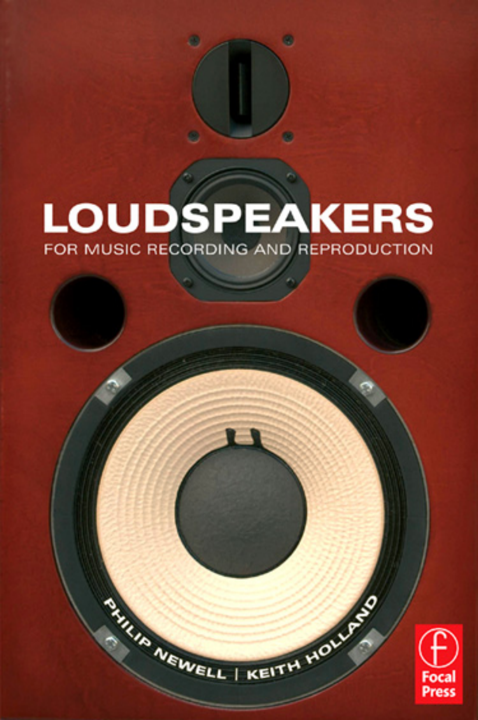


LOUDSPEAKERS

FOR MUSIC RECORDING AND REPRODUCTION

PHILIP NEWELL | KEITH HOLLAND

Focal Press

Loudspeakers

This page intentionally left blank

Loudspeakers

For Music Recording and Reproduction

Philip Newell and Keith Holland

ELSEVIER

AMSTERDAM • BOSTON • HEIDELBERG • LONDON
NEW YORK • OXFORD • PARIS • SAN DIEGO
SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO

Focal Press is an imprint of Elsevier

Focal Press

Focal Press is an imprint of Elsevier
Linacre House, Jordan Hill, Oxford OX2 8DP, UK
30 Corporate Drive, Suite 400, Burlington, MA 01803, USA

First edition 2007

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloguing in Publication Data
A catalogue record for this book is available from the Library of Congress

ISBN-13: 978-0-240-52014-8
ISBN-10: 0-2405-2014-9

For information on all Focal Press publications
visit our website at www.focalpress.com

Printed and bound in Great Britain
07 08 09 10 10 9 8 7 6 5 4 3 2 1

Typeset by Integra Software Services Pvt. Ltd, Pondicherry, India
www.integra-india.com

Contents

This page intentionally left blank

About the authors

Philip Newell began working professionally with loudspeakers in 1966, in the maintenance department of a shop selling high fidelity sound reproduction equipment in the town of his birth, Blackburn, England. Within a year he had begun to work for the Mecca chain of dance halls as a live-sound engineer. By 1970 he was working at a recording studio in south London, where he designed his first studio monitoring system. Philip moved to Pye Records in late 1970, when Pye was one of the UK's premier record companies with a large recording complex near Marble Arch, in central London. He worked primarily as a studio maintenance engineer, but was also involved in many recordings, and then moved to a fledgling Virgin Records organisation in late 1971 as chief recording engineer. In 1973 Philip co-founded The Manor Mobile with partners Richard Branson and Nik Powell, putting on the road what was probably at the time the world's most advanced mobile recording studio. From 1974 to 1982 he was technical director of the whole recording divisions of Virgin, but remained working as a recording engineer and record producer during the entire period, feeling that it was better to keep in practice than to concentrate solely on administration if the most balanced decisions were to be made.

After selling his shares in Virgin in 1982, to concentrate on flying seaplanes, a chance meeting in London in 1983 with Alex Weeks, the then owner of Reflexion Arts, led them to deciding to start an acoustics and monitoring branch of the company, for which Philip designed a series of studio monitor loudspeakers during a period of building many recording studios. It was during this time that he met Keith Holland at the Institute of Sound and Vibration Research (ISVR), a department of Southampton University, in the UK. Together they worked on the design of improved mid-range horn loudspeakers for studio monitoring systems. Although Philip left Reflexion Arts in 1988 to pursue a free-lance career, his collaboration with them *and* Keith Holland and the ISVR has lasted the twenty years to the publication of this book.

Philip Newell's 40 year career with loudspeakers has encompassed their use in domestic hi-fi, live sound, musical instrument amplification, music recording and mixing, film studios, television studios, video post-production rooms and many other uses. He has worked with them as a live sound engineer, recording and mixing engineer and record producer. The majority of his work since leaving Virgin has been as a designer of music recording studios, cinema mixing theatres and live performance spaces, and he is still involved in the design and development of high performance loudspeaker systems.

Philip is a Fellow if the Institute of Acoustics, a Member of the Audio Engineering Society, a member of the Seaplane Pilot's Association and

British Mensa. He has written five books on acoustics and electro-acoustics, and has published over 100 related articles, journal papers and conference papers. Since 1992 he has lived in Spain, and during the course of his work, he has travelled to over 30 countries. His recording career was very musically varied, from the Duke Ellington Orchestra to Queen; from The Who to The Warsaw Philharmonic Orchestra; from Mike Oldfield to John Cale; to English brass bands and Welsh choirs.

Dr Keith Holland is currently a Lecturer at the Institute of Sound and Vibration Research (ISVR), University of Southampton, UK, where he has been in full-time employment since 1993, and from where he obtaining a BSc in Engineering Acoustics and Vibration in 1987, and a PhD on horn loudspeakers in 1993. Since 1990, Keith has taught Electroacoustics and Audio Systems to under- and post-graduate students at the ISVR, and for ten of those years, also to the Tonmeister students at the University of Surrey.

As a researcher and academic at the ISVR, Keith has been involved in a large number of acoustic research projects on a wide variety of topics, which include: acoustic source location, advanced measurement techniques, aircraft cabin noise, cathedral acoustics, crossovers, drive-unit characterisation, duct acoustics, engine exhaust noise, fluid dynamics, guitar amplifiers, horn loudspeakers, inverse methods, jet noise, loudspeaker arrays, loudspeaker cables, loudspeaker directivity, microphone array processing, musical acoustics, nonlinear acoustics, numerical acoustic modelling, psychoacoustics, recording studio acoustics, room acoustic measurement, sound absorption, spacial imaging, tyre noise, vibration transducer development, vibroacoustic reciprocity and virtual audio. He is the author/co-author of over 60 papers, of which more than 30 are audio-related, and was the author of a series of 36 monthly objective monitor loudspeaker reviews published in Studio Sound magazine.

A healthy interest in all things audio, and loudspeakers in particular, was inherited from his father, Peter Holland, who, throughout the 1950s, 60s and beyond, spent many hours tinkering and experimenting with valves, transistors, tape recorders, loudspeakers and a lot of wire - with sometimes remarkable results! Keith followed suit, and built his first complete audio system in about 1967, using a Garrard turntable, home-made amplifiers (3 watts per channel) and speakers based around 9 inch × 5 inch elliptical drivers which came out of an old radiogram. By the early 1970s, this system had evolved into a room-dominating monster, including a 9 cubic foot sand-filled corner cabinet from a book by Gilbert Briggs (the founder of Wharfedale). While experimenting with crossovers for the giant woofer, it became very obvious that the more components that were used in the crossover, the worse the bass sounded. Using a 10-channel graphic equaliser, with the 30 and 60 Hz sliders fully up and the rest down on the left channel, and the opposite on the right, a makeshift mono active system was created. The benefits of active crossovers were immediately heard as the sound quality of this system was vastly superior to that from any attempt to use passive crossovers.

A few years later, a chance meeting with Ian Piper of ICP Electronics resulted in a collaboration on the design and construction of an actively-driven PA loudspeaker system which, to the amazement of many clients, delivered a sound quality and level far beyond what was expected using such apparently modest components. This system evolved over a period of about 10 years, during which time Ian taught Keith a great deal about the skills of live sound mixing, and between them they set up and mixed hundreds of live acts, some very good, some awful, and some were even quite famous!

By 1984, Keith had spent six years working in the manufacturing industry since being awarded a HND in Mechanical Engineering at (the then) Bournemouth College of Technology in 1978. He had become a skilled machinist, but his strong interest and curiosity about all things acoustic drew him to leave a well-paid job and go back to school at the ISVR to learn more. Three years later he was awarded a BSc in Engineering Acoustics and Vibration and received the prize for academic performance in his final year. It was during this year that Keith first met Philip Newell.

Keith had mentioned to his academic supervisor, Professor Frank Fahy, that he was interested in horn loudspeakers, and, quite by chance, Philip Newell was developing monitor loudspeaker systems and had been making enquiries at Southampton University to find out if anyone was doing research into horns. Philip was put in touch with Frank, and within a very short time Keith was beginning a three-year doctoral research project on horn loudspeakers, sponsored by Philip. A key result of the research is the AX2 horn used in the current Reflexion Arts monitor loudspeakers. To date, the work has produced, or inspired a total of 26 papers jointly authored by Philip Newell and Keith Holland, and, as this book proves, their collaboration is still as strong as ever.

Keith is a Member of the Institute of Acoustics and is a regular contributor to the Institute's Reproduced Sound series of conferences. He is also a member of the Audio Engineering Society. He continues to maintain and build on his interest in loudspeakers through many teaching, research, consultancy and hobby activities.

Keith is married to Sharon, whom he met in 1984. They live in their native southern England, and together they have two children, Bethany and Thomas. As a family, they enjoy camping, boating, walking and, of course, listening to music.

Sergio Castro, who orchestrated all the figures for this book, was born in 1955, in Oporto, Portugal and soon found his great interest in music. At the age of 12 he bought his first acoustic guitar and a few months later, with a turntable ceramic cartridge, he added electric amplification to the instrument through an old Telefunken radio set.

When he was 13, he was playing drums professionally with a local rock band and from then on, and simultaneously with his high school and his university studies later, he has been a full time professional musician, playing bass and guitar with some of the most relevant bands both in Portugal and in Spain until the early 90s.

During 1983 he built and operated the first multitrack recording studio in Oporto, and in 1985 he initiated the Planta Sonica Studios project, later to be designed and built by Philip Newell. During the following years, in this studio, he recorded and/or produced most of the Pop, Folk and Rock acts in the region (Galicia).

His interest in loudspeakers goes back to the time he had to choose his bass and guitar amplification, when he started experimenting with different driver types in order to tailor the timbre of the instruments he played. Modifying some of the off-the-shelf available amplifiers, and experimenting with bigger speaker boxes in order to achieve extended bass, he attempted to understand more about loudspeaker behaviour, aiming to improve his live gear without huge money investment.

In the late 70s he contributed to the PA loudspeaker designs at SEC, one of the pioneering pro-audio systems manufacturers in Portugal.

His involvement in the studio business, both as a producer and as a recording engineer, led him to investigate further the concept of studio monitors, trying to understand the audible differences he could then find when travelling among different studios and different listening environments.

Since then, his interest in the matter has grown and he became an Associate Member of the Institute of Acoustics in the UK and a member of the AES (currently a Member of the Board of Directors of the Spanish section). He studied acoustics at Vigo University, Spain, where he took a degree in Applied Acoustics.

Today he shares his time between being the head designer and co-owner of Artesania de BluesBox, the manufacturer of the recently introduced and successful brand of guitar, bass and installation PA speakers cabinets, as well as being the managing director of Reflexion Arts. Reflexion Arts is a company dedicated to the acoustic design and installation of music-related spaces, who also manufacture one of the highest definition range of studio monitor loudspeakers. Apart from that he keeps playing guitar live and in the studio through excellent amplifiers and loudspeakers.

Acknowledgements

Special thanks to Janet Payne, who assembled all the text of this book from the thousands of pages of manuscript.

Many thanks to Christian Haselev, at Middle Tennessee State University, for many helpful comments after reading the first draught of the book.

This page intentionally left blank

Preface

I was building a studio, some years ago, in the Basque region of Spain, close to Bilbao. We were using local labour, but I used a Portuguese foreman who had previously worked with me on other constructions in order to provide some experienced guidance on the specialised day-to-day work. He understood Castilian Spanish quite well, but many of the people working on the construction were speaking Euskera, the Basque national language. After a few days, the foreman said to me, “This language is very similar to Arabic” (In fact, apart from ‘Coca Cola’ and ‘Windows’ there are probably no other words or structures in common), I asked how he came to this conclusion, and he replied, “A few years ago I was working in Morocco, where they speak Arabic, and I couldn’t understand anything. Well, I can’t understand anything when they speak Basque either, so the two languages must be very similar”.

It must have taken a full week for my brain to recover from the intellectual offence which it had suffered, and yet, at times, within the recording industry I am assaulted by opinions and reasoning about loudspeakers which bear little more logic than the aforementioned foreman’s linguistic conclusions. In an attempt to throw some more light on these matters, Dr Keith Holland and I therefore decided that we should write a book on the subject of the fundamental differences between the Basque and Arabic languages. However, once we realised that we knew no more than half a dozen words between us (‘Coca Cola’ and ‘Windows’ included) we resolved to write a book on loudspeakers, instead.

Philip Newell
Moaña
Spain
2006

This page intentionally left blank

Introduction

Every day, around the world, millions of people use loudspeakers as some sort of reference in their place of work. These people include those who work in film studios, radio stations, live sound events, discothèques, clubs, music recording studios, television studios, theatres, cinemas and many other associated professions. For many musicians who play electrified or electronic instruments, their loudspeakers are an integral part of their instruments, and as such can significantly affect the *interpretation* of a performance, as well as its sound. Amongst all of the professionals there are only a very few who actually know much about how loudspeakers work, other than a little knowledge gained in a cursory way that can be as deceiving as it is revealing. They rely on loudspeakers as 'black-boxes' which somehow transform electrical signals into sound. Upon opening a box, they see very little inside, except a metal or plastic chassis, cardboard of plastic diaphragms, a magnet, some fluffy material, and perhaps a few simple electronic components and wires that may make up a crossover circuit.

It is often a source of frustration to these people – and to countless millions more of domestic hi-fi enthusiasts – that these simple little boxes fail to deliver an accurate and repeatable sound in a variety of circumstances. It is a further source of frustration that they all seem to sound different, because it seriously complicates the compatibility problems when their work travels from system to system, or room to room. They feel frustrated because it seems that surely the knowledge exists to sort out such uncomplicated devices. This situation is often exacerbated by the powers of marketing, when so many advertisements from so many manufacturers all claim that *their* loudspeakers have the ability to tell the truth; but obviously the reality cannot be quite so simple. In fact, loudspeakers are electro-mechanico-acoustic systems whose behaviour is complex to a degree that seems totally disproportionate to the simplicity of their appearance.

Drivers of racing cars, helicopter pilots or submarine captains are people who we cannot imagine as not having a through knowledge of the vehicles that they command. It almost goes without saying that when they return to base they would be able to communicate to the technicians, mechanics and engineers in a way that was clear and concise about any technical failings or handling difficulties that had occurred. Conversely, many people who use loudspeakers, professionally, have remarkably little insight into what is going on inside the equipment, and when problems arise with the sound, they are at a total loss to explain either what the symptoms are or what the problem might be. They work by trial and error, and often try to restrict themselves to working in familiar environments where any problems, even if not understood, are at least known.

In all fairness, some professional users of loudspeakers do try to learn something more about the devices which are so important to their daily work, but they are often faced with a choice of two equally blind alleys. The first is to look at some text books on the subject, and the other is to search through popular magazines which publish articles about loudspeakers, such as the hi-fi and home recording publications. The text book approach often grinds to a halt somewhere before the end of page 1 as they become overwhelmed by the complexity of the electro-acoustic theory which, even for specialised loudspeaker engineers, is not always entirely straightforward. [In fact, one of the most authoritative books on loudspeaker theory and application was written by fourteen different people, because such a work would almost certainly be beyond the ability of any, one person[1].] The popular press, on the other hand, is largely concerned with filling pages with text and selling advertising space, which is perfectly understandable because their primary raison d'être is to entertain the readers. However, conjecture and opinions are often passed off as authoritative fact, and contradictions are commonplace. It can thus become very difficult for the non-specialist to separate the fact from the fiction, so the avid reading of such publications is liable to result in an information overload, but with no clear facts being apparent in any unequivocal manner.

The object of *this* book therefore is to try to fill the gap which currently exists between the text books and the popular press. It will try to describe the theory behind, and application of, the loudspeakers which are used for music recording and reproduction in a way that is accessible to those who would benefit from a greater understanding of the concepts, but who do not have anything more than a basic understanding of general science. Nevertheless it is intended that the facts and descriptions will be both accurate and thorough, and where subjective aspects of loudspeaker performance is discussed, it will be backed up with objective and perceptual justification.

It is inevitable that subjective perceptions must be discussed, because ultimately it is the ear of each individual listener which acts as judge and jury, despite what the measurement may say, but recent research has shown that subjective assessments *can* be reliable and quite precise as long as the variable peripheral factors are minimised and understood. The authors are very aware of the pitfalls, and have had a great deal of experience in dealing with them. One of the authors is a designer of recording studios, film dubbing theatres, television studios and concert halls, who for many years was a recording engineer, live sound engineer, record producer and monitor system designer. The other is a Doctor of Acoustics, and a university lecturer in electro-acoustics who has had much experience in live sound as a front of house engineer. Both are members of the UK's Institute of Acoustics and the Audio Engineering Society, and both are experienced at teaching audio technology from very basic levels. Hopefully, therefore, ways will be found to explain complex ideas via understandable but nevertheless accurate analogies, which should enable the readers to grasp an intuitive feel for the subject, especially where mathematical explanations would elude them.

Whilst it has to be accepted that the majority of loudspeaker users neither have the time nor the inclination to formally study electro-acoustics, it is still extremely useful for them to understand much more than most of them currently do, in order to help them make more informed decisions about matters which affect their working lives. People beginning to work in the recording world or people with a keen interest in sound reproduction will also probably find this book useful. Inevitably, from time to time, things may need to get a little deep. This will be essential when the discussion requires it, but hopefully it can be done in a series of steps which will not leave the less technical readers too isolated.

The book will begin with a brief history of loudspeaker development in the early days, and will look at the basic concepts of just what a loudspeaker is, and what it must do. Some basic principles of sound radiation will then be introduced, in order to give a better understanding of the principles before looking at the wide range of motor systems technologies that are available, such as moving coil, electrostatic, piezoelectric, ionic, magnaplanar and various other concepts. The pros and cons of different diaphragm technologies will be discussed, as will those of the loading techniques, such as with horns and various cabinets and baffles.

Loudspeakers, of course, require an electrical drive signal, so Chapter 5 will look at the whole concept of crossovers, discussing why they are needed and how they can be realised in practice. Active and passive designs will be investigated, as will various slopes, shapes, phase effects and reconstruction difficulties. The electrical, acoustical and mechanical (physical) factors which affect crossover performance will be dealt with in a thorough, yet understandable way, before the following chapter discusses the amplifiers and cables which are necessary to complete any monitoring system. Different amplifier topologies will be discussed, and their suitability for different specialised uses will be indicated. Without getting into the subjective minefield of loudspeaker cable audibility, an objective presentation will be made of the ways in which it has been shown that cables can, and do, affect system performance. There is no point in paying ten times more than necessary for a special loudspeaker cable if a standard cable will sound exactly the same, but in sensitive circumstances a more esoteric cable may be beneficial. It is the first time that most of this work has been published outside of the proceedings of international electro-acoustic conferences.

The book will then go on to discuss the consequences of the interaction between the loudspeaker systems and the rooms in which they are sited, and how, in so many cases, the rooms can dominate the overall response. Guidance will be given as to where to find out more about the room acoustics, but the consequences of different mounting regimes will be dealt with, here.

Chapter 8 is a wide-ranging analysis of the reasons why different loudspeakers appear to be more appropriate during different phases of the recording/mixing/mastering/listening process. Many concepts of loudspeaker system design will be analysed, and their application to the environments in which they are most likely to be used will be assessed. Motor systems and cabinet options will be brought together in ways that they can be applied to the physical, electro-acoustic and psychoacoustic requirements of each stage of the work. With the introduction of the question of perception, highlighted by the fact that different loudspeakers tend to be chosen for different phases of the work, Chapter 9 analyses the different measurements which can be used to define the performance of a loudspeaker system, and how the objective measurements relate to different aspects of the subjective perception of the music. Chapter 10 will then reverse the concept, and will discuss how the musical arrangements can be the culprits for many system-to-system compatibility problems for which the loudspeakers usually take the blame. This is a crucial subject, but one which one very rarely sees discussed in print.

The final two chapters will deal with subjects which are, in general, very poorly understood by the vast majority of people who work with loudspeakers. Chapter 11 discusses the fundamental requirements of the low frequency radiation from small loudspeakers. It explains why, in accuracy terms, loudspeaker designers compromise the performance in order to reduce box sizes and/or extend the responses. Brand new measurement concepts will be introduced which demonstrate the degree to which these performance enhancements reduce the reproduction accuracy in an exchange of quantity for quality. The chapter then goes on to look at the transient performances, which are so important for the realistic perception of music, but which are consistently ignored by many loudspeaker manufactures, partly because they do not feature in most performance specifications which are used for publicity purposes. They are often compromised for the betterment of some less significant responses benefits which do carry more weight in terms of advertising.

The book concludes with a chapter on surround-sound application; not only in the more physical realms of room interfacing that were discussed in Chapter 7, but also in terms of the application problem that confound the day-to-day use of surround sound, such as format compatibility, the suitability of certain loudspeaker radiation patterns to specific mounting conditions, and the appropriateness of loudspeaker choices for different musical programme.

Hopefully, therefore, this book will fill the gap left between the textbooks and the magazines. It deals with the application and use of the technology and science, justifying ideas with hard facts rather than conjecture, and in a way that should be accessible to anybody with a general level of experience in the use of loudspeakers, whether for work or leisure purposes.

Reference

1 Borwick, J., ‘Loudspeaker and Headphone Handbook’, Third Edition, Focal Press, Oxford, UK (2001)

Chapter 1

What is a loudspeaker?

1.1 A brief look at the concept

Before answering the question posed by the title of this chapter, perhaps we had better begin with the question "What is sound?" According to Fahy [1] "sound may be defined as a time-varying disturbance of the density of a fluid from its equilibrium value, which is accompanied by a proportional local pressure, and is associated with small oscillatory movements of the fluid particles". The difference between the equilibrium (static) pressure and the local, oscillating pressure is known as the *sound pressure*.

Normally, for human beings, the fluid in which sound propagates is air, which is heavier than most people think – it has a mass of about 1.2 kg per cubic metre at a temperature of 20 degrees C at sea level. It is also interesting to note that sound propagation in air is by no means typical of its propagation in all substances, especially in that the speed that sound propagates in air is relatively slow, and is constant for all frequencies. For music lovers, this latter fact is quite fortunate, because it would be hard to enjoy a musical performance at the back of a concert hall if the notes arrived jumbled-up, with the harmonics arriving before the fundamentals, or vice versa. Conversely, as we shall see later, most of the materials from which loudspeakers are made do *not* pass all frequencies at the same speed of sound, a fact which can, at times, make design work rather complicated.

The speed of sound in air is about 343 metres per second (m/s or ms^{-1}) at 20 degrees C and varies proportionally with temperature at the rate of about 0.6 m/s for every degree Kelvin. (In fact, the speed of sound in air is *only* dependent upon temperature, because the changes that would occur due to changes in atmospheric pressure are equal and opposite to the accompanying changes in density, and the two serve to cancel each other out). Air therefore has some clearly defined characteristic properties, and our perception of sound in general, and music in particular, has developed around these characteristic properties.

The job of a loudspeaker is to set up vibrations in the air which are acoustic representations of the waveforms of the electrical signals that are being supplied to the input terminals. A loudspeaker is therefore an electro-mechanico-acoustic transducer. Loudspeakers transform the electrical drive signals into mechanical movements which, normally via a vibrating diaphragm, couple those vibrations to the air and thus propagate acoustic waves. Once these acoustic waves are perceived by the ear, we experience a sensation of sound.

To a casual observer, a typical moving coil loudspeaker (or 'driver' if you wish to restrict the use of 'loudspeaker' to an entire system) seems to be a simple enough device. There is a wire 'voice-coil' in a magnetic field. The coil is wound on a cylindrical former which is connected to a cardboard cone, and the whole thing is held together by a metal or plastic chassis. The varying electrical input gives rise to vibrations in the cone as the electromagnetic field in the voice coil interacts with the static field of the (usually) permanent magnet. The cone thus responds to the electrical input, and there you have it, sound! It is all as simple as that! . . . . . . Or is it?

Well, if the aim is to make a sound from a small, portable radio that fits into your pocket, then maybe that concept will just about suffice, but if full frequency range, high fidelity sound is the object of the exercise, then things become fiendishly complicated at an alarming speed. In reality, in order to be able to reproduce the subtle structures of fine musical instruments, loudspeakers have a very difficult task to perform.

1.2 A little history and some background

When Rice and Kellogg [2] developed the moving coil cone loudspeaker in the early 1920s, [and no; they did not also invent Kellogg's Rice Krispies!] they were already well aware of the complexity of radiating an even frequency balance of sound from such a device. Although Sir Oliver Lodge had patented the concept in 1898 (following on from earlier work by Ernst Werner Siemens in the 1870s at the Siemens company in Germany), it was not until Rice and Kellogg that practical devices began to evolve. Sir Oliver had had no means of electrical amplification – the thermionic valve (or vacuum tube) had still not been invented, and the transistor was not to follow for 50 years. Remarkably, the concept of loudspeakers was worked out from fundamental principles; it was not a case of men playing with bits of wire and cardboard and developing things by trial and error. Indeed, what Rice and Kellogg developed is still the essence of the modern moving coil loudspeaker. Although they lacked the benefit of modern materials and technology, they had the basic principles very well within their understanding, but their goals at the time were not involved with achieving a flat frequency response from below 20 Hz to above 20 kHz at sound pressure levels in excess of 110 dB SPL. Such responses were not required because they did not even have signal *sources* of such wide bandwidth or dynamic range. It was not until the 1940s that microphones could capture the full frequency range, and the 1950s before it could be delivered commercially to the public via the microgroove, vinyl record.

Prior to 1925, the maximum output available from a radio set was in the order of milliwatts, normally only used for listening via earphones, so the earliest 'speakers' only needed to handle a limited frequency range at low power levels. The six inch, rubber surround device of Rice and Kellogg used a powerful electro-magnet (not a permanent magnet), and as it could 'speak' to a whole room-full of people, as opposed to just one person at a time via an earpiece, it became known as a *loud* speaker. The inventers were employed by the General Electric Company, in the USA, and they began by building a mains-driven power amplifier which could supply the then huge power of *one watt*. This massive increase in the available drive power meant that they no longer needed to rely on resonances and rudimentary horn loading, which typically gave very coloured responses. With a whole watt of amplified power, the stage was set to go for a flatter, cleaner response. The result became the Radiola Model 104, which with its built in power amplifier sold for the then enormous price of 250 US dollars. [So there is nothing new about the concept of self-powered loudspeakers: they *began* that way!] Marconi later patented the idea of passing the DC supply current through the energising coil of the loudspeaker, to use it instead of the usual, separate smoothing choke to filter out the mains hum from the amplifier. Therefore right from the early days it made sense to put the amplifier and loudspeaker in the same box.

Concurrently with the work going on at General Electric, Paul Voight was busy developing somewhat similar systems at the Edison Bell company. By 1924 he had developed a huge electro-magnet assembly weighing over 35 kg and using 250 watts of energising power. By 1926 he had coupled this to his Tractrix horn, which rejuvenated interest in horn loudspeakers because it enormously improved the sensitivity and acoustic output of the moving coil loudspeakers, and when properly designed did not produce the 'honk' sound associated with the older horns. Voight then moved on to use permanent magnets, with up to 3.5 kg of Ticonal and 9 kg of soft iron, paving the way for the permanent magnet devices and the much higher acoustic outputs that we have today.

Gilbert Briggs, the founder of Wharfedale loudspeakers, wrote in his book of 1955[3] "It is fairly easy to make a moving-coil loudspeaker to cover 80 to 8,000 cycles [Hz] without serious loss, but to extend the range to 30 cycles in the bass and 15,000 cycles in the extreme top presents quite a few problems. Inefficiency in the bass is due mainly to low radiation resistance, whilst the mass of the vibrating system reduces efficiency in the extreme top". The problem in the bass was, and still is, that with the cone moving so relatively slowly, the air in contact with it simply keeps moving out of the way, and then returning when the cone direction reverses, so only relatively weak, low efficiency pressure waves are being propagated. The only way to efficiently couple the air to a cone at low frequencies is to either make the cone very big, so that the air cannot get out of the way so easily, or to constrain the air in a gradually flaring horn, mounted directly in front of the diaphragm. Unfortunately, both of these methods can have highly detrimental effects on the high frequency response of the loudspeakers. For a loudspeaker cone to vibrate at 20 kHz it must change direction forty thousand times a second. If the cone has the mass of a big diaphragm needed for the low frequencies, its momentum would be too great to respond to so many rapid accelerations and decelerations without enormous electrical input power – hence the loss of efficiency alluded to by Briggs. Large surfaces are also problematical in terms of the directivity of the high frequency response, but we will come to that later. So, we can now begin to see how life becomes more complicated once we begin to extend the frequency range from 20 Hz to 20 kHz – the requirements for effective radiation become conflicting at the opposing frequency extremes.

The wavelength of a 20 Hz tone in air is about 17 metres, whereas the wavelength at 20 *k*Hz is only 1.7 *centi*metres, a ratio of 1000 to 1, and for high quality audio applications we want our loudspeakers to produce all the frequencies in-between at a uniform level. We also need them to radiate the same waveforms, differing only in size (but not shape) over a power range of at least 10,000,000,000 to 1 *and* with no more than one part in a hundred of spurious signals (non-linear distortion). It is a tall order! Indeed, for a single drive unit, it still cannot be achieved at any realistic SPL (sound pressure level) if the full frequency range is required.

1.3 Some other problems

There are also many mechanical concepts which must be considered in loudspeaker design. For example, the more that one pushes on a spring, the more one needs to push in order to make the same change in length. If the force is limited to less than what is necessary to fully compress the spring, equilibrium will be reached where the applied force and the reaction of the spring balance each other. This is useful when we go to bed, because it prevents the suitably chosen springs from bottoming out, and allows the mattress to adapt to our shape yet retain its springiness. The suspension systems of loudspeaker diaphragms are also springs, but they must try to maintain a consistent opposition to movement or they would compress the acoustical output. If, for example, the first volt of input moved the diaphragm x millimetres, and a second volt moved it only 0.8 x mm further, this would be no good for high fidelity, because when we double the voltage we expect to see the same linear increase in motion. Otherwise, the acoustic output would not be linearly following the electrical input signal, and the non-linear movement would introduce distortion. Therefore, to keep the diaphragm well centred, but to still allow it to move linearly with the input signal, suspension systems must be used which do not exhibit a nonlinear restorative force as the drive signal increases, at least not until the rated excursion limit is reached. (However, as we shall see in Chapter 8, certain non-linear loudspeaker characteristics may actually be *desirable* for musical instrument amplification.) We tend complicate this suspension problem further when we put a loudspeaker drive unit into a box, because the air in the box acts as an additional spring which is also not entirely linear.

Electrically we can also run into similar problems. Whenever an electric current flows through a wire, the wire will heat up. It is also a property of voice coil wires that as they heat up their resistance increases. As the resistance increases, the signal voltage supplied by the amplifier will proportionally drive less current through the coil. So, as the drive force depends on the current flowing through the wire which is immersed in the magnetic field, if that current reduces, the movement of the diaphragm will reduce correspondingly. We therefore can encounter a situation where the amplifier sends out an accurate drive signal *voltage*, but how the loudspeaker diaphragm responds to it can, depending on level, change with the voice coil temperature and the springiness of both the air and the suspension system. Even the very magnetic field of the permanent magnet, against which the drive force is developed, can be modulated by the magnetic field given rise to by the signal in the voice coil. Notwithstanding, all of these effects, we still need our diaphragm to move exactly as instructed by the drive voltage from the amplifier, because in reality modern amplifiers are usually *voltage* sources. This is despite the fact that the loudspeaker motor is *current* driven, and that the voice coil resistance and reactance (together they form the impedance) will not remain constant over the whole of the frequency range.

Clearly, things are beginning to get complicated, and already we have seen problems begin to pile up on each other. The concept of a loudspeaker being simply an electromagnet, coupled to a moveable cone and placed into a box is obviously not going to produce the high fidelity sounds needed for music recording and reproduction. Good loudspeakers are complex devices which depend on the thorough application of some very rigid principles of electroacoustics in order to perform their very complex tasks. In the minds of most people the concept of a loudspeaker, if it exists at all, is usually grossly over-simplified, and unfortunately this is the case even with most *professional* users of loudspeakers.

Typical loudspeaker systems consisting of one or more vibrating diaphragms, either on one side of a rectangular cabinet or flush mounted into a wall, represent physical systems of sufficient complexity that accurate and reliable predictions of their sound radiation are rare, if not non-existent, even with the aid of modern computer technology. Despite the fact that we inevitably have to deal with a degree of artistry and subjectivism in the final assessment, at the design and development stages we must stick close to the objective facts. So, to begin to understand the mechanisms of sound radiation it is necessary to establish the means by which a sound 'signal' is transported from a source, through the air, and to our ears.

1.4 Some basic facts

As explained in the opening paragraphs of this chapter, acoustic waves are essentially small local changes in the physical properties of the air which propagate through it at a finite speed. The mechanisms involved in the propagation of acoustic waves can be described in a number of different ways, depending upon the particular cause or source of the sound. With conventional loudspeakers that source is the movement of a diaphragm, so it is appropriate here to begin with a description of sound propagation away from a simple moving diaphragm.

1.4.1 Acoustic wave propagation

The process of sound propagation is illustrated in Figure 1.1. For simplicity, the figure depicts a diaphragm mounted in the end of a uniform pipe, the walls of which constrain the acoustic waves to propagate in one dimension only. Before the diaphragm moves (Figure 1.1(a)), the pressure in the pipe is the same everywhere and equal to the static (atmospheric) pressure P_0. As the diaphragm moves forwards (Figure 1.1(b)), it causes the air in

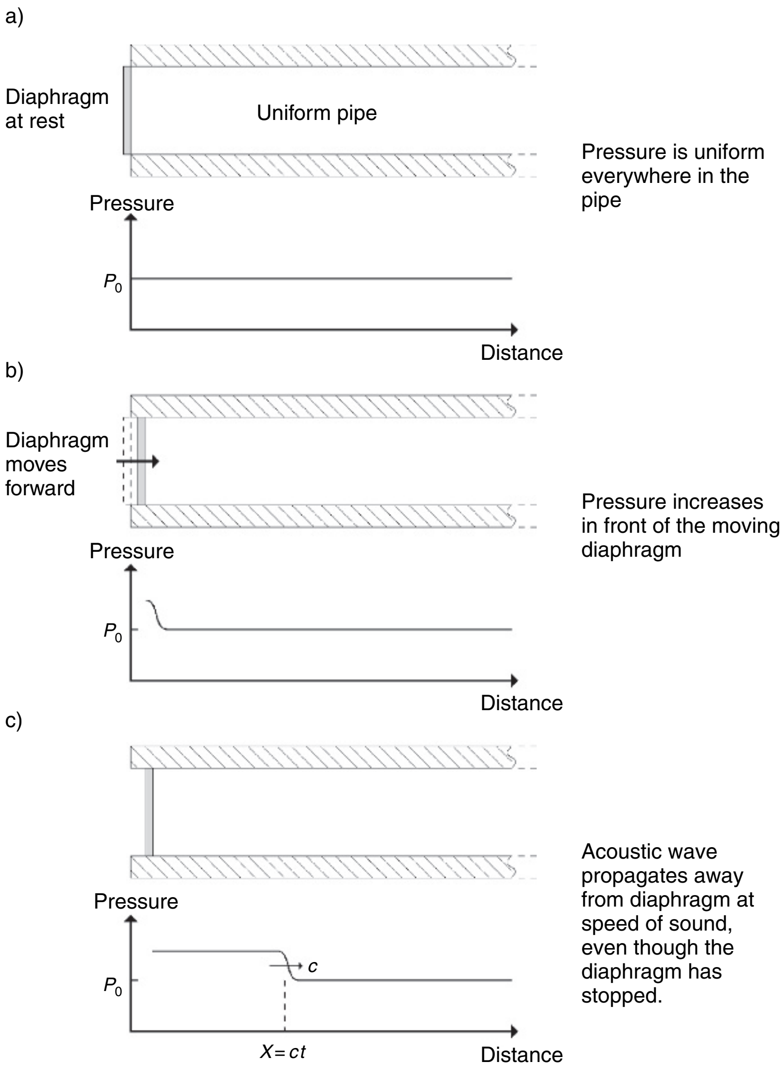


Figure 1.1 The generation and propagation of an acoustic wave in a uniform pipe

contact with it to move, compressing the air adjacent to it and bringing about an increase in the local air pressure and density. The difference between the pressure in the disturbed air and that of the still air in the rest of the pipe gives rise to a force which causes the air to move from the region of high pressure towards the region of low pressure. This process then continues forwards, and the disturbance is seen to propagate away from the source in the form of an acoustic wave. Because air has mass, and hence inertia, it takes a finite time for the disturbance to propagate through the air; a disturbance 'leaves' a source and 'arrives' at another point in space some time later (Figure 1.1(c)). The rate at which disturbances propagate through the air is known as the *speed of sound,* which has the symbol '*c*', and after a time of t seconds, the wave has propagated a distance of $x = ct$ metres. Note, though, that the speed of propagation is *not* related to the velocity of the diaphragm. In the great tsunami of December 2004, a five metre displacement of the ocean floor caused a tidal wave to travel at over 500 kilometres per hour, a speed much more rapid than that of the displacement which caused it. For most purposes, the speed of sound in air can be considered to be constant, and independent of the particular nature of the disturbance, although as previously stated it does vary with temperature. A one-dimensional wave, such as that shown in Figure 1.1, is known as a *plane wave*. A wave propagating in one direction only (e.g. left-to-right) is known as a *progressive wave.*

1.4.2 Mechanical and acoustic impedance

The description of sound propagation in Section 1.4.1 mentioned the motion of the air in response to local pressure differences. This localised motion is often described in terms of acoustic particle velocity, where the term ‘particle’ here refers to a small quantity of air that is assumed to move as a whole. Although we tend to think of a sound field as a distribution of pressure fluctuations, any sound field may be equally well described in terms of a distribution of particle velocity, and there is a relationship between the distribution of pressure in a sound field and the distribution of particle velocity. At a given frequency, the ratio of pressure to particle velocity at any point (and direction) in a sound field is known as acoustic impedance (strictly *specific* acoustic impedance), $Z_a = p/u$, and it is very important when considering the sound power radiated by a source. Acoustic impedance can be thought of as a quantity that expresses how difficult the air is to move. A low value of impedance tells us that the air moves easily in response to an applied pressure (low pressure, high velocity), and a high value of impedance tells us that it is hard to move (high pressure, low velocity). Mechanical impedance is directly equivalent to acoustic impedance, but with pressure replaced by force (pressure is force per unit area) and particle velocity replaced by velocity: $Z_m = F/u$.

With acoustic radiators, as well as electrical circuits and mechanical systems, there is a need to match impedances for good energy transfer. If a microphone needs to be connected to a 600 ohm input, then it will not sound as intended by its designers or exhibit its quoted sensitivity if connected to a 30 ohm or 10,000 ohm input. An amplifier which is optimised to function into a 4 ohm load will not produce its maximum power output capability into a load of 16 ohms. Loudspeakers are effectively ‘plugged into’ the air, so if the air load impedance does not match the electromechanical output impedance of the loudspeaker, the radiated power will be less than optimal. Impedance changes with frequency, so, for example, a resistor and capacitor in parallel have a frequency dependent impedance which is the combination of the purely resistive, frequency *independent* characteristic of the resistor, and the reactive, frequency *dependent* characteristic of the capacitor. This is the basis of electrical filter design.

Impedance (Z), whether it be electrical, mechanical or acoustic, can be divided into two components, resistance (R) and reactance (X). Reactive impedances represent systems which store input energy, but which later

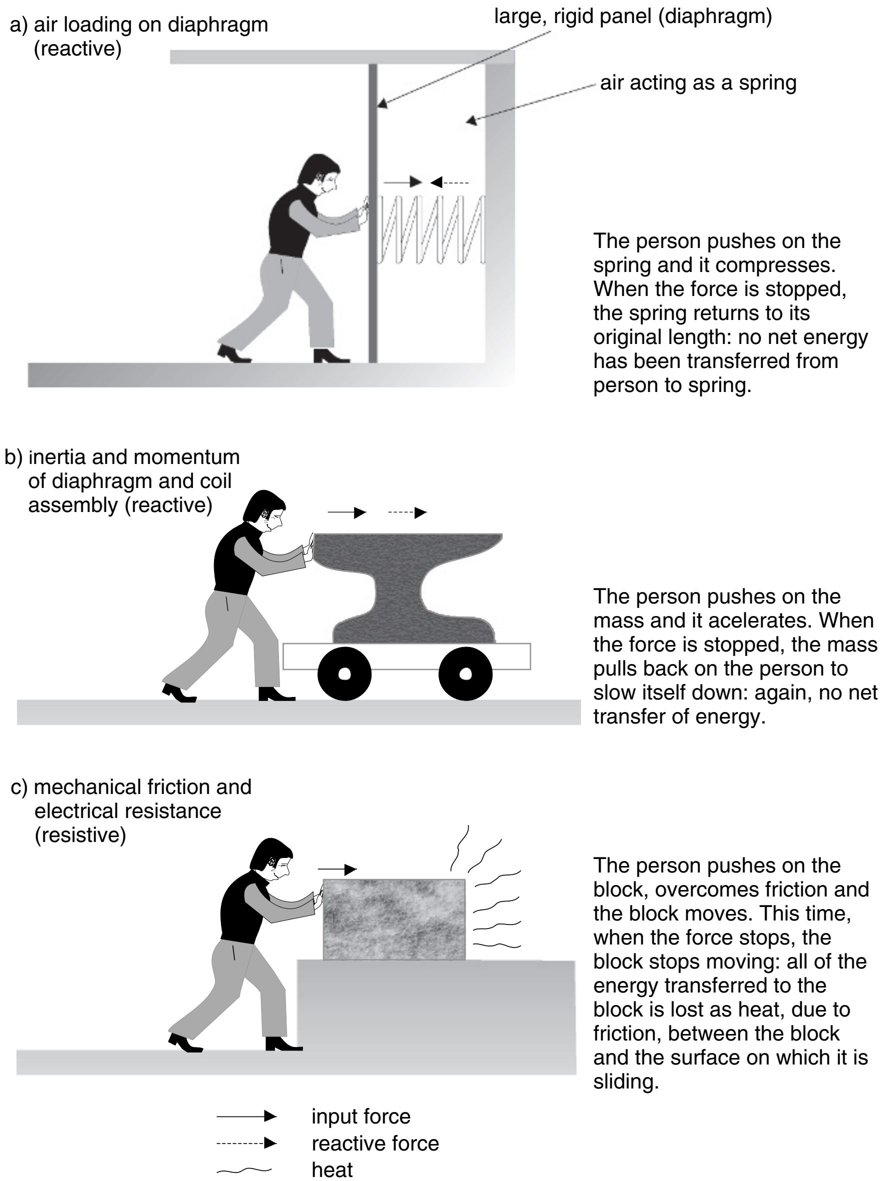


Figure 1.2 Three characteristic properties of a moving coil loudspeaker depicted as three components of its impedance. In electrical terms, these can be related to capacitance, inductance and resistance

give it back, whereas resistive impedances represent systems which transfer energy away from the input, never to return. Figure 1.2 shows three mechanical systems which can be used to demonstrate the three different components of mechanical impedance and the way in which they relate to conventional loudspeakers. In Figure 1.2(a), the person applies a force to compress a spring. When the applied force is stopped, the spring returns to its original length and all of the (potential) energy applied to the spring is returned to the person as the spring pushes back. If the person pushes back and forth on the spring in an oscillatory manner, the energy flows from person to spring and back again in each half-cycle with an overall zero transfer of energy. A spring represents a purely reactive mechanical impedance. Figure 1.2(b) shows the person applying a force to a mass on a trolley. The force acts to accelerate the mass from rest, but when the force is stopped the mass tries to continue moving at a steady velocity. If the person then applies a force to slow the mass down, the (kinetic) energy possessed by the mass is returned as the mass pulls back on the person. Again, if the force is applied in an oscillatory manner, there is zero overall transfer of energy from the person to the mass. A mass also represents a purely reactive mechanical impedance, but with the opposite sign to that of a spring. (Note the different direction of the reactive force arrows in the figure.) Figure 1.2 c) shows the person pushing a block along a table. The applied force overcomes the friction between the block and the table and the block moves with a constant velocity. When the force is stopped the block also stops, and none of the energy supplied to the block is returned to the person. If the force is applied in an oscillatory manner in this case, the flow of energy is always from person to block, regardless of the direction of motion. All of the energy is 'lost' to friction as heat, and none is returned to the person. The friction block represents a resistive mechanical impedance which is the mechanism by which power can be transferred from one system to another (in this case, from the person to heat). There are acoustic counterparts to each of these mechanical components, for example, a small sealed cavity driven by a piston is an acoustic spring. The electrical counterparts will be further discussed in Section 1.6.

1.4.3 Impedance in loudspeakers

In a cone loudspeaker, we have all of these forms of impedance present at the same time. The diaphragm radiates useful sound power through its motion via an acoustic radiation *resistance*. The mass of the voice coil and diaphragm, together with the stiffness of the suspension, produce a *reactive* impedance which merely serves to reduce the diaphragm motion. The *reactive* inductance of the voice coil reduces the current, and hence the applied force, at higher frequencies. And finally, the *resistive* frictional losses in the suspension and the electrical *resistance* of the voice coil simply waste power by turning it into heat.

1.5 The practical moving-coil cone loudspeaker

The majority of all loudspeakers are moving coil devices employing a cone-shaped radiating diaphragm, but the mechanisms of sound radiation from these devices are not as straightforward as they may initially seem to the casual onlooker. This type of loudspeaker is essentially a 'volume velocity' source. In other words it creates a pressure wave equivalent to injecting air from a point source at a rate of injection measured in cubic metres per second (or *extracting* the air on the rarefaction half cycle). However, unlike the piston shown in Figure 1.1, most cabinet-mounted cone loudspeakers, direct radiating into a room (i.e. not radiating via a horn), do not couple effectively with the air. Instead, the cone finds itself punching into thin air, with much of the potential load being lost as most of the air adjacent to it simply moves out of the way, then returns when the direction of the cone movement reverses. Efficiency is therefore often extremely low, with less than 1% of the energy being supplied to the voice coil resulting in the radiation of sound. The remaining energy either gets lost by friction in the moving system, or by being burned up as heat by the resistance of the voice coil, or even by being reflected back into the output stages of the power amplifier. To complicate matters, many of these things are frequency dependent, so, it is little wonder that many things must be considered and balanced before there is any chance of such a device having a flat frequency response.

As long as the circumference of the diaphragm remains small with respect to the wavelength, the radiation will be omnidirectional, but when the wavelength starts to become small compared to the circumference of the source, the radiation begins to beam directly ahead. [The wavelength, in metres, can be calculated simply by dividing the speed of sound in metres per second by the frequency in hertz.] This is a result of the interference field where the different parts of the diaphragm, radiating in phase, become significantly different in their path lengths to an off axis listening position. The cause is depicted in Figure 1.3, and the effect is shown in Figure 1.4. The *ka* values referred to in the latter figure are derived from the wave number k, which is simply 2π divided by the wavelength (λ) in metres, and a, which is the radius of the radiating surface. In practice we can think of ka as being the number of wavelengths around the circumference of the diaphragm: $ka = 2\pi a/\lambda$.

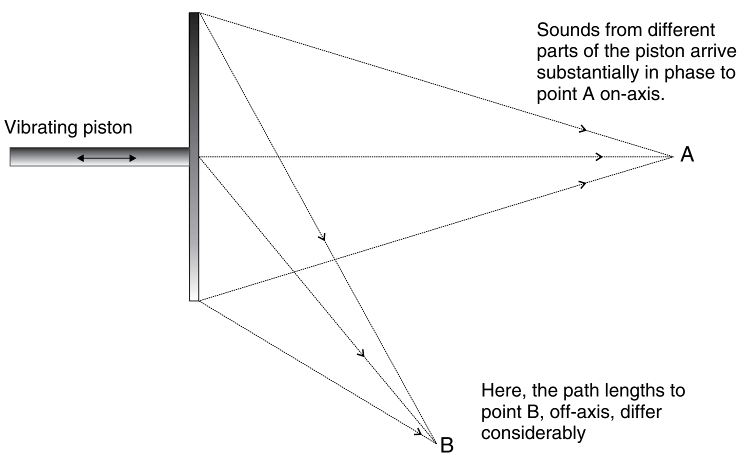


Figure 1.3 The cause of the off-axis interference effects that give rise to the directivity shown in Figure 1.4

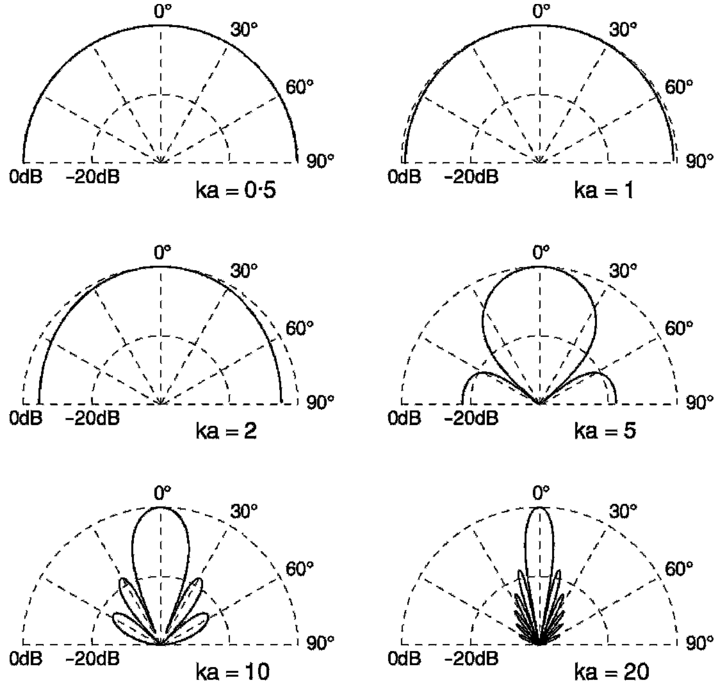


Figure 1.4 The directivity of a vibrating piston due to off-axis interference effects at different frequencies. Note the narrowing of the main lobe as the frequency increases

We would tend instinctively to think of the whole diaphragm as being the radiating surface, however, this is often not the case, and a large diaphragm driven at high frequencies will almost certainly *not* move uniformly. The outer parts would lag with respect to the movement of the central area, in which case the outer parts would radiate with a phase shift relative to the central area of the diaphragm. In fact they may even simply stay still, in which case they would not radiate at all. In either case the radiation may not correspond with the *ka* value taken from the *physical* measurements of the diaphragm.

Over the years, many 'solutions' have been tried in the search for perfect pistonic motion in a diaphragm, such as using ultra-rigid cone materials, or solid, conical 'plugs' as diaphragms, but internal losses and differential sound speeds *within* the materials have tended to confound all efforts. Remember, it was stated at the beginning of this chapter that air was not typical of all materials in terms of the speed with which sound propagates though it being equal for all frequencies. Many types of solid materials propagate high frequencies faster than low frequencies, a property known as phase dispersions, and exhibit sound speeds very different from that in air.

When the sound propagates from the voice coil *through* the material of the cone, in a radially outward direction, the different sound speeds may give rise to interference between the sound waves travelling to the edge of the cone and the forward radiation into the air by the pistonic action of the whole cone being driven by the voice coil. Resonances can also be set up within the cone itself, and one job of the cone surround is to absorb as much as possible of the waves propagating *radially* through the cone material, to prevent them from being reflected back and causing even more interference. Complex cone surrounds are often as much to do with suppressing these waves as with linearly suspending the cone itself.

1.5.1 The combined response

It was stated earlier that a loudspeaker was an electro-mechanico-acoustic transducer. Well, we have just looked at a few of the acoustic properties, but the electrical and mechanical aspects also present their own complications. The voice coil is a mixture of resistance and inductance, the latter being a form of frequency dependent reactance. The reactance, in ohms, of an inductor is given by $2\pi f L$, where f is the frequency in hertz and L is the inductance in henries. An 8 ohm voice coil is only nominally 8 ohms, and approximates that value only over a limited band of frequencies. The inductive part of the impedance (impedance here being resistance plus reactance) rises with frequency, whereas any capacitive reactance effects *decrease* with frequency. The reactance(c) of a capacitor, in ohms, is given by $1/\ 2\pi f c$, so the frequency component, being in the denominator in this case, reduces the reactance as it rises. An overall impedance plot of a typical low frequency loudspeaker with respect to frequency is shown in Figure 1.5.

Fortunately, however, some effects counterbalance each other. Figure 1.6 shows how a flat frequency response can be achieved above the resonance frequency of a cone driver, mounted on a theoretical infinite baffle, where the roll-off in the diaphragm velocity compensates for its rising relationship with the pressure radiated to a point far away on the cone axis. However, at very close distances some rather more complex relationships manifest themselves. Academically speaking, this region is known as the near-field,

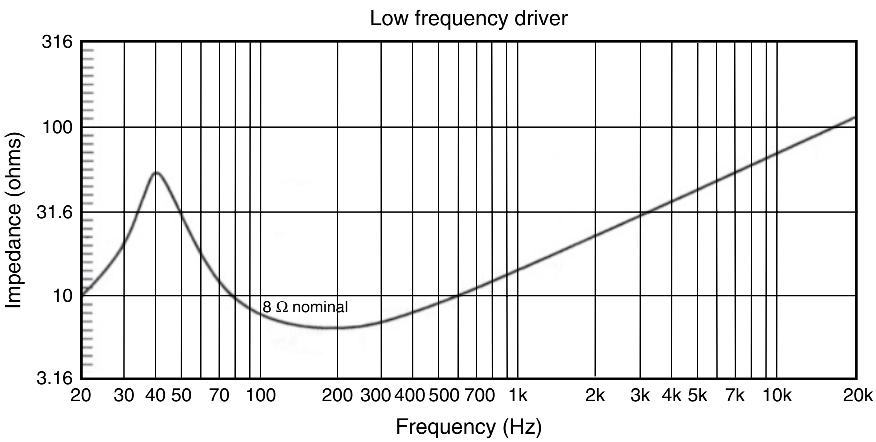


Figure 1.5 Magnitude of electrical impedance of a typical loudspeaker drive unit

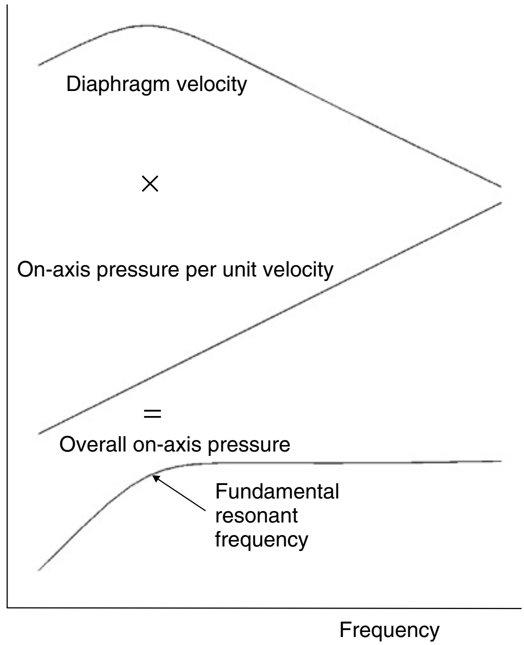


Figure 1.6 How a flat frequency response results from a falling velocity and a rising radiated sound pressure

but it should not be confused with the more colloquial ‘near-field’ that many recording personnel speak of. In fact, it is better to use the term ‘close-field’ for desk-top monitors, because the true near-fields are not good distances to listen within as the frequency balance can be strange in those regions.

For people conversant with electrical circuit diagrams, equivalent circuits can be devised which represent the electrical, mechanical and acoustic properties of a loudspeaker as a single electric circuit. This type of representation is very common in traditional text books on electroacoustics. In these circuits, the mechanical and acoustic components such as springs and masses are replaced by their electrical equivalents. One way in which this can be realised is to replace mechanical forces (and acoustic pressures) with electrical currents, and mechanical velocities with electrical voltages. It follows then that mechanical springs are replaced by electrical inductors, masses are replaced by capacitors, and mechanical resistances are replaced by resistors. This arrangement is known as the *mobility analogy,* although it should be noted that a different, but equally valid, *impedance analogy* exists, where voltages replace forces and currents replace velocities. Each analogy has its strengths and weaknesses and may better suit different aspects of loudspeaker analysis. The equivalent mobility analogy circuit of a typical moving coil loudspeaker is shown in Figure 1.7. The transformers represent the coupling between the electrical and mechanical, and the mechanical and acoustical domains. This type of circuit is what an amplifier

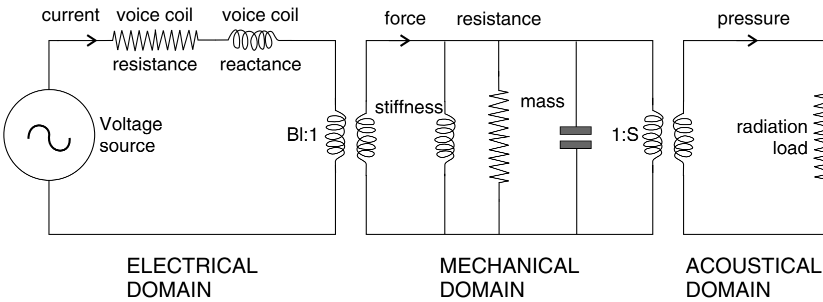


Figure 1.7 The equivalent electrical circuit of a moving coil loudspeaker drive unit (mobility analogy): Bl = the product of the magnetic flux density and the length of wire in the magnetic field – the force factor – T_m (tesla × metres): S = area of the cone

really 'sees' when terminated by a typical moving coil loudspeaker, which is a far cry from a resistive 8 ohm load.

