of the calculations needed for one second-order section must be rounded from 40 to 24 bits. Because the number of bits of the accumulator is higher than that of the rounded results, special attention is paid to deciding which poles and zeros have to be paired within one section. Fig. 5.3 and 5.4 depict the effect of a wrong choice for the pairing. It shows the transfer function of a filter that consists of two sections in cascade with individual transfer functions H_1 and H_2, respectively. The total transfer function H_1H_2 in Fig. 5.3 is very smooth, but H_1 (see Fig. 5.4 solid curve)

and H_2 (dashed curve) both have a large quality factor Q. To select a good pole-zero pairing a rather heuristic approach is followed. All the possible combinations to form second-order sections are generated, forming a set S1. The transfer function of each section is calculated and the combination is rejected when the maximum gain exceeds a certain prescribed level, the remainders forming a set S2. A new set S3 is formed containing all the possible combinations of elements out of S2 necessary to construct the desired filter response. Usually S3 contains about 10 elements. A final choice is made out of S3 to form the complete filter. The best choice is the one with the highest signal-to-noise ratio, and the least sensitivity to limit cycles.

5.4.2 Limit cycles

Limit cycles are fluctuations in the output of a digital filter when the input signal is kept constant. In an audio environment this phenomenon must be avoided. At 'digital silence' input, limit cycles are particularly inconvenient, and even under normal music conditions care must be taken to avoid limit cycles. Jackson's triangle [157] shows a safe area in which no limit cycles occur. As mentioned in paragraph 5.4.1, the avoidance of limit cycles is one of the considerations for making the final choice out of set S3.

5.4.3 Scaling

When the signal level at the input of the filter is high, overflow may occur at any point in the filter. At places with a high Q, appropriate scaling must be applied to avoid overflow. However the drive level in the filter must be as large as possible to obtain the maximum signal-to-noise ratio.

5.4.4 A case study

As a vehicle we chose a three-way loudspeaker system, equipped with a passive radiator. First the box was measured in an anechoic room to obtain the pressure response of each individual loudspeaker (mounted in the box). The power responses were measured in a reverberation room, and finally the impedances of the loudspeakers were measured. All the data were measured using a digital two-channel signal analyzer. With the aid of the (analog) crossover filter optimization package [144], a number of crossover filter sets were calculated. For different initial values different loudspeaker polarities and different weighting values for the target functions, multiple solutions were found. From these solutions, a selection was

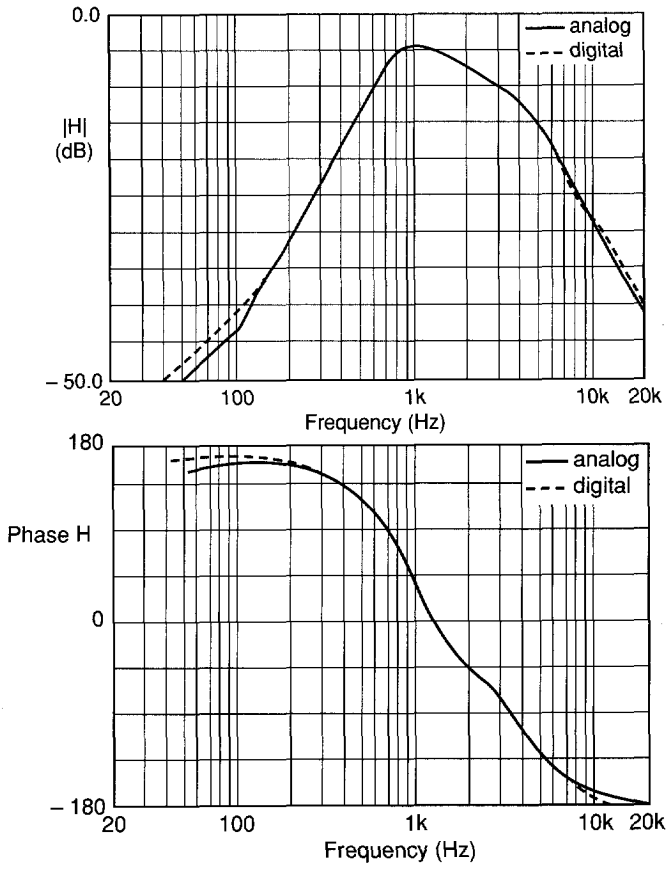

Figure 5.5: Upper panel:(a) Amplitude responses. Lower panel:(b) Phase responses. Solid curve- mid-range crossover filter loaded with the driver (the target function); dashed curves-$H(z)$ with truncated coefficients.

made for actual listening tests. In this case study we will focus on the squawker part of a particular set. The amplitude and phase of the transfer function of this filter, loaded with the loudspeaker, are shown in Fig. 5.5 (solid curves). This complex-valued analog transfer function is provided by the crossover filter optimization program and it is the target function for the optimization of the z-domain transfer function. This optimization actually minimizes the differences between the analog target function and the z-domain transfer function. The output of the optimization is a z-domain transfer function which fits the desired (analog) transfer function as well as possible. However, the coefficients of the z-domain function are represented in 12 bits in the DSP hardware, which is much less accurate

than in the mainframe computer. The coefficient truncation can have a considerable influence on the transfer function. Depending on different initial values, number of points in the optimization, etc., one finds multiple solutions that simulate the same (analog) transfer function. However, all these solutions exhibit large differences for coefficient truncation. To evaluate the influence of the 12 bits coefficient representation or coefficient truncation, we calculate the transfer function with truncated coefficients, for two different sets of coefficients. Fig. 5.6 shows that the influence of

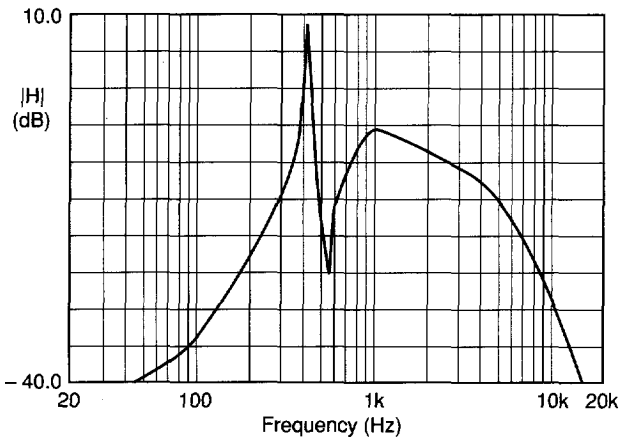

Figure 5.6: Amplitude response of the transfer function with truncated coefficients.

the coefficient truncation can be rather dramatic, but the solution (set of z-domain function coefficients) shown in Fig. 5.5 (dashed curve) is hardly affected by the coefficient truncation.

5.5 Conclusions

It has been shown that a loudspeaker crossover filter can be simulated by a digital signal processor. This technique can be used for evaluating both passive and active crossover designs. The combination of a numerical optimization of the crossover filters and a digital simulating device is a flexible and powerful design tool, which is being used for the design of new series of commercially available loudspeaker systems.

Chapter 6

Analysis of loudspeaker filters [1]

A symbolic analysis program for RLC-loudspeaker-crossover filters is designed and implemented using computer algebra. The program generates Pascal code that can be used to compute the input impedance and the voltage transfer function of the filter as a function of frequency, filter components and loudspeaker impedance as is discussed in the preceding chapter.

6.1 Introduction

In this chapter is discussed how computer algebra can be used to compute the symbolic transfer function of a filter from the topological description of the filter only.

Using this transfer function (Pascal) code can be generated, which can be used to compute the input impedance and the voltage transfer function of the filter as a function of frequency, filter components and loudspeaker impedance.

6.2 The problem

Using a software package previously made for the numerical optimization of a multi-way loudspeaker system crossover filter (FOPT) [64], a set of component values is given for the crossover filter that satisfies the designer's demands and constraints to the greatest extent possible. The inputs for

[1] Parts of this chapter have been published in [12].

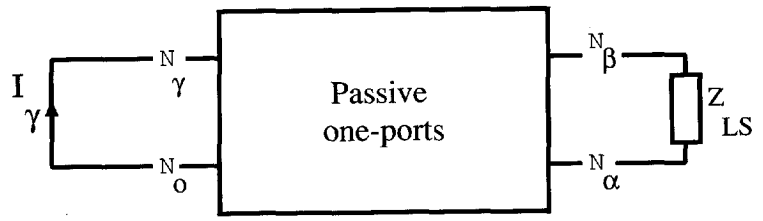

Figure 6.1: General two-port network.

FOPT are loudspeaker data (e.g. loudspeaker characteristics), target data (e.g. desired sound pressure response and power response) and a Pascal procedure, which is used to compute the input impedance and the voltage transfer function of the filter as a function of frequency f, filter components $\mathbf{C}$ and the loudspeaker impedance Z_{LS}.

The numerical analysis of the filter is done by the user, i.e., the user must write the Pascal code based on the topology of the filter. The symbolic analysis of the filter is also performed by the user and results in the Laplace domain coefficients of the voltage transfer function as a function of filter components $\mathbf{C}$ and the lumped-element parameters $\mathbf{L}$ of the s-domain model used to represent the loudspeaker.

The drawbacks of this program are clear: there is user intervention required for both the numerical and the symbolic analysis on a case by case basis. The (Pascal) code to compute the numerical transfer function can be computed from the symbolic transfer function of the RLC-filter. Therefore, the main task is to use computer algebra to compute the symbolic transfer function of a filter.

6.3 The mathematical model

We will restrict ourselves to linear time-invariant two-port networks that consist of resistance R, conductance G, inductor L and capacitor C elements only. Fig. 6.1 shows a general two-port network. It is driven by an ideal unity current source I_γ and loaded with a complex impedance Z_{LS}. The network consists of one-ports only. The nodes n_i are numbered from zero to N. The input nodes are n_0 and n_γ. The output nodes are n_α and n_β. They are referred to as I/O-nodes. For the network of Fig. 6.1 with node n_0 grounded and a unity current source, the input impedance is defined as

$$Z_\gamma(s,x) = V_\gamma(s,x), \tag{6.1}$$

where s is the frequency variable in the Laplace domain, x is the parameter vector containing load and network parameters and V_γ is the input node voltage with respect to the ground. The voltage transfer function for the network of Fig. 6.1 is given by

$$H_{\beta\alpha} = \frac{V_\beta(s,x) - V_\alpha(s,x)}{V_\gamma(s,x)}, \tag{6.2}$$

where n_0, n_γ, n_α and n_β belong to $\{0, 1, \ldots, N\}$ such that n_0, n_γ are input and n_α, n_β are output nodes and V_i is the node voltage with respect to the ground.

The indefinite nodal method (see [180]) provides a set of non-singular linear equations from which all node voltages can be solved. Let $Y = (y_{ij})$ be a matrix such that

- The diagonal entries of Y are given by $y_{jj} = \sum$ admittances connected to n_j.
- The off diagonal entries of Y are given by $y_{jk} = -\sum$ admittances connected to n_j and n_k.

The indefinite nodal equation is defined as follows

$$YV = J,$$

where $V = (V_0, \ldots, V_N)^T$ denotes the node voltage vector and $J = (J_0, \ldots, J_N)$,

$$J_j = \sum \text{ currents from independent sources entering } n_j.$$

A two-port network consisting of one-ports only has a symmetric admittance matrix ($Y = Y^T$) and an admittance matrix of a one-port connected between nodes k and l has entries $y_{kk} = y_{ll} = Y$ and $y_{kl} = y_{lk} = -Y$, where Y denotes the admittance connected to k. To transform the indefinite nodal equation into a linear nonsingular system that can be solved, we shall use that grounding the k^{th} node is equal to deleting the k^{th} row and column of Y.

6.3.1 A loudspeaker model

A loudspeaker can be modeled by a lumped-element network see Section 1.2.1. The impedance of a lumped-element model is given by Eq. 1.2.

Given the symbolic transfer function $H(s)$ of the corresponding filter, we make the following transformation

$$s = 2f_s \frac{(1 - z^{-1})}{(1 + z^{-1})}, \tag{6.3}$$

where f_s denotes the sampling frequency. Once the z-domain transfer function is known it can be made suitable for implementation on a Digital Signal Processor. It is commonly known [157] that high order direct form digital filters have bad coefficient sensitivity properties. For this reason the high order transfer function is divided into a cascade of second order sections (biquads).

6.4 The mathematics

To minimize the number of node voltages and network parameters, we shall use the following reduction rules [109]:

- The P-reduction rule. In a two-port network M parallel one-port connections can be replaced by one new one-port with new admittance Y_M given by the sum of the old admittances

$$y_M = \sum_{k=1}^{M} y_k. \tag{6.4}$$

- The S_M-reduction rule. An internal node voltage can be removed from the nodal equation by a star to maze transformation. If a star shape with $M + 1$ nodes and M branches is transformed into a maze structure with M nodes and $M(M - 1)/2$ branches, the number of network branches increases by $M(M - 3)/2$, but the number of node voltages is reduced by one. If the center node is n_l, then the new admittances are given by

$$y_{ij} = \frac{y_{il} y_{jl}}{\sum_{k=1}^{M} y_{kl}}. \tag{6.5}$$

A star to maze transformation can introduce new parallel one-ports. Remark that, if the reduction of a star structure with $M \geq 2$ branches into a maze structure, is followed by a parallel reduction, then the resulting structure has precisely M new star shaped structures with at least $M - 1$ branches. This implies that a $P * S_M$ reduction should be followed by a

$P * S_{M-1}$ reduction if $M > 2$ and by $P * S_M$ reduction if $M = 2$. This yields the following reduction algorithm that eliminates the internal node voltages.

```
M := 2;
apply P reduction rule
while 'there are internal node voltages'
do
   apply P*S_M reduction rule;
   if M >2 then
      M := M-1;
   else
      M := M+1;
   fi
od
```

After the algorithm comes to a stop, the (definite) nodal equation can be solved and we obtain the symbolic voltage transfer function $H_{\beta\alpha}$ for the network.

6.5 The computer algebra involvement

The above described method to compute the symbolic transfer function of linear time-invariant two-port networks is implemented using the REDUCE computer algebra system.

The program is called netprogram and has two main functions

- From the filter specification, the program generates Pascal code that can be used to calculate the input impedance and symbolic s-domain voltage transfer function of the filter;
- From the filter component data, lumped-elements parameters of the loudspeaker and the symbolic s-domain voltage transfer function, the program generates the z-domain biquad coefficients that can be used as a start value for the curve-fitting optimization package CURFIT.

6.6 Interpretation of the results

In this section we illustrate the results with a simple example. Consider a low-pass loudspeaker crossover filter (Fig. 6.2), where all one-ports are represented as admittances. Let

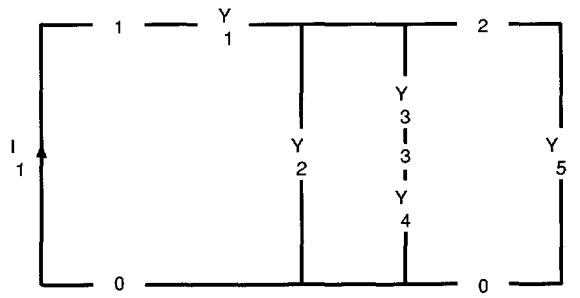

Figure 6.2: A low-pass crossover filter loaded with a loudspeaker.

$$\begin{aligned} Y_1 &= 1/(sL_1 + R_{L_1}) \\ Y_2 &= sC_1 \\ Y_3 &= sC_2 \\ Y_4 &= 1/R_1 \\ Y_5 &= 1/Z_{LS}. \end{aligned} \tag{6.6}$$

Using the definitions we find $\gamma = 1$, $\alpha = 0$, $\beta = 2$ and

$$x = (L_1, R_{L_1}, C_1, C_2, R_1, Z_{LS})^T.$$

The indefinite nodal equation is given by

$$\begin{pmatrix} Y_2+Y_4+Y_5 & 0 & -Y_2-Y_5 & -Y_4 \\ 0 & Y_1 & -Y_1 & 0 \\ -Y_2-Y_5 & -Y_1 & Y_1+Y_2+Y_3+Y_5 & -Y_3 \\ -Y_4 & 0 & -Y_3 & Y_3+Y_4 \end{pmatrix} \begin{pmatrix} V_0 \\ V_1 \\ V_2 \\ V_3 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

Minimization of the indefinite nodal equation using the reduction algorithm proceeds as follows. First the P-cycle will detect a parallel port. A new admittance element is created, say Y_6 and the nodal equation becomes

$$\begin{pmatrix} Y_4+Y_6 & 0 & -Y_6 & -Y_4 \\ 0 & Y_1 & -Y_1 & 0 \\ -Y_6 & -Y_1 & Y_1+Y_3+Y_6 & -Y_3 \\ -Y_4 & 0 & -Y_3 & Y_3+Y_4 \end{pmatrix} \begin{pmatrix} V_0 \\ V_1 \\ V_2 \\ V_3 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

Next the $P * S_2$ step will detect the internal node voltages V_3 as candidate for minimizing and the nodal equation becomes

$$\begin{pmatrix} Y_8 & 0 & -Y_8 \\ 0 & Y_1 & -Y_1 \\ -Y_8 & -Y_1 & Y_1+Y_8 \end{pmatrix} \begin{pmatrix} V_0 \\ V_1 \\ V_2 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix},$$

where

$$\begin{aligned} Y_6 &= Y_2 + Y_5 \\ Y_7 &= Y_3 Y_4/(Y_3+Y_4) \\ Y_8 &= Y_6 + Y_7. \end{aligned} \tag{6.7}$$

There are no internal nodes nor parallel structures left and the algorithm has come to a stop. To solve the resulting nodal equation, we ground n_0 and then solve the remaining system.

The transfer functions of the filter are

$$Z_1 = V_1/I_1 = \frac{Y_1+Y_8}{Y_1 Y_8} \tag{6.8}$$

$$H_{20} = V_2/V_1 = \frac{Y_1}{Y_1+Y_8}. \tag{6.9}$$

Substituting the values of Y_i and the numerical value of x,

$$x = (1,1,1,1,1,Z)^T,$$

we obtain the transfer function as a function of s and the impedance $Z_{LS} = Z$:

$$H_{20}(s,Z) = \frac{Z}{1+Z+(1+2Z)s+Zs^2}.$$

6.7 Conclusion and discussion

The use of computer algebra enables us to write a program that computes the symbolic transfer function of a filter based on the topology of the filter. Furthermore, given the symbolic transfer function, it is easy to compute the numerical transfer function. The computer algebra program can be used to generate the (Pascal) code, which can be used to calculate the voltage transfer function and input impedance of a filter directly.

Chapter 7

Crossover filters design [1]

A new method is presented for the design and evaluation of loudspeaker crossover filters. The desired system characteristic can be prescribed by a (complex) acoustic transfer function rather than an electrical one only. It may be derived from conventional filters or based on a measured one from a reference (favorable) system. Double blind listening tests are performed to verify subjectively the similarity between the reference system and its experimental counterpart. The drivers of the experimental loudspeaker are preceded by digital filters, enabling the imitation of several different favorable loudspeakers. Multidimensional scaling techniques are applied to represent the results of the listening tests. These results affirm the strength of the design method.

7.1 Introduction

The crossover or frequency-dividing network plays an important role in the performance of high-quality loudspeaker systems. Individual drivers do not have the capability of reproducing all frequencies in the entire audio range with sufficient directional properties. The crossover network allows various drivers, each suited to a particular frequency range, to be combined into a system covering a wide frequency range. Because of its importance, since the early paper by Hilliard [103], a great deal has been published on this subject during the past fifty years. See [32, 39] for an overview of the many designs available. The design of crossover networks can be divided into three approaches.

[1] Parts of this chapter have been published in [2, 3].

- The first approach is to use one of the well-known filter families as crossover filter, e.g. the Butterworth types or the celebrated Linkwitz-Riley [127, 128] filters. Other possible arrangements are the delay-derived high slope filters [129], the constant voltage filters [181], all-pass systems [83], magnitude complementary filters [160] and the filler-driver design [16, 132]. In contrast to the other filters, the filler-driver is more than a filter only. It uses an additional driver.

  However, most analyses are concerned with the electrical part of the filter, thereby ignoring the acoustic behavior of the drivers and the cabinet. Usually, the drivers are considered to be ideal; the peaks and dips in the operating frequency range are ignored, as is the frequency dependence of the electrical input impedance. The design of a crossover network can be a difficult and tedious affair when these imperfections are taken into account.

- The second approach is by numerical optimization [14, 64, 173, 206, 220]. These programs have proved their usefulness in the design of crossover networks when the imperfections mentioned are taken into account. Filters known from network theory or existing crossover designs are used as a first approximation or seed value for the numerical algorithm. During the optimization the filter component values are determined in such a way that the difference between the desired and the calculated acoustic response of the loudspeaker system is minimized.

- The third approach, which will be the subject of this chapter, uses a real existing or a hypothetical (favorable) system as a reference to supply a target function for the system under design. Each driver plus crossover filter of that reference system acts individually as an electro-acoustic target function for the crossover filter to be optimized. The optimization process yields a set of coefficients for the digital filters preceding each driver of the experimental loudspeaker. By only changing the coefficients of these digital filters, it is possible to let one and the same experimental loudspeaker sound very similar to different reference systems.

In the following we will discuss the design of these digital filters and present examples with two different reference systems. The similarity between the reference system and its experimental counterpart is verified both by objective and subjective means. The latter is performed by listen-

ing tests using triadic comparisons. The results are represented using multidimensional scaling techniques.

7.2 Aim

In Chapter 5 and [11] a tool was described for the simulation of real existing or hypothetical crossover filters. It consists of a number of computer programs and a digital signal processor. A set of filters was calculated by a computer program [64] that performs the (non-linear) optimization of several functions. These functions are usually the sound pressure response and the sound power response. Listening experiments in which the subjects could select from these filters underscored again the powerful influence of crossover filters. With one and the same system, i.e. the cabinet and three drivers, very different acoustical images could be effected. After listening tests a final choice was made for the construction of the analog filter. Another possible approach is to use a real existing loudspeaker set as an object function. The aim is to make the experimental system response equal to that of the reference system. A simple method would be to place an equalizer before the experimental system, and adjust it in such a way that both the reference and experimental systems will give the same response at a certain field point. One step further is to equalize each driver instead of the whole system. One could consider this method as a combination of an equalizer and crossover filter. In the new case each driver has its 'own' equalizer, thereby giving the possibility of varying the directional behavior with only a minimal effect on the direct or on-axis sound pressure response. In order to have the same acoustical image for both the reference and the experimental system, the crossovers of the experimental box have to be chosen in such a way that the total transfer function (acoustic and electric) of each driver is the same as that of the corresponding reference drivers. Let the acoustical reference transfer function of the jth driver be defined by

$$H_j^a(s, \mathbf{x}) = \frac{p_j(s, \mathbf{x})}{V_j^t(s)} \tag{7.1}$$

where p_j is the sound pressure level at position $\mathbf{x}$, V_j^t is the voltage across the terminals of the jth driver and s is the complex frequency variable. The electrical transfer function of the jth filter is defined by

$$H_j^e(s) = \frac{V_j^t(s)}{V_{in}(s)} \tag{7.2}$$

where V_{in} is the input voltage of the whole system. Then the total driver response for the jth driver yields

$$H_j^t(s, \mathbf{x}) = H_j^a(s, \mathbf{x}) H_j^e(s), \tag{7.3}$$

and, assuming linear responses, the total system response for n drivers yields

$$H_t(s, \mathbf{x}) = \sum_{j=1}^{n} H_j^a(s, \mathbf{x}) H_j^e(s). \tag{7.4}$$

The intention is to make the experimental system response $\hat{H}_j^e(s, \mathbf{x})$ equal to that of the reference system. This can be accomplished if the drivers of the experimental system are preceded by a filter having

$$\hat{H}_j^e(s, \mathbf{x}) = \frac{H_j^t(s, \mathbf{x})}{\hat{H}_j^a(s, \mathbf{x})} \tag{7.5}$$

as transfer function. As Eq. 7.5 shows, $\hat{H}_j^e(s)$ depends on $\mathbf{x}$. To cancel this dependence we require the physical position of the drivers on the baffle and the size of the drivers and the cabinet to be comparable for the reference system and the experimental counterpart. If this requirement is satisfied we assume that the radiational behavior (polar response) of both systems have been made similar when the responses at one particular position are equal.

The filter $\hat{H}_j^e(s, \mathbf{x})$ is approximated for a chosen $\mathbf{x} = \mathbf{x_0}$ by a digital filter $H_j^d(s)$. The coefficients $\mathbf{c}$ of that filter have to be calculated by means of an optimization procedure i.e. by minimizing the functional

$$\epsilon'_j(\mathbf{c}) = \int_{\omega_b}^{\omega_e} |\hat{H}_j^e(i\omega, \mathbf{x} = \mathbf{x}_0) - H_j^d(i\omega)|^2 d\omega \tag{7.6}$$

where ω_b and ω_e are the begin and end points of the frequency band of interest. In practice, however, the reference transfer functions are obtained by measurement and are therefore known at r discrete frequencies only. This replaces the functional in a form

$$\epsilon_j(\mathbf{c}) = \sum_{k=1}^{k=r} W_k |\hat{H}_j^e(i\omega_k, \mathbf{x} = \mathbf{x}_0) - H_j^d(i\omega_k)|^2 \tag{7.7}$$

where W_k are (frequency dependent) weighting factors. The coefficients of the digital filter are calculated by means of an optimization procedure based on a modified Gauss-Newton method [86]. As the variables in Eq. 7.7 are complex, both amplitude and phase are taken into account.

7.3 Design examples

In order to test the practical outcome of the design method discussed, two different brands of 3-way loudspeaker systems were chosen as a vehicle. These systems are comparable with each other with respect to the cabinet volume, the driver size and the commercial price. However, their sound was very different. This is due to a different on-axis (anechoic chamber) response, as well as a different polar response. Due to the latter, the final perceived sound depends on the listening environment. First we will consider the anechoic chamber response, and then a normal listening environment response.

7.3.1 Anechoic chamber response

For each driver of both systems, labeled brand A and B respectively, the sound pressure responses were measured at a distance of 3 m right in front of a point between the tweeter and squawker. The total system responses were determined too, and are shown in Fig. 7.1-a (solid curve) for system A and in Fig. 7.1-b (solid curve) for system B. The optimization procedure as discussed in the previous section yields a set of coefficients A_a if system A is used as a target, and a set B_a if system B is used as a target. The subscript a denotes the anechoic chamber environment. With the experimental set-up depicted in Fig. 7.2, the transfer functions were calculated for the digital filters loaded with either set A_a or B_a. The results are plotted in Fig. 7.3. Fig. 7.3-a gives the responses of the filters for the woofer, squawker and tweeter using set A_a, Fig. 7.3-b the responses using set B_a. As these figures show, the SPL responses are quite different from the responses known from the classic designs.

Each filter is placed before the appropriate driver of the experimental system. The sound pressure of the total system as well as that of the individual drivers were measured under the same conditions as those of the target systems. The responses are drawn in Fig. 7.1-a (dashed curve) using set A_a, and Fig. 7.1-b (dashed curve) using set B_a. It should be emphasized that both (dashed curve) responses were measured from one and the same experimental loudspeaker system, only the coefficients of the filters being changed. As the figures show, the responses of target and experimental system are very similar. But even more important, this also holds for the corresponding individual drivers, both for the magnitude and phase. The phase equality is crucial because the summation in Eq. 7.4 can be very sensitive to phase errors. Although Fig. 7.1-a and b show

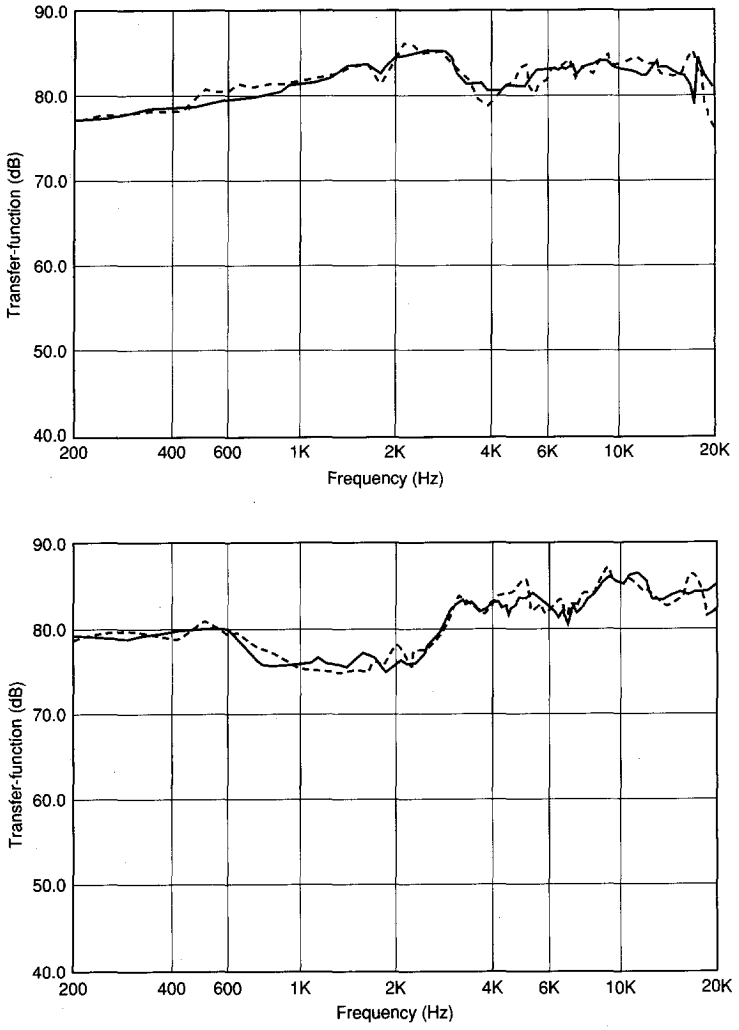

Figure 7.1: SPL of system A and B. Upper panel:(a) system A and Lower panel:(b) system B (solid). Dashed lines in both panels represent the performance of the simulated systems A and B through the experimental system.

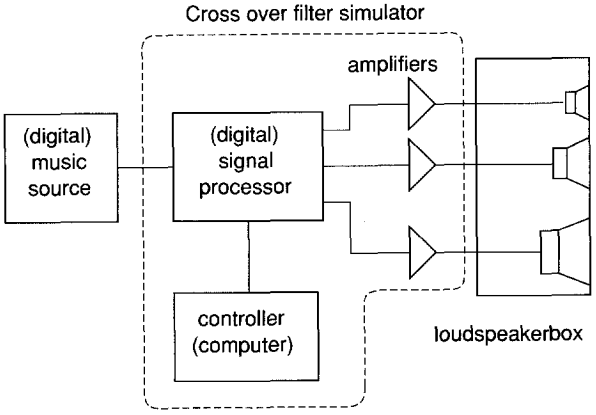

Figure 7.2: Experimental set-up.

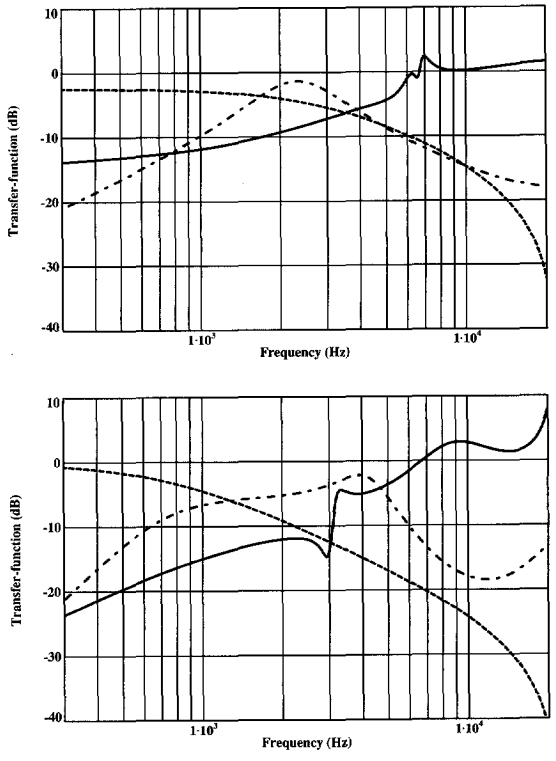

Figure 7.3: Filter responses of two coefficient sets. Upper panel:(a) using set A_a. Lower panel:(b) Filter responses using set B_a.

good agreement, subjective comparisons were also made and these will be discussed in Section 7.4.

7.3.2 Listening room response

As pointed out by several others, the total acoustic power of a loudspeaker system is very important, especially in relation to crossover designs [190, 211]. Systems with the same on-axis anechoic response can sound quite different in a living room. Staffeldt [183] noticed the resemblance between the average living room response and the total acoustic power response measured e.g. in a reverberation room. Because of the important influence on the response exerted by the listening environment, the response of system A was determined in a listening room (V = 210 m^3 and at 500 Hz, T_{60} = 0.4 s). Target system A was placed on the ground. With a microphone at 3 m distance and at 1 m height, the responses of the individual drivers (preceded by their filters) and that of the total system were measured. Thereafter, the responses of the experimental system were measured under the same conditions. Following the same procedure as described in the previous section, a set of coefficients A_l was determined. The subscript l denotes the listening room as measuring environment. Due to the acoustic modes in the listening room, the measured transfer functions exhibit at high frequencies a rather high variance. To compare the target and experimental loudspeaker responses, some spatial averaging is applied. The same microphone used to measure the target functions was placed on a rotating boom. The orbit of the microphone described a circle of 25 cm diameter parallel to the floor at 1 m height. The spectra obtained in this manner are given in Fig. 7.4. The solid curve is the measured response of the target system A, while the chain-dotted one represents the response for the experimental system using filters with coefficient set A_l. As the figures show, there is a good similarity between both systems (the erratic behavior is mainly due to the room modes). The corresponding driver responses display a good agreement too. This is shown in Fig. 7.5 for the squawker response of the target A (solid curve) and the experimental one (dotted curve). As noted above, to affirm the similarity between both systems listening tests were carried out, and these will be discussed in the following.

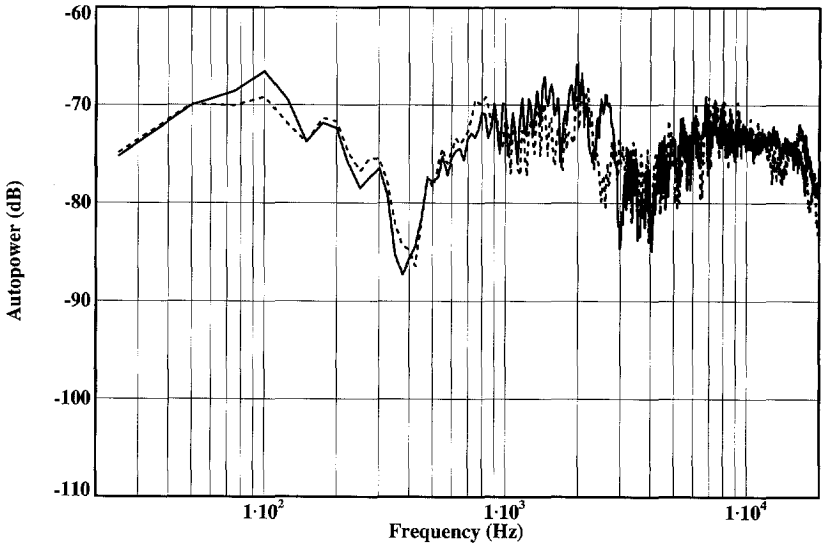

Figure 7.4: SPL of system A (solid) and experimental system (dotted).

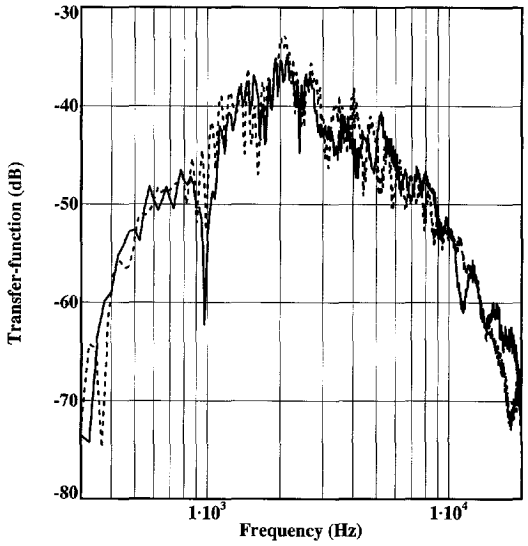

Figure 7.5: Squawker SPL of system A (solid) and experimental system (dotted).

7.4 Subjective evaluation

At present subjective evaluation of loudspeaker systems is still a crucial part of the design process. No good criteria are yet available for translating objective measurements into a prediction of the perceived sound. Although several others [81, 130, 174, 183, 197, 199–201] have tried this, and achieved some success, listening remains important and is fortunately not the most boring part of the design procedure.

The aim was to make the experimental system similar to several other reference systems by only changing certain filter parameters. An obvious method to verify the result would be to place the reference and experimental systems next to each other, and subsequently to listen to them to discern whether they are indistinguishable, or, if they are not, how much they differ. However, an insuperable problem arises, as follows. If one switches between the two systems, it is easy to identify both systems because of the readily perceived change in the direction of the sound. A fair comparison would only be possible if the systems to be compared were placed in the same physical position. This makes a fast comparison, which is considered to be very important, impossible. We agree with Shanefield's opinion [174]: 'Fast comparisons provide sensitivity to real differences, and long-term listening simply encourages imaginary ones. But both should be tried.' Yet fast comparisons are possible by listening not directly to the systems, but to recordings of them, as will be discussed in the next sections.

7.4.1 First experiment

Four (mono) music pieces, with a duration of two minutes each, as well as some test signals, were successively recorded on track one of an eight-track tape recorder. On the other tracks recordings were made in an anechoic chamber of the various systems described in Section 7.1. One loudspeaker system at a time was present in the anechoic chamber, so that each system occupied the same place during the recordings. The microphone and cabinet position for the systems under test were identical to those described in Section 7.3.1. To have some additional reference material, recordings were made of a third brand-name reference loudspeaker labeled C, and of the experimental loudspeaker equipped with filters using coefficient set D_a. The latter was intended to imitate a hypothetical ideal system, having a flat pressure response and a gradually decreasing (acoustic) power response. The first track of the tape recorder was used as source for the other six recordings. At each recording session the tape was rewound, and because

Table 7.1: Track allocation for the first experiment.

track	
1	direct recording from a Compact Disc player
2	experimental system using coefficient set D_a
3	system C
4	experimental system using coefficient set A_a
5	system A
6	experimental system using coefficient set B_a
7	system B

the recorder has a combined playback and record head, each track can be played back synchronously with the others. Finally, it was then possible, to listen to each system, and switch instantaneously between them. The tape can be played back in any environment, without the need of an anechoic environment, using high quality loudspeaker systems or headphones.

Aim

This experiment used similarity judgements between music pieces reproduced by different reference loudspeaker systems and their experimental counterparts. Multidimensional scaling was used, to examine how many subjective dimensions are needed to describe the perceptual space, and to scale the measured perceptual differences.

Method

The method of triadic comparisons (see Appendix C) was used to obtain perceptual distances between the stimuli. One can form 35 different triads from 7 stimuli. Presentation of all triads would have loaded the subjects too much. Therefore a subset of 14 triads was taken, which still presented all loudspeaker systems an equal number of times to the subjects. Such a subset is a Balanced Incomplete Block Design (see Appendix B). Twelve subjects participated in the experiment. All stimuli were presented by headphones under diotic (both ears the same signal) conditions. Using the experimental set-up as depicted in Fig. 7.6, a triad was selected and remained fixed as long as the subject wished. The triad components were always labeled 1,2 and 3. The subjects did not know the relationship between these numbers and the real channel numbers of the tape recorder. The subjects could freely select one out of the three and switch between them as

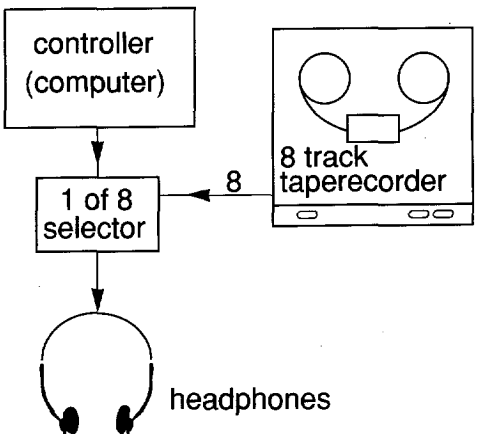

Figure 7.6: Experimental set-up for the listening tests.

much as they wanted. After answers were given to the questions: 'Which are the most similar?' and 'which are the least similar pairs of the three ?', the test was stopped and a new triad was chosen. This was repeated until all 14 triads had been considered. In the meanwhile the tape recorder played the music pieces sequentially, and then rewound automatically. The stimuli were presented through high-quality headphones. The subjects could adjust the playback level to their own preference, but they were allowed to do so only once. A session of 14 triads took approximately 40 minutes.

Results

The experiment resulted in a triangular similarity matrix for each subject. In order to generate a more graded matrix, the similarity indices of all the subjects were gathered and normalized, the normalization factor being the total number of occurrences of that pair. So the maximum of an entry is two. The results are listed in Table 7.2. Consequently, as Table 7.2 shows, there was a consensus between all the subjects for the pair 6,7. The multidimensional scaling program KYST2a was used to scale these data (see Appendix C). The program was used in a non-metrical mode, using a normal Euclidian space (Minkowski parameter = 2). The program calculated, with Stress1 = 0.008 in two dimensions, a configuration as depicted in Fig. 7.7. The Stress1 value for one dimension was 0.231. Using higher dimensions than two did not give a significant improvement, indicating that a two-dimensional representation describes the perceptual space

Table 7.2: Normalized similarity half-matrix for all the 12 subjects in the first experiment.

1	-						
2	1.6	-					
3	1.9	1.6	-				
4	0.4	1.2	1.1	-			
5	1.2	1.1	0.8	1.9	-		
6	0.4	1.1	0.1	0.9	0.9	-	
7	0.4	1.0	0.6	0.3	0.6	2.0	-
	1	2	3	4	5	6	7

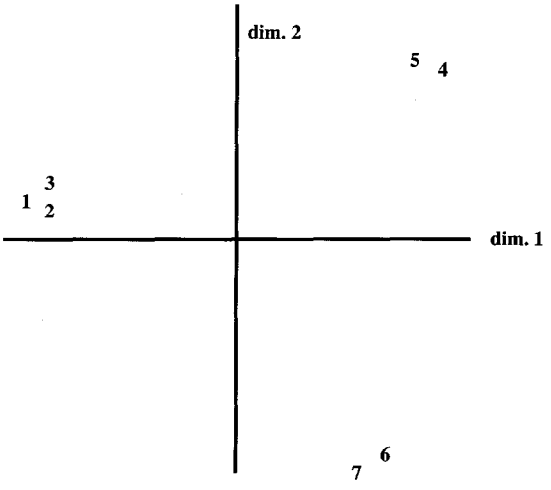

Figure 7.7: Perceptual configuration, (see Table 7.1 for the legend).

Table 7.3: Track allocation for the second experiment.

track	
1	direct recording from a Compact Disc player
2	system A, pos. M_2
3	experimental system coeff. set A_l, position M_2
4	system A, pos. M_3
5	experimental system coeff. set A_l, position M_3

very well. The labels in the figure correspond with the channel numbers of the tape recorder as in Table 7.1. From Fig. 7.7 three clusters can be distinguished. The similarity between system A and its experimental counterpart, and similarly system B and the experimental system, is obvious. The third cluster led to the conclusion to denote the experimental system using coefficient set D_a as a true High-Fidelity system: it is very similar to the original Compact Disc music.

7.4.2 Second experiment

In the first experiment only one microphone position in an anechoic chamber was used to record the music on tape. In order to examine the influence of a deviation of the microphone position with respect to the position used to calculate the filters, some other recordings were made with the microphone displaced. According to the theory of Section 7.2 only a minor influence should be observed. Another difference compared with the first experiment concerns the recording room; this time a listening environment was used instead of an anechoic chamber.

One loudspeaker system at a time was present in the listening room, so that each system occupied the same place during the recordings. The others were out of the room to avoid them disturbing the soundfield. The cabinet position for the systems under test was identical to that described in Section 7.3.2, but the microphone position was changed to positions M2 and M3. Position M2 was right in front of the system at a distance of 4 m, M3 was at the same distance but with a deviation of 25°. As described in Section 7.4.1, track one of the tape recorder was used as the source for the recordings made according to Table 7.3. This experiment enabled the subject, while listening to a system in a particular room, to move from a chair at position M2 to one at position M3 instantaneously

Table 7.4: Normalized similarity half-matrix for all the 15 subjects in the second experiment.

2	-			
3	1.9	-		
4	0.7	0.4	-	
5	0.6	0.5	1.8	-
	2	3	4	5

without breaking arms or legs, which probably would have been the result of such a jump. As in the first experiment, triadic comparisons were used. Track one of the tape was not included this time. This means that there were four possible triads. In order to increase the number of triads each triad was repeated three times, each time with a different order within the triad, yielding a total of 12 triads. Fifteen subjects participated in this experiment. They used on average 35 minutes per session.

Results

The similarity matrices obtained from the subjects were pooled to form one matrix, and normalized as in the previous experiment. The result is shown in Table 7.4. Again the MDS program was used in a non-metrical mode, using a normal Euclidian space to calculate a configuration in two dimensions out of this matrix. The result is depicted in Fig. 7.8. The labels in the figure correspond with the channel numbers of the tape recorder as shown in Table 7.3. The similarity between system A and the experimental counterpart, both at position M2, is obvious; the same holds for both systems at position M3. The perceptual distance for changing the microphone position while using the same system is larger than the distance for changing to the other system while keeping the microphone position fixed. So we conclude that a deviation of the listening (microphone) position does not influence the similarity.

7.5 Conclusions

A new technique is presented for the design of crossover filters. It is shown to be a powerful and versatile one. With these filters preceding (non-ideal) drivers, the acoustical (on-axis and polar) responses can be made

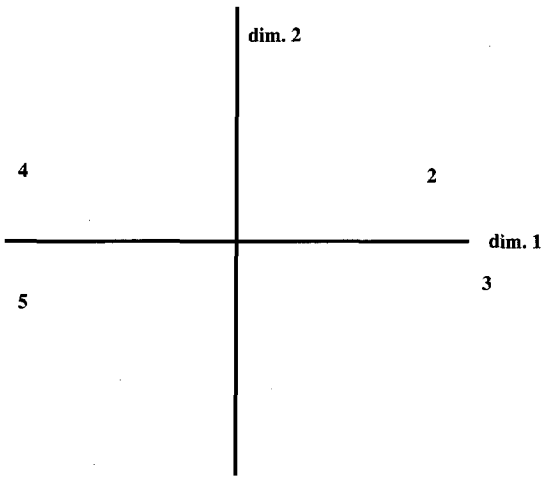

Figure 7.8: Perceptual configuration (see Table 7.3 for the legend).

similar to prescribed ones. The prescribed responses can be obtained from a hypothetical system or from the measurement of an existing one. The similarity holds not only for the total system but also for the individual drivers. Listening tests, using triadic comparisons, were used to affirm the claimed similarity. The influence of extraneous factors was excluded as far as possible by using a special listening arrangement. This involved the subjects not listening directly to the loudspeaker systems, but to tape recordings of them. Multidimensional scaling, shown to be a very useful technique, was applied to represent the results of the listening tests.

Chapter 8

Enlarging the listening area for stereophony by Time-intensity trading [1]

The correction to compensate for the degradation of the stereophonic illusion due to off-center listening is investigated, both for a small and a wide loudspeaker base. Data obtained by psychoacoustic experiments are derived, which can be used to establish electronic filters driving a loudspeaker array, such that the required correction is achieved. The experimental results for video and (multimedia-) computer monitor applications are compared with the case of a much wider loudspeaker base, as in audio-only applications. The outcome of these comparisons shows that the optimal radiation patterns for audio and video are in general incompatible.

8.1 Introduction

Since the first demonstration of an electro-acoustic reproduction system [228] with the aim of achieving a realistic sound reproduction, increasing efforts have been directed at improving the stereophonic effect [18, 29, 59] and the description of it [26, 36, 61, 69, 74]. Generally, it is considered as a serious artefact of the present stereo system that the listener is aware of the stereophonic illusion only within a limited region. If the head is moved away from the perpendicular on the center of the base of the loudspeakers, then

[1] Parts of this chapter have been published in [7, 8].

the stereo effect decreases. This off-center listening problem becomes even more serious when the distance between the loudspeakers is not very large in comparison with the deviations from the center position, as in multimedia PC monitor and (HD)TV applications; the latter normally has a wider base but in some cases a smaller base is desired. In the following experiments are described with the aim of investigating the influence of off-center listening for normal stereo set-ups and for small-base stereo set-ups, as in HDTV and multimedia applications. Differences between the ear input signals are necessary in order for auditory events to appear laterally, away from the median plane. Interaural time (ITD) and level differences (ILD) are the main parameters for evoking the illusion of a phantom source. A frequently used technique for the measurement of the equivalence between ILD and ITD are trading experiments, that is asking a subject what time difference is equivalent to a certain level difference. A large number of experiments has been reported, see [26] for an overview. However, most are performed by using headphones, or if loudspeakers are used, by applying an artificial delay between them and leaving the subject in the center position. It is the aim of the present experiments to apply natural time differences only, by using a set-up such as occurs in practical situations for normal audio listening and for a (HD)TV sound environment.

8.2 Experiments

The set-up of the experiment is shown in Fig. 8.1. In an anechoic chamber a pair of two-way (in the horizontal plane) omni-directional loudspeakers were positioned 2.5 m apart. These loudspeakers were constructed by mounting the drivers coaxially, with their axes perpendicular to the horizontal plane. Between the two loudspeakers a third loudspeaker was mounted, hereafter referred to as the center loudspeaker. The three speakers were mounted at 1.62 m height from the floor, the average ear height of the standing subjects. An arc of seven numbered positions for the subject to stand in was marked in such a way that the subject was always 2.19 m away from the central loudspeaker. With a DSP and a small computer the set-up, as in Fig. 8.2, was used to control the experiment. The computer drew up a random order of positions for the subject to stand in, in such a way that the experiment was balanced (i.e. all angles had an equal number of occurrences per subject). The difference in SPL of the right and left loudspeaker at the start of each run was randomly chosen by the computer in such a way that the total electrical power at the center position was

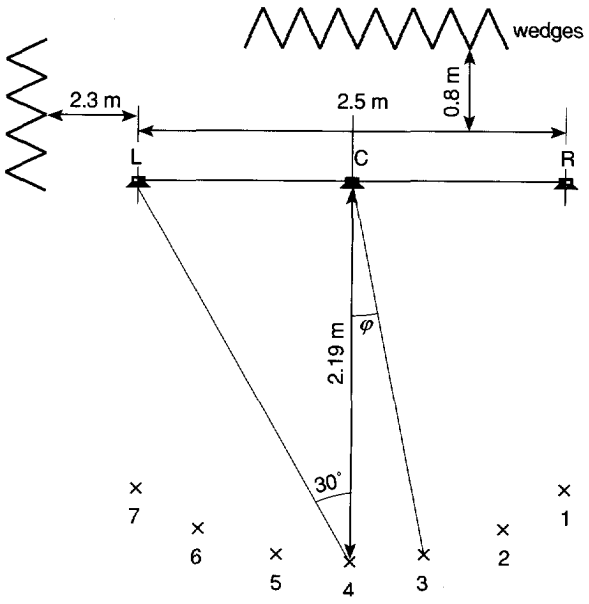

Figure 8.1: Set-up of the experiment in the anechoic chamber.

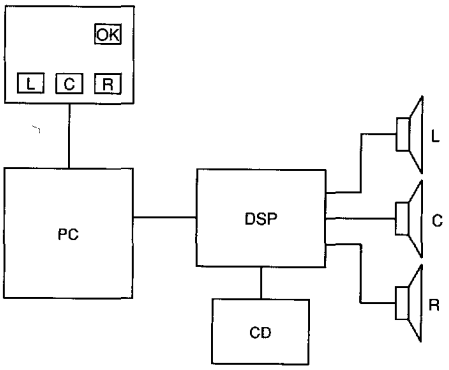

Figure 8.2: Apparatus for controlling the experiments.

always the same. Throughout the experiment the subject kept his head pointing straight forward (i.e. with his shoulders parallel to the wall). The subjects were given a note about the experiment, and the procedure to be followed was made clear to them before the experiment was started. The task for the subjects was to adjust the level of the right and left loudspeakers until the phantom source was located in the center position. A switching facility was included that sent the signal to the left and right loudspeakers,

or to the central loudspeaker, so that at any time during the experiment the subject could switch to the signal coming from the central loudspeaker as a reference of the center position.

The experiment leader indicated to the subject which position to start at. The subject was presented with a box and familiarized with it. Using this box he could control the SPL of the left and right loudspeakers, while the total electrical power fed to both loudspeakers was kept constant. The subject could switch to the signal coming from the central loudspeaker only, and back to the left and right loudspeakers, and indicate when he thought the phantom source was at the same position as the center loudspeaker. The subject was given as much time as was required. When the SPL had been adjusted so that the subject felt that the sound was as though it was coming from the central position, he pressed the button marked 'OK' and the difference in level between the right and left loudspeaker was stored in the computer. This run was then repeated another three times. At the end of the fourth run the experiment leader indicated the new position for the subject and the whole process was repeated until all 7 positions had been taken up.

8.3 Experiment 1

8.3.1 Conditions

The experiment was carried out using pink noise from a CD. The level at the center position at ear height was about 75 dB. The experiment was carried out in an anechoic chamber, with sufficient light for the three loudspeakers to be clearly visible. Ten subjects with normal hearing, four female and six male, participated in the experiment. The average age of the subjects was 25 (std. dev. 4.4 years). Before the experiments, all the subjects' threshold levels were determined by pure tone audiometry. All the subjects had a hearing loss of less than 10 dB; the threshold differences between the left and right ears of each subject were also less than 10 dB. They were carefully and precisely instructed about their task.

8.3.2 Results

The level differences between the right (R) and left (L) loudspeakers are plotted in Fig. 8.3. The vertical bars indicate the variance over all 10 subjects with the four replications per subject. In looking at the data two features appear. Firstly, the variance per angle is rather large. This is

caused by the large variability among the subjects. The data were analyzed using ANOVA, considering the subjects and the different angles as fixed factors. The results are shown in Table 8.1. The entry R^2 shows that

Table 8.1: Analysis of Variance Experiment 1.

SOURCE	SS	DF	F RATIO	SIG	GM	R^2
					0.59	0.99
angles	27070	6	409	0.0		
subjects	151	9	13	0.6E-14		
interaction	958	54	13	0.4E-42		
error	278	210				
TOTAL	28457	279				

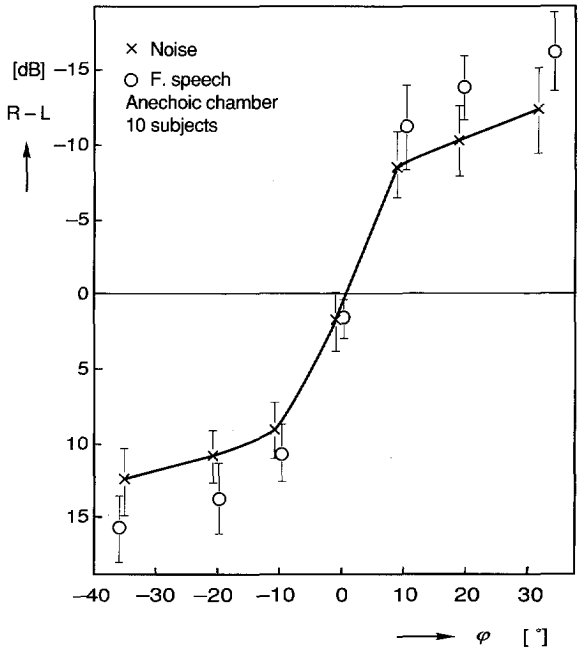

Figure 8.3: Required level difference between right and left loudspeakers, as a function of the subjects' positions, in order to perceive the phantom image in the center.

99% of the variance can be explained by the factors angles and subjects and their interaction. The variance of each of the subjects is remarkably

low. They can reproduce the task very well, but do not agree with each other. The entry SIG shows the significance level and appears to be high for all factors.

The second feature is the deviation of 2 dB at the center position, as can be seen in Fig. 8.3. One would expect this to be 0 dB. Initially a psychological reason was expected, because the experiment was performed in the corner of an anechoic chamber with the corner to the left of the subject. However, in the third experiment the set-up was made symmetrical, but the difference remained the same. This level difference is also reported in [133, 165, 216]. The latter considered it to be an experimental artefact. The entry GM in Table 8.1 is the grand mean of all 280 observations and appears to be very close to 0 dB, as one might expect for equal deviations to the right and left of the median plane, indicating that there is a good symmetry between the left and right plane, except for the center position itself.

All subjects reported a broadening of the phantom source for increasing angles from the center position. At the extreme positions of the subjects, the space between the loudspeakers was filled by the image. This effect is known as image broadening, see [26, 84].

8.4 Experiment 2

The same set-up and conditions as in the former experiment were used. The only difference was the stimulus. A female voice from CD [236] was used as the stimulus, the duration of the excerpt was about 20 s and was then repeated over and over. The results are depicted in Fig. 8.3. The figure shows that the results are very similar to those obtained from the pink noise stimulus. The data were analyzed using ANOVA, considering the subjects and the different angles as fixed factors. The results are shown in Table 8.2.

8.5 Experiment 3

The aim of this experiment was to study the time-intensity trading at a very close listening distance and a relatively small loudspeaker base, as would occur in multimedia PC and TV applications. The set-up is depicted in Fig. 8.4. The switching facility was the same as described before, see Fig. 8.2.

Table 8.2: Analysis of Variance Experiment 2.

SOURCE	SS	DF	F RATIO	SIG	GM	R^2
					0.23	0.99
angles	44878	6	5777	0.00		
subjects	140	9	12	1.0E-13		
interaction	1028	54	14	0.6E-45		
error	272	210				
TOTAL	46318	279				

8.5.1 Conditions

The conditions were adapted to the monitor application. A 16:9-aspect-ratio monitor with a 36" screen was used, without displaying a picture. This was done to make the localization more precise than with a picture, because it is known that when the seen and heard locations of an object are to be discrepant, the sound is perceived near or at the seen location. The experiment was performed in a studio using the same female voice as in the second experiment. The subjects were seated in a chair in five different positions. Each position was repeated four times. Nine subjects out of the set of the previous experiments performed the test.

8.5.2 Results

The level differences between the right and left loudspeakers were the outcome of the experiment and are plotted in Fig. 8.5. Again, a large variance can be observed and a bias at the center position of almost 2 dB appears. The data were analyzed using ANOVA, considering the subjects and the different angles as fixed factors. The results are shown in Table 8.3. The variance within a run for each subject is for this experiment even less than for the previous experiment, 99.6% of the variance being caused by the factors and their interaction. The subjects are consistent, although there is not a broad consensus among the subjects.