Getting a frequency-independent acoustic output from such a combination of components is no simple task. It can only be achieved by very careful balancing of the values of the components, and even then the perfect balance is usually only achievable over a limited bandwidth. It is apparent that any extra series resistance *increase*, or any parallel resistance *decrease*, will serve to reduce the power supplied to the load for any given input power, and so will reduce the efficiency of the system. What is more, with such a finely balanced system, any change to any component part(s) may require a counterbalancing change to many other component parts. As true perfection can never be achieved, there are an infinite number of close approximations which are possible, and this fact contributes to the diversity of loudspeaker designs that we have available in the marketplace.

1.6 Resistive and reactive loads

Before ending this introductory chapter, it may be worth looking a little more closely at the concepts of resistive and reactive loading. Figure 1.8 shows three potential-divider networks. If we consider a constant voltage source to be applied between terminals A and C we can consider what voltages occur between points B and C, and how much power will be dissipated in each component. In the case of the resistor-resistor network shown in Figure 1.8 (a), the resistance of the two components R_1 and R_2 is equal. In a resistor, as the voltage is increased across its terminals the current will also increase in direct proportion, as given by Ohm's law, which may be written in the following ways:

$$I = \frac{V}{R} \tag{1.1}$$

$$V = IR \tag{1.2}$$

$$R = \frac{V}{I} \tag{1.3}$$

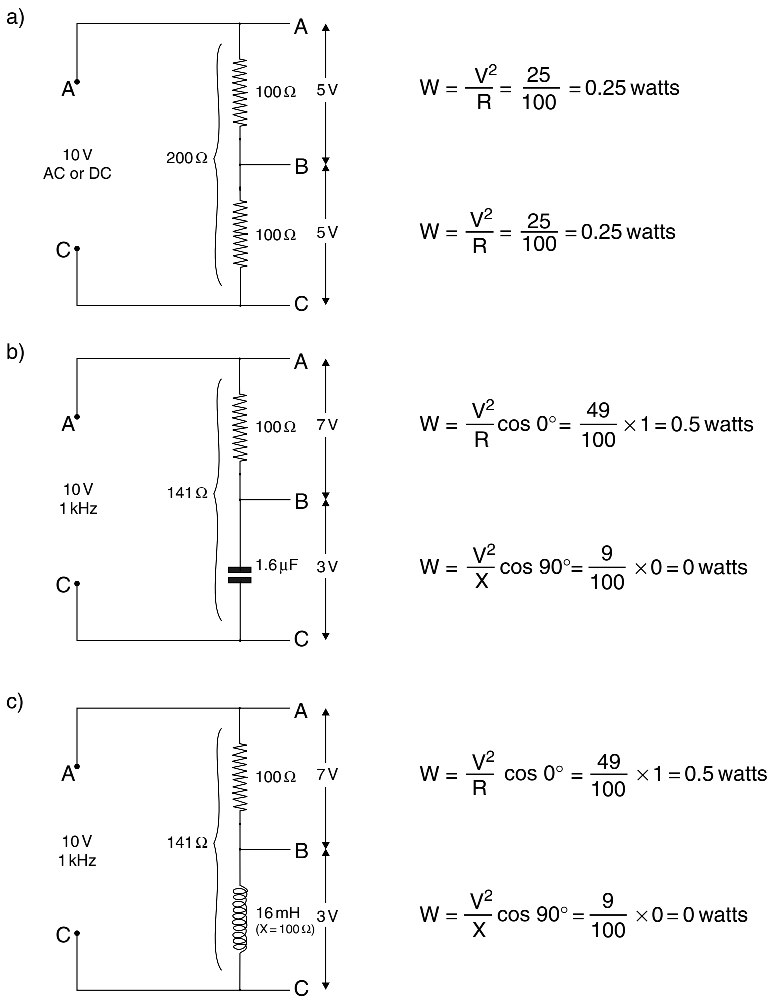


Figure 1.8 Three potential divider networks. W = watts, V = volts, R = resistance (ohms), X = reactance (ohms), Cos 0 degrees = 1, Cos 90 degrees = 0. In each case the total power dissipation is the same, at 0.5 watts, but it is *only* dissipated in the resistors. The inductors and capacitors merely store the energy then release it into the resistors. Note that a *total* of 0.5 watts is dissipated in each case

Where I = current in amps
V = voltage in volts
R = resistance in ohms

The voltage and current are always in phase, so they will produce power, in watts (W), according to a simple multiplication:

$$V \times I = W \tag{1.4}$$

Strictly speaking, this should be multiplied by the cosine of the phase angle, but as this is 0 degrees in resistive circuits, and the cosine of 0 degrees is 1, then a multiplication by one has no effect, so it is traditionally omitted. Whether the current is direct or alternating is also of no consequence in resistors, and the resistance remains the same irrespective of frequency.

In the case of capacitors, there *is* a frequency dependent effect. Capacitors will not pass DC, because the impedance (reactance in this case) at 0 Hz is infinite, but as the frequency rises, the reactance lowers. Reactance (X), like impedance (Z) and resistance (R), is measured in ohms, but unlike resistance it is frequency dependent. In the case of capacitors, the reactance is *inversely* proportional to the frequency: as the frequency rises the reactance lowers. The formula for the reactance (X) of a capacitor is given by:

$$X = \frac{1}{2\pi fc} \tag{1.5}$$

where f = frequency in hertz

c = capacitance in farads

$\pi = 3.142$

or,

$$X = \frac{1{,}000{,}000}{2\pi fc} \tag{1.6}$$

where c is in microfarads

In capacitors, the current and voltage are *not* in phase, but are shifted by 90 degrees. In the case of Figure 1.8(b) an application of a DC voltage across A-C will initially see an uncharged capacitor behaving like a short circuit (zero ohms). A current will then flow through R and the plates of the capacitor will begin to charge. As the voltage rises across B-C, the current will reduce until, once the plates are fully charged, the voltage will rise to a maximum (as across A-C) and the current will reduce to zero. All conduction will then cease. Thus it can be seen how the *current leads the voltage*: the current flows first, then it falls as the voltage rises. The *electrostatic* charge is a *voltage effect*.

In the case of Figure 1.8(c) the effect is the reverse. Inductors work on an *electromagnetic* principle, which is a *current effect*. The formula for the reactance (X) of an inductor is given by:

$$X = 2\pi fL \tag{1.7}$$

Where: f = frequency in hertz (Hz)
L = inductance in henries (L)

In this case, the reactance is *directly* proportional to the frequency. As the frequency rises, so does the reactance, and the *voltage leads the current*. In the cases of both the capacitor and the inductor, the voltage and current are 90 degrees out of phase. Equation 1.4 showed the formula for calculating the power from the voltage and the current, and it was noted that for AC currents and voltages there should be a phase angle multiplier, cos θ (theta). The cosine of 90 degrees is *zero*, therefore whatever values of voltage and current exist in the circuit, the power dissipation (heating effect) in inductors and capacitors is always zero, (except for losses due to imperfections). This is *wattless power*, and is why AC power circuits are measured in VA (volt-amps) and not in watts. Electricity meters measure kVA because in heavily inductive loads, such as electrical motors and machinery, the kW value would be less, and the electricity company would not be charging for all the current and voltage that they were supplying.

In Figure 1.8(a), a 10 volt, 1 kHz voltage at terminals A-C would give rise to a voltage of 5 volts across B-C if R_2 was equal to R_1, at say 100 ohms each, and the resistors would heat up with the power dissipation. In the case of Figure 1.8(b), we could select a capacitor to have a reactance of 100 ohms at 1 kHz, but the circuit would not behave in the same way. The capacitor would be selected from the formula

$$X = \frac{1{,}000{,}000}{2\pi fc}$$

which transposes to $$c = \frac{1{,}000{,}000}{2\pi fx}$$

so for 100 ohms at 1 kHz $$c = \frac{1{,}000{,}000}{6.284(2\pi) \times 1{,}000 \times 100}$$

$$c = \frac{1{,}000{,}000}{628400}$$

$$c = 1.6 \text{ microfarads } (\mu F)$$

However, despite the resistance and reactance in the circuit both being equal at 100 ohms, the total impedance (resistance *plus* reactance) would *not* be 200 ohms. The same current would flow through both components, but whereas the voltage across the resistor would be in phase with the current, the voltage across the capacitor would be 90 degrees out of phase with the current. We can draw a right-angled triangle, as in Figure 1.9, with one side representing the resistance and the other side, at 90 degrees, representing the reactance. The total impedance (Z) would be represented by the hypotenuse. From Pythagorus' theorem, the square of the hypotenuse is equal to the sum of the squares of the other two sides. Therefore:

$$100^2 + 100^2 = 10{,}000 + 10{,}000 = 20{,}000$$

$$Z^2 = 20{,}000$$

$$Z = \sqrt{20{,}000}$$

$$Z = 141$$

total impedance = 141 ohms

The resistor would dissipate power in the form of heat, but the capacitor would dissipate no power, and would not heat up. The inductive circuit

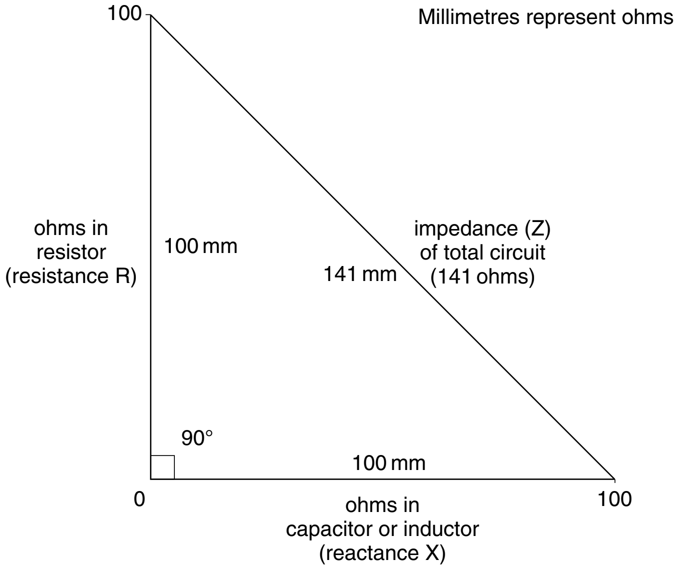


Figure 1.9 Phase angle vector triangle, using 1 mm to represent 1 ohm

in Figure 1.8(c) would behave similarly, except that the phase of the voltage across the inductor would be 90 degrees out with the voltage across the resistor in the opposite direction to the phase shift across the capacitor. The 90 degree differences in opposite directions leads to the 180 degree phase shift between capacitors and inductors, which gives rise to the resonance in tuned circuits. A capacitor and inductor in series are the electrical equivalent to the mass and the spring shown in Figure 1.10. Both are tuned circuits, and both work in the same way, by transferring energy backwards and forwards and dissipating very little of it. The mass takes in energy when one tries to move it, but releases it again when one tries to stop it – this is why the brakes get hot when a car stops. A spring takes in energy when it is stretched or compressed, but releases it when it is released. A mass, basically, 'wants' to stay free of accelerations and decelerations. A spring, basically, does not 'want' to change its length. The two together, given their different phase relationships, take in and release energy alternately, and so remain in oscillation until frictional losses eventually turn the energy into heat.

At the risk of labouring the point (but it is really not well understood by most loudspeaker users) if we have a mass and a spring as shown in Figure 1.10 (a) at equilibrium, the spring is 'happy' because it is at its equilibrium length given the forces acting upon it. The mass is also 'happy' because it is at rest; it is neither accelerating nor decelerating. If somebody then pulls down on the mass, and holds it there, the mass is still happy, but the spring is not, because it is stretched as in Figure 1.10(b). If the mass is released, the spring will act to overcome the inertia of the mass so that it can return to its equilibrium length. When the spring reaches its equilibrium length the mass is in motion, so the mass times velocity gives it momentum, which will carry it through the rest position of the spring, and will begin to compress it, as shown in Figure 1.10(c). The 'unhappy' spring

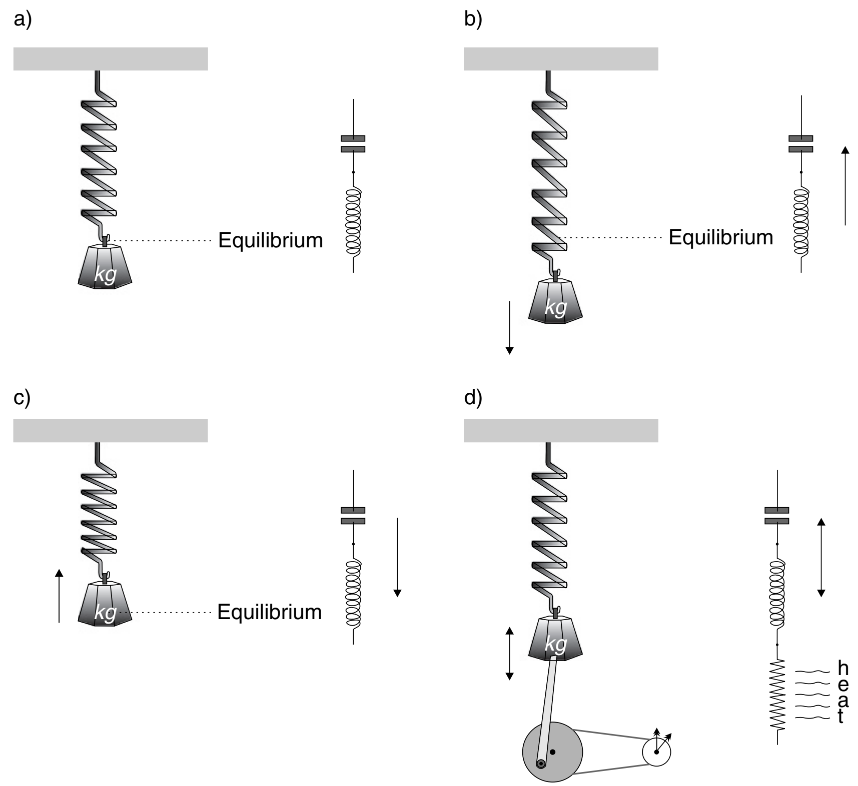


Figure 1.10 Masses, springs and resistances vis-à-vis inductors, capacitors and resistors

therefore begins to slow the mass, so that it (the spring) can once again return to its 'happy' equilibrium position, but it overshoots once again, and the oscillation continues. This is a reactive system like a capacitor and inductor, and the only energy loss is due to its imperfection; which in the mechanical case is friction and air resistance, and in the electrical system is electrical resistance due to the less than perfect conductors – the coil of an inductor will always have a small resistance due to the wire.

Now, if we add some resistance to the mechanical system, we can make it do some work, like moving the hands of a clock, as in Figure 1.10(d), but the work takes energy, so the oscillations will be reduced more rapidly – the oscillations will be *damped*. The electrical circuit equivalent would be to put a resistor in series with the inductor and the capacitor, so that the current flowing through the circuit will produce some heat, and damp the electrical oscillation.

When we load a loudspeaker diaphragm with a horn, the air in front of the loudspeaker cannot escape to the side of the diaphragm, and the stretching pressure due to the sideways component of the wave is restricted. The reactive conditions are therefore minimised, and the air load in the horn provides a substantially resistive load, with the particle velocity and pressure in phase, so more useful work can be done by the diaphragm, such as producing sound instead of just flapping backwards and forwards.

These concepts are relative to horns and direct radiators with their predominantly resistive and reactive loadings respectively. The resistive horn loading tends to be efficient because it gives rise to *acoustic power* being radiated. The largely reactive loading on a direct radiator gives rise to little power being radiated, just as little heat is produced in a capacitor or an inductor.

1.7 The bigger picture

This chapter has tried to set out the fundamental characteristics of sound radiation from moving coil loudspeakers, which represent 99% of all drive units manufactured, but up to now we have only been referring to single drive-units, which as we have previously discussed cannot be realistically expected to cover the full frequency range. Once we are forced to consider multi-driver systems with their obligatory crossover filters, the combined system can take on a further considerable degree of complexity. Many entire loudspeaker systems, some of great complexity, are made only from drivers using this motor technique. However, there are many other types of drive systems which will be discussed in the following chapter, and which we will need to look at in order to get a better appreciation of the characteristics which they can offer. Chapter 3 will then discuss the enclosures in which we must mount them, because the cabinets give rise to their own, extra complications as they affect the air loading on the diaphragms of open-framed drive units. Once we put the whole loudspeaker assemblies into rooms, as described in Chapter 7, a further set of complications arise. As all of these aspects of loudspeaker design and use compound each other, achieving a uniform response in all cases is not possible. For this reason, as perfection is not possible, much of this book will discuss the aspects of what we hear and what we may *need* to hear at various different stages of the music recording process; and, of course during the ultimate objective of domestic listening enjoyment. It will become apparent that it is all about compromise, but the optimum compromise points cannot be chosen without a thorough understanding of the overriding priorities at each stage of the process. A single driver is only a very small component part of a complex system that takes a musical signal from the amplifier to the ear.

References

1 Fahy, F., Walker, J., ‘Fundamentals of Noise and Vibration’, Chapter 5, Spon Press, London, UK, (1998)
2 Rice, C., Kellogg, E., ‘Notes on the Development of a New Type of Hornless Loudspeaker’, Transactions, American Institute of Electrical Engineers, Vol 44, pp 461–475, (1925)
3 Briggs, G., ‘Loudspeakers’ Fourth Edition, Wharfedale Wireless Works Ltd, Bradford, England (1955) – Reprinted by Audio Amateur Publications Inc, Peterborough, NH, USA (1990)

Bibliography

1 Borwick, J., ‘Loudspeaker and Headphone Handbook’, Third Edition, Focal Press, Oxford, UK (2001)
2 Colloms, M., ‘High Performance Loudspeakers’, 6th Edition, John Wiley and Sons, Chichester, UK (2005)
3 Eargle, J. M., ‘Loudspeaker Handbook’, Chapman and Hall, New York, USA (1997)
4 Briggs, G. A., ‘Loudspeakers’, Fifth Edition, Rank Wharfedale Ltd, Bradford, UK (1958) – Reprinted until 1972
5 Jordan, E., ‘Loudspeakers’, Focal Press, Oxford, UK (1963)
6 Borwick, J., ‘Loudspeaker and Headphone Handbook’, Second Edition, Focal Press, Oxford, UK (1994) (significantly different in content from the aforementioned Third Edition.)

Chapter 2

Diversity of design

Although the original moving-coil, cone loudspeaker of Rice and Kellogg was the first true loudspeaker of a type that we know today, it was, itself, a development of ideas which had gone before, principally relating to the design of telephone earpieces, which were *not* very loud speakers. The moving coil direct radiator, along with amplifiers as great as 15 watts output – which was then huge – soon opened a door to room-filling sound levels, and, within only a couple of years, talking pictures at the cinema. The need to fill larger and larger theatres with sound led to horn designs, and the need for greater bandwidth led to the separation of the drive units into frequency ranges where they could operate more efficiently. Thus began a refinement and specialisation of designs which continues to this day, with ever more ideas, magnet materials, diaphragm materials and radiator concepts all designed essentially to do the same thing – convert electrical energy into sound waves. What follows in this chapter is a discussion of some of the various ways in which this conversion can be made to take place, and the strengths and weaknesses of the various approaches.

2.1 Moving-coil cone loudspeakers

Of all types of drive units, there is probably none so varied in size, shape, materials of construction or performance as the moving coil cone loudspeaker. They basically all follow the concept shown in Figure 2.1, and little has changed in the underlying principles of their operation in the 80 years of existence so far. They all need a magnet, which was often an electro-magnet in the early years before permanent magnets of sufficient strength were developed. In this case, a 'field coil' was supplied with a DC current sufficient to generate the required strength of magnetic field for the 'voice coil' (which was fed with the output signal from the amplifier), to drive the cone with the required level of sensitivity. Early permanent magnets were often made from iron and chromium. Aluminium, nickel and cobalt were variously used in the early 1930s, alloyed with iron in different combinations, and the three together gave rise to the name Alnico. In the 1970s, the civil war in the Congo (then Zaire) created a big hole in the production of cobalt, whose price rose astronomically in a very short period of time. This led to the use of ferrite materials, known as ceramic magnets, which had their strengths and weaknesses which will be discussed in Section 2.1.6. More recently, 'rare earth' magnets, principally made from

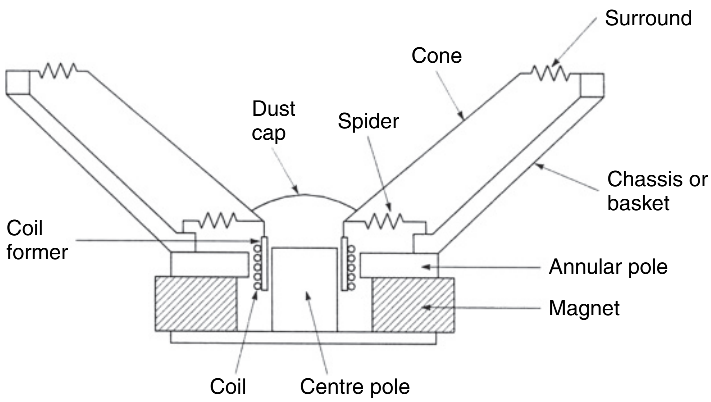


Figure 2.1 The components of a moving coil loudspeaker

neodymium and samarium based alloys, such as neodymium with iron and boron, or samarium and cobalt, have led to very light weight magnets, and opened a door to new magnet shapes and magnetic field designs. The basic concept of two different magnet structures is shown in Figure 2.2.

The magnetic circuits are designed to concentrate the magnetic field in a circular gap, as shown in Figure 2.3. In this gap is inserted the voice coil, which receives the electrical drive current from the power amplifier. This current produces its own, alternating magnetic field, whose phase and amplitude depend on the drive signal. The variable field interacts with the static field in the circular gap, and creates a force which either causes the voice coil to move into or out of the gap. Of course, a means is required to maintain the coil centralised in the gap, and this is achieved by the use of centring device, or inner suspension, which is still often referred to as a spider for reasons which should be clear from an inspection of Figure 2.4. A more typical modern device is shown in Figure 2.5. A chassis, also known as a frame or basket, supports the whole assembly and enables it to be mounted on a front baffle. The cone is connected rigidly to the former upon which the voice coils is wound, and is also connected more or less at the same point to the inner suspension. At the chassis' outer edge the cone is attached via a flexible outer suspension, or surround, which may take the form of half-rolls, corrugations, or pleats. These will be discussed in more detail in Section 2.1.2. A dust cap is then normally placed in the apex of the cone in order to prevent the ingress of dust and any abrasive dirt, and may also be used as an air pump to cool the voice coil and gap when the cone assembly moves in and out.

2.1.1 Cones

A three-way loudspeaker system consisting entirely of cone drivers is shown in Figure 2.6. Although the cone drivers all follow the above principles of construction, their designs are very different. In Chapter 1 it was explained how, at low frequencies, the electro-acoustic conversion efficiency is low, because the air tends to move out of the way of the vibrating

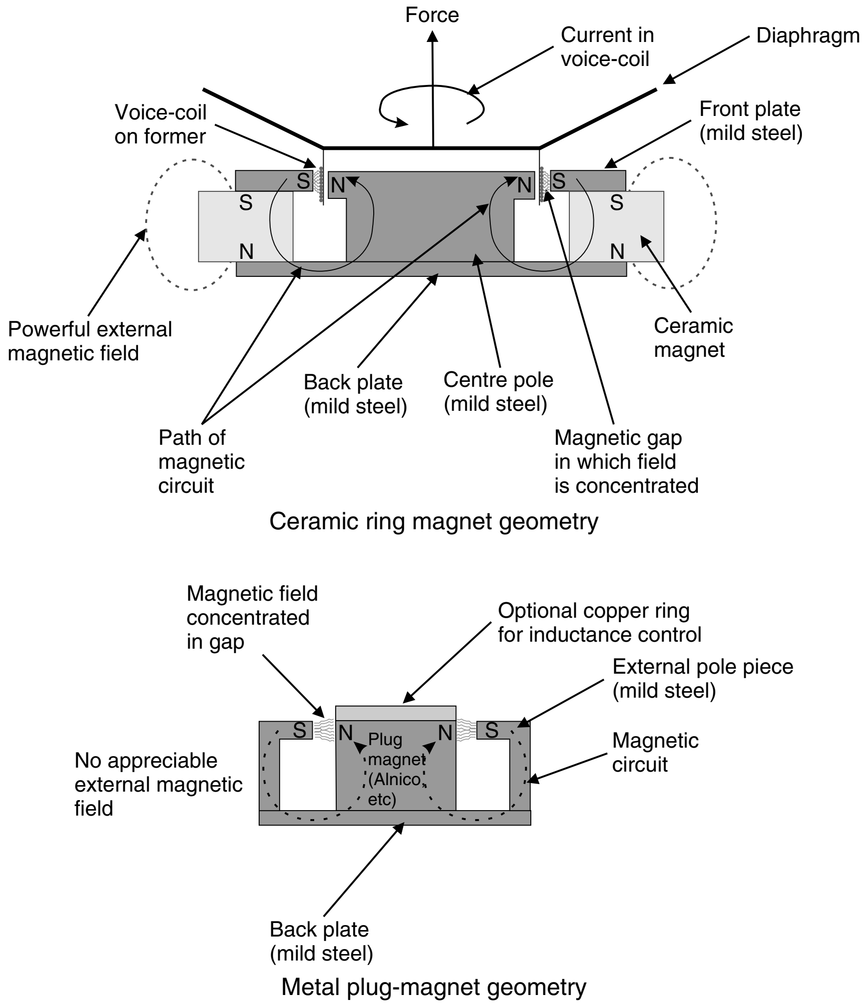
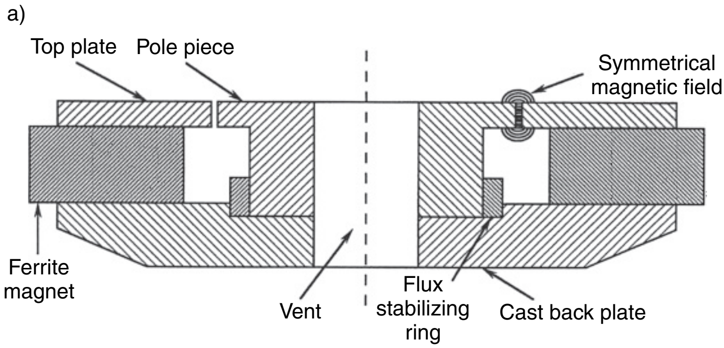


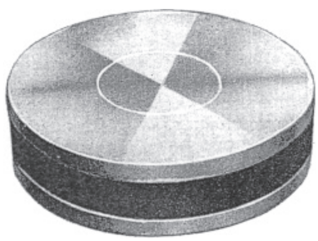
Figure 2.2 Typical motor topologies

cone. The result of this is that for a reasonable on-axis sensitivity, the cone needs to be quite large. In the loudspeaker shown in Figure 2.6, the low frequency cone is of nominally 12 inch diameter (300 mm), although the effective radiating area is only just over 10 inches because the surround does not contribute much to the radiation. The cone needs to be rigid, because if it breaks up into non-uniform movement, phase cancellations will occur at some frequencies and the subsequent frequency response will not be flat. In some cases, cone break up can be used to extend the frequency response, and can be used in musical instrument loudspeakers to create desirable colouration in the sound, but for flat, uncoloured low-frequency responses, the piston which is pumping the air needs to maintain its rigidity. Some of the ways in which a cone can break up are shown in Figure 2.7.



Magnet geometry for concentrating the flux in the voice-coil gap – section view (courtesy JBL inc.)

b)



The circular voice-coil gap – perspective view

Figure 2.3 The voice-coil gap

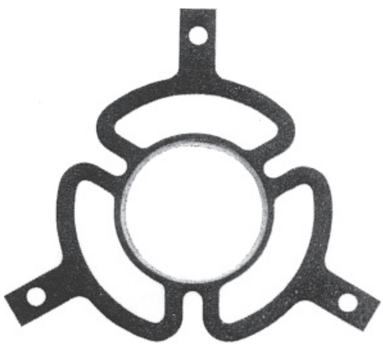


Figure 2.4 An early centring device – a 'spider', some of which had more legs than the one shown

Once cones exceed a diameter of about 18 inches (460 mm) it can become difficult to maintain their rigidity. The gain in efficiency due to the large radiating areas of big drivers can rapidly be offset by the greater proportional weight needed to keep them rigid. Many designers favour multiple, smaller drivers to the use of single larger drivers partly because they feel

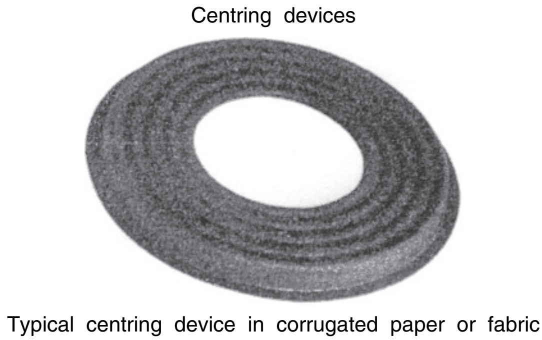


Figure 2.5 A typical, modern centring device, or inner suspension

Figure 2.6 A full-range, all cone driver loudspeaker system – the JBL L100

that they can keep these better controlled. It is unusual to see drivers of greater than 18 inch diameter, although they do exist, as shown in Figure 2.8. Sandwich cones, honeycomb cones and Kevlar and carbon fibre and metal cones have all been employed in attempts to maintain cone rigidity, and consequently the pistonic movement. In each case, the cones exhibit different characteristics above certain frequencies, so the suitability of each material or construction may depend upon the upper frequency limit to which a driver will be used. Solid cones have also been used, but they can introduce as many problems as they solve, and they are not so obviously beneficial as they may at first appear to be. One problem with many of these approaches has been that the near-perfect rigidity has improved matters at low frequencies but has only pushed the resonances up in frequency, rather than eliminating them. When stiff structures do

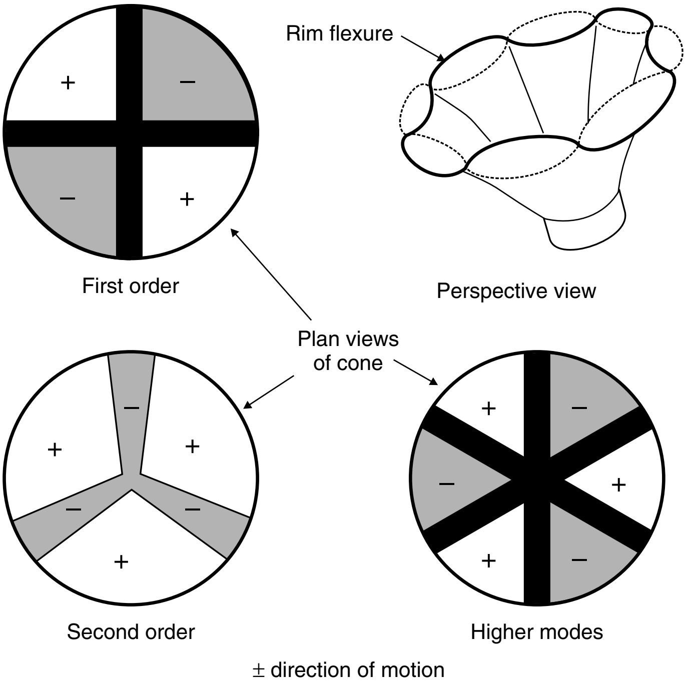


Figure 2.7 Bell-mode break-up in cones

Figure 2.8 A very large loudspeaker – a 30 inch (800 mm) low frequency loudspeaker with radial reinforcing ribs to augment the cone rigidity

break up they tend to do so much more severely, so crossover frequencies must be chosen well away from the break-up frequencies if colouration is to be avoided. It has proved difficult to achieve uncoloured mid-range responses from highly rigid low frequency driver cones, so they are often best restricted to use at low frequencies only.

Bextrene, a mixture of polystyrene and neoprene, was pioneered as a cone material by the BBC, in the UK, as far back as the 1960s. This was originally researched largely to find a solution to the inconsistency problems encountered in the manufacture of paper pulp (cardboard) cones.

Paper, being made from wood, which is a natural material, can suffer from the problem of all natural organic materials; they are not homogeneous substances so they tend to vary from batch to batch. Bextrene was well-damped and resisted break-up to relatively high frequencies. Designs have been employed using Bextrene cones which have used 12 inch (300 mm) bass drivers up to crossover frequencies beyond 1.5 kHz with little mid-range colouration. Polypropylene has since been developed for use as a cone material, offering even more consistency, long term stability and sensitivity. However, opinions vary about the sonic neutrality of polypropylene-coned drivers.

The original loudspeaker of Rice and Kellogg used a paper cone. Quite remarkably, despite all of the modern developments in materials and construction, paper pulp cones, and even folded and seamed paper cones, are still in use in all levels of performance ranges. Paper pulp cones are made by drawing a slurry (wet mix) of paper fibres through a fine screen in the required shape. The resulting cone is then cured and dried before being cut to the exact size required. This 'old fashioned' material still exhibits excellent characteristics of high rigidity and high internal damping, and these are two things which normally are contradictory inasmuch as the augmentation of either one usually tends to reduce the other. At low frequencies the rigidity is necessary to maintain piston action, but once any break-up does begin, the internal damping of the cone material needs to suppress the waves which travel as shown in Figure 2.9, which would cause peaks and dips in the frequency response. The bass cone shown in Figure 2.6 has been further treated on both sides with a damping material known as Aquaplas, which also adds some mass, and this lowers the free-air resonance of the driver.

Despite the fact that many drive units of similar specification which employ different cone materials may perform very closely in objective measurement, there is strong evidence that some very similarly performing drivers do not *sound* the same. In very high quality loudspeakers, paper pulp is still a favoured material, despite its sensitivity to humidity and batch-to-batch variation. Colloms[1] refers to the well-balanced characteristics of paper pulp, together with the fact that its properties and manufacturing techniques are well-understood, as strong justification for its continued use. He notes how some high-loss materials also may tend to lose, or mask, fine musical detail. The authors of this book have noticed a loss of reverberation detail when substituting some synthetic cones for paper cones used up to 1 kHz, and have received comments from professional users about guitar strings not sounding as new when heard via the synthetic cones as when heard via high quality paper pulp cones. The Celestion loudspeakers company still produce the exact model of guitar loudspeaker which was made famous in the Vox AC30 guitar amplifier of the 1960s. This blue-chassised driver has resisted all efforts to update its construction, yet still maintain its highly desirable sound qualities. A great number of musicians claim to hear their guitars more 'clearly' via paper cones.

The observations about guitar strings could suggest a harmonic enhancement due to non-linear distortion products enriching the sound, but harmonic distortion could not explain the increased sensation of low-level detail and reverberation. Synthesising natural sounding reverberation is

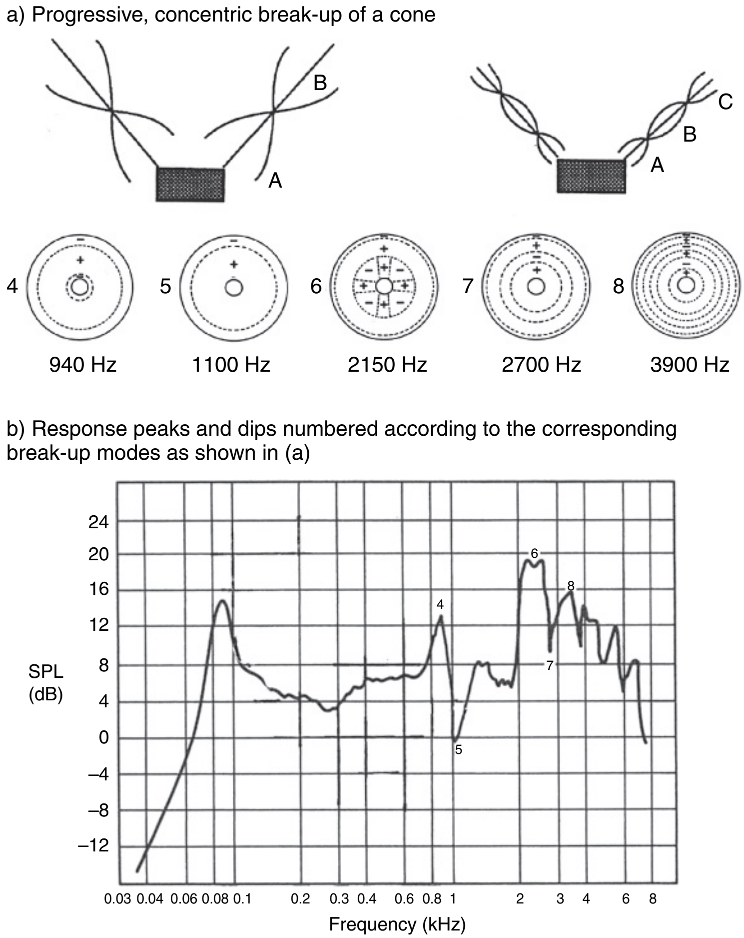


Figure 2.9 Concentric modes in loudspeaker cones

not something which one would expect from the addition of harmonic distortion. It seems probable that the guitar strings are benefiting from the same characteristics which are enabling the greater resolution of reverberation and room effects, and paper-pulp seems to be a good performer in this respect. Work is currently under way to investigate the possibility of intermodulation distortion contributing to the low level detail loss with certain materials, given rise to by non-linear hysteresis effects connected with the damping action. Intermodulation distortions will be discussed in more detail in Chapter 9, but it results in harmonically *and* non-harmonically related products which together tend to produce a noise signal below the music.

As the frequency rises, two things begin to affect the performance of a cone driver. It was shown in Figures 1.3. and 1.4 how the directivity of a cone, or any pistonic radiating surface, narrows as the wavelength become small compared to the circumference of the radiating area. However, as frequencies rise, the mass of the moving assembly gradually tends to oppose more strongly the force which is trying to move it. There comes a point where the increasing efficiency of radiation due to the greater radiation resistance provided by the air as the frequency rises can no longer compensate for the mass effects of the moving parts, and the power response of the driver begins to roll off. For a 15 inch (380 mm) loudspeaker, 1 kHz is about the upper limit of either its flat response range or the acceptability of its narrowing directivity. Nevertheless, the directivity narrowing may not be too severe if the centre section of the cone begins to decouple itself from the outer section, as often happens – either by design or accident. Eight inch loudspeakers (200 mm) can work well up to 2 kHz, or more, but the compromises which must be made if their responses are to be extended at the bottom end may begin to degrade the higher frequency performance. In the loudspeaker shown in Figure 2.6, the 12 inch (300 mm) low frequency driver, having a free-air resonance of 25 Hz, is used up to a frequency of around 1 kHz, at which point the crossover begins to divert the signal towards the 5 inch (125 mm) cone. Five octaves is just about the limit of the bandwidth of a cone driver if very high quality, wide directivity, minimum-compromise sonic performance is required.

In smaller loudspeakers, cone rigidity is much easier to achieve, therefore much lighter moving assemblies can be employed which can exhibit good efficiency without the massive magnet assemblies needed for the low frequency drivers. The low frequency driver shown in Figure 2.6 has a sensitivity of 89 dB for 1 watt input at 1 metre distance, yet with a much smaller magnet but a lighter moving assembly, the 5 inch mid-range driver has a corresponding sensitivity of 94 dB. This sensitivity increase can be important, because smaller loudspeakers cannot lose the waste heat as easily as can large loudspeakers, which was one driving force behind the developments of domes, as will be discussed in Section 2.2. The tiny tweeter cone shown in Figure 2.6 is only 1½ inches (38 mm) in diameter, but handles frequencies from 4 kHz to almost 20 kHz, so in this design of loudspeaker cabinet all frequencies from below 40 Hz to almost 20 kHz are handled by paper cones.

2.1.2 Surrounds

The outer suspension, or cone surround, serves two functions. The most obvious is to maintain the outer edge of the cone stable during the rapid movements along the front-back axis, but surrounds also serve to absorb vibrational waves which propagate from the voice coil in the manner shown in Figure 2.9. A selection of surround designs are shown in Figure 2.10. The surrounds are variously formed from a continuation of the cone material or from separate materials attached with adhesives. Polyurethane foam, butyl rubber, nitryl rubber, cambric (pronounced Kāymbrik - a woven linen or cotton fabric), or other treated fabrics are frequently used as cone surround materials. PVC is also sometimes used. Figure 2.10(a) shows a half-roll surround. These are usually made from synthetic foams or rubbers. They allow long travel because they tend to stretch in an elastic manner over a considerable range of movement, and hence give rise to

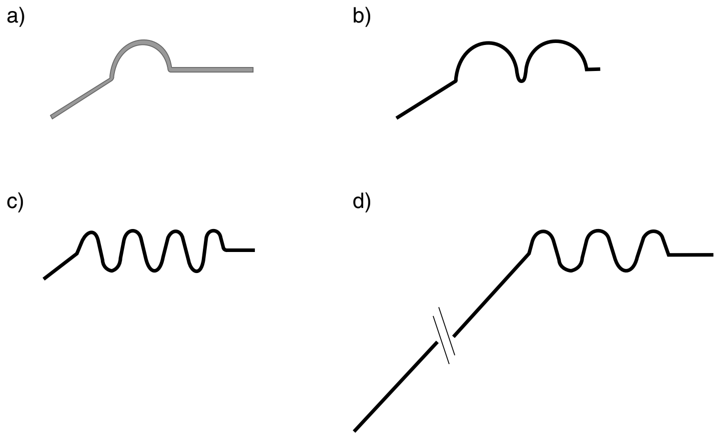


Figure 2.10 Cone surround variants. a) Half-roll of polyurethane foam – low stiffness (high compliance) for long travel, but requires precise choice of centring device for controlled linearity. b) Double half-roll cloth – shape of rolls can precisely tune the stiffness characteristic. c) Multiple-roll accordion pleat – long travel but prone to rim-resonance response dip problems (see Figure 2.11). d) One piece cone/surround with treated edge – stiff, non linear suspension, provides HF resonance peak. Principally used for musical instrument loudspeakers to prevent over-excursions of the cone

little distortion which would be caused by restraining the cone travel. When the materials are carefully chosen they can effectively absorb the resonant modes which pass radially along the cone, thus avoiding standing waves within the materials of the cones. The *double* half-roll surrounds, shown in Figure 2.10(b) are usually made from treated cloths. They are more rugged than the single half-rolls, and find much use in sound reinforcement and musical instrument amplification, but tend in general to be used in less long-throw loudspeakers with higher resonant frequencies than those which usually employ the single half-rolls. The extra stiffness of the double half-roll surrounds both adds to their ruggedness and increases the resonant frequency of the moving system as compared to single half-roll devices. Figure 2.10(c) shows a concertina (accordion) surround. These are often pressed or moulded as part of the cone. They can allow extended travel, but unless very carefully damped can give rise to resonance problems. They also tend to be stiffer than the single or double half-roll surrounds, and are therefore not often found on low-resonance designs.

Surround *materials* are also a specialised subject. Some of the foams which are commonly used can deteriorate much quicker than expected in sunlight or polluted atmospheres, and can also suffer from insect damage. Nevertheless, in clean, temperature-controlled environments such as exist in many sound control rooms they can easily last for 20 years or more, and can be replaced by skilled artisans without removing the cone from the chassis. Plasticised PVC is another material which has been employed with success, and is found on some mid-range cones where its properties efficiently absorb the waves reaching the cone edges. In sealed cabinets, the surrounds must also resist the differential pressures between the inside

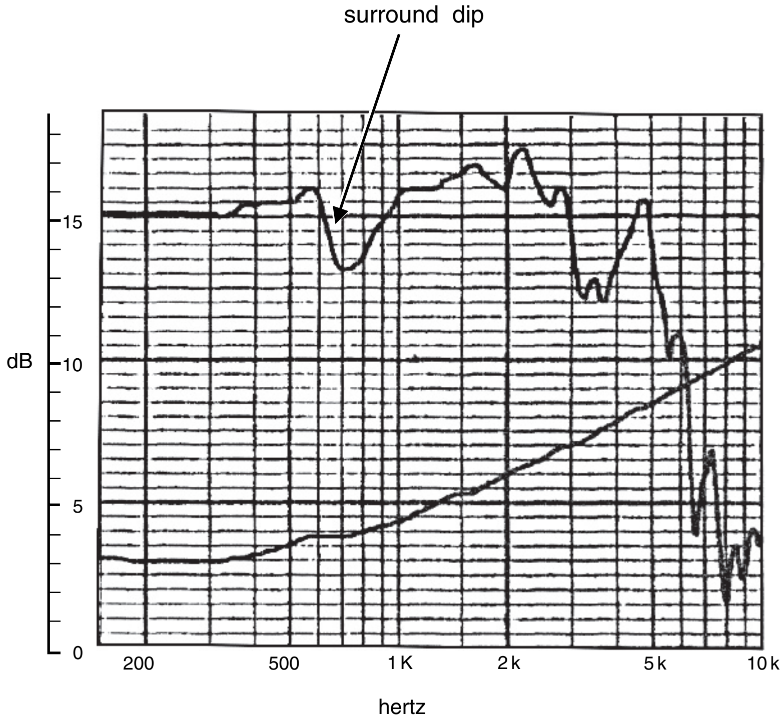


Figure 2.11 Response dip due to a surround resonance

and the outside of the cabinet, and this fact can preclude the use of some foams in certain cases. In order to fulfil all of the demands made of them, surrounds are quite specialised devices. The surrounds not only need to be selected according to the density of the cone material, but also the principal frequency ranges over which a driver will work, and the excursion limits within which they must allow relatively unrestricted movement. For wide-range drive units, the compromise choices are not simple. Figure 2.11 shows a response dip due to an antiphase movement of a surround. The fact that the dip exists in the response of such a high quality driver suggests that solutions to the dip problem would have compromised the overall response to a greater degree.

2.1.3 Rear suspensions

The prime function of the rear suspension, or spider, is to maintain the coil centralised in the magnetic gap. On its inner edge it is connected to the voice coil former, and at its outer edge it is glued to the chassis. Modern suspensions, such as the one shown in Figure 2.5, are usually made from phenolic resin impregnated cloth, hot pressed into shape. Care has to be taken in the design of the corrugations to ensure that movement in one direction is not favoured over the other direction, because an asymmetrical movement would give rise to non-linear distortion. Some designs have employed double spiders, mirror imaged, in order to ensure symmetrical linear travel of the cone. A double spider arrangement is shown in Figure 2.12. These are fine in vented magnet designs, but a complication in double suspension designs arises if they must allow air to pass through to cool the voice coil in designs that do not have vented magnet systems. As with the corrugated, concertina surrounds, the corrugated inner suspensions can suffer from resonance problems unless they are carefully 'tuned'.

The stability of the inner suspensions needs to be very good because the gap between the voice coil and the magnet can be less than half a millimetre, even with a relatively large cone and coil. With large, heavy cone/coil assemblies, the suspensions can become stretched if the drivers are stored in a horizontal position without adequate support for the cone (which effectively means in the manufacturers' shipping boxes). Likewise, complete loudspeaker systems should not be stored on their backs or 'cone sag' is likely to result. Once mounted vertical again, the suspension may have 'set' to a new equilibrium position which is not in the centre of the cone's axial travel, hence the cone excursion will be prematurely limited in one direction. In many cases, this cone sag cannot be corrected by any simple means, and so storage conditions likely to give rise to it should be avoided.

The inner suspension is very critical because it usually provides the main restoring force for centralising the cone in the axial as well as the radial directions. Over 50 years ago, Briggs recognised not only the third-harmonic distortion-producing mechanisms of badly designed suspensions, but also the fact that inadequate suspensions could give rise to distorted transient responses[2]. Over 40 years later, Colloms wrote about work done at KEF which correlated well with his own experiences that some inner suspensions could give excellent, low distortion results on sine waves and the more open bass waveforms of orchestral music, but could be slow in responding to the more percussive bass sounds found in much modern

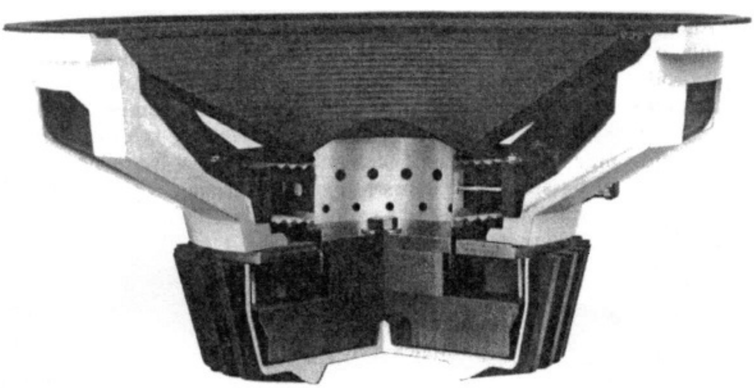


Figure 2.12 Cut-away view of a 380 mm loudspeaker with double spider (centring device) construction (Cetec-Gauss Inc)

music[1]. The KEF findings had emerged from investigations into the different low frequency measurements obtained via the use of steady state or impulsive signal sources. [These test methods are described further in Chapter 9.] The differences have been attributed to hysteresis in the suspensions, shown diagrammatically in Figure 2.13.

In some very small cone loudspeakers, designed only for high frequency use, the additional complexity of the use of an inner suspension can often be omitted, the external surround being sufficient to maintain the cone in a central position, and thus avoiding all the inner-suspension-related problems which larger, heavier cones must endure.

Although the suspension systems, both surrounds and spiders, are mechanically essential in low and mid frequency cone loudspeakers, towards their excursion limits they all begin to give rise to third harmonic distortion due to the approaching elasticity limits where they become less compliant (i.e. more stiff). Once they no longer obey Hooke's Law (the law governing the relationship between force and compression [or expansion] of a simple spring) they can produce quite a number of undesirable artefacts.

2.1.4 The chassis

Loudspeaker chassis provide a frame for the mounting of the magnet system and inner and outer suspensions. They can be made of plastic, or pressed or cast metal. Metal is preferred when power levels are high because it helps to conduct the heat away from the magnet assembly. Cast

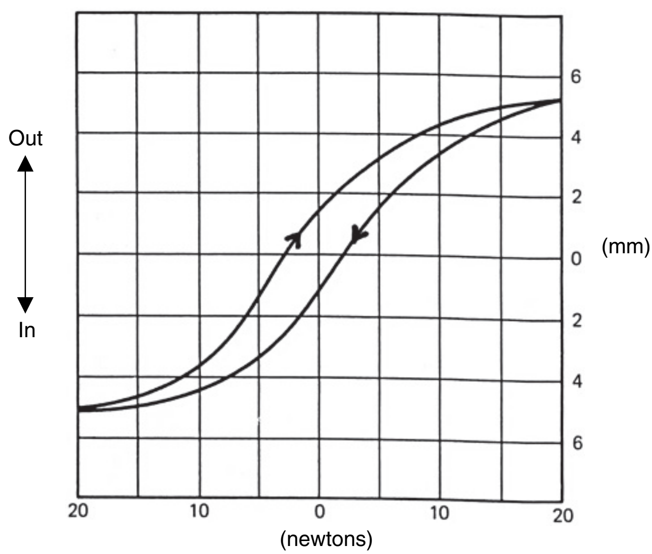


Figure 2.13 Hysteresis curves – the hysteresis curves represent processes which are cyclic but where the forward and backward processes do not follow the same path. 'Hysteresis' is from the Greek word meaning 'to lag behind'

metal is usually preferable to pressed metal on grounds of stability and dimensional accuracy; cast aluminium being the material of choice for most large professional bass units. With magnet assemblies weighing 10 kg or more on some 18 inch loudspeakers, the chassis ('frames', 'baskets') need to be strong to withstand shipping shocks without disturbing the centralisation of the sub-half-millimetre coil clearances. Also, coil temperatures of 250 degrees C, or over, need to be withstood and dissipated without warping. Unfortunately, there is a conflicting requirement between the strength of the chassis and the need not to impede the free movement of air between the cone and the inside of the box (or outside the box in the case of external chassis designs, as shown in Figure 2.14). The chassis therefore needs to be as strong as necessary for support, whilst being as open as required for non obstruction of the air adjacent to the cone. It is also important that it should not suffer from resonance problems, and so needs to be well acoustically damped.

Once a cone and coil assembly is mounted in a chassis at the surround and the inner suppression attachment points, its mass will resonate with the compliance of the suspension systems to determine the free air resonance of the driver. The free air resonance normally defines the lower response limit of a loudspeaker system in any given volume or design of cabinet. However, there do exist some special designs of loudspeaker systems which drive the bass units through their resonances, but they are rare, and need electronic compensation.

2.1.5 The voice-coil assembly

The voice-coil former is normally attached to a cone at some point between its apex and a point mid-way between the apex and the perimeter. The coil former, of cylindrical design, must be mechanically stable under high degrees of vibration and temperature changes, neither deforming in circularity nor expanding or contracting to any significant degree. Without the required stability, the coil or the former could rub on the sides of

Figure 2.14 A 15 inch (380 mm) loudspeaker manufactured by the British company Volt Loudspeakers Ltd, employing an external chassis for improved heat dissipation

the magnetic gap and make undesirable scraping noises. Although paper was used to great effect on early, low-power loudspeakers, modern day, high-temperature voice coils need to be bonded to thermally stable formers with thermally stable adhesives. Glass fibre has been used as a former material, but polyamides are now more normal, (Nomex, Kapton etc). Aluminium has also been used, but metal formers can suffer from eddy current problems by acting as a shorted turn in the alternating magnetic field, and can also conduct very high temperatures to the necks of the cones, where charring, melting or softening can occur, and adhesives may also be caused to fail.

The coils are almost exclusively made from copper or aluminium, although silver has also been used, reportedly to some sonic benefit[1]. Copper offers lower resistance, but aluminium offers lighter weight. Which one is the most appropriate depends on many other design factors, but both materials have been used at either frequency extreme – there are no hard and fast rules. Copper clad aluminium is another option, which simplifies the soldering problems which may be encountered with pure aluminium. Round wire and rectangular section (ribbon) wire are also options. The ribbon wire packs more densely, eliminating the gaps between the adjacent round wires, but round wires offer much simpler winding processes. To prevent short circuits between turns, the copper wire is insulated with a heat resistant lacquer, and aluminium wires are anodised, which creates a layer of non-conducting aluminium oxide on the surface. In either case, the wires may also be coated with a thermosetting adhesive before winding, which, after curing, helps to render the entire assembly more rigid.

The most appropriate diameter of a voice coil is also the subject of compromise. Larger diameter voice coils have more surface area for any given length of wire than coils of less diameter, at least for any given gap depth, and can therefore lose heat more easily. They can also help to stiffen a cone by driving it in a more evenly distributed manner. However, if the coil is too big in diameter, the centre of the cone can begin to decouple from the coil, which can cause strange frequency response and directivity problems. Proponents of small diameter voice coils cite advantages of deeper coils in longer gaps giving them design advantages such as deeper gaps with less magnetic material at no thermal dissipation cost. As with so many things in loudspeaker design, the art of the science is finding the best compromise for any given situation. Low frequency efficiency versus high frequency extension, for example, can dictate the optimum choice of coil material as copper or aluminium. Some manufacturers have also great expertise in using particular design concepts or manufacturing processes which suit certain ways of doing things better than others. Electro-Voice, for decades, continued to make excellent 15 inch (380 mm) loudspeakers with 2½ inch (62 mm) voice coils, whilst JBL had long since moved to 4 inch (100 mm) coils for their equivalent designs, but using different magnet topologies.

2.1.6 Magnet systems

The voice coil and the magnet form the motor system of a moving-coil loudspeaker. As mentioned above, either one depends on the other in order to produce the required force to drive the cone in the required direction at the required speed, as instructed by the electrical input signal.

Magnets are a huge and complex subject in themselves, so only some of the more fundamental aspects of their behaviour can be dealt with here, but the Bibliography at the end of this chapter gives references to some excellent further reading. Cost, however, perhaps plays a bigger part in the choice of magnet systems than in any other aspect of the design of cone loudspeakers. Some of the best magnetic materials ever developed for loudspeaker use used cobalt, which, as mentioned in Section 2.1, rose in price by over 2000% in a very short period of time when the civil war in Zaire (the former Belgian Congo) began in the 1970s, because Zaire was the world's largest producer of cobalt. This drove many manufacturers to use ferrite materials, predominantly barium ferrite, which had been developed for deriving the static magnetic fields necessary around cathode ray tubes in television sets.

The design of typical Alnico (*Al*uminium, *ni*ckel, *co*balt and iron), Ferrite (ceramic), neodymium and Alcomax magnets are shown in Figure 2.15, from which the geometrical differences are obvious, (although the Alcomax and Alnico geometries shown are interchangeable). Many modern loudspeakers use neodymium alloys, and an alloy of samarium and cobalt is also finding use in loudspeaker designs. These materials give enormous magnetic strength for their weight, and they have given rise to further changes in magnetic geometry. The ferrite materials are very resistant to loss of magnetism due to time or heat stresses, but they exhibit powerful stray magnetic fields and can pose some difficulties in achieving the desired magnetic field geometries. Alnico is somewhat less durable, but can allow designs enabling very compact and concentrated magnetic fields. Strength for strength, neodymium magnets are much lighter than either ferrite or cobalt alloy magnets, but can be relatively easily demagnetised at relatively low temperatures, and cannot withstand 250 degrees C voice coil temperatures without permanently losing some of their magnetic strength. The metal magnets are good conductors of electricity, but the ferrite magnets are ceramic materials, and hence are electrically non-conductive. The non-conducting nature of the ferrite materials can be a problem unless careful measures are taken to use other means to avoid unnecessary and undesirable flux modulation effects. Iron of high magnetic permeability is used in the magnetic structures shown in Figure 2.15 to complete the magnetic circuit, and to achieve the correct shape of field and density of magnetic flux in the gap in which the coil is positioned. The type of iron used is normally a mild steel of low carbon content, but when very high flux densities are required, a material known as Permendur is often used, especially in compression drivers. Permendur is an iron-cobalt-vanadium alloy, which is very hard and difficult to work, but its magnetic properties, when required, may demand its use.