8.6 Experiment 4

The aim of this experiment was to study the trading for three reference positions, so the subject's task was to direct the auditory event in three directions, $+30°$, $0°$ and $-30°$. The set-up of the experiment is similar

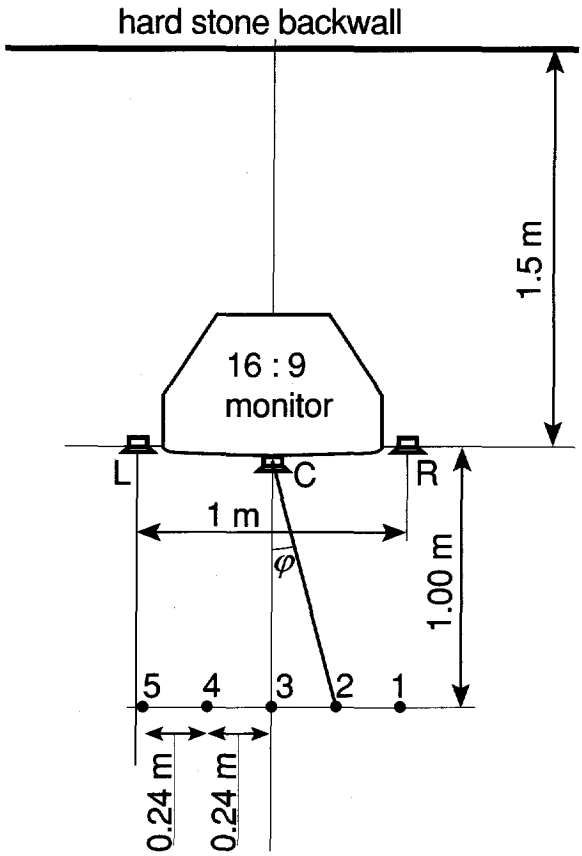

Figure 8.4: Set-up of the experiment in the studio.

to that shown in Fig. 8.1, which was used in the first two experiments. However, the subject position was fixed in front of the center loudspeaker and the delay between the loudspeakers was performed by electronic means.

The subjects were seated in a chair, position 4 in Fig. 8.1, with their head clamped. The task was similar to that in the first three experiments. One of the three loudspeakers (randomly chosen) started radiating pink noise. It was the reference loudspeaker/direction for that run and was switched off after two seconds. The subject's task was to adjust the balance between the outer loudspeakers in such a way that it gave the same direction as the reference one. There were 2 runs per subject per delay value and 4 delay values per angle, giving 24 runs per subject.

The level differences between the right and left loudspeakers were the outcome of the experiment and are plotted in Fig. 8.6. Although the head

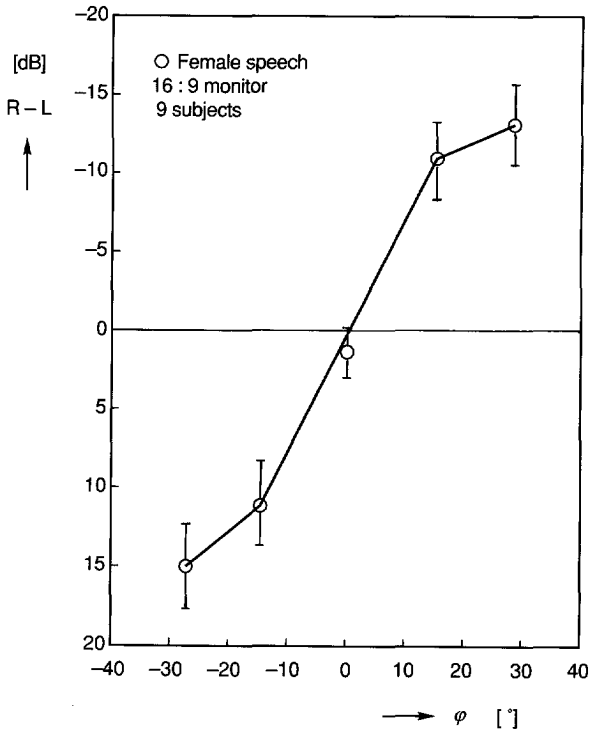

Figure 8.5: Required level difference between right and left loudspeakers, as a function of the subjects' positions, in order to perceive the phantom image in the center of a TV monitor.

Table 8.3: Analysis of Variance Experiment 3.

SOURCE	SS	DF	F RATIO	SIG	GM	R^2
					0.36	0.996
angles	23392	4	9027	0.0E+00		
subjects	191	8	37	0.2E-29		
interaction	757	32	37	0.7E-51		
error	87	135				
TOTAL	24427	179				

was clamped, a very large variance can be observed, especially for the two outer reference positions. Again, a bias at the center position appears.

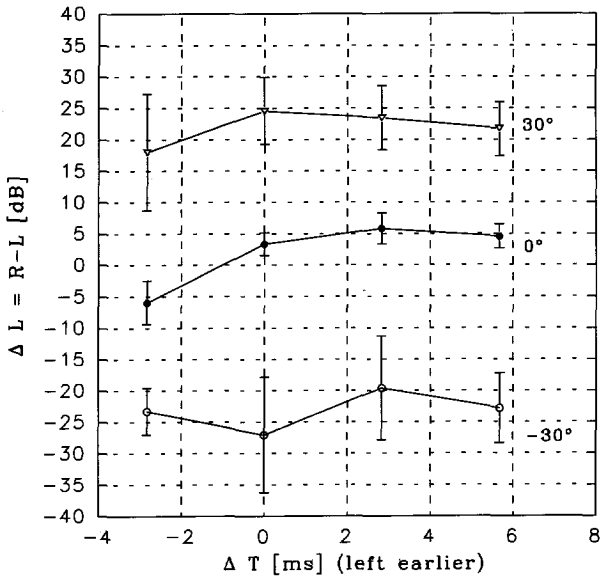

Figure 8.6: Required level difference between right and left loudspeakers, as a function of the time difference between the loudspeakers, in order to perceive the phantom image in the direction of $+30°$, $0°$ or $-30°$.

8.7 Discussion

The data of the first three experiments are converted to exclude the intensity difference caused by different path lengths from left and right loudspeaker to the head, r_{-1} and r_1 respectively. This is done by adding the factor $20\log(r_{-1}/r_1)$ to the R-L values. According to the time-intensity trading concept, the resulting level differences are caused by time differences only. They are plotted in Fig. 8.7, where $\Delta T = (r_1 - r_{-1})/c$. In order to compare the $\Delta T/\Delta I$ values, some other values are also plotted in Fig. 8.7. A convenient way to parametrize this set of curves is

$$\Delta L = b\tanh(a\Delta T) \tag{8.1}$$

where $a \sim 0.8$ and $b \sim 10$. An enormous range of trading ratios has been reported, ranging from 1.7 μsec/dB [176] to 500 μsec/dB [89]. It is known that many factors affect the trading relation. The ILD and the ITD interact in a complex fashion: Some of the features of this interaction may be predicted on physiological grounds. The nature of the stimulus employed and the level at which it is presented are probably the most influential. The smallest ratios are reported for low-frequency tones at high levels, the

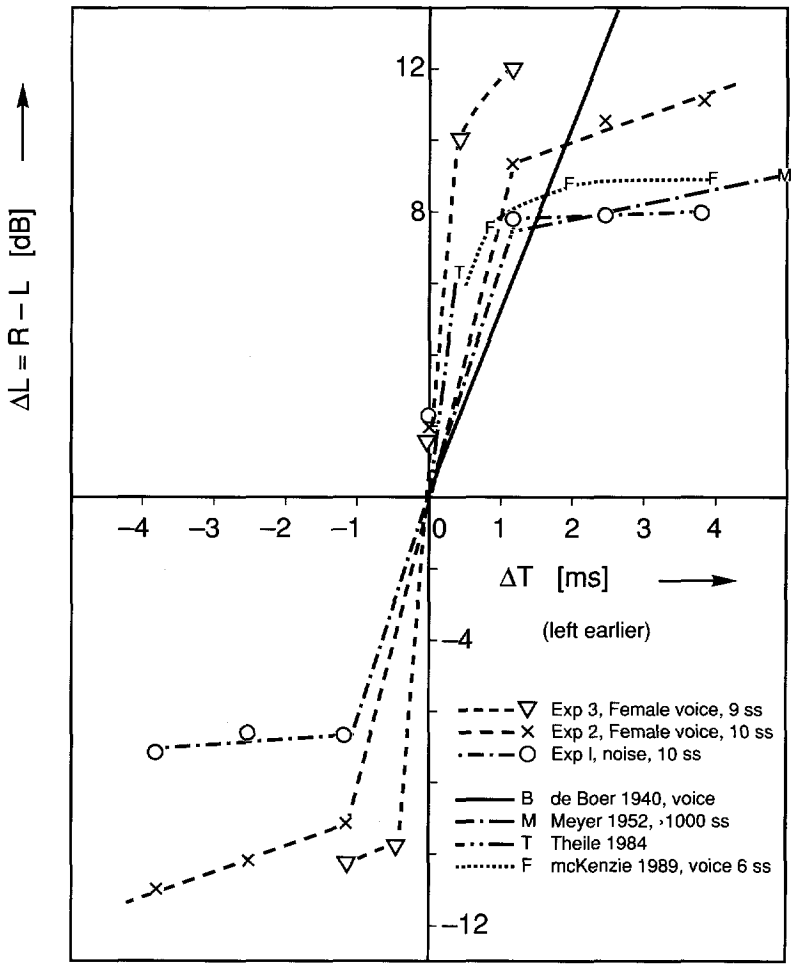

Figure 8.7: Time-intensity trading results of the experiments in comparison with other experiments from the literature.

largest with pulses at low intensity levels. However, there are also important psychological factors. One of these may be conflicting intersensory information. When spatial information in two modalities is available, the information in one of the modalities is partly or wholly discarded in favor of that in the other. When spatial information in light and sound is made to conflict, vision dominates. This phenomenon has been called the ventriloquism effect [104]. It can be important for sound for television, see for example [100, 121, 215], but may also influence the trading experiments; at large angles the subjects can see only one loudspeaker when they are facing

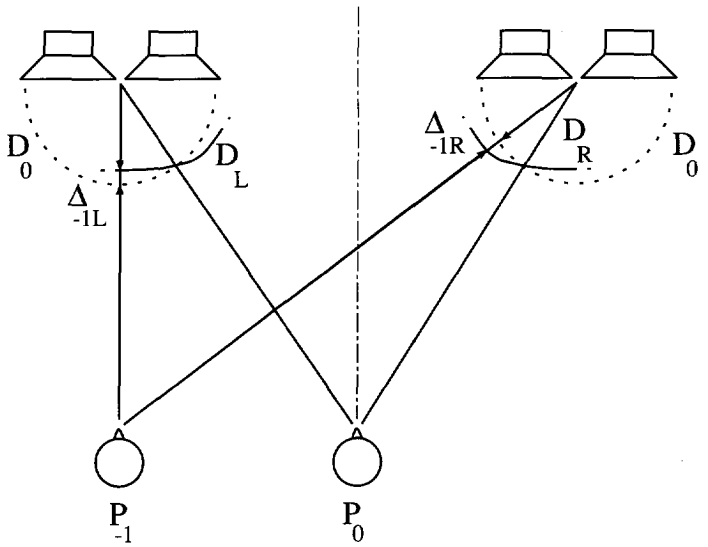

Figure 8.8: Off-center listening effect. If the subject moves from P_0 to P_{-1}, the sound intensity due to the non-uniform polar pattern D_L of the left array decreases at P_{-1} by Δ_{-1L} while the intensity due to the right array increases by Δ_{-1R}.

forward. The paradigm of the experiment may affect the results, even if the ILDs and ITDs are the same as in another experiment. This was the reason for using a third loudspeaker in the center, so it could serve both as an acoustic and as a visual reference.

8.8 Application

The data gathered by the experiments discussed before was used to derive an optimal polar pattern for loudspeakers. The set-up of the experiment is shown in Fig. 8.8. The loudspeaker arrays do not exhibit a uniform polar pattern D_0 (see Fig. 8.8), but D_L for the left, and D_R for the right array respectively. Therefore it seems for a subject moving from position P_0 to P_{-1} that the right loudspeaker array is playing at a higher and the left array at a lower level, by amounts of Δ_{-1R} and Δ_{-1L} respectively. A pair of loudspeaker cabinets was equipped with a pair of drivers for each frequency range. The drivers were fed by digital filters, with coefficients which were calculated by means of a numerical optimization procedure, in such a way that both loudspeakers together radiated sound in a broad range of listening positions in accordance with the time-intensity trading

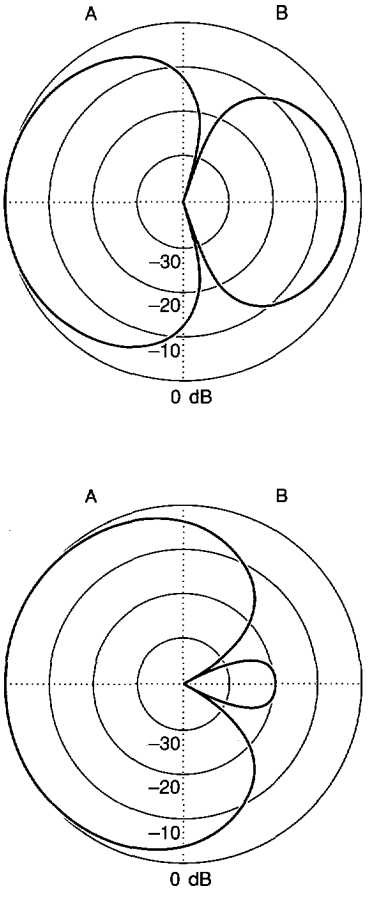

Figure 8.9: Polar pattern of the optimized loudspeaker array for two different frequencies.

results. An example of a calculated pattern is plotted in Fig. 8.9. The other loudspeaker has the same pattern, but mirrored on the main axis. Informal listening tests showed a pronounced enlargement of the listening area for the optimized patterns in comparison with ordinary loudspeakers.

8.9 Conclusions

The results for pink noise and the female voice as stimuli are very similar. Time-intensity trading beyond about 1 ms is not possible, the phantom source then broadens as the time difference increases. The results exhibit a large variability between the subjects, but a low variance within the subjects. The time-intensity curves for (HD)TV applications at a close viewing distance and a small loudspeaker base are much steeper than for a wider stereo-base in the case of normal audio applications. The trading can be usefully applied to the determination of electronic filters steering a loudspeaker (array), in such a way that the polar pattern of the array ensures an enlargement of the listening area for a stereophonic loudspeaker set-up.

Chapter 9

Phantom sources applied to stereo-base widening [1]

A stereo-base widening system producing a good effect, achievable with a simple analog circuit as well as with a digital one, is presented. The system is derived from the combination of the traditional approach using HRTFs (head-related transfer functions), and by using a simple model where ideal loudspeakers and an acoustically transparent subject's head are assumed. Both the HRTF approach as well as the simple model are discussed, and some special cases for the latter are considered. Merging these models, with some (experimentally determined) adaptations led to a practical widening system. The system produces a pleasant and natural sound, particularly for voices, dialogues and 'natural' mixed recordings. The widening circuit is tuned in such a way that it gives the same tonal balance for common stereo recordings as for conventional stereo.

9.1 Introduction

Since the first stereophonic demonstration at the Paris Opera [228], a whole range of improvements to the system have been suggested, of which Blumlein's patent [29] is an early example. Other early examples of interest are de Boer's papers [61–63], of which [63] is concerned with the prob-

[1] Parts of this chapter have been published in [9].
The system is registered as 'Incredible Sound' and patents are pending, however, the system is hereafter referred to as 'the widening circuit'.

lem of the distance between the loudspeakers, which in 1946!, had to be minimized 'in order to obtain a compact arrangement'. This served as a mechanical solution, but numerous electronic solutions have also been proposed; see e.g. [30, 49, 50, 148]. It is the purpose of this chapter to discuss a method enabling a virtual widening of the loudspeaker separation, with less artifacts than the known systems. Although there is a vast amount of literature on general stereophony, of which [61, 75, 102, 136, 182, 184] are early publications and [153, 185, 186, 195] are more contemporary, it is beyond the scope of this section to discuss stereophony in general. The idea of creating virtual sources was introduced over 30 years ago. Schroeder and Atal [171] outlined a method for generating a phantom source. These ideas were elaborated by Damaske [55–57] and are known as TRADIS. The aim of the following sections is to derive a practical solution, replacing the traditional approach generally used. Several approaches for the widening problem will be discussed. A digital filter (using many multiplications) with the HRTF approach in Section 9.3, and a second one, by using a simple model where ideal loudspeakers and an acoustically transparent subject's head are assumed. Finally, both the HRTF approach as well as the simple model are merged (via experimentally determined adaptations), to a third system. This system appeared to be very practical to implement and tolerant against head-movements. This third widening system is implemented by analog circuitry discussed in Section 9.5.1 and a (recursive) digital filter (using only a few multiplications) in Appendix F. Finally, the illusion producing the widening effect will be compared with 'real' loudspeakers in Section 9.6.

9.2 General approach

In order to create phantom sources LS_{LPh} and LS_{RPh}, the left and right source signals, denoted as V_L are V_R, were processed as shown in Fig. 9.1, where for reason of simplicity, only LS_{RPh} is drawn. The notations for a left phantom source follow analogously. The rationale of this processing is, that the same pressure is obtained in both ears when the real loudspeakers LS_L and LS_R are playing as in the situation when the phantom loudspeakers LS_{Lph} LS_{Rph} are playing. If the pressure at the eardrums generated by real loudspeakers is the same as that which would be generated by the phantom sources, and there are no other (e.g. visual) cues, then the subject will perceive no difference, and a virtual enlargement of the aperture α to β is realized. The transfer function H_{Lr}^{α} is from the left loudspeaker to the

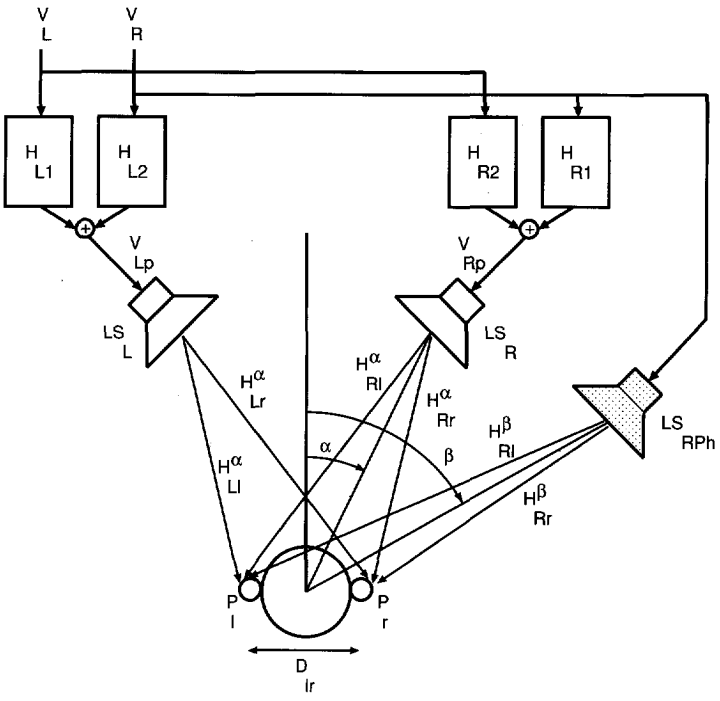

Figure 9.1: Set-up for generating phantom sources. For reason of simplicity, only one phantom source (LS_{RPh}) is drawn, the notations for a left phantom source follow analogously.

right ear, with an angle α between the loudspeaker and median plane. The other functions used are as indicated in Fig. 9.1, with the first (uppercase) index corresponding to the source and the second (lowercase) index to the ear. The pressures in the left and right ears caused by the real loudspeakers LS_L and LS_R are

$$p_l|_{L,R} = V_{Rp} H^{\alpha}_{Rl} + V_{Lp} H^{\alpha}_{Ll} \tag{9.1}$$

and

$$p_r|_{L,R} = V_{Lp} H^{\alpha}_{Lr} + V_{Rp} H^{\alpha}_{Rr}, \tag{9.2}$$

respectively, where V_{Lp} and V_{Rp} are the filtered source signals fed to loudspeakers LS_L and LS_R, respectively. The pressure at the ears caused by the phantom source LS_{RPh} is

$$p_l|_{RPh} = V_R H^{\beta}_{Rl} \tag{9.3}$$

and

$$p_r|_{RPh} = V_R H^{\beta}_{Rr}. \tag{9.4}$$

The combination of Eqs. 9.1–9.4 and the loudspeaker voltages relations given by

$$V_{Lp} = H_{L1}V_L + H_{L2}V_R \tag{9.5}$$

and

$$V_{Rp} = H_{R1}V_R + H_{R2}V_L \tag{9.6}$$

yield the filter functions H_{R1}, H_{L1}, H_{R2} and H_{L2}

$$H_{R1} = \frac{H^{\beta}_{Rl}H^{\alpha}_{Lr} - H^{\beta}_{Rr}H^{\alpha}_{Ll}}{\Delta} \tag{9.7}$$

$$H_{L1} = \frac{H^{\beta}_{Lr}H^{\alpha}_{Rl} - H^{\beta}_{Ll}H^{\alpha}_{Rr}}{\Delta} \tag{9.8}$$

$$H_{R2} = \frac{H^{\beta}_{Ll}H^{\alpha}_{Lr} - H^{\beta}_{Lr}H^{\alpha}_{Ll}}{\Delta} \tag{9.9}$$

$$H_{L2} = \frac{H^{\beta}_{Rr}H^{\alpha}_{Rl} - H^{\beta}_{Rl}H^{\alpha}_{Rr}}{\Delta} \tag{9.10}$$

where $\Delta = H^{\alpha}_{Rl}H^{\alpha}_{Lr} - H^{\alpha}_{Rr}H^{\alpha}_{Ll}$.

If these transfer functions (given by Eqs. 9.7 – 9.10), are implemented, the same pressure is obtained in both ears when the real loudspeakers LS_L and LS_R are playing as in the situation when the phantom loudspeakers LS_{Lph} LS_{Rph} are playing. Then the subject will perceive no difference, and a virtual enlargement of the aperture α to β is realized.

If there is symmetry in the median plane, $H_{Lr} = H_{Rl}$, and $H_{Ll} = H_{Rr}$. Then let $H_{R1} = H_{L1} = H_1$, $H_{R2} = H_{L2} = H_2$ and Eqs. 9.7 – 9.10 reduce to

$$H_1 = \frac{H^{\beta}_{Rl}H^{\alpha}_{Rl} - H^{\beta}_{Rr}H^{\alpha}_{Rr}}{\Delta_s} \tag{9.11}$$

$$H_2 = \frac{H^{\beta}_{Rr}H^{\alpha}_{Rl} - H^{\beta}_{Rl}H^{\alpha}_{Rr}}{\Delta_s} \tag{9.12}$$

$$\Delta_s = (H^{\alpha}_{Rl})^2 - (H^{\alpha}_{Rr})^2. \tag{9.13}$$

Depending on the model used for the transfer functions H, various solutions can be obtained, as will be discussed below.

9.3 Measured transfer functions

The transfer functions used in Eqs. 9.1 – 9.4, were measured with the aid of the set-up shown in Figs. 9.2 and 9.3, in an anechoic chamber. The results

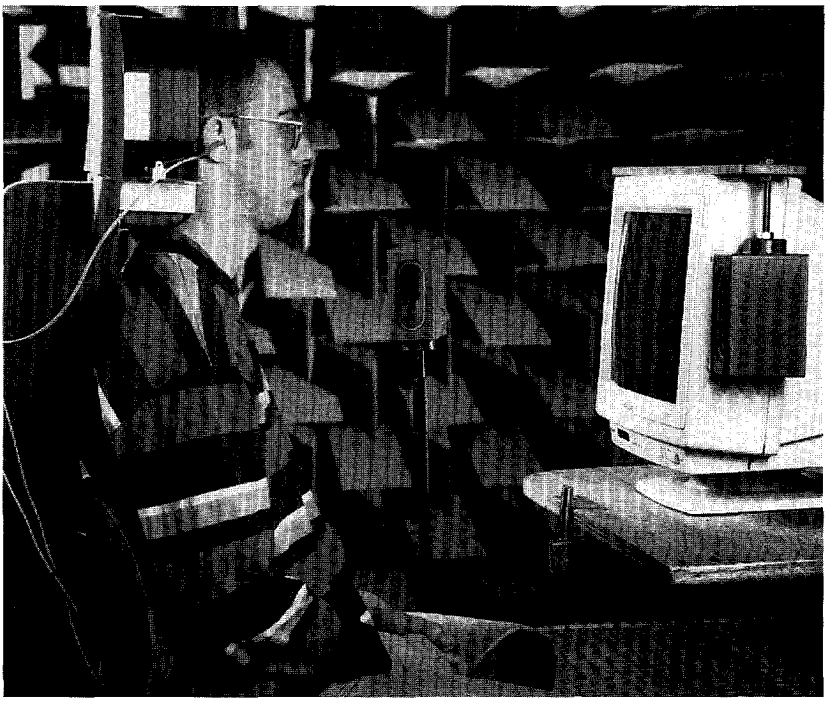

Figure 9.2: Set-up for measuring HRTFs (head-related transfer functions). For verification and measurement purpose is at the place of the aimed phantom source LS_{LPh} a real source present (about the center of the picture). The left source LS_L is present at the left side of the monitor, however not visible in the picture.

are plotted in Fig. 9.4. Invoking Eqs. 9.11 - 9.13 functions H_1 and H_2 were computed; the results are shown in Figs. 9.5 and 9.6, respectively. As Figs. 9.5 and 9.6 show, the required filter characteristics are very complex and expensive to implement. Due to possible changes in the position of the subject's head during sound reproduction with respect to the head position during the measurement of the HRTFs, severe errors may be introduced in functions H_1 and H_2. An adaptive method would hence be more suitable for computing H_1 and H_2. The adaptation process is described as follows. Initially three loudspeakers are used, one un-filtered at the position of the phantom source LS_{RPh}, the other two filtered. Microphones positioned near the subject's ears are used to control the filter's coefficients until the pressure measured is at a minimum level. The coefficients are then frozen and the third loudspeaker is removed. When the two loudspeakers are driven through the phase inverse of the function of the filter, only the illusion of a third source results. The method described above can hence

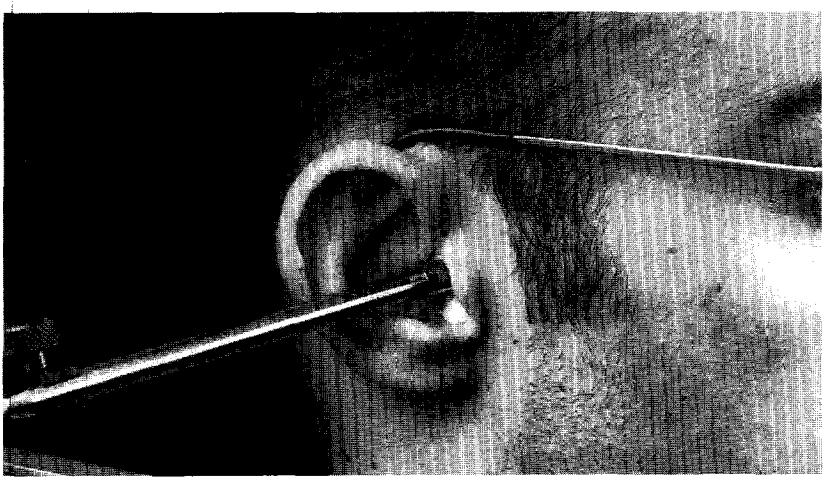

Figure 9.3: Details of the set-up for measuring transfer functions.

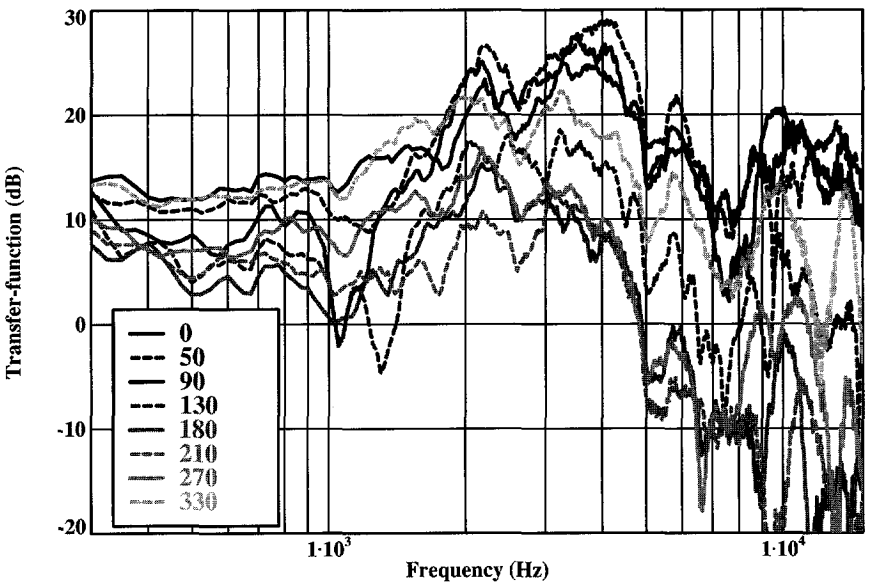

Figure 9.4: Measured HRTFs (head-related transfer functions), from right loudspeaker to right ear. The legend indicates the angle of the loudspeaker with respect to the median plane.

be regarded as an active noise-cancellation solution. There is not elaborated on this noise-cancellation principle, but a simple model will be discussed in the following section.

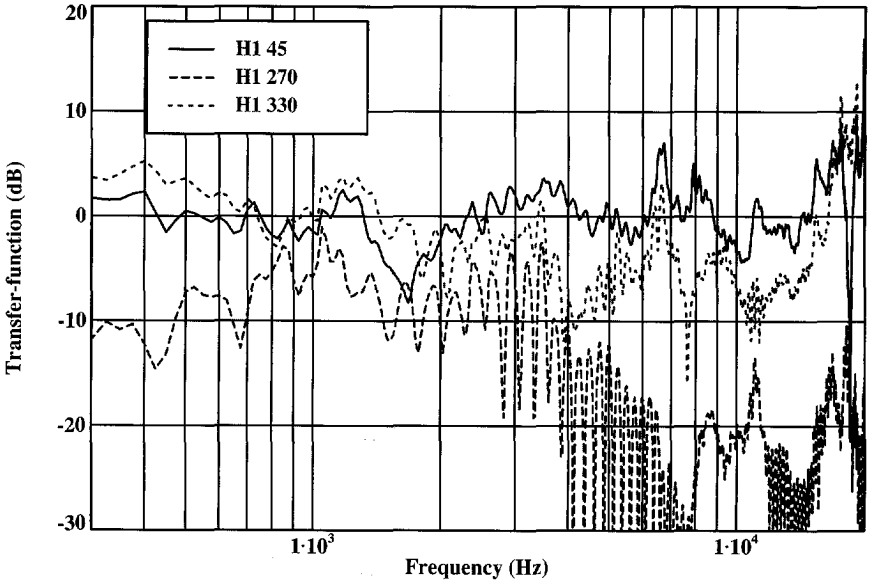

Figure 9.5: Transfer functions of filter H_1 using HRTFs and Eq. 9.11. The legend indicates the angle β of the phantom loudspeaker with respect to the median plane, while the 'real' sources were at $\alpha = \pm 10°$.

9.4 Simple transfer functions

In order to create phantom sources LS_{LPh} and LS_{RPh}, the left and right signals V_L and V_R were processed as indicated in Fig. 9.1. When ideal loudspeakers and an acoustically transparent subject's head (with a diameter $D_{lr} = 2r$) are assumed, then the transfer function H^{α}_{Lr}, from the left loudspeaker to the right ear, with an angle α between the loudspeaker and median plane, is modeled as

$$H^{\alpha}_{Lr} = \exp(-jkD^{\alpha}_{Lr})/D^{\alpha}_{Lr} \tag{9.14}$$

where k is the wave number and D^{α}_{Lr} is the distance between the left loudspeaker and the right ear. If symmetry around the median plane is assumed, all loudspeakers are arranged in an arc with a radius R, $R \gg r$, and the common phase term $\exp(-jkR)$ may be omitted. Therefore, the pressure in the left and right ears generated by the real loudspeakers LS_L and LS_R is,

$$p_l|_{L,R} = V_{Rp} \exp(-jkr\sin(\alpha))/R + V_{Lp} \exp(jkr\sin(\alpha))/R \tag{9.15}$$

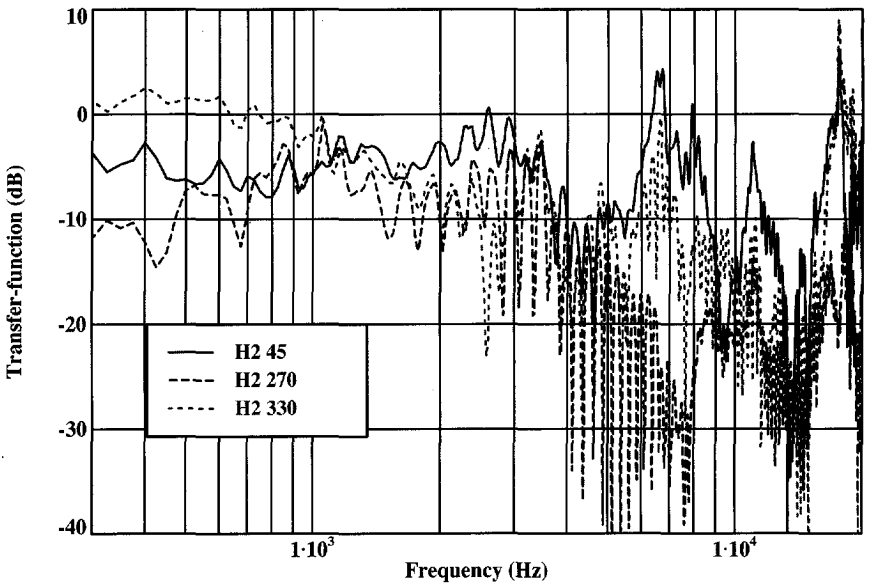

Figure 9.6: Transfer functions of filter H_2 using HRTFs and Eq. 9.12. The legend indicates the angle β of the phantom loudspeaker with respect to the median plane, while the 'real' sources were at $\alpha = \pm 10°$.

and

$$p_r|_{L,R} = V_{Rp} \exp(jkr \sin(\alpha))/R + V_{Lp} \exp(-jkr \sin(\alpha))/R, \tag{9.16}$$

respectively, where V_{Lp} and V_{Rp} are the voltages applied to the two loudspeakers, supplied by the filters H_1 and H_2, where, as in the assumed symmetrical set-up, $H_1 = H_{L1} = H_{R1}$ and $H_2 = H_{L2} = H_{R2}$.

The pressure generated by the phantom source LS_{RPh} is

$$p_l|_{RPh} = V_R \exp(-jkr \sin(\beta))/R \tag{9.17}$$

and

$$p_r|_{RPh} = V_R \exp(jkr \sin(\beta))/R. \tag{9.18}$$

If Eq. 9.15 is equated to 9.17 and Eq. 9.16 is equated to 9.18 and Eqs. 9.5, 9.6 are implemented, then the following holds for filters H_1 and H_2

$$H_1 = \frac{\sin[kr(\sin(\alpha) + \sin(\beta))]}{\sin[2kr(\sin(\alpha))]} \tag{9.19}$$

and

$$H_2 = \frac{\sin[kr(\sin(\alpha) - \sin(\beta))]}{\sin[2kr(\sin(\alpha))]}. \tag{9.20}$$

9.4.1 Special cases

If $A = kr\sin(\alpha)$ and $h = \sin(\beta)/\sin(\alpha)$ then

$$H_1 = \frac{\sin[(1+h)A]}{\sin(2A)} \tag{9.21}$$

and

$$H_2 = \frac{\sin[(1-h)A]}{\sin(2A)}. \tag{9.22}$$

For odd integer values of h the denominators of Eqs. 9.21 and 9.22 become equal to a factor in their numerators as can be shown by using the identity

$$\frac{\sin(2\gamma(2N+1))}{\sin(2\gamma)} = 1 + 2\sum_{l=1}^{N}\cos(4l\gamma). \tag{9.23}$$

As an example two values for h will be used.

For $h = 3$ we have

$$H_{1|_{h=3}} = 2\cos(2A) \tag{9.24}$$

and

$$H_{2|_{h=3}} = -1. \tag{9.25}$$

Here we observe the usual, well-known effect of the introduction of a 180° phase-shifted 'cross-talk' between the channels; see e.g. [99]. The 'main path' ($H_{1|_{h=3}}$) then becomes less dominant than the 'cross-path' ($H_{2|_{h=3}}$), which can produce annoying effects, such as an instable stereo image and extreme sensitivity to subjects' head movements. These equations are real. However, after multiplication by a common phase term, they can be easily implemented with the Finite Impulse Response filters. The same result (Eqs. 9.24, 9.25) would be obtained for a small-angle and low-frequency approximation of Eqs. 9.19 and 9.20.

For $h = 5$ we have

$$H_{1|_{h=5}} = 0.5 + \cos(4A) \tag{9.26}$$

and

$$H_{2|_{h=5}} = -2\cos(2A). \tag{9.27}$$

For the value $h = 5$ the cross-term H_2 becomes frequency-dependent. Thus, for odd integer values of h Eqs. 9.21 and 9.22 can be realized with the aid of simple FIR filters as will be discussed in the following.

9.5 Implementations of stereo widening

Two electronic circuits realizing the widening system have been designed, an analog component version and a digital one. The analog will be discussed in the following section, the other will be discussed in Appendix F. Both circuits are based on the simple model of Section 9.4. According to this model the required filters can be implemented by FIR filters using Eqs. 9.21 and 9.22 as shown in Fig. 9.7. The transfer functions of these functions are plotted in Figs. 9.8 and 9.9, for $D_{lr} = 2r = 17.5$ cm and $\alpha = 5°$. However, the model of Section 9.4 does not take into account the influence of the pinna, the ear canal, the head and torso etc. Therefore the filter shown in Fig. 9.7 (for $h = 3$) was tuned 'by hand', in such a way that the timbre was as close as possible to that for normal stereo reproduction for various repertoire of music; while the widening remained. The so obtained transfer function was transferred to a circuit using analog components, as discussed in the next section.

9.5.1 Simple analog filter

The tuned filter response as described above, was fitted experimentally to a circuit using analog components. The bass response was slightly increased, which was previously not possible in the (digital) FIR filter (Fig. 9.7) due to its short length. The final circuit diagram is shown in Fig. 9.10. Only the right channel of this circuit diagram of the analog circuit is shown. The left output channel is the same, except for the fact that the left and right input labels are reversed. The amplitude response of the widening circuit is shown in Fig. 9.11 and the phase response in Fig. 9.12. The label 'Hsum' refers to the output of one of the channels, while both channels are driven by the same signal. The line marked 'S' is the level at which the circuit has the same loudness as ordinary stereo. This level was determined empirically by several listeners using several types of music.

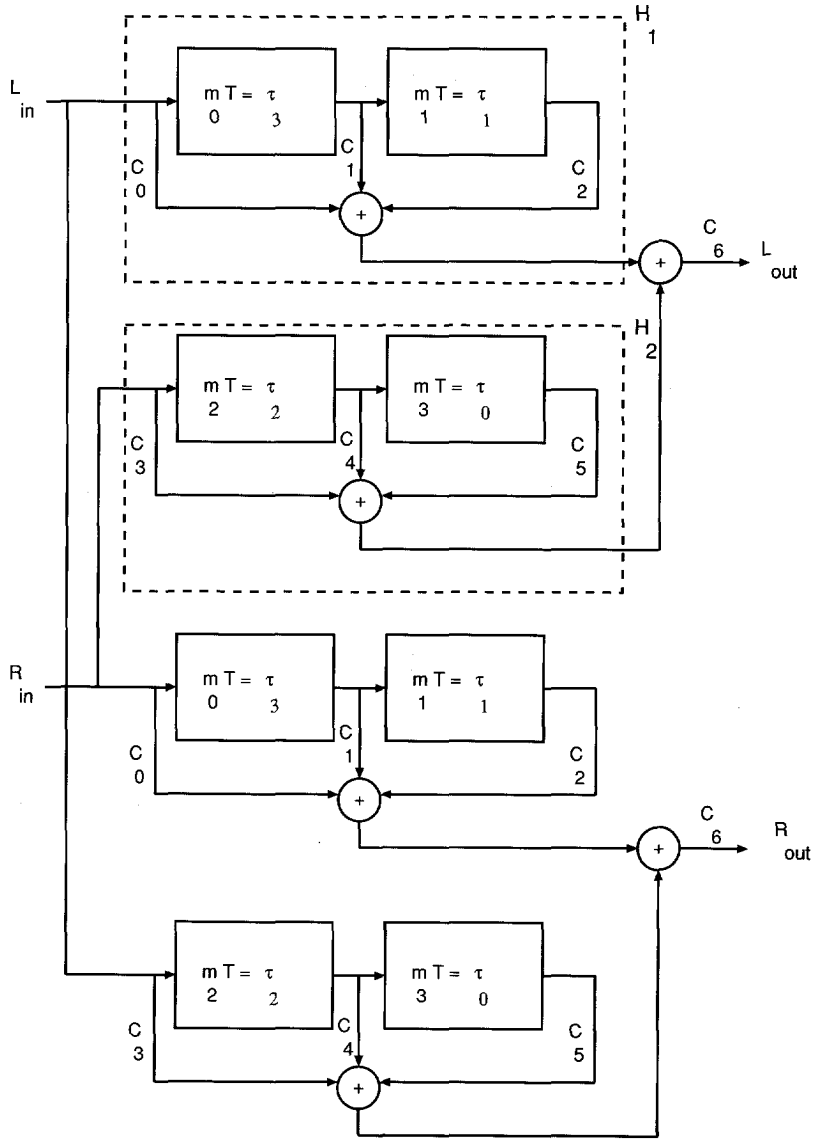

Figure 9.7: FIR filter set-up for H_1 and H_2 given by Eqs. 9.21 and 9.22 using the simple model where ideal loudspeakers and an acoustically transparent subject's head are assumed.

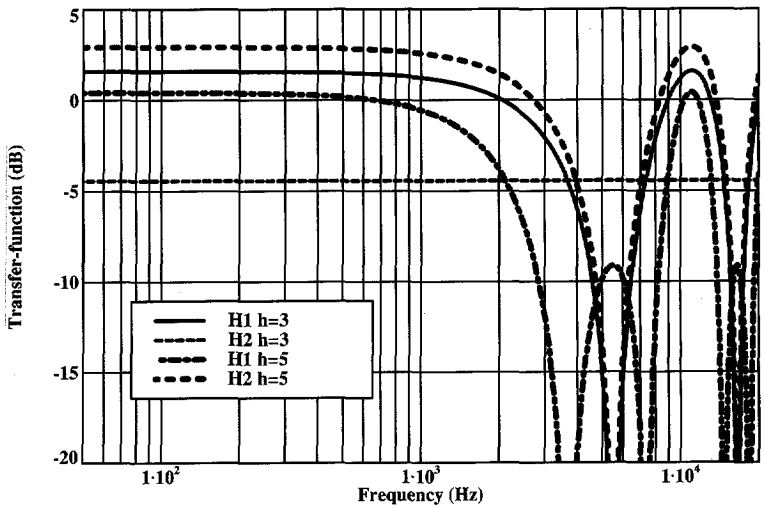

Figure 9.8: FIR filter transfer function H_1 and H_2 given by Eqs. 9.21 and 9.22 using the simple model for $h = 3$ and $h = 5$ where ideal loudspeakers and an acoustically transparent subject's head are assumed, for $D_{lr} = 2r = 17.5$ cm and $\alpha = 5°$.

9.6 Psychoacoustic validation

A listening test was conducted to determine the apparent angle with respect to the median plane at which the phantom source is perceived. The listeners were surrounded by loudspeakers placed in an arc with a radius of two meters ranging from $-90°$ to $10°$, with a numbered loudspeaker at every $10°$. The loudspeakers at $\pm 10°$ were driven by the circuit shown in Fig. 9.10; the others were not connected to the amplifier. Three subjects, one at a time were seated in a studio. They were asked to indicate the position of the apparent source of left-channel band-limited noise played at 1/3 octave band center frequencies ranging from 250 Hz to 16 kHz. The noise was periodically gated by a gaussian multiplier with a total window length of 2 s. The subjects were looking straight ahead (0°), as if they were watching TV; their heads were not clamped. They indicated the number of the loudspeaker which they thought was playing. The subject was given as much time as was required. After completion of all frequency bands, the test was repeated once more. The averaged angle over six points per frequency band is plotted in Fig. 9.13. As the figure reveals, although the sound-producing loudspeakers were only at $\pm 10°$, the apparent angle was much larger.

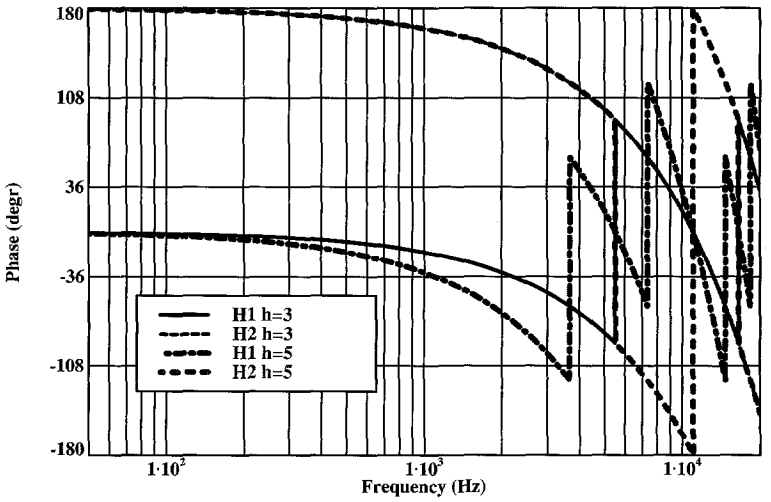

Figure 9.9: FIR filter phase response H_1 and H_2 given by Eqs. 9.21 and 9.22 using the simple model for $h = 3$ and $h = 5$ where ideal loudspeakers and an acoustically transparent subject's head are assumed, for $D_{lr} = 2r = 17.5$ cm and $\alpha = 5°$.

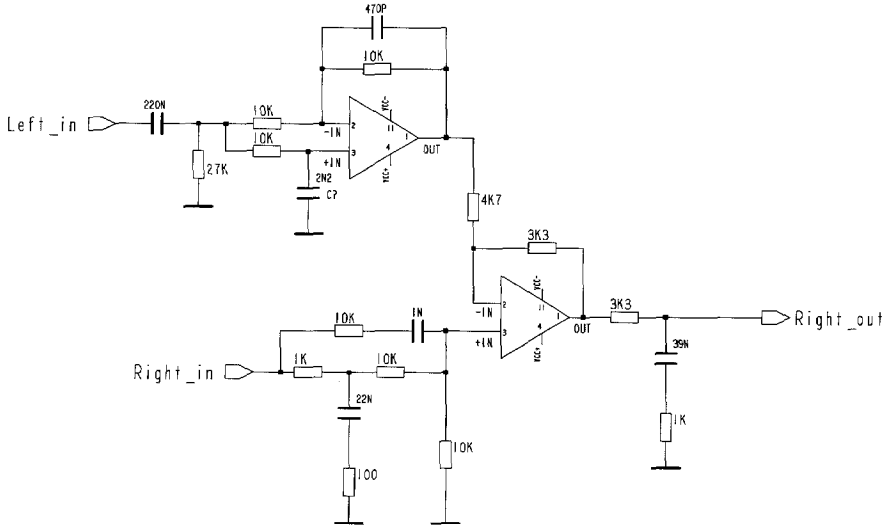

Figure 9.10: The widening circuit, only the right channel is shown.

It was noted by Trahiotis [205] that there is an intriguing similarity between Fig. 9.13 and Figs. 2–5 from Schiano et al. [167]. Although the present experiment and those of Schiano are quite different, they used head-

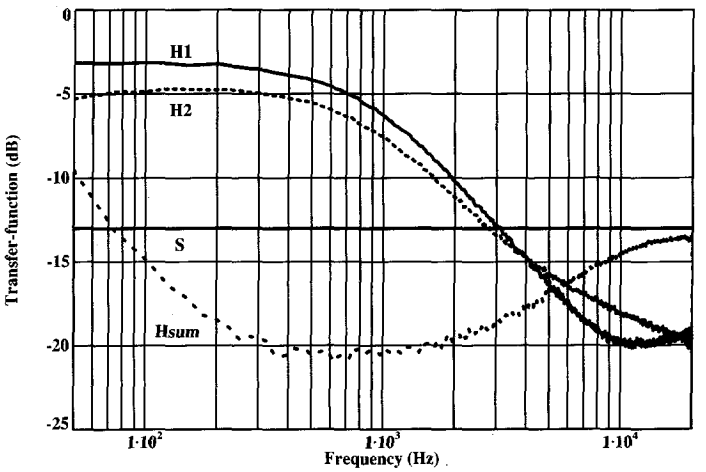

Figure 9.11: Amplitude response of the widening circuit of Fig. 9.10.

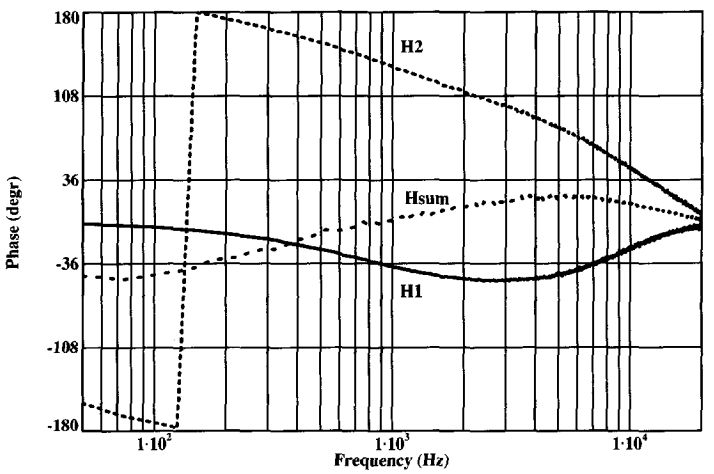

Figure 9.12: Phase response of the widening circuit of Fig. 9.10.

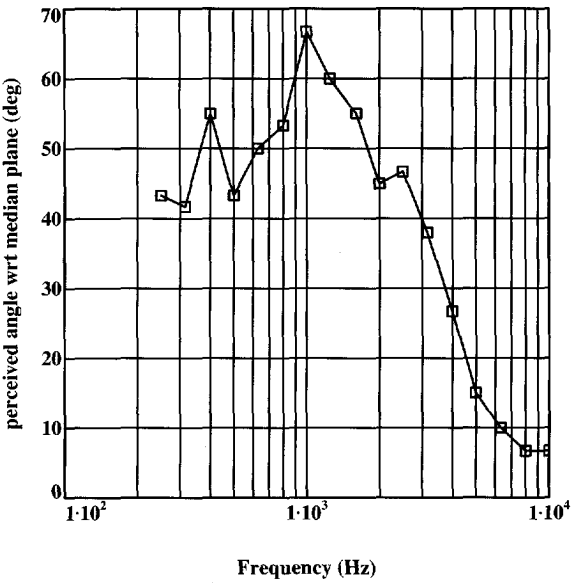

Figure 9.13: Averaged apparent angle of the phantom source indicated by three subjects, using the widening circuit of Fig. 9.10. The real sources were at $\pm 10°$.

phones and stimuli with a fixed interaural time difference, the underlying principles are the same. At frequencies above say 1.5 kHz interaural level differences are the main parameters for localization.

9.7 Conclusions

A stereo-base widening system producing a good effect has been presented. The 'averaging' between the traditional HRTF approach and a simple model appeared to be a practical method to derive such a system. The system is applicable to TV/monitor-sets (multimedia), portable audio and midi-sets. It produces a pleasant and natural sound, particularly for voices, dialogues and 'natural' mixed recordings and, for normal stereo recordings, it produces the same tonal balance as is obtained without the circuit.

Chapter 10

Adapting loudspeaker signals for headphones [1]

Reproduction of loudspeaker signals via headphones leads to in-head localization. By convolving these signals with the impulse responses measured from loudspeakers to ear canals and deconvolving them with impulse responses measured from headphones to ear canals, binaural signals are generated, which simulate a loudspeaker set-up and make better localization possible. In general the inverse of the system describing an outer ear and a headphone is not stable and a modification is necessary to make implementation in a digital signal processor possible.

In this chapter a new technique is described to modify the measured functions and allow for implementation with minimal errors in the magnitude of the functions.

First the discrete-time impulse response of a measured transfer function is transformed into the z-domain. All zeros are then determined. The positions of the zeros outside the unit circle show deviations from a minimum phase system, indicating the frequency regions where errors in the phase characteristic are introduced when the function is inverted and made stable. It will be shown that when stabilizing the quotient of loudspeaker to ear and headphone to ear functions some of these errors may be avoided. Advantages of analysis and processing using z-transform are discussed.

[1] Parts of this chapter have been published in [31].

10.1 Introduction

Binaural signals contain directional information and are intended to be reproduced in left and right ears to create a natural hearing event. Using an artificial head directional information of sound sources is included in the recorded signals.

However, when headphones are used to reproduce signals intended for a standard stereo loudspeaker set-up, then directional information is not transferred and as a consequence all sound sources are localized inside the listener's head. In much the same way as in a binaural mixing console [26] binaural correction functions containing directional information and compensating for the headphone to ear transfer may be generated electronically and used to improve headphone sound reproduction, see also [198, 217, 218]. This chapter describes a method to design the binaural correction functions with minimal errors in the magnitude characteristics and which can be implemented in a digital signal processor. There is no attention paid to the psychoacoustic validation of the results, these are considered in [214]. We assume that our hearing is more sensitive to magnitude errors (common to both ears) of the binaural correction functions than to errors in the phase. For this reason methods, which trade magnitude errors against phase errors are not considered in this chapter.

First a transfer function is measured from a sound source to a point in the ear canal (for instance $H_{Lr}(\omega)$ from left loudspeaker to right ear). This function contains the directional information of one sound source. A second transfer function $H_{Hr}(\omega)$ is measured from the headphone to the same point in the ear canal, the inverse of which must be used to compensate for the headphone when the binaural signals are reproduced. The application directional information $H_{Lr}(\omega)$ and compensation for the headphone $H_{Hr}^{-1}(\omega)$ is done in one step using a binaural correction function $H'_{Lr}(\omega) = H_{Lr}(\omega)H_{Hr}^{-1}(\omega)$. Four of these binaural correction functions are used to add directional information to left and right stereo loudspeaker signals L and R as shown in Fig. 10.1. Because of the similarity in the calculation of the four binaural correction functions only the processing of one of them ($H'_{Lr}(\omega)$) will further be discussed. When reproducing the processed signals L_{out} and R_{out} via headphones the direct sound from a loudspeaker set-up in an anechoic environment is simulated. Therefore this simulation creates no impression of spaciousness, unless another processing step is performed where the reflection pattern of a listening room is added. This step will not be discussed here. In Section 10.2 the measurements of the outer ear transfer functions of loudspeakers and headphones to ear

canal are described. In Section 10.3 a new method to determine the binaural correction functions in the z-domain is described. A discussion follows in Section 10.4.

10.2 Filters for loudspeaker simulation

In a standard stereo set-up two loudspeakers are positioned in $+30°$ and $-30°$ directions at 2.5 m distance from the listener. The transfer functions from loudspeakers to ears can be measured in an anechoic room by positioning a subject and a loudspeaker in a corresponding position. A probe microphone is placed at a reference point in the ear canal to measure the signals radiated by the loudspeaker [134], the ear canal was not blocked as was done by others [138]. We thus obtain the function

$$H_{Lr}M(\omega), \tag{10.1}$$

where $H_{Lr}(\omega)$ is the transfer function from the left loudspeaker to the right ear canal and $M(\omega)$ is the acoustic disturbance due to the microphone. Likewise, using the headphones as a source a measurement is performed to the same reference point in the ear canal

$$H_{Hr}M'(\omega). \tag{10.2}$$

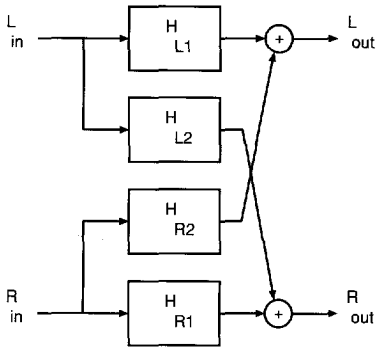

Figure 10.1: Scheme to change left (L_{in}) and right (R_{in}) loudspeaker signals to binaural signals L_{out} and R_{out} using four binaural correction functions. Where $H_{L1} = H_{Ll}/H_{Hl}$, $H_{L2} = H_{Lr}/H_{Hr}$, $H_{R1} = H_{Rr}/H_{Hr}$ and $H_{R2} = H_{Rl}/H_{Hl}$. The transfer functions are denoted by a capital subscript for the sound source (L loudspeaker, H headphone) and a second subscript for the receiving ear (l left, r right). The binaural correction functions each consist of a quotient of two transfer functions. The numerator contains directional information of one loudspeaker to one ear and the denominator compensates for the headphone to ear transfer.

Here $H_{Hr}(\omega)$ is the transfer function from headphones to ear canal. $M'(\omega)$ is the disturbance by the microphone if the headphone is present. If a source signal $S(\omega)$ is reproduced using the binaural correction function then the signal at the reference point is equal to

$$S'(\omega) = S(\omega)H_{Hr}(\omega)\frac{H_{Lr}(\omega)M(\omega)}{H_{Hr}(\omega)M'(\omega)}. \tag{10.3}$$

We call the portion of Eq. 10.3 that is written as a fraction the binaural correction function. Schröter et al. proved [172] that if:

- the microphone is small enough not to change the acoustic input impedance of the ear canal, or
- the headphones have an open structure and do not change the radiation impedance of the ear canal,

then $M'(\omega) = M(\omega)$. Thus for the two loudspeaker directions $+30°$ and $-30°$ the transfer functions are measured at both ears and from these measurements and the two headphone to ear transfer functions four binaural correction functions are calculated.

10.3 Calculation of the correction functions

A correction function is determined by multiplication of the complex functions $H_{Lr}(\omega)$ and $H_{Hr}^{-1}(\omega)$. The convolution of the source signal with the impulse response of the correction function is carried out in a finite impulse response filter [157]. From a number of binaural correction functions determined for various ears and headphones it appears that due to the limited number of coefficients in a practical implementation a considerable truncation error has to be made.

One reason is that the inversion of $H_{Hr}(\omega)$ increases the length of its impulse response from the length determined by the measurement of the function to infinite.

The second reason is that if $H_{Hr}(\omega)$ does not describe a minimum phase system then the inverted function $H_{Hr}^{-1}(\omega)$ belongs to an unstable system. The truncation errors have a disturbing influence on the frequency response. We found that, if the function to be inverted is made minimum phase avoidable errors may be introduced in the correction function, as will be shown below.

Unlike the conventional methods, which perform processing in the time domain or in the frequency domain, we prefer proceeding in the z-domain.

The z-transform is a useful technique for representing and manipulating discrete-time sequences. Given an impulse response $h(k)$, which is zero outside the interval $0 \leq k \leq N-1$, its z-transform [112, 157] is

$$H(z) = \sum_{k=0}^{N-1} h(k) z^{-k} = h(0) \prod_{k=1}^{N-1} (1 - z_k z^{-1}). \tag{10.4}$$

In this way a system with an impulse response of length N is completely described by the $N-1$ zeros z_k. The system described has a real impulse response and consequently zeros appear in complex conjugate pairs or are real. Therefore their position in the z-plane has a mirror symmetry with respect to the real z-axis. From the distribution of zeros in the z-plane we can make some observations.