There are many people who consider the metal magnets to be capable of better sonic performance than the ferrite magnets, citing better resolution of fine detail with materials such as Alnico and the neodymium alloys. It is hard to find evidence of tests which rigorously compare the differences in the magnetic materials only, because the required structural differences needed to get exactly the same magnetic field in the same gap can lead to other changes being necessary. Nevertheless, there is a tendency for many of the highest resolution devices to use metal magnets, and explanations

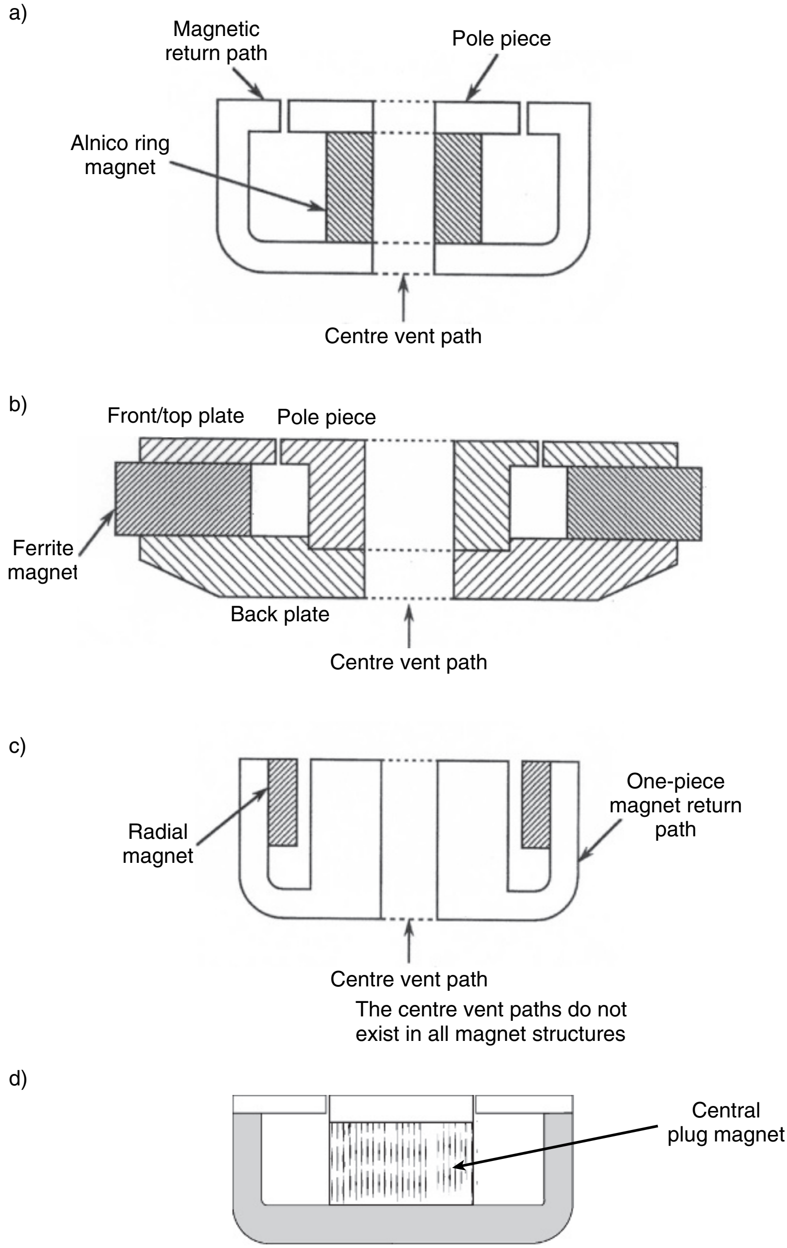


Figure 2.15 Basic magnet structures. a) Alnico ring magnet. b) Ferrite ring magnet. c) Radial, high energy magnet (neodymium etc). d) Alcomax plug magnet a), c) and d) do not have any appreciable external magnetic fields

have been put forward to suggest that the magnetic domain jumps which take place in non-conducting materials can give rise to effects not dissimilar to digital quantising distortion. These jumps are smoothed out by large eddy currents flowing in the electrically conducting magnets. In some loudspeaker designs the central pole-piece of the magnetic assembly is fitted with a copper ring to provide a very low electrical resistance – less than that of steel – to effectively short out any flux-modulation currents.

As shown in Figure 2.15(a), (b) and (c), the entire magnet assemblies in many high-power, low-frequency drivers have cylindrical holes through their central axes to allow air to be pumped through by the dust cover (dome) which caps the voice-coil former on the outer face of the cone. In some other designs, the spider (inner suspension) and the dust cover are of an open weave nature, to allow hot air to pass to the outside of the cabinet. As mentioned earlier, the voice coils can get to temperatures above 250 degrees C in some cases, and this heat needs to be dissipated as quickly as possible, not only to prevent the burn-out of the coil but also to avoid overheating and weakening of the magnet. Furthermore, the resistance of a copper wire rises by about 0.6% for every degree of temperature rise, so a changing voice coil resistance will affect the sensitivity of the motor system and may lead to signal compression.

2.1.7 Ferrofluids

The problem of the conduction of heat from the voice coil to the pole piece of the magnet assembly can, in some instances, be augmented by the use of a ferrofluid. Air is not a good conductor of heat, so much of the heat is transferred from the voice-coil to the magnet assembly by radiation alone. In the 1970s, ferrofluids began to be introduced which were liquids with magnetic particles in colloidal suspension. The magnetic field holds the ferrofluid in the gap, and the good heat-conduction properties of the ferrofluids aids the cooling of the voice-coil by means of a thermal bridge to the magnet assembly. The viscosities of the ferrofluids can be adjusted according to the circumstances of use, a fast-moving tweeter needing lower viscosity than a mid-range driver if viscous damping effects are to be avoided. However, ferrofluids are rarely used in high-excursion low frequency drivers, because the shearing effects of the large axial movements of the coil tend to create non-linear movement due to the non-laminar flow of the fluids. In high frequency drive units, the ferrofluids can be advantageous in damping some mechanical resonances if the viscosities are chosen appropriately.

2.1.8 The complete system

The moving coil cone loudspeaker is quite remarkable in its degree of versatility of application, and has formed the backbone of loudspeaker system design since its first application in 1925. Despite all the technological developments and improvements in materials, Rice and Kellogg, if still alive today, would almost certainly be able to explain the workings of any modern moving coil loudspeakers from simple inspection. When they filed their patent, they already had described, in principle, almost everything that we can find in a modern driver. Loudspeaker design is a science; but there is art in deciding about the best compromise points for what are imperfect devices. For example, whether a complete system should use wider range drivers and fewer crossover points, or narrow range drivers and more crossover points, is a question that may depend more on the circumstances of use than any single measurement at a fixed position. Things such as the room acoustics, the music, the required timbral fidelity and the listening distance may all influence a design, but these things will all be discussed later.

2.2 Dome loudspeakers

In general principle, a dome loudspeaker is a cone loudspeaker with the voice coil having the same diameter as the diaphragm. The diaphragm is also usually inverted, to be convex rather than concave to the exterior. A mid-range dome loudspeaker is shown in Figure 2.16. The development of dome loudspeakers largely grew out of the problems surrounding how to lose heat from the small voice coils of mid and high frequency drivers at high power levels. A 1½ inch (38 mm) cone tweeter, with a coil of only 12 mm diameter simply cannot lose heat quickly enough to prevent itself from burning up at power levels much above 10 watts, because the heat production is all confined in a small space. However, if the coil were to be made the same diameter as the diaphragm, a much greater surface area would be available for heat loss. As light weight is important for sensitivity, the coil former can be kept to a minimum length if the dome diaphragm is convex because it will remain clear of the pole piece of the magnet assembly, as shown in Figure 2.17.

Despite the 'common sense' belief of many people that domes radiate over a wider angle than cones, this is not so. It is important not to confuse

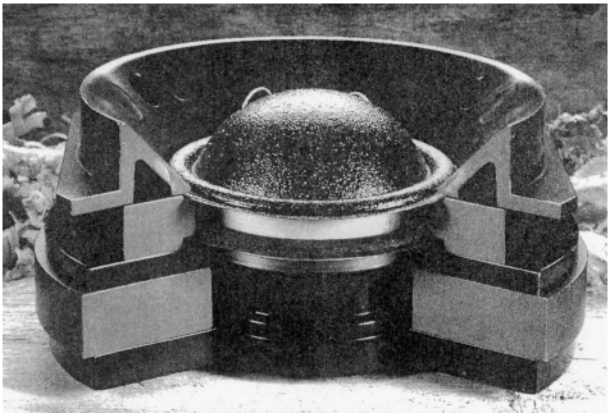


Figure 2.16 A cut-away view of an ATC 3 inch (75 mm) soft dome driver. The British company ATC pioneered the development of this type of high-output mid-range dome

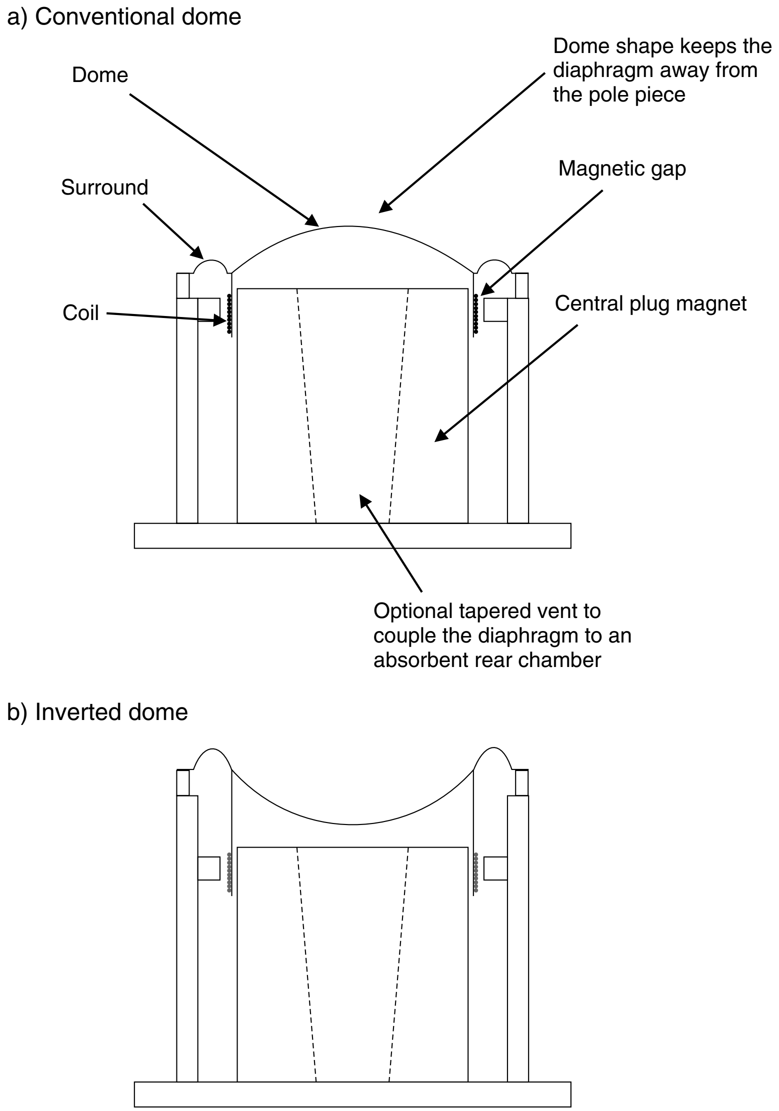


Figure 2.17 A dome as a piston – in principle, if the voice-coil former were to be extended, and the dome inverted (as in b) the sensitivity would drop due to the extra weight, but no material change would take place in terms of the radiation pattern (directivity)

domes with pulsating spheres. As shown in Figure 2.18, a pulsating sphere, which would radiate radially, moves by expanding and contracting in three dimensions, whereas a dome simply radiates as a piston, because it moves in one direction only – along its axis of movement. What is more, when a cone begins to decouple from its voice coil, it does so from the outer parts first. The central part, nearest to the voice coil, always remains under the control of the coil, so the radiating area concentrates towards the centre

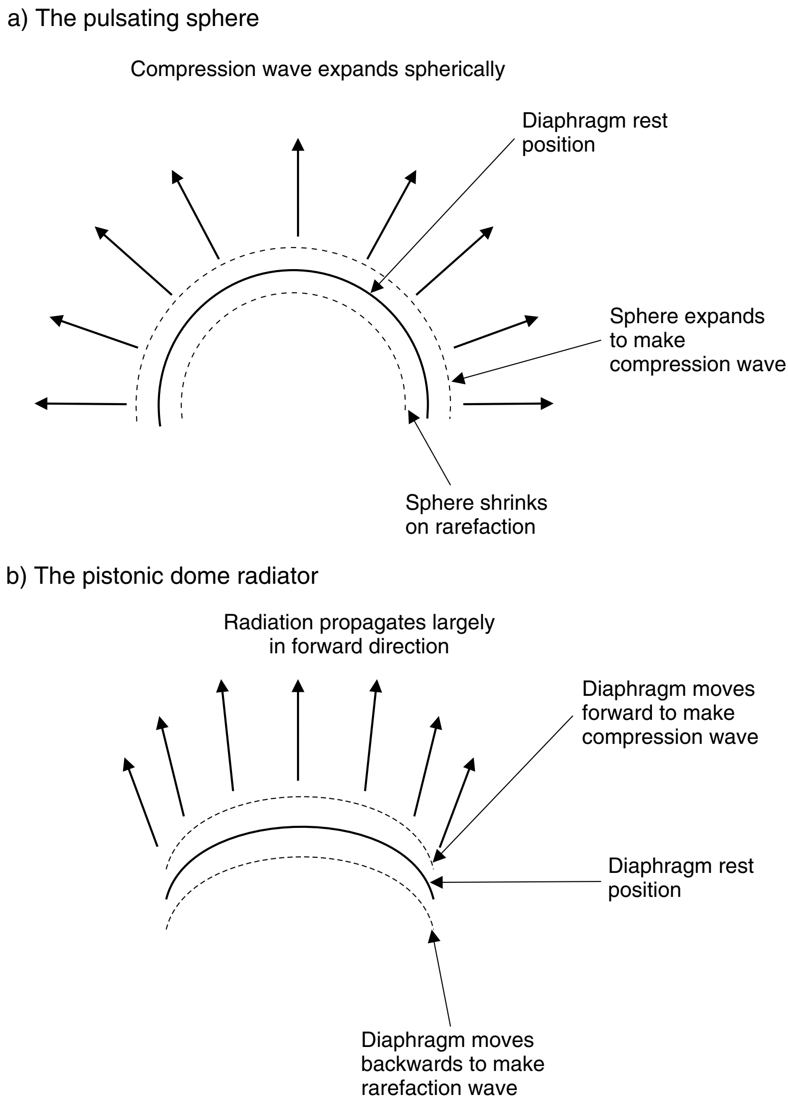


Figure 2.18 Comparison of radiation from pulsating spheres and domes

of the diaphragm as the frequency rises, which is exactly what is needed to maintain its directivity. That is, the radiating area reduces in diameter as the frequency rises, which can be desirable. Conversely, with a dome, it is the *centre* of the radiating area which is furthest from the coil, so as the frequency rises the tendency is for the voice coil to keep control over the *outer perimeter* of the radiating area, whilst the centre of the diaphragm decouples itself, as shown in Figure 2.19. This leads to a ring radiator, which has very peculiar directivity properties, so domes, if not applied

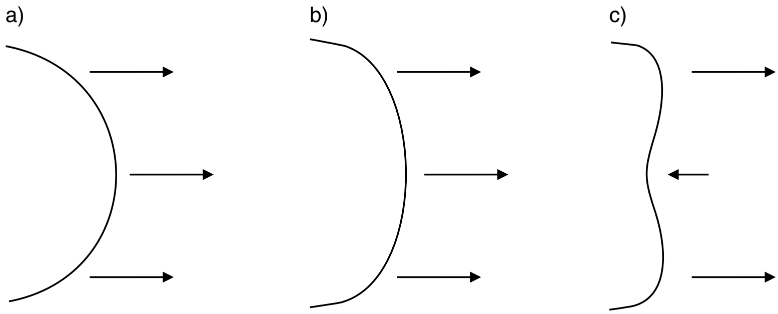


Figure 2.19 A dome in break-up. a) Radiation equal in phase and amplitude from all parts of the dome. b) First bending begins. The radiation amplitude increased from the dome edges and begins to reduce from the centre. c) The dome breaks up. Radiation from the centre becomes out of phase with the edge radiation

very carefully and below their break-up frequencies can actually have less smooth directivity than cones.

2.2.1 Hard and soft domes

The diaphragms of hard domes are usually either made from phenolic-resin impregnated cloth, aluminium, titanium, beryllium, carbon fibre or other, similar materials with very high strength to weight ratios. Soft domes are typically made from moulded cloth which has been treated with a viscous damping material, usually a synthetic rubber. Other types of domes are also sometimes used, as shown in Figure 8.8. Here, a rigid 7 inch (175 mm) dome of polyurethane is used in the lower mid-frequency range. Hard domes are usually restricted to use at high frequencies, above 3 or 4 kHz, because their resonances are difficult to control in diaphragms of sufficient diameter to radiate useful power at much lower frequencies. In rare cases, hard domes of 3 inch (75 cm) or more can be found operating down to frequencies as low as 800 Hz, but in order to maintain piston action up to very high frequencies before the first break-up modes occur, materials such as titanium or beryllium need to be employed. As beryllium is difficult to work with and its vapour is highly toxic, the production of diaphragms out of this metal is an expensive process. Such units have found favour in the design of domestic high-fidelity loudspeakers, but are rarely to be seen in studios. At higher SPL, when they *do* break up into separately radiating sections, they tend to do so suddenly and in a sonically most unpleasant manner, and produce non-linear distortion products quite differently from soft domes.

Soft domes, on the other hand, can be used down to around 400 Hz, but many exhibit a hysteresis type of response, as discussed in Section 2.1.3 dealing with suspensions, and the same comments apply. The lagging response shown in Figure 2.14 can tend to mask low level detail. The ‘rigid’ domes have been developed partly in response to this problem.

Because of their nature, domes only have an outer suspension. The surround may be formed from the material of the dome in the case of soft diaphragms, but hard domes often employ a separate, bonded suspension material. Rocking motion can be a problem in some designs, and solutions to remedy this are not always practical if they add weight to the moving system, and hence lower the sensitivity, because they may introduce other problems as a result. Lower sensitivity, for example, means more heat in the coil for the same SPL, and can lead to various problems. Ferrofluids can be beneficial in some cases by damping the rocking modes. On rare occasions, double outer suspensions are used. A cross-section of such a construction is shown in Figure 2.20. The choice of material for surrounds can be quite an arduous task in the mid frequencies because the ideal damping properties and compliance (the reciprocal of stiffness) may be conflicting for frequencies two or three octaves apart, yet which must all be radiated by the same diaphragm. At low mid frequencies, where the diaphragm excursion may still be considerable, a cone driver may need to be in a separate enclosure with at least half a litre of air. Dome diaphragms are no different, but the magnetic assembly is an obstruction to the air that is trapped behind the diaphragm, and which tends to push up the resonant frequency. The hollow cavity between the diaphragm and the magnet centre pole, clearly visible in Figure 2.20, is often filled with absorbent material to reduce the cavity resonances. However, if this cavity is too small, it can create problems by way of excessive back-pressure on the diaphragm. Relieving the back pressure may require quite complex boring of the magnetic system if resonances in the tubes and cavities are to be kept out of the working range of the drivers. At high frequencies, these problems are less complex because low resonance frequencies are not required, so neither are such large air cavities required.

Dome tweeters have become very widespread in use, and now probably account for the majority of high frequency drivers. Composite diaphragms are also not uncommon, such as polyester bonded to PVC. Unfortunately, dome tweeters tend to be rather low in sensitivity. This in some ways is ironic, because one of the driving forces behind the development of domes was to overcome the thermal failure problems due to the very small coil surface area in small cone loudspeakers when used at high SPLs. The lower

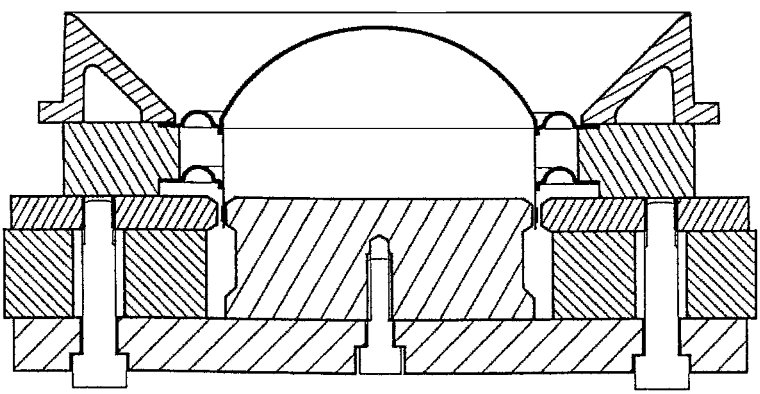


Figure 2.20 Sectional view of an ATC soft-dome driver showing the double suspension which helps to eliminate rocking motion

sensitivity of an equivalent dome driver needs more power to drive it, and hence produces more heat. Nevertheless, in many cases, the balance is still in favour of the dome.

2.3 Compression drivers

Compression drivers are almost always used with horns, except in some rare cases where their internal throats are sufficient to act as horns for very high frequency use. Essentially, the diaphragm, coil and surround assemblies of compression drivers are rather similar to those of dome drivers. The principle difference lies in the way in which the diaphragm couples to the outside air. In the case of the compression driver, the acoustic radiation passes through a restricted aperture, the ratio of its area to the area of the diaphragm being known as the compression ratio. This puts a highly resistive air load on the diaphragm, which is then passed through a horn of roughly exponential flare rate which prevents the air from moving out of the way of the gradually expanding radiated sound wave. Figure 2.21 shows the cross-section of a typical compression driver, and due to the very resistive load, electro-acoustic efficiencies of up to 50% can be achieved. This means that for every watt of electrical input, only half a watt will be dissipated as heat, and the other half a watt will be radiated as sound. It is thus not unusual for mid and high frequency compression driver/horn combinations to reach sensitivities of over 110 dB SPL for one watt input at one metre distance. Domes, on the other hand, can rarely convert more than 5% of the electrical input into radiated sound, so the other 95% serves only to heat up the voice coil.

Horns are often shunned by many people who have not heard the best examples. Much of this negative attitude has arisen from the days when studio loudspeakers using compression drivers and horns were virtual transplants from the world of sound reinforcement and public address, and once a bad reputation sticks it can be very difficult to lose. To far too many people, because they have heard some bad horns, all horns must be bad. This is the absolute opposite to the general perception of soft dome mid-range drivers, where many people have heard some excellent ones and thus think that they all must be excellent. It is sometimes difficult to understand human reactions to these situations. In neither case is the point of view either logical or correct. It is also worth noting how so many people who state that they do not like horn loudspeakers in the mid range will also say how they like the classic Tannoy Dual Concentric loudspeakers (shown cross-section in Figure 8.7) which, above 1 kHz, are exactly no more and no less than compression drivers and horns.

One problem which does always plague compression drivers is that the sound pressure levels within their throats can reach levels where the air itself cannot linearly propagate sound waves. Air at these SPLs does not compress and rarefy to the same degree under the same applied force in each direction, and so gives rise to harmonic distortion. However, at recording studio SPLs, which are way below live sound and cinema SPLs because of the much closer listening distances, air overload often does not become a problem until levels where direct radiating loudspeakers are

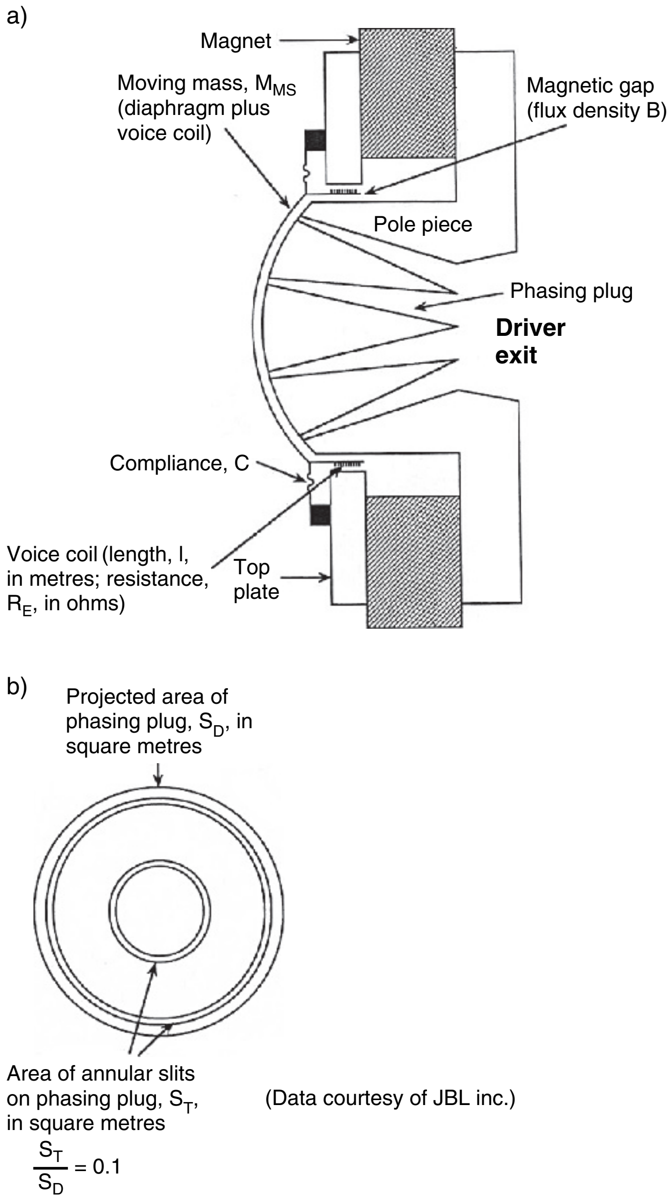


Figure 2.21 A high frequency compression driver (Data courtesy of JBL Inc). a) Section view of a JBL high-frequency compression driver. b) Plan view of the diaphragm side of the phasing plug

suffering from mechanical and thermal non-linearities of their own. When a compression driver made from 5 kg of metal is dissipating half a watt of heat and producing 100 dB SPL for the people behind the mixing console, thermal problems absolutely do not exist, and neither do mechanical stress problems because the diaphragms are moving over such short distances. Non-linear distortions can remain remarkably low, and transient attack can be second to none. The problems with compression driver/horn combinations usually arise when designers fail to respect the physical realities of how to couple the horn to the air, but that will be discussed further in Chapter 4. High sensitivity loudspeakers, in general, also enjoy another benefit in that the lower current in the voice coil for any given output SPL gives rise to much less disturbance of the static magnetic field in the gap, and thus avoid some intermodulation distortion products that are largely unavoidable in less sensitive drive systems when passing high currents through their voice coils.

Unfortunately, good compression drivers are not cheap to manufacture, because they require precision, low tolerance engineering. It is therefore futile judging compression drivers in general by listening to cheap examples. Tolerances of less than 50 micrometres are not unusual in manufacturing specifications, and the magnetic flux density required in the gap can be so high that special materials may be needed which are difficult to cut and need heat treatment afterwards. Diaphragms also need to be made to very high standards of uniformity whilst often only being about 50 micrometres in thickness. Some of the finest diaphragms are made out of beryllium, because of its enormous strength-to-weight ratio, but its melting point of 1600 degrees C is too high to be accurately moulded in any practicable manner, and rigidity is such that it would shatter like glass if stamped to shape in a press. Some diaphragms are therefore made by a time-consuming and laborious in-vacuo vapour deposition process which is definitely not suited to mass production and is obviously expensive, but 2 inch (50 mm) diameter diaphragms can be made in this manner weighing less than 0.15 grammes, and with frequency responses from 500 Hz to well over 20 kHz.

In order to avoid phase cancellations at high frequencies in compression drivers, a phasing plug is usually incorporated which guides the pressure from different parts of the diaphragm down a series of tubes or concentric slits (as can be seen in Figure 2.21), in order to bring a phase-coherent wavefront to the driver exit, even at the highest frequencies of use. Alternatively, for very high frequency horns used largely only above about 5 kHz, ring diaphragms may be used, as shown in Figure 2.22. These drivers also usually incorporate a short exponential horn as a part of the driver itself, and hence require no external horn. The diaphragms are clamped at the outer and inner edges, and radiate into a ring-shaped aperture which gradually flares into a single exit by means of some sort of central 'nose'. The sensitivities of these devices range up to about 108 dB SPL for 1 watt input at one metre distance. With a typical power handling capacity of 40 watts, they are virtually indestructible in recording studio or domestic playback use when operating only above about 7 kHz.

Compression drivers tend to work best from about 500 Hz upwards. Below that frequency they *can* be used, but the horns required to couple

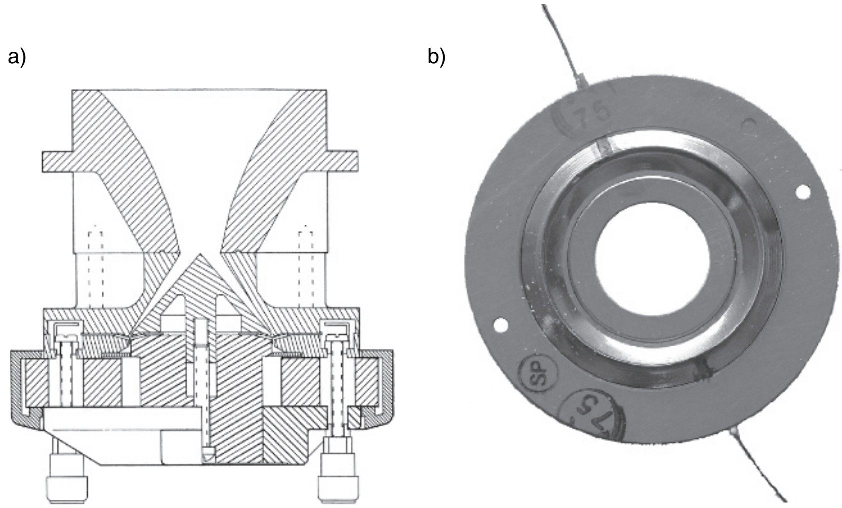


Figure 2.22 Ring diaphragm drivers. a) Section view of a JBL high-frequency compression driver employing a centrally and peripherally clamped ring diaphragm (Drawing courtesy of JBL Inc). b) Photograph of a ring diaphragm

them optimally to the outside air tend to become impractically large. The horn shown in Figure 8.2(a) uses a 4 inch (100 mm) diaphragm driving into a 2 inch (50 mm) throat, and is used from around 300 Hz to 20 kHz. This is an exceptionally wide frequency range, although stress loads on the diaphragms can be high at high SPLs. The designer, Shozo Kinoshita, cites smooth directivity and lack of crossover points in the sensitive range of hearing to be great benefits. Certainly the subjective impression from the loudspeakers is one of great low level detail and a rapid transient response which is effortless even at high levels.

One of the principal differences between the design of compression drivers and dome loudspeakers is that a soft diaphragm is not an option for compression devices. The diaphragms must be light and they must be rigid if high sensitivity and low distortion are required. The rear cavities in compression drivers are small, and the space between the diaphragm and the phasing plug is so small that any flexing of the diaphragm would be likely to make contact with the plug. Leaving more space between the diaphragm and the phasing plug would lead to a loss of sensitivity at high frequencies.

More will be said about horn/driver applications in Chapter 4, because many of the relevant points relate more to the horns than they do to the driver.

2.4 Ribbon loudspeakers

The origin of the ribbon loudspeaker actually predates the moving coil cone loudspeaker, with Schottky and Gerlack filing their patent two years before the moving coil loudspeaker patent application. However, in practice it was rather disastrous, with a response of about two octaves between

250 Hz and 1000 Hz. Eight years later, Olson and Massa made use of the concept when they reversed it and turned it into a ribbon *microphone*. Nevertheless, the concept of a ‘current sheet’ had emerged, with the diaphragm being suspended between the extended poles of a magnet system, and with the diaphragm itself passing the current. The idea is shown graphically in Figure 2.23. Stanley Kelly patented a much superior device in the 1950s, and, in his own words; “The ideal radiator is one which

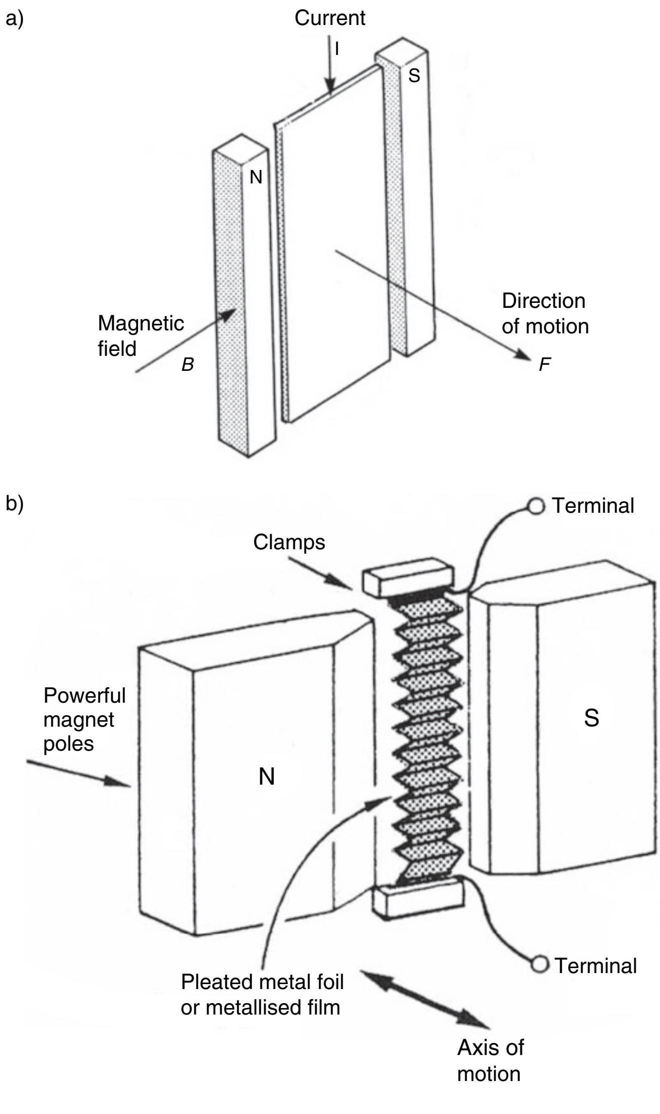


Figure 2.23 The ribbon driver. a) The basic concept of a ribbon driver. The current flowing through the diaphragm reacts with the static magnetic field between the north (N) and south (S) poles, and gives rise to a force, and hence a movement of the diaphragm in the direction shown. b) A more practical realisation

a) vibrates in phase over its whole surface, b) has a mass comparable to the air load, and c) has only resonances which are outside the working frequency band. In order to meet these requirements, the radiator must be subject to a mechanical force equal in amplitude and phase over its whole surface. There are only two commercial systems which meet this requirement, viz, the constant-charge electrostatic and the ribbon electromagnetic loudspeakers".[3]

In practice, the ribbon is corrugated to give it more rigidity. In order to avoid giving rise to resonances in its frequency band of operation it is supported at each end but it is not stretched. The support of the ribbon in the gaps is normally by means of elastomers and silicone rubber. As a current flows through the diaphragm, the corresponding magnetic field interacts with the static magnetic field from the magnets parallel to the plane of the ribbon. This generates a force at right angles to the plane of both the magnetic field and the ribbon, which moves the ribbon back and forth, but parallel to the direction of the magnets. To maintain adequate efficiency, the mass and electrical resistance must be kept as low as possible, the latter being typically as low as 0.2 ohms and requiring an impedance matching transformer in order to be useable with normal amplifiers. The low impedances also imply high currents, so a conflict can arise between the thickness of the diaphragm being low enough to keep the weight down but high enough to keep the electrical resistance down, otherwise the diaphragm might melt with the signal current. The classic Decca/Kelly 'London' ribbon loudspeaker had a horn attached to it, had a sensitivity of 92 dB SPL for 1 watt at 1 metre distance, and covered a frequency range from 1 to 30 kHz with a power handling of 25 watts. Sonically, it was widely appreciated.

Modern technical advances have led to printed circuit current sheets, with the copper tracks on polyimide sheets of around 12 micrometres thickness. The sheets can be made with overlapping copper tracks on each surface, thus allowing the whole sheet area to be conductive. In conventional ribbons, however, the diaphragm material is usually aluminium, because it has the best compromise between resistance and mass. In recent years, the American company SLS Loudspeakers has made a big feature of the use of ribbon loudspeakers beyond about 2 kHz. Ribbons, traditionally have been delicate, and difficult to manufacture, but sonically they have always had many friends.

2.5 Heil air-motion transformers

These devices often get mistaken for ribbon loudspeakers when people see the folded diaphragms set in short horns, but they are definitely *not* ribbon loudspeakers. A ribbon radiates sound by the whole diaphragm moving backwards and forwards in a uniform manner, with the pleats never changing their angles. The air-motion transformer, quite differently, moves its diaphragm in a concertina movement, drawing the air into the folds as they expand, and expelling the air as they contract. The German company A.D.A.M Audio have recently featured this technology rather in the way that SLS have made big use of ribbons. The air motion transformer was designed by Dr Oskar Heil and its general outline is shown in Figure 2.24. The current flows through a flat conducting track which is bonded to the diaphragm, and which is folded such that the conductive strip lies parallel to itself on the adjacent fold. When the current flows in one direction through the entire circuit, it travels in different directions in the conductors on adjacent folds, and the magnetic field is either attracted to or repelled from the nearby permanent magnets. When the current in the circuit reverses, the folds which were opened are then closed, and vice versa. It is called an air motion transformer because there is a ratio of about four to one between the air particle velocity in and out of the folds relative to the velocity of movement of the pleated diaphragm. The magnet structures need to be very large because the entire pleated diaphragm must fit in the gap between the poles. The diaphragms are made from plastics such as p.t.f.e. or polyethylene, which have good damping. Current units can work from about 500 Hz to 20 kHz. Some models have been shown to produce quite high levels of second harmonic distortion above about 5 kHz, but the subjective audibility of this does not seem to be significant as the distortion products are all above 10 kHz and about 30 dB down relative to the signal.

2.6 Distributed mode loudspeakers

These are the flat panels developed under licences from NXT and its subsidiary New Transducers Ltd in the UK. The principal patent is held by the British Ministry of Defence, on whose behalf Dr Ken Heron was not actually trying to develop a loudspeaker at all – he was trying to build lighter helicopters and stumbled upon an aluminium honeycomb panel that radiated sound quite efficiently.

At first glance, the concept of a distributed mode loudspeaker (DML) seems to be a total contradiction. It is a mess of resonances, when, in almost all other aspects of loudspeaker design diaphragm, resonances are taboo. The drive points where the electromagnetic exciters couple to the panel are chosen so that they couple to as many of the vibrational modes as possible, in order to excite the panel in the most uniform manner. The panels which are currently used are typically made from resin impregnated glass-fibre sheets bonded to a 3 mm honeycomb core of 'Nomex' - a polyamide which is often used to make loudspeaker voice-coil formers. The panels are not driven by a voice coil attached to the panel which is then connected to the same frame/ chassis as the magnet, but rather the lightweight panel and coil react against the mass of a much heavier, freely suspended magnet. Panels are commonly excited by two or four coils, to more evenly distribute the drive force. An example of a commercial panel is shown in Figure 2.25.

As the name 'distributed mode loudspeaker' suggests, they radiate all frequencies from all parts of the panel, and so are naturally diffuse sources. As such, they can work well in both studios and domestic circumstances as the rear channels of a surround system, where they can create excellent ambient diffusion effects. Their sonic colouration, though not unpleasant, tends to render them inappropriate as main front channels where serious listening is the goal, but their performance on the surround channels can

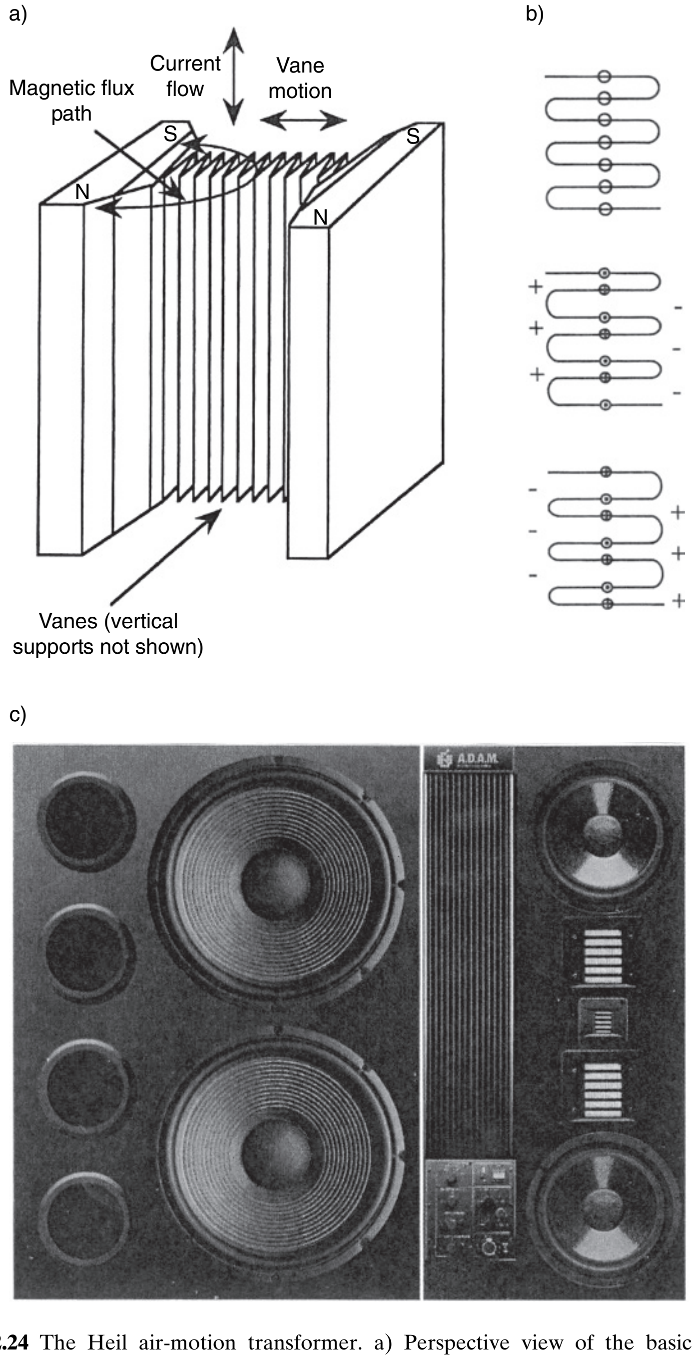
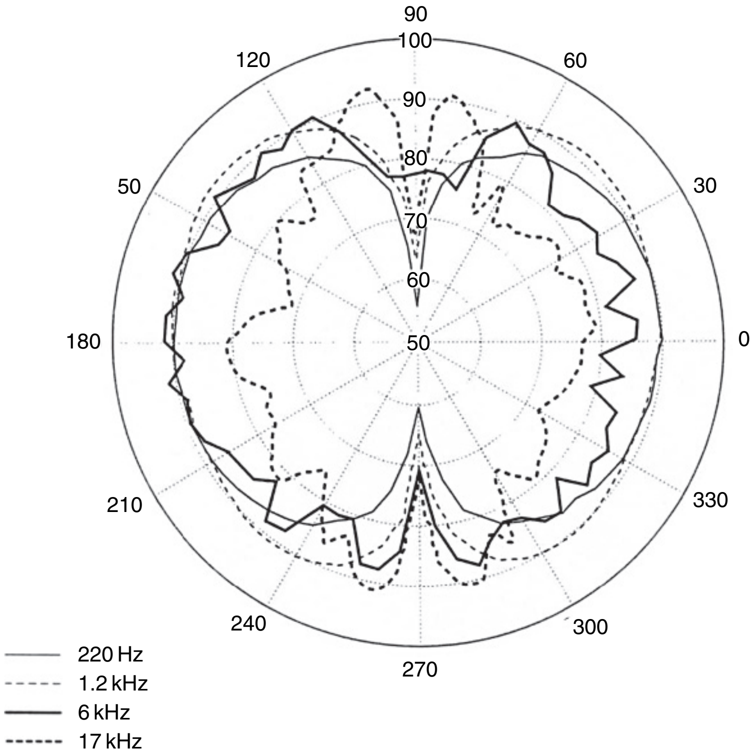


Figure 2.24 The Heil air-motion transformer. a) Perspective view of the basic concept. b) Polarity changes in the conductors cause the opening and closing of alternating folds, drawing in and expelling air on alternate half-cycles. c) A full-range loudspeaker system using Heil air-motion transformers for the mid and high frequencies in a symmetrical, D'Appolito layout – an A.D.A.M. S7

Figure 2.25 A distributed-mode loudspeaker (DML)

be very involving, and numerous professional installations have used them in this role. It is also interesting to note that the radiation from both sides of the panel (if the rear is not enclosed) couples to the room neither in an omnidirectional way nor as a figure-of-eight pattern (or dipole) like electrostatic panel loudspeakers, but as something more akin to a bi-pole, where the radiation from each side is only partially correlated. As such, and if spaced away from a reflective wall, they can fill a room with reflexions which have very little tendency towards showing typical summation and cancellation effects (peaks and dips in the response) in different places in the room. This characteristic can add still further to the beneficial way that they can be used for ambience channels, where their inherent sonic colouration appears to be little disadvantage.

A constant problem for DML loudspeakers has been the lack of low frequency output, but designs are now emerging, such as the Fane Minipro, which extend reasonably well down to 60 Hz. For surround use this is often quite adequate, especially since many of these systems will be used with bass management systems which will pass the low frequencies to a separate (sub-) woofer. Panels with dimensions of about 40 cm × 60 cm are sufficient for such purposes, but smaller panels suffer from much higher roll-off frequencies. The low frequency response can be extended by mounting the panels on a shallow, absorbent-lined box, of about 10 cm depth, but care must be taken to avoid over-filling the box with absorbent material or the damping on the rear of the panel can become too great. When the open-backed panels are hung against walls they should be hung at an angle, in order to prevent resonances in the parallel cavity formed between the panels and the wall. As the general radiation directivity is very broad at all frequencies, as shown in Figure 2.26(a), it is not necessary to point the panel at a normal to the listening area. From almost whatever angle a panel is



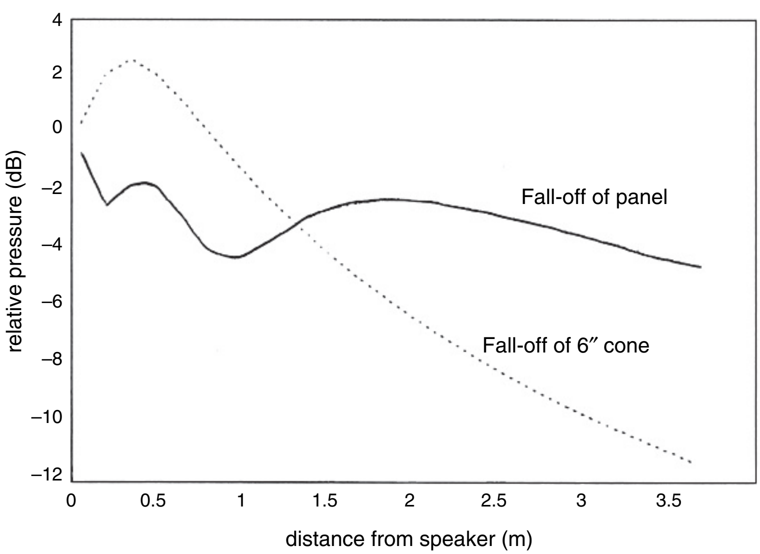


Figure 2.26 Radiation characteristics from a typical DML. a) DML polar response at 220 Hz, 1.2 kHz, 6 kHz and 17 kHz. b) Computed comparison of loudness with distance – a distributed mode panel versus a 6 inch (150 mm) piston. (From data published by NXT)

radiating, and no matter where it is in a room, it will excite the whole room with its full frequency range. Only from a position in line with the panel edge is there a region of a few degrees where low frequency cancellation takes place. The high frequencies radiate in an almost omnidirectional manner, though from a spacially diffuse source.

As a result of the diffuse nature of the radiation, the fall-off of SPL with distance is initially less than from a conventional loudspeaker, as shown in Figure 2.26(b). It is more typical of larger, planar radiators, as discussed in Sections 2.8 and 2.9. This can also be a benefit when filling a room with ambient sounds, because the left and right signals, from positions more or less laterally alongside the listeners, are more evenly distributed across the room, even when the room acoustics are relatively dead. The coupling to the room modes (see Chapter 7 if necessary) is also accomplished in a different manner to the way in which conventional radiators couple with modes. This again has been shown to give rise to fewer peaks and dips in the room. When all the characteristics are taken together, the DMLs do offer some interesting opportunities for surround applications. In fact, the low frequency response can be noticeably better if panels larger than those referred to above are used, but very large flat panels tend to become unwieldy and can introduce resonance and reflexion problems into the room when used in conjunction with other sources.

2.6.1 Panel/piston combinations

Since the mid 1990s, various efforts have been made to develop combinations of DMLs and conventional loudspeakers in such manners as to take advantage of their uncorrelated and correlated radiation characteristics respectively. The thinking behind these ideas is that in live music situations the direct propagation from the instruments to the ears tends to be highly correlated, whilst the reflected energy from walls, ceilings, floors and other hard surfaces tends towards being uncorrelated. In good concert halls it has been found that the ones with low levels of inter-aural cross-correlation tend to produce the generally most desirable sensations of spaciousness. In domestic situations, the reflexions from walls do not tend to be as diffuse and uncorrelated as in good concert halls, and their frequency response is inevitably affected by the directivity characteristics of the loudspeakers. To help to combat these deficiencies, the concept of creating a sensation of diffuse reflexions has been pursued by the use of relatively diffuse sources in combination with a conventional stereo pair of loudspeakers. The relative level of the two types of sources can be adjusted to taste.

The KEF company in the UK has patented a concept of using DMLs behind conventional loudspeakers with the axes of the DMLs at right angles to the axes of the conventional loudspeakers. The conventional loudspeakers generally point towards the listeners, whilst the DMLs work more omnidirectionally (although with their weak lateral nulls towards the listening position) in an attempt to excite the rooms with diffuse reflexions from the surround channel(s) of multichannel recordings.

Another system, marketed under the trade name of Layered Sound, was patented by Dr Shelley Katz, a Canadian pianist, who initially researched the concept as a means to make electric pianos sound less 'stiff' and more acoustic. This technology was licensed for research purposes to the Japanese company Korg. In the domestic reproduction or sound reinforcement formats, the panels are placed closely, above or behind the conventional loudspeakers, but usually with the axes parallel to each other, and not necessarily at 90 degrees as with KEF systems. The panels in this system are fed with the same signal as the conventional loudspeakers, and can also fed via a delay, with the delay time and the relative SPLs from the different sources being used to control the overall effect.

By definition, such systems are not high-fidelity in the classical sense, because they seek to introduce artefacts which are not in the original drive signals. Nevertheless, that overall sense of realism which they can help to generate may be considered in many cases to be highly faithful to the *sensations* of the performance spaces or the wishes of the recording personnel or musicians. Proponents claim that as the current recording processes via conventional microphones and the reproduction via conventional loudspeakers are still limited and compromised by the inherent short-falls of their performance, then piston/panel combinations may be able to realistically add, globally, more than is lost in conventional reproduction, and so can be considered to be more than making up for those short-falls. However, all assessments of these types of loudspeaker systems currently need to be made subjectively, because there are still no measurement systems which can reliably define the performances of such combined systems in any meaningful manner.

Therefore, whilst technical accuracy in terms of conventional reproduction might be compromised, the developers of such combined systems can reasonably claim that the perceptual fidelity, in terms of overall realism, may be superior, at least on certain types of music, to reproduction on systems with more measurable fidelity in the conventional sense. Whilst it would seem to be unwise to use the combined systems for music recording quality monitoring, the beneficial effects of Layered Sound for electric piano amplification seems to be established. Of course, if recordings were being specially made for reproduction on these composite systems or their derivatives, then the monitoring of the recordings via such systems may also be justified. Development of DMLs is still in progress, however, and some new designs are already emerging with much less coloured responses than have previously been achievable.

2.7 Beyond magnetics

All of the loudspeaker drive systems discussed in this chapter so far have been electromagnetic transducers. One way or another, all of them have employed the magnetic field generated by an alternating music signal current in a moveable conductor to react with a static magnetic field. The force generated at right angles to the current and the static magnetic field has then been applied to a diaphragm of some sort or other which has been designed to move air and radiate sound. They are all, basically, variations on the same theme. There are, however, various other means by which loudspeakers can be made.

2.7.1 Piezoelectric devices

There are certain materials which can be made to twist and bend when electrical signals are applied to opposing surfaces, and, in general, there is a useful proportional relationship between the applied voltage and the degree of movement of the material. Such transducers have found use as high frequency loudspeakers of very robust design, which have in turn found use in guitar amplifiers and sound reinforcement systems where conventional tweeters have been considered to be too fragile. Quartz, Rochelle salt and some ceramics such as barium titanate have piezoelectric properties, as do some high polymer plastics, such as polyvinylidene fluoride. Direct radiator and horn loaded piezoelectric radiators are available, the horn-loaded Motorola device being quite widespread and has an axial response within ±3 dB from 4 kHz to 20 kHz. Pioneer have also developed a cylindrical piezo radiator. In these, a thin film of high polymer plastic is made into a cylindrical shape which is then caused to pulsate with the applied signal voltage. The response is respectably flat from 2 kHz to 20 kHz, with 360° horizontal radiation.

The piezoelectric units are rugged largely because they are self-protecting. The impedance tends to rise as the frequency lowers, so driving low frequency signals through them is difficult, even with no crossover. They also, effectively, have nothing to burn out and nothing to go off-centre. Although they are very rarely encountered in loudspeakers used for music monitoring, they can be found in domestic system as well as in music amplification systems. The principles of their construction are shown in Figure 2.27. Piezoelectric drivers tend to be mid-sensitivity devices, offering the low 90s of decibels SPL for one watt (or at least its voltage equivalent – 2.83 volts – into their varying impedance) at one metre distance.

2.7.2 Ionic loudspeakers

It is unlikely that anybody will find these in use today because production ceased around 1968, but the concept is interesting. Radio frequency interference and the production of irritating ozone were unwanted side-effects that helped towards their demise, along with low output, but there is widespread agreement that the sound of these devices was true high fidelity. In ionic drivers a 27 MHz high voltage signal is fed to the electrodes of a quartz cell. A corona discharge results, giving off a blue light as the air is ionised. When the radio frequency voltage is modulated by an applied audio frequency signal the volume of ionised air will vary and produce pressure fluctuations in the air. The frequency response was given as 3 kHz to 50 kHz ± 2 dB, with only 0.5% distortion at 93 dB SPL. Absolute peak output was around 98 dB SPL for the 'Ionofane' version. This was the cutting edge of high fidelity in the 1950s and '60s, and is still good today, however, at higher SPLs, above about 96 dB at 1 metre, compression soon set in, seriously limiting the output. They also cost 28 guineas each (just under 30 pounds, sterling) in the early 1960s, which was an entire month's salary for many working people in those days.