- If a zero is located outside the unit circle the function is not minimum phase and after inversion the function becomes unstable,
- if in $H_{Lr}(\omega)$ and $H_{Hr}(\omega)$ similar resonances exist with coinciding zeros outside the unit circle these zeros will cancel in the quotient in Eq. 10.3 and will not give rise to instability,
- resonant dips with a high quality factor, which may appear due to a special headphone construction in combination with the outer ear can be recognized from a zero close to the unit circle.

The correction function can now be constructed in a few steps. First we invert $H_{Hr}(\omega)$ by replacing all zeros by poles, which are then combined with the zeros of $H_{Lr}(\omega)$. Fig. 10.2 shows this configuration in the z-plane. It appears that the poles and zeros are distributed close to the circle $|z| <\sim 1$, which will be discussed in Appendix G in more detail. If any poles are situated outside the unit circle, as is the case shown in Fig. 10.2, then additional steps are necessary. The function must be modified to obtain a stable system, the impulse response of which must not be longer than the available finite impulse response filter. If the magnitude function is to be left unchanged the only displacement of zeros and poles allowed is a mirroring with respect to the unit circle. Thus the following options exist:

- position all zeros and poles inside the unit circle to make the correction function minimum phase and stable,
- position all poles inside the unit circle to make the correction function stable,

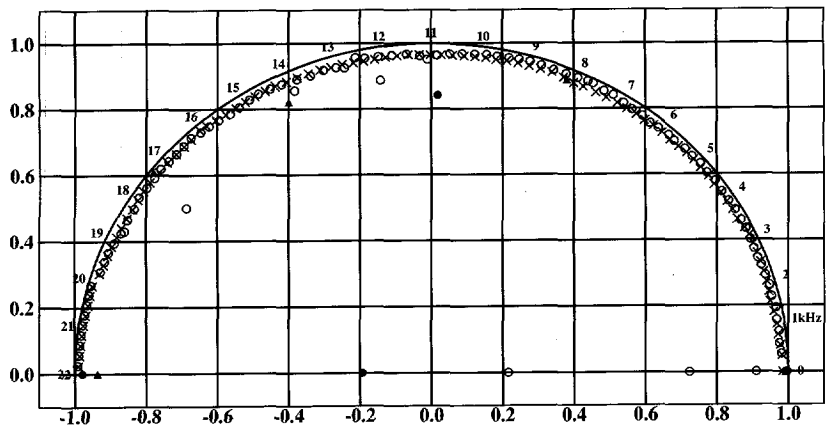

Figure 10.2: Complex z-plane with 200 zeros (o and •) of the transfer function from a Quad ESL63 loudspeaker in direction of +30° to the nearest ear of the artificial head KU81i. Real frequencies from 0 to half the sampling frequency (22.05 kHz) are plotted along the unit circle. The filled circles • show zeros outside the unit cycle but drawn mirrored with respect to the unit circle. The inverse of the transfer function from a Stax-λ headphone to the same ear is described with 200 poles (x). This function is unstable due to poles outside the unit circle (mirrored and plotted as filled triangles) near frequencies of 0, 8.2, 14.2, 17.5 and 22 kHz.

- position only the poles that do not coincide with zeros inside the unit circle to make the correction function stable.

The first method changes the phase characteristic but not the magnitude. A disadvantage is that also all zeros are mirrored introducing unnecessary errors in the phase characteristic. However, this method results in a correction function with an impulse response of minimal length. If the system described by all zeros has an impulse response longer than the available finite impulse response filter, then this method limits the errors in the magnitude of the correction function, which would be caused by truncation during the implementation.

The second method gives better results by preventing errors in the phase characteristic due to displacement of zeros as mentioned above. The third method shows an advantage of the description with zeros and poles. Now the correction function is made stable by mirroring with respect to the unit circle only those poles, which lie outside the unit circle and are not compensated by zeros. All superfluous errors in the phase characteristic are avoided and the magnitude remains unchanged. Stabilizing the correction function this way is therefore preferable. Fig. 10.3 - 10.4 show the complex

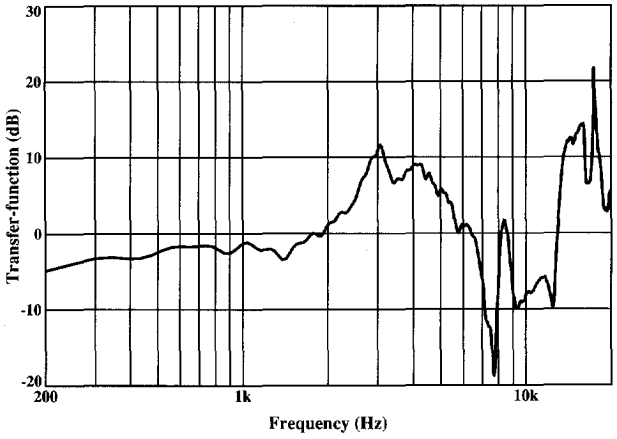

Figure 10.3: The magnitude of the unstable function $H_{Rr}(\omega)H_{Hr}^{-1}(\omega)$ as obtained from the same measurements as in Fig. 10.2 and of the (stabilized) processed version to be implemented in a finite impulse response filter of 200 taps. The two curves are not distinguishable.

quotient of a loudspeaker and a headphone measurement, prior to any processing (solid lines) and the correction function obtained using the third method (dashed lines). The magnitude characteristics, which are equal before and after processing are shown in Fig. 10.3; as the figure shows the two curves are not distinguishable. Fig. 10.4 shows the errors appearing in the group delay characteristic due to the mirroring of some of the poles. The group delay is proportional to the derivative of phase with respect to frequency and shows processing errors more clearly. When the pole at $z = -1.07$ is mirrored with respect to the unit circle we see that its influence does not extend into the audible frequency region. Three poles at 8.2, 14.4 and 17.5 kHz are mirrored with respect to the unit circle to make the correction function stable. The difference between the two curves is due to stabilizing the polynomial. The influence of the mirroring of the poles can be seen at the mentioned frequencies. One pole at $z = 1.02$ is compensated by a zero at the same position and not mirrored. Thus neither in the phase nor in the magnitude characteristic an error at low frequencies is introduced.

Another advantage of the zeros and poles description appears when we have to deal with resonances of high quality factor. These give rise to sharp dips or peaks in the frequency characteristic and thus must have a zero or pole close to the unit circle. If some damping of this resonance is desired, this can be realized easily by moving the corresponding zero or pole away

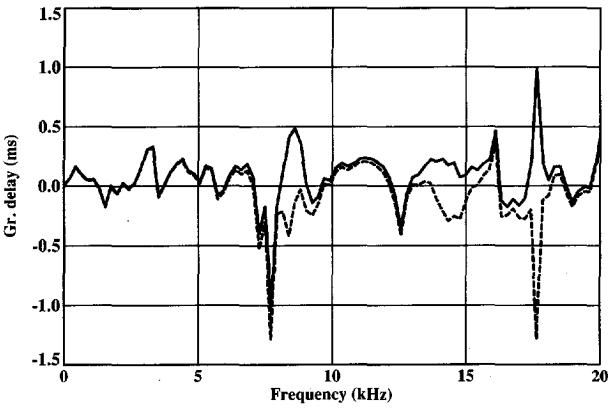

Figure 10.4: The group delay of the same functions as shown in Fig. 10.3. The difference between the two curves is due to stabilizing the polynomial. This can be seen in the frequency regions where a pole has been mirrored to make the function stable (8.2, 14.2 and 17.5 kHz). The pole outside the unit circle at $z = 1.02$ (Fig. 10.2) was not mirrored to prevent the introduction of phase errors in the low frequency region.

from the circle.

Another problem sometimes arises when the loudspeakers reproduce low frequencies better than the headphones. In this case the correction function will rise towards low frequencies to boost the headphone sound reproduction. This may result in an unacceptable distortion. Moving a zero or pole in this frequency region allows us to shape the low frequency characteristic and thus decrease the boosting.

After the processing in the z-domain we obtain the correction function from the zeros and poles by substituting in the system function:

$$H(z) = \frac{\prod_{k=1}^{N-1}(1 - z_k z^{-1})}{\prod_{k=1}^{M-1}(1 - p_k z^{-1})}, \tag{10.5}$$

for z_k and p_k the calculated N zeros and M poles and $z = e^{j\theta}$. Here θ is the real angular frequency normalized to the sampling frequency. Using a discrete Fourier transformation we then find the impulse response, which can be used in a finite impulse response filter to perform the convolutions.

10.4 Conclusions

A method is described based on the z-transform to analyze and manipulate functions consisting of the quotient of two transfer functions the result of

which must be implemented in a finite impulse response filter. Following this method a binaural correction function is found, which can be used to simulate a loudspeaker set-up when a normal stereo signal is reproduced via headphones. Another application could be the equalization of an artificial head for a sound source in a specific direction or for a specific headphone.

We assume that our hearing is more sensitive to errors in the magnitude of the binaural correction function than to errors in the phase characteristic. For this reason we describe a method resulting in minimal errors in the magnitude characteristic and trading of magnitude errors against errors in the phase are not considered.

Using the z-transform a function can be described with zeros only. A zero finding algorithm was used in combination with a high precision number representation (see Appendix G). A description using zeros is useful to gain insight in functions to be processed. The configuration of zeros in the z-plane shows in what frequency regions the function is not minimum phase and where errors in the phase characteristic are introduced if the function has to be inverted and made stable. Inversion of a function is easily performed by replacing all zeros by poles and vice versa. When poles exist outside the unit circle then a function is made stable by mirroring all poles with respect to the unit circle.

When an implementation in a finite impulse response filter is desired then truncation errors will occur. It was shown that poles in the inverted function, which are outside the circle and coincide with zeros should not be mirrored when the function is made stable. This will lead to minimal errors in the magnitude characteristic and only limited errors in the phase characteristic.

Appendix A

Details of lumped-element parameters and efficiency

The lumped parameters for some loudspeakers are given in Table A.1. From these values and Eq. 1.13, the efficiency η is calculated and plotted in Fig. A.1, showing that the efficiency is in the range of 0.2–1.0%.

The equivalent volume of a loudspeaker is given by

$$V_{eq} = \rho c^2 (\pi a^2)^2 / k_t \tag{A.1}$$

or for a given volume of the enclosure, the corresponding k_t of the 'air-spring' can likewise be calculated. Mounting a loudspeaker in a cabinet

Table A.1: The lumped parameters for a 4" and some 7", 8", 10" and 12" loudspeakers, each with a low and high Q_e, (see Table 1.1 for abbreviations).

Type	R_e	Bl	k_t	m_t	R_m	S	f_0	Q_m	Q_e
	Ω	Tm	N/m	gr.	Ns/m	cm^2	Hz		
AD44510	6.6	3.5	839	4	0.86	54	72	2.2	1.02
AD70652	7.5	6.5	885	13.2	1.48	123	41	2.3	0.61
AD70801	6.9	2.9	1075	6.3	0.81	123	66	3.2	2.13
AD80110	6	9.0	971	16.5	1.38	200	39	2.9	0.29
AD80605	6.8	5.1	1205	13.4	0.84	200	48	4.8	1.05
AD10250	6.6	13	1124	38.5	2.74	315	27	2.4	0.25
AD10600	6.8	5.9	909	28.5	1.06	315	29	4.8	0.99
AD12250	6.6	13	1429	54	2.93	490	26	3.0	0.34
AD12600	6.9	6	1205	33	0.76	490	31	8.2	1.21

will increase the total spring constant by an amount given by Eq. A.1 and subsequently increase the bass cut-off frequency of the system. To compensate for this bass loss, the moving mass has to be increased, thus $\sqrt{k_t m}$ is increased, changing Q_e (see Eq. 1.3). Therefore, Bl must be increased in order to preserve its original value. The original frequency response is then achieved, but at the cost of a more expensive magnet and a loss in efficiency. This is the designer's dilemma: high efficiency or small enclosure? To meet the demand for a certain cut-off frequency, the enclosure volume must be greater. Alternatively, the efficiency for a given volume will be less than for a system with a higher cut-off frequency.

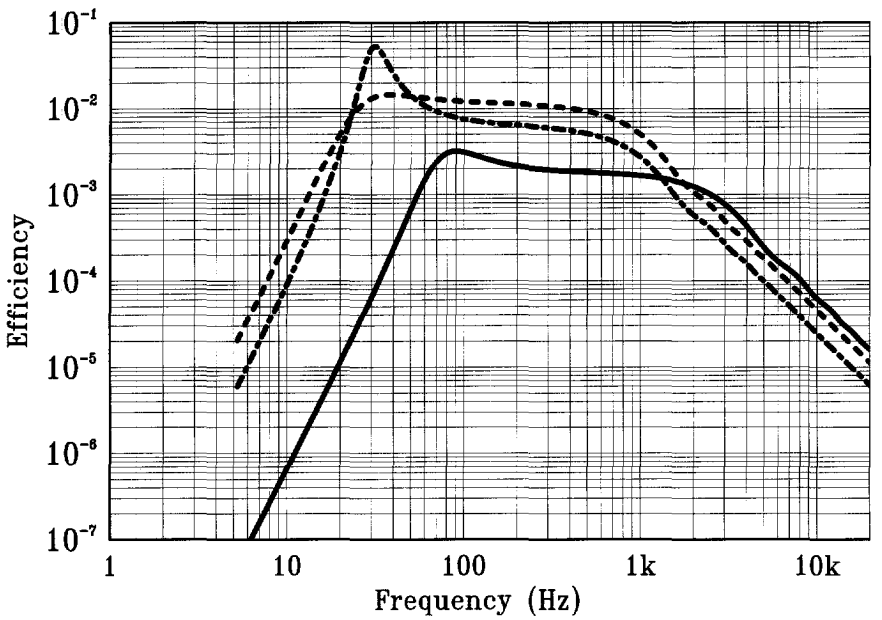

Figure A.1: Efficiency of loudspeakers AD44510 (solid), AD12250 (dashed) and AD12600 (dot- dashed). See Table A.1 for the parameters.

Appendix B

Balanced Incomplete Block Designs

In this section some notes on Balanced Incomplete Block Designs (BIBDs) are discussed and a bibliography is given on this topic. Very little attention is paid to the theory of BIBDs. However, some designs are presented here; these are mainly intended for use in listening tests in order to gather information for scaling purposes.

B.1 Introduction

A lady wishes to invite her seven friends to a series of dinner parties. Her table provides room for three other persons. In addition, the lady wants each pair of friends to meet only once at her table. The question arises: How should she arrange her invitations over the days [17]? Such an arrangement is a Balanced Incomplete Block Design (BIBD). It is incomplete because not all possible combinations of 3 out of 7 are used, and it is balanced because all pairs occur (once). Balanced Incomplete Block Designs are important for experimental design. One application might be in the use of triadic comparisons[1] , see [126, 168]. The gain of using a BIBD for triadic comparisons with v stimuli can be a factor

$$G_1 = 3 \left(\begin{array}{c} v \\ 3 \end{array} \right) / \left(\begin{array}{c} v \\ 2 \end{array} \right) = v - 2, \tag{B.1}$$

[1]The question raised in [126] whether in their BIBD more than four blocks are possible, can be answered affirmative. The solution is presented in Table B.3.

it is based on the combination of 3 pairs in one block. However, as Roskam [164] pointed out, there are some disadvantages. He showed that it is in general impossible to rank all the stimuli when only all pairs of stimuli are considered, and that by using a BIBD this problem even gets worse. A direct but laborious and cumbersome method to overcome this problem would be to collect the required data by asking a subject to rank all pairs of stimuli according to the similarity of the stimuli in each pair. The important advantage of this method is that the data will be transitive by themselves. This means that the triangle equality is always fulfilled, so when three stimuli are perceived as $A > B$ and $B > C$, then the comparison of A and C must give $A > C$. Because this method of ranking the pairs turns out not to be very practical, one may use paired comparisons of pairs of stimuli. There are $m = \binom{v}{2}$ pairs and thus $N_{pp} = \binom{m}{2}$ pairs of pairs or

$$N_{pp} = \binom{\binom{v}{2}}{2} = (n/8)(n^3 - 2n^2 - n + 2). \tag{B.2}$$

The gain of using triadic comparisons instead of pairs of pairs is

$$G_2 = N_{pp}/(3\binom{v}{3}) = (v+1)/4. \tag{B.3}$$

The price that has to be paid for this benefit is the omission of information. It is obvious that the enormous advantage of using a BIBD instead of comparing pairs of pairs

$$G_3 = G_1 G_2 = (v-2)(v+1)/4 \tag{B.4}$$

can sometimes take its toll. It appears that for the construction of the BIBD unnecessarily exhaustive searches are sometimes performed, or that one adjusts the number of stimuli to be equal to that of other researchers (using their tables). In the following we discuss some constructions of BIBD with triadic comparisons as the application. Many of the contributions to the construction of BIBDs have been made by mathematicians and sometimes antedate those by psychometricians. An initial introduction to this subject may be [17].

B.2 Definition of BIBD

In a block design there are v objects (or varieties), e.g. loudspeakers, arranged into b blocks of k distinct objects each. Each object occurs r times

and each unordered pair occurs λ times. For triadic comparisons $k = 3$. An example is given by the design in Table B.1. The numbers correspond to the $v = 7$ stimuli (or loudspeakers in particular). The $b = 7$ blocks each with $k = 3$ stimuli are presented to the subject. Each stimulus is presented $r = 3$ times to the subject and each pair of stimuli is presented $\lambda = 1$ times. The elementary conditions on the parameters are

$$bk = vr \tag{B.5}$$

and

$$r(k-1) = \lambda(v-1). \tag{B.6}$$

If a design exists we write $B(k, \lambda, v)$. Combining Eq. B.5 - B.6 gives

$$b = \lambda(v-1)v/((k-1)k). \tag{B.7}$$

If $b = v$, $k = r$ the design is called a symmetric design. Although the construction of BIBDs is very interesting, see e.g. [94], we will focus on the case $k = 3$, for use in triadic comparisons. The combinatorial problems arising in connection with BIBDs were surveyed in papers by Bose [33] and Hanani [95].

A special case occurs for $k = 3$ and $\lambda = 1$, which is called a Steiner triple system. T. Kirkman proved in 1847 that if $v \equiv 1, 3 \bmod 6$, Steiner triple systems do exist for all $v \geq 1$. Over 100 years later, Hanani [96] proved that the necessary conditions

$$\lambda(v-1) \bmod (k-1) = 0 \tag{B.8}$$

and

$$\lambda v(v-1) \bmod k(k-1) = 0 \tag{B.9}$$

are also sufficient for the existence of a BIBD with $k = 3, 4, 5$ and any λ, with the exception of the non-existent design $k = 5, \lambda = 2, v = 15$ [95]. Finally, Wilson proved that the necessary condition of Eqs. B.8 - B.9 is also sufficient for every k and every λ if v is sufficiently large.

A special case of a Steiner system is a resolvable design or Kirkman triple system, which is a Steiner system partitioned into r sub-designs, so that each element occurs once and once only in the sub-design. It was originally formulated as: 'A school-mistress is in the habit of taking her girls for a daily walk. The girls are fifteen in number, and are arranged in five rows of three each, so that each girl might have two companions. The problem is to dispose them so that for seven consecutive days no girl will

walk with any of her school-fellows in any triplet more than once.' It was only in 1969 that the complete solution was given by D.K. Ray-Chaudhuri and R.M. Wilson [158]. They proved that a Kirkman triple system exists for $v \equiv 3 \bmod 6$ (e.g. v = 3,9,15,21).

A group divisible design (GD) is a partitioning of a BIBD so that the design is divided into sub-designs of m blocks each, with each sub-design containing all the elements. A Kirkman triple system is the case where $m = 5$; this corresponds to an extension of the original schoolgirl problem (as Sylvester suggested) by: 'to make the school walk every week in the quarter so that each three walk together' [66].

B.3 Tables

There are several tables for BIBDs e.g. [4]. In [48] constructions or tables are given for: $k = 3$ and b, v = 4,4; 7,7; 10,5; 10,6; 12,9; 26,13; 30,10; 35,15; 57,19; 70,21. The 35,15 design is partitioned so that it forms a Kirkman design. Cochran [48] gives the same b, v pairs and also 20,6. Cole et al. [51] give all 80 possible Steiner systems for $v = 15$.

For $v = 9$ there is 1 Kirkman design and for $v = 15$ there are 13 disjoint Kirkman designs [66] listed in Table B.3.

B.4 Construction

The construction of BIBDs can be performed in many many ways, e.g. based on finite geometries or on difference sets. The latter will be discussed next. If v is a prime, a simple construction can be the following. Let G be an Abelian group of v elements $x_i = x_0, x_1, \ldots, x_{v-1}$, with addition as the group operation. We can associate each treatment with a different element in the group. The elements are associated with the integers $0, 1, \ldots, v-1$, and all sums are reduced mod$v$. Let $D = d_1, d_2, \ldots, d_k$ be a proper subset of G such that every element of G , except x_0, can be expressed in exactly λ ways as a difference $d_i - d_j$. Then D is called a (perfect) difference set. If $D_i = \{d_1 + x_i, \ldots, d_k + x_i\}, x_i \in G$, then the collection of all v sets D_i yields a symmetric ($b = v, r = k$) BIBD. For example, Table B.1 shows the construction of two designs with B(3,1,7) as parameters, and (0,1,3) and (0,1,5) as difference sets. When $x_i = i$, and i starts at 0 and is incremented to 6 in steps of 1.

When two designs have no block in common, as in the foregoing example, they are orthogonal or disjoint. It is not possible, as found by ex-

Table B.1: $k = 3, \lambda = 1, v = 7, r = 3, b = 7$.

0 1 3	0 1 5
1 2 4	1 2 6
2 3 5	2 3 0
3 4 6	3 4 1
4 5 0	4 5 2
5 6 1	5 6 3
6 0 2	6 0 4

perimental result, to construct a Steiner system for $v = 3$ without a block in common with two other disjoint Steiner systems. It is proved [111], that for $v > 7$ and $v \equiv 1, 3 \bmod 6$, the maximum number of orthogonal Steiner systems is equal to $v - 2$, except possibly v =141, 283, 501, 789, 1501, 2365. A simple upper-bound is also $v - 2$, since $G_1 = v - 2$. Obviously, it is easy to construct a B(3,3,7) by removing both orthogonal sets from the complete design consisting of $\begin{pmatrix} 7 \\ 3 \end{pmatrix}$ blocks. Clearly, at most v blocks can be constructed by one difference set, but e.g. B(3,1,15) can be constructed by using more difference sets: (0,3,4) generates the first 15 blocks, (0,6,8) generates the second part and (0,5,10) generates the remaining part (only the first 5 blocks). This design can be partitioned in the way of a Kirkman design, the sub-design generated by (0,5,10) is Kirkman-partitioned by itself. The same holds for (0,1,4), (0,7,13) and (0,5,10). However, this one and the preceding one are not disjoint (e.g. (0,7,13) are in common). A construction for Steiner triple systems for all $v = 6n + 1$ or $v = 6n + 3$ is given by Moore [94].

Two disjoint (the maximum achievable) Steiner sets for $k = 3, \lambda = 1, v = 7, r = 3, b = 7$ are given in Table B.1. Designs partitioned in such a way that each sub-design has at most p different elements and each sub-design differs at most by w elements from its predecessor and successor, are Minimal Distance BIBDs. For triadic comparisons of loudspeakers these are important designs. One wants the least possible number of loudspeakers simultaneous on their listening position (parameter p), while the number of triads with these loudspeakers must be maximum and the effort to bring new loudspeakers from the depot to the listening position and vice versa must be minimal (number of changes of parameter w). It appears impossible to partition a BIBD (for $\lambda = 1$) into parts greater than two blocks if each part has at most 5 different elements in common.

Table B.2: $k = 3, \lambda = 1, v = 9, r = 4, b = 12$.

0	1	2	0	1	3	3	4	6	6	7	0
3	4	5	4	5	6	7	8	0	1	2	3
6	7	8	7	8	2	1	2	5	4	5	8
0	3	6	0	4	7	3	7	1	6	1	4
1	4	7	1	5	8	4	8	2	7	2	5
2	5	8	3	6	2	6	0	5	0	3	8
0	5	7	0	6	8	3	0	2	6	3	5
6	2	4	7	3	5	1	6	8	4	0	2
3	1	8	4	1	2	7	4	5	1	7	8
0	4	8	0	2	5	3	8	5	6	2	8
3	2	7	4	3	8	7	2	6	1	0	5
6	1	5	7	1	6	1	0	4	4	3	7
6	0	1	0	3	4	3	6	7			
2	3	4	5	6	7	8	0	1			
5	7	8	8	1	2	2	4	5			
6	2	5	0	5	8	3	8	2			
0	3	7	3	6	1	6	0	4			
1	4	8	4	7	2	7	1	5			
6	4	7	0	1	7	3	1	4			
5	1	3	8	4	6	2	0	7			
2	0	8	5	2	3	8	5	6			
6	3	8	0	2	6	3	0	5			
2	1	7	5	1	4	8	4	7			
5	0	4	8	3	7	2	1	6			

The maximum number of disjoint Steiner designs for $v = 9$ is 7. They are constructed according to the method of [68]. The triplets within a design are permuted so that each sub-design contains all the elements. The results are listed in Table B.2. Thirteen disjoint Kirkman sets (the maximal achievable) [66] for $D60v2 : k = 3, \lambda = 1, v = 15, r = 7, b = 35$ are listed in Table B.3.

Table B.3: $k = 3, \lambda = 1, v = 15, r = 7, b = 35$ (the solution to Kirman's 15 schoolgirls problem).

0	1	9	2	4	12	5	10	11	7	8	13	3	6	14
0	2	7	3	4	8	5	6	12	9	11	13	1	10	14
0	3	11	1	7	12	6	8	10	2	5	13	4	9	14
0	4	6	1	8	11	2	9	10	3	12	13	5	7	14
0	5	8	1	2	3	6	7	9	4	10	13	11	12	14
0	10	12	3	5	9	4	7	11	1	6	13	2	8	14
1	4	5	2	6	11	3	7	10	8	9	12	0	13	14
1	2	10	0	3	5	6	11	12	8	9	13	4	7	14
1	3	8	4	5	9	0	6	7	10	12	13	2	11	14
1	4	12	0	2	8	7	9	11	3	6	13	5	10	14
1	5	7	2	9	12	3	10	11	0	4	13	6	8	14
1	6	9	2	3	4	7	8	10	5	11	13	0	12	14
0	1	11	4	6	10	5	8	12	2	7	13	3	9	14
2	5	6	3	7	12	4	8	11	0	9	10	1	13	14
2	3	11	1	4	6	0	7	12	9	10	13	5	8	14
2	4	9	5	6	10	1	7	8	0	11	13	3	12	14
0	2	5	1	3	9	8	10	12	4	7	13	6	11	14
2	6	8	0	3	10	4	11	12	1	5	13	7	9	14
2	7	10	3	4	5	8	9	11	6	12	13	0	1	14
1	2	12	5	7	11	0	6	9	3	8	13	4	10	14
3	6	7	0	4	8	5	9	12	1	10	11	2	13	14
3	4	12	2	5	7	0	1	8	10	11	13	6	9	14
3	5	10	6	7	11	2	8	9	1	12	13	0	4	14
1	3	6	2	4	10	0	9	11	5	8	13	7	12	14
3	7	9	1	4	11	0	5	12	2	6	13	8	10	14
3	8	11	4	5	6	9	10	12	0	7	13	1	2	14
0	2	3	6	8	12	1	7	10	4	9	13	5	11	14
4	7	8	1	5	9	0	6	10	2	11	12	3	13	14
0	4	5	3	6	8	1	2	9	11	12	13	7	10	14
4	6	11	7	8	12	3	9	10	0	2	13	1	5	14
2	4	7	3	5	11	1	10	12	6	9	13	0	8	14
4	8	10	2	5	12	0	1	6	3	7	13	9	11	14
4	9	12	5	6	7	0	10	11	1	8	13	2	3	14
1	3	4	0	7	9	2	8	11	5	10	13	6	12	14
5	8	9	2	6	10	1	7	11	0	3	12	4	13	14
1	5	6	4	7	9	2	3	10	0	12	13	8	11	14
5	7	12	0	8	9	4	10	11	1	3	13	2	6	14
3	5	8	4	6	12	0	2	11	7	10	13	1	9	14
5	9	11	0	3	6	1	2	7	4	8	13	10	12	14
0	5	10	6	7	8	1	11	12	2	9	13	3	4	14
2	4	5	1	8	10	3	9	12	6	11	13	0	7	14
6	9	10	3	7	11	2	8	12	0	1	4	5	13	14

2	6	7	5	8	10	3	4	11	0	1	13	9	12	14
0	6	8	1	9	10	5	11	12	2	4	13	3	7	14
4	6	9	0	5	7	1	3	12	8	11	13	2	10	14
6	10	12	1	4	7	2	3	8	5	9	13	0	11	14
1	6	11	7	8	9	0	2	12	3	10	13	4	5	14
3	5	6	2	9	11	0	4	10	7	12	13	1	8	14
7	10	11	4	8	12	0	3	9	1	2	5	6	13	14
3	7	8	6	9	11	4	5	12	1	2	13	0	10	14
1	7	9	2	10	11	0	6	12	3	5	13	4	8	14
5	7	10	1	6	8	0	2	4	9	12	13	3	11	14
0	7	11	2	5	8	3	4	9	6	10	13	1	12	14
2	7	12	8	9	10	0	1	3	4	11	13	5	6	14
4	6	7	3	10	12	1	5	11	0	8	13	2	9	14
8	11	12	0	5	9	1	4	10	2	3	6	7	13	14
4	8	9	7	10	12	0	5	6	2	3	13	1	11	14
2	8	10	3	11	12	0	1	7	4	6	13	5	9	14
6	8	11	2	7	9	1	3	5	0	10	13	4	12	14
1	8	12	3	6	9	4	5	10	7	11	13	0	2	14
0	3	8	9	10	11	1	2	4	5	12	13	6	7	14
5	7	8	0	4	11	2	6	12	1	9	13	3	10	14
0	9	12	1	6	10	2	5	11	3	4	7	8	13	14
5	9	10	0	8	11	1	6	7	3	4	13	2	12	14
3	9	11	0	4	12	1	2	8	5	7	13	6	10	14
7	9	12	3	8	10	2	4	6	1	11	13	0	5	14
0	2	9	4	7	10	5	6	11	8	12	13	1	3	14
1	4	9	10	11	12	2	3	5	0	6	13	7	8	14
6	8	9	1	5	12	0	3	7	2	10	13	4	11	14
0	1	10	2	7	11	3	6	12	4	5	8	9	13	14
6	10	11	1	9	12	2	7	8	4	5	13	0	3	14
4	10	12	0	1	5	2	3	9	6	8	13	7	11	14
0	8	10	4	9	11	3	5	7	2	12	13	1	6	14
1	3	10	5	8	11	6	7	12	0	9	13	2	4	14
2	5	10	0	11	12	3	4	6	1	7	13	8	9	14
7	9	10	0	2	6	1	4	8	3	11	13	5	12	14
1	2	11	3	8	12	0	4	7	5	6	9	10	13	14
7	11	12	0	2	10	3	8	9	5	6	13	1	4	14
0	5	11	1	2	6	3	4	10	7	9	13	8	12	14
1	9	11	5	10	12	4	6	8	0	3	13	2	7	14
2	4	11	6	9	12	0	7	8	1	10	13	3	5	14
3	6	11	0	1	12	4	5	7	2	8	13	9	10	14
8	10	11	1	3	7	2	5	9	4	12	13	0	6	14
2	3	12	0	4	9	1	5	8	6	7	10	11	13	14
0	8	12	1	3	11	4	9	10	6	7	13	2	5	14
1	6	12	2	3	7	4	5	11	8	10	13	0	9	14
2	10	12	0	6	11	5	7	9	1	4	13	3	8	14
3	5	12	0	7	10	1	8	9	2	11	13	4	6	14
4	7	12	0	1	2	5	6	8	3	9	13	10	11	14
9	11	12	2	4	8	3	6	10	0	5	13	1	7	14
0	3	4	1	5	10	2	6	9	7	8	11	12	13	14

Appendix C

Multidimensional scaling

C.1 Introduction

A problem encountered in many disciplines is how to measure and interpret the relationships between objects. A second problem is the general lack of a mathematical relationship between the perceived response and the actual physical measure. Sometimes relationships are rather vague. How much does the character of one person resemble that of another? Or in the case of this thesis, to what extent are different loudspeakers alike? How do we measure and what scale do we need? In the following we discuss some scales and techniques and give some examples.

C.2 Scaling

The purpose of scaling is to quantify qualitative data. Scaling procedures attempt to do this by using rules that assign numbers to qualities of things or events. Multidimensional scaling (MDS) is an extension of univariate scaling; it is simply a useful mathematical tool that enables us to represent the similarities of objects spatially as in a map.
MDS models may be either metric or non-metric, depending on the scale of measurement used in collecting the data. For metric scaling the collected data should be measured using an interval or ratio scale. In the former case the unit of the 'yardstick', used for measuring the phenomenon, as well as the zero point (offset), are unknown. For a ratio scale the zero point is known but there is an unknown scaling factor. For non-metric MDS only the ranking order of the data values is used; the data are or are used at an ordinal level. However, it is sometimes possible to recover metric distances

obtained by non-metric MDS, as will be shown in an example later on.
In order to obtain a spatial map from an MDS computer program, we only need to apply a set of numbers as the input. To all (or most) combinations of pairs out of a group of objects, a number is assigned which expresses the similarity between the objects of that pair. Such numbers are sometimes referred to as proximities. MDS procedures will then represent objects judged similar to one another as points close to each other in the resulting spatial map. Objects judged to be dissimilar are represented as points distant from each other.

MDS programs which use direct similarity measures as input have the advantage of being low in experimental contamination. They do not require a prior knowledge of the attributes of the stimuli to be scaled.

C.3 Example

An obvious procedure for obtaining similarity data is to ask people directly to judge the 'psychological distance' of the stimulus objects. Another way is the method of triadic comparisons [126]. This has the advantage that it simplifies the subject's task, because only the ranking order of three presented stimuli is asked for. However, there can be some drawbacks, as pointed out by Roskam [164]. A practical problem arises when for a complete experiment the number of triads (all possible combinations of three stimuli out of the set of all stimuli) is considered too large. It can be reduced by using an incomplete balanced block design or BIBD (see Appendix B).

As an example of MDS using triadic comparisons, consider the following. Suppose someone with a good topographical knowledge of the Netherlands is asked to give the nearest and the most distant two cities out of 3 given cities. The same question is asked for 3 other cities and so on, so each distance from one city to another (each out of a total list of e.g. 14 cities) is considered. A matrix M can be constructed so that for the 3 cities (i,j,k) the two closest together e.g. (i,j) contribute 0 points to the matrix element M(i,j), the next closest pair, e.g. (j,k), add 1 point and the remaining pair add two points to M(i,k). The (dissimilarity) matrix obtained in this way resembles an ordinary distance table. If the phrases most distant and nearest in the question are interchanged, one obtains a similarity (data) matrix.
Instead of relying on a topographer we used an ordinary distance table as input for the program. The program we applied was KYST-2a, pronounced

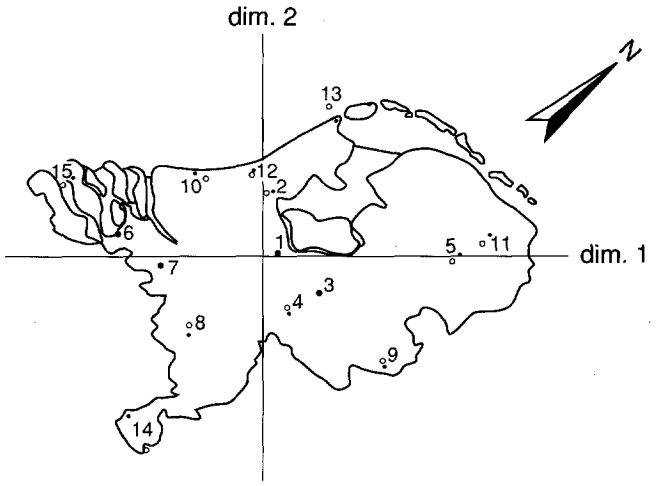

Figure C.1: Configuration with real (solid) and calculated (circles) locations.

'kissed', formed from the names Kruskal, Young, Shepard and Torgerson. It gives the coordinates of the cities in one or more dimensions. The analysis was carried out for both the metric case (with linear regression) and the non-metric case. In the latter, the actual number of miles was not used. However, the ranking order of the calculated interpoint distances should be, as far as possible, the same as the ranking order of the interpoint distances in the given distance matrix. The results of both the metric and the non-metric case were practically the same. Only the results of the latter will be discussed in the following.

All calculations were carried out in the Euclidian space (Minkowski's parameter = 2). A measure of the goodness of fit between both rankings is called stress, which can to some extent be compared with a least-squares sum in an ordinary fitting procedure. The stress value in this particular case is 0.249 for 1 dimension, 0.028 for 2 dimensions and 0.013 for 3 dimensions. It appears that a 2-dimensional fit is a good one. The decrease of stress in 3 dimensions is rather weak, while the deterioration due to a low dimensionality is obvious. The results of the calculations are plotted in Fig. C.1, the solid points are the real locations, whereas the small circles represent the calculated places. As the figure shows, in this particular case it is possible to derive metric data from non-metric input data. The orientation of the map is arbitrary, there is no real North-South axis. For convenience only, the contour of the Netherlands and a compass needle are drawn.

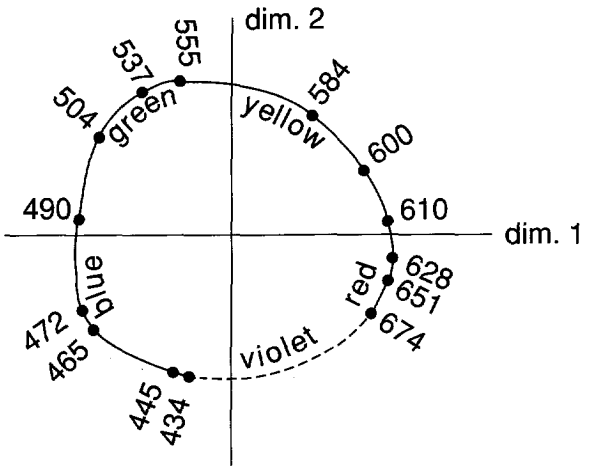

Figure C.2: Color circle (From Shepard's analysis of Ekman's data [177]).

A second example is from Ekman's [70] similarity judgement among 14 colors varying in hue. Subjects made ratings of qualitative similarity for each pair of combinations of colors ranging in wavelength from 434 nm to 674 nm. Shepard [177] applied a non-metric MDS procedure to the similarity ratings and extracted the underlying structure depicted by Fig. C.2. The underlying structure recovered from Ekman's similarity data was simply the conventional color circle, with the colors arranged along the smooth contour in order of increasing wavelength.

C.4 Notation

The input data, δ_{ij} or *proximities*, are numbers that indicate how similar or how different two objects are, or are perceived to be. The distance between the points i and j in the configuration which reflects the 'hidden structure' is denoted by d_{ij}. The basic concept takes the form that

$$f(\delta_{ij}) = d_{ij}, \tag{C.1}$$

where f is of some specified type. The discrepancy between $f(\delta_{ij})$ and d_{ij} is then:

$$f(\delta_{ij}) - d_{ij}. \tag{C.2}$$

An objective function which is called stress is:

$$S_1 = \sqrt{\frac{\sum_i \sum_j [f(\delta_{ij}) - d_{ij}]^2}{\sum_i \sum_j {d_{ij}}^2}} \tag{C.3}$$

The values $f(\delta_{ij})$ are often called fitted distances and denoted by $\hat{d_{ij}}$ (read 'd-hat-i-j'); they are also sometimes 'disparities'. When the *only* restriction for f is that it has to be monotonous, then the procedure is of the non-metric type.

C.5 Precautions concerning the solution

The interpretation and generation of the configuration map should both be monitored carefully, as undesirable results can occur which render any use of the configuration inadvisable [123]. *Before* attempting to interpret the configuration, the user should *always* check for these possibilities, beginning with the inspection of $\delta - d$ or the Shepard diagram (i.e., the scatter diagram of recovered distances versus data values). Some anomalous $\delta - d$ diagrams are discussed below. The relation between stress and significance is studied [213] and is discussed in the next section.

jaggedness of the fitted function. The function relating distances to data values will always be somewhat jagged. However, this function should in fact approximate a smooth and continuous curve. Since the user is assuming an underlying continuous function, a configuration associated with a step function is undesirable. There are, however, two possible remedies for step-functions. One is the possibility of a local minimum, hence different initial values have to be tried. The second remedy is to specify a stronger form of regression.

clumping of stimulus points. This is when several distinct objects occur at the same position in the Shepard diagram. The phenomenon is associated with undesirable behavior of the fitted function (i.e. the $\hat{d}$ values) in the region of the smallest dissimilarities. The antidote for undesirable clumping is to declare a different form of regression, perhaps with a preliminary transformation of the data as well.

degeneracy. A degenerate solution is an extreme case of both clumping (clustering) and jaggedness. Usually the stress is in this case zero, or nearly so. Hence, a very small stress value can indicate un utterly

useless solution instead of an exceptionally good one. This result often occurs when, for one or more subsets of the stimuli, the dissimilarities within that subset are smaller than the dissimilarities between stimuli in that subset and the remaining stimuli [177]. Possible solutions are, one, separate the clusters, two, scale them separately and combine them in a later run, three, use the FIX option, or, four, a form of regression stronger than monotone can be specified.

C.6 Significance of stress

For the representation of an arbitrary dissimilarity matrix by distances between n points in m dimensions, a probability $p(s)$, exists that a stress value $\leq s$ will be obtained by chance. For the determination of the probability distributions $p(s)$, the dissimilarity matrices contained the numbers from 1 to $0.5n(n-1)$ and were attributed randomly to the cells. In this way 100 'random scatters' were produced and analyzed by the MDS technique in various dimensions. The results are in Table C.1. The empty cells in Table C.1 correspond to conditions where more than 5% of the scatters have a stress smaller than 0.5%; it is advisable never to use MDS in these conditions.

Table C.1: The maximum stress in percentages which can be accepted at a significance level of $\alpha = 0.05$ for n points in m dimensions, from Table III of Wagenaar and Padmos [213].

	m=1	2	3	4	5
n=7	20	7	-	-	-
8	27.5	10	1.5	-	-
9	30.5	13	5.5	1	-
10	34	15	7	3	-
11	35	18	9.5	4.5	1
12	39.5	20.5	10	6.5	3.5

C.7 Bibliographic notes

A short but authoritative introduction to MDS is Kruskal's book [122]. A comprehensive survey of the development of MDS is that of Caroll & Arabie [45], which cites 334 references, mostly published during the nineteen

seventies. More recent are the surveys by Young [223], and (on general scaling) by Gescheider [85]. The latter is more about sensory and cognitive factors that affect psychophysical behavior than about measurement and computational aspects. A review intended for a wide, generally scientific audience, concerning the models and applications of MDS and cluster analysis, has been provided by Shepard [178].

Appendix D

Stimuli of loudness experiment

The program material of the loudness experiment described in Chapter 2 consisted of excerpts from Compact Discs, which, besides the quality of reproduction, has the advantage of allowing repetition of a fragment as often as required with only a brief interruption.

1. Pink noise, from the audio frequency test sample No. 3 Philips No. 410 055-2, which was also the noise source used in the first experiment.

2. Pop1, excerpt from: Early in the mornin' (0:0–0:25), from the album Step by Step, by Eddie Rabbit, Mercury No. 800 046-2. This fragment was used because of its wide spectrum and its rather large dynamic range in a short interval.

3. Pop2, excerpt from: It's only love (0:24–0:41), from the album: Let's Stick Together, by Bryan Ferry, E'G No. 821 521-2.

4. Jazz, excerpt from: Down home rag (0:13–0:31), by The Dutch Swing College Band, (track #8) from a demo disk of Philips No. 810 027-2.

5. Classic, excerpt from: part II of Tchaikovsky's 5th symphony (7:15–7:44) Chicago Symphony Orchestra (cond. by G. Solti), Decca No. 425 516-2.

Appendix E

Computation of A–D Weighting

A computer procedure is listed in order to compute the A, B, C and D-weighting functions [1]. The time constants for the filters for the A, B and C-weighting are from [231] and for the D-weighting from [230].

```
PROCEDURE abcdweight(f: REAL; VAR aw, bw, cw, dw: REAL);
(* f = freq. (Hz) *)
CONST
  ca = 8.0002266419162E-01; cb = 9.8767069950664E-01;
  cc = 6.6709544848173E-09; cd = 6.8966888496476E-05;
  p1c = 20.6; p2c = 12200.0; p1a = 107.7; p2a = 737.9;
  p1b = 158.5; p1d  = 282.7; p2d  = 1160.0; p3d1 = 1712.0;
  p3d2 = 2628.0; z1d1 = 519.8; z1d2 = 876.2;
  ps1c = sqr(p1c); ps2c = sqr(p2c);
  ps1a = sqr(p1a); ps2a = sqr(p2a); ps1b = sqr(p1b);
  ps1d  = sqr(p1d); ps2d  = sqr(p2d); ps3d1 = sqr(p3d1);
  ps3d2 = sqr(p3d2); zs1d1 = sqr(z1d1); zs1d2 = sqr(z1d2);
VAR f2, hw: REAL;
BEGIN
  f2:= sqr(f);
  cw:= f2 / (f2 + ps1c) / (f2 + ps2c) / cc;
  aw:= cw * f2 / sqrt((f2 + ps1a) * (f2 + ps2a)) / ca ;
  bw:= cw * f / sqrt(f2 + ps1b) / cb ;
  hw:= 1 / (sqr(ps3d1 + ps3d2 - f2) + 4 * f2 * ps3d1);
  hw:= hw * (sqr(zs1d1 + zs1d2 - f2) + 4 * f2 * zs1d1);
  dw:= f * sqrt(hw/(f2 + ps1d) / (f2 + ps2d)) / cd;
END;
```

Appendix F

The design of the digital stereo-base widening filters [1]

A rational approximation is calculated for a given measured impulse response by using a modified Prony's method. The method is incorporated in a computer program ('fit'). The main applications are the modeling of acoustic transfer functions and their equalization. As an example, a digital version of the analog widening circuit filters, discussed in Chapter 9, is given.

F.1 Introduction

The purpose of this Appendix is twofold. Firstly, the implementation of Prony's method [156, 175] in a computer program ('fit') is discussed in Section F.1.1. A second program ('curfit'), based on a modified Gauss-Newton method [86] (also used in Chapter 7), meant as a fine tuning tool, is used.

Secondly, as a vehicle of the program(s), a digital version of the widening circuit shown in Fig. 9.10 is used. Coefficients of the required filters, for three different sample frequencies (32 kHz, 44.1 kHz and 48 kHz) and two different topologies, are calculated.

F.1.1 Modeling using Prony's method

Direct modeling of a given impulse response $h(n)$ (the time-sampled version of the continuous version $h(t)$) is a rather complex affair due to notorious

[1] Parts of this Appendix have been published in [10].

computational requirements and stability problems. Therefore, an indirect method is applied, which is based on Prony's method [156, 175], which fits damped exponentials to the model.

Suppose there is a causal impulse response $h(n)$ with its z-transformed function $H(z^{-1})$, which has to be modeled by $h_a(n)$, as a rational model

$$H_a(z^{-1}) = \frac{B(z^{-1})}{A(z^{-1})} = \frac{\sum_{k=0}^{q} b_k z^{-k}}{1 + \sum_{k=1}^{p} a_k z^{-k}}, \tag{F.1}$$

or written as the ratio of products

$$H_a(z^{-1}) = h(0) \frac{\prod_{k=1}^{q}(1 - z_k z^{-1})}{\prod_{k=1}^{p}(1 - z_k z^{-1})}. \tag{F.2}$$

The error

$$E = \sum_{n=0}^{\infty} |h_a(n) - h(n)|^2 \tag{F.3}$$

has to be minimized. The resulting equations turn out to be nonlinear in the parameters a_k and for that reason the problem is uncoupled. Cross-multiplying in Eq. F.1 $H_a(z^{-1})A(z^{-1}) = B(z^{-1})$ yields a convolution of $h_a(n)$ and a_n. This can be written as the matrix product

$$\mathbf{Xa} = \mathbf{b}, \tag{F.4}$$

where

$$\mathbf{X} = \begin{pmatrix} h(0) & 0 & \dots & 0 \\ h(1) & h(0) & \dots & 0 \\ h(2) & h(1) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ h(q) & h(q-1) & \dots & h(q-p) \\ h(q+1) & h(q) & \dots & h(q-p+1) \\ \vdots & \vdots & \ddots & \vdots \\ h(q+p) & h(q+p-1) & \dots & h(q) \\ h(N-1) & h(N-2) & \dots & h(N-p-1) \end{pmatrix} \tag{F.5}$$

$$\mathbf{a^T} = (1\, a_1\, a_2\, \dots a_p) = (1\, \mathbf{a_1}), \tag{F.6}$$

$$\mathbf{b^T} = (b_0\, b_1\, b_2 \dots b_q\, 0 \dots 0), \tag{F.7}$$

and N is the number of known samples of $h(n)$. The basic idea is that after q samples the numerator of Eq. F.1 has no influence on $h_a(n)$. The matrix $\mathbf{X}$ can be partitioned as

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}'_{\mathbf{a}} \\ \mathbf{X}'_{\mathbf{b}} \end{pmatrix} = \begin{pmatrix} \mathbf{h}_{\mathbf{a}} & \mathbf{X}_{\mathbf{a}} \\ \mathbf{h}_{\mathbf{b}} & \mathbf{X}_{\mathbf{b}} \end{pmatrix}, \tag{F.8}$$

where

$$\mathbf{h}_{\mathbf{a}}^{\mathbf{T}} = (h(0) \ldots h(q))$$

and

$$\mathbf{h}_{\mathbf{b}}^{\mathbf{T}} = (h(q+1) \ldots h(N-1)).$$

Using Eq. F.4, Eq. F.6 and Eq. F.8, the following holds

$$\mathbf{X}'_{\mathbf{b}}\mathbf{a} = \mathbf{0}, \tag{F.9}$$

using the partitioning in Eq. F.8, Eq. F.9 can be written as

$$\mathbf{X}_{\mathbf{b}}\mathbf{a}_{\mathbf{1}} = -\mathbf{h}_{\mathbf{b}}. \tag{F.10}$$

This set of $N - q - 1$ equations with p unknowns is over-determined if $(N - q - 1) > p$ and can then be solved by a least-squares solution, e.g. using the method of normal equations [88]

$$\mathbf{C}\mathbf{a}_{\mathbf{1}} = \mathbf{X}_{\mathbf{b}}^{\mathbf{T}}\mathbf{X}_{\mathbf{b}}\mathbf{a}_{\mathbf{1}} = -\mathbf{X}_{\mathbf{b}}^{\mathbf{T}}\mathbf{h}_{\mathbf{b}}. \tag{F.11}$$

where $\mathbf{C}$ is the covariance matrix of $\mathbf{X}_{\mathbf{b}}$. Because $\mathbf{C}$ is positive-definite, Eq. F.11 can be solved for $\mathbf{a}_{\mathbf{1}}$ by Cholesky factorization of $\mathbf{C}$ and then applying forward and backward substitution. Once $\mathbf{a}_{\mathbf{1}}$ is known, $\mathbf{b}$ is calculated using

$$\mathbf{X}'_{\mathbf{a}}\mathbf{a} = \mathbf{b}. \tag{F.12}$$

Some special cases might arise. If $p = 0$, the model is an all-zero model, if $q = 0$, then it is an all-pole model. The all-zero model is known (in the statistical literature) as the moving average (MA) model, and the all-pole model is known as the autoregressive (AR) model.

Given the impulse response of an arbitrary signal, it is generally not possible to ascertain the identity of the system that generated the signal in terms of a set of poles and zeros. The problem is inherently nondeterministic, for a zero can be approximated arbitrarily closely by a large number of poles and vice versa.

F.2 Program results

The program flow chart is as shown in Fig. F.1. The input can be either direct to 'curfit', or to 'fit' with q, the order of the numerator of Eq. F.2, p the order of the denominator, m the length of the target impulse response, which makes part of the calculations and $h(t)$ the target impulse response. The target function, $h(t)$, serves as input for 'fit' and 'curfit'.

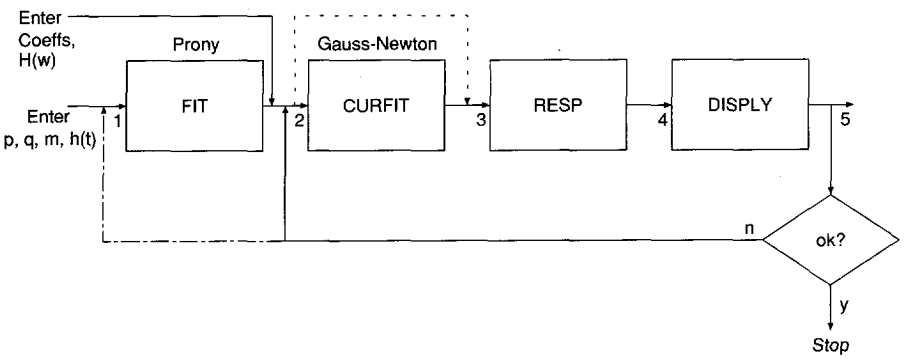

Figure F.1: Program calling sequence. The program 'RESP' evaluates the frequency response of the computed filters and the program 'DISPLY' displays the response.

F.2.1 The digital version of the widening circuit

The analog widening circuit filters given by Fig. 9.10 in Chapter 9, are fitted onto digital ones using the programs 'fit' and 'curfit' for three different sample frequencies and two filter topologies (Fig. F.4). The filter can be realized with FIR filters as well, however, the H_2 require a rather large number of taps. The filters H_1 and H_2 can be made more efficient by using the lattice shuffler, as shown in Fig. F.4-b. The magnitude and phase responses for the 32 kHz sampling frequency case are shown in Fig. F.2 and Fig. F.3. Informal listening tests revealed that there is virtually no difference between the digital filters and the analog version, besides a slightly higher bass response for the digital filters. However, this was a deliberately chosen design goal.

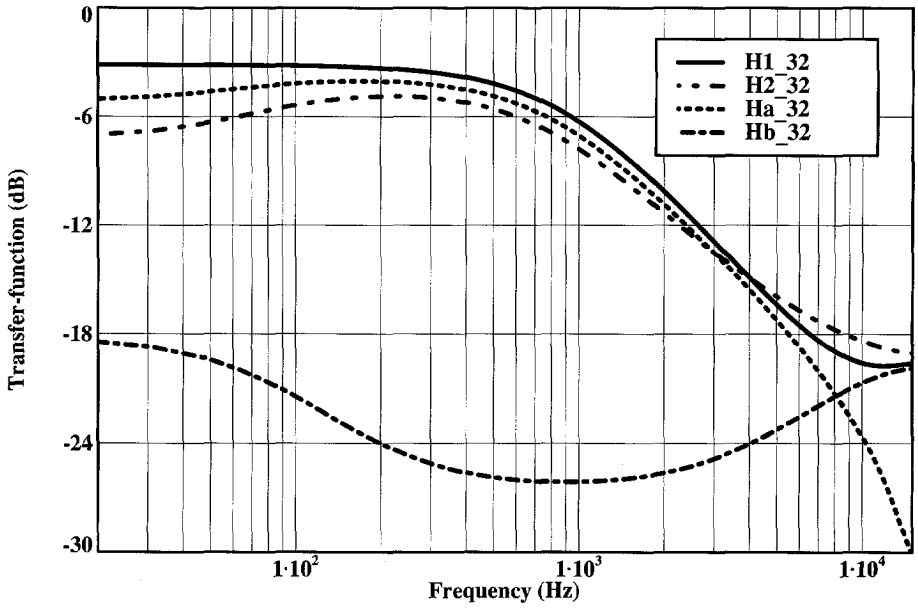

Figure F.2: Magnitude response of the widening circuit filters H1 ($p = q = 3$), H2 ($p = q = 4$), Ha ($p = q = 4$) and Hb ($p = q = 2$) for 32 kHz sampling frequency.

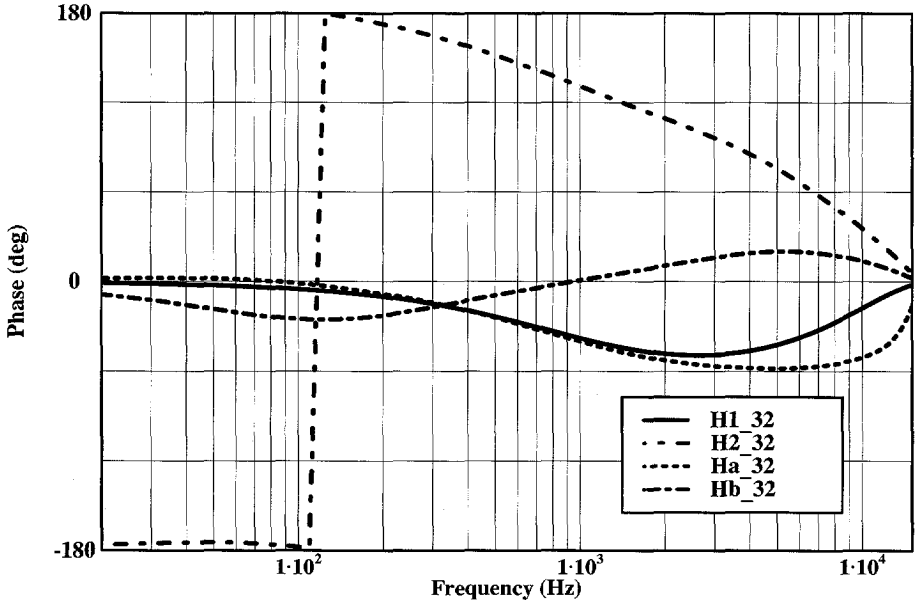

Figure F.3: Phase response of the widening circuit filters H1, H2, Ha and Hb for 32 kHz sampling frequency.

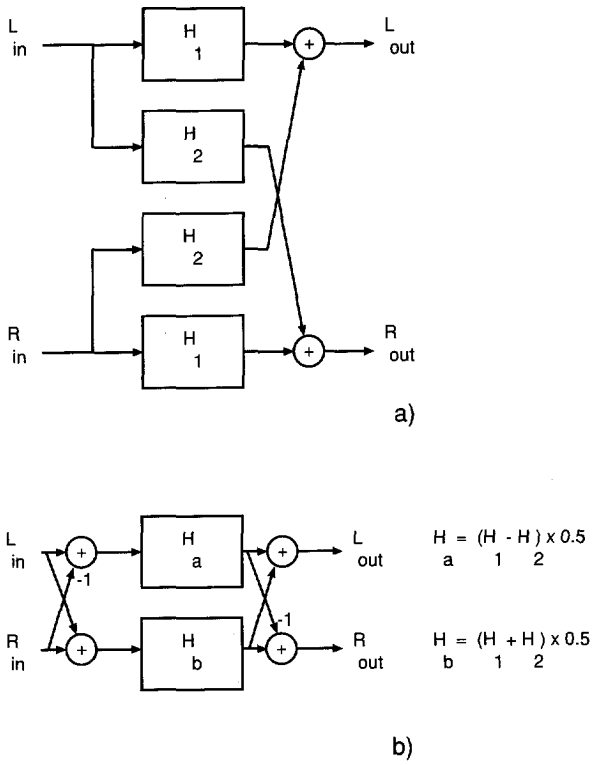

Figure F.4: Equivalence of two different system topologies for the widening circuit. a) The standard widening circuit topology. b) The lattice shuffler.