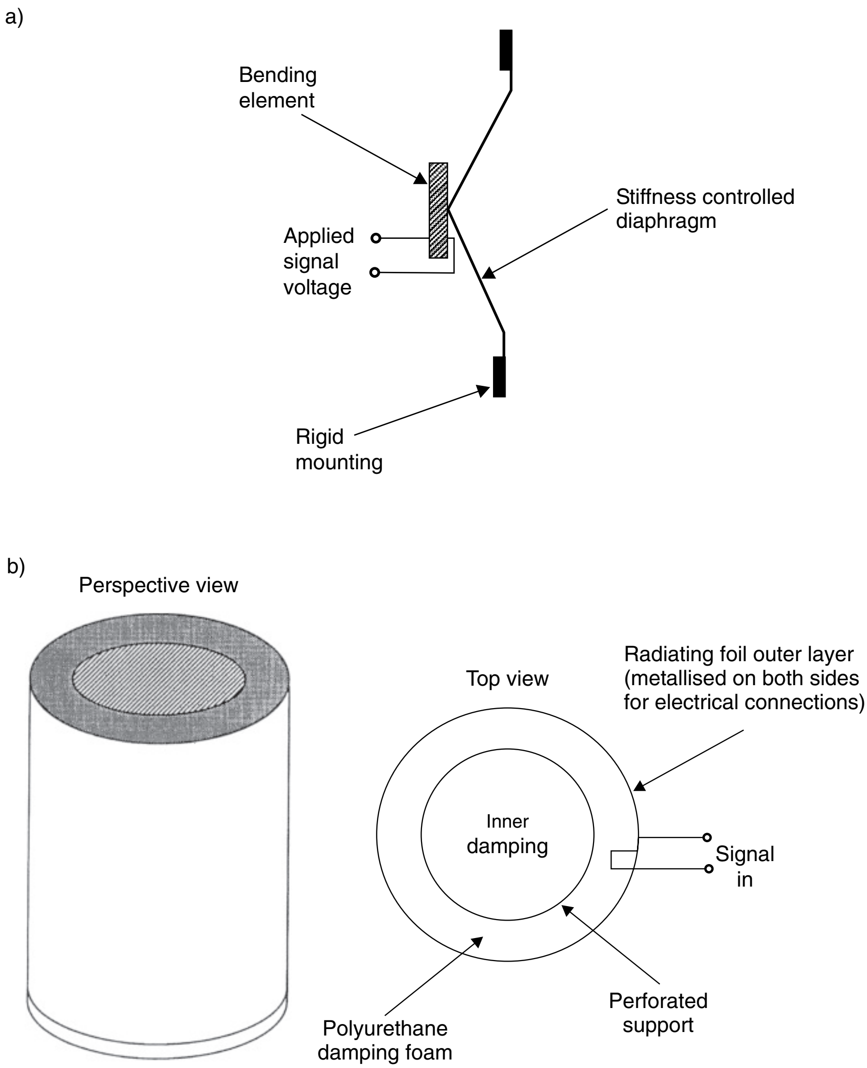


Figure 2.27 Piezo-electric radiators. a) Section view of a typical piezo-electric HF radiator. b) The Pioneer cylindrical piezo-electric driver – the High Polymer Radiator

2.8 Electrostatic loudspeakers

Just as dynamic microphones, such as ribbons and moving coils, have their equivalent loudspeakers, and as piezoelectric loudspeakers relate to crystal microphones, electrostatic loudspeakers are the counterparts to condenser (capacitor) microphones. [Condenser is the old term, capacitor is the modern term, but the old terms often take root.] Within their limitations, electrostatic loudspeakers can produce a sound quality which, in microphone terms, we would only associated with the finest condenser

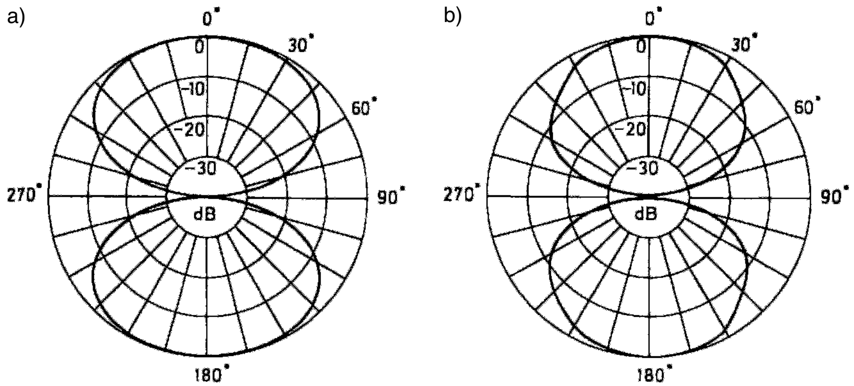


Figure 2.28 Typical directivity pattern from a dipole radiator. a) Low to mid frequency response. b) High frequency response

microphones. Sonically they can be astounding. The limitations are size, (because they need large surface areas of diaphragm), relatively low maximum output SPL, and their figure-of-eight radiation pattern which does not suit all room layouts. The typical radiation pattern from a full-range electrostatic loudspeaker (ESL) is shown in Figure 2.28. The absorption of the rear radiation is not a practicable solution unless the box is enormous because the air load would inhibit the movement of the extremely light diaphragm, whose mass is critically chosen to match that of the free air surrounding it. The loudspeakers therefore need to be placed away from walls, and the nature of the walls close to the sides and behind them need to be duly taken into account, acoustically, when the response in front of the loudspeakers is being considered. Full-range electrostatic loudspeakers may therefore be less forgiving in terms of where they can be placed.

These devices do not act as volume-velocity pumps like cone loudspeakers, but radiate as pressure gradient sources. This means that they do not couple to the *pressure* anti-nodes of the room modes, but to the *velocity* anti-nodes, which are the pressure *nodes* (see Chapter 7). Other than in anechoic chambers, this means that the optimal siting of electrostatic loudspeakers will be different to that for moving coil loudspeakers. The maximum SPL is limited (although 95 dB SPL should be no problem) because at higher SPLs the polarising voltages would need to be so great that highly specialised materials and techniques would need to be used in the manufacture of the devices, and also because the air would reach its own electrical breakdown limits. This is rather similar to the situation in compression drivers, where the air itself begins to be the limiting factor at higher SPLs. Air is not just something that you can do what you want with. It has its own characteristic properties and it can impose its own limits on what it can be made to do.

The basic concept of an ESL is shown in Figure 2.29. A polarising voltage of around 3000 volts is applied to the diaphragm whilst the two, outer, perforated electrodes, are grounded via resistances, and spaced away by about 2 mm. The charge on the diaphragm keeps it centralised, in

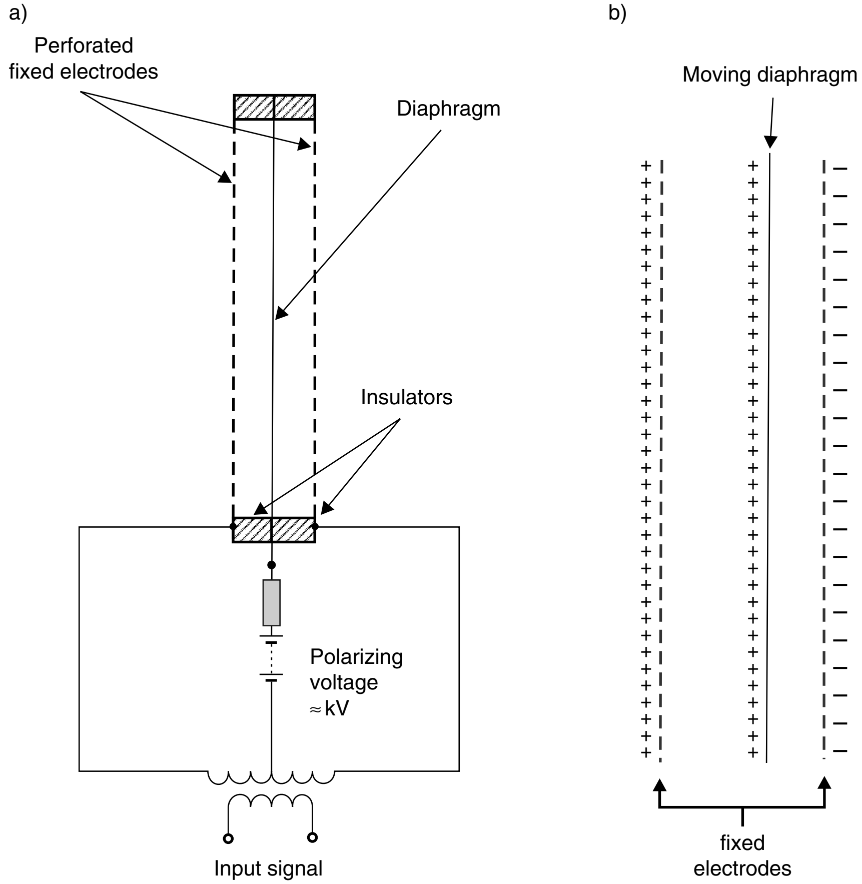


Figure 2.29 The electrostatic radiator. a) The basic principle of operation of an electrostatic loudspeaker. b) When the charge becomes opposite on the fixed electrodes the diaphragm moves to take up a new equilibrium position. If the charge between the fixed electrodes changes, the diaphragm position will change correspondingly

equilibrium, in the absence of signal. If, via an input transformer, a signal voltage is applied across the electrodes, the equilibrium point will shift and the diaphragm will move to chase it. The whole device operates on high voltages and high impedances, so the signals have to be fed from the power amplifiers via a step-up transformer, which itself requires careful design if it is to pass all frequencies equally. Nevertheless, this type of loudspeaker is largely a capacitor, so it still presents a predominantly capacitive load to the amplifier, which therefore needs to be able to supply high currents even when voltages are low, due to the phase angle difference between the voltage and the current, (described previously in Section 1.6). The choice of amplifier may therefore depend on its ability to supply current more than its ability to supply power.

Because of the inevitably small distances between the electrodes, necessary to maintain useful sensitivity with polarising voltages which will not cause air breakdown, the distance that a low frequency diaphragm can travel is severely limited. Therefore, the only way that the diaphragm can move quantities of air sufficient to generate the required SPLs is to be large. However, the diaphragm size of a single unit is limited by its ability to maintain an even tension, and not to sag in places, so this also restricts the SPL achievable by single units.

The large source area would, as with the DML, give a rather irregular directivity at high frequencies, (but without the DMLs ability to produce so many irregularities that they become almost regular again). Full range ESLs therefore tend to be made as two or three-way devices, as with moving coil loudspeakers, using a much smaller radiating area for the higher frequencies. The Quad ESL63 uses a series of concentric rings, as shown in Figure 2.30, in order to mimic a point source situated some way behind the diaphragm. These loudspeakers have also been made available with dipole

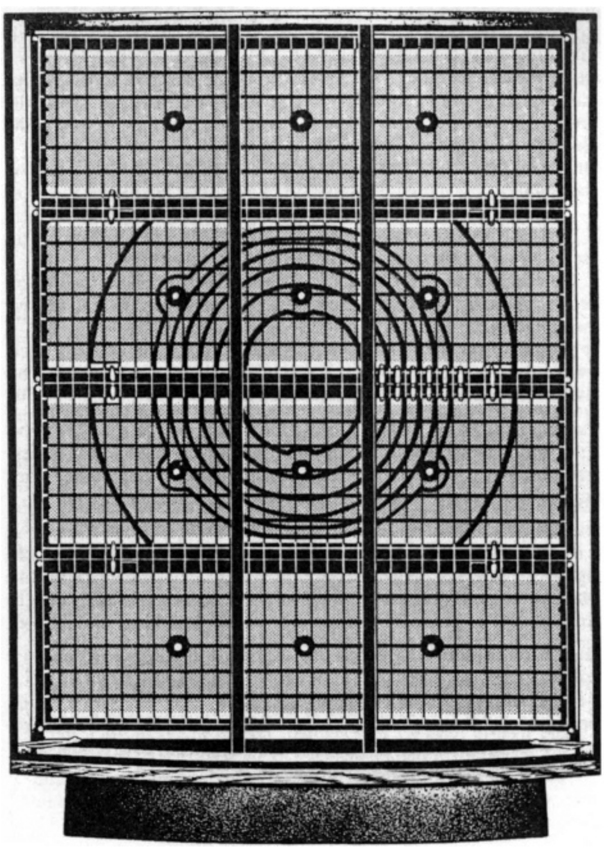


Figure 2.30 The Quad ESL 63 electrostatic loudspeaker with its concentric diaphragms. The higher frequencies radiate only from the central sections

moving-coil sub-woofers mounted beneath them to extend their rather limited low-frequency responses, or rather, their limited low-frequency *output* capability due to the limited maximum excursion and size of the diaphragms. Electrostatics are therefore not the ideal loudspeakers for monitoring a solo bass drum in a large control room, but they do find use in critical listening rooms and audiophile high fidelity applications, where their natural sound and resolution of low-level detail are highly valued. When heard in an anechoic room reproducing recordings of acoustic instruments recorded in the same room, and with the instruments alongside for reference, their ability to mimic the original sound can be quite startling. Granted, the anechoic response is not the be-all and end-all of loudspeaker reproduction, but the general tendency is for anything which can work so well in such circumstances to have a good start when transposed to other circumstances. Figure 2.31 shows the step function response of a Quad Electrostatic Loudspeaker: the attack of the signal is exemplary, and very hard to beat with other loudspeakers. [The step function responses are further discussed in Chapter 9.]

Although the electrostatic loudspeaker principle was experimented with as early as the 1920s, it was not until 1957 that the first really viable design was put into production. It took the advent of the concept of a constant charge and the development of new plastic foils before it could be fully realised. However, once all the pieces of the jig-saw were in place, the progress was remarkable. A pair of ELSs from the 1950s can, even 50 years later, put many of the latest loudspeakers to shame in terms of low colouration, low distortion, transient response, frequency response flatness and, perhaps most of all, perceived sound quality. What Walker and Williamson did when they developed the Quad ESL was to take a step forwards to a degree that has rarely been equalled in the world of sound reproduction, and this is especially so considering all the technical difficulties which they had to overcome.

Whereas the moving coil loudspeaker exhibits non-linearities in its inner suspension, outer suspension, magnetic flux disturbances, magnetic field asymmetries and various other sources, the electrostatics more or less only

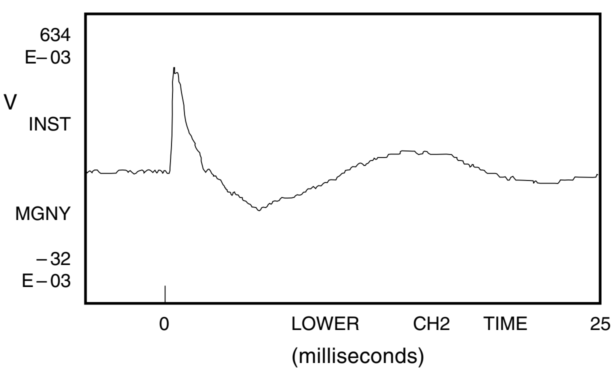


Figure 2.31 Step function response of an electrostatic loudspeaker, showing the exemplary attack (rise time)

exhibit non-linearities due to very small asymmetries in construction, which can be minimised by careful quality control. In general, the non-linear distortion production by electrostatic loudspeakers is much lower than that produced by most moving coil devices. The authors of this book have, for decades, used full-range electrostatic loudspeakers as benchmarks against which to judge other loudspeakers, both objectively and subjectively. This is not to say that they cannot be surpassed on individual aspects of their performance, but their global performance is hard to beat.

Occasionally, electrostatic mid-range drivers and/or high frequency drivers can also be found in compound,electromagnetic/electrostatic designs of domestic loudspeaker systems.

2.9 Electromagnetic planar loudspeakers

As if to rise to the electrostatic challenge, one of the electro-dynamic (electromagnetic) responses was the planar loudspeaker. These use light, thin, tensioned plastic film diaphragms which have voice coil circuits printed on them, rather in the manner of a very thin printed circuit board. The diaphragms are stretched over frames with many openings and a large number of small magnets dispersed over their area. There is no attempt to concentrate the flux in any area, but just to set up a field of fringe flux in the vicinity of the magnets. In this way, a diaphragm is caused to move by the reaction of the signal current in the printed tracks with the static magnetic field, which results in the diaphragm being more or less uniformly driven over its entire surface. As with large electrostatic diaphragms, they must cross over at higher frequencies into drivers of smaller radiating area if strange directivity problems are to be avoided. They do not have the ‘random’ distribution of high frequency sources as exhibited by the DMLs, but neither do they tend to have as much colouration.

2.10 Summary

There are a considerable number of different ways to transform electrical drive signals into sound waves, and no one system has all of the advantages to itself. In fact, all the drive systems are electro-mechanico-acoustic transducers; that is, they must first convert the electrical signals into mechanical forces which are then used to drive sound radiating diaphragms of some sort or another. (The one exception perhaps, being the now defunct ionic tweeter.) The necessary double conversion tends to involve a number of non-linear processes, and it is largely the mechanical components which are the main offenders. It is therefore unfortunate that we often refer to loudspeakers as simply electro-acoustic transducers because it fails to recognise the existence of the principal culprit for our problems.

At the limit, the air itself is non-linear, so when we try to reproduce loud sounds from small sources that were originally produced by large sources, local concentrations of high air pressures close to the small sources will give rise to non-linearities, which our ears will recognise as not being the real thing. The reason for describing this wide range of loudspeaker drive unit concepts so early in the book (and there are other, less common ones) is to establish the point that at the very heart of all loudspeaker systems are imperfect components, and, as mentioned earlier, that the art of the science of the designs is to find the best compromise for any individual requirement.

References

1 Colloms, M., 'High Performance Loudspeakers', 5th Edition, John Wiley & Sons, Chichester, UK (1997)
2 Briggs, G., 'Loudspeakers, The Why and How of Good Reproduction', Fourth Edition, Wharfedale Wireless Works Ltd, Bradford, UK (1955)
3 Borwick, J., 'Loudspeaker and Headphone Handbook' Second Edition, Chapter 2 (By Stanley Kelly), Focal Press, Oxford, UK (1994)

Bibliography

1 Chapter 3 of Reference 3, above, contains what is perhaps the definitive work on electrostatic loudspeakers, written by the late Peter Baxandall. In the Third Edition of the book, published in 2001, Peter Walker somewhat modified the text. Either edition, in its entirety, is recommended reading for anybody wishing to delve deeper into the world of loudspeakers and headphones.
2 The books mentioned in References 1 and 3 above
3 Eargle, J., 'Loudspeaker Handbook', Chapman and Hall, New York, USA and London, UK (1997)
4 Borwick, J., 'Loudspeaker and Headphone Handbook', Third Edition, Focal Press, Oxford, UK (2001)

Chapter 3

Loudspeaker cabinets

3.1 The concept of the infinite baffle

When the diaphragm of an open-framed driver moves forwards, the compression of the air at the face of the diaphragm is accompanied by a rarefaction at the other side of the diaphragm, and the natural tendency is for the pressure difference to equalise itself by a movement of air around the sides of the driver. At frequencies whose wavelengths are large compared to the circumference of the diaphragm, the equalisation is almost perfectly accomplished, and so almost no sound is radiated. It is therefore necessary to discourage this pressure equalisation if low frequencies are to be radiated. The simplest means of accomplishing this is to mount the loudspeaker in a large, rigid board, or baffle, as shown in Figure 3.1. If the board were to extend in all directions to infinity, it would be a true infinite baffle. It would cause no change in the air loading on each side of the diaphragm, it would exhibit no resonances, it could cause no diffraction, and, with a good quality driver (or drivers) would sound excellent. Unfortunately, its great drawback is that it is a rather impractical concept.

The two practical realisations of this idea are the finite baffle, where a baffle of perhaps a metre square is employed, or the *so-called* infinite baffle, which is, in fact, a sealed box. The radiation pattern of the finite baffle is shown in Figure 3.2(a). The cancellation around the sides of the extended plane of the driver cause response nulls to the sides, in the direction of the plane of the baffle, resulting in a three-dimensional figure-of-eight pattern in free space. The low frequency cut-off is determined by the size of the baffle. The final rate of low frequency roll-off is 18 dB per octave, but some measures can affect the nature of the entry to the roll-off. Varying the Q of the driver resonance, by mechanical and/or magnetic changes, can yield response shapes such as those shown in Figure 3.2(b). By placing the driver off-centre, the cut-off can be made more gradual due to the distance from the driver to each edge of the baffle being different. Open baffles are rarely used in recording studio control rooms because of the problems of where to site them and how to control the rear radiation, but they find use in listening rooms and domestic high-fidelity systems. In these instances the baffles can be sited somewhat more flexibly than in an equipment-loaded control room, and the loudspeaker and listener positions can usually be found which give good results. Subjectively, open baffles tend to sound very clean and, not surprisingly, open. They are largely free of resonances, so their time-domain responses are limited only by the drivers and the

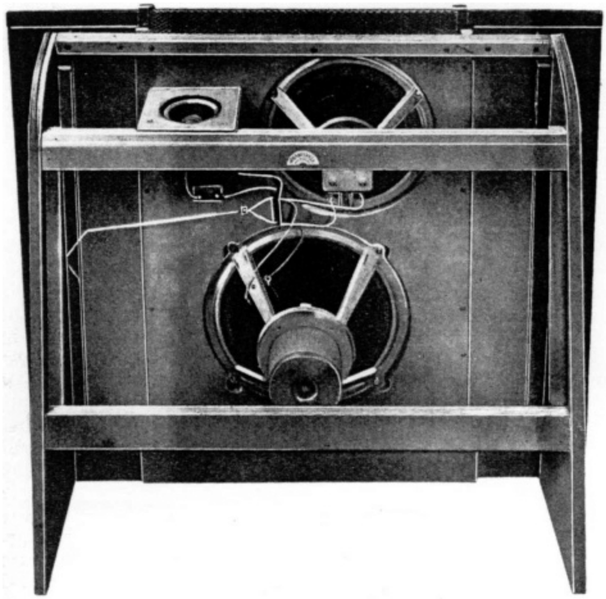


Figure 3.1 An open baffle of Wharfedale design from the 1950s. The front panel was a sand-filled plywood sandwich, to damp resonances. The upward-pointing tweeter was to generate a more diffuse high-frequency response

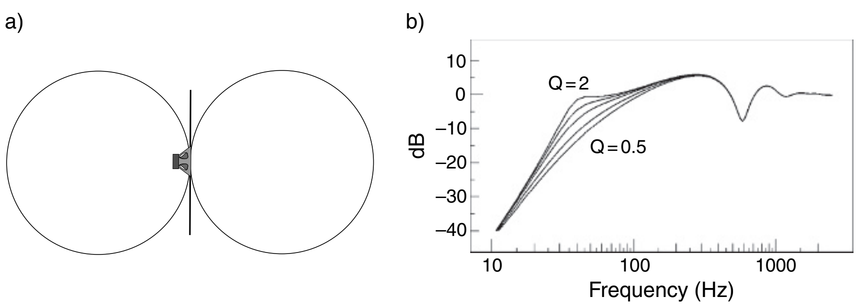


Figure 3.2 Directivity and roll-off of open baffles. a) Radiation pattern polar plot of an open, finite baffle. b) Low frequency response roll-off of an open baffle, showing the effect of the Q (degree of sharpness of resonance) of the driver. The final roll-off tends towards 18 dB per octave below the driver resonance

rooms in which they are placed. The ways in which they couple to the rooms will be discussed in Chapter 7.

When mounted on the floor, the solid surface below the open baffle acts like an acoustic mirror, so a baffle of one square metre placed on the floor behaves like a baffle of two square metres in free space. This enables baffles of practical size to be useful down to frequencies of 40 Hz or below, but the lack of anything other than atmospheric loading on the rear of the diaphragms and poor efficiency of radiation may lead to over-excursion problems with high sound pressure levels at low frequencies. The resonance frequency of the driver on an open baffle will be that of its free-air resonance. Because the open baffle mounting does not push up the free air resonance of the driver, and the back-pressures are not augmented by any constraint of the air behind the diaphragm, lighter moving assemblies may be used. Driver cooling is also something that poses no problem with open baffles, so power compression problems are rarely encountered. The open baffle, in the hi-fi world, still enjoys a devoted following of aficionados.

3.2 The sealed box

The practical realisation of an infinite baffle (the sealed box) is rarely large enough to avoid significantly loading the rear of the loudspeaker, so is best called what it really is, *a sealed box*. Just as open baffles tend to sound 'open', sealed boxes often tend to sound 'boxy'. However, this need not be the case if the box and driver are of adequate size, well-matched, and if sufficient attention is paid to the suppression of resonances within the box. The constraint of the air within a sealed box causes it to act like a spring, which reacts against the movement of the diaphragm in either direction. This effectively stiffens the suspension of the drive unit, and raises its resonant frequency. The smaller the box, the stiffer will be the spring, so the higher will be the resonant frequency of the driver/box system. As the system resonance defines the frequency at which the low frequency roll-off will begin, then for any given driver the low frequency response will become progressively more curtailed as the box size reduces. The only way to counteract this tendency is to use drivers of lower free-air resonance frequency, which means using a driver with a heavier moving assembly or, to a lesser extent, a more compliant suspension, but both of these characteristics have their drawbacks.

A heavier cone takes more energy to move it, so it will need more amplifier power to drive it to produce the same SPL as a lighter cone. A more compliant suspension will be much less rugged than a stiffer suspension, and will tend to be much more easily damaged in the event of an overload. What is more, a very loose, flexible suspension may not be able to adequately resist the pressure changes inside the box at high SPLs, and may physically deform, giving rise to non-linearities in its travel and non-linear distortions in its radiation. This will all be discussed in much more detail in Chapter 11, but suffice it to say here that a small sealed box must suffer from either poor system sensitivity (due to its poor, overall electro-acoustic conversion efficiency) or a low frequency roll-off that begins well into the musical spectrum. The roll-off exhibits a rate of 12 dB per octave below its frequency of resonance, but considerable roll-off may begin well above this frequency, depending on the system Q. Some typical roll-off curves are shown in Figure 3.3. Nevertheless, the time responses (transient responses) of well-designed sealed boxes with correctly matched drivers and adequate damping can be very accurate. Largely for this reason, sealed boxes have a strong following, and large sealed boxes can be the bases of excellent loudspeaker systems.

A sealed box system is said to be critically damped when its size and the driver resonant frequency are matched such that the overall response

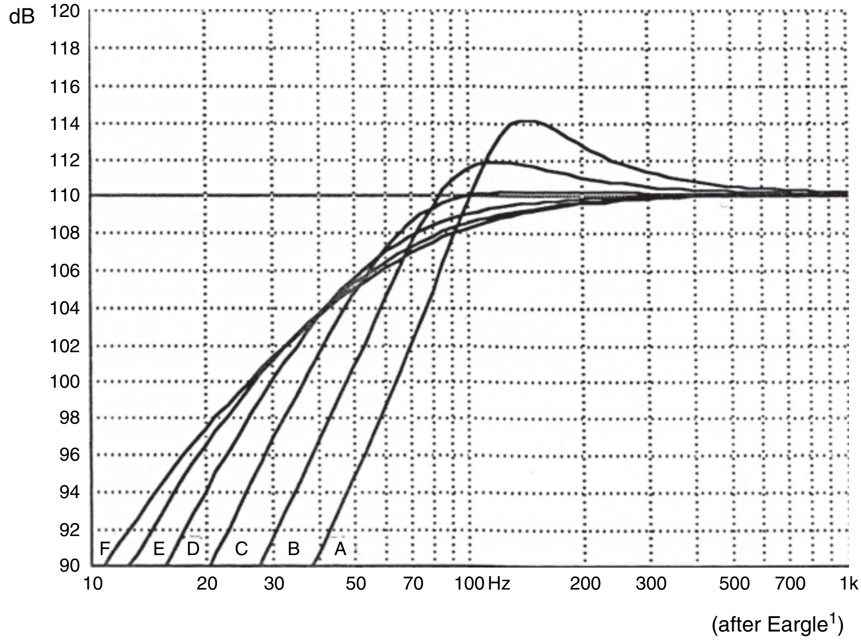


Figure 3.3 Typical low frequency roll-off behaviour of a sealed box loudspeaker, showing the responses of the same driver in cabinets of six different volumes. All measurements in free-field conditions
A, 7L. B, 14L. C, 28L. D, 56L. E, 112L. F, 224L
Loudspeaker free-air resonance 20 Hz

is already 6 dB down at the resonant frequency. With this alignment, the transient response can be exemplary, with no perceptible ringing. The total system Q_{TC} (or quality factor of resonance) is 0.5. The Butterworth 'B2' (maximally flat) alignment is very popular, with a Q_{TC} of 0.7. This exhibits a system response which is 3 dB down at the resonant frequency, and still has a transient response which is extremely well controlled. The low frequency responses can be extended downwards with alignments where the Q_{TC} is set at 1, or even up to 2, but as the Q increases, so does the tendency for the transient response to become extended, and for audible ringing or 'boominess' to become obtrusive. The outcome of these relationships is that if the low frequency −3 dB point is to be dropped to 30 Hz, and a fast, well-damped transient response is required at the same time, then the box must be big. If high SPLs are required, then the only solution to the compromise of a low resonance driver with an adequately robust construction and a good sensitivity is that the driver must also be big.

A 15 inch (380 mm) driver, of high quality, with a 20 Hz free air resonance in a 500 litre enclosure can yield some very impressive bass. However, 'impressive' in this context means full, flat, fast and low distortion – in other words, 'accurate'. Unfortunately, many sealed boxes get themselves a bad name by trying to use ‘boomy’ alignments in forlorn efforts to keep the size down whilst seeking to extend the low frequency response to frequencies that the box size cannot really support. The penalty paid is in terms of low sensitivity and poor transient response. It must be thoroughly understood that there is no clever computer program which can solve this problem. The restrictions that we must accept are deeply entrenched in the physical laws of the universe in which we live. They are that fundamental!

Some manufacturers have tried to sacrifice system sensitivity by lowering the magnet flux in order to lower the system Q. There is a strong ‘amplifier power in cheap’ lobby, who believe that lower efficiency systems can exhibit higher Qs, and hence can be extended in their low frequency range. What they often seem to fail to realise is that a heavier current in the voice coil and a lower power magnet will drastically alter the ratio of the fixed magnetic field to the variable magnet field. The much higher variable field due to the voice coil current can severely distort the position of the flux lines of the weak, permanent magnet, and give rise to loss of low level detail in the sound and increased levels of intermodulation distortion. This highlights perhaps one of the worse aspects of the use of programmable calculators or computers in the wrong hands – they can lead to good results on paper, but they can give rise to unpleasant side-effects in practice. Figure 3.4 gives a graphic illustration of the connection between system Q (Q_{TC}) and the transient response. The Q_{TC} is derived from the electrical, magnetic, mechanical and acoustical properties of the total system – electro

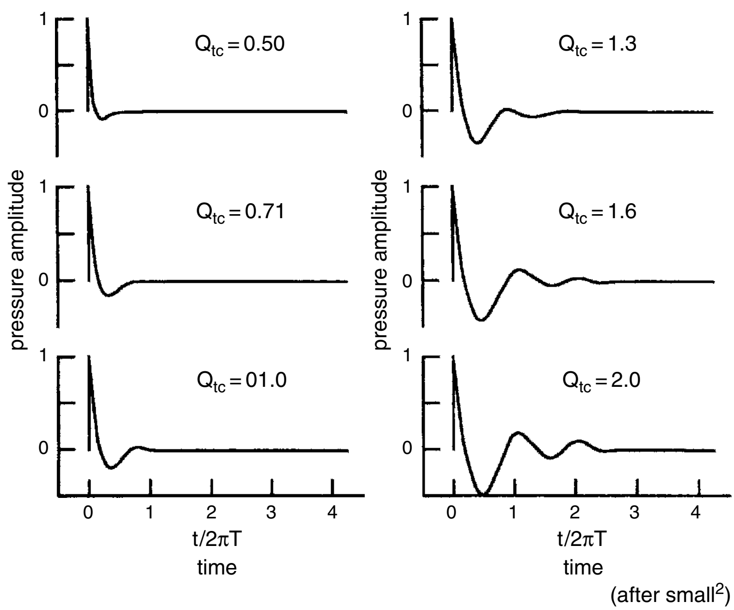


Figure 3.4 Transient response of a sealed box enclosure as a function of the total system Q (Q_{TC}). As the Q_{TC} increases, the transient decay time also increases

magnetic damping, mechanical stiffness and air loading. Note that as the Q_{TC} increases, the transient response becomes longer. This is perfectly logical because the transient response becomes more resonant as the system Q_{TC} becomes more resonant. The amplitude response is boosted and extended downwards by keeping the energy responding for a longer time, and not by instantaneously boosting the level. As stated before, in order to boost the level, and nothing else, a bigger box and driver are needed.

Small sealed boxes with relatively high rates of roll-off can be mounted near to room boundaries, where the constraint of their radiation angle can boost their low frequency output, acoustically, without suffering time penalties (boundary effects are discussed in Chapter 7), but on pedestals in the centre of a room, the low frequencies from small sealed boxes will be found to be either weak, resonant or both. On the metre bridge of a solidly-built mixing console they can also receive some low frequency support, but colouration problems due to the reflective surface being *between* the small loudspeaker and the listening position can be a problem. This 'acoustic mirror' concept was discussed in the previous section, but when applied to floor standing open baffles, the mid and high frequency drive units are usually mounted well clear of the floor. When a small sealed box is placed on top of a mixing console, the sources of mid-range and high frequencies are inevitably close to the reflecting surface, so comb-filtering of the response is the likely result.

To put things into proportion with respect to size, a cabinet which is 3 dB down with a given low frequency driver at 80 Hz would need to be 4 times larger if it were to be 3 dB down at 40 Hz and 16 times larger to be 3 dB down at 20 Hz, so sealed box sizes do tend to get larger very quickly if lower roll-of frequencies are required.

One advantage of sealed box systems is that they are relatively self-protecting in terms of excessive cone excursions. Compared to the open baffle, which offers almost no protection, an input signal below the resonant frequency of a sealed box system will tend to drive the cone at a constant excursion for any given input level, independent of frequency. (See Note 1 at end of chapter.) The thermal overload of the coil is therefore the biggest risk factor in terms of driver integrity at input level extremes.

The lining materials in the boxes also have an effect on the low frequency response. Although they are primarily intended to prevent cabinet resonances at mid frequencies, which may colour the sound by passing to the outside via a relatively acoustically transparent cone, the lining materials can also affect the low frequency damping and total system Q. They should not be too tightly packed, or effectively they will be more or less solid and will reduce the enclosure volume. Neither should they be able to move en masse, or they can introduce non-linear distortion due to their somewhat erratic movement. Given just the right quantity, however, they can not only reduce box resonances but can also make the boxes appear to be up to around 20% acoustically bigger due to their ability to act as heat sinks and slow down the speed of sound by absorbing heat on compression half cycles and releasing it on rarefaction half cycles. The tortuosity of the path through the pores or fibres also gives rise to sound absorption. The density and quantity of the absorbent material inside a sealed box are therefore chosen for the parts that they play in the air loading and damping calculations for the whole system.

3.2.1 Acoustic suspensions

Developed in the 1950s by Edgar Villchur, in the USA. The principal is to use a very low resonance loudspeaker in a small sealed box. The air in the box may push up the resonance by an octave, or more, and is the predominant restoring force for centralising the diaphragm, because the low resonance suspension is very weak. Generally, although an acoustic suspension system is a sealed box, the specialised term is normally only used when the ratio of the cabinet (air) compliance to the driver's compliance exceeds a factor of about 4 to 1[1]. In the late '50s and early '60s, the company Acoustic Research enjoyed very great success with these designs by making huge improvements in the low frequency fidelity of small loudspeaker systems.

3.3 Reflex enclosures

Also known as ported enclosures, vented boxes or phase inverters, reflex enclosures use openings, or ports, to tune the cabinet resonance to a desired frequency. Effectively, the air in the port, which may be a simple hole or a tube, acts as a mass which resonates with the spring created by the air inside the cabinet. In Figure 3.5, a mass is shown suspended below a spring. Almost everybody will intuitively know what would happen if they were to pull down on the weight and then let go – the weight would spring back and the system would go into oscillation until the energy was finally

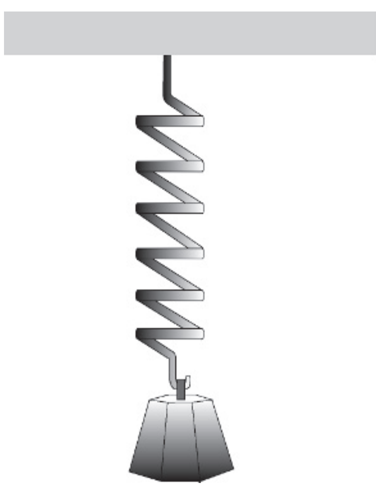


Figure 3.5 A mass/spring system. A mass suspended beneath a spring. It is easy to imagine how pulling down on the weight and releasing it would set up an oscillation due to the mass-spring interaction

dissipated. Adding more weight would cause the oscillation to slow down, as would using a weaker spring. Therefore:

more weight)
weaker spring) slower oscillation (lower frequency)

less weight)
stronger spring faster oscillation (higher frequency)

In the case of a reflex cabinet, a bigger box provides a weaker spring, because the enclosed air is compressed or rarefied proportionately less than in a small box for any given diaphragm displacement. For any given diameter of hole (port), extending it with a tube will lower the resonant frequency because a greater mass of air will be trapped within it. For any given cabinet volume and mass of air in the port, changing the *area* of the port will also change the resonant frequency. *Increasing* the area will *increase* the resonant frequency. This is because there is more surface area in contact with the air-spring, so more force acts upon the air mass, effectively stiffening the spring. There are therefore three variables in the equation, the cabinet volume, the length of the port, and the area of the port – the latter two defining the volume of air in the port, and hence its mass. Air weighs about 1.2 kg per cubic metre, and thus about 1.2 grams per litre.

The cabinet tuning frequency can therefore be calculated approximately from either of the two following equations, the first in imperial measure and the second in metric units.

$$f_v{}^2 = \frac{2700A}{V(L + \sqrt{A})} \tag{3.1}$$

where:

f_v = resonant frequency of box (Hz)

A = area of port in square inches

V = volume of box in cubic feet

L = length of port in inches

OR

$$f = \frac{1}{2\pi} \times \sqrt{\frac{C^2 A}{V\,L_e}} \tag{3.2}$$

where:

f = resonant frequency of box (Hz)
c = speed of sound in air – 340 m/s
A = area of port in *square metres*
V = volume of box in *cubic metres*
L_e = effective length of port in *metres*

Note: L_e allows for an end correction. The effective length of a port tube is, in reality, somewhat longer than the physical length, but for many calculations the actual, physical length can be used.

The formulae are not precise, because there are always variables such as the quantity of air displaced by the drive units themselves, the air displaced by the port tubes, the air displaced by internal bracing, and the effect of the absorbent material inside the enclosure. Nevertheless, the formulae give good working approximations or starting points for calculations. Of course, the cabinet volume is calculated from the *interior* dimensions of the cabinet, not the exterior dimensions.

In practice, when the frequency of resonance of the driver in the cabinet is just above the resonant frequency of the box, the port resonance gives rise to a high load on the diaphragm and greatly reduces the diaphragm excursion. In this way, the ported cabinet can protect the driver from excessive travel while still maintaining a flat response. Below this frequency, the driver output falls, but the port, itself, begins to radiate, thus extending downwards the frequency response. At still lower frequencies the port and loudspeaker outputs occur in opposite polarity, so the response falls off rapidly at 24 dB per octave. Moreover, below the port resonance, air simply pumps in and out of the port under the influence of the driver. At these frequencies, the cabinet is just a box with a big air leak, and it can provide no loading on the driver diaphragm, which then behaves as if it were in an open baffle with no air loading protection, so over-excursions are easy to encounter in reflex enclosures unless the low frequency drive signal is filtered, or has no natural content, below the resonant frequency of the box.

A comparison of the performance of two different low frequency drivers in an open baffle, a sealed box, and a reflex enclosure is shown in Figure 3.6. In practice, a driver would be specifically designed for each type of loading, because the different cabinets or baffles match more optimally with drivers of specific Q_{TC} values. The Q_{TS}, which can be found in many formulae and reference texts, is the sum of the Q_{MS} and the Q_{ES} , which are the mechanical and electrical system quality factors (sharpness of resonance), respectively. The higher the Q, in each case, the more highly tuned is the resonance, as shown in Figure 3.7. For reference, the Q terms commonly found in loudspeaker texts are as follows[2]:

Q_{MS} is the mechanical system Q. It is the ratio of the electrical equivalent of the frictional resistance of the moving parts of the driver to the reflected motional reactance at the free-air resonance frequency of the driver.

Q_{ES} is the electrical system Q, which is given by the ratio of the voice coil DC resistance to the reflected motional reactance at the free-air resonance frequency of the driver.

Q_{TS} is the parallel combination of the Q_{MS} and Q_{ES}, and the equation takes the same form as that for two, parallel, electrical resistors:

$$Q_{TS} = \frac{Q_{MS} \times Q_{ES}}{Q_{MS} + Q_{ES}} \tag{3.3}$$

Q_{TC} is the total system Q of the driver and the cabinet.

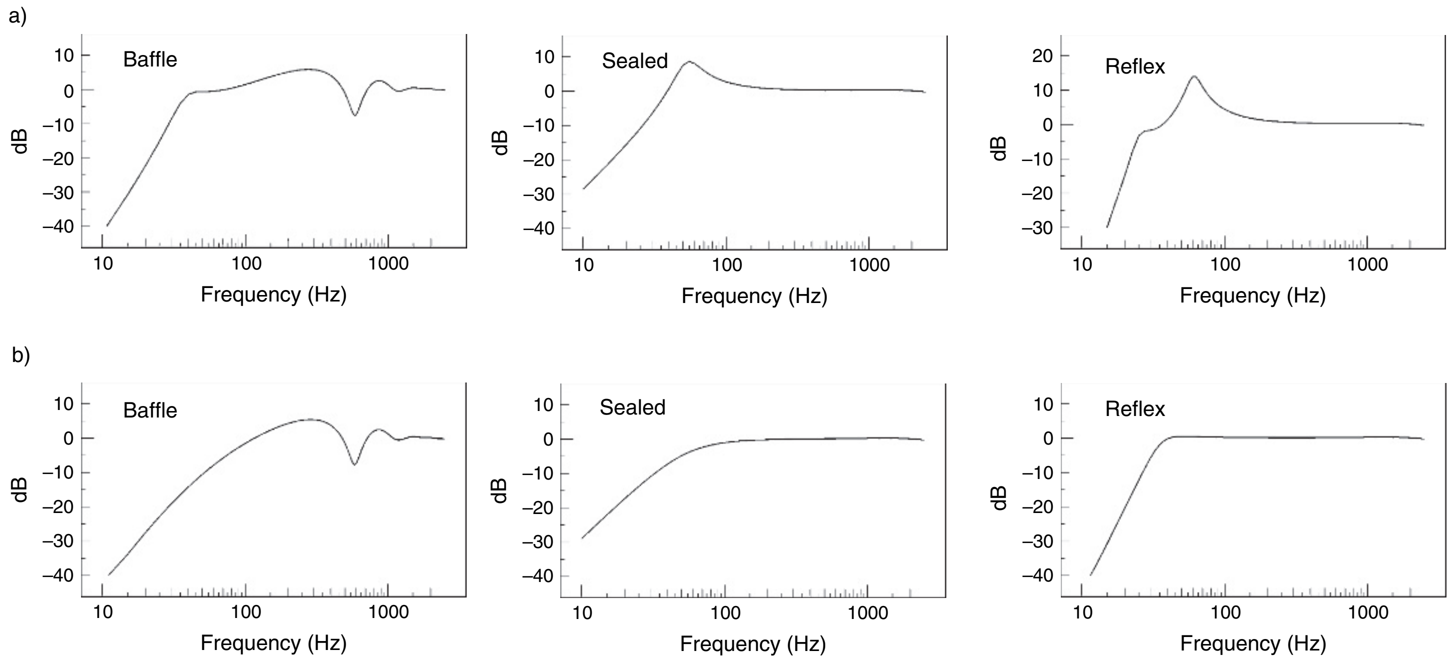


Figure 3.6 a) Response of one driver mounted in an open baffle, a 50L sealed box, and a 50L reflex enclosure with the driver Q optimised for the open baffle response. b) As in a), but with the driver Q optimised for the reflex enclosure. Note how as the driver is optimised for one type of loading, the response may suffer with other types of loading

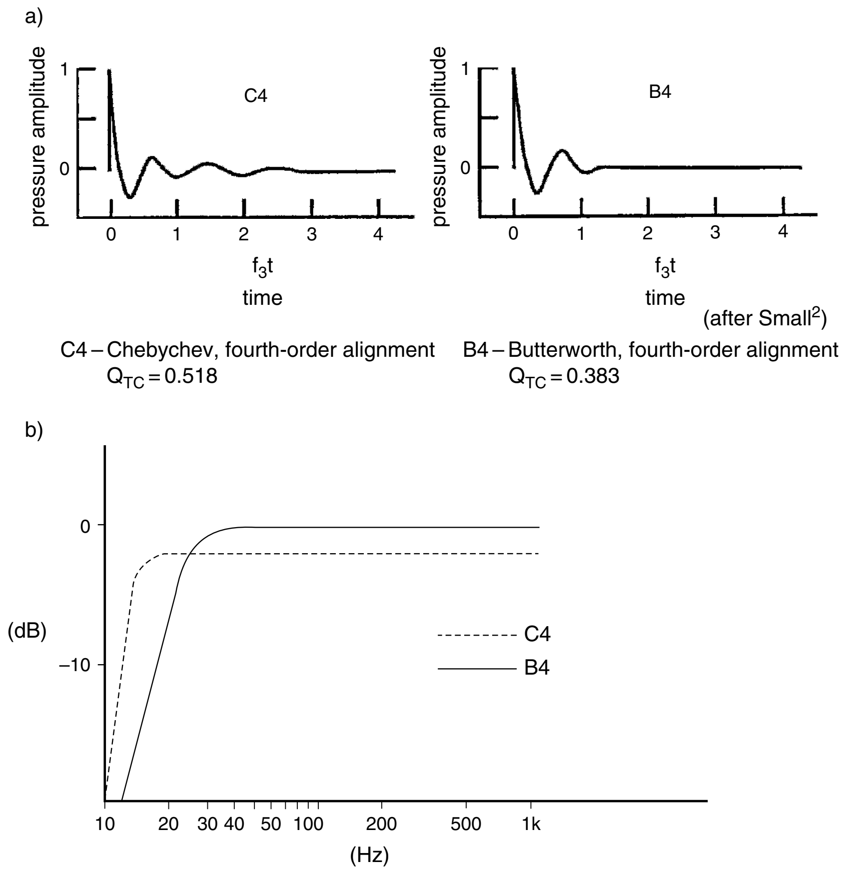


Figure 3.7 a) Transient response of a reflex enclosure as a function of the total system Q (Q_{TC}). Compare with Figure 3.4. b) Pressure amplitude response of the C4 and B4 alignments. Note how the response extension of the C4 alignment corresponds with an increase in the decay time, as shown in a). Transient response is traded for response extension

The free-air resonant frequency of the drivers for reflex enclosures may also need to be different to the optimum resonant frequencies for sealed enclosures of similar size or covering a similar frequency range. The modern tendency is to tailor complete driver designs to given box sizes and system concepts, and the programs now available for computer analysis are very powerful and very accurate. However, careful listening tests are still a *sine qua non* because concentration on the optimisation of one aspect of driver design may unexpectedly change for the worse some other aspect of performance that was not under such close scrutiny. Unfortunately, listening rooms are expensive to build and listening panels can be an expensive luxury. Computer time on the other hand is cheap, and quick, and there has developed a strong tendency to design systems ever more by computer and ever less by careful listening.

3.4 Acoustic labyrinths

These are sometimes referred to as transmission lines, but at low frequencies they are usually *not* transmission lines. A true transmission line needs a rear cavity, straight or folded, at least a quarter of a wavelength long. At 30 Hz, with a wavelength of about 11 metres, the line would need to be around 3 metres long, and lines of this length are rare indeed. A true transmission line works by presenting the correct acoustic impedance at the rear of a driver so that all the backwards radiation propagates away from the driver, never to return. This can be achieved by loading the rear of the driver with an infinitely long pipe (which is rather impractical) or by some other system that absorbs all the sound energy. A finite length pipe can therefore be made to operate as a transmission line if it contains sound absorbing material strategically placed to give the correct, purely resistive acoustic impedance, such as an anechoic wedge. However, in order to work at low frequencies this pipe still needs to be very long, so some form of low frequency tuning is often employed.

If, instead of an infinite pipe we think of an organ pipe, it would exhibit a series of resonant frequencies determined by its length. If we attach it to the rear of the driver and fold it round to the front, there will be interference between the sound from the front of the driver and that from the open end of the pipe. When the pipe is one quarter wavelength long, there will be a high acoustic pressure at the rear of the driver and a high acoustic velocity at the open end, which combines with a phase difference of 90 degrees with the acoustic velocity from the front of the driver, and this provides a useful boost in output. As the frequency is lowered, the output from the pipe increases in phase difference with respect to the direct output from the driver, and so tends towards cancelling the combined output. This yields a 24 dB per octave roll-off below the tuning frequency, which leads to transient problems not unlike those of a conventional reflex cabinet. A finite length open pipe, which is what the vast majority of so-called transmission lines certainly are, is clearly *not* a transmission line, as it works on a completely different principle and yields a different acoustic performance. A typical cross-section of such a design is shown in Figure 3.8(a).

In order to tame the strong resonant behaviour exhibited by the open pipe, absorbent material is introduced into the pipe to add damping. As the amount of absorbent material is increased, the acoustic performance tends towards that of a transmission line except at very low frequencies, where there is insufficient absorption. A carefully lined, or filled, open-ended pipe may thus exhibit some of the properties of a transmission line, but may also rely on the quarter wavelength resonance to supplement the total low frequency sound output. Most commercial 'transmission lines' are therefore really something between the two extremes of a transmission line and an open pipe, depending on the amount of absorption present at any given frequency.

Some versions of 'transmission lines' are closed. In these designs the line is made to be as absorbent as possible. In reality it is a sealed box, but it differs from the sealed box in that there is no added air-spring effect, and therefore it does not raise the free-air resonance of the driver.

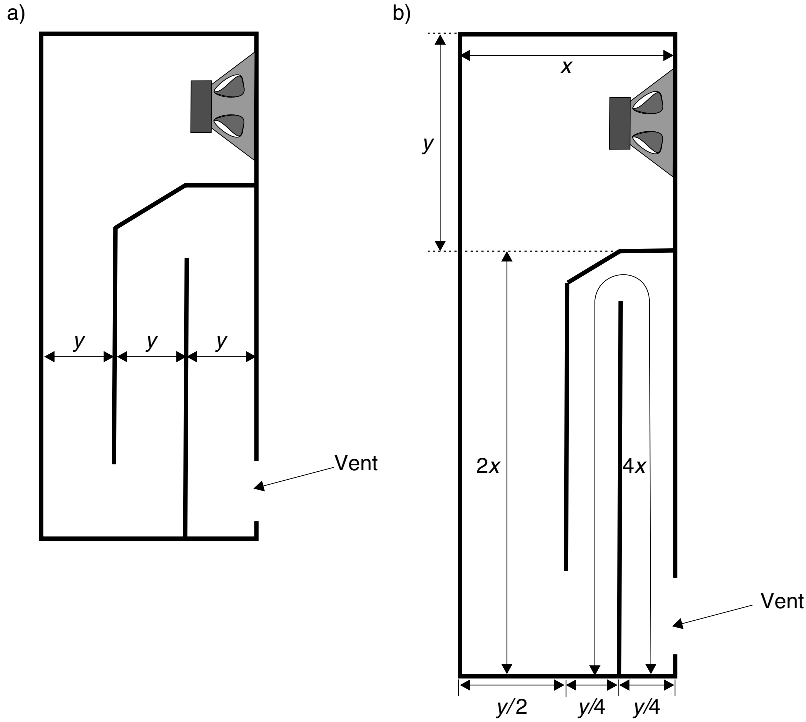


Figure 3.8 Acoustic labyrinths. a) A parallel labyrinth. b) A tapered labyrinth – each time the section of the labyrinth changes, the length (x) doubles and the width (y) halves. The entire labyrinth would be lined with absorbent material in both a) and b)

In practice, the low frequency 'pipe', whether straight or folded, must not only be a quarter wavelength long at the lowest frequencies, but must also be wide enough so as not to obstruct the rear radiation from the driver. In some cabinets, as shown in Figure 3.8(b), the line narrows along its length, each section doubling the length of the previous section and halving its cross-section. Inevitably, in order to maintain close to ideal working, transmission line enclosures need to be large, therefore a small, low frequency transmission line tends to be a contradiction in terms. The smaller cabinets at low frequencies may act as either sealed boxes or reflex enclosures, dependent upon the density and distribution of the absorbent material. The labyrinth enclosures with open-end terminations tend to act like reflex enclosures in that the cone excursions are much reduced at the resonant frequency of the enclosure tuning, which may result in lower distortion than might be expected from an approximately equivalent sealed box at similar SPLs.

One other interesting aspect of acoustic labyrinths is that the highly resistive rear loading can actually *lower* the driver's free-air resonance due to the mass of air which acts directly on the cone, effectively making it heavier when in movement.

Labyrinths/transmission-lines have their following, but many designers feel that they are a complex way of achieving rather little. In lines which almost totally lose the rear radiation, there is no out of phase output at low frequencies, so they exhibit 12 dB per octave roll-offs like sealed boxes.

3.4.1 Modern transmission lines

In the early 1990s, the British company PMC (Professional Monitor Company) began manufacturing a range of 'transmission line' loudspeakers which have since achieved considerable acclaim and commercial success. The subject of what is and what is not a transmission line has been rather controversial in recent years, so Peter Thomas, the principal design engineer and managing director was asked by the authors of this book to try to clarify the matter, and he subsequently supplied the following paragraphs of this section.

The birth of the modern transmission line speaker design came about in 1965 with the publication of A R Bailey's article in Wireless World, "A Non-resonant Loudspeaker Enclosure Design"[3], detailing a working transmission line. Radford Audio took up this innovative design and briefly manufactured the first commercial transmission line loudspeaker. Shortly thereafter John Wright of IMF Electronics designed a range of transmission line designs and made them popular through his refinement and development of Bailey's theory. Although acknowledged as the father of the transmission line, Bailey's work drew on the work on labyrinth design, dating back as early as the 1930s. His design, however, differed significantly in the way in which he filled the cabinet with absorbent materials. Bailey hit upon the idea of absorbing all the energy generated by the bass unit inside the cabinet, providing an inert platform for the drive unit to work from. Unchecked, this energy produces spurious resonances in the cabinet and its structure, adding distortion to the original signal.

The transmission line (TL) is the theoretical ideal and most complex construction with which to load a moving coil drive unit. The most practical implementation is to fit a drive unit to the end of a long duct that is open ended. In practice, the duct is folded inside a conventional shaped cabinet with the open end of the duct usually appearing as a vent on the front of the cabinet. There are many ways in which the duct can be folded, and Figure 3.8 illustrates two typical forms. The line is often tapered in cross-section to avoid parallel internal surfaces that encourage standing waves. Depending upon the drive unit and quantity – and various physical properties – of absorbent material, the amount of taper will be adjusted during the design process to tune the duct to remove irregularities in its response. The internal partitioning provides substantial bracing for the entire structure, reducing cabinet flexing and colouration. The inside faces of the duct or line are treated with an absorbent material to provide the correct termination with frequency to load the drive unit as a TL. A theoretically perfect TL would absorb all frequencies entering the line from the rear of the drive unit, but remains theoretical as it would have to be infinitely long. The physical constraints of the real world demand that the length of the line must often be less than 4 meters before the cabinet becomes too large for any practical applications, so not all the rear energy can be absorbed by the line. In a realised TL, only the upper bass is TL loaded in the true sense of the term (i.e. fully absorbed); the low bass is allowed to freely radiate from the vent in the cabinet. The line therefore effectively works as a low pass filter, another crossover point in fact, achieved acoustically by the line and its absorbent filling. Below this 'crossover point' the low bass is loaded by the column of air formed by the length of the line. The length is specified to reverse the phase of the rear output of the drive unit as it exits the vent. This energy combines with the output of the bass unit, extending its response and effectively creating a second driver.

Phase inversion is achieved by selecting a length of line that is equal to the quarter wavelength of the target lowest frequency. The effect is illustrated in Figure 3.9(a), which shows a hard boundary at one end (the speaker) and the open-ended line vent at the other. The phase relationship between the bass driver and vent is in phase in the pass band until the frequency approaches the quarter wavelength, when the relationship reaches 90 degrees as shown. However by this time the vent is producing most of the output – as shown in Figure 3.9(b). Because the line is operating over several octaves with the drive unit, cone excursion is reduced, providing higher SPL's and lower distortion levels compared with reflex and sealed box designs.

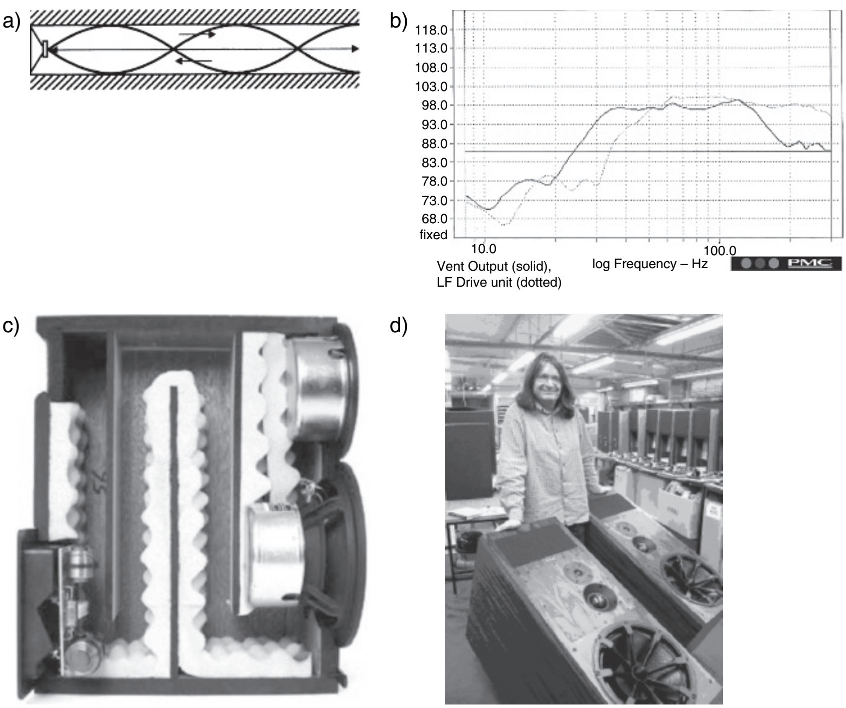


Figure 3.9 PMC transmission lines. a) Phase relationship between the driver and the vent in the quarter wave transmission line. b) The drive unit and vent contributions to the overall output. c) Cut-away view of a PMC transmission line. d) A pair of large PMC cabinets, with their designer

The calculation of the length of the line required for a certain bass extension appears to be straightforward, based on a simple formula:

$$\lambda = 344/4 \times f \qquad (3.4)$$

where:

f is the quarter wavelength frequency
344 ms is the speed of sound in air at 20 degrees C
λ is the length of the transmission line

However the introduction of the absorption materials reduces the velocity of sound through the line, as discovered by Bailey in his original work. Bradbury published his extensive tests to determine this effect in an AES Journal in 1976[4] and his results agreed that heavily damped lines could reduce the velocity of sound by as much as 50%, although 35% is typical in medium damped lines. Bradbury's tests were carried out using fibrous materials, typically longhaired wool and glass fibre. These kinds of materials however produce highly variable effects that are not consistently repeatable for production purposes. They are also liable to produce inconsistencies due to movement, climatic factors and effects over time. High specification acoustic foams, developed by PMC with similar characteristics to longhaired wool, provide repeatable results for consistent production. The density of the polymer, the diameter of the pores and the sculptured profiling are all specified to provide the correct absorption for each speaker model. Quantity and position of the foam is critical to engineer a low pass acoustic filter that provides adequate attenuation of the upper bass frequencies, whilst allowing an unimpeded path for the low bass frequencies.

There are therefore two distinct forms of bass loading employed in a TL, which historically and confusingly have been amalgamated in the TL description. Separating the upper and lower bass analysis reveals why the TL has so many advantages over reflex and sealed box designs. The upper bass is almost completely absorbed by the line allowing a clean and neutral response. The lower bass is extended effortlessly and distortion is lowered by the line's control over the drive unit's excursion. One great advantage of the low frequency extension provided by transmission lines is the perception of deep bass even at low listening levels, due to the extended flatness of the response.

The complex loading of the bass drive unit demands specific Thiele-Small driver parameters to realise the full benefits of a TL design. Most drive units in the marketplace are developed for the more common reflex and sealed box designs and are usually not suitable for TL loading. To design a high efficiency woofer with extended low frequency ability, one tends to need cones which are often extremely light and flexible with very compliant suspensions. Whilst performing well in a reflex design, these characteristics do not match the demands of a TL design. The drive unit is effectively coupled to a long column of air which has mass. This lowers the resonant frequency of the drive unit, negating the need for a highly compliant device. Furthermore, the control of this column of air requires an extremely rigid cone, to avoid deformation and consequent distortion. The lack of available suitable drive units created the necessity for PMC to design a series of drivers employing a flat, 6mm thick diaphragm, manufactured from aerospace materials, that provide extraordinary stiffness whilst maintaining a relatively low mass.

The combination of extended frequency response, higher sound pressure levels and lower distortion afforded by TLs, separates them from reflex and sealed box models. In addition, phase accuracy is superior to many other moving coil designs as a result of the absorption provided by the line in the upper bass range. The low frequency roll off can be as low as 12 dB per octave in highly damped lines, matching the sealed box arrangement and avoiding the large phase changes inherent in reflex designs. A cutaway section of a PMC transmission line is shown in Figure 3.9(c), and a pair of complete cabinets in Figure 3.9(d).

3.5 ABR systems

A variation on the use of reflex enclosures is to use an auxiliary bass radiator (ABR) instead of a port. The ABR often takes the form of a loudspeaker with no magnet assembly. They are often also referred to as passive radiators or drone cones, and are normally of approximately the same size as the cone of the driver. One of the primary advantages of ABRs is that there is no air-flow associated with them, so there is no turbulence or wind noise, which can be a problem when small diameter ports are the only means of tuning a cabinet. On the negative side, the compliance (springiness) of the suspension system of an ABR can be nonlinear at high excursions, and hence can be a source of distortion which can be more subjectively noticeable than port distortion.

ABRs can be used in small boxes, giving them a reflex-type response when the tuning frequency would require ports which would be too small to be practical. In general, ABRs offer reflex-type advantages over sealed boxes, such as 4–6 dB more output on typical programme material and a lower −3 dB point at low frequencies, but the cut-off is more steep once it begins, and, as with any resonant system, the transient response gets smeared. The transient response of a resonant system depends upon the Q of the resonance, as shown in Figure 3.7. There are some alignments (tunings) known as quasi-Butterworth third order (QB3) and sub-Chebychev fourth order (SC4) which use lower Qs. They exhibit transient characteristics rather like sealed boxes of $Q = 1$ (see Figure 3.4) but maintain the reduced diaphragm excursions of the reflex enclosures. However, the choice of musical programme and room acoustics may well be the determining factor as to whether these alignments are beneficial or not. Electronic music in a highly damped control room is much more revealing of tuning resonances than would be romantic music in a relatively live domestic room. In many ways it is fortunate that we have at our disposal a range of loudspeaker performances, because none are perfect, and we have a similar range of musical and acoustical requirements.

ABR systems have not had a totally continuous development. They tend to emerge from time to time as solutions to specific problems. When air is used in a resonator, its density is fixed, so if not enough volume is available in the cabinet for the necessary sized port, or if excessively long tubes are called for (which suffer from viscous losses), then low tuning frequencies

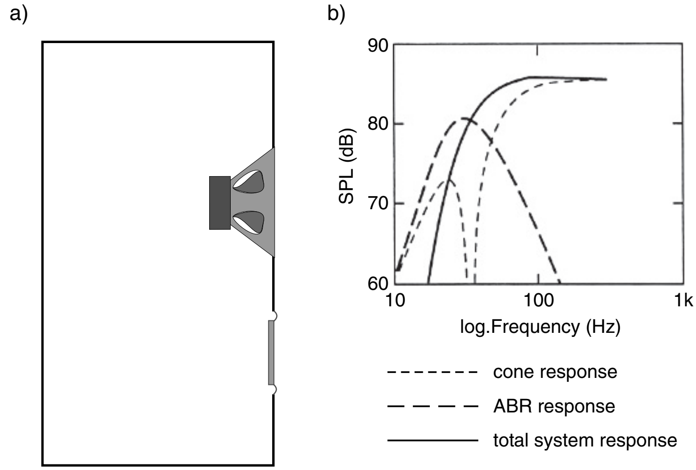


Figure 3.10 The ABR. a) A passive radiator (ABR) loudspeaker system. b) Cone and ABR contributions to the overall response

cannot be accomplished. An ABR offers the use of a number of different materials of varying weights (masses, densities) and also offers the possibilities of lower box tuning when size is limited. Polystyrene diaphragms of the correct weight, suspended in a loudspeaker-type surround (usually a half-roll) are now often the choice of the designers who use ABRs, as shown in Figure 3.10(a). The total acoustic output is the sum of the volume velocity of the front of the driver diaphragm and the out of phase volume velocity of the ABR. The ABR is, of course, driven by the rear of the driver via the air in the cabinet. The relative contributions of the driver and the ABR to the combined output of the system is shown in Figure 3.10(b). ABRs are also sometimes chosen for use in systems where an air-tight cabinet is required, such as for outdoor use.

3.6 Bandpass cabinets

If the low frequency driver is enclosed with air volumes on each side of the cone, and the only radiation to the outside air is via a port in one of the air volumes, the result is a bandpass loudspeaker, as shown in Figure 3.11. Bandpass cabinets are usually restricted to use as sub-woofers because the pass-band is very limited – rarely more than one octave of flat response. The roll-offs at either end of their spectrum of use tends to be 12 dB per octave, because there is no direct radiation to give rise to the phase differences that lead to the 24 dB per octave lower roll-off rate of reflex enclosures, unterminated transmission lines or ABR systems. In the design shown in Figure 3.11, the lower roll-off frequency is governed largely by the inner, sealed chamber, and the upper roll-off frequency is governed by the outer ported chamber.

However, there do exist some designs with both chambers vented, generally in an attempt to gain overall system efficiency in the pass band.

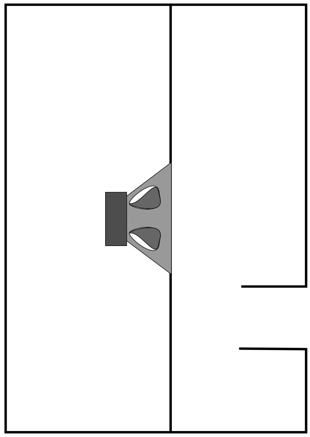


Figure 3.11 A typical bandpass enclosure

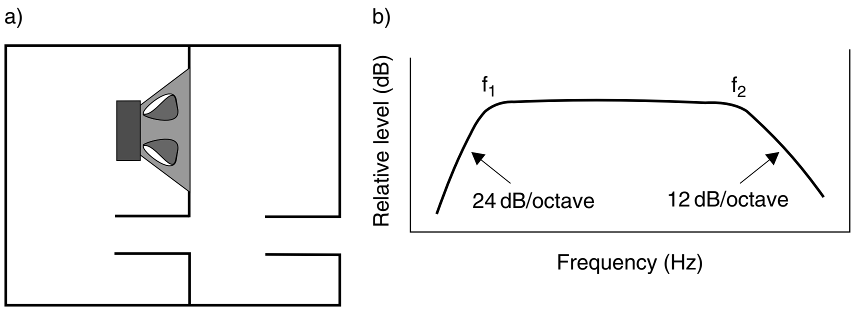


Figure 3.12 a) A bandpass enclosure with the inner chamber ported to the outer chamber. b) Roll-off responses of a)

Bandpass enclosures can normally be much physically smaller than other configurations for any given output capability, but the responses, not surprisingly, tend to be resonant and thus exhibit poor transient responses. Figure 3.12 shows the design of a bandpass enclosure in which both chambers are ported, one to the outside and the other to the outer chamber. This results in a 24 dB per octave roll-off at the lower frequency. A further development is shown in Figure 3.13. In this instance, both chambers are ported to the outside, resulting in a 30 dB per octave roll-off at the lower frequencies and an 18 dB per octave roll-off at the higher frequencies. Both of the latter two designs were developed by the Bose Corporation.[1]

Care must be taken in the acoustic treatment of the chambers which vent to the outside because internal resonances at high frequencies can escape through the ports if not dealt with internally. Care must also be taken when siting bandpass enclosures, because their proximity to solid boundaries may severely affect their response, and the high velocity ports must be free from obstructions. Some designers claim better response linearities due to reduced cone excursions for any given output SPL as compared to other

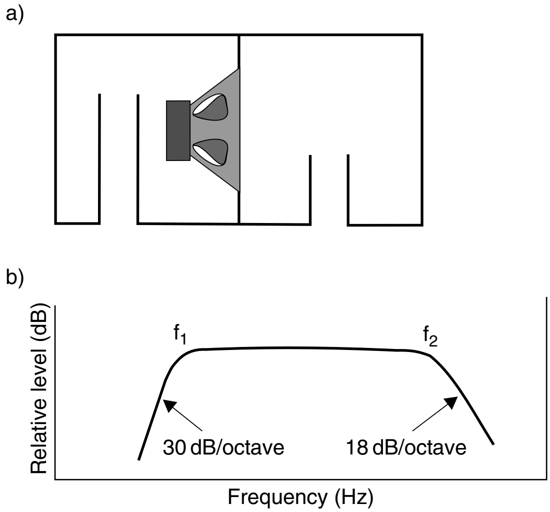


Figure 3.13 a) A band pass enclosure with both chambers ported to the outside. b) Roll-off response of a)

low frequency loudspeaker systems, but their use with true high-fidelity systems is very limited due to transient response anomalies.

3.7 Series driver operation and isobaric loudspeakers

If the port in the outer chamber shown in Figure 3.11 were to be replaced by another drive unit, a system would result as shown in Figure 3.14, with the two drivers in acoustic series, (but with the drivers still connected in

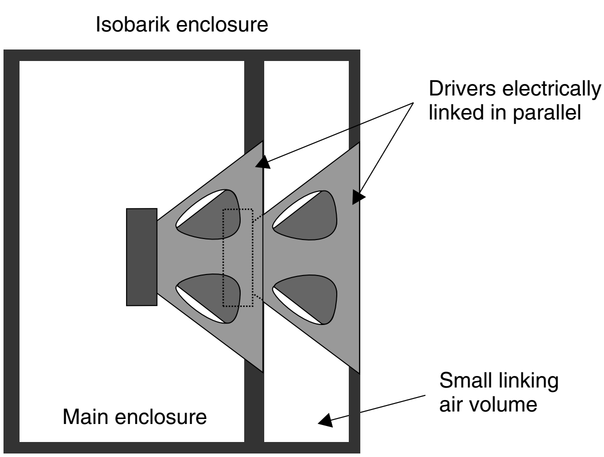


Figure 3.14 The Linn Isobarik enclosure concept. The loudspeaker drivers are connected in acoustic series (cascade), but electrically in parallel

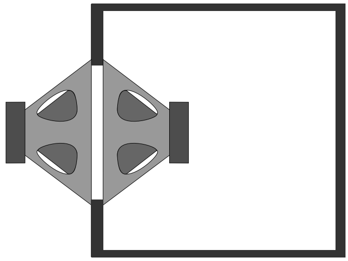


Figure 3.15 A variation on the theme of Figure 3.14

parallel, electrically). At low frequencies, the two drivers behave like a single driver with twice the moving mass. The result is a downward shift in resonant frequency to about 70% of that of just one of the drivers in the sealed enclosure. The reduced back-pressure exerted by the enclosure on the externally radiating driver will result in lower distortion because the inner loudspeaker is tending to keep the pressure in the outer chamber constant. However, in practice, the two drivers also tend to function as one due to the strong coupling by the trapped air mass. Teifenbrun first used the term isobaric for this type of operation – 'isobaric' meaning 'same pressure' (UK patent 1,500,711). Another variation on the theme is shown in Figure 3.15.

3.8 General discussion

Although there are many esoteric designs of low frequency loudspeaker systems on the market, the ones described so far in this chapter are the ones which will cover 99.9% of the loudspeakers to be found in the mainstream music recording and domestic reproduction environments. In general, high levels of fast, flat responses are not available from small enclosures. This subject is dealt with much more thoroughly in Chapter 11, but it seems obvious that nobody would choose to use large, expensive, unwieldy cabinets if more compact solutions were available. Nevertheless, despite the barrage of marketing claims about revolutionary low frequency sources, and a widespread tendency for many people to believe that computer control and signal processing can resolve any problems, the fact remains that the radiation of low frequencies is something that resides firmly in the domain of the laws of acoustics, and they tend to be somewhat inflexible. It should be noted that in nature, only large objects radiate low frequency sounds.

The tendency towards using compact, single, mono low-frequency sources is something that can greatly improve the overall response, both subjectively and objectively, in poor-to-quite-good circumstances of use. However, as the room acoustics become better controlled and the signal path, including the loudspeakers, becomes higher in resolution, the optimum choice for low frequency reproduction tends to return to favour stereo sources in integrated loudspeaker systems – i.e. without physically separated sub-woofers. What is optimum in the mid-to-reasonably-high quality range of loudspeaker systems may not necessarily be extrapolated as being best at the highest quality levels. One must be very careful not to generalise about things which are for specific applications. For example, an orchestral recording with phase differences in the left and right channels at low frequencies would benefit, in a good room, from stereo bass. It has now been determined that only by restricting the crossover frequency to less than 50 Hz can the bass be summed into mono without loosing the spaciousness in the sound[5]. However, in a poorly controlled room, a mono bass may lead to less general confusion if crossed over and combined into mono an octave higher. Such choices can be very circumstantially dependent.

3.9 Cabinet lining materials

Any hard-surfaced box will suffer from internal reflexions and resonances when excited by a loudspeaker drive unit mounted in one of its surfaces. The nature of cone loudspeakers is such that they are relatively transparent to sound at mid frequencies, so any reflexions and resonances occurring within the box are likely to pass to the outside via the cone, and combine with the directly radiating sound in a way that will give rise to undesirable colouration. One of the fundamental reasons for applying absorbent linings of foams or fibrous materials to the inside of loudspeaker cabinets is to reduce to inconsequentially low levels the colouration effects by rendering the boxes as acoustically non-reflective as is reasonably possible. A further advantage of the application of porous or fibrous materials is that they can slow down the speed of sound, and thus make the cabinets appear to be acoustically larger than they are physically. The practical limit to this size increase appears to be around 20% (which is still a useful gain) because the thermal transfer characteristics of the materials normally used are not sufficient to achieve the theoretical 41% maximum. As briefly mentioned in Section 3.2, air heats up when compressed and cools on rarefaction. Both of these effects tend to augment the speed of sound on the successive half-cycles because the air itself is a poor conductor of heat, so the thermal changes are trapped within the waves. However, if the air is in close contact with another material, distributed throughout its volume, which *can* conduct the heat, the augmentation of the speed of sound due to the heat of compression and the cold of rarefaction will not be apparent. The ratio of the speed of sound with and without this augmentation is about 1.41 to 1, hence the approximately 40% difference in apparent box size if the heat conduction were total.

Partly for this reason, but also because fibrous materials are better absorbers where the particle velocities of the air movement are at their highest (they must be zero at rigid boundaries, so they are at their lowest close to the boundaries), the absorbent materials are best placed in the volume of the box, lightly packed, and not only against the sides of the box. Reticulated (open cell) foams, glass fibre, mineral wool, bonded acetate fibres, polyester fibres and cotton-waste felt are all common lining materials. The cut-away view of the 'transmission-line' cabinet shown in Figure 3.9 illustrates the use of a synthetic foam lining which the manufactures specifically chose for its ability to maximally damp the line. Material types and densities are normally chosen with care for specific applications.

The KEF loudspeaker company, in the 1980s, noticed some differences in their loudspeaker frequency responses depending upon whether the excitation signal was of a steady state or transient nature. The discrepancy turned out to be due to non-uniform movement of some of the internal lining materials, which was somewhat uncontrolled after the shock excitation of a transient signal. The lining materials should therefore not be in panels which can vibrate en masse, or non-linear effects may be sufficient to be noticeable in the sound from the loudspeakers. Vibrating lining materials may settle into regular patterns on relatively steady signals, but can be excited rather unpredictably by transient shocks. At high SPLs, the linings can move in rather erratic manners, but the effects are usually swamped by the higher SPLs of the radiation direct from the driver. Nevertheless the ideal lining would be relatively inert. Colloms[6] claims that unstable lining materials can impair the sense of ‘rhythm’ from a loudspeaker system.

At low frequencies, the thicknesses of the lining materials are far too little to provide much absorption, but the cabinets are normally also far too small to support any resonant modes (100 Hz would require an internal dimension of at least 1.6 metres) so the lack of absorption rarely becomes a practical problem. However, at higher frequencies, the absorption is important in order to reduce internal resonances which could powerfully excite structural resonances in the cabinet walls, and which would then radiate into the listening rooms.

3.10 Cabinet constructions

Above all, loudspeaker cabinets should be either rigid or heavy or both. A non-rigid (and/or lightweight) cabinet will be excited into vibration by the drive unit(s). In most cases, rigid materials tend to be heavy or expensive, so light, cheap loudspeakers always tend to be sonically suspect because they will probably suffer from cabinet vibration colouration. Any part of a cabinet which vibrates in sympathy with the driver diaphragm will, itself, act as a diaphragm and interfere with the driver output, leading to colouration of the overall sound output. In order to prevent the structural vibration of cabinets, sandwich constructions can be used, with lead sheet or plasticised deadsheets between the layers. Internal bracing is also an option, which pushes up the resonant frequencies into regions that tend to be more easily damped. Phenolic materials are another common choice because they can produce wood composites of very high density and rigidity. Lighter weight, highly rigid materials are sometimes to be found, in the nature of honeycombs or matrices, but they tend to be expensive and difficult to manufacture. However, in all cases, the goals are similar – rigidity and high vibration damping. It is important that the materials do not ring when struck.

The panel radiation is proportional to the size of the cabinet, so the problem of resonance avoidance becomes greater as cabinet size increases. For this reason, a given thickness of material for the wall of a small cabinet may need to be significantly increased as the cabinet size increases. The weight therefore increases greatly, because not only are the panel sizes larger, but they must also be thicker if the same vibrational insensitivity is to be maintained. Large loudspeaker systems of high quality tend to be very heavy indeed.

In many cases, a stiff, large panel, which is well-behaved at low frequencies may still ring at mid-frequencies, and a well-damped panel at mid frequencies may flex at low frequencies. Finding solutions which are dead at all frequencies is often difficult. In some of the more esoteric designs, mineral-loaded acrylics and Melamine are used as panel materials, but they can be very difficult to work with.

3.11 Cabinet shapes and diffraction effects

In many cases, modern loudspeakers, although basically of rectangular shape, have rounded or chamfered edges. Figure 3.16 reproduces the classic work of Olson[7] on the subject of the diffraction effects on the overall loudspeaker response due to cabinet shape. The responses are for an identical drive unit in each cabinet. The sphere looks attractive, but difficulty of manufacture and problems due to all the internal axial reflexion path lengths being the same lead to practical problems in its implementation. Figures 3.16 J and L are reminiscent of many modern designs, and their validity is borne out by the response plots. Of course, with built-in/flush-mounted loudspeakers, the diffraction problem is nullified, which is one reason why so many professional monitors are so mounted.

At low frequencies, the sound from a loudspeaker cabinet is radiated spherically if the loudspeaker is mounted in free space. At higher frequencies, where the wavelengths are small with respect to the front face of the cabinet, the *cabinet* will tend to act like an infinite baffle, and the radiation will be hemispherical. At still higher frequencies, where the wavelengths are small compared to the *radiating diaphragm*, the sound will be beamed forwards, regardless of whether it is mounted on a baffle, or not. The diffraction effects largely arise at the transition between the first two zones of radiation, i.e. between the spherical and hemispherical radiation zones.

As a sound wave in this transition zone radiates away from a source on a finite-sized cabinet wall, it spreads out as it propagates in the manner of half of a spherical wave. When the wave reaches the edge of the wall, it suddenly has to expand more rapidly to fill the space where there is no wall (see Figure 3.17). There are two consequences of this sudden expansion. First, some of the sound effectively ‘turns’ the corner, around the edge, and carries on propagating into the region behind the plane of the source. Second, the sudden increase in expansion rate of the wave creates a lower sound pressure in front of the wall, near the edge, than would exist if the edge were not there. This drop in pressure then propagates away from the edge into the region in front of the plane of the source. The sound wave that propagates behind the plane of the source is in phase with the wave that is incident on the edge, but the one that propagates to the front is in phase opposition. These two ‘secondary’ sound waves are known as diffracted waves and they ‘appear’ to emanate from the edge; the total sound field may then be thought of as being the sum of the direct wave from the source (as if it were on an infinite baffle) and the diffracted waves.

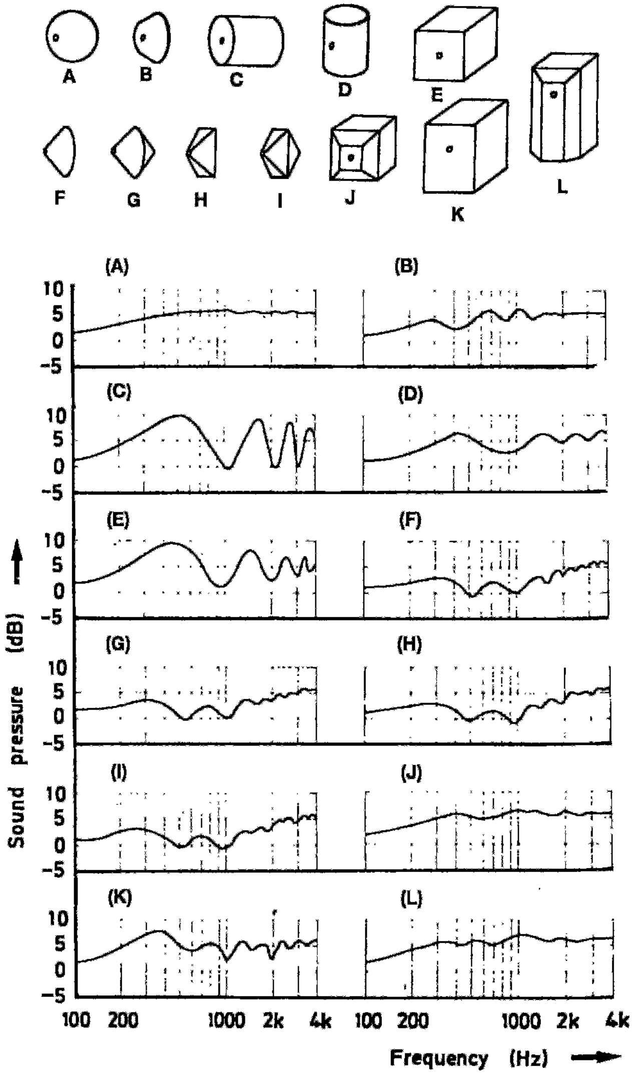


Figure 3.16 Olson's classic work on the effects of cabinet shapes on driver responses[5]

The edges of the cabinets can be thought of as small loudspeaker radiating in antiphase to the real driver. The direct wave exists only in front of the baffle; the region behind is know as the 'shadow' region where only the diffracted wave exists.

At low frequencies, the diffracted waves from all of the edges of the finite-sized cabinet sum to yield a sound field with almost exactly one half of the pressure radiated by the source of an infinite baffle. Thus behind the cabinet there is pressure due to the diffracted wave only, and in front of the cabinet there is the direct wave plus the negative-phased diffracted wave. Assuming that the edge is infinitely sharp (has no radius of curvature),

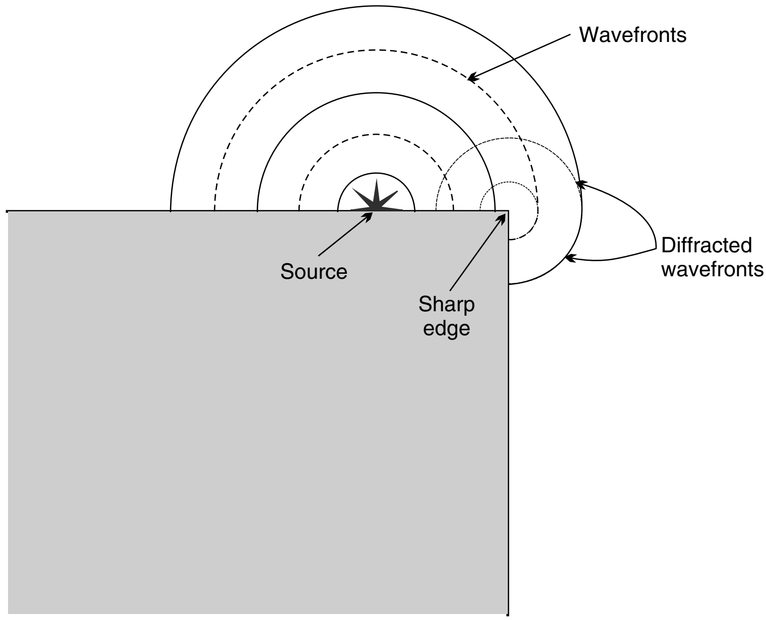


Figure 3.17 Graphical representation of the sudden increase in the rate of expansion of a wavefront at a sharp edge. The diffracted wave in the shadow region behind the source plane has the same effect as the wave incident on the edge; the diffracted wave in front of the source plane is phase reversed

there can be no difference between the strength of the diffracted wave at low frequencies and that at high frequencies (the edge remains sharp regardless of scale). The only difference, therefore, between the diffracted waves at low frequencies and those at higher frequencies is the effect that the path length differences between the source and different parts of the edge has on the radiated field. The diffracted waves from those parts of the edge further away from the source will be delayed relative to those from the nearer parts, giving rise to significant phase differences at high frequencies but not at low frequencies. The net result is a strong diffracted sound field at low frequencies and a weak diffracted sound field at high frequencies.

Figure 3.18 shows the results of a computer simulation of the typical effect that a finite-sized cabinet has on the frequency response of a loudspeaker. Figure 3.18(a) is the on-axis frequency response of an idealised loudspeaker drive-unit mounted in a true infinite baffle. The response is seen to be uniform over a wide range of frequencies. Figure 3.18(b) is the frequency response of the same drive-unit mounted on the front of a cabinet of dimensions 400 mm high by 300 mm wide by 250 mm deep. The 6 dB decrease in response at low frequencies, due to the change in radiation from baffled to unbaffled, is evident from a comparison between Figures 3.18(a) and (b). Also evident is an unevenness in the response in the mid-range of frequencies. These response irregularities are due to path length differences from the diaphragm to the different parts of the diffracting edges and on to the on-axis observation point. Unlike the low-frequency

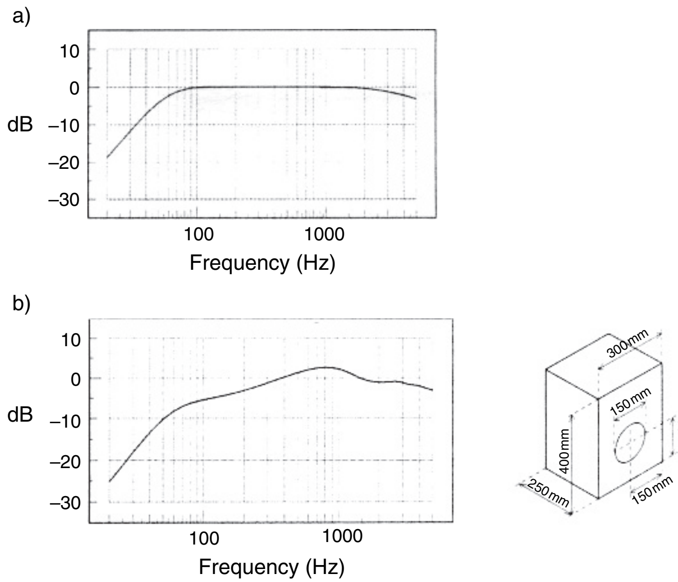


Figure 3.18 a) On-axis frequency response of an idealised loudspeaker diaphragm mounted in an infinite baffle: the response is uniform over a wide range of frequencies. b) On-axis frequency response of the same loudspeaker diaphragm mounted on the front face of a finite-sized cabinet (the rear enclosure size is assumed to be the same in both cases). The response has reduced by 6 dB at low frequencies and is uneven at higher frequencies. The differences between this and Figure a), above, are due to the diffraction from the edges of the cabinet

behaviour, these are dependent upon the detailed geometry of the driver and cabinet and the position of the observation point. Therefore, in order to try to ameliorate these response irregularities, many loudspeakers have contoured edges. Although this does not eliminate diffraction, it tends to make the transitions from the baffled to the unbaffled conditions occur in a less abrupt manner, and thus have a less disturbing effect on the axial frequency response. The off-axis responses are also improved.

Although the combined effects of loudspeaker and rooms is the subject of Chapter 7, it is worth noting here the combined effect of cabinet diffraction and nearby surfaces. Figure 3.19 shows the effect on the axial response when the loudspeaker shown in Figure 3.18(b) is placed against, and close to, a wall. The different distances to the wall, behind the loudspeaker cabinet, give rise to different reflected paths for the diffracted waves, and hence different disturbance patterns in the on-axis, forward response. The clear advantage of flush-mounting the loudspeakers into a wall can be seen by comparison to Figure 3.18(a). Nevertheless, it should be notes that loudspeakers designed for free-standing may have had their low frequency responses engineered for a higher low frequency output, and flush-mounting them may result in an excess of low frequencies. Active loudspeakers often have filter controls which can compensate for this, and reflex cabinets can sometimes have their ports reduced in size such that a flat response can easily be restored. In other cases, a slight bass boost may be deemed to be more acceptable than the irregular response resulting

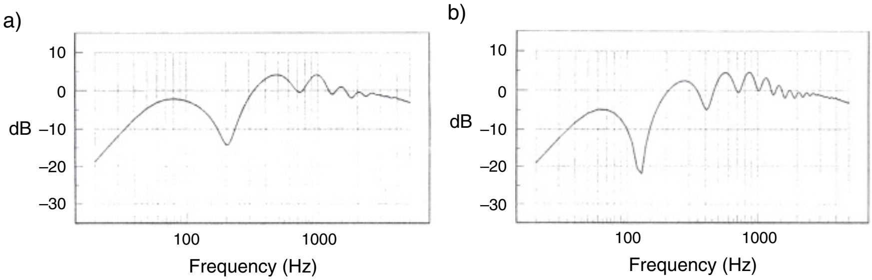


Figure 3.19 a) On-axis frequency response of the same diaphragm and cabinet as in Figure 3.18(b) but with the rear of the cabinet against a rigid wall. Interference between the direct sound from the loudspeakers and that reflected from the wall produces a comb-filtered response, but the response level at low frequencies is restored to that for the infinite baffle case (Figure 3.18(a)). b) As in a), above, but with the rear of the cabinet 0.25 metres from the rigid wall

from the diffraction. However, the bass rise due to the flush mounting is much more easily equalised than the diffraction irregularities, which tend not to be equalisable because of the complexity of the response delays from the reflexions.

In Figure 3.16, most of the drive unit positions are symmetrically placed. In practice, the diffraction effects can often be reduced by the non-symmetrical positions of the drivers with respect to the cabinet boundaries, but the necessity for this will depend upon the nature of the drive unit and the size of the cabinet. Some modern loudspeakers have the mid and high frequency drivers mounted in shallow horns, sometimes referred to as waveguides, which can project the wave in a more forward direction, and greatly reduce the effects of edge diffraction. In general, it is also important to keep the front surface of the loudspeaker system as smooth as possible by recessing the drivers and screw heads.

3.12 Front grilles

It is somewhat rare, these days, to see front grilles on loudspeakers for professional use. The fact is that there are no truly transparent grille materials, but in domestic use their use is strongly justified, not only for aesthetic reasons but also for protection against children and over-zealous cleaners. Fabric grilles usually require wooden frames, and these can become diffraction sources as they provide a step at the edge of the baffle. Foam grilles can avoid this problem, but self-supporting foams of sufficient thickness will exhibit different distances through which the sound must pass, dependent upon the angle of radiation. They can also obstruct the air flow in the ports of reflex enclosures (and open transmission lines) where the exits are on the front face of the cabinets.

However, in some cases, the grille losses have been taken into account in the design of the loudspeaker system, so the removal of grilles should not be undertaken without careful listening to the effects. The diffusive effects of some grilles have been reported to impair the stereo imaging of some loudspeaker designs. In general, the best grille is no grille from a purely sonic viewpoint.

3.13 Cabinet mounting

Figure 3.19 shows the effects of placing a loudspeaker near to reflective surfaces. Obviously, a floor standing loudspeaker should have been designed for standing on the floor, but the intended mounting conditions for small cabinets is not always so obvious. The well-known Yamaha NS10 was originally designed as a bookshelf loudspeaker for domestic use. Its response was therefore tailored such that its bass/mid/treble response was best balanced when the loudspeaker was placed with its back against a wall. This fact was partly responsible for its success when mounted on top of the meter bridges of mixing consoles, because the flat surface below the loudspeaker tended to reinforce the low frequencies in the same way as a wall behind the cabinet. Mounting this type of loudspeaker on a stand in free space will lead to a reduction in the low frequency response, but mounting the loudspeaker on a mixing console will cause time-smearing and comb filtering, the causes of which are highlighted in Figure 3.20.

Mounting loudspeakers on table tops or work surfaces is something that should be avoided, because the effects shown in Figure 3.20 will be exaggerated and colouration of the sound will be inevitable. Far too many people fail to realise that the mounting of a loudspeaker forms part of the loudspeaker system itself, and no loudspeaker's sound is independent of its mounting conditions. A poor loudspeaker, well mounted, may sound better than a good loudspeaker poorly mounted. The number of television and video studios in which the loudspeakers are really appallingly mounted indicates how little many of their users care about the sound, despite their protestations to the contrary.

Wall mounting of loudspeakers will need to take into account the nature of the wall, as not all walls can be considered to be rigid at low frequencies. The nature of the wall – plasterboard, hollow bricks, solid bricks, stone, concrete etc – will affect the loudspeaker response, so if the wall in the showroom was not the same as the wall at home, the same low frequency response will not be heard.

Placing loudspeakers on pieces of furniture is not recommended because of the vibrational coupling which will lead to sound being radiated by the furniture. Even so-called bookshelf loudspeakers should only be mounted on substantial bookshelves, which should ideally also be rather full of books, to add mass, damping, and reduce diffraction effects.

Heavy, narrow, floor stands are recommended, with wide bases for stability, and coupled to the floors (either wood, or carpeted) by spikes. Soft rubber pads are not usually a good idea because they can give rise to rocking; hard rubber is a better option. Broad, columnar stands can induce floor reflexion problems by obstructing the free passage of the sound below the cabinets. Under all circumstances, wobbly mounting systems should be avoided, because the imaging can be severely impaired, even by small movements of the cabinets.

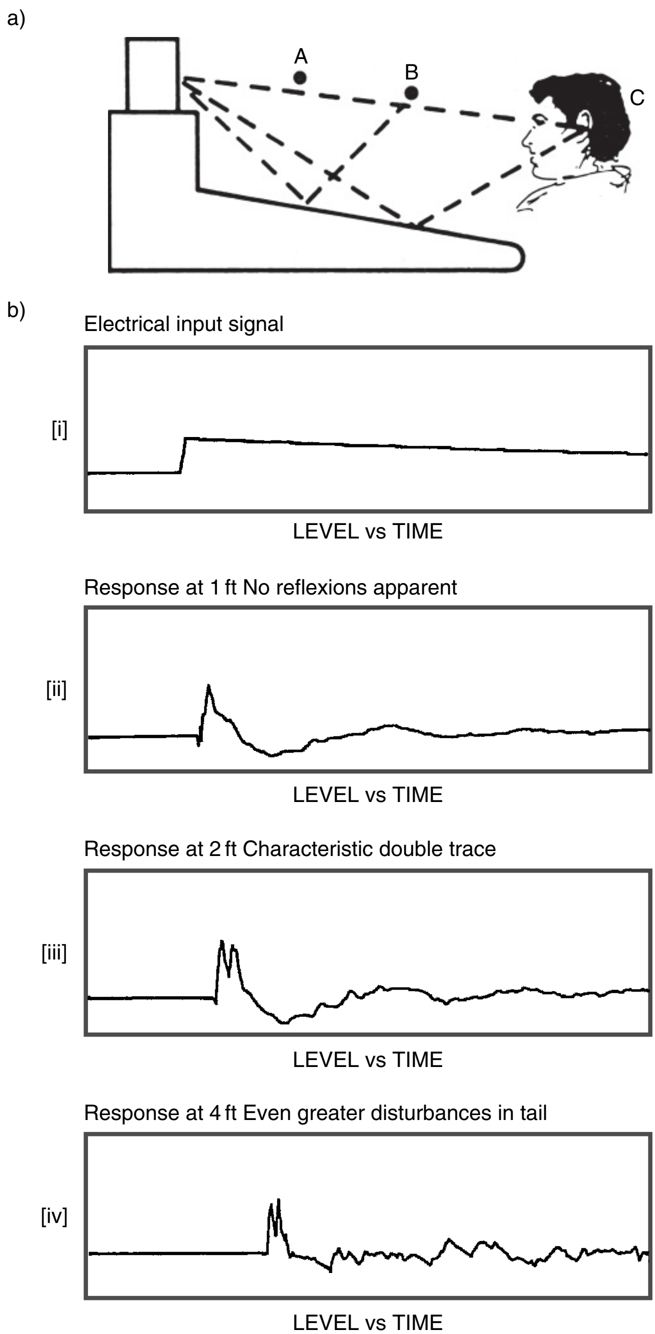


Figure 3.20 Effects of desk-top reflexions on transients. a) Positions A, B and C relate to the 1 foot (30 cm) 2 foot (60 cm) and 4 foot (120 cm)) responses, shown below. b) Transient responses at the positions shown in a), above

In effect, all of these different mounting regimes can be thought of as extensions to the loudspeaker cabinets, because they directly affect the loading on the diaphragms and the radiated sound output. They are not just simply different places to put a loudspeaker cabinet.

Note 1

In order to maintain a constant output SPL, a diaghragm needs to move four times the distance each time the frequency falls by an octave. The fact that the cone excursions are independent of frequency below the resonant frequency of a sealed box loudspeaker system is what gives rise to the 12dB/octave roll-off below that frequency.

References

1 Eargle, J. M., 'Loudspeaker Handbook', Chapman and Hall, New York, USA, (1997)
2 Small, R., 'Direct Radiator Loudspeaker System Analysis and Synthesis – (Parts 1 and 2)'. Journal of the Audio Engineering Society, Vol 20, No 5, and Vol 21, No 1, (1972 and 1973)
3 Bailey, A. R., 'A Non-Resonant Loudspeaker Enclosure Design', Wireless World p 483–486, (October 1965)
4 Bradbury, L. J. S., 'The Use of Fibrous Materials in Loudspeaker Enclosures', Journal of the Audio Engineering Society, Vol 24, pp 404–412, (April 1976)
5 Martens, W., Braasch, J., Woszczyk, W., 'Identification and Discrimination of Listener Envelopment Percepts Associated with Multiple Low-Frequency Signals in Multi-Channel Sound Reproduction', AES 117th Convention, Pre-print No 6229, (October 2004)
6 Colloms, M., 'High Performance Loudspeakers', 6th Edition John Wiley & Sons, Chichester, UK (2005)
7 Olson, H. F., 'Direct Radiator Loudspeaker Enclosures', Journal of the Audio Engineering Society, Vol 17, No 1, pp 22–29, (January 1969)

Chapter 4

Horns

No practical direct radiating loudspeaker can achieve high radiation efficiency at low frequencies. For example, a diaphragm with a diameter of 250 mm has a radiation efficiency of just 0.7% at 50 Hz when mounted in an infinite baffle, and half that when mounted in a cabinet. Sound power output is proportional to the product of the mean-squared diaphragm velocity and the radiation efficiency, so a low radiation efficiency means that a high diaphragm velocity is required to radiate a given sound power. The only way in which the radiation efficiency can be increased is to increase the size of the radiating area, but larger diaphragms have more mass (if rigidity is to be maintained) which means that greater input forces are required to generate the necessary diaphragm velocity. (This is discussed further in Chapter 11.) The *electroacoustic* efficiency is defined as the sound power output radiated by a loudspeaker per unit electrical power input. Because of the relatively high mass and small radiating area, electroacoustic efficiencies for typical loudspeaker drive-units in baffles or cabinets are of the order of only 1-5%. However, horn loudspeakers can combine the high radiation efficiency of a large radiating area with the low mass of a small diaphragm in a single unit. This is achieved by coupling a small diaphragm to a large area via a gradually tapering flare. This arrangement can result in electroacoustic efficiencies of 10-50%, or ten times the power output of the direct-radiating loudspeaker for the same electrical input. Additionally, horns can be employed to control the directivity of a loudspeaker and this, along with the high sound power output capability, is why they are used extensively in public address and sound reinforcement loudspeaker systems.

The following sections describe, in a conceptual rather than mathematical way, how horns increase the radiation efficiency of loudspeakers, how they control directivity, and why there is often the need to compromise one aspect of the performance of a horn to enhance another.

4.1 The horn as a transformer

Close to the diaphragm, in the hydrodynamic near-field, the change in area of an acoustic wave as it propagates gives rise to a ‘stretching pressure’ which is additional to the pressure required for sound propagation. In other words, imagine a balloon being inflated. As well as the outward movement of the skin, radially, the surface is also expanding laterally. Dots painted on the surface of the balloon would move apart as they moved away from the centre. The dots moving apart represent the stretching pressure which does not contribute to sound propagation as it is in phase quadrature (90 degrees) with the radial velocity, so the acoustic impedance in the near-field is dominated by reactance (for readers unfamiliar with the concepts of impedance, reactance and resistance see 'Impedance' in the glossary). As a consequence, large particle velocities are required to generate small sound pressures when the rate of change of area with distance of the acoustic wave is significant. It is this stretching phenomenon that is responsible for the low radiation efficiency of direct-radiating loudspeakers at low frequencies. Physically, one can imagine the air moving sideways out of the way, in response to the motion of the loudspeaker diaphragm, instead of moving backwards and forwards. In the hydrodynamic far-field, the stretching pressure is minimal, the acoustic impedance is dominated by resistance, and efficient sound propagation takes place. The only difference between the sound fields in the near- and far-fields is the rate of change of area with distance of the acoustic wave; the flare of a horn is a device for controlling this rate of change of area with distance, and hence the efficiency of sound propagation.

Horns are waveguides that have a cross-sectional area which increases, steadily or otherwise, from a small throat at one end to a large mouth at the other. An acoustic wave within a horn therefore has to expand as it propagates from throat to mouth. The manner in which acoustic waves propagate along a horn is so dependent upon the exact nature of this expansion that the acoustic performance of a horn can be radically changed by quite small changes in flare-shape. It is usually assumed in acoustics that changes in geometry that are small compared to the wavelength of the sound of interest do not have a large effect on the behaviour of the sound waves, so why should horns be any different? The answer lies in the stretching pressure argument above. The concept of a stretching pressure can be applied to horns by considering *flare-rate*. Flare-rate is defined as the rate of change of area with distance, divided by the area, and usually has the symbol m.

$$m(x) = \frac{1}{\mathrm{S}(x)}\frac{\mathrm{dS}(x)}{\mathrm{d}x} \tag{4.1}$$

Where $S(x)$is the cross sectional area at axial position x. The simplest flare shape is the conical horn, as shown in Figure 4.1, which has straight sides in cross-section, and where $S(_0)$ is the area of the throat (*at* $x = 0$) and x_0 is the distance from the apex of the horn to the throat. The sound field within a conical horn can be thought of as part of a spherical wave field, and has a flare-rate which is dependent on distance from the apex.

The flare rate in a conical horn (and in a spherical expanding wave) is therefore highest close to the throat, decreasing with increasing distance *from* the throat.

The radial dependence of the flare-rate in a conical horn (and a spherical wave) gives rise to a gradual transition from the reactive, near-field dominated behaviour associated with the stretching pressure, to the resistive radiating, far-field dominated propagation as a wave propagates from

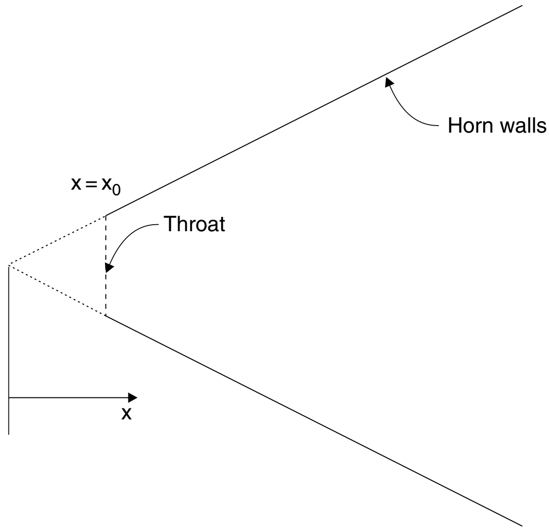


Figure 4.1 Geometry of a conical horn
The origin for the axial coordinates is usually considered to be the imagined apex of the cone

throat to mouth. The transition from near- to far-field dominance is gradual with increasing frequency and/or distance from apex, so distinct 'zones' of propagation are not clearly evident.

However, a more common flare shape for loudspeaker horns is the exponential. The flare shape of the exponential horn is shown in Figure. 4.2. The flare-rate of an exponential horn is constant along the length of the horn, giving rise to a behaviour that is quite different from the conical horn. At low frequencies, and throughout the entire length of the horn, the reactive, near-field-type propagation dominates, and, if the horn is sufficiently long, an almost totally reactive impedance exists everywhere.

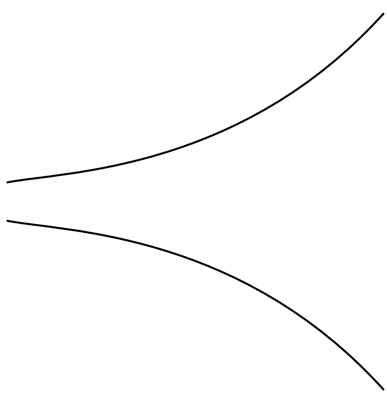


Figure 4.2 The flare shape of an exponential horn

Above a given frequency, known as the cut-off frequency, throughout the entire length of the horn, the far-field-type propagation dominates, leading to an almost totally *resistive* impedance everywhere. The cut-off frequency of an exponential horn marks a sudden transition from inefficient sound propagation within the horn to efficient sound propagation. The cut-off frequency for any given horn is dependent upon its rate of flare. As the flare rate goes up (the horn expands more rapidly) the cut-off frequency also goes up, therefore rapidly flaring horns cannot be used at low frequencies.

Physically, propagation within an exponential horn above cut-off is similar to a spherical wave of large radius, with minimal stretching pressure. Below cut-off it is similar to a spherical wave of small radius, dominated by the stretching pressure. The sharp cut-off phenomenon clearly occurs because the transition from one type of propagation to the other occurs simultaneously throughout the entire length of the horn as the frequency is raised through cut-off. The acoustic impedance at the throat of an infinite-length exponential horn is shown in Figure 4.3, which clearly illustrates that, at frequencies below cut-off, the resistive part of the acoustic impedance is zero, which means that a source at the throat can generate no acoustic power. At frequencies above the cut-off frequency, the resistive part of the acoustic impedance is close to the characteristic impedance of air: a source at the throat therefore generates acoustic power with a radiation efficiency of 100% due to the perfect match.

In practice, horns have a finite length and so, unless the mouth of the horn is large compared to a wavelength, an acoustic wave propagating towards the mouth sees a sudden change in acoustic impedance from that within the horn to that outside, and some of the wave is reflected back down the horn. A standing-wave field is set up between the forward propagating wave and its reflexion, which leads to comb-filtering in the acoustic impedance. Figure 4.4 shows the radiation efficiency at the throat of a typical finite-length exponential horn. Also shown are the radiation efficiency of a conical horn having the same overall dimensions, and that of a piston the size of the throat mounted on an infinite baffle. The frequency

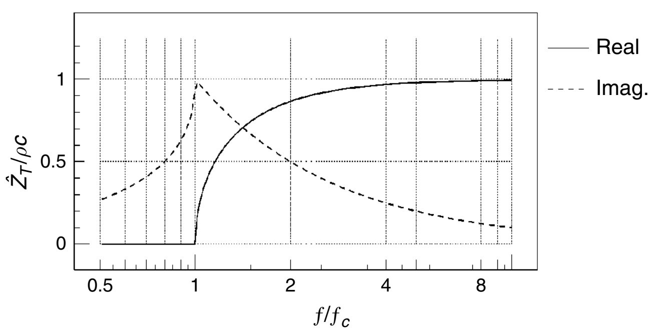


Figure 4.3 The acoustic impedance at the throat of an infinite length exponential horn f/fc is the ratio of frequency to cut-off frequency, and ρc is the characteristic impedance of air. No acoustic power can be radiated below the cut-off frequency as the real part of the acoustic impedance is zero

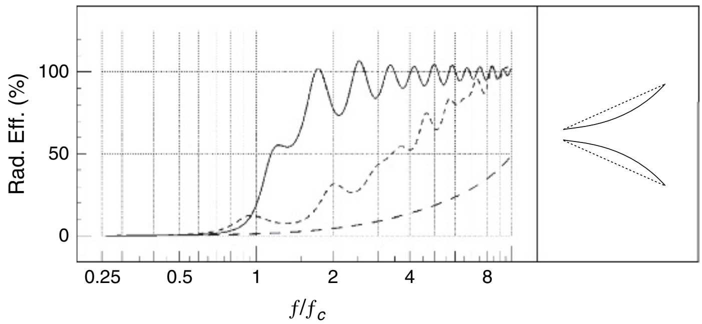


Figure 4.4 The radiation efficiency of an exponential horn compared to that of a conical horn (short-dashed line) of the same overall size. Relatively small changes in the flare shape of a horn can have a large effect on the efficiency at low frequencies. The third curve (long-dashed line) is the radiation efficiency of a baffled piston having the same size as the throats of the horns

scale is normalised to the cut-off frequency of the exponential horn. The comb-filtering, due to the standing wave field within the horn can be seen, as can the improvement in radiation efficiency of the conical horn over the baffled piston, and of the exponential horn over the conical horn at frequencies above cut-off.

The exponential horn acts as an efficient impedance matching transformer at frequencies above cut-off by giving the small throat approximately the radiation efficiency of the large mouth. The power output of a source mounted at the throat of a horn is proportional to the product of its volume velocity and the radiation efficiency at the throat; thus, a small loudspeaker diaphragm mounted at the throat of an exponential horn can radiate low frequencies with high efficiency. Below cut-off, however, the horn flare effectively does nothing, and the radiation efficiency is then similar to the diaphragm mounted on an infinite baffle. In practice, however, this seemingly ideal situation is marred somewhat by the sheer physical size of horn flare required for the efficient radiation of low frequencies.

The cut-off frequency is proportional to the flare rate of a horn, which in turn is a function of the throat and mouth sizes and the length of the horn. Therefore, for a given cut-off frequency and throat size, the length of the horn is determined by the size of the mouth. To avoid gross reflexions from the mouth, leading to a strong standing wave field within the horn, and consequently an uneven frequency response, the mouth has to be sufficiently large to act as an efficient radiator of the lowest frequency of interest. In practice, this will be the case if the circumference of the mouth is larger than a wavelength. For the efficient radiation of low frequencies, the mouth is then very large. Also, a low cut-off frequency requires a low flare-rate which, along with the large mouth, requires a long horn. By way of example, a horn required to radiate sound efficiently down to 50 Hz from a loudspeaker with a diaphragm diameter of 200 mm would need a mouth diameter of over 2 metres, and would need to be over 3 metres long! Compromises in the flare-rate raise the cut-off frequency, and compromises in the mouth size gives rise to an uneven frequency response. Reference 1 is a classic paper on the optimum matching of mouth size and flare-rate.

A radiation efficiency of 100% is not usually sufficient to yield the very high electroacoustic efficiencies of 10% to 50% quoted in the introduction of this section. However, unlike 'real' efficiency figures, which compare power output with power input, the radiation efficiency can be greater than 100% because the figure is relative to the radiation of acoustic power into the characteristic impedance of air ρ. Arranging for a source to see a radiation resistance greater than that of the characteristic impedance results in radiation efficiencies greater than 100%. A technique known as compression is used to increase the radiation efficiency of many horn drivers; all that is required is for the horn to have a throat that is smaller than the diaphragm of the driver, as shown in Figure 4.5

Assuming that the cavity between the diaphragm and the throat is small compared to a wavelength, it can be shown that the acoustic impedance at the diaphragm is approximately that at the throat multiplied by the ratio of the diaphragm area to the throat area, known as the compression ratio.

A compression ratio of 4:1 thus gives a radiation efficiency of 400% at the diaphragm. The 'trick' to achieving optimum electroacoustic efficiency is to match the acoustic impedance to the mechanical impedance (mass, damping, compliance, etc.) of the driver. If the compression ratio is too high, the velocity of the diaphragm will be reduced by the additional acoustic load and the gain in efficiency is reduced. This can, however, have the benefit of 'smoothing' the frequency response irregularities brought about by insufficient mouth size. Some dedicated compression drivers operate with compression ratios of 10:1 or more.

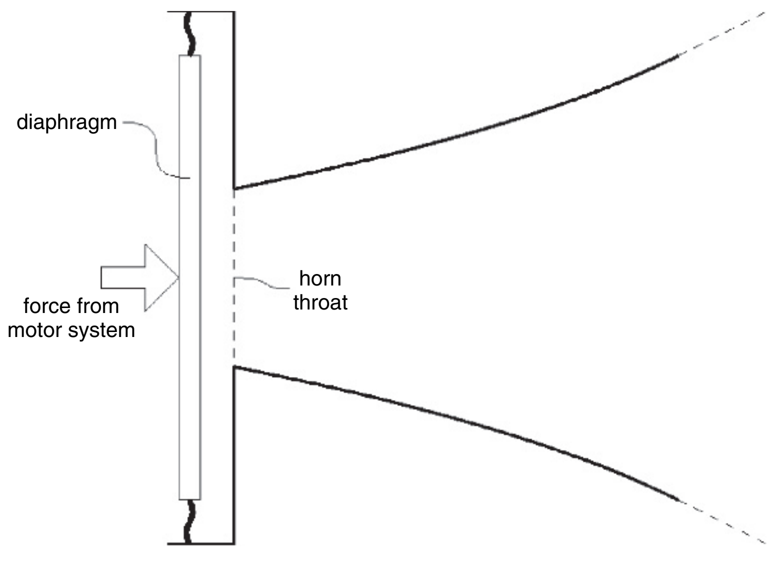


Figure 4.5 Representation of the principle behind the compression driver. Radiation efficiencies of greater than 100% can be achieved by making the horn throat smaller than the diaphragm

4.2 Directivity control

In addition to their usefulness as acoustic transformers, horns can be used to control the directivity of a loudspeaker. The directivity of a piston in a baffle narrows as frequency is raised, as was shown in Figure 1.4. For many loudspeaker applications, this frequency-dependent directivity is undesirable. In a public address system, for example, the sound radiated from a loudspeaker may be required to 'cover' a region of an audience without too much sound being radiated in other directions where it may increase reverberation. What is required in these circumstances is a loudspeaker with a directivity pattern that can be specified and that is independent of frequency. By attaching a specifically designed horn flare to a loudspeaker driver, this goal can be achieved over a wide range of frequencies.

Consider the simple, straight-sided horn shown in Figure 4.1. The directivity of this horn can be divided into three frequency regions as shown in Figure 4.6. At low frequencies, the coverage angle reduces with increasing frequency in a manner determined by the size of the horn mouth, *similar to a piston with the dimensions of the mouth.* Above a certain frequency, the coverage angle is essentially constant with frequency and is equal to the angle of the horn walls. At high frequencies, the coverage angle again decreases with increasing frequency in a manner determined by the size of the throat, *similar to a piston with dimensions of the throat.* Thus the frequency range over which the coverage angle is constant is determined by the sizes of the mouth and of the throat of the horn. The coverage angle within this frequency range is determined by the angle of the horn walls. This behaviour is best understood by considering what happens as frequency is reduced. At very high frequencies, the throat beams with a coverage angle which is narrower than the horn walls, as if the horn were not there. As frequency is lowered, the coverage angle (of the throat) widens to that of the horn walls and can go no wider. As frequency is further lowered, the coverage angle remains essentially the same as the horn walls until the mouth (as a source) begins to become 'compact' compared

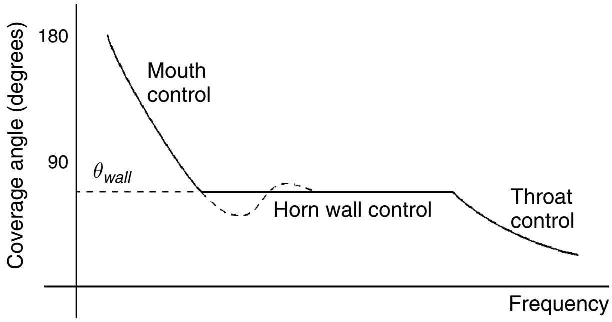


Figure 4.6 Simplified representation of the coverage angle of a straight-sided horn. At low frequencies, the coverage angle is determined by the size of the mouth, and at high frequencies by the size of the throat. The coverage angle in the frequency range between the two is fairly even with frequency, and is roughly equal to the angle between the horn walls (θ wall). The dashed line shows a narrowing of the coverage angle at the lower end of the wall-control frequency range, which is often encountered in real horn designs

to a wavelength and the coverage angle is further increased, eventually becoming omni-directional at very low frequencies. The coverage angle shown in Figure 4.6 is, of course, a simplification of the actual coverage angle of a horn. In practice, the mouth does not behave as a piston and there is almost always some narrowing of the directivity at the transition frequency between mouth control and horn wall control. A typical example of this is shown as a dashed line in Figure 4.6. Different coverage angles in the vertical and horizontal planes can be achieved by setting the horn walls to different angles in the two planes.