Appendix G

Calculation of zeros of a high-degree polynomial

A very large number of papers has been written describing iterative methods for finding the roots of polynomials. There are perhaps two major fields of interest associated with such methods. The first is concerned with the nature of convergence to a root or roots. The second is concerned with strictly practical matters and owes its existence mainly to the effect of rounding errors, which are inevitably made during the execution of an algorithm. Very striking examples may be found in [151, 219]. However, in some cases, when there is correlation among the coefficients of the polynomial, there is less sensitivity of the roots to errors in the coefficients, as is shown in [93]. As is shown at the end of this section, there can be a high correlation among the coefficients and the determination of the zeros is more robust than in the general case.

The currently used program is built around the routine C02AGF of the NAG [233]. The NAG routine itself is modified to a higher accuracy than the standard one. The program attempts to find all the roots of an n degree real polynomial equation

$$P(z) = \sum_{i=0}^{n} a_i z^{n-i} = \sum_{i=0}^{n} b_i z^{-i}. \tag{G.1}$$

It uses a modified Laguerre method described by the following iterative scheme

$$L(z_k) = z_{k+1} - z_k = \frac{-nP(z_k)}{P'(z_k) \pm \sqrt{H(z_k)}}, \tag{G.2}$$

where $H(z_k) = (n-1)[(n-1)(P'(z_k))^2 - nP(z_k)P''(z_k)]$, and z_0 is spe-

cified. The sign in the denominator is chosen so that the modulus of the Laguerre step at z_k, viz. $|L(z_k)|$, is as small as possible. The method can be shown to be cubically convergent for isolated (real or complex) roots and linearly convergent for multiple roots. The program is capable of finding all the zeros of a polynomial. The maximum degree depends on the machine accuracy and dynamic range, as is discussed in the following.

The machine hardware can only represent a subset of real numbers. This subset is denoted by **F**. A floating-point number system on a particular computer is characterized by four integers: the base β, the precision t, and the exponent range $[L, U]$. In particular, **F** consists of all numbers f of the form

$$f = \pm .d_1 d_2 \ldots d_t * \beta^e \quad 0 \leq d_i < \beta, \quad d_1 \neq 0, \quad L \leq e \leq U \tag{G.3}$$

together with zero. For a non-zero $f \in \mathbf{F}$ we have $m \leq |f| \leq M$ where

$$m = \beta^{L-1} \text{ and } M = \beta^U(1 - \beta^{-t}). \tag{G.4}$$

For three computers with various parameters, see Table G.1, the program was adapted. The maximum degree that could be achieved, n_{max}, depends on the floating-point system. The execution time can be calculated as

$$T_{exec} = \alpha n^2, \tag{G.5}$$

where n is the degree of the polynomial.

G.1 Location of zeros

For practical transfer functions with large n the locations of the zeros appear to be very close to $|z| = 1$. This can be shown analytically by the following example. If a damped cosine, $h(n) = a^n \cos \pi n$ is the system's impulse response, then the corresponding z-transform is

$$H(z) = \frac{z}{z + a}. \tag{G.6}$$

Synthetic division of Eq. G.6 yields

$$H(z) = \sum_{i=0}^{\infty} (-a/z)^i. \tag{G.7}$$

Truncating this series after K terms yields

$$H_K(z) = \sum_{i=0}^{K-1} (-a/z)^i = \frac{1 - (-a/z)^K}{1 + (a/z)}, \text{ for } |a/z| < 1. \tag{G.8}$$

The error due to the approximation is

$$|\epsilon| = |(H(z) - H_K(z))/H(z)| = |a|^K . \tag{G.9}$$

The zeros of $H_K(z)$ are

$$z_i = -ae^{-2ji\pi/K},\ i = 1, \ldots, K-2 \tag{G.10}$$

which are all on a circle of radius $|a|$ and exhibit a behavior similar to that in Fig. 10.2. In general, a is complex and, as an example, a pole at 0.707+j0.592, as shown in Fig. G.1, is approximated by 100 zeros, as plotted in Fig. G.2 (note the missing zero at the pole location).

Table G.1: Floating-point parameters and performance of some computers for the factorization of a polynomial and the type used in the FORTRAN program.

computer	t	U	L	β	M	m	n_{max}
IBM	56	252	-259	2	$7.23 * 10^{75}$	$5.40 * 10^{-79}$	~ 200
PC	52	1024	-1021	2	$1.79 * 10^{308}$	$2.23 * 10^{-308}$	~ 600
HP-1	52	1024	-1021	2	$1.79 * 10^{308}$	$2.23 * 10^{-308}$	~ 600
HP-2	112	16384	-16381	2	$1.19 * 10^{4932}$	$3.36 * 10^{-4932}$	> 2000

computer		α	Mflop/s [67]
IBM	IBM3081/K32, `REAL*8`	$2 * 10^{-5}$	~ 6
PC	386SX+387, f_{clock}= 20 MHz, `REAL*8`	$3 * 10^{-4}$	~ 0.2
HP-1	HP 9000/720, `REAL*8`, 50 MHz	$6 * 10^{-6}$	18
HP-2	HP 9000/735, `REAL*16`, 66 MHz	$7 * 10^{-4}$	24

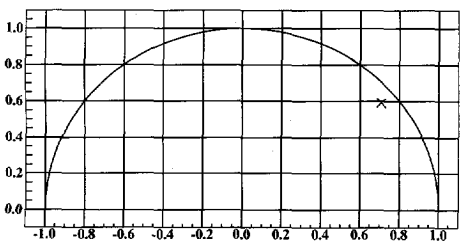

Figure G.1: Complex z-plane with one pole (x) at 0.707+j0.592.

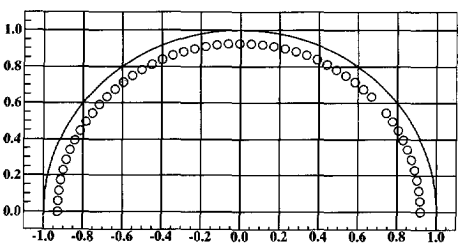

Figure G.2: Complex z-plane with 100 zeros (o), approximating one pole at 0.707+j0.592.

Bibliography

[1] R.M. Aarts. Algebraic expression for A-weighting response. *J. of Sound and Vibration*, 115(2):372, 1987.

[2] R.M. Aarts. Een nieuwe ontwerpmethode voor luidspreker overname-filters (A new method for the design of crossover filters). *NAG journaal*, (97):27–40, March 1989.

[3] R.M. Aarts. A new method for the design of crossover filters. *J. Audio Eng. Soc.*, 37(6):445–454, June 1989.

[4] R.M. Aarts. Some balanced incomplete block designs intended for loudspeaker listening tests. Technical report, Philips Research Labs, 1990. Nat.Lab. Report 6502.

[5] R.M. Aarts. Calculation of the loudness of loudspeakers during listening tests. *J. Audio Eng. Soc.*, 39(1/2):27–38, January/February 1991.

[6] R.M. Aarts. A comparison of some loudness measures for loudspeaker listening tests. *J. Audio Eng. Soc.*, 40(3):142–146, March 1992.

[7] R.M. Aarts. Time/intensity trading stereophony for (HD)TV and audio applications. In *14th International Congres on Acoustics (Beijing, China)*, September 1992. (session L8-3).

[8] R.M. Aarts. Enlarging the sweet-spot for stereophony by time-intensity trading. In *preprint 3473(C1-1) of the 94th AES Convention Berlin*, pages 1–15. Audio Eng. Soc., March 1993.

[9] R.M. Aarts. The generation of phantom sources. Stereo base widening ('incredible stereo') for sound reproduction. Technical report, Philips Research Labs, 1994. Nat.Lab. Report 6794.

[10] R.M. Aarts. Modelling of acoustic transfer functions and the realisation of digital filters: As applied to incredible stereo/sound. Technical report, Philips Research Labs, 1994. Nat.Lab. Report 6716.

[11] R.M. Aarts and A.J.M. Kaizer. Simulation of loudspeaker crossover filters with a digital signal processor. *J. Audio Eng. Soc.*, 36(3):115–121, March 1988.

[12] R.M. Aarts and S.M. Verduyn Lunel. *Symbolic Analysis of loudspeaker crossover filters.* J. Wiley, 1993. In: Computer Algebra in Industry Edited by: A.M. Cohen.

[13] M. Abramowitz and I.A. Stegun. *Handbook of mathematical functions.* Dover, 1972.

[14] G. Adams and S. Roe. Computer aided design of loudspeaker crossover networks. *J. Audio Eng. Soc.*, 30(7/8):496, July/Aug. 1982.

[15] F.J. Ampel and T. Uzzle. The history of audio and sound measurement. In *Preprint 3481 (L1-1) of the 94th AES Convention Berlin.* Audio Eng. Soc., March 1993.

[16] E. Bækgaard. A novel approach to linear phase loudspeakers using passive crossover networks. *J. Audio Eng. Soc.*, 25(5):284, May 1977.

[17] W.W. Rouse Ball and H.S.M. Coxeter. *Mathematical recreations and essays.* Dover, 13th edition, (1st ed. 1898) 1987.

[18] B.B. Bauer. Broadening the area of stereophonic perception. *J. Audio Eng. Soc.*, 8(2):91–94, April 1960.

[19] B.B. Bauer. Audibility of phase distortion. *Wireless World*, 80:27–28, March 1974.

[20] S. Bech. Perception of timbre of reproduced sound in small rooms: influence of room and loudspeaker position. *J. Audio Eng. Soc.*, 42(12):999–1007, December 1994.

[21] G.L. Beers and H. Belar. Frequency-modulation distortion in loudspeakers. *Proc. of the IRE*, 31(4):132–138, April 1943. Reprinted in J. Audio Eng. Soc. Vol. 29, No. 5, pp. 320-326, 1981.

[22] Alexander Graham Bell. Telegraphy, (filed morning of 14 February 1876) issued 7 March 1876. US patent 174 465.

[23] L.L. Beranek. *Acoustics.* McGraw-Hill, New York, 1954. (Reprinted by ASA 1986).

[24] R. Berkovitz and B. Edvardsen. Phase sensitivity in music perception. In *preprint 1294(J-4) of the 58th AES Convention New York.* Audio Eng. Soc., November 1977.

[25] F.A. Bilsen. On the influence of the number and phase of harmonics on the perceptibility of the pitch of complex signals. *Acustica,* 28(1):60–65, January 1973.

[26] J. Blauert. *Spatial hearing: The Psychophysics of Human Sound Localization.* The MIT Press, 1983.

[27] J. Blauert and P. Laws. Group delay distortions in electroacoustical systems. *J. Acoust. Soc. Am.,* 63(5):1478–1483, May 1978.

[28] P.J. Bloom and D. Preis. Perceptual identification and discrimination of phase distortion. In *ICASSP 83, Boston,* pages 1396–1399. IEEE, 1983.

[29] A.D. Blumlein. Improvements in and relating to sound-transmission, sound-recording and sound-reproducing systems. *J. Audio Eng. Soc.,* 6(2):91–99, 1958. British Patent Specification 394,325 (1931).

[30] P.M. Boers. The influence of antiphase crosstalk on the localization cues in stereo signals. In *preprint 1967(A-5) of the 73rd AES Convention Eindhoven,* pages 1–17. Audio Eng. Soc., March 1983.

[31] P.M. Boers and R.M. Aarts. Signal processing for binaural applications: Adapting loudspeaker signals for headphone reproduction. *NAG journaal,* (104):71–79, September 1990.

[32] J. Borwick, editor. *Loudspeaker and Headphone Handbook.* Butterworths London, 1988. Ch.4: Multiple driver loudspeakers by L. Fincham.

[33] R.C. Bose. Combinatorial problems of experimental design I incomplete block designs. *Proc. Symposia in Pure Mathematics (Am. Math. Soc.),* 34:47–68, 1979.

[34] F.H. Brittain. The loudness of continuous spectrum noise and its application to loudness measurements. *J. Acoust. Soc. Am.,* 11:113–117, 1939.

[35] F.H. Brittain. Loudspeakers: Relations between subjective and objective tests. *Journ. of the British Inst. of Radio Engineers*, 13:105–109, 1953.

[36] F.H. Brittain and D.M. Leaky. Two-channel stereophonic sound systems. *Wireless World*, pages 206–210, May 1956.

[37] P. Bruel. The enigma of sound power measurements at low frequencies. Technical report, B&K Technical review, 1978. No. 3.

[38] R.B. Bücklein. Hörbarkeit von Unregelmäßigkeiten in Frequenzgängen bei akustischer Übertragung. *Frequentz*, 16:103–108, 1962. Translated into English in J. Audio Eng. Soc. , Vol. 29, No. 3, March 1981, pp. 126-131.

[39] R. Bullock. Satisfying loudspeaker crossover constraints with conventional networks - old and new designs. *J. Audio Eng. Soc.*, 31(7):489, July/Aug. 1983.

[40] H.J. Butterweck. About the doppler effect in acoustic radiation from loudspeakers. *Acustica*, 63:77–79, 1987.

[41] T.J.F. Buunen. Two hypotheses on monaural phase effects. *Acustica*, 34:98–105, 1975.

[42] T.J.F. Buunen. *On the perception of phase differences in acoustic signals.* PhD thesis, Delft University of Technology, 1976.

[43] T.J.F. Buunen, J.M. Festen, F.A. Bilsen, and G. van den Brink. Phase effects in a three-component signal. *J. Acoust. Soc. Am.*, 55(2):297–303, February 1974.

[44] R.C. Cabot. Perception of nonlinear distortion. In *preprint C1004 of the 2nd AES International Conference Anaheim, California*, May 1984.

[45] J.D. Carroll and P. Arabie. Multidimensional scaling. *Ann. Rev. Psychol.*, 31:607–649, 1980.

[46] M. Clark. Audio technology in the United States to 1943 and its relationship to magnetic recording. In *Preprint 3481 (H2-1) of the 94th AES Convention Berlin.* Audio Eng. Soc., March 1993.

[47] P.M. Clarkson, J. Mourjopoulos, and J.K. Hammond. Spectral, phase, and transient equalization for audio systems. *J. Audio Eng. Soc.*, 33(3):127–132, March 1985.

[48] W.G. Cochran and G.M. Cox. *Experimental Designs.* John Wiley, 1957.

[49] J.M. Cohen. A new image enhancement technique for stereo reproduction. In *preprint 1671(D-7) of the 67th AES Convention New York*, pages 1–5. Audio Eng. Soc., October 1980.

[50] J.M. Cohen. Stereo image separation and perimeter enhancement, 1981. US patent 4,308,423.

[51] F. Cole, A. White, and L. Cummings. Complete classification of triad systems on fifteen elements. *Mem. Nat. Acad. Sci.*, XIV:57, 1925.

[52] J.H. Craig and L.A. Jeffress. Effect of phase on the quality of a two-component tone. *J. Acoust. Soc. Am.*, 34(11):1752–1760, November 1962.

[53] L. Cremer and H.A. Müller. *Principles and applications of Room Acoustics*, volume 1, Ch.II.1. Applied Science Publ. New York, 1982.

[54] Charles Cuttriss and Jerome Redding. Telephone, (filed 28 November 1877) issued 14 June 1881. US patent 242816.

[55] P. Damaske. Head-related two-channel stereophony with loudspeaker reproduction. *J. Acoust. Soc. Am.*, 50(4 (part 2)):1109–1115, 1971.

[56] P. Damaske. Zur richtungstreuen stereophonen Zweikanalübertragung. *Acustica*, 24:222–225, 1971.

[57] P. Damaske and V. Mellert. Ein Verfahren zur richtungstreuen Schallabbildung des oberen Halbraumes über zwei Lautsprecher. *Acustica*, 22:153–162, 1969/70.

[58] H.A. David. *The method of paired comparisons.* Ch. Griffin, London, 2nd edition, 1988.

[59] M.F. Davis. Loudspeaker systems with optimized wide-listening-area imaging. *J. Audio Eng. Soc.*, 35(11):888–896, 1987.

[60] E. de Boer. A note on phase distortion and hearing. *Acustica*, 11:182–184, 1961.

[61] K. de Boer. *Stereofonische geluidsweergave (Stereophonic sound reproduction).* PhD thesis, Delft University of Technology, December 1940.

[62] K. de Boer. The formation of stereophonic images. *Philips Technical Review*, 8(2):51–56, February 1946.

[63] K. de Boer. Stereophonic reproduction apparatus, 1946. US patent 2,610,694.

[64] R. de Wit, A.J.M. Kaizer, and F.J. Op de Beek. Numerical optimization of the crossover filters in a multiway loudspeaker system. *J. Audio Eng. Soc.*, 34(3):115, March 1986.

[65] J.A. Deer, P.J. Bloom, and D. Preis. Perception of phase distortion in all-pass filters. *J. Audio Eng. Soc.*, 33(10):782–786, October 1985.

[66] R.H.F. Denniston. Sylvester's problem of the 15 schoolgirls. *Discrete Mathematics*, 9:229–233, 1974.

[67] J.J. Dongarra. Performance of various computers using standard linear equations software. Technical report, Computer Science Department Univ. of Tennessee Knoxville, 1985. CS-89-85.

[68] J. Doyen. Constructions of disjoint Steiner triple systems. *Proc. of the Am. Math. Soc.*, 32(2):409–416, 1972.

[69] J. Eargle, editor. *Stereophonic techniques: An anthology of reprinted articles on stereophonic techniques.* Audio Eng. Soc., 1986. ISBN 0-937803-08-1.

[70] G. Ekman. Dimensions of color vision. *J. Psychology*, 38:467–474, 1954.

[71] L.R. Fincham. The subjective importance of uniform group delay at low frequencies. *J. Audio Eng. Soc.*, 33(6):436–439, June 1985.

[72] H. Fletcher and W.A. Munson. Loudness, its definition, measurement and calculation. *J. Acoust. Soc. Am.*, 5:82–108, 1933.

[73] F.J.M. Frankort. *Vibration and sound radiation of loudspeaker cones.* PhD thesis, Delft University of Technology, 1975.

[74] N.V. Franssen. *Some considerations on the mechanism of directional hearing.* PhD thesis, Delft University of Technology, July 1960.

[75] N.V. Franssen. *Stereofonica.* Philips Technische Bibliotheek, 1962. Available in Dutch, French, English and Spanish.

[76] A. Gabrielsson. Statistical treatment of data from listening tests on sound-reproducing systems. Technical report, Report TA No. 92, Dept. of Technical Audiology, Karolinska Institutet, Sweden, 1979.

[77] A. Gabrielsson. Perceived sound quality of reproductions with different sound levels. Technical report, Report TA No. 123, Dept. of Technical Audiology, Karolinska Institutet, Sweden, 1991.

[78] A. Gabrielsson et al. Perceived sound quality of reproductions with different frequency responses and sound levels. *J. Acoust. Soc. Am.*, 88(3):1359, September 1990.

[79] A. Gabrielsson and B. Lindström. Perceived sound quality of high-fidelity loudspeakers. *J. Audio Eng. Soc.*, 33(1/2):33, 1985.

[80] A. Gabrielsson, B. Lindström, and O. Till. Loudspeaker frequency response and perceived sound quality. *J. Acoust. Soc. Am.*, 90(2):707–719, August 1991.

[81] A. Gabrielsson, U. Rosenberg, and H. Sjögren. Judgements and dimension analyses of perceived sound quality of sound reproducing systems. *J. Acoust. Soc. Am.*, 5(4):854, April 1974.

[82] A. Gabrielsson and H. Sjögren. Perceived sound quality of sound reproduction systems. *J. Acoust. Soc. Am.*, 65(4):1019–1033, 1979.

[83] P. Garde. All-pass crossover systems. *J. Audio Eng. Soc.*, 28(9):575, Sept. 1980.

[84] M.B. Gardner. Image fusion, broadening, and displacement in sound location. *J. Acoust. Soc. Am.*, 46(2 (part 2)):339–349, 1969.

[85] G.A. Gescheider. Psychological scaling. *Ann. Rev. Psychol.*, 39:169–200, 1988.

[86] P.E. Gill and W. Murray. Algorithms for the solution of the nonlinear least-squares problem. *SIAM J. Numer. Anal.*, 15(5):977–922, Oct. 1978.

[87] J.L. Goldstein. Auditory spectral filtering and monaural phase perception. *J. Acoust. Soc. Am.*, 41(2):458–479, 1967.

[88] G.H. Golub and C.F. van Loan. *Matrix computations.* J. Hopkins University Press, 1989.

[89] P. Goossens. Position-independent stereo for HDTV by time/intensity trading. Technical report, Philips Research Labs, 1990. Nat.Lab. Report 6431.

[90] D.M. Green. *Profile analysis. Auditory discrimination.* Oxford University Press, 1988.

[91] R.A. Greiner and D.E. Melton. Observations on the audibility of acoustic polarity. *J. Audio Eng. Soc.*, 42(4):245–253, April 1994.

[92] J.P. Guilford. *Psychometric Methods.* McGraw-Hill New York, 1954.

[93] P. Guillaume, J. Schoukens, and R. Pintelon. Sensitivity of roots to errors in the coefficient of polynomials obtained by frequency-domain estimation methods. *IEEE Trans. on Instrumentation and Measurement*, 38(6):1050–1056, December 1989.

[94] M. Hall. *Combinatorial Theory.* J.Wiley, New York, 2nd. ed. edition, 1986.

[95] H. Hanani. On balanced incomplete block designs with blocks having five elements. *J. Combinatorial Theory (A)*, 12:184–201, 1972.

[96] H. Hanani. Balanced incomplete block designs. *Discrete Mathematics*, 11:225–369, 1975.

[97] V. Hansen and E.R. Madsen. On aural phase detection. *J. Audio Eng. Soc.*, 22(1):10–14, Jan./Feb. 1974.

[98] V. Hansen and E.R. Madsen. On aural phase detection: Part II. *J. Audio Eng. Soc.*, 22(10):783–788, December 1974.

[99] R.L. Hanson and W.E. Kock. Interesting effect produced by two loudspeakers under free space conditions. *J. Acoust. Soc. Am.*, 29(1):145, 1957. See Comment of Jeffress in: JASA Vol. 29(5), p. 655 and Comment of Leakey in: JASA Vol. 29(8), p. 966.

[100] T. Hatada, H. Sakata, and H. Kusaka. Psychophysical analysis of the "sensation of reality" induced by a visual wide-field display. *SMPTE Journal*, 89:560–569, August 1980.

[101] H. Helmholtz. *Die Lehre von den Tonempfindungen.* Friedr. Vieweg & Sohn, 1913 (6th printing). The 1st printing: 1863, English translation of the 4th German edition of 1877: On the Sensations of Tone, Dover, 1954.

[102] J.K. Hilliard. The history of stereophonic sound reproduction. *Proc. of the IRE*, 50:776–780, May 1962.

[103] J.K. Hilliard and H.R. Kimball. Dividing networks for loud speaker systems. *J. Audio Eng. Soc.*, 26(11):1–28, Nov. 1978. (Original in: Academy Research Council Technical Bulletin, Vol.26, No.11, March 1936).

[104] I.P. Howard and W.B. Templeton. *Human spatial orientation.* J. Wiley, 1966.

[105] F.V. Hunt. *Electroacoustics.* J. Wiley, 1954.

[106] F.V. Hunt. *Origins in acoustics: The science of sound from antiquity to the age of Newton.* Ac. Soc. of Am. through the Am. Inst. of Physics, 1992.

[107] A. Illényi and P. Korpássy. Correlation between loudness and quality of stereophonic loudspeakers. *Acustica*, 49(4):334–336, Dec. 1981.

[108] M. Jahn. Subjektive und objektive Bewertung von Maschinengeräuschen. *Acustica*, 16(3):175–186, 1965/1966.

[109] R.G.J. Janmaat. Analysis and simulation of RLC-loudspeaker-cross-over-filters, 1990. Master thesis with R.M. Aarts.

[110] C.P. Janse and J.M. van Nieuwland. Theoretical and experimental evaluation of the Nat. Lab. reverberation room. Technical report, Philips Research Labs, 1980. Nat.Lab. Report 5613.

[111] Lu Jia-Xi. On large sets of disjoint Steiner triples. *J. of Combinatorial Theory (Series A)*, 37:189–192, 1984.

[112] E.I. Jury. *Theory and application of the z-transform method.* Wiley, New York, 1964.

[113] A.J.M. Kaizer. *On the design of broadband electrodynamical loudspeakers and multiway loudspeaker systems.* PhD thesis, Eindhoven University of Technology, 1986.

[114] A.J.M. Kaizer. Modelling of the nonlinear response of an electrodynamic loudspeaker by a Volterra series expansion. *J. Audio Eng. Soc.*, 35(6):421–433, June 1987.

[115] L.E. Kinsler, A.R. Frey, A.B. Coppens, and J.V. Sanders. *Fundamentals of acoustics.* J. Wiley, 1982.

[116] R.E. Kirk. Learning, a major factor influencing preferences High-Fidelity reproducing systems. *J. Audio Eng. Soc.*, 5(4):238–241, 1957. Reprinted from the J. Acoust. Soc. Am. Nov. 1956, see also Letters to the Editor J. Audio Eng. Soc., Oct. 1960 p.269.

[117] J.A. Klaassen and S.H. de Koning. Motional feedback with loudspeakers. *Philips Technical Review*, 29(5):148–157, 1968.

[118] W. Klippel. Dynamic measurement and interpretation of the nonlinear parameters of electrodynamic loudspeakers. *J. Audio Eng. Soc.*, 38(12):944–955, December 1990.

[119] W. Klippel. The mirror filter – A new basis for reducing nonlinear distortion and equalizing response in woofer systems. *J. Audio Eng. Soc.*, 40(9):675–691, September 1992.

[120] W. Klippel. The nonlinear large-signal behavior of electrodynamic loudspeakers at low frequencies. *J. Audio Eng. Soc.*, 40(6):483–496, June 1992.

[121] S. Komiyama. Subjective evaluation of angular displacement between picture and sound directions for HDTV sound systems. *J. Audio Eng. Soc.*, 37(4):210–214, April 1989.

[122] J.B. Kruskal and M. Wish. *Multidimensional scaling, Quantitative applications in the social sciences.* Sage Publ., 10th ed. edition, 1983.

[123] J.B. Kruskal, F.W. Young, and J.B. Seery. How to use the KYST2A, a very flexible program to do multidimensional scaling and unfolding. BELL Telephone Laboratories, 1978.

[124] H. Kuttruff. On the audibility of phase distortions in rooms and its significance for sound reproduction and digital simulation in room acoustics. *Acustica*, 74:3–7, 1991.

[125] H. Larsen. The reverberation process at low frequencies. Technical report, B&K Technical review, 1978. No. 4.

[126] W.J.M. Levelt, J.P. van de Geer, and R. Plomp. Triadic comparisons of musical intervals. *The British Journal of Mathematical and Statistical Psychology*, 19 Part 2:163–179, November 1966.

[127] S.H. Linkwitz. Active crossover networks for noncoincident drivers. *J. Audio Eng. Soc.*, 24(1):2, January/February 1976.

[128] S.H. Linkwitz. Passive crossover networks for noncoincident drivers. *J. Audio Eng. Soc.*, 26(3):149, 1978.

[129] S. Lipshitz and J. Vanderkooy. A family of linear-phase crossover networks of high slope derived by time delay. *J. Audio Eng. Soc.*, 31(1/2):2, Jan. 1983.

[130] S.P. Lipshitz and J. Vanderkooy. The great debate: Subjective evaluation. *J. Audio Eng. Soc.*, 29(7/8):482–491, July/Aug. 1981.

[131] E. Lübcke, G. Mittag, and E. Port. Subjektive und Objektive Bewertung von Machinengeräuschen. *Acustica*, 14(2):105–114, 1964.

[132] F. Maffioli and U. Nicolao. Another approach to the ideal crossover: the energy filler. In *Preprint 2642 (L-6) of the 84th AES Convention Paris.* Audio Eng. Soc., March 1988.

[133] A.R. McKenzie and I.H. Flindell. Time-intensity trading in stereophony. In *Proceedings I.O.A.*, pages 117–133. Institute of Acoustics (UK), 1989. Volume 11, part 7.

[134] S. Mehrgardt and V. Mellert. Transformation characteristics of the external human ear. *J. Acoust. Soc. Am.*, 61(3):157–167, 1977.

[135] J. Merhaut. *Theory of electroacoustics.* McGraw-Hill, 1981.

[136] E. Meyer. Über das stereoakustische Hören. *Elektrotechnische Zeitschrift*, 22:805–807, 1925.

[137] J. Moir. Doppler distortion in loudspeakers. *Wireless World*, pages 65–69, April 1974.

[138] H. Møller, D. Hammershøi, C.B. Jensen, and M.F. Sørensen. Transfer characteristics of headphones measured on human ears. *J. Audio Eng. Soc.*, 43(4):203–217, April 1995.

[139] B.C.J. Moore. *An introduction to the psychology of hearing.* Academic Press Inc., London New York, 1982.

[140] B.C.J. Moore. Psychoacoustic considerations in the design and specification of sound-reproduction equipment. In E. Ozimek, editor, *Proc. of the International Symposium on Subjective and objective evaluation of sound*, Poznań, Poland, 1990. World Scientific.

[141] B.C.J. Moore and B.R. Glasberg. Formulae describing frequency selectivity as a function of frequency and level, and their use in calculating excitation patterns. *Hearing Research*, 28:209–225, 1987.

[142] B.C.J. Moore, S.R. Oldfield, and G.J. Dooley. Detection and discrimination of spectral peaks and notches at 1 and 8 kHz. *J. Acoust. Soc. Am.*, 85(2):820–836, February 1989.

[143] P.M. Morse and K.U. Ingard. *Theoretical acoustics.* McGraw-Hill, 1968.

[144] Y. Nomura. Estimation method of direct-radiator loudspeaker system parameters in low frequency range by nonlinear optimization technique. *Trans. Inst. Electr. and Communic. Eng.*, 368 A(5):504–511, 1986. (in Japanese).

[145] K. Ohgushi. Recent research on hearing in Japan. *J. Acoust. Soc. Japan (E)*, 5(3):127–133, 1984.

[146] H.F. Olson. *Acoustical engineering.* Van Nostrand, 1957.