4.3 Horn design compromises

Sections 4.1 and 4.2 describe two different attributes of horn loudspeakers. Ideally, a horn would be designed to take advantage of both attributes, resulting in a high-efficiency loudspeaker with a smooth frequency response and constant directivity over a wide frequency range. However, very often a horn designed to optimise one aspect of performance must compromise other aspects. For example, the straight-sided horn in Figure 4.1 may exhibit good directivity control but, being a conical-type horn, will not have the radiation efficiency of an exponential horn of the same size. The curved walls of an exponential horn, on the other hand, do not control directivity as well as straight-sides horns. Early attempts at achieving high efficiency *and* directivity control in one plane led to the design of the so-called *sectoral horn* or *radial horn* shown in Figure 4.7. In this design, the two side walls of the horn are straight, and set to the desired horizontal coverage angle. The vertical dimensions of the horn are then adjusted to yield an overall exponential flare. Whereas the goals of high efficiency and good horizontal directivity control can be achieved with a sectoral horn, the severely compromised vertical directivity can be a problem. Given that a minimum mouth dimension is required for directivity control down to a low frequency, setting the horizontal and vertical walls to

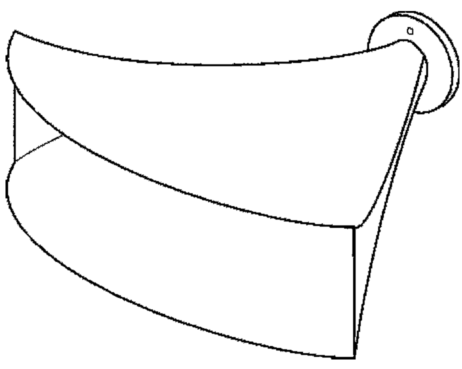


Figure 4.7 The sectoral (or radial) horn

The walls controlling the horizontal directivity are set to the desired coverage angle. The shape of the other two walls is adjusted to maintain an overall exponential flare, resulting in less-than-ideal vertical directivity

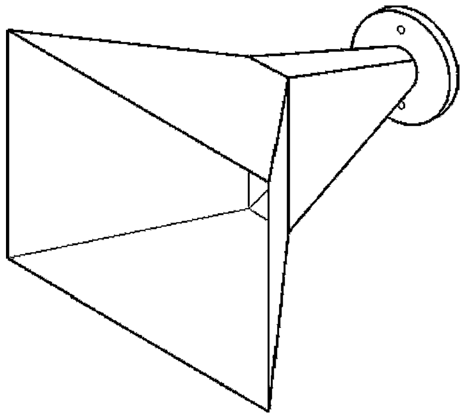


Figure 4.8 The constant directivity horn
Different horn wall angles in the two planes can be achieved using compound flares. Sharp discontinuities within the flare can set up strong standing-wave fields, leading to an uneven frequency response

different angles, for example 90 degrees by 60 degrees, means that different horn lengths are required in the two planes. To overcome this problem, later designs used compound flares[2] so that the exit angles of the horn walls can be different in the two planes, but the mouth dimensions and overall horn length remain the same. The so-called *constant directivity horn* (CD) is shown in Figure 4.8. The sudden flare discontinuities introduced into the horn with these designs result in strong standing wave fields within the flare which can compromise frequency response smoothness. In fact, this is true of almost any flare discontinuity in almost any horn. Modern public address horn designs employ smooth transitions between the different flare sections and exponential throat sections to achieve a good overall compromise, but constant directivity horns, because of the reflexion problems, tend not to be used on the highest fidelity loudspeaker systems.

The control of directivity down to low frequencies requires a very large horn. For example, in a horn designed to communicate speech, directivity control may be desirable down to 250 Hz at a coverage angle of 60 degrees.This can only be achieved with a horn mouth greater than 1.5m across. The same horn may have an upper frequency limit of 8 kHz, which needs a throat no greater than 35 mm across. Maintaining 60 degrees walls between throat and mouth then requires a horn length of about 1.3 m. Attempts to control directivity with smaller devices will almost always fail.

4.4 Non-linear acoustics

In the vast majority of studies in acoustics, and loudspeakers in particular, the acoustic pressures and particle velocities encountered are sufficiently small that the processes of sound radiation and propagation can be assumed to be linear. If a system or process is linear, then there are several rules that govern what happens to signals when they pass through the system or process. These rules include the *principal of superposition*, which states that the response to signal (A + B) is equal to the response to signal (A) + the response to signal (B). Most of the analysis tools and methods, such as Fourier analysis and the frequency response function, rely entirely on the principal of superposition, and hence linearity. When a system or process is non-linear, the principal of superposition no longer applies, and the usual analysis methods cannot be used. In this section, the conditions under which acoustic radiation and propagation may become non-linear are discussed, along with some examples of the degree of non-linear acoustic behaviour encountered in loudspeakers.

The speed of sound in air is dependent upon the thermodynamic properties of the air.

An acoustic wave consists of alternate positive and negative pressures above and below the static pressure and, as this is an isentropic process, the relationship between the instantaneous pressure and the density is progressively more non-linear as SPLs rise.

In linear acoustic theory, the relationship between pressure and density is assumed to be linear, which is a good approximation if the changes in pressure are small compared to the static pressure. A linear relationship between pressure and density means that the temperature does not change, so neither does the speed of sound. However, when the changes in pressure are significant compared to the static pressure, changes in instantaneous temperature, and hence the speed of sound, cannot be ignored.

In addition, when an acoustic wave exists in flowing air, the speed of propagation is increased in the direction of the flow, and decreased in the direction against the flow; the acoustic wave is 'convected' along with the flow. Although steady air flow is not usually encountered where loudspeakers are operated, the particle velocity associated with acoustic wave propagation can be thought of as an alternating, unsteady flow. Again, if the particle velocities are small compared to the speed of sound, the effect can be neglected, but in situations where the particle velocities are significant compared to the speed of sound, the dependence of the speed of propagation on the particle velocity must be taken into account.

The result of all of this is that the speed of propagation increases within increasing pressure and particle velocity, and decreases with decreases in pressure and particle velocity. For a plane progressive wave, positive pressures are accompanied by positive particle velocities, and the speed of propagation is therefore higher in the positive half-cycle of an acoustic wave than it is in the negative half-cycle. The positive half-cycle then propagates faster than the negative half cycle and the waveform distorts as it propagates. Figure 4.9 shows the distortion, known as *waveform steepening*, that occurs in the propagation of sound when the acoustic pressures are significant compared to the static pressure and/or the acoustic particle velocities are significant compared to the static speed of sound.

4.5 Examples of non-linear acoustics in loudspeakers

At the sound levels typically encountered when loudspeakers are operated, the effect of pressure and particle velocity on the instantaneous speed of

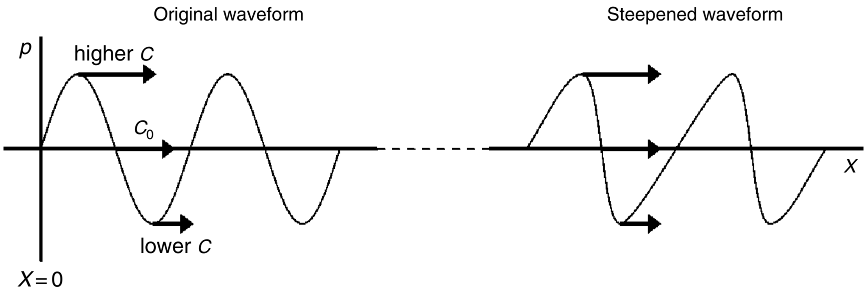


Figure 4.9 Waveform steepening due to acoustic pressures that are significant compared to the static pressure and/or acoustic particle velocities that are significant compared to the speed of sound, c_0

The steepened waveform is no longer a sine wave, and therefore *must* contain harmonic distortion

sound is so small as to be negligible, and the resultant linear approximation is sufficiently accurate. However, there are some situations where this is not the case. Two common examples are the high sound pressures in the throats of horn loudspeakers, and the high diaphragm velocities of long-throw low-frequency drive-units.

When horn loudspeakers equipped with compression drivers are used to generate high output levels, the pressure in the throat of the horn can exceed 160 dB SPL, with even higher levels at the diaphragm. Sound propagation is non-linear at these levels and the acoustic waveform distorts as it propagates along the horn. If the horn flares rapidly away from the throat, then these levels are maintained only over a short distance and the distortion is minimised. Horns having throat sections that flare slowly suffer greater waveform distortion (it is interesting to note that the rich harmonic content of a trombone at fortissimo is due to this phenomenon). Nevertheless, investigations have shown that the distortion produced by high-quality horn loudspeakers only exceeds that from high-quality conventional loudspeakers when the horn system is producing output levels beyond the capability of most conventional loudspeakers.

The use of small, long-throw woofers in compact, high-power loudspeaker systems can also introduce non-linear distortion. The power output of a loudspeaker diaphragm is proportional to the square of the volume velocity of the diaphragm, so for a given sound power output the required diaphragm velocity is therefore proportional to the inverse of the diaphragm area. Consider two loudspeakers, one with a diaphragm diameter of 260 mm, the other with a diameter of 65 mm. In order to radiate the same amount of acoustic power at low frequencies, the smaller loudspeaker requires a velocity of 16 times that of the large loudspeaker, as it has 1/16 of the area. The rms velocity of the large loudspeaker when radiating a sound pressure level of 104 dB at 1 m at a frequency of 100 Hz is approximately 0.5 m/s. The same sound pressure level from the smaller loudspeaker requires 8 m/s. Whereas 0.5 m/s may be considered insignificant compared to the speed of sound (= 340 m/s), 8 m/s represents peak-to-peak changes in the speed of sound of around 8%.

A secondary effect, which is a direct consequence of particle velocities that are significant compared to the speed of sound, is the so-called Doppler distortion. If, at the same time as radiating the 100 Hz signal above, the small loudspeaker were also radiating a 1 kHz signal, the cyclic approach and recession of the diaphragm due to the low-frequency signal would frequency modulate the radiation of the higher-frequency signal by approximately 70 Hz. Hence long-throw woofers which are to be used at high SPLs need to cross over to mid-range drivers at relatively low-frequencies if Doppler distortions are to be avoided.

4.6 Practical horns in studios and homes

The previous sections have outlined in some depth the concept of the way in which horn loudspeakers work. This has been necessary before any meaningful discussions of horns can be undertaken because they are so poorly understood by the vast majority of people who use them. The normal state of affairs is that discussions about horns are based on hearsay, bad experiences due to the abundance of bad designs, and a widespread misapplication of good designs. In the worlds of public address and sound reinforcement, the directivity pattern control of horns is a very useful tool. The benefits of good pattern control normally greatly outweigh the more subtle aspects of their sound quality. Unfortunately, in the past decades, many horns which have been used in studio monitoring and domestic hi-fi systems have been based on sound reinforcement technology, and there has also existed a lack of understanding about how horn cross-sectional shapes can have detrimental repercussions on sonic purity.

The non-linear acoustics discussed in Section 4.4 can be a problem in horns. However, at the SPLs found in most music recording control rooms and homes, the non-linear propagation region is only occasionally reached on transient peaks at loud listening levels, and the audibility of such distortion on transient signals is unlikely to be very great. In fact, when other types of drive units are pushed to the same levels, they too may well be suffering from their own forms of non-linearities. Misunderstanding and poor designs have, over the years, created a lot of bad publicity for horns, yet some horns, such as the famous Tannoy 15 inch Reds and Golds have been the very antithesis of all that has been said against horns. As previously mentioned in Chapter 2, above about 1 kHz, these loudspeakers are nothing more and nothing less than compression drivers and axisymmetric horns. It has been quite remarkable how so many people have been heard to say that they cannot work with horns and yet have quoted the Tannoys amongst their favourite loudspeakers. Figure 4.10 shows a studio monitoring horn of the highest quality. It bears a remarkable similarity to the high frequency section of the Tannoy Dual Concentric shown in Figure 5.10. The development work which led to the horn in Figure 4.10 took four years to complete. The following summary is taken from Reference 3, which was based on the aforementioned research.

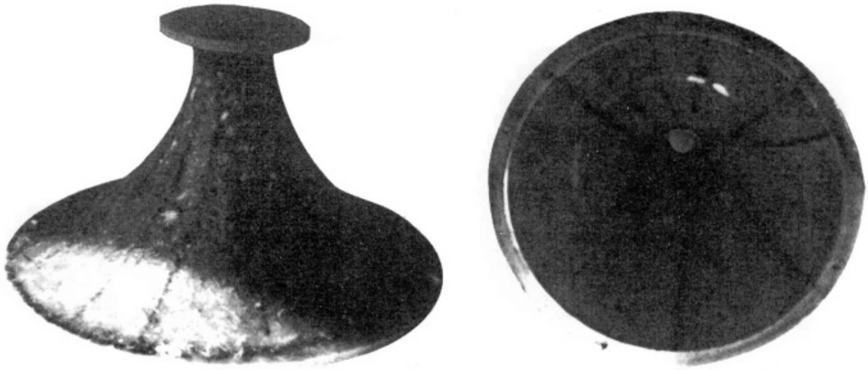


Figure 4.10 Axisymmetric horn geometry
The AX2 (see also Figures 8.6(a) and 8.16)

4.7 Implications for practical horn design parameters

By the very careful choice of design parameters and construction, horns can be produced for use above 1 kHz, or thereabouts, which approach the performance of electrostatic loudspeakers with their very low levels of deviation from their intended amplitude and phase (and hence time) responses. It would appear, however, that there are finite practical limits to the performance ranges over which horns can produce near optimum results:

1) The cut-off frequency of a horn is a function of its rate of flare. A low cut-off frequency demands a slow rate of flare.
2) If 'horn-like' sound characteristics are to be avoided in a practical horn, the length should not exceed 12 inches (300 mm) or thereabouts.
3) Taking 1 and 2 together, if a horn has a low flare rate, and cannot exceed 12 inches in length, then given a 1 inch throat diameter, and, say a 250 Hz cut-off frequency, the horn will inevitably have a small mouth area. There will consequently be an abrupt change in cross-sectional area when it meets the outside air. Samples 4 and 11 in Figure 4.11 highlight this point, [both were short, low flare-rate horns with small mouths] showing poor throat impedance linearity, or smoothness, especially near cut-off. Subjectively, although almost always being grouped with the direct radiators in the listening tests, as musical reproduction devices they were not considered smooth, flat, or natural, despite not sounding typically horn-like.
4) In order to achieve a smooth and trouble-free mouth termination, from a 1 inch throat of a horn not exceeding 12 inches (300 mm) in length (diaphragm to mouth), a mouth diameter of around 12 inches would seem to be the smallest practical size. This dictates a flare rate which results in a cut-off frequency in the order of 1 kHz, but can yield exceptionally smooth performance through cut-off if carefully designed; even allowing use *through* cut-off, and utilising the acoustic roll-off as part

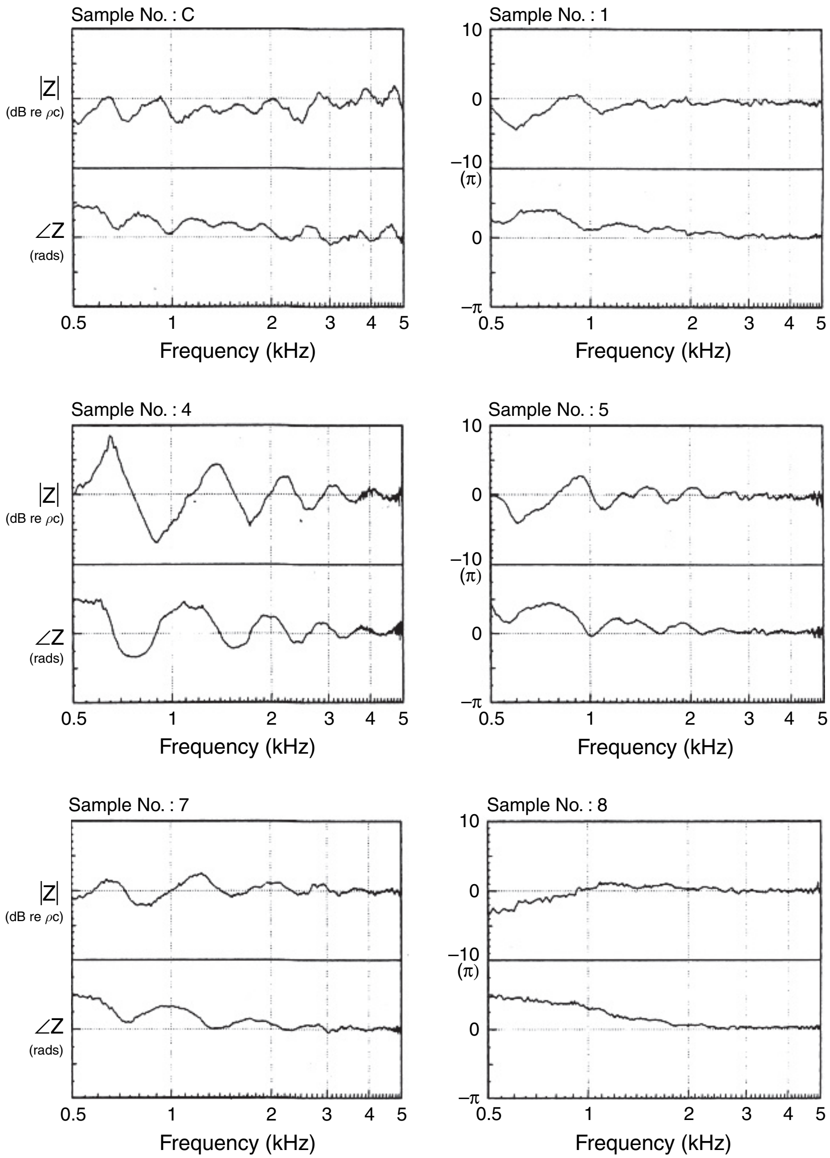


Figure 4.11 The logarithmic throat impedance plots of a selection of different mid-range horns. (This figure is continued on the next page)

of the electroacoustic crossover (see Figure 4.12). [Sample 8 in Figure 4.11 shows the throat impedance plot of such a device.]

5) To minimise internal disturbances which can cause disruption to both the on- and off-axis responses, all corners, angles and obstructions should be removed, rendering axial symmetry and smoothly contouring surfaces. Figure 4.11 shows the logarithmic throat impedance plots of all

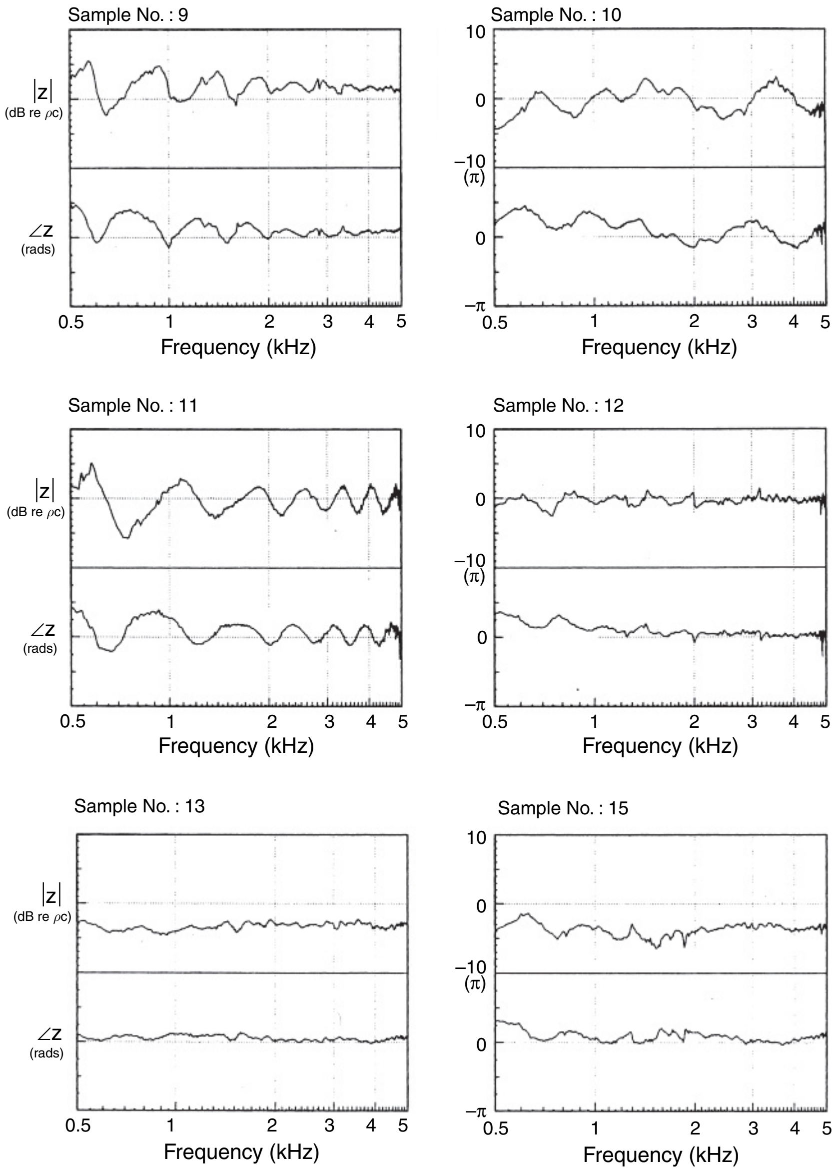


Figure 4.11 Continued

but one of the horns used in the tests. The vastly superior characteristics of the AX2 (Sample 8) can be readily seen.

6) ‘Squashing’ the axisymmetric shape into an ellipse would perhaps allow some change in directivity pattern without undue disturbance of the time response.

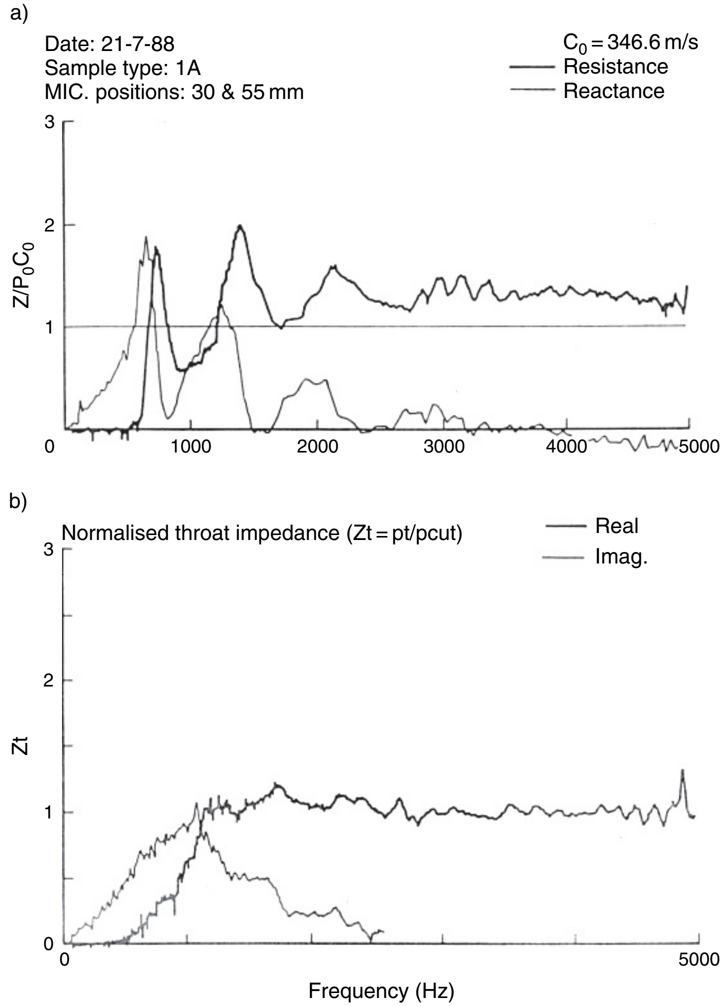


Figure 4.12 a) Throat impedance of a horn which has been widely used in studio monitor systems. The systems were generally considered to be typical of horn-loaded systems b) Throat impedance of the AX2 horn, showing an absence of reflexion-induced irregularities and a smooth impedance through cut-off. Systems using this horn have generally *not* been considered to sound horn-like

Sample 8, mated with the TAD 2001 drive unit produced what would seem to be a near optimum response in terms of both phase and amplitude (and hence time), smooth directivity, a very smooth overall performance from 1 kHz to beyond 20 kHz, and was deemed to be very musical, natural, transparent, and definitely not horn-like. It is interesting that in physical dimensions, though not in its drive system, it strongly resembles the Tannoy Dual Concentrics from around 40 years before. The fact that the shape and size were so similar to the Tannoy 15 inch Dual Concentric HF horn served to explain why the latter had enjoyed 40 years of use without being considered to sound horn-like. It would appear that the Tannoy, all those years ago, defined the physical limits for accurate performance, beyond which horns will begin to run into trouble.

Whether the engineers at Tannoy knew all of this at the time when the Dual Concentrics were first designed, or whether some of them merely ‘saw the logic’ of using a duly contoured bass cone for the horn of a co-axial system, and recognised that it sounded good, maybe we shall never truly know. In general, however, there is now no doubt that carefully designed horns and drivers can produce both sonic and measured performances as good if not better than the finest dynamic direct radiators, without any hint of a ‘horn-like’ sound. Indeed, to run a truly seamless 1 kHz to 20 kHz response within tight limits, whilst producing a very smoothly controlled directivity pattern (courtesy of a horn emanating a section of a spherical expanding wave, unlike a piston), a well designed horn and driver combination can be superior to the vast majority of direct radiators. If very high sound pressure levels are added to the list of requirements, there are few alternatives to horns.

4.8 Summary of results

The tests leading to the above conclusions are fully documented, eminently repeatable, and open to inspection. The initial findings suggest that short horns can be produced having high efficiency, wide frequency range and benign distortion levels, which are not sonically horn-like, and can be grouped as audibly similar to typical direct radiators. Much of the audible similarity of loudspeakers would appear to be in their time histories, and where a mouth reflexion effect of a horn is in the same order as any inherent reflexions in direct radiator units, then general audible grouping can be expected. Long horns produce longer reflexion delays, and sonically tend to group together, whilst electrostatics group together due to their rapid and accurate transient responses. In fact, much of the general audible similarity of all loudspeaker drive units of similar frequency range and general overall quality, irrespective of generic type, lies not in the non-linearities, nor solely in the pressure amplitude response, but in the time domain response as specified by the linear distortions of the convolution of the amplitude and phase responses.

The AX2 was developed for the listening tests which led to the above conclusions at the Institute of Sound and Vibration Research (Southampton, UK) in 1989[4,5]. It was designed to have the minimum number of characteristics which previous research had suggested would lead to a horn-like ‘honky’ sound. The AX1 – Sample 4 in Figure 4.11 – was designed to maximise the horn-like characteristics. Neither the AX1 nor the AX2 were modelled on other loudspeakers, but were designed from first principles. As it turned out, the AX1 was physically remarkably similar to a horn of widespread use and of well-known manufacture which was often criticised for its harsh, nasal characteristics, and the AX2 resembled, physically, the sweeter sounding Tannoys. As mentioned in the earlier sections of this chapter, when mouths are designed from the requirement of associated components and directivity control only, the acoustic termination to the outside air is often very poor. This gives rise to strong reflexions from the mouths, which in turn give rise to the roller-coaster impedance plots typical of so many of the horns in Figure 4.11.

Conversely, the AX2 was designed to minimise the impedance irregularities, which was clearly successful (Sample 8 in Figure 4.11) but the size and directivity were the natural result of the design, and could not be controlled. Larger mouths could be smooth down to lower frequencies, but could give rise to difficulties in closely locating the adjacent drivers at crossover points. The practice of mounting high frequency horns and/or drivers on the central axis of the larger horn may ruin the mouth termination of the latter. The AX2 just about defines the practical limit for a MF/HF horn if the highest audio quality is the sole aim.

The fact that such strong evidence exists which discourages the use of horns below 1 kHz for the highest quality audio systems helps to explain why so many horn-loaded systems have received bad reviews.

Figure 12.18 shows a low frequency horn in a discotheque. This horn is flat down to 20 Hz. The floor mounting provides an acoustic mirror, so the mouth is effectively $4\,m \times 2\,m$ and the length is about 2 m. The loudspeaker cabinets behind the $1\,m^2$ throat are over a metre deep, and in the throat are four 18 inch drivers. The sheer size of this horn precludes its use in studios or homes, but its size cannot be reduced without compromising its performance. In the discotheque it will produce 140 dB SPL, but the power level also has nothing to do with its response. Even if it was only required to produce 80 dB it would still need to be the same size. When a large array of relatively small horn-loaded bass cabinets are seen at a rock concert, the mouth areas combine to form one, composite mouth, equal to the sum of the sizes of the individual mouths. It is no use buying just one of the cabinets for a small venue and expecting the same response. So, the large arrays are not just for more volume, but also serve to extend the low frequency response by augmenting the overall mouth size and making a better mouth termination to the outside air.

4.9 General horn characteristics

The resistive load which the air presents to a diaphragm when loaded by a horn gives rise to rapid transient responses. Subjectively, good horns sound ‘fast’. The high sensitivity which is characteristic of most horn designs leads to small voice coil currents in powerful static magnetic fields. This tends to lead to less Bl profile distortion, because the integrity of the static magnetic field against which the voice coil field must push is much less disturbed as compared to a high voice coil current distorting the field of a weaker magnet in a heavier diaphragm, low sensitivity design. Despite many common beliefs, horn-loaded loudspeakers are capable of very low distortion reproduction if only the appropriate rules are respected and they are not pushed beyond the limits where the air, itself, is non-linear. In too many cases horns have been used which are too small for the frequency range in which they have been operated, and flare shapes have been used

Figure 4.13 The Meyer HD2, using an axisymmetric horn and a non-compression driver

Figure 4.14 A Genelec 1038A loudspeaker system which employs vestigal horns, or 'waveguides', to load the mid-range and high-frequency drivers

for their directivity performance with little consideration for the need for low colouration through the use of smooth contours.

Figure 4.13 shows a loudspeaker system which effectively horn loads a dome tweeter. In this case no compression driver is used. This is only one step away from the design shown in Figure 4.14, where vestigial horns are used as 'waveguides' (another name for horns) in a very popular studio monitoring loudspeaker. In fact, a direct radiator on an infinite baffle is mounted on a 180 degree horn of infinite flare rate. There is therefore no dividing line between direct radiation and horn loading. They are part of the same continuum which extends from the infinite 180 degree baffle to the infinite parallel pipe. It therefore follows that if somebody says that they do not like horns, one has to ask the question "Which horns?" The scope of what constitutes a horn is enormous, and so is the scope of the compromise points for any designs. Equally enormous, therefore, is the range of possibilities for making inappropriate compromises. Unfortunately, what we generally consider to be horns are relatively complex, precise devices, and hence are normally quite expensive when compared to direct radiators for use at relatively low SPLs. So, when corners are cut in the production process in attempts to reduce the costs, the sound quality usually suffers. Again, horns *in general* cannot be criticised because of the failings of badly engineered, cheap products.

The Tannoy Dual concentric concept has already been noted as not sounding horn-like, however the principle does suffer from two drawbacks. The high frequency horn is the cone of the bass driver, and it therefore moves, axially, with the low frequency signal. At high SPLs, the movement, peak to peak, can be a significant proportion of a wavelength at the highest audible frequencies, so some modulation effects are only to be expected, the degree of which are both level and frequency dependent. The Tannoys therefore tend to sound sweeter at low levels than at high levels. The second problem is that the voice coil gap, seen clearly in Figure 5.10, is a discontinuity in the throat of the horn, which on the larger devices can be a source of reflexions whose amplitude again depends upon the excursion of the bass cone. Nevertheless, despite these drawbacks, the concentric nature of the high and low frequency drivers facilitates the engineering of a relatively seamless crossover and a radiation pattern uniformity which many people consider to be sufficiently beneficial to offset the drawbacks of the designs, especially when listening levels are relatively low.

Another characteristic of the Tannoy design is that it uses a low-compression driver. As described in Section 4.1, high compression ratios are often used in order to increase the sensitivity of the driver by increasing the radiation resistance. However, if a high frequency horn loudspeaker is only to be used with a relatively inefficient low frequency driver, and at relatively low SPLs, there would seem to be no reason for compromising its sonic purity by using unnecessarily high compression. The choice of Tannoy has been to use low compression, low sensitivity, high frequency horn systems in most of their Dual Concentric designs.

4.10 Phasing plugs

A horn loudspeaker is shown in Figure 4.13 in which the diaphragm is mounted directly at the throat of the horn. No compression is used in this design as the diaphragm and throat are of the same diameter. By contrast, Figure 4.5 shows the concept of compression, where the diaphragm area is much greater than the throat area. Unfortunately, except for use as

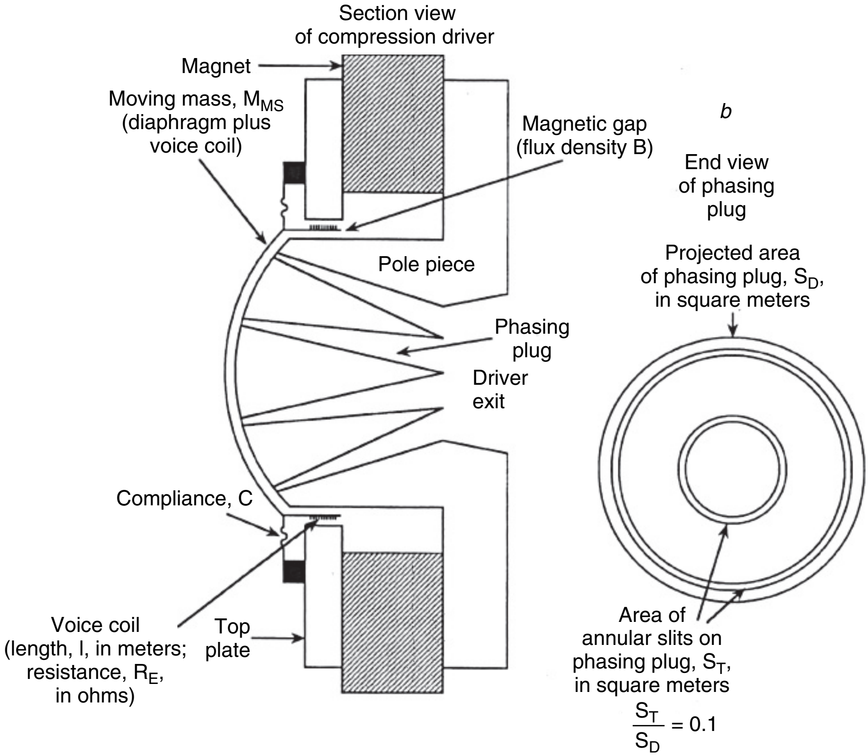


Figure 4.15 A typical phasing plug (courtesy of JBL Professional)

an electric klaxon or evacuation alarm, the design of Figure 4.5 would be of little use, because the pressure fluctuation from the perimeter of the diaphragm would take longer to reach the throat of the horn than the fluctuation from the centre of the diaphragm. The different distances from the different parts of the diaphragm to the throat of the horn would give rise to phase differences and severe losses at high frequencies. To overcome this problem, compression drivers use *phasing plugs* as shown in Figure 4.15.

There are various types of phasing plugs, using either a series of tubes or axial slits to connect the various parts of the diaphragm to the horn throat via equal length pathways. The geometry of these needs to be very carefully controlled, because with a wavelength of only about 17 mm at 20 kHz, a path-length difference of only about 8 mm would reverse the polarity of the wave at the exit, and cancellation would result.

Because of the extreme sound pressure levels that can exist in the tubes, they represent one of the few situations in audio where actual air *flow* can result, leading us into the domain of aerodynamics rather than acoustics. Care must therefore be taken to avoid sources of turbulence, which can give rise to modulation dependent noises and gross distortion. The small size of the phasing plug tubes can also make visco-thermal losses significant.

4.11 Acoustics lenses

Figure 4.16 shows a selection of acoustic lenses, also known as a slant plates, chip-cutters, crinkle plates and pepper-pots. These were in vogue in the late 1960s and the 1970s and were developed as solutions to the problem of how to better match a short, slow-flaring horn to the outside air. They first came into use in the 1950s, in cinema loudspeaker systems, but took some time to find their way into the music recording studios. In all cases, the path through the centre of the lens is shorter than the path through the outer sections, and thus the wavefront was bent into a wider directivity pattern. Acoustic lenses worked well from a technical point of view, but could give rise to audible colouration. The time during which JBL used them for their studio monitor systems was quite short, but the high respect for JBL and the high profile of their users led to other companies copying them, and using them for many years after JBL had already abandoned them for studio quality monitoring.

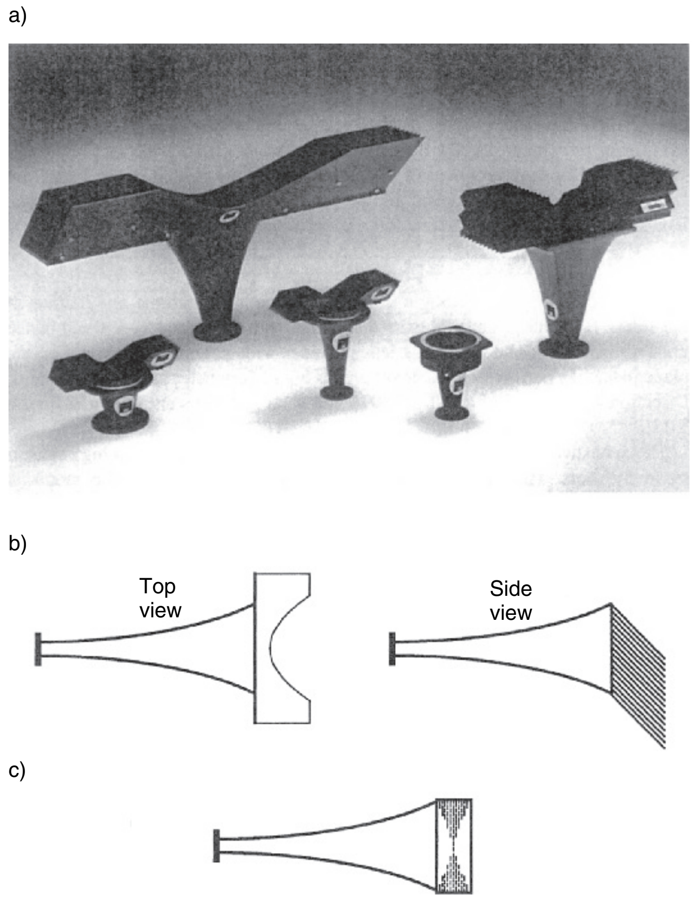


Figure 4.16 A family of acoustic lenses (courtesy of JBL Professional)

4.12 Horn types

Already mentioned in Section 4.3, and illustrated in Figure 4.7 was the radial horn, also known as the sectoral horn. The first name comes from the fact that the straight side walls in one direction form the radial lines of a circle, which meet at an imaginary apex behind the throat. They also have the appearance in plan view of a sector of a 'pie-chart', hence the name 'sectoral'. For the highest quality subjective listening they have three basic problems. First, the cross-sectional shape must somehow be made to match the circular orifice of a compression driver. In some examples, even expensive ones, this was rather abrupt, leading to uncontrolled wave-expansion in the sensitive throat area of the horn. Such practices are simply begging for colouration due to reflexion problems. It has been shown that *any* discontinuity in the flare, especially near the throat, will give rise to reflexion-induced colouration [3,5]. Figure 4.17 shows a cepstrum plot of an AX2 horn mated to an Emilar EK175 driver. The cepstrum plot is useful for identifying echoes or reflexions in signals or responses. [See Chapter 9.] The flare rate of the throat tube of the EK175 was not a precise match to the flare rate at the throat of the horn, even though both were perfectly circular, and the series of reflexion spikes before 2 ms was the result. When the AX2 was mated to the TAD TD2001 compression driver, the flare from throat tube to horn was continuous, and the reflexions disappeared.

The second problem with the radial horns is that the geometrical discontinuities at the junctions between the top, sides and bottom of the horns, even though not affecting the flare rate, can lead to off-axis colourations at normals to the discontinuities. In non-reflective environments this may not be a problem, but it can lead to coloured reflexions in more lively environments. The third problem involves the mouth termination, which is often achieved by means of rounded lips which project from the front

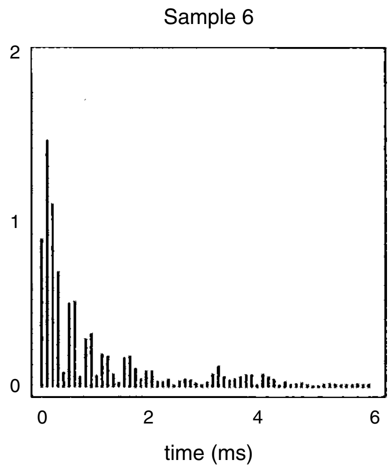


Figure 4.17 Power cepstrum of the Emilar EK175/AX2 combination. The series of diminishing echoes in the first 2 ms result from the flare-rate changes at the driver/horn junction

baffle. These can lead not only to diffraction problems from the horn itself, but also from the obstructions which they may cause to the expanding wavefronts from the other drivers of the loudspeaker system. Dividers, which sectionalise some of these horns for improved directivity control also can act as diffraction sources. The general rule for horns for the highest fidelity is that *nothing* must disturb the flare rate, the flare should blend smoothly into the baffle, and no geometrical changes should exist other than the defined expansion rate. All of these things can detract from the sonic purity of the horn.

Diffraction horns are another type of horn. They have mouths that are wide in one direction and narrow in the other, the narrow dimensions relying on diffraction from a relatively sharp edge to widen the directivity. Unfortunately, diffraction necessarily introduces reflexions, and, as has just been stated in the last paragraph, for the highest fidelity, everything should be done to avoid reflexions in the flare of a horn. The constant directivity/uniform coverage horns also rely on the diffraction principle for the wide directivity of the higher frequencies, so the same comments apply. As discussed previously, attempts to take tight control of directivity come at a cost in terms of other aspects of horn response. In this case, the evenness of the throat impedance suffers from the reflexions given rise to by the diffraction. In general, when listening in rooms with highly controlled acoustics, the reflexion-induced colouration resulting from diffraction horns *is* noticeable. The exception to this is on very high frequency loudspeakers, say 7 kHz and above, where the reflexion delays are very rapid and the sensitivity to this type of colouration at such high frequencies is minimal.

Multicellular horns are rarely used nowadays, neither in studios nor cinemas. A multicellular horn, as shown in Figure 4.18, is a grouping of smaller horns which is designed to deliver an even frequency distribution over a desired area by means of the use of many narrow horns of uniform coverage angle. The mouth sizes sum, but unlike in the case of a radial horn of the same sized mouth, the high frequencies are less likely to loose contact with the walls because the individual horns are so narrow. Nevertheless, similar to the situation with the radial horns, the question arises as to how to match the throats of the compound horn to the throat of the circular driver without any discontinuities. However, some modern multicellular horns combine the horn cells with the phasing plug, thereby eliminating problems caused by the manifolds used in earlier designs. For a good compromise between sound quality and directivity control, a well-designed multicellular horn is hard to beat.

4.13 Materials of construction

There should not really be sonic differences between horn construction materials as long as they are sufficiently rigid and highly damped. However, if these conditions are not met, the horns can be excited into resonance, particularly noticeable on transient signals. Aluminium, solid wood, plywood, plastics and glass-fibre are all commonly used materials. The horn

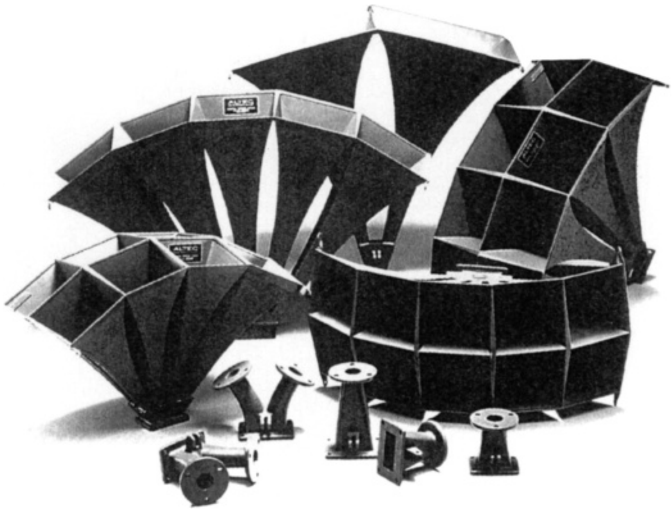


Figure 4.18 A family of multicellular horns showing the throat adaptors which, after changing section from round to square, must adapt the single driver to the multiple horn throats. This is very difficult to achieve in any precise manner (photo by Altec)

shown in Figure 8.2 is made from a specially selected Japanese apitong plywood, which is dense and highly damped. The horn shown in Figure 8.6(a) is made from glass-fibre which has been loaded with powdered slate. Some metallic horns have exhibited characteristic rings, but these have rarely found their way into studio monitoring or audiophile hi-fi loudspeakers. Nevertheless, in musical instrument amplifier use they have been used with no detrimental effect.

In many cases, the materials which have been used to make horns have been chosen for reasons such as ease of manufacturing or conditions of use. Massively heavy horns, for example, would be a poor choice for mobile sound reinforcement equipment where the benefits of their subtleties would be totally lost, so lighter weight materials have been the norm. Complex shapes have traditionally been easier to mould than to machine, so they have tended to be made from mouldable materials for manufacturing reasons, rather than for acoustic reasons. When material choices have been made for purely acoustic reasons, the results, such as the horns made from apitong plywood, have often been very expensive, and are actually only found in limited quantities. One result of this situation is that relatively few people have experienced what horns can achieve when the manufacture and marketing constraints are removed. Horns have suffered from a terrible history of compromise.

4.14 Vestigial horns and 'waveguides'

Shown in Figure 8.6(c) is the sculpted baffle of a large Genelec monitor system. The fact that Genelec refer to them as Directivity Control Waveguides (DCWs) suggests that their prime design function is just that, but nevertheless they are horns. In the Genelec case, they only slightly increase the sensitivity of the drive unit. The principal difference between the AX2 in Figure 4.10 and the DCWs is that the compromise points are different. The AX2 aims at increasing sensitivity, and the waveguide role is secondary, whereas the priorities are reversed for the DCW. It is also true that the DCW does not use a compression horn, yet neither does the high frequency horn shown in Figure 4.13, which places the diaphragm of a tweeter somewhat similar to that which is used for the high frequencies in Figure 8.6(b) at the throat of a horn similar to the AX2. In the latter case, the extra sensitivity afforded by a compression driver was not necessary for a system so small as the Meyer HD2. The system in Figure 8.6(b) uses a titanium dome mid-range unit which is clearly mounted in a short horn for reasons of improved sensitivity *and* directivity control.

To many people, the mid-range loudspeakers shown in Figure 8.6 a) and d) are horn loaded, and the systems in Figure 8.6(b) and (c) are not horn loaded, yet the truth is that they are *all* horn loaded at the mid and high frequencies. In Morfey's 'Dictionary of Acoustics' [6] the definition of a horn is:

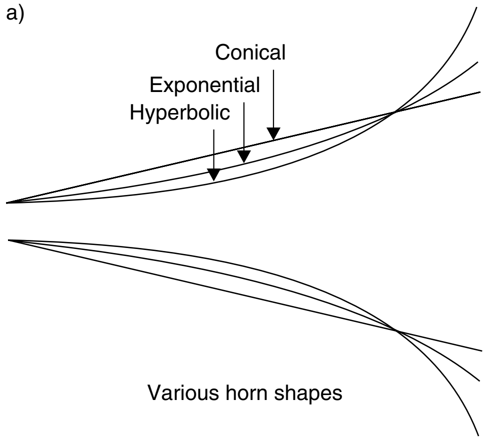
> '– a waveguide whose properties are arranged to vary monotonically from one end to the other, in order to produce a resistance in the lowest order waveguide mode.... that is many times larger at the driver end than at the opposite end'.

How many times larger is not specified. Therefore, by definition, a horn is a waveguide, and the continuum shown through Figures 8.6(c), 8.6(b), 4.13, 8.7 (the low compression Tannoy), 8.6(a) and 8.6(d) demonstrates a very wide range of increasing loading, but without any differences in the operating principles of the waveguide/horn effects. Even a conventional loudspeaker cabinet placed on the floor in the corner of a room is being loaded by a triangular cross-section low frequency horn, as shown in Figure 7.9.

Sculpted baffles are therefore nothing magic, but are based on sound acoustic principles. However, they were late arriving because until the advent of computer modelling and CNC (computer numerically controlled) machinery, their precise design and construction was difficult. Despite their complex shape, the desired overall flare rate must be maintained up to the point where they blend into the front panel.

4.15 Flare rates

Although the majority of horns are based upon exponential flares for the reasons given in Section 4.1, other flare rates are also used. Hyperbolic horns appear from time to time, but catenoidal and hypex horns have also been used. In fact the horn shown in Figure 4.10 begins its flare on the conical side of exponential, then passes through exponential. The design aim was to maintain the optimum phase response in the expanding wave



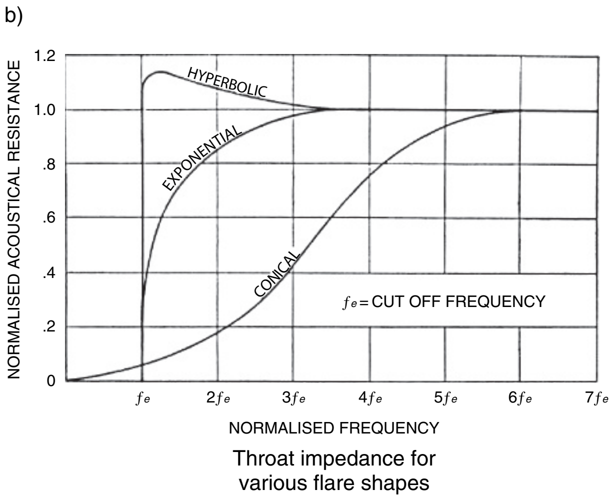


Figure 4.19 Comparison of flare geometrics. a) An infinite variety of shapes is possible between conical and hyperbolic. b) The drawing shows how the throat impedances for the three main types of horn varies with frequency

leaving the mouth of the horn. The relationship between the three most common shapes can be seen from Figure 4.19.

The throat cut-off frequency is determined by the rate of expansion of the horn. Doubling the horn length for a given mouth and throat size will halve the flare rate and lower the cut-off frequency by an octave. Halving the horn length will double the cut-off frequency, (raise it by an octave) if the throat and mouth sizes are kept constant.

As can be seen from Figure 4.19(b), exponential flares begin to loose their efficiency well above the normal cut-off frequency, which means that they are normally only used above about half an octave above cut-off. The hyperbolic horns function almost all the way to the theoretical cut-off, but the abrupt nature of their response fall-off can give rise to phase distortions which can lead to coloured responses. Nevertheless, for efficient public address applications this may not be a problem, and their amplitude benefits can be exploited.

At low frequencies, some loudspeaker cabinets use folded horns, but it is very difficult to maintain a constant flare rate around all the folds, so reflexions and resonances can be problems and response flatness can be difficult to achieve. Some rather complex shapes have historically been applied to such horns with varying degrees of success. Nowadays, however, with much more powerful low frequency drivers and amplifiers being available, the use of low frequency horns has become a rarity, except for sound reinforcement uses. Nonetheless, due to the resistive loading which the horn provides on the face of the driver, horn loaded bass systems are frequently liked for their characteristically fast and detailed subjective low frequency response.

Richard Small, in his 1970 paper 'Constant-Voltage Crossover Network Design[7]' highlighted this difference between horn-loaded and direct-radiating devices, and considered its implication for crossover design when used between horns and cones. He stated that whilst direct radiation diaphragm motion is largely mass-controlled, horn diaphragms are resistance controlled, and that the result is a constant phase difference of 90 degrees between the transfer characteristics of the two types of drivers. As previously stated in Section 4.1, the reactive loading (due to mass control) and the resistive loading (due to horn loading) are the mechanisms primarily responsible for the sensitivity differences – more power being radiated by the drivers which are resistively loaded, given the same electrical input.

References

1 Keele, D. B.Jnr., 'Optimum Horn Mouth Sizes', Presented at the 46th Convention of the Audio Engineering Society, Preprint No 933 (1973)

2 Keele, D. B.Jnr., 'What Is So Sacred About Exponential Horns?' Presented at the 51st Convention of the Audio Engineering Society, Preprint No 1038 (1975)

3 Newell, P., 'Studio Monitoring Design', Focal Press, Oxford, UK (1995)

4 Holland, K.R., Fahy, F. J, Newell, P. R., 'Axi-symmetric Horns for Studio Monitoring Systems', Proceedings of the Institute of Acoustics, Vol. 12, Part 8, pp 121–128, Reproduced Sound 6 Conference, Windermere, UK (1990)

5 Newell, P. R., Holland, K.R., ' Do All Mid-Range Horn Loudspeakers Have A Recognisable Characteristic Sound?' Proceedings of the Institute of Acoustics, Vol. 12, Part 8, pp 249–258, Reproduced Sound 6 Conference, Windermere, UK (1990)

6 Morfey, C. L., 'Dictionary of Acoustics', Academic Press, London, UK and San Diego, USA (2001)

7 Small, R. H., 'Constant-Voltage Crossover Network Design', Journal of the Audio Engineering Society, Vol. 18 pp 172–180 (1970)

Chapter 5

Crossovers

5.1 What is a crossover?

The term 'crossover' appears to have been originally used to describe the relationship of the filter slopes – crossover filters – as shown in Figure 5.1. In reality, and in many languages other than English, they are better described as frequency dividing networks, or words to that effect, though the name crossover has generally stuck. The fact that no loudspeaker drive unit suitable for music monitoring or serious listening can provide a flat response over the entire musical frequency range requires that the multiple drivers in a system need to be fed by signals which are only appropriate to their designed performance range. The two normal ways to apply these filtered signals are via high level, passive crossovers – where the filter components are placed between the power amplifier and the loudspeaker drive units, or low level active crossovers – where the filters are placed in the line level signal circuits, ahead of the amplifier inputs. In the latter case, each filter output feeds a separate amplifier, which is then directly connected to the corresponding drive unit(s). In some cases, mixtures of the two concepts are applied to one system, such as an active crossover between the bass and mid drivers, and a high level passive crossover between the mid and high frequency drivers, as shown in Figure 5.2.

Other forms of crossover also exist, such as simple, low-level passive crossovers, though they are rarely used because the filters can be more precisely tailored when the components are part of the feedback path in an electronic circuit. Mechanical crossovers are another type of filter. These can take the form of aluminium domes in the centre of the cones, which decouple from the main cone at higher frequencies and radiate separately, extending the frequency response above that which could be achieved by the main cone, alone. However, the response tends to be somewhat irregular, but this type of high frequency extension can find use in loudspeakers for music *production* – as opposed to reproduction – and the technique is extensively used in loudspeakers for musical instrument amplification, such as guitar amplifiers. An example is shown in Figure 5.3. Figure 5.4 shows a 'parasitic cone' or 'whizzer cone'. The concept is generally the same in principle as that of the metal dome – the small cone decouples from the main cone at high frequencies – although the response of the parasitic cone tends to be more controlled, and a flatter frequency response can normally be achieved. Another concept, although not widely used, is inductive coupling, where the high frequency cone is not electrically connected to the amplifier. In fact, the coil can be the single, shorted turn

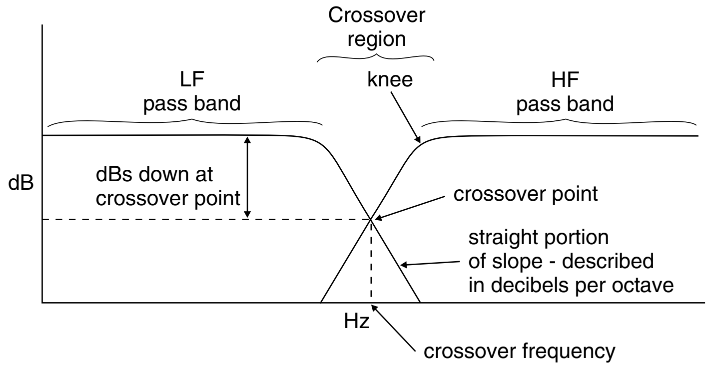


Figure 5.1 The basic concept of a pair of crossover filters

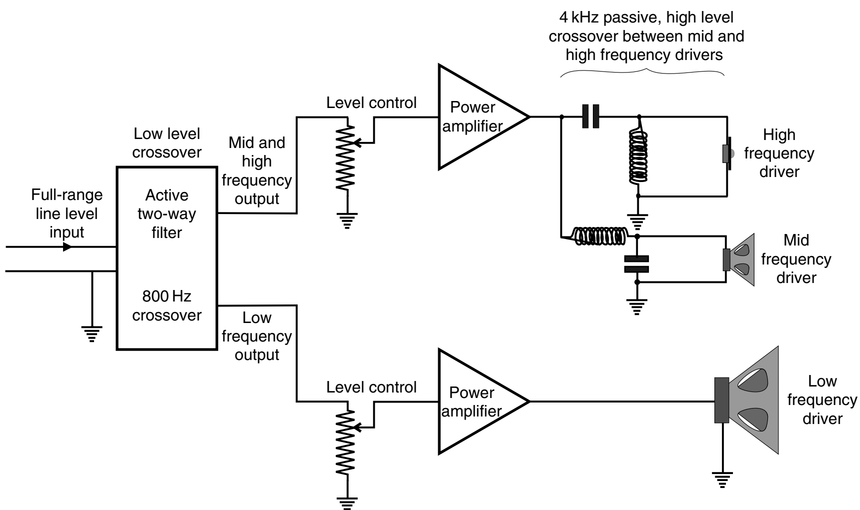


Figure 5.2 Example of a crossover system employing both active, low-level and passive, high-level filters

formed by the metal dome itself. The dome and a former are simply placed over the centre pole of the magnet assembly, sharing the same gap as the LF/MF cone assembly. Such inductively coupled transducers, or ICTs, are operated by the modulated magnetic coupling between the 'coil' and the magnetic circuit. This type of 'crossover' is neither electrical, electronic nor mechanical, but is simply a magnetic-inductive effect.

5.2 Reconstruction problems

Unfortunately, the division of the frequencies is not all that a crossover must achieve. They must divide the frequencies in a way that the individual drive units can re-construct in the acoustic far-field of the loudspeaker

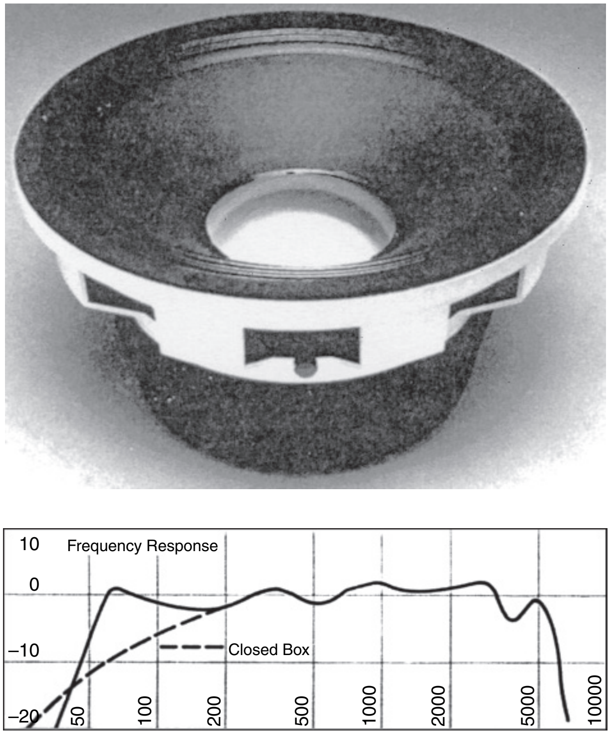


Figure 5.3 The Gauss model 4281, 12 inch (300 mm) drive unit for musical instrument use, which used an aluminium dome to extend the high-frequency response to over 5 kHz

a representation of the waveform which was electrically applied to the electronic amplification system, and it is not an easy task to do so. Figure 5.5 shows a representation of a typical, two-way loudspeaker system. Note how, due to the physical requirement of the radiation of the different frequency bands, the sizes of the drive units give rise to a displacement of the voice coils if the front faces of the drives share a plane, common baffle. If the displacement were to be 10 cm, then a frequency with a wavelength of 20 cm would be received on the axis between the two drivers with its polarity reversed from either driver with respect to the other. The wavelength (λ) can be calculated by dividing the speed of sound (c), in metres per second, by the frequency (f) in Hz, so we arrive at the formula:

$$\lambda = \frac{c}{f} \tag{5.1}$$

To find the frequency with a wavelength of 20 cm, we can re-arrange the formula as:

$$f = \frac{c}{\lambda}$$

Therefore: $f = \dfrac{340}{0.2} = 1700\,\text{Hz}$

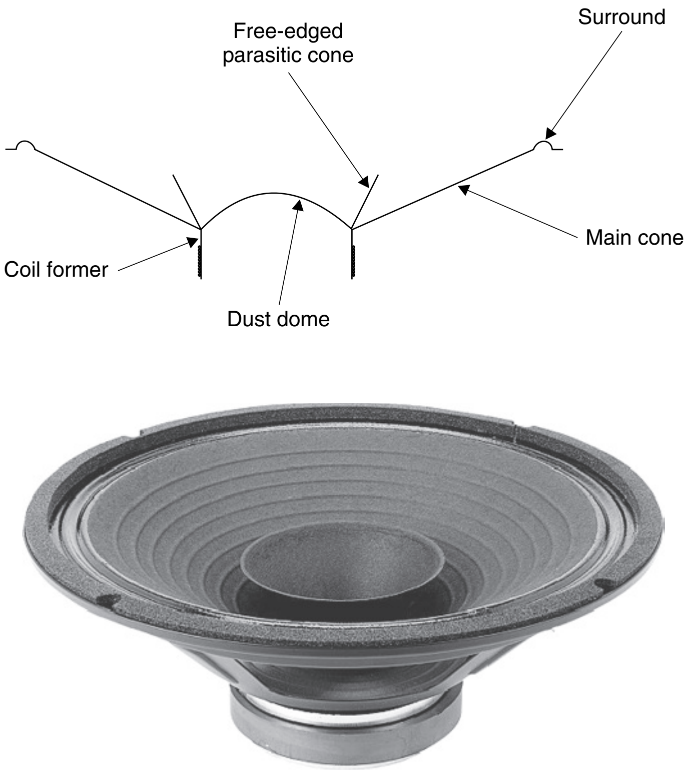


Figure 5.4 Parasitic, free-edged cone for high-frequency extension

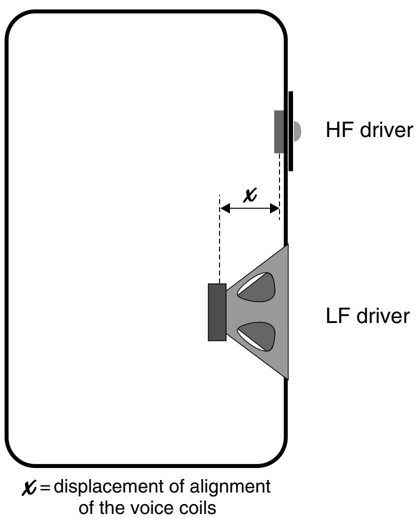


Figure 5.5 Elevation of a two-way loudspeaker system showing the typical offset of the voice coils in the axial plane

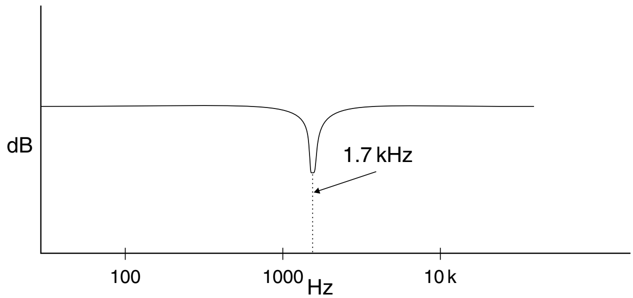


Figure 5.6 Composite response of the loudspeaker shown in Figure 5.5 on its central axis if the distance *x* were to be set at 10 cm

At 1700 Hz we would have cancellation on axis, producing a response as shown in Figure 5.6

However, this is not the only complication which arises in the reconstruction. All conventional filters exhibit the property of 'group delay'. There is a finite time necessary for the information in a signal waveform to pass through a filter, which is a function of the slope of the filter and its cut-off frequency. As the frequency drops, the delay increases. As the steepness of the filter slope increases, so does the group delay. A filter of 24 dB/octave at 300 Hz would exhibit a group delay of around one millisecond. With a speed of sound of 340 m/s, one millisecond would represent 340/1000 m/s, or a 34 cm equivalent physical displacement. If the loudspeaker represented in Figure 5.5 were to be fed via such a crossover, the *real* radiation from the low frequency driver would be delayed by the equivalent of being mounted 44 cm behind the HF driver, (34 cm due to the filter and 10 cm due to the physical misalignment). Figure 5.7 shows the step-function responses of four loudspeaker systems with different degrees of arrival synchronisation from the drive units. Figure 5.7(d) shows the step response of a commercial 3-way loudspeaker which clearly exhibits delays between the arrival times of the individual drive units.

In effect, if not compensated for, the flat-response-axis of the loudspeaker shown in Figure 5.5 would be tilted, as shown in Figure 5.8, for moderate low frequency signal delays. Conversely, though, if we engineer a flat response on axis by compensating for the delays, it follows that the frequency responses *off*-axis must be incorrect, as shown in Figure 5.9. So, it can be appreciated that once the drive units have been physically separated, the problem of reconstructing the waveform of the original signal can become very difficult indeed. The problem can be partially solved by the use of concentric loudspeakers, but these concepts bring their own problems with them. For example, with the Tannoy, Dual Concentric approach, or the KEF Uni Q for that matter, the low frequency cone serves as a horn/waveguide for the high frequency driver. Modulating the LF cone with high levels of low frequencies can hardly be expected *not* to affect the high frequencies. In the case of the Tannoys, the LF coil gap also serves as an unwanted discontinuity in the HF horn. The general concepts of these drivers are shown in Figure 5.10, along with the Altec/UREI approach. In

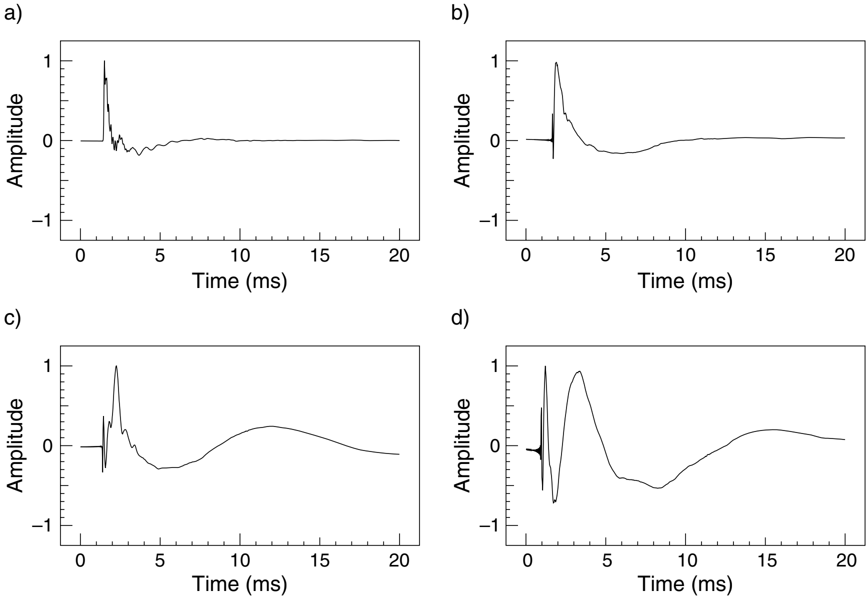


Figure 5.7 Step function responses with one, two and three drivers. a) Integrated attack of a relatively wideband, single driver. b) Integrated attack of two-way system with excellent time alignment. c) Separate rise times visible from the two drivers of a system with slightly delayed low frequency driver response. d) 3-way system with a clearly delayed response from the bass driver

In the latter case, the delay was a consequence of designing the system so that the main lobe of irregular response was forced upwards into a generally inoffensive direction. The trade-off was considered to be beneficial to the overall performance in typical situations in which is was expected to be used

the latter case, the separate, concentrically mounted horn is left hanging in free air, but this method of mounting is really, too abrupt for proper mouth termination at the 1 kHz crossover frequency. The termination problem was discussed in detail in the previous chapter. Therefore, as so very often is the case with loudspeaker design, the tendency is to be trading one problem for another, rather than solving them – finding the best compromise for each situation – but that is often the reality of loudspeakers, which is something that will form the basis of discussion in Chapter 8.

5.3 Orders, slopes and shapes

Despite the different solutions on offer, electrical filters are overwhelmingly the most common manner of dividing the frequency bands. Whether this is done at high level, low level, actively or passively, the same basic filter concepts apply. Figure 5.11 shows a simple high pass filter, (a) to (d) showing first, second, third and fourth order roll-offs, respectively. Each inductor or capacitor adds 6 dB per octave of roll-off, and each 6 dB is known as an

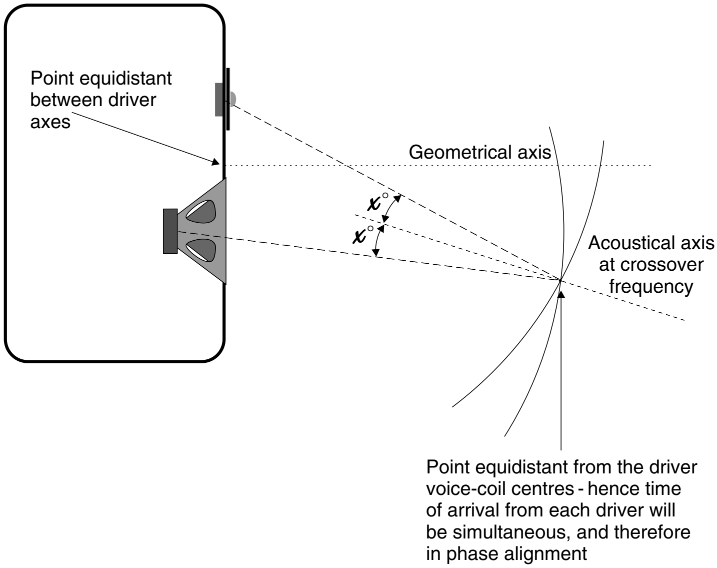


Figure 5.8 Tilting of the acoustic axis at the crossover frequency due to voice-coil physical displacement. On this axis, the response dip shown in Figure 5.6 would not be observed – the response would be flat

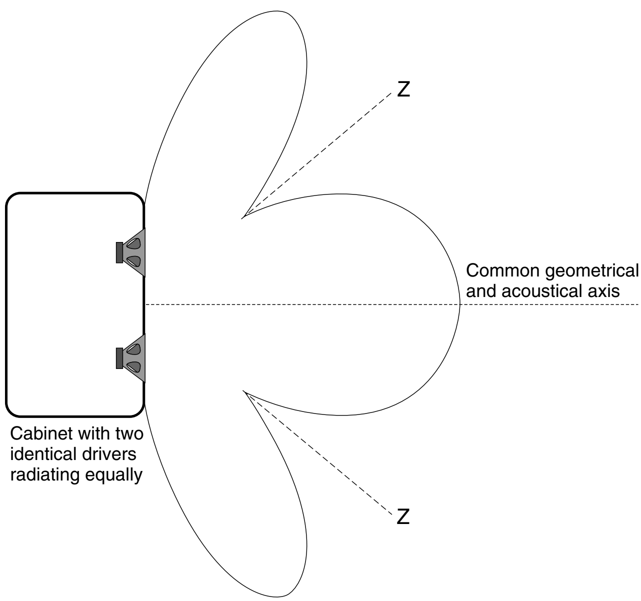


Figure 5.9 Lobing of the response when physically displaced drivers radiate a common frequency whose wavelength is close to, or smaller than, the distance between the drivers Z = axis of phase cancellation – varies with frequency – cancellation occurs whenever the distance to the two drive units varies by half a wavelength. The pattern shown therefore represents the situation at one frequency, only

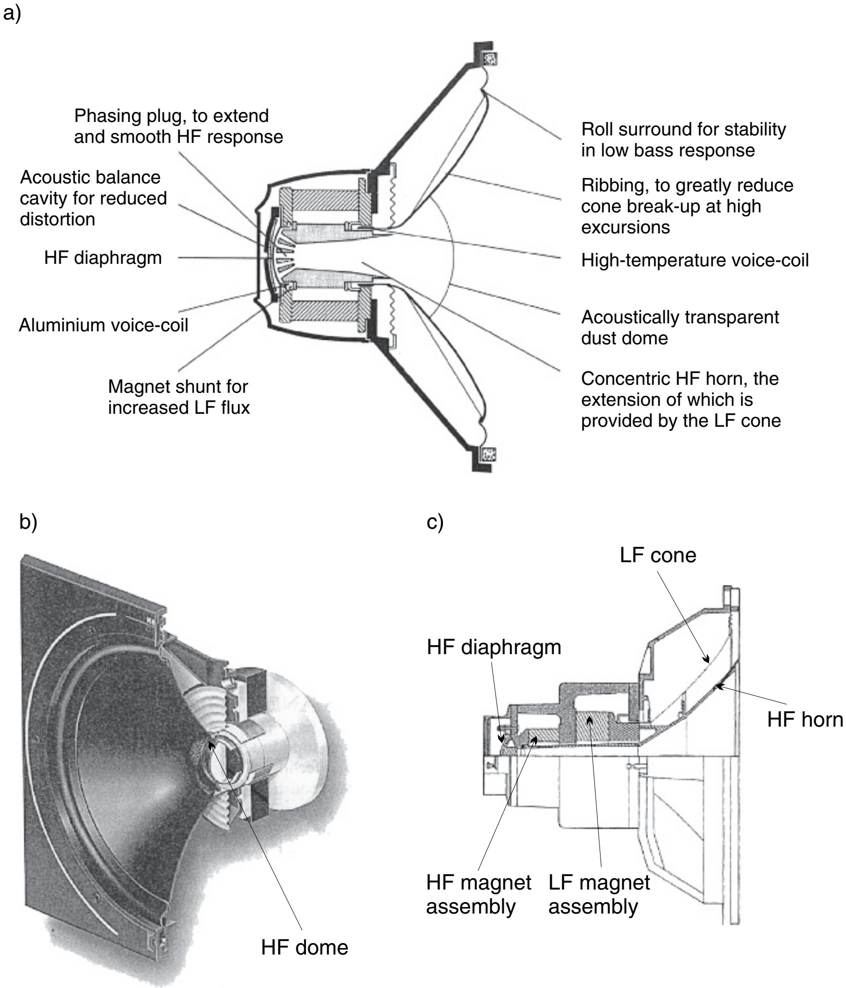


Figure 5.10 Some concentric drive units. a) Tannoy Dual Concentric – with Alnico magnet structure (courtesy of Tannoy Ltd) b) The KEF Uni-Q. c) The Altec 604

order of roll-off, a term which comes from the mathematical application of filter theory. An alternative approach to the inductor/capacitor (LC) design is a resistor/capacitor (RC) method shown in Figure 5.12. The LC approach is preferred for high level crossover in the loudspeaker/amplifier interface because the power losses are much less, but the RC approach is preferred in low level circuitry because of its simplicity (perfect inductors are not easy to make) and its relative insensitivity to drift and interference pick-up. In the active circuits, where gain is plentifully available, the higher losses of the RC circuits are of little consequence.

First order crossovers are rarely used, because the low rate of roll-off requires the individual drivers to have respectably flat responses for at least

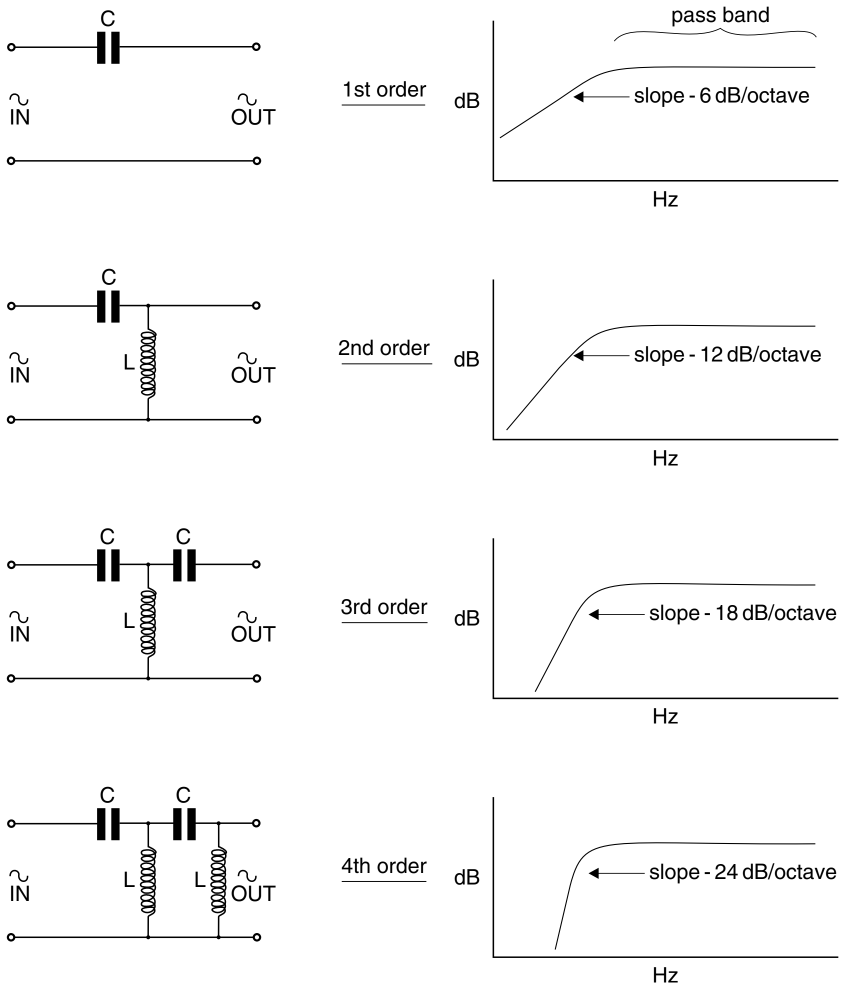


Figure 5.11 Capacitor/inductor filters – circuits and slopes. Values C and L depend upon the turnover frequency and the load impedance

two, if not three, octaves each side of the crossover point, which is usually not practicable. Nonetheless, when they *are* able to be used, they have the advantage that they are the only conventional crossovers whose combined outputs reconstruct the input waveform. This is shown in Figure 5.13, and results from the fact that although each side of the filter is only 3 dB down at the crossover frequency (*voltage* summing would normally require that they should be 6 dB down to sum back to a flat response) the +45 degrees phase shift through one half of the crossover and the −45 degrees phase shift through the other half give rise to a 90 degrees combined shift which leads to another 3 dB of attenuation. The combination of amplitude roll-off and phase shifts leads to the perfect reconstruction shown in Figure 5.13.

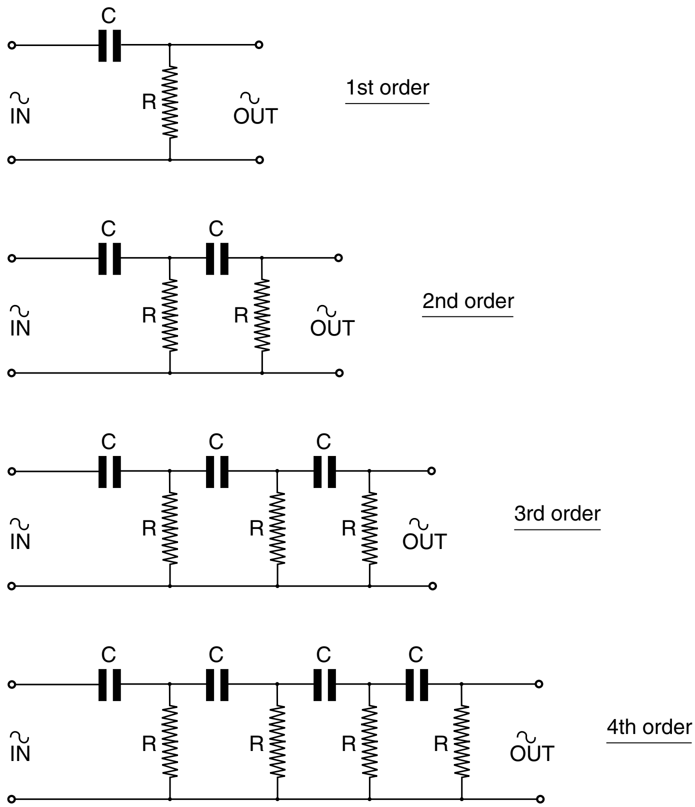


Figure 5.12 Resistor/capacitor equivalents of the filters shown in Figure 5.11
Except for the first-order filter, and where the R is supplied by the loudspeaker load impedance, these filters are not suitable for use as high-level filters because of the excessive power dissipated in the resistors. Values of C and R depend on the turnover frequency and the load impedance

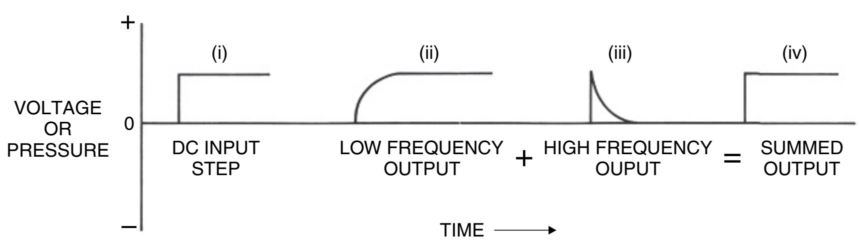


Figure 5.13 6 dB/octave crossover waveform reconstruction

Second order crossovers, with their 12 dB/octave roll-offs, are very popular with the manufacturers of small, two-way cabinet loudspeakers. They are relatively cheap to construct and the power losses through the filters are quite small, but, as can be seen from Figure 5.14, they will *not* reconstruct

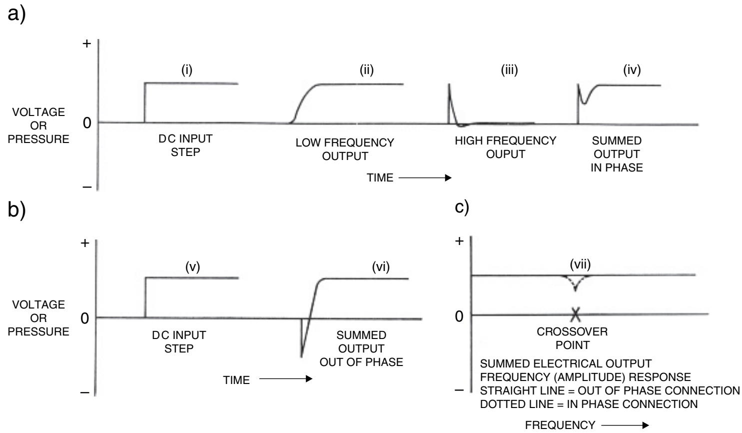


Figure 5.14 12 dB/octave crossover waveform reconstruction. a) In-phase. b) Reversed polarity. c) Amplitude responses

Note that whichever polarity is applied, the output waveform is not a true representation of the input waveform. (The waveform in Figure 5.13 reconstructed perfectly)

the original waveform. The group delays associated with the phase shifts cause a temporal offset between the two halves, so the reconstruction is not summing the two outputs at the same instant. This 'latency' in one half of the filter, relative to the other, creates a phase shift at the crossover frequency which does *not* compensate for the amplitude summation; hence the time and amplitude summations shown in Figure 5.14(a). Despite the fact that the reversing of the polarity of one of the outputs yields a flat amplitude response, the time response (waveform) becomes even more distorted. This is shown in Figure 5.14(b). There are ways of 'juggling' this arrangement by offsetting the 3 dB down points so that the filter sections overlap, but this can introduce other irregularities unless it is very carefully implemented. However, in general, a standard second order crossover yields *either* a flat frequency response *or* a synchronous time response, but cannot exhibit both properties at the same time.

Third order crossovers, with slopes of 18 dB/octave are popular in more expensive passive loudspeaker systems, and are not infrequently used in active crossovers. They are more expensive in the passive form than the second order types, not only because they use more components, but also because the components may need to be less lossy. Otherwise, with so many components between the amplifier and the drive units, they would begin to sap considerable power from the signal. Typical circuit diagrams are shown in Figure 5.15, and a response summation is shown in Figure 5.16. The 18 dB/octave roll-off is useful in reducing the disturbance from out of band irregularity of the driver responses, because the driver responses can be almost 20 dB down an octave beyond the crossover frequency. This allows drivers to be used over almost all of their flat response region, but with third-order crossovers there is no 'correct' polarity relationship between the outputs. The phase responses for normal and reversed polarity are shown in Figure 5.17, and the reverse polarity connection actually

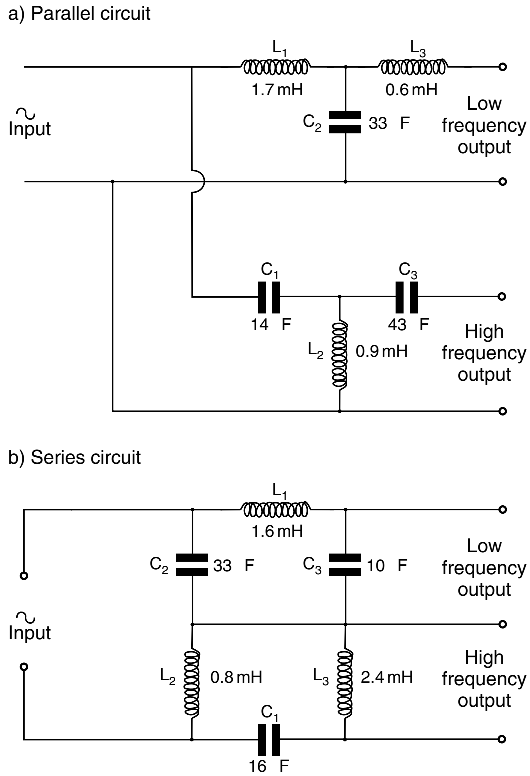


Figure 5.15 Typical, 2-way, 18 dB/octave, 3rd order passive crossover circuits
Approximate values shown for 1 kHz crossover frequency and a uniform loading of 8 ohms

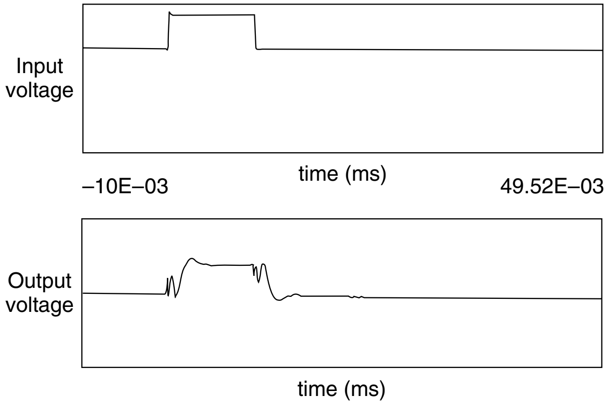


Figure 5.16 Summed output of a 3-way, 18 dB/octave crossover
The signal path delays through the individual filter sections are clearly evident

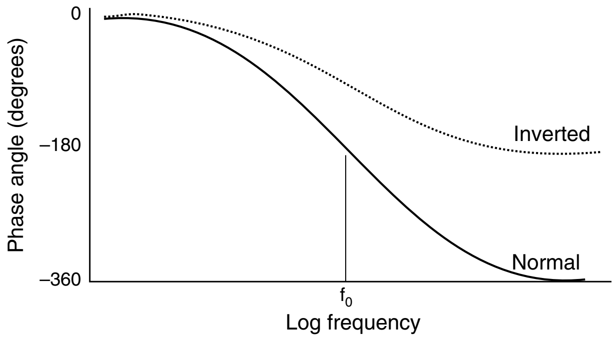


Figure 5.17 Phase relationships through the crossover region for the normal and inverted polarity connection of a third-order Butterworth crossover – f_0 is the nominal crossover frequency

exhibits less phase shift through the crossover region after the summation of the outputs. However, this is the effect of simply summing the electrical outputs. Once those outputs are connected to physically displaced loudspeakers the story can be rather different.

One of the most popular types of crossover filter for use in active designs is the fourth order Linkwitz-Riley. The 24 dB/octave slope is achieved by cascading a pair of Butterworth (see next section) 12 dB/octave crossovers. The 24 dB/octave slopes are beneficial in high-power loudspeaker systems, where out-of-band energy is rapidly cut off, and for use with drivers whose out-of-band response is also irregular. The *power* response exhibits a dip at the crossover frequency, but the width of the dip is so narrow as to be, in many cases, almost inconsequential. The on-axis amplitude response is flat, because each section is normally designed to be 6 dB down at the crossover frequency. The two outputs sum to unity (−6 dB = half voltage [or half pressure], therefore ½ + ½ = 1) because each section of the crossover is rotated 180 degrees, and the overall response therefore remains in phase. However, due to the shape of the polar response, there are dips *off axis* at the crossover frequency, hence the dip in the total power response. The degree of audibility of this effect depends on how much reflected energy is returned to the listening position. In typical highly damped control rooms, the effect is usually imperceptible close to the axis and around the typical principal listening positions. With these crossovers, the narrowness of the band of frequencies over which the drivers on either side of the crossover frequency are simultaneously radiating ensures that the interference effects are kept small.

Filter orders higher than fourth are not normally used in crossover pairs, but they can sometimes be found in asymmetrical designs, such as a sixth order with a second order, to help to compensate for group delays or physical alignment delays. These individual filters sometimes also have their 3 dB or 6 dB- down points offset in frequency, in order to achieve a flat on-axis response or flat power response, depending upon the circumstances. In practice, filters above sixth order are rarely used because they tend to serve no useful purpose, and due to their tendency to introduce time response anomalies may actually create more problems than they can solve.

5.4 Filter shapes

Until now we have been looking at standard filters, where, for generating higher orders, the filters are simply repeating the characteristics of the first order sections. However, different entry slopes and different response shapes can be contoured by adjusting the components which are used to derive the higher orders. Therefore, by adjusting the Q of the filter, the 'knee' of the curve – between the flat response and the uniform slope – can be adjusted in shape to produce either a more abrupt or more gradual entry to the final roll-off. The modified shapes can be used to help to achieve the desired, overall electro-acoustic response when the individual responses of the drivers are taken into account, or when system summation is affected by physical displacement problems. Many different Q factors – or quality factors – are in use. Butterworth filters are 'maximally flat' until the roll-off begins. Bessel filters are more abrupt in their transition from flat to sloping response, but exhibit alternating up and down responses before the main roll-off begins. Figure 5.18(a) shows a typical 6 dB/octave roll-off produced by a simple capacitor-resistor filter, but Figure 5.18(b)

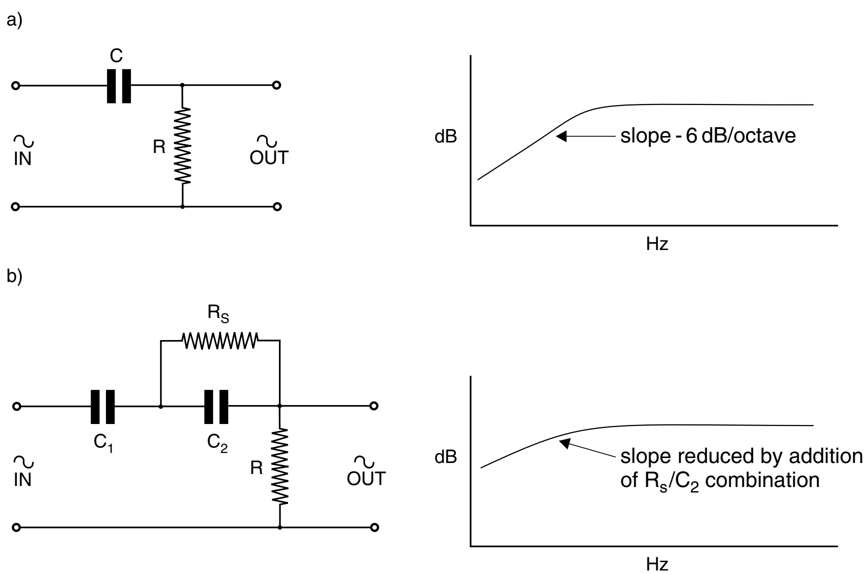


Figure 5.18 Response contouring. a) A simple high-pass filter of first order characteristic roll-off. The response will be 3 dB down at the frequency where the reactance of the crossover equals the resistance (R). Note, a resistor having the same value as R, substituted for C, would give rise to a 6 dB reduction of level at the output. The 90 degree phase shift through the capacitor is the reason for the 3 dB difference. b) Added components for reducing the slope

shows how the addition of an extra resistor can limit the rate of roll-off. There are therefore means at our disposal to modify the slopes of the curves of the filters, which is extremely useful when we have to tailor curves to mirror the responses of real drivers, which never behave like pure resistors, and so can never ideally terminate the standard filters which were discussed in Section 5.3.

Figure 5.19 shows an example of a conjugate network. These are used to flatten the impedance curves of drive units by compensating for the reactive properties of their electro-mechanico-acoustic characteristics. However, the resistors in these networks do dissipate power, so the overall efficiency of the system may reduce when conjugate networks are applied. Sonically, they can be questionable, and some complicated passive crossovers applying such technology for electrical and response flattening purposes have been considered to sound worse than simpler networks, but the overall effect may also depend on whether the power amplifiers with which they are used are capable of driving complex loads, or not. If they are not, then the conjugate networks may be helpful, but in professional situations, the choice of a more load-tolerant amplifier is often the preferred solution. Complex passive crossovers are not entirely distortion free, because the inductors, in particular, can produce non-linear distortion. Cases tend to need to be judged individually, partly because, as will be discussed in the next chapter, amplifier outputs stages and power supplies vary so widely in design concept.

Clearly, where complex circuitry is involved, it is essential to use components whose values remain stable if the performance of the overall system is not to change as time passes. Unfortunately, the large electrolytic capacitors which tend to be required by high-power, low frequency passive crossovers are notoriously prone to changing their capacitance as they age. Likewise, changes in mechanical compliances of the drive units with

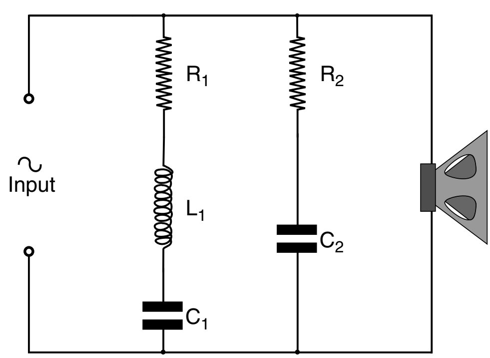


Figure 5.19 A conjugate network

Figure 1.5 shows the impedance curve of a typical low frequency driver. In the circuit above, R_1, C_1, and L_1, provide compensation for the impedance rise at the resonant frequency of the driver, whilst R_2 and C_2 compensate for the rising impedance at higher frequencies due to the voice-coil inductance. Opinions vary about the sonic effects of this type of impedance equalisation, but such circuits are undoubtedly useful to flatten the response of systems using hi-order, passive crossovers age can also cause a drift in the circuit parameters. These factors tend to put high-level, passive crossovers at a disadvantage when compared to low level active circuitry, especially when precise tailoring of the response is required. The higher impedance, active circuitry of the latter case can avoid the use of large value capacitors, and can therefore employ smaller value components of much greater long-term stability. This is an important concern when we need to produce very precise response curves. Active circuitry can also eliminate the need for lossy, and potentially non-linear inductors.

5.5 Target functions

Until as late as the 1970s it was commonplace to treat a moving coil loudspeaker as if it were a resistor when designing crossover circuitry, but, as we saw in Chapter 1, this is usually a long way from reality. In fairness, before the 1970s, and the seminal work of Thiele and Small[1,2,3,4,5,6], it was not always very easy to find the necessary electro-mechanical information about drive unit characteristics, so design and development were often two separate processes – design it ideally, and then select components during tests, to modify the response. Since the 1980s, with much more data available, it has been customary to look at a drive unit as a complex impedance, and to view its response for what it actually is, and not as an ideal response. There has subsequently emerged the concept of the 'target function'.

Whereas in Figure 5.1 the plots represent the responses of the filter circuits which presume that the driver is of constant impedance and has a flat acoustic response, we can also view the same plots as the *target* response for our combined electro-mechanico-acoustic system. In order to achieve this, the desired *filter* response becomes whatever is required to combine with the actual *driver* response to realise the *target* response. Figure 5.20(a) shows a target response; (b) shows the response of an actual driver, and (c) shows the response of a filter which, when used with (b) will yield the response shown in (a). This 'new' approach obviously calls into question the value of purchasing ready-made crossovers as stand-alone devices, be they active or passive. They may be of value for sound reinforcement systems, where each loudspeaker box has been engineered to be more or less flat, and multi-band system equalisation is de rigueur, but their use may be over-simplistic when applied to many other loudspeaker systems. Also, many modern drive units are not necessarily *engineered* for flat responses if other benefits can be gained by sacrificing the flatness. Computer-aided filter design can then be applied in order to design a crossover which will achieve the required target function from the complete system. In fact, in the current cost-conscious world, it is often less expensive to use electronic means to flatten a response rather than electro-mechanical means. Nonetheless, despite having said that, there is still in the experience of many listeners a certain *je ne sais quoi* about the purity of sound of an inherently flat driver. It should also be acknowledged that the more complex filter shapes such as that shown in Figure 5.20(c) are much more easily implemented with active crossover designs.

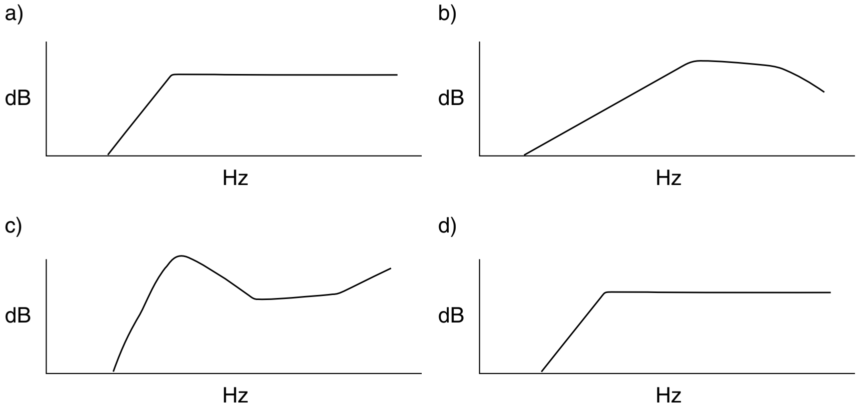


Figure 5.20 Target functions –practical realisation. a) Desired target function. b) Measured driver response. c) Electrical filter response – allows for a non-flat driver response. d) System response as measured – equal to a), i.e. it achieves its target

5.5.1 Minimum and non-minimum phase effects

The selection of the target function is somewhat more complicated than it may initially appear, because the phase shifts and group delays associated with the filters, *and* the physically induced delays due to the different drive units occupying different points in space, lead to non-minimum-phase responses. A minimum-phase response is one where the correction of the amplitude towards a flat response also leads to a corresponding flattening of the phase response, or vice versa. In the case of a non-minimum phase response, the correction of either the amplitude or the phase response does not automatically correct the other. Non-minimum-phase responses give rise to situations where a flat amplitude response cannot be accompanied by an accurate transient (time) response. The amplitude and phase responses are defined by the time response, and vice versa, which is why the Fourier Transform and Inverse Fourier Transform can be used to derive the frequency response from the impulse response or the impulse response from the frequency response. [For more on this subject, see Chapter 9.] The frequency response, in this case, is referring to the *complete* frequency response, i.e. the amplitude *and* the phase. Non-minimum-phase effects are typically associated with the recombination of non time-synchronous signals, such as a recombination of a reflexion with a direct signal, or the summation of signals where different group delays or digital latency have been incurred.

Unfortunately, for loudspeaker designers and users, the non-minimum-phase effects often prevent the perfect reconstruction of a waveform from a multi-drive unit loudspeaker system. When this fact is coupled with the fact that no *single* drive unit can cover the whole frequency range in a manner which is flat in frequency response and adequate in its directivity pattern (at least not at useful sound pressure levels) we are condemned to compromise. The time-shifted sections of the overall response cannot be equalised flat in a minimum-phase manner, and so the transient responses and directivity patterns will be altered. Even when using concentric drivers, where the vertical and lateral displacements can be avoided, the problem of the crossover group delays still dogs the design process, although, if crossover frequencies are carefully matched to distances on the front/back plane, they can sometimes almost be eliminated. The effects of trying to equalise non-minimum phase responses are clearly shown in Figures 11.17 and 11.18, in which the time response aberrations are very evident.

5.5.2 Corrective measures and side-effects

In some cases, the electrical crossover filters can be overlapped to some degree, in order to take into account certain physical aspects of the relative driver positions and in order to vary the polar pattern of the complete system. Lobes in certain directions may or may not be problematical dependent upon the intended position in which the loudspeaker is expected to be used, for example, see Figure 5.21. If a loudspeaker system does exhibit a lobed radiation pattern with an aberration in the frequency response, then that lobe is best directed towards places where there is least likelihood of returning a reflexion to the listening position which would be detrimental

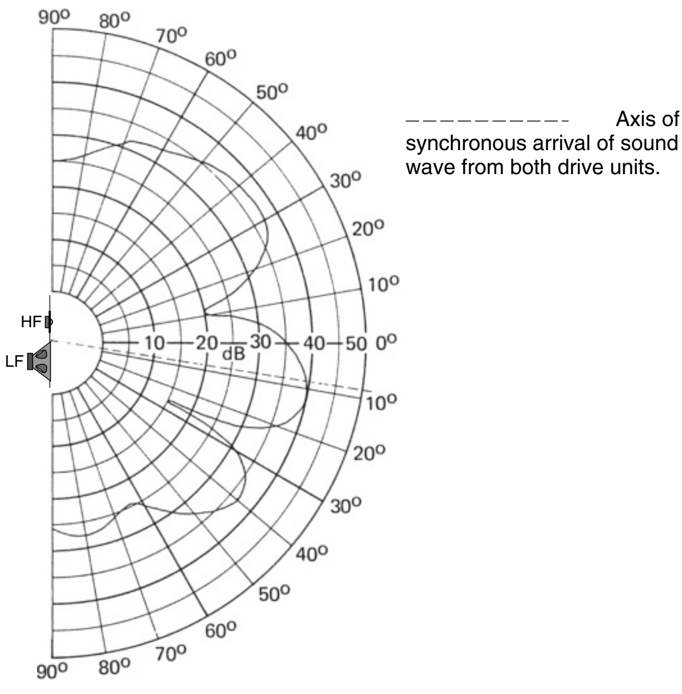


Figure 5.21 This type of lobing could be problematical if an irregular frequency response returns towards the listener from a reflective floor. Some designers choose to invert the cabinet in order to direct the irregular response above the listeners' heads, especially if the ceiling is high, or more absorbent than the floor

to the overall perception of the music. Decisions must also be made about whether the most likely movements of the listener to the loudspeaker will be in the horizontal or vertical direction when the loudspeakers are in critical use. For example, either sitting in a chair whilst moving left or right along a mixing console, or working predominantly in one place but at various times either sitting down or standing up.

In Section 5.3 it was described how a first-order crossover could reconstruct a perfect waveform. However, if the crossover is used with physically non-coincident drivers, then even a first order crossover will suffer from the above problems except on its acoustic axis, which may not always correspond with its physical axis. This means that off-axis reflexions would *not* have time-coincident origins, and there would be a lobe which was tilted, either upwards or downwards. One reason for the great popularity of the fourth order Linkwitz-Riley filters for high quality monitoring is that the in-phase relationship between the two drivers on either side of the crossover gives rise to the main lobe being symmetrical about the central axis between the drivers.

A very important point to emphasise here is that the design of crossover filters is not by any means the easy task which many people believe it to be. Many things must be taken into consideration before the filter functions can be decided upon, and the design of suitable filters can be work of a very specialised nature. Computer aided design of filters has been a great boon to loudspeaker engineers.

5.6 Active versus passive crossovers

For high quality loudspeaker applications, the consensus is almost universally in favour of active crossovers. By virtue of their feedback loops they can remain remarkably stable over very many years, and complex filter shapes can be devised without any loss of power efficiency. Conversely, passive crossovers rely entirely on the long term stability of each component part for their overall stability, which is not easy to achieve when the low impedances of loudspeaker circuits call for high value capacitors – both in terms of capacitance *and* working voltage – which in turn call for capacitors of types which may not be able to provide good, long-term stability. Complex filter shapes may need to use many components, which when placed in the power circuitry will probably lead to lower system efficiency, wasting many watts of amplifier output. If large electrolytic capacitors are needed, their stability can be questionable, but, if the much larger, solid dielectric capacitors are used, their physical construction and large size can lead to them having considerable unwanted inductance, which can upset the crossover operation.

Active filters are free from these problems, and in fact require no inductors at all. What is more, if state-variable filters are used, any drift which does occur can reflect equally in both halves of the crossover response, thus only slightly varying the crossover frequency as opposed to opening a gap or causing an overlap. The overall frequency response of the system is therefore unlikely to be affected. There can be no equivalent to this type of stability or self-correction with passive crossovers, and neither can

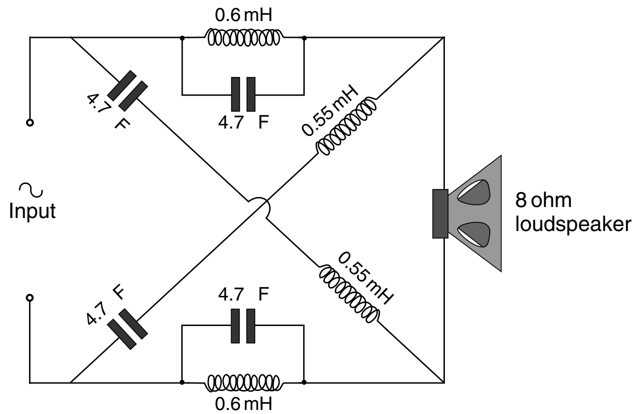


Figure 5.22 An all-pass delay circuit

Tannoy developed a delay circuit similar to the one shown, but unwanted phase-shifts, response ripples and mistermination problems tend to prevent the theoretical benefits of its approximately 150 microsecond delay from being fully realised. One hundred and fifty microseconds of delay would ideally compensate for the distance of about 5 cm between the voice coil planes

passive crossovers easily compensate for group delays. Figure 5.22 shows a passive all-pass circuit (an analogue delay circuit) of a type which has been applied commercially to loudspeaker systems, but there are people who feel that this type of circuitry between an amplifier and a loudspeaker can again cause as many problems as it solves; if not more!

Conversely, active filters can *easily* incorporate delay compensation for driver mounting offsets, such as when a horn driver is set behind the woofers. Response tailoring is independent of the loudspeaker impedance complexities. The list of advantages in favour of active crossovers and multi-amplification is impressive:

1) Loudspeaker drive units of different sensitivities may be used in one system without the need for lossy resistive networks or transformers. This can be advantageous because drive units of *sonic* compatibility may be electronically incompatible in passive systems.
2) Distortions due to overload in any one band are captive within that band, and cannot affect any of the other drivers.
3) Occasional low frequency overloads do not pass distortion products into the high-frequency drivers, and instead of being objectionable may, if slight, be inaudible.
4) Amplifier power and distortion characteristics can be optimally matched to the drive unit sensitivities and frequency ranges.
5) Driver protection, if required, can be precisely tailored to the needs of each driver.
6) Complex frequency response curves can easily be realised in the electronics to deliver flat (or as required) acoustic responses in front of the loudspeakers. Driver irregularities can, except if too sharp, be easily regularised.

7) There are no complex load impedances as found in passive crossovers, making amplifier performance (and the whole system performance) more dynamically predictable.
8) System intermodulation distortion can be significantly reduced.
9) Cable problems can be dramatically reduced.
10) If *mild* low frequency clipping or limiting can be tolerated, much higher SPLs can be generated from the same drive units (vis-à-vis their use in passive systems) without subjective quality impairment. (See 2) and 3) above.)
11) Modelling of thermal time constants can be incorporated into the drive amplifiers, helping to compensate for thermal compression in the drive units, although they cannot totally eliminate its effects.
12) Low source impedances at the amplifier outputs can damp out-of-band resonances in drive units, which otherwise may be uncontrolled due to the passive crossover effectively buffering them away from the amplifier.
13) Drive units are essentially voltage-controlled, which means that when coupled directly to a power amplifier, (most of which act like voltage sources) they can be more optimally driven than when impedances are placed between the source and load, such as by passive crossover components. When 'seen' from the point of view of a voice coil, the crossover components represent an irregularity in the amplifier output impedance.
14) Direct connection of the amplifier and loudspeaker is a useful distortion reducing system. It can eliminate the strange currents which can often flow in complex passive crossovers.
15) Higher order filter slopes can easily be achieved without loss of system efficiency.
16) Low frequency cabinet/driver alignments can be made possible which, by passive means, would be more or less out of the question.
17) Drive unit production tolerances can easily be trimmed out.
18) Driver ageing drift can easily be trimmed out.
19) Subjectively, clarity and dynamic range are generally considered to be better on an active system compared to the passive equivalent (i.e. same box, same drive units). [See also Figure 5.23.]
20) Out of band filters can easily be accommodated, if required.
21) Amplifier design may be able to be simplified, sometimes to sonic benefit.
22) In passive loudspeakers used at high levels, voice-coil heating will change the impedance of the drive units, which in turn will affect the crossover termination. Crossover frequencies, as well as levels, may dynamically shift. Actively crossed-over loudspeakers are immune to such crossover frequency changes.
23) Problems of inductor siting (to minimise interaction with drive unit voice coils at high current levels) do not occur.
24) Active systems have the potential for the relatively simple application of motional feedback, which may come more into vogue as time passes.

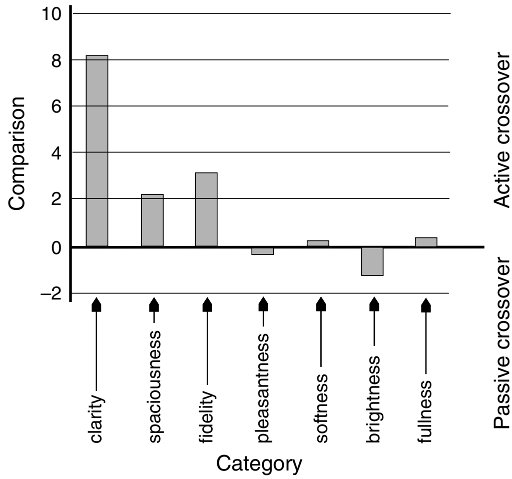


Figure 5.23 Active versus passive crossovers – subjective data (after Campbell)[7]
The results show a clear overall favour for the active crossover. The assessment of the extra brightness from the passive crossover could well be due to greater levels of non-linear distortion, which may in fact *not* be a result in its favour

Conversely, the list of benefits for the use of passive, high level crossovers for studio monitors would typically consist of:

1) Reduced cost? Not necessarily, because several limited bandwidth amplifiers may be cheaper to produce than one large amplifier capable of driving complex loads. What is more, the passive crossovers for the 1000 watt Kinoshita studio monitors shown in Figure 8.2(a) cost over 3000 euros each.
2) Passive crossovers are less prone to being misadjusted by misinformed users, who think that crossovers are some sort of 'adjust to taste' tone controls. On the other hand, passive systems have a tendency to mis-adjust themselves with age.
3) Simplicity? Not really, because very high quality, passive, high level crossovers can be hellishly complicated to implement, not to mention the amplifiers which are needed to drive them.
4) Ruggedness? No, because the electrolytic capacitors (necessary for the large values) are notorious for ageing, and gradually changing their values.

However, it must be stated that in less demanding circumstances than studio monitor loudspeakers passive crossovers obviously have their appropriate applications, but the above lists highlight the benefits of active designs where the *highest* system performance levels are required.

Clearly, the advantages of active, low level crossovers completely eclipse those of passive, high level crossovers, yet it was only around the late 1980s that dedicated active crossovers began to be seriously used on large scale monitor systems. Prior art used stock electronic crossovers, and perhaps these caused some delay in the acceptance of totally active designs because they were usually made with fixed slopes on all the filter bands. This tended to necessitate the use of multi-band equalisers, many of which were of dubious sonic quality. It took some time before people generally began to accept the need to buy a specific crossover with a monitor system, which was relatively useless in any other application. There was still a mix-and-match mentality towards components parts, each of which was expected to function as a 'stand alone' device. Once attitudes like this become established it can be very difficult to introduce new concepts. In fact, it took a long time before self-powered, actively crossed-over *small* monitors could establish their place in studio use, but domestic resistance to their acceptance has been even more pronounced. Established practices die hard, and they can be remarkably difficult to change, even in the face of clearly superior technology.

In 2003 and 2004, Alex Campbell worked on a performance comparison between active and passively crossed-over *domestic* loudspeaker systems at the Institute of Sound and Vibration Research, in the UK. He had an identical amount of money to spend on each design. His findings were presented to an international conference of the UK's Institute of Acoustics.[7] Even at this modest level of engineering, aimed at the *retail* price range of £400 – £500 (€600 – €750) per pair, the subjective assessment, made under ISVR control and using a panel of 30 subjects, came out heavily in favour of the active design. The results are reproduced in Figure 5.23, with the 'clarity' and 'fidelity' ratings being strongly in favour of the active designs, [probably also as a result of greatly reduced intermodulation distortion]. In fact, the only tendency for the passive design to show a more positive result than the active design was in 'brightness', which could perhaps be a result of higher non-linear distortion levels. Perhaps the only real block to the general acceptance of the superiority of active crossovers and multi-amplification in the world of domestic hi-fi is the fact that so many hi-fi enthusiasts want to select their own favourite amplifiers and loudspeakers as separate items, but this perhaps has more to do with human psychology rather than audio engineering. Also, of course, choosing your own system with future up-grades in mind, as extra money becomes available, is a fun part of building up a hi-fi system, and perhaps a necessity in some smaller studios.

5.7 Physical derivation of crossover delay

Crossovers are more than simple filters. As we have seen, due to the group delays which exist in filter circuits, the outputs of the various filter sections are time-shifted by virtue of the phase shifts which are inextricably linked to the roll-offs. The result is that when the outputs are re-combined, either electrically or acoustically, there are often non-minimum-phase response irregularities in either the amplitude or phase responses. Essentially, they cannot be corrected by analogue, *electrical* means. Digital crossovers *can* be made to provide summing outputs, but the sonic benefits are not necessarily worth the efforts. For example, digital crossovers, unless they employ high sampling rates, (96 kHz or more) may introduce limitations in such a way that would make it difficult to accurately monitor analogue or high sampling rate digital recordings. What is more, a three-way stereo crossover would require six D to A converters on the outputs, and two A to D converters if they were being fed from analogue sources. If these were to be of the highest quality (and in a high quality monitoring system they should be expected to be nothing less) the cost of the unit could be exorbitant. Using anything less than the finest converters would make a mockery of trying to monitor *recordings* made through the best converters. This subject is discussed further in the following section.