[147] H.F. Olson. Analysis of the effects of nonlinear elements upon the performance of a back-enclosed, direct radiator loudspeaker mechanism. *J. Audio Eng. Soc.*, 10(2):156–162, 1962.

[148] R. Orban. Stereophonic image widening circuit, 1989. US patent 4,837,824.

[149] R.D. Patterson. A pulse ribbon model of monaural phase perception. *J. Acoust. Soc. Am.*, 82(5):1560–1586, November 1987.

[150] E. Paulus and E. Zwicker. Programme zur automatischen Bestimmung der Lautheit aus Tertzpegeln oder Frequenzgruppenpegeln. *Acustica*, 27(5):253–266, 1972.

[151] G. Peters and J.H. Wilkinson. Practical problems arising in the solution of polynomial equations. *J. Inst. Maths Applics*, 8:16–35, 1971.

[152] A.D. Pierce. *Acoustics, An Introduction to Its Physical Principles and Applications.* ASA, 1989.

[153] G. Plenge. On the behavior of listeners to stereophonic sound reproduction and the consequences for the theory of sound perception in a stereophonic sound field. In *preprint 2532(C-7) of the 83rd AES Convention New York*, pages 1–28. Audio Eng. Soc., October 1987.

[154] R. Plomp. *Timbre as a multidimensional attribute of complex tones*, chapter Frequency analysis and periodicity detection in hearing, pages 397–414. Sijthoff, 1970. R. Plomp and G.F. Smoorenburg, eds.

[155] R. Plomp and H.J.M. Steeneken. Effect of phase on the timbre of complex tones. *J. Acoust. Soc. Am.*, 46(2 (Part 2)):409–421, 1969.

[156] R. Prony. Essai expérimental et analytique. Sur les lois de la dilatabilité des fluides élastiques et sur celles de la force expansive de la vapeur de l'eau et de la vapeur de l'alkool, à différentes températures. *Journal de l'Ecole Polytechnique (Paris)*, 1(2):24–76, 1795.

[157] L.W. Rabiner and B. Gold. *Theory and Application of Digital Signal Processing.* Prentice-Hall, Englewood Cliffs, NJ, 1975.

[158] D.K. Ray-Chaudhuri and R.M. Wilson. Solution of Kirkman's schoolgirl problem. *Proc. Symposia in Pure Mathematics (Am. Math. Soc.)*, 19:187–204, 1971.

[159] J.W. Strutt (Lord Rayleigh). On the production and distribution of sound. *Philosophical Magazine*, VI:289–305, 1903. Reprinted in: Scientific Papers by Baron Rayleigh Vol. V, p.137, 1902-1910, 1912.

[160] P. Regalia and S. Mitra. A class of magnitude complementary loudspeaker crossovers. *IEEE ASSP*, 35(11):1509, Nov. 1987.

[161] C.W. Rice and E.W. Kellog. Notes on the development of a new type of hornless loud speaker. *Trans. Am. Inst. Elec. Engrs.*, 44:982–991, September 1925.

[162] L.F. Richardson and J.S. Ross. Loudness and telephone current. *J. Gen. Psychol.*, pages 121–164, 1916.

[163] S. Rosen. Monaural phase sensitivity: Frequency selectivity and temporal processes. In B.C.J. Moore and R.D. Patterson, editors, *Auditory frequency selectivity*, pages 419–427. Plenum, New York, 1986.

[164] E.E. Roskam. The method of triads for nonmetric multidimensional scaling. *Ned. Tijdschrift voor Psychol. en haar Grensgeb.*, 25, 1970.

[165] T.T. Sandel et al. localization of sound from single and paired sources. *J. Acoust. Soc. Am.*, 27(5):842–852, September 1955.

[166] B. Scharf. *Handbook of Perception Vol. IV, (ed. E.C. Carterette).* Academic Press New York, 1978.

[167] J.L. Schiano, C. Trahiotis, and L.R. Bernstein. Lateralization of low-frequency tones and narrow bands of noise. *J. Acoust. Soc. Am.*, 79(5):1563–1570, May 1986.

[168] S.S. Schiffman, M.L. Reynolds, and F.W. Young. *Introduction to Multidimensional scaling, Theory, Methods and Applications.* Academic Press, 1981.

[169] M. Schildbach. Audio technology in Berlin to 1943: Headphones and loudspeakers. In *Preprint 3485 (H2-6) of the 94th AES Convention Berlin.* Audio Eng. Soc., March 1993.

[170] M.R. Schroeder. Die statistischen Parameter der Frequenzkurven von Großen Räumen (The statistical parameters of frequency curves of large rooms). *Acustica*, 4(2):594–600, 1954.

[171] M.R. Schroeder and B.S. Atal. Computer simulation of sound transmission in rooms. *IEEE Intern. Conv. record pt.7*, pages 150–155, 1963.

[172] J. Schröter, G. Spikofski, and G. Theile. Messung des Diffusfeldübertragungsmaßes von Kopfhörern am Kunstkopf (Measurement of diffuse-field sensitivity level of headphones on a dummy-head). *Acustica*, 60(2):105–116, April 1986.

[173] P. Schuck. Design of optimized loudspeaker crossover networks using a personal computer. *J. Audio Eng. Soc.*, 34(3):124, March 1986. corr. 35(9):667, 1987 and 34(7/8):563, 1986.

[174] D. Shanefield. The great ego crunchers: Equalized, double-blind tests. *High Fidelity Magazine*, pages 57–61, March 1980.

[175] J.L. Shanks. Recursion filters for digital processing. *Geophysics*, 32(1):33–51, February 1967.

[176] J.H. Shaxby and F.H. Gage. Studies in the localization of sound. M.R.C. Spec. Rep. Ser. No.166, 1932.

[177] R.N. Shepard. The analysis of proximities: Multidimensional scaling with an unknown distance function. *Psychometrica*, 27:part I 125–140, part II 219–246, 1962.

[178] R.N. Shepard. Multidimensional scaling, tree-fitting and clustering. *Science*, 210:390–398, Oct. 1980.

[179] D.E.L. Shorter. A survey of performance criteria and design considerations for high-quality monitoring systems. *J. Audio Eng. Soc.*, 7(1):13–27, January 1959.

[180] K. Singhal and J. Vlach. *Computer Methods for Circuit Analysis and Design.* Van Nostrand Reinhold, 1983.

[181] R.H. Small. Constant-voltage crossover network design. *J. Audio Eng. Soc.*, 19(1):12–19, Jan. 1971. (Original in Proc. IREE Aust. Vol. 31 pp. 66-73, 1970).

[182] W.B. Snow. Basic principles of stereophonic sound. *J. Soc.of Motion Picture and Television Engineers*, 61:567–589, November 1953.

[183] H. Staffeldt. Correlation between subjective and objective data for quality loudspeakers. *J. Audio Eng. Soc.*, 22(6):402–415, July/Aug. 1974.

[184] J.C. Steinberg and W.B. Snow. Auditory perspective–Physical factors. *Electrical Engineering, (Journal and Transactions of the A.I.E.E)*, 53(1):12–17, January 1934. Reprinted in [69].

[185] G. Steinke. Stand und Enwicklungstendenzen der Stereofonie Teil I: Grundlagen – Probleme der konventionellen Wiedergabetechnologie. *Technische Mitteilungen des RFZ*, 28(1):1–10, 1984.

[186] G. Steinke. Stand und Enwicklungstendenzen der Stereofonie Teil II: Bifone Wiedergabeverfahren – Wiedergabe von Raumsignalen – neuere Systemlösungen. *Technische Mitteilungen des RFZ*, 28(2):25–32, 1984.

[187] S.S. Stevens. Procedure for calculating loudness: Mark VI. *J. Acoust. Soc. Am.*, 33(11):1577–1585, 1961.

[188] S.S. Stevens. Perceived level of noise by Mark VII and Decibels (E). *J. Acoust. Soc. Am.*, 51(2 (Part 2)):575–601, 1972.

[189] D.S. Stodolsky. The standardization of monaural phase. *IEEE Trans. on Audio and Electroacoustics*, AU-18(3):288–299, September 1970.

[190] H. Suzuki. Mutual radiation impedance of a double-disk source and its effect on the radiated power. *J. Audio Eng. Soc.*, 34(10):780–788, Oct 1986.

[191] H. Suzuki and J. Tichy. Sound radiation from convex and concave domes in an infinite baffle. *J. Acoust. Soc. Am.*, 69(1):41–49, 1981.

[192] H. Suzuki and J. Tichy. Sound radiation from an axissymmetric radiator in an infinite baffle. *J. Acoust. Soc. Japan (E)*, 3(3):167–172, 1982.

[193] Y. Tannaka and T. Koshikawa. Correlations between soundfield characteristics and subjective ratings on reproduced music quality. *J. Acoust. Soc. Am.*, 86(2):603–620, August 1989.

[194] Y. Tannaka, K. Muramori, M. Kohashi, and T. Koshikawa. Correlations between harmonic distortion, sound field characteristics and reproduced sound quality change in listening tests for loudspeakers. *J. Acoust. Soc. Japan (E)*, 11(1):29–42, 1990.

[195] G. Theile. On the naturalness of two-channel stereo sound. *J. Audio Eng. Soc.*, 39(10):761–767, October 1991.

[196] L.L. Thurstone. A law of comparative judgement. *Psychol. Review*, 34:273–286, 1927.

[197] F.E. Toole. Listening tests - turning opinion into fact. *J. Audio Eng. Soc.*, 30(6):431–445, June 1982.

[198] F.E. Toole. The acoustics and psychoacoustics of headphones. In *preprint C1006 of the 2nd AES International Conference Anaheim, California*, May 1984.

[199] F.E. Toole. Subjective measurements of loudspeaker sound quality and listener performance. *J. Audio Eng. Soc.*, 33(1/2):2–32, Jan./Feb. 1985.

[200] F.E. Toole. Loudspeaker measurements and their relationship to listener preferences Part I. *J. Audio Eng. Soc.*, 34(4):227–235, April 1986.

[201] F.E. Toole. Loudspeaker measurements and their relationship to listener preferences Part II. *J. Audio Eng. Soc.*, 34(5):323–348, May 1986.

[202] F.E. Toole. Subjective evaluation: Identifying and controlling the variables. In *The Proceedings of the AES 8th International Conference Washington D.C.*, pages 95–100. Audio Engineering Society, 1990.

[203] F.E. Toole and S.E. Olive. The modification of timbre by resonances: Perception and measurement. *J. Audio Eng. Soc.*, 36(3):122–142, March 1988.

[204] W.S. Torgerson. *Theory and methods of scaling.* J. Wiley, 1958.

[205] C. Trahiotis. Private communication, April 1995.

[206] T. Ueno and S. Ishii. Optimisation system for speaker network design. *Journal of the Acoustical Society of Japan*, 38(7), 1982. (in Japanese).

[207] L.J. van der Pauw. The trapping of acoustical energy by a conical membrane and its implications for loudspeaker cones. *J. Acoust. Soc. Am.*, 68(4):1163–1168, October 1980.

[208] J.M. van Nieuwland and C. Weber. Eigenmodes of nonrectangular reverberation rooms. *Noise Control Engineering Journal*, 13(3), 1979.

[209] D.W. van Wulfften Palthe. Doppler effect in loudspeakers. *Acustica*, 28:5–11, 1973.

[210] C. Vandenbulcke et al. An integrated digital audio signal processor. In *77th AES Convention 1985 March Hamburg.* Audio Eng. Soc., March 1985.

[211] J. Vanderkooy and S. Lipshitz. Power response of loudspeakers with noncoincident drivers - the influence of crossover design. *J. Audio Eng. Soc.*, 34(4):236–244, April 1986.

[212] E. Villchur and R.F. Allison. The audibility of doppler distortion in loudspeakers. *J. Acoust. Soc. Am.*, 68(6):1561–1569, December 1980.

[213] W.A. Wagenaar and P. Padmos. Quantitative interpretation of stress in Kruskal's multidimensional scaling technique. *British Journal of Mathematical Statistical Psychology*, 24:101–110, 1971.

[214] W.M. Wagenaars. Localization of sound in a room with reflecting walls. Institute for Perception Research Annual Progress report, 24, pp. 15–24, 1989.

[215] R.B. Welch and D.H. Warren. Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, 88(3):638–667, 1980.

[216] K. Wendt. *Das Richtungshören bei der Überlagerung zweier Schallfelder bei Intensitäts- und Laufzeitstereophonie.* PhD thesis, Technische Hochshule Aachen, February 1963.

[217] F.L. Wightman and D. Kistler. Headphone simulation of free-field listening. I: Stimulus synthesis. *J. Acoust. Soc. Am.*, 85(2):858–867, February 1989.

[218] F.L. Wightman and D. Kistler. Headphone simulation of free-field listening. II: Psychophysical validation. *J. Acoust. Soc. Am.*, 85(2):868–878, February 1989.

[219] J.H. Wilkinson. *Rounding errors in algebraic processes.* Prentice-Hall, N.J., 1963.

[220] R.J. Wilson, G.J. Adams, and J.B. Scott. Application of digital filters to loudspeaker crossover networks. *J. Audio Eng. Soc.*, 37(6):455–464, June 1989.

[221] B.J. Winer. *Statistical principles in experimental design.* McGraw-Hill Book Company, New York, 1962.

[222] J.R. Wolberg. *Prediction Analysis.* D. Van Nostrand Co. Inc., New York, 1967.

[223] F.W. Young. Scaling. *Ann. Rev. Psychol.*, 35:55–81, 1984.

[224] E. Zwicker. Ein Verfahren zur Berechnung der Lautstärke. *Acustica*, 10, 1960.

[225] E. Zwicker. *Psychoakustik.* Springer-Verlag, New York, 1982.

[226] E. Zwicker, H. Fastl, and C. Dallmayr. Letter to the editors: BASIC-program for calculating the loudness of sounds from their 1/3-oct band spectra according to ISO 532 B. *Acustica*, 55(1):63–67, April 1984.

[227] E. Zwicker and R. Feldtkeller. *Das Ohr als Nachrichtenempfänger.* S. Hirzel Verlag, Stuttgart, 1967.

[228] The telephone at the Paris Opera. Scientific American. pp. 422–423, December 1881.

[229] International standard ISO 532-1975(E), Acoustics - method for calculating loudness level, (1st ed.1975) 1977.

[230] International Electrotechnical Commission standard IEC 537, Frequency weighting for measurements of aircraft noise (D-weighting), 1976.

[231] International Electrotechnical Commission standard IEC 651, Sound level meters, 1979.

[232] NAG library manual. Numerical Algorithm Group Ltd., Oxford, 1983. Mark 10, Chapter E04.

[233] NAG library manual. Numerical Algorithm Group Ltd., Oxford, 1983. Mark 14, Chapter C02.

[234] International Electrotechnical Commission (IEC), Publication 268-13, Sound system equipment, part 13: Listening tests on loudspeakers., Geneva, Switzerland, 1985.

[235] International standard ISO 226-1987(E), Acoustics - normal equal-loudness level contours, 1987.

[236] Sound Quality Assessment Material, (recordings for subjective tests). European Broadcasting Union, 1988. no. 422 204-2.

List of Figures

List of Tables

Summary

On the design and psychophysical assessment of loudspeaker systems

This thesis discusses the design and assessment of loudspeaker[1] systems. In judging the quality of a loudspeaker, it is important to listen to it, either alongside other loudspeakers or by itself, using different parameters. In these listening tests, opinions that are formed about sound quality and stereo imaging are influenced by many factors in addition to the one that may be of specific interest. In the first chapter an overview is given of the relevant mechanical and psychoacoustic parameters which influence the assessment of loudspeakers.

One of these parameters is loudness. The loudness balancing of loudspeakers during listening tests is considered to be very important. One has to ensure that all the systems being tested have the same loudness levels. In Chapter 2, two loudness measurement methods will be discussed. A computer program is used to calculate loudness levels in order to determine the reproduction levels of loudspeakers during listening tests. The calculation results are compared with the results of listening tests.

In Chapter 3, simple weighting methods and ISO loudness models are compared with listener-adjusted loudness levels. For a loudness level of about 80 phon, the B-weighting appeared to be the best method, while the commonly used A-weighting is unreliable.

Another important parameter is the sound power response of loudspeakers. A particular sound power response of a loudspeaker may be more important than the traditional flat sound-pressure-level (SPL) design aim. Therefore, methods for measuring a loudspeaker's sound power response

[1] 'Loudspeaker' is understood here in a wide sense. In general, a loudspeaker consists of one or more transducers, or drivers, a cabinet with one or more volumes and either active or passive electronic circuitry.

will be considered in Chapter 4. One driver alone cannot radiate sufficient acoustical power over a large frequency range. To overcome this problem, a loudspeaker is usually constructed using two or more drivers, each optimized for a restricted frequency range. An electronic network (crossover filter) is used to combine these drivers for the full audio band (usually 20 Hz–20 kHz). The influence of the crossover filters on the sound power response will also be discussed.

By numerical optimization of the crossover filters of a multi-way loudspeaker, a set of component values is given for the crossover filter that satisfies the designer's demands and constraints to the greatest extent possible. In practice, however, many sets are found, e.g. because of different specifications of the optimization targets, different filter topologies or different initial values. These various sets can be differentiated by examining their numerical quality or by comparing their sensitivities to component value tolerances. After such a differentiation, several almost optimum solutions (sets of component values) remain, and such solutions with almost identical numerical quality can sound quite different. Therefore, a final check is needed to select the best-sounding solution. Obviously, such a final check is a listening test. In such a listening test it is desirable for the designer to switch between different crossover-filter candidates, without changing the loudspeakers or the location of the loudspeaker system. In Chapter 5, it will be discussed how these filters can be simulated in real-time, enabling the design engineer to select the best from a set of possible filters.

A method for investigating loudspeaker crossover filters in an analytical way will be presented in Chapter 6. A symbolic analysis program for RLC-loudspeaker-crossover filters has been designed and implemented using computer algebra. The program generates Pascal code that can be used to compute the input impedance and the voltage transfer function of the filter as a function of frequency, filter components and loudspeaker impedance.

In Chapter 7, a new method for designing crossover filters will be presented, in which they are both used as crossover and equalizing filters, which enables extra design properties. The desired system characteristic can be prescribed by a (complex) acoustic transfer function rather than an electrical one only. It may be derived from conventional filters or based on a measured one from a reference (favorable) system. The drivers of the experimental loudspeaker are preceded by digital filters, enabling the imitation of several different favorable loudspeakers. Double blind listening tests have been performed to verify subjectively the similarity between the

reference system and its experimental counterpart. Multidimensional scaling techniques are applied to represent the results of the listening tests.

It is impossible to imagine sound reproduction today without stereophonic techniques. Various kinds of improvements for stereophonic sound reproduction will be discussed. The listening area for good stereophonic sound reproduction is limited, but can be enlarged using time-intensity trading, as will be described in Chapter 8. The correction to compensate for the degradation of the stereophonic illusion due to off-center listening is investigated, both for small and wide loudspeaker bases. Data obtained by psychoacoustic experiments can be used to establish electronic filters driving a loudspeaker array, such that the required correction is achieved. The experimental results for video and (multimedia-) computer monitor applications are compared with the case of a much wider loudspeaker base, as in audio-only applications.

For some stereophonic set-ups the necessary distance between the loudspeakers cannot be realized. Chapter 9 will discuss how this problem can be solved to some extent. A stereo-base widening system producing a good effect, achievable with a simple analog circuit as well as with a digital one, is presented. The system is derived from the combination of the traditional approach using HRTFs (head-related transfer functions), and by using a simple model where ideal loudspeakers and an acoustically transparent subject's head are assumed. Both the HRTF approach as well as the simple model are discussed, and some special cases for the latter are considered. Merging these models, with some (experimentally determined) adaptations, led to a practical widening system. The system produces a pleasant and natural sound, particularly for voices, dialogues and 'natural' mixed recordings. The widening circuit is tuned in such a way that it gives the same tonal balance for common stereo recordings as for conventional stereo.

In some cases it is preferable to listen via headphones rather than via loudspeakers, but if one wishes to have the same perception as when listening via loudspeakers, some signal processing is required, as will be discussed in Chapter 10. Reproduction of loudspeaker signals via headphones leads to in-head localization. By convolving these signals with the impulse responses measured from loudspeakers to ear canals and deconvolving them with impulse responses measured from headphones to ear canals, binaural signals are generated which simulate a loudspeaker set-up and make better localization possible. In general, the inverse of the system describing an outer ear and a headphone is not stable and a modification is necessary to make implementation in a digital signal processor possible. A new tech-

nique is described for modifying the measured functions and for allowing implementation with minimal errors in the magnitude of the functions.

Ronald M. Aarts

Samenvatting

Het ontwerpen en psycho-fysisch beoordelen van luidsprekersystemen

Dit proefschrift behandelt het ontwerpen en het beoordelen van de kwaliteit van luidsprekersystemen. Bij het beoordelen van luidsprekers[1] door middel van luisterproeven is het van belang om deze te vergelijken met andere luidsprekers, of steeds dezelfde luidspreker te gebruiken terwijl verschillende parameters worden gewijzigd. Bij deze proeven worden subjectieve oordelen, zoals de geluidskwaliteit en het stereobeeld, beïnvloed door veel factoren.

In het eerste hoofdstuk is een overzicht gegeven van de relevante mechanische en psycho-akoestische parameters die de kwaliteitsbeoordeling beïnvloeden. Een van deze parameters is luidheid. Het balanceren van de luidheid tussen de luidsprekers onderling voorafgaand aan een luisterproef is erg belangrijk: men moet ervoor zorgen dat alle systemen even luid klinken.

In hoofdstuk 2 worden twee luidheidsmeetmethodes besproken. Een computerprogramma is gebruikt om de luidheidsniveaus te berekenen die tijdens een luisterproef worden gebruikt. De berekende niveaus zijn vergeleken met de resultaten van luisterproeven.

In hoofdstuk 3 worden simpele weegfunctie-methodes en de ISO-luidheidsmodellen vergeleken met door luisteraars ingestelde luidheidsniveaus. Voor een luidheidsniveau van ongeveer 80 phon, is de B-wegingsmethode de beste gebleken, terwijl de gewoonlijk gebruikte A-wegingsmethode onbetrouwbaar is gebleken.

Een andere belangrijke parameter is de geluidsvermogensresponsie van

[1] 'Luidspreker' is hier in brede zin bedoeld. In het algemeen bestaat een luidspreker uit één of meer transducenten, een behuizing met één of meer afzonderlijke volumes en actieve of passieve elektronica.

luidsprekers. Er zijn aanwijzingen dat een bepaalde geluidsvermogensresponsie van een luidspreker belangrijker is dan het gangbare ontwerpcriterium van een vlakke geluidsdrukkarakteristiek. Daarom worden methodes om de vermogenskarakteristiek te meten behandeld in hoofdstuk 4. Een enkele transducent kan in het algemeen onvoldoende akoestisch vermogen afstralen over een groot frequentiegebied. Om dit probleem op te lossen wordt een luidspreker meestal opgebouwd uit twee of meer transducenten, elk geoptimaliseerd voor een beperkt frequentiegebied. Een elektronisch netwerk (overname- of wisselfilter) wordt gebruikt om de verschillende transducenten te combineren zodat zij samen de gehele audio-frequentieband bestrijken (gebruikelijk van 20 Hz tot 20 kHz). De invloed van deze overnamefilters op de geluidsvermogensresponsie wordt eveneens behandeld.

Door numerieke optimalisatie van een overnamefilter voor een meerwegsluidsprekersysteem worden de componentwaarden bepaald die zo goed mogelijk aan de eisen en randvoorwaarden van de ontwerper voldoen. In de praktijk worden echter vele verschillende verzamelingen van oplossingen gevonden, bijvoorbeeld door verschillende specificaties van de ontwerpcriteria, verschillende filtertopologieën of verschillende startwaarden. Deze verschillende verzamelingen onderscheiden zich door hun numerieke kwaliteit of door hun verschillende gevoeligheden voor de toleranties van de componenten. Na een selectie blijven een aantal bijna optimale oplossingen (verzameling van componentenwaarden) over, die ondanks een bijna gelijke numerieke kwaliteit toch verschillend kunnen klinken. Daarom is een luisterproef nodig om de beste oplossing te kiezen. Tijdens een luisterproef is het voor de ontwerper wenselijk om tussen verschillende kandidaatovernamefilters te kunnen omschakelen zonder daarbij de luidsprekers of hun positie te veranderen. In hoofdstuk 5 wordt besproken hoe deze filters kunnen worden gesimuleerd, zodat de ontwerper de beste kan kiezen uit een verzameling van kandidaat-filters.

In hoofdstuk 6 wordt een analytische onderzoeksmethode voor overnamefilters behandeld. Een symbolisch analyse-programma voor *RLC*-luidspreker-overnamefilters is ontworpen en geïmplementeerd door gebruik te maken van computeralgebra. Het programma genereert een (Pascal-) computerprogramma dat de ingangsimpedantie en de overdrachtsfunctie als functie van de frequentie, overnamefiltercomponenten en luidsprekerimpedantie bepaalt.

In hoofdstuk 7 wordt een nieuwe ontwerpmethode voor luidsprekerovernamefilters behandeld, waarbij deze filters niet alleen als overnamefilters maar ook als egaliserende filters worden gebruikt, hetgeen extra ontwerpvrijheden biedt. De vereiste systeemkarakteristiek kan worden gegeven

in de vorm van een (complexe) akoestische overdrachtsfunctie, een prototype of een referentiesysteem, in plaats van alleen in de vorm van een elektrische overdrachtsfunctie. De gewenste karakteristiek kan worden afgeleid van gangbare filters of gebaseerd zijn op een gemeten karakteristiek van een goed klinkend referentiesysteem. De transducenten worden dan voorafgegaan door (digitale) filters, zodat verschillende goed klinkende systemen worden gesimuleerd, gebruikmakend van één en dezelfde luidspreker. Dubbelblinde luisterproeven zijn uitgevoerd om langs subjectieve weg de gelijkenis van het referentiesysteem en de experimentele systemen te verifiëren. Meerdimensionale schalingstechnieken zijn gebruikt om de resultaten van die luisterproeven te representeren.

Men kan zich tegenwoordig moeilijk geluidreproduktiesystemen voorstellen zonder stereofonische technieken. Diverse methodes ter verbetering van stereofonische weergave worden geïntroduceerd. Het luistergebied voor een goede stereofonische weergave is beperkt maar kan, zoals beschreven in hoofdstuk 8, door gebruik te maken van tijd/intensiteit-uitwisseling worden vergroot. De correctie die nodig is om te compenseren voor het luisteren op een plaats buiten de luisteras, is onderzocht zowel voor een kleine als voor een grote afstand tussen beide luidsprekers. De zo bepaalde gegevens kunnen worden gebruikt om elektronische filters te construeren waarmee een groep van luidsprekers zo wordt aangedreven dat de nodige correctie wordt bereikt. De resultaten van video- en (multimedia-) computermonitortoepassingen zijn vergeleken met die van systemen met een veel grotere luidsprekerafstand, zoals in audio-toepassingen.

Voor sommige stereofonische opstellingen kan de nodige afstand tussen de luidsprekers niet gerealiseerd worden. Hoofdstuk 9 behandelt hoe dit probleem kan worden ondervangen. Een stereo-basisverbredingssysteem met een goed effect, dat zowel met een eenvoudige analoge schakeling als met een digitaal filter kan worden uitgevoerd, wordt gepresenteerd. Het systeem is afgeleid van een benadering met hoofd-gerelateerde overdrachtsfuncties met gebruikmaking van een eenvoudig model waarbij de luidsprekers ideaal en het hoofd van de luisteraar akoestisch transparant worden verondersteld. Beide modellen worden behandeld en sommige speciale gevallen worden beschouwd. Een tussenvorm van deze modellen, met enige langs experimentele weg bepaalde aanpassingen, heeft geleid tot een praktisch verbredingssysteem. Het systeem produceert een prettig en natuurlijk klinkend geluid, speciaal voor stemgeluiden en 'natuurlijk' gemengde opnames. Het verbredingssysteem is dusdanig afgeregeld dat het timbre hetzelfde is als bij gewone stereofonische weergave.

In sommige gevallen is het prettiger om via hoofdtelefoons te luisteren

dan via luidsprekers, terwijl men dezelfde geluidsperceptie wenst als bij het luisteren via luidsprekers. Hiervoor is signaalbewerking nodig, die behandeld wordt in hoofdstuk 10. Geluidsweergave van luidsprekersignalen via hoofdtelefoons leidt tot de zgn. 'in-hoofd-lokalisering'. Door het convolueren van de signalen met de impulsresponsies gemeten van luidspreker naar de oorkanalen en deze te deconvolueren met de impulsresponsie gemeten van hoofdtelefoon naar de oorkanalen, kunnen binaurale signalen worden afgeleid die de luidsprekerweergave-situatie simuleren en daardoor een betere lokalisering mogelijk maken. In het algemeen is de inverse overdrachtsfunctie van hoofdtelefoon naar oor niet stabiel. Om de implementatie in een digitaal signaalbewerkingssysteem mogelijk te maken is daarom een modificatie van de overdrachtsfunctie nodig. Een nieuwe techniek wordt behandeld die het modificeren van de gemeten functies mogelijk maakt met minimale fouten in de absolute waarden van de te realiseren functies.

Ronald M. Aarts

Biography

Ronald Aarts was born in 1956, in Amsterdam, the Netherlands. He received a BSc degree in electrical engineering, from the College of Advanced Technology of Alkmaar, the Netherlands, in 1977, where he was involved in a project regarding ultrasonic bloodflow measurement. In 1977 he joined the Optics group of Philips Research Laboratories, Eindhoven, the Netherlands. There he was engaged in research into servos and signal processing for use in both Video Long Play players and Compact Disc players. In 1984 he joined the Acoustics group of the Philips Research Laboratories and was engaged in the development of CAD tools and signal processing for loudspeaker systems. He has published a number of technical papers and reports and is the holder of several patents in the afore mentioned fields. He was a member of the organizing committee and chairman for various conventions.