There is, however, an analogue means of deriving delays. The loudspeaker cabinets, themselves, can be stepped, or the different drivers can be mounted in separate enclosures which are then mounted at different distances from the listeners. These two systems are depicted in Figure 5.24, but care needs to be taken to ensure that diffraction effects do not become problematical due to the increased number of cabinet edges. Figure 5.25 shows how a composite system using direct radiating bass drivers and a horn-loaded mid/high frequency driver can compensate for the delayed output from the bass units. In this case, the 'natural' position of the high frequency voice coil is *behind* the coil of the bass driver. All of these means achieve the same result by delaying the high frequency signals with respect to the low frequency signals. The distances between the voice-coils of the different drivers can be further physically off-set to whatever degree necessary to compensate for the electrically derived group delays. Given the speed of sound at 20 degrees C, each centimetre that a driver is mounted behind another (relative to their voice-coils) would give rise to a delay of 29.4 microseconds. Ten centimetres would therefore give rise to a delay of 294 microseconds, which would be sufficient to reverse the polarity of a wave at 1.7 kHz. (Look again at Figure 5.6.) Incidentally, the voice coils are used as an approximate reference for the source of the sound propogation because even though the diaphragms are ahead of the voice coils there is a finite time of propogation from the coil to the diaphragm face which, with most loudspeaker constructions, is roughly of the same order as the speed of sound in air.

5.8 Digital crossovers

As time passes, digital crossovers have become more commonplace in professional loudspeaker systems, although their use in domestic circumstances is still largely restricted to home recording facilities. They are particularly attractive because of the easy implementation of almost any amplitude response, phase response, signal delay, driver compensation and even room compensation. However, all this flexibility comes at a considerable cost.

Sonically, in terms of 'hi-end' hi-fi or high resolution studio monitoring, the highest fidelity can only be achieved if the sample rate and bit rate used in the crossovers at least equal those of the recording medium, or exceed the resolution of the ear and produce no audible artefacts. It may be difficult to hear the difference between a 20 bit/96 kHz recording and a 24 bit/192 kHz recording when listening to a crossover based on

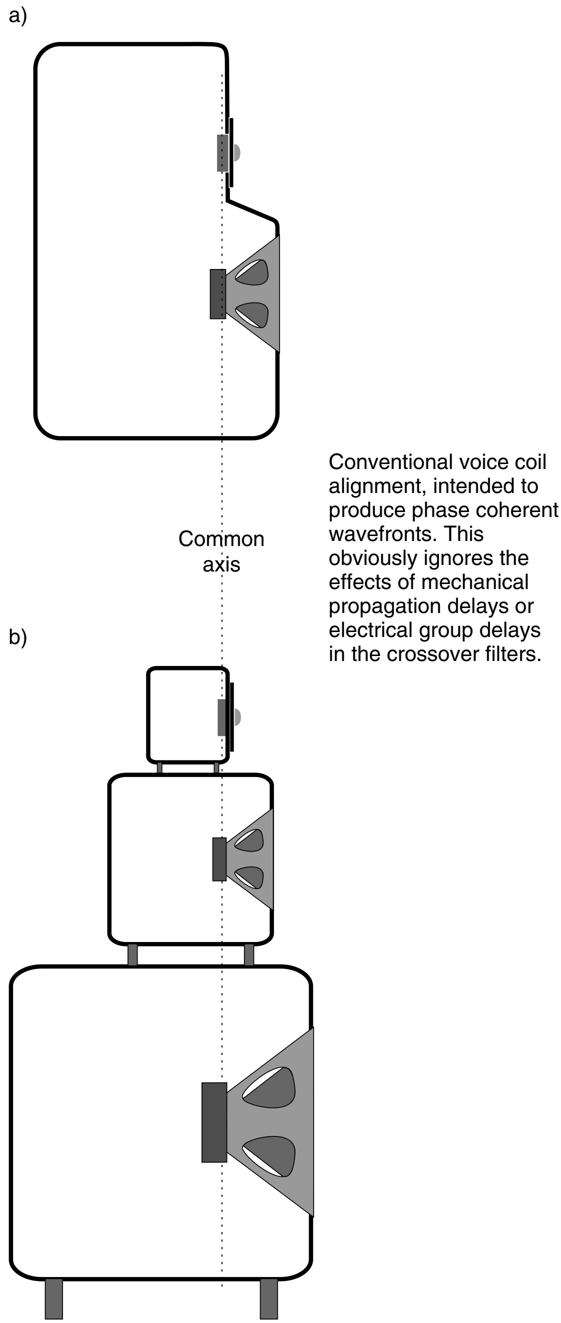


Figure 5.24 Methods for the physical compensation of propagation delays
a) The stepped baffle. b) Separate boxes

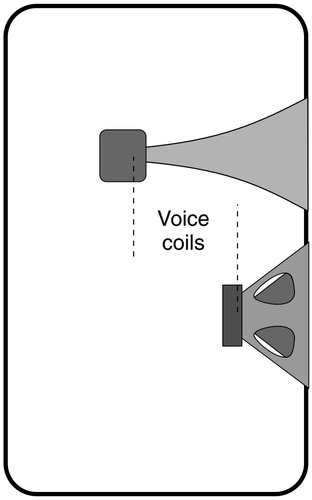


Figure 5.25 Flat baffle with a horn-loaded high frequency driver

In many cases it is possible for horn mounted compression drivers to align themselves very closely to the ideal whilst sharing a common, flat fronted baffle with a low frequency driver when the electrical group delays of the crossover filters are also taken into account, especially when steep-slope crossovers are used

16 bit/48 kHz processing. Even if a crossover has internal processing which seems higher than necessary, the resultant output after signal manipulation has taken place may not be as great as the marketing figures would suggest.

The main problem, however, is concerned with the converters. As the finest amplifiers are still of the analogue variety, D to A (digital to analogue) conversion must take place in each output of the crossover, and, for either professional monitoring or 'high-end' high fidelity, these crossover D to As must be of higher resolution than any other part of the signal chain. If they are not, then they will limit the quality and the resolution of the chain. As with the power amplifiers and loudspeaker cables (which will be discussed in the next chapter) the splitting of the frequency bands does offer some respite from the demands of handling the full audio bandwidth. Nonetheless, to achieve the highest levels of sonic quality, D to A converters are expensive, and a stereo, three-way digital crossover would require six of them. At the time of writing, and whilst analogue power amplifiers are the general order of the day, it would be reasonable to expect to pay 3000 to 5000 euros for those six converters. If analogue *inputs* were required, then the A to D (analogue to digital) converters may add another 1000 euros or so to the price of the crossover if the equivalent quality was to be expected.

Under less critical circumstances, digital crossovers can be very useful tools, but when they are user-programmable, they run the risk of being inappropriately applied. One must be very careful when trying to 'solve' amplitude/phase problems that the solutions do not go against the laws of nature. Straightening out the phase associated with an amplitude roll-off may be very tempting, especially where the application is at the extremes of the frequency bands, but the results can sound unnatural because they *are* unnatural. On the other hand, in concert sound applications, where subtleties are by no means as important as solving the normal problems faced by such events, digital crossovers have been an enormous step forwards. Their real drawback is their cost when they must operate at the highest levels of sonic transparency. In many such cases, analogue crossovers may simultaneously be simpler, cheaper, more robust, *and better*.

References

1 Theile, A. N., 'Loudspeakers in Vented Boxes', Part 1, Journal of the Audio Engineering Society, Vol 19, No 5, pp 382–392 (1971)
2 Theile, A. N., 'Loudspeakers in Vented Boxes', Part 2, Journal of the Audio Engineering Society, Vol 19, No 6, pp 471–483 (1971)
3 Small, R. H., 'Vented Box Loudspeaker System', Journal of the Audio Engineering Society,
Part I, 'Small signal analysis', Vol 21, No 5, pp 363–372 (1973)
Part II, 'Large signal analysis', Vol 21, No 6, pp 438–444 (1973)
Part III, 'Synthesis', Vol 21, No 7, pp 549–554 (1973)
4 Small, R. H., 'Closed Box Loudspeaker Systems', Journal of the Audio Engineering Society,
Part I, 'Analysis', Vol 20, No 10 (1972)
Part II, 'Synthesis', Vol 21, No 1 (1973)
5 Small, R. H., 'Passive Radiator Loudspeaker Systems', Journal of the Audio Engineering Society,
Part I, 'Analysis', Vol 22, No 8 (1974)
Part II, 'Synthesis', Vol 22, No 9 (1974)
6 Small, R. H., 'Direct Radiator Loudspeaker System Analysis and Synthesis', Journal of the Audio Engineering Society, Vol 20, No 5 (1972)
7 Campbell, A. M., Holland, K. R., 'Active vs Passive Crossovers for Mid-Priced Hi Fi Loudspeakers', Proceedings of the Institute of Acoustics, Vol 26, Part 8, pp 116–123, Reproduced Sound 20 conference, Oxford, UK (Oct 2004)

Bibliography

1 Colloms, M., 'High Performance Loudspeakers', 6th Edition [Chapter 6], John Wiley & Sons, Chichester, UK (2005)
2 Borwick, J., 'Loudspeaker and Headphone Handbook', Third Edition [Chapters 5 and 6], Focal Press, Oxford, UK (2001)

Chapter 6

Effects of amplifiers and cables

Clearly, no loudspeaker can do its job unless it is connected to a suitable power amplifier, and to make that connection, an appropriate cable is required. As no cable can *improve* any signal which it is passing (unless it is filtering out some other problem that should not be there) it must be concluded that the best cable is no cable. Moreover, as no amplifier can improve the accuracy of the signal which it is passing, the only conclusion which can be drawn is that the combination of amplifier and cable can only degrade the signal. However, as we *must* use an amplifier and a cable, art and science are both employed in order to minimise the inevitable degradation.

Well, at least that is the situation as far as accurate monitoring is concerned. In the case of amplifiers which are used for musical instrument amplification, the amplifier is actually a part of the instrument, so whatever sounds right *is* right. The gross distortion produced by an old, valve, Leslie tone-cabinet, when amplifying a Hammond organ beyond the point of overload, was one of the most emotive sounds in the history of rock music, but we will come to that in Chapter 8. Similarly, the popularity of valve amplifiers in many domestic hi-fi systems is widely attributed to the pleasing 'musical' sounding even harmonics which they tend to produce. Again, this will be discussed further in Chapter 8, but as these effects are totally subjective in terms of their desirability, it is very difficult to deal with the subject in any definitive way. Therefore, what we will attempt to do in this chapter is look at the problems of amplifiers and loudspeaker cables when *accuracy* of reproduction is the goal, and which we *can* deal with in an objective manner.

6.1 Amplifiers – an over-view

The purpose of a power amplifier is to take a voltage signal, usually in the order of up to one or two volts peak, and deliver the waveform of that signal in a way that a proportionally larger voltage can be supplied to drive current through a loudspeaker motor system. This, in turn, is used to generate a radiated acoustic output which is also the best replica possible, in terms of pressure fluctuations, of the input signal to the amplifier. Driving power into a resistor is a reasonably simple exercise, but driving the signal into a load similar to that shown in Figure 1.7 is a different matter. Driving power into a load such as that shown in Figure 6.1, an electrostatic loudspeaker,

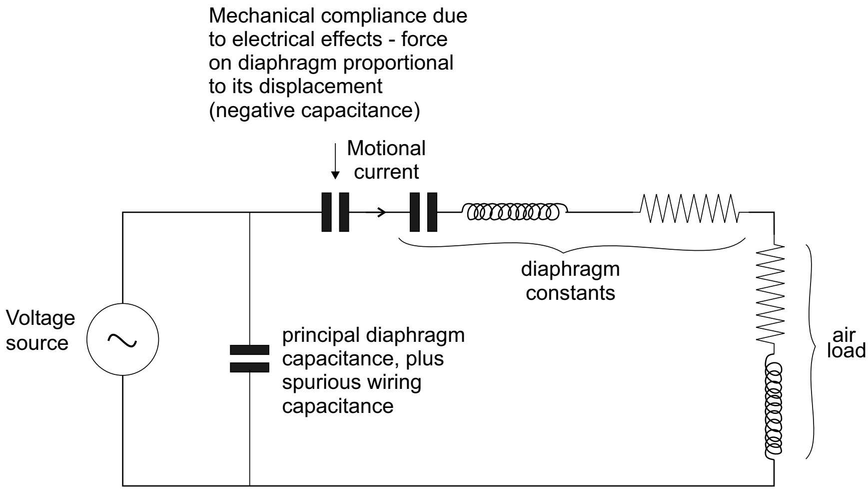


Figure 6.1 Load impedance presented by an electrostatic loudspeaker

can be an even more difficult problem to overcome, because the first thing that an amplifier 'sees' is a large capacitor across its output terminals. As explained in the previous chapter, the current and voltage are not in phase when passing through loads which are predominantly either inductive or capacitive.

Traditionally, power amplifiers have been rated into resistive loads. This makes sense, even though it is unrealistic, because there is no standard loudspeaker load – the range of variability is just too great – but driving a loudspeaker nevertheless tends to be very different from driving a resistor. There was, and perhaps still is to some extent, an objectivist lobby who claim that all well-engineered amplifiers exhibiting adequate output power and minimal distortion will sound identical. In the opinions of the authors of this book, this is absolutely not the case. Amplifiers *do* sound different, at least once they can be heard through high resolution loudspeakers in well controlled rooms. Undoubtedly, part of the reason *why* they sound different is due to their performance when driving loads which are typically represented by Figures 1.7 and 6.1, although numerous other factors also play their parts.

An interesting incident occurred at the Tannoy factory in Scotland[1], in the early 1990s, where a group of people of some repute in terms of their auditory acuity were assembled to select an amplifier to offer as standard with a new range of loudspeakers. There were four loudspeakers in the range, and four different amplifiers had been short-listed for audition. The intention was to select one amplifier from the four. However, in the blind tests, one amplifier was selected as sounding most accurate on one of the loudspeakers, another amplifier was chosen for another loudspeaker, and a third amplifier was deemed to sound most accurate on the remaining two loudspeakers of the range. The loudspeakers in question were all of similar design concept, but varying in size. The participants in the tests were all highly experienced and respected engineers, and all had been expecting that they would choose one amplifier for the whole range of the loudspeakers.

It could have been the case that certain characteristics of the individual amplifiers counterbalanced opposing characteristics in the different loudspeakers, and this could also have been due to the different load characteristics presented by each loudspeaker. The component values in terms of the circuit shown in Figure 1.8 could all have been different. However, we do not have anything like a perfect loudspeaker to use as a reference, so it is difficult to say which of two amplifiers is definitely better or worse when the differences are subtle and there is no 'audible quality meter' that we can connect to their outputs. In fact, we listen to amplifier/loudspeaker combinations, rather than just to amplifiers, and we can add listening rooms to those combinations in most circumstances. Given the fact that the rooms constrain the air which loads the loudspeaker diaphragms, it can be seen from Figure 1.8 that the rooms will also reflect back their presence into the loading of the amplifiers.

Ultimately, the ear is the only judge, and as it listens to the systems, it often cannot differentiate the sources of some audible effects from the sources of others. We also all have different ears, with different perception of sound, and as no microphone is perfect, we have no perfect sources of sound to compare anything with. To further compound all of this, unlike the optic nerve, which carries a measurable and recognisable signal from the eye to the brain, the auditory nerves disappear into about six different parts of the brain, and our understanding of how the mind puts the whole sound together is still not well advanced.

So just what *can* we say about amplifiers if they have to perform in such a murky world of variable loading and variable perception? This is a question that we must address in a very careful way if we are not to descend into the realms of misguided subjectivism, which may be appropriate in some circumstances, but which cannot lead to a robust consensus in professional circumstances.

6.2 Basic requirements for current and voltage output

Contrary to much popular thinking, amplifier power should not be matched to the power rating of the loudspeakers or individual drivers to which they are connected. The amplifiers should have a margin of at least 6 dB (four times the power) above the peaks of the highest level that they will be likely to be called upon to deliver. If it is to be presumed (somewhat sarcastically but nonetheless realistically) that most loudspeakers will at some times be used to their limits, then some common sense needs to be applied to their use because this means using amplifiers which can maintain an undistorted output beyond the peak level rating of the loudspeaker. However, the last requirement is not always easy to assess from any simple Ohm's law calculation. The reactive components in the impedance of a loudspeaker can give rise to significant phase shifts between the current and the voltage. If the phase shifts become significant, currents can be drawn which are totally beyond any expectations which would be calculated from the treatment of the load impedance as being largely resistive. Many amplifiers of seemingly adequate power rating have been shown to fail to be able to supply adequate current with difficult loads, even though their voltage headroom has been easily sufficient for their proposed use.

Such amplifiers may be found to be seriously lacking in the quality of sound that they can achieve with a loudspeaker which presents a highly non-uniform impedance to its output terminals, yet many of these amplifiers may be considered to give excellent reproduction when driving a more benign load. In general, it is loudspeakers with passive crossovers which present the greatest problems, at least when we are considering moving coil loudspeakers, but attempts to flatten the impedance curves by the addition of compensation networks in the crossovers has often been found to detrimentally affect the overall sonic performance of the system. In the previous chapter a strong case was made against the use of passive crossovers for very high quality loudspeaker systems, but many popular monitor loudspeakers still use this approach, so the problems which they give rise to still need to be considered.

6.3 Transient response

In the perception of music, the leading edge of a waveform is highly significant for the recognition of instruments. With the leading edges removed it can be difficult to tell the sound of a guitar from the sound of a violin, for example. It follows that the subtle differences between different guitars and different violins – and most other instruments for that matter – can be very dependent upon the accuracy of the transient waveform of the onset of the note. Apart from the need to supply adequate current when a bass drum is played loudly through the loudspeakers, slew rate and frequency response are also important factors. The former is measured in volts per microsecond, and is a measure of the speed with which the output of an amplifier can respond to the input signal. Forty volts per microsecond would be the typical slew rate capability of a good monitoring amplifier, but this same rate needs to be maintained at high levels, and not just at low levels, or slew limiting will occur, which can produce some very odd waveform distortion.

The transient response is also greatly affected by the frequency response of the amplifier. An electrical step function is shown in Figure 6.2. This waveform is also known as a Heaviside function, named after Oliver Heaviside. (Discoverer of the Heaviside layer which surrounds the Earth.) The generalised function is defined by

$$\mathrm{H}(x) = 0 \quad \text{for } x < 0$$

$$\mathrm{H}(x) = 1 \quad \text{for } x > 0$$

Its value at $x = 0$ is not defined

It is an all or nothing function, also known as a step function, and it contains all frequencies. A battery switched on and off, with a few seconds between switching, is a form of step function generation. If connected to the input of a spectrum analyser, a 1½ volt battery will demonstrate

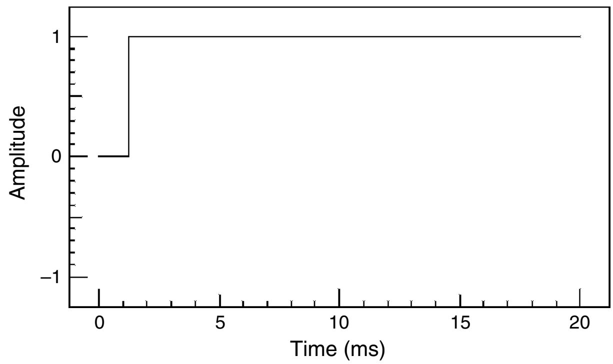


Figure 6.2 Waveform of a step-function (Heaviside function)

the wide frequency distribution. When connected or disconnected, all the columns of the analyser will be seen to be excited, giving rise to a straight line response with a roll-off of 3 dB per octave relative to the flat response of pink noise.

If an amplifier has high and low frequency roll-offs which begin too close to the audible frequency extremes, it will exhibit phase shifts as shown in Figure 6.3. The instantaneous onset of the step function requires that all the frequencies begin at time zero. Any phase shifts present in the system will give rise to signal delays, which will smear the waveform. Figure 6.4 shows a step function after passing through an amplifier whose response is 3 dB down at 15 Hz and 30 kHz in (a), then 3 dB down at 30 Hz and 15 kHz in (b), with 12 dB/octave roll-offs in each instance. Rates of change of phase beyond about 5 degrees per octave in the audible range are noticeable to most people, at least when demonstrated on monitoring systems which are themselves relatively phase-accurate. However, many monitor loudspeaker systems have phase responses that are so appalling

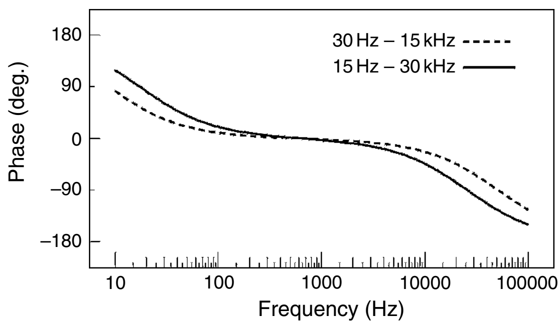


Figure 6.3 Roll-offs and their corresponding phase shifts

Changing phase-shifts associated with different bandwidths. In the case shown, the roll-offs are second order – 12 dB/octave. When the rate of change of phase exceeds about 5 degrees per octave in the audio band, noticeable changes in the timbre of musical instruments become noticeable in high quality monitoring conditions (especially with loudspeakers exhibiting fast transient responses in well-damped rooms)

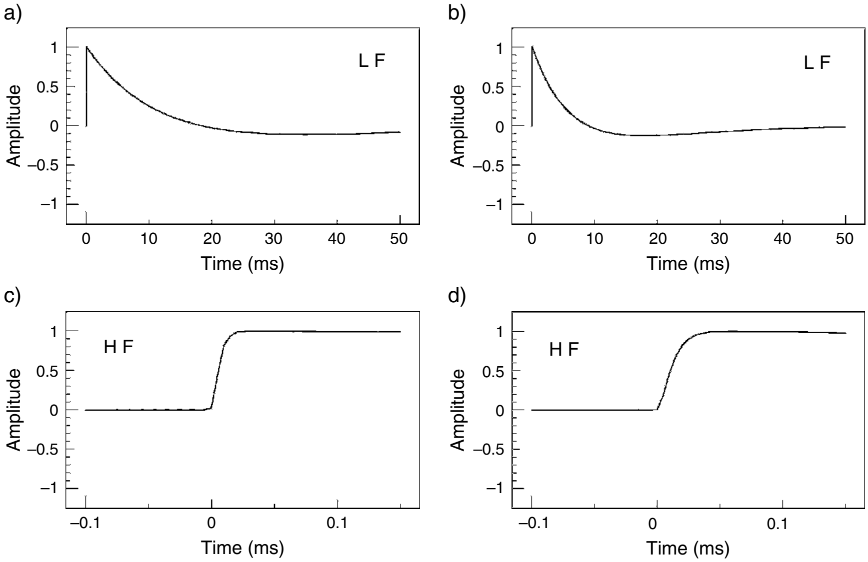


Figure 6.4 Effects of bandwidth on transient responses. Note different time scales

The result of passing the waveform shown in Figure 6.2 through the filtered bandwidths shown in Figure 6.3. The plots (a) and (b) show the decay of the waveform when subjected to second-order roll-offs (12 dB/octave) at 15 Hz and 30 Hz respectively. Note the tendency for the 30 Hz roll-off (b) to overshoot the zero axis as it decays. The effect manifests itself as 'ringing', or resonance. The time (transient) response takes longer to return to a flat, zero line than is the case with the 15 Hz response shown in (a)

The plots (c) and (d) show the effect of filtering the response at 30 kHz and 15 kHz respectively. It can clearly be seen how the response filtered at 15 kHz (d) takes longer to rise to its maximum value than in the case of the 30 kHz roll-off shown in (c)

The higher frequency of roll-off at the low frequencies therefore can be seen to extend the decay of a signal, whilst the lower frequencies of roll-off at high frequencies can be seen to extend the attack time of a signal. In either case, as the bandwidth is restricted, the time (transient) response is lengthened

that they can swamp the audible effects of phase shifts elsewhere in the system, rendering them inaudible, but this can hardly be considered to be 'monitoring'. In general, it is desirable for a good monitor amplifier to have a flat amplitude/frequency response between around 5 Hz and 80 kHz – two octaves either side of the audible frequency range – if good transient responses are to be expected.

It should be noted that phase shift, however, is something which exists only in the frequency domain, because it is the rate of change of phase with frequency. In the time domain we should speak of the phase slope, which gives rise to the phase distortion that changes the waveform shown in Figure 6.2 to those shown in Figure 6.4. Under good monitoring conditions the transient response of a step function can be heard to change as the bandwidth is reduced. The bandwidth can clearly be demonstrated to have an effect on the attack of a tight sounding bass drum or tom-tom, although, as mentioned above, it takes a good loudspeaker system to show the effects clearly.

This brings us to another 'chicken and egg' situation, because when using loudspeakers with poor transient responses the benefits of using a fast responding amplifier may not be appreciated. This can lead to people drawing the conclusion that a wideband response of an amplifier is unnecessary, and reasons may then be found to curtail the bandwidth, applying the philosophy that "the wider the window is opened, the more dirt blows in". Given the ever-increasing amount of electromagnetic radiation in the atmosphere, there can be a great temptation to filter it out, stage by stage, but the phase shifts through the filters of a recording chain are cumulative. Closing the window unnecessarily in the monitor system is unwise if the monitor system is to be capable of revealing transient problems elsewhere in the system. Maintaining wideband frequency responses throughout the electronic system is very important for maintaining the transient accuracy of the chain.

6.4 Non-linear distortions

Until now we have been speaking about the *linear* distortions of the amplifiers. A linear distortion involves a change in the waveshape, or the frequency response, but where no new frequencies are introduced that are not present in the input signal. If either the amplitude or phase of the frequency response are distorted, they are linear distortions, and will affect the waveform. Academically speaking, the full frequency response includes both the amplitude *and* the phase responses, although in more general conversation the term 'frequency response' usually simply relates to the amplitude portion, variously known as the magnitude or modulus of the frequency response.

Unlike the linear distortions, a *non*-linear distortion is one where new frequencies *are* introduced into the output of a system. In the case of *harmonic* distortion, these extra frequencies are multiples of the input frequencies. *Intermodulation* distortion occurs when the various frequencies in the input signal interact to produce sum and difference frequencies which may have no harmonic relationship whatsoever to the input frequencies. Noises and rattles also constitute forms of non-linear distortion in mechanical systems. In general, the non-linear distortion in well-designed electronic systems is way below the levels produced in loudspeakers. This situation has given rise to suggestions that the non-linear distortion in amplifiers will be insignificant when heard below the much higher levels of loudspeaker non-linearities. However, electronic and electro-mechanical distortions are produced by very different mechanisms. Loudspeaker distortion tends to be benign compared to electronically-produced non-linear distortion, which tends to be much more subjectively disagreeable.

Intermodulation distortion has always been an elusive property to measure. It depends on:

1) the signal level
2) the bandwidth of the signal
3) the complexity of the signal
4) the peak to mean ratio of the signal

5) the signal waveform
6) the interaction between any of the above, and a number of other factors

Historically, *and* currently, harmonic distortion is still the measured quantity, but harmonic distortion, alone, is not necessarily either unmusical or unpleasant. Referred to the above list, harmonic distortion for any given frequency is only dependent upon level, because a sine wave:

1) has no bandwidth
2) has no complexity
3) has a fixed peak to mean ratio
4) has a defined waveform
5) cannot interact with itself

Chapter 9 will deal with the deeper aspects of the relevance of distortion more closely, but these points need to be made now because the ways in which amplifiers produce non-linear distortions is much more varied than would be expected simply from reading the brochures, where all good quality transistor amplifiers tend to produce approximately similar levels of harmonic distortion in the range of 0.01%. As such low levels of *purely* harmonic distortion are almost certainly not detectable by the human ear, it implies that other forms of non-linear distortion are the culprits where there is harshness, lack of clarity, lack of transparency, lack of 'air', or where other similar descriptions are used to qualify the less than desirable sound of any amplifier, or the difference between amplifiers. The situation is that from published distortion figures alone, there is little that can be implied about the musical accuracy of an amplifier. Conclusions can only be drawn about the sound of an amplifier by listening to it in the specific system with which it is intended to be used. However, the 'biasing class' can give some guide to possible performance under specific circumstances.

6.5 Amplifier classes and modes of operation

There are many different designs of amplifier output stages, but they are all usually grouped into a system of classification relating to their output biasing or switching. Classes A, B and C are unswitched stages, whilst Classes D, E, G and H are switched stages. As with so many things, there are pros and cons to each design, and amplifier designers must use their experience to decide which class seems to be most appropriate for the intended use of the amplifier. Although this is an enormous subject in its own right, we need to at least outline the concepts here, because the choice of the type of amplifier can significantly affect the characteristics of a loudspeaker system in ways which may not be apparent from simple, traditional measurement techniques.

As discussed in the last sections, things such as intermodulation distortion and instantaneous current capacity into reactive loads can give rise to distinct sonic differences between amplifiers of different design implementations. As the above two problems are unlikely to be shown up by any normal static tests of amplifiers performance, they may well not appear on any specification sheets. Until now, there has not arisen any standardised test for intermodulation distortion with any close correlation to subjective listening assessments. Although proposals have been discussed[1,2], the number of variables listed in Section 6.4 may still mean that, in many cases, only certain types of music at certain levels and with certain combinations of instruments will lead to problems. The question then arises as to how to cover all eventualities. Clearly, amplifiers should be free of problems to the greatest degree possible, but at what cost? One approach may be to build an amplifier with big reserves of performance, but if its construction is too big and heavy to fit in a small, self-powered monitor system, it is no option at all. However, in this discussion we will try to stick to performance, and only mention practical considerations when they can be seen to affect the overall thinking.

6.5.1 Class A amplifiers

Class A biasing has always been highly regarded. In high quality *line* output stages, such as in microphone pre-amplifiers, equalisers or compressors, it is popular, but as these devices are rarely capable of supplying more than about one watt, the problems of the inefficiency of Class A do not usually arise. Class A refers to continuous conduction through the output device(s), none of which are ever cut off. To allow for symmetrical clipping, and to maximise the output voltage for a normal musical signal, the no-signal current passing through the devices is set to 50% of the maximum output current. However, by using this means, *twice* the rated output power of the amplifier will be dissipated whether any signal is present or not. As discussed in Section 6.1, many loudspeakers require large, transient voltage and current swings due to the reactive components of their impedance. This may require an output capability in watts that is rarely, if ever, used, just to have the individual swings available when necessary. Nevertheless, in Class A, *double* that output power would always be being dissipated by the amplifier during all the time when it was switched on, except when it was supplying a portion of it into the loudspeaker load. For this reason, beyond about 50 or 100 watts, Class A power amplifiers usually become impractical.

In a recording studio, for example, where perhaps a 500 watt amplifier would be needed for each loudspeaker, a pair of Class A amplifiers would be dissipating 2000 watts all the time that they were switched on. If the amplifiers were in the control room, additional air-conditioning capacity would be needed to remove that heat. More probably, such hot devices would be mounted outside the control room, because they would be likely to need cooling fans, the noise of which would not be desirable in any listening room. Unfortunately, this would perhaps be in conflict with the need for short loudspeaker cables, as will be discussed later in this chapter. Therefore, with the amplifiers being asked to supply a maximum of about 100 watts, average, of music signal, during maybe 10% of the working day, perhaps 3 kW of electricity would be needed to both supply the amplifiers *and* keep them cool. Convection (fan-free) cooling *is* possible on 500 watt Class A devices, but the heat sinks would need to be huge and the amplifiers would need to be mounted in an open, well-ventilated place. The power supply components would also need to be larger than for an amplifier of similar output power but of one of the other classes of output stage. Large Class A amplifiers therefore tend to be expensive to build, expensive to run, difficult to site conveniently and generally very wasteful of electricity.

Despite these drawbacks, it must still be appreciated that Class A designs do have advantages over many other amplifier classes.

1) There is no crossover distortion where signal is transferred from one output device to another because no such transfer occurs. This type of distortion can occur in some power amplifiers at very low levels, but is almost non-existent at the high levels at which most amplifier distortion measurements are taken. For this reason, the authors have applied Class A amplifiers to the highly sensitive compression drivers in some studio monitoring loudspeakers, where, at a level of 70 dB at the listening position, only around 1 *milli*watt of power is being taken from the amplifier. In such instances, with sensitivities reaching 110 dB SPL for 1 watt input at one metre, any low-level distortion in the amplifier could be highly detrimental to the sonic transparency of a monitor system.
2) Given the constant total current draw of the amplifier, irrespective of signal levels, the power supplies, both AC and DC, are not subjected to the current surges which can inject distortion into the low level stages of the amplifiers, or other surrounding equipment.
3) There is no doubt that Class A circuits are simpler to realise, and there is a general tendency for simple circuits to do less sonic damage to the signal. However, simpler circuits may take longer to stabilise, and for this reason there has been a tendency to switch-on Class A amplifiers well in advance of their expected time of use.
4) The harmonic structure of any distortion products which do arise tends to be more benign, sonically, than those produced by many other classes of amplifiers.
5) Being inherently of lower distortion than many other amplifier designs, the Class A amplifiers can often exhibit greater tolerance of complex loudspeaker and cable loads, because the lesser degree, or non-existence, of global negative feedback may render the amplifiers less prone to disturbance by the complex back EMFs generated in the complex loads.

6.5.2 Class A derivatives

A number of ways have been devised to try to maintain as many of the advantages of Class A whilst reducing the total power consumption, and hence the weight, heat, cost and inconvenience of pure Class A designs. *Class A sliding bias* is one such means, where small signals experience Class A conditions, but where, as the signal increases, the larger signals experience Class A/B conditions. *Super Class A* is perhaps not what it at first appears to be, but is in fact a low level Class A amplifier in series with a Class B amplifier for the higher level signals. It is, in fact, a Class A + B design. In reality, although the maximum dissipation is lowered, some of the beneficial aspects of pure Class A operation are lost. The name smacks more of marketing than engineering.

On the other hand, *Dynamic Class A* maintains most of the benefits, if not all, of pure Class A. With these circuits, even in the late 1970s, distortion levels of as low as 0.03% were being achieved from transistor amplifiers without global negative feedback, which is in the order of only 1% of the open-loop distortion of many Class AB designs. Intuitively, such a low distortion basic design always seems to be a good starting point.

6.5.3 Class AB

Pure Class B was originally developed to save battery life in portable amplifiers, such as in battery powered radios of either valves or transistors. In this mode, only one transistor of a push-pull pair is ever conducting, depending on whether the signal waveform is positive or negative. Class B amplifier produce only around 10% of the waste heat of Class A amplifiers, but, as each half of the output stage is baised almost to cut-off on no signal conditions, there can be audible unpleasantness arising from the discontinuities in the signal waveform as the handling of the signal passes from one transistor to another. Pure Class B amplifiers are therefore not used in high fidelity audio applications. Nonetheless, by adjusting the bias on the transistors in the direction of Class A biasing, Class AB working can be achieved. Unlike in the cases of Dynamic Class A or Sliding Bias Class A, Class AB amplifiers are not operating in true Class A conditions when operating at low levels. Particularly after loud passages of music, the thermal delays in the output devices can give rise to bias changes, which can lead to a form of crossover distortion in some designs when one side of the output devices does not perfectly mirror the other side. The potential unpleasantness of this type of distortion relates to the production of high order harmonic and inharmonic spectral products.

On the other hand, Class AB amplifiers can be very practical devices, and advancing technology has continually sought ways to overcome the limitations of Class AB. In many cases, these limitations *have* been overcome to a very great degree, and Class AB amplifiers are, at the time of writing, by far the most numerous in use in professional audio. Output stage biasing can be critical, and after heavy use and many thermal shocks, the transistor parameters tend to change as time passes. Class A circuits are relatively immune to this process, partly because of less sensitivity to precise biasing, and partly due to their more constant temperature of operation, albeit high for much of the time. (Constant high temperatures tending to be less damaging in the long term than temperatures which are always changing.)

6.5.4 Class D

Class C amplifiers have no place in audio, but do find use in radio transmitters. Class D, on the other hand, is an emerging technology. Class D amplifiers are frequently referred to as *digital* amplifiers, but some designs are better described as switching amplifiers. They essentially consist of a switch-mode power supply, supplying current into the load (loudspeaker) under the control of the audio input waveform, with the output being low-pass filtered in a similar manner to digital-to-analogue converters. Class D amplifiers are light in weight and can be relatively cheap to build. However, when extreme high fidelity is required, things can become more difficult to achieve. They are also very energy efficient, beginning from about 75% at 5% power to beyond 95% at full power. Sonic performance improves year by year. In self-powered loudspeakers they can find willing partners because of their small size, low cost, and low heat generation. However, they still can be prone to the emission of troublesome electromagnetic interference (EMI) because of the high switching frequency used, (and the whole concept of switching, itself). Of course, they all must meet current electromagnetic compatibility (EMC) regulations, just like the office fax machine and digital radio/alarm clock, but, as this is being written, those devices must often be switched off in order to clearly hear the BBC World Service on a small, portable radio. Fast switching always generates harmonics into the megahertz regions, and such emissions have a great potential to interfere with nearby electronic equipment.

In other words, compliance with EMC regulations is one thing, but being a good neighbour with the rest of the sensitive equipment in a recording studio is another thing. This is especially so when ‘vintage’ equipment is in use, designed before there was a need to even think about digital switching transients. Nevertheless, as time goes on, improvements will be made, and Class D amplifiers are steadily progressing. The promise of an efficient, cheap, lightweight *and* high quality amplifier is a strong spur to commercial development. At the time of writing, some questions still exist about the sonic performance of the top two octaves of the audio frequency range in Class D circuitry. The very low level switching artefacts which many designs exhibit can prove to be problematical when using high sensitivity loudspeakers, such as mid-range and high frequency horns. Achieving low noise and distortion at both low and high power still presents many design difficulties.

In effect, there are two types of Class D amplifiers. Early designs used analogue inputs, which were compared to a triangle wave running at the switching frequency. This type of switching, when the audio signal crossing the triangle wave causes the output to switch, gives rise to a PWM (pulse width modulation) output. Some later designs, however, accept a linear PCM (pulse code modulation) digital input, and use more sophisticated modulation techniques. As such, these designs are effectively digital until the output filtering which removes the unwanted switching artefacts. The output transistors are switched on and off typically at a rate from 100 kHz to 1 MHz, so switching noise is only to be expected. It also leads to quantisation noise, because the output switching rate is finite.

However, factors such as the need for very low-jitter clocking and the use of air-cored output filter inductors are complications which are not always easy to solve. Clocking errors and ferrite inductor cores can lead to non-linear distortion. The filtering is necessary because the direct output waveform is a high frequency square wave, which, if left unfiltered, could radiate large amounts of radio frequency (RF) interference from the loudspeaker cables, which can act as transmitter aerials. Output inductors that behave well at these frequencies but which do not exhibit much loss at audio frequencies can be difficult to make. The required clocking accuracy of less than 100 picoseconds may also be hard to achieve.

In order to reduce the potential for clashing clocks, it can be necessary to synchronise the clock of the switch-mode power supply with the clock of the amplifier, which can also be beneficial in ensuring that the maximum current is available exactly when needed. The demands made of the power supplies are quite exacting. Any changes in power supply voltage causes a proportional change in the output signal, whereas in Class A, the current drawn is relatively constant, and in Class AB designs the audio feedback circuitry tends to compensate for the moderate fluctuations. The power supplies for Class D amplifiers may also have to deal with absorbing the power which can be reflected back from the output filter inductors.

6.5.5 Class G and H

There seems to exist some geographical confusion about the application of the terms G and H. However, in both cases, they involve multiple or variable supply rails, the higher voltages only being brought into play via fast switching circuitry when high output is called for. Thus, a Class AG or AH amplifier can operate in Class A at low signal levels, without excessively wasting heat, but higher-voltage power supplies are brought into service when the required output level of the signal exceeds the capabilities of the first rail. The high voltage supplies are therefore only used when needed, greatly improving the overall efficiency of the amplifier. Overall performance can be excellent, with the waste heat, weight and component costs being almost as low as Class D, but with the distortion products being almost as low as pure, simple Class A.

6.6 MOSFET or BJT?

The bipolar junction transistor (BJT) has long been generally preferred to the MOSFET for high quality, professional power amplifiers, but some excellent MOSFET designs do exist. [MOSFET, according to the "Power Mos Fet Data Handbook, Page 7, published by Hitachi in August 1985, standing for *M*etal, *O*xide and *S*ilicon, *F*ield *E*ffect *T*ransistor – although other meanings of S are sometimes to be seen in print.] Hitachi was the first company in the world to produce 100 watt, complimentary power MOSFETs, in 1977. Somewhat like valves (tubes) MOSFETs are voltage controlled devices. BJTs, on the other hand, are relatively low impedance *current* controlled devices, so the driver stages which feed them need, themselves, to be small power amplifiers. The concept of operation is therefore very different, as can be the circuitry which surrounds them, so it is hardly surprising that there can be sonic differences between the amplifiers which employ the different devices. Another difference exists in the minimum resistance through the devices when they are turned full on. The BJT tends to have lower resistance than the *lateral* MOSFETs used in audio. There are vertical MOSFETS, which do have lower resistance, but they are not suitable for high power audio use. The MOSFETs also exhibit higher input capacitance than BJTs, and this must be taken into account when designing the driver stage which comes before the power output devices. All of these factors conspire to change the circuit concepts in ways which may have audible repercussions.

When MOSFETs fail, they tend to do so in ways which are less disastrous to loudspeakers and driver stages than do BJTs. The MOSFETs themselves are generally also more tolerant to abuse than BJTs. They tend to thermally protect themselves, and are not prone to the 'thermal runaway' that can let BJTs get totally out of control. Theoretically, power MOSFETs also exhibit a higher bandwidth than BJTs, though to some peoples' ears, the BJTs sound sweeter. However, none of these apparent advantages and disadvantages are clear-cut situations. There are simply so many circuit possibilities for each type of device that many other factors contribute more to the sound of an amplifier than the simple choice of MOSFETs or BJTs. Furthermore, both types of device are still being developed, and the pendulum can swing with the arrival of new devices or circuit concepts.

6.7 Choosing an amplifier

In many ways it is advantageous that so many design options are available, because different types of loudspeaker drivers, different types of enclosures, differing requirements in terms of power levels, different loudspeaker sensitivities, different musical styles, and a host of other variables means that there is no, single amplifier to fit all needs. There is much casual talk about the pros and cons of different amplifier components or characteristics, but often it is of little value because it is applying specific circumstances to the general cases – 'Cats have tails, my dog has a tail, therefore my dog is a cat. Fact!' Or so go many of the arguments!

In general, well-designed, well-engineered amplifiers, specifically tailored for their purposes, can give excellent results despite employing many different technological approaches. Difficult load impedances may be handled more easily by some designs than by other designs; transients with heavy low frequency content may favour one amplifier whilst smooth string sections may favour another. Clearly, what we all would like is the perfect amplifier to suit all applications, but such is not the nature of compromise, and in the commercial world, the marketing and business people now hold more sway over the designers and engineers than ever before. Even in the case of small, powered monitor loudspeakers which unashamedly market themselves as fully-professional monitors, there is usually nothing in those amplifiers which could be considered to be 'over the top'. They are almost always engineered to a price and performance which is no more than necessary to satisfy the majority of their customers. One perceived problem with powered monitors, therefore, is the lack of ability to upgrade the system. Indeed, a 12,000 dollar Krell amplifier can 'improve' the sound of a pair of Yamaha NS10s (600 dollars) compared with their use with a more modest amplifier, but the two could hardly be marketed together as a viable package.

Inevitably, therefore, a question of balance tends to pervade the subject of design selection. Questions such as whether mass sales are envisaged, or whether the designs are to be tailored to a very specific use, all need to be taken into account, but the chosen points of balance may also be influenced by the personal philosophies and the order of the priorities of the individual designers. We have already discussed how high sensitivity compression drivers can expose any low-level distortion in the amplifier, such as crossover distortion, whereas an HF driver 20 dB less sensitive, needing 100 times the power for the same acoustical output, may highlight other aspects of an amplifier's failings. The choice of which amplifier to use for the high frequencies of a system can therefore depend very much on the choice of tweeter, even if the desired maximum output SPL is the same in each case.

Undoubtedly, the most difficult task to ask an amplifier to perform is to deliver the highest quality when delivering a high power, full frequency range, ultra low distortion signal into a loudspeaker with a passive crossover having a 'difficult', widely varying input impedance. An amplifier to suit this purpose is probably going to be big, heavy and expensive. On the other hand, equal sonic quality may be achieved from amplifiers of much greater simplicity by splitting the frequency range with an active crossover, then driving each loudspeaker motor separately. In this case, the amplifiers would probably need to be of lower power, and as they would each be driving a much more limited frequency range, and hence a less complex waveform, the ultra low distortion would be easier to achieve from smaller, lighter, less expensive amplifiers. The concept of the system therefore has to be defined before the most appropriate amplifiers can be chosen.

What *is* most appropriate in each case is very much governed by what an amplifier is being called upon to do. In general, at low frequencies, a high current capacity is beneficial, and an amplifier which can handle large current surges with ease will tend to produce a tight and effortless-sounding bass. At higher frequencies, low levels of intermodulation distortion are essential for maintaining an open, sweet, transparent sound. Wide, flat frequency response bandwidth is also essential if the transient response of the entire system is not to be compromised, though it should be obvious that extending the -3 dB bandwidth from 30 Hz down to 10 Hz would hardly be relevant for an amplifier which was only being used on mid and high frequencies. Neither would the response above 20 kHz be too meaningful for an amplifier which was only being asked to drive bass frequencies. In general, the job of any amplifier is greatly simplified when its range of use can be restricted to four or five octaves, rather than ten or eleven. In itself, the band-splitting results in greatly lowered intermodulation distortion.

The same is somewhat more obviously true for loudspeaker drive motors, where the high power, low distortion, wide directivity, high sensitivity, wide bandwidth loudspeaker driver does not exist. Many of the advantages of band-splitting were discussed in the previous chapter, and more will be said about it in the final sections of this chapter, but it is now well-recognised that for any part of an audio system which either delivers power, or has mechanico-acoustic properties, eleven octaves is a very wide range of frequencies to handle together. Amplifiers which are best suited to very low frequency reproduction may not be the best for reproducing high frequency subtleties. *Or*, perhaps one amplifier *could* do both jobs equally well, but perhaps not both at the same time. In fact, the same concepts of band-splitting and the subsequent selection of appropriate characteristics for each band can also be applied to the choice of loudspeaker cables.

Whilst this book is primarily about loudspeakers, and not amplifiers, the problems given rise to by the complex impedance which many loudspeakers present to the amplifiers has inevitably required at least a brief look at amplifier technology, or the repercussions of the impedance problem would not be fully understood. Likewise, perhaps now we need to consider a little more about the component part which connects the amplifier to the loudspeaker – the loudspeaker cable – because the subject is not quite so simple as some people would have us believe. (Although neither is it quite as shrouded in Voodoo as *others* may have us believe!)

6.8 Loudspeaker cables and their effect on system performance

Few subjects in the world of audio systems excite so much controversy and heated debate as the subject of loudspeaker cables. A truly enormous amount of pseudo-science has been written about this subject, and, once again, so many cases have been reported which have tried to extrapolate from the specific to the general case, which is clearly nonsense. In the world of high fidelity, there are many esoteric designs of amplifiers and loudspeakers which do not show the robustness of more typically professional equipment, and they tend to be used in domestic circumstances where electrical installations will not have been made with the type of attention to detail that would be found in a top professional recording studio. In some of these cases, a certain cable may clean up a sound, but the same cable in different circumstances, such as when used with professional equipment with clean electrical supplies, may not lead to any sonic improvement whatsoever. Granted, an inadequate cable can certainly degrade a system, but there is no justification for using inadequate cables if serious listening is contemplated. So, perhaps we should look at what we mean by adequate.

6.8.1 The bare minimum

A loudspeaker cable, like an electrical power cable, needs to have a current carrying capacity such that it will not overheat, and a voltage insulation such that it will not arc, but the current carrying capacity of a loudspeaker cable cannot be calculated from the simple $W = I^2R$ formula that would apply to the wiring of an electric heater. What is more, loudspeaker cables may be carrying 11 octaves of frequency range and not just the single frequency of an electrical supply, so what happens over the whole range of operation is of interest, and all frequencies must be passed as uniformly as possible.

In the case of the vast majority of transistor amplifiers we are dealing with what amounts to a constant voltage source. This means that from the minimum rated load impedance up to infinity, the output voltage of the amplifier, for any given input voltage and at any frequency within its range, will be independent of the load to which it is connected. However, this is not the case with valve (tube) amplifiers, whose output loads must be critically matched to the output impedance of the valves, usually via a relatively complex output transformer, but for now we will restrict our discussion to the more typical transistor amplifiers.

In order to behave as a voltage source, and remain independent of load, the output impedance of the amplifier needs to be very low indeed – typically hundredths of an ohm. The ratio of the load impedance to the output impedance gives us the *damping factor,* which is an indication of the ability of the amplifier to suppress the natural resonances within a loudspeaker. This electrical damping effect can be easily tested by lightly tapping the cone of a low frequency loudspeaker with the amplifier connected but turned *off*, then comparing the sound by tapping again with the amplifier switched *on*. A significantly deader sound should be heard in the latter case, when the near zero output impedance of the amplifier short circuits the voice coil. A similar effect can be demonstrated without an amplifier, simply by connecting a wire between the loudspeaker terminals, thus short-circuiting the coil.

When a loudspeaker diaphragm is struck, it behaves like a microphone. In fact, when a loudspeaker is connected to a microphone pre-amplifier input it makes quite a *good* microphone. The movement of the diaphragm in response to vibrations in the air moves the coil within the field of the magnet, which generates a voltage across the terminals. If the terminals are short-circuited, a current flows which tends to drive the voice coil in a direction which opposes the resonant movement of the diaphragm. The shorted coil acts as a dynamic brake. The damping effect of an amplifier can greatly 'tighten up' the sound of the bass, because it allows the amplifier to better control the resonant movement of the low frequency drivers. In other words, the transient response is improved. The need for a very good short circuit is important, so if the resistance/impedance of a loudspeaker cable is not close to zero, the damping will not be as close to perfect as possible, and therefore the effect of the amplifier on the loudspeaker resonance will be reduced.

When dealing with such low impedances as are found in typical loudspeaker circuits, such as 4 ohms or less, we are dealing with impedances much lower than would normally be encountered in general electrical wiring. In order to achieve a moderately good damping factor of 40 on a 4 ohm load, the total of the series impedance exhibited by both the amplifier output impedance *and* the cable impedance could not exceed one tenth of an ohm.

The principal concern of an electrical engineer, when choosing a gauge of cable, is that it will pass the required current without either overheating and producing a fire risk, or causing a voltage drop due to its predominantly resistive impedance (at 50 or 60 Hz) which would reduce the effectiveness of the device to which it is connected. So, if the motor is running at its correct speed and the cable is stone cold, then that is more or less the end of the story from an electrician's point of view.

Conversely, far from dealing with a fixed voltage at a single frequency, a loudspeaker cable has an impedance which may be dominated by its resistance only at low frequencies. At high frequencies the *inductance* can be contributing more to the impedance than the resistance, and a loudspeaker working on the end of a 10 metre cable producing the same SPL as when connected to the output of the amplifier with 50 cm of cable is no indication that the cable is lossless. If a cable has a *resistance* of 1 ohm and is connected to an 8 ohm loudspeaker, the voltage drop across the cable would be one part in 9 (i.e. across 1 ohm out of 9 ohms total). This represents about 11%, or about 1 dB, which would be barely audible. For a voice announcement installation, such a cable may be entirely acceptable, but for a high quality music system it would impose a serious limit on the damping factor, and hence on the accuracy of the transient response at low frequencies. The subsequent resonance due to the loss of damping may even restore the 1 dB loss of perceived volume, but the nature of the sound would have changed.

From the point of view of the loudspeaker, the cable is a part of the output impedance of the amplifier, and the damping factor is defined by:

$$\frac{\text{Z load}}{\text{Z source}}$$

In other words, the load impedance divided by the output impedance. Therefore, if an 8 ohm load was connected directly to an amplifier having an output impedance (source impedance) of 0.1 ohm, the damping factor would be 8/0.1 or 80. If the two were now connected by a cable having a resistance of 1 ohm, the damping factor would be dependent on the combined resistance of the source and the cable: $1+0.1$ ohms. The resulting damping factor would be:

$$\frac{8}{1.1}=7.2$$

which in audio engineering terms is poor, and likely to lead to subjectively woolly bass.

At higher frequencies, the drive units which are used do not generally have the mass or compliance to freely resonate, and are additionally resistively damped by the air load. However, at these frequencies the cable *inductance* can behave like a frequency dependent resistor, and act as a filter. When used with loudspeakers which present variable impedance loads with relation to frequency, both the resistance and the reactive components of the impedance can act as potential dividers, giving rise to a frequency response that varies according to the relative values of cable impedance and loudspeaker input impedance. The loudspeaker response will then vary in different ways as cable length is varied.

Cables also have capacitance between the conductors, but this is not usually problematical because the capacitive reactance is so low compared to the impedances of loudspeaker circuits that its effect would not normally become apparent until hundreds of kilohertz. However, it has been known to affect some marginally stable amplifiers, although it could be said that these problems should be solved at source. Some cables are intentionally capacitive, up to 0.2 microfarads, to maintain the high frequency response at the loudspeaker terminals, but they may unpredictably alter the performance and stability of amplifiers.

6.8.2 The status quo

In terms of professional use, a marginally stable amplifier has little to justify its use. In fact, a marginally stable professional amplifier is almost an oxymoron – a contradiction in itself. However, in the world of hi-fi, some designs exist which are so esoteric that practicality and justifiable engineering are not high on the list of design priorities. Some of these are more works of art than works of science, and they are designed to be pampered and appreciated rather than to be bolted into a rack and forgotten about. Nevertheless, without doubt, some of these specialised hi-fi amplifiers do perform extremely well under the limited circumstances of their intended use, but some are so minimalistic in their design, (even if not in their price) that they *can* sometimes be only marginally stable (although very few), and often their unbalanced input circuits and high sensitivity (100 mV as opposed to 1 volt) make them more prone to disturbance than the less sensitive and often more robust professional designs. It is therefore not surprising that such amplifiers may show a higher degree of sensitivity to both the input *and* output cables with which they are connected than do the more robustly designed amplifiers for professional use, which must work even in relatively hostile environments. However, the term ‘professional’ does not ensure sonic transparency, and some supposedly professional amplifiers which are more robust than transparent would not be very sensitive to cable differences due to their own limited performance.

In all fairness, it must be stated that domestic high fidelity and professional recording are two different worlds. Despite the fact that they have a lot in common, they also have many differences. Whilst professionals tend to work with standardised, known, and objectively designed equipment, domestic equipment tends to be individualistic, and marked by diversity more than commonality. Often, in the home of a hi-fi enthusiast, the equipment has pride of place, where aesthetic design can be almost as important as sonic design, and where minimalism and purity at domestic listening levels take precedence over the tolerance of hard-driving and abuse which may be needed by professional equipment. An idiosyncratic, 8 watts per channel valve amplifier has rarely found a home in a professional recording studio, and especially not if it cost 5000 euros or more. It is therefore worth re-emphasising that the innumerable stories about either input or output cables magically changing the sound of domestic equipment (whilst a giant and highly respected organisation such as the BBC has ‘no policy’ on esoteric cables) are more a testament to the sensitivity of much domestic equipment to minor changes in termination than to the *general* importance of esoteric cable design or materials of construction. However, that is not to say that cables are cables, and that any cable will suffice for connecting a loudspeaker as long as it manages not to catch fire at full volume. So, perhaps we can now look at some of the more important aspects of professional loudspeaker cable design. [Although we should perhaps note, here, that professional recording engineers *do* tend to be rather more conscious of *microphone* cable design].

6.8.3 Cable designs for loudspeaker use

The first way to combat the resistance problem is to shorten the cable; halving the length of the cable will halve *all* the impedance components. Another way to halve the resistance would be to double the cross-section of the cable, but whilst this may be effective on the resistive part of the impedance, the increased spacing between the centres of the conductors will *increase* the inductance. The effect may therefore be beneficial at low frequencies but detrimental at high frequencies. A way to overcome this problem is to use a co-axial cable, where the two conductors share the same axis. This can minimise the inductance by effectively cancelling the opposing magnetic fields in the two conductors, but as a result of this construction there is more opposing surface area *between* the conductors, so the *capacitance* can increase moderately, although this will usually not be a problem. What *is* important is to always keep the pairs of loudspeaker wires as close and parallel as possible. This enables the magnetic fields around each core to cancel as much as possible of the inductance. Twisting the wires is another way to achieve this, but twisted wires are inevitably slightly longer than straight wires for any given overall length of cable. Well designed twisted cables have proved to be successful in high quality applications. What should be avoided is the use of single wires, each following its own route to the loudspeaker. Such configurations can act as effective aerials, and can introduce RF interference into amplifier circuitry. Figure 6.5 sums up the general philosophy.

There is also a phenomenon known as skin effect. It is a controversial subject as to what effect it has at audio frequencies, but, as was shown in Figure 6.3, if high frequency roll-offs give rise to significant phase shifts below 20 kHz, then their effects *may* be audible. Skin effect is the tendency for high frequencies to travel through the outer skin of a conductor, and not through the centre of the core. The whole cross-section of the conductor is therefore not used, so the resistance rises as the conducting section of the cable reduces, introducing a high-frequency roll-off. Once again, the shorter the cable, the less the problem.

Some manufactures have opted to address the problem by plating the outside of the conductors with a lower resistance metal. Another approach is to use Litz-wire, where multiple, individually insulated, hair-like wires are twisted together. They thus have a much greater ratio of surface area to volume. The evidence seems to suggest that on lengths of 10 metres or more, this type of cable can exhibit improved results when compared with ‘ordinary’ loudspeaker cables, but in professional situations, placing the amplifiers 10 metres from the loudspeakers would not normally be considered to be good engineering practice, for other reasons which will hopefully become apparent from the latter sections of this chapter. But first let us look at some detailed measurements which were made on loudspeaker cables

6.9 The amplifier/loudspeaker interface

It must be clearly understood that when a loudspeaker is used with an amplifier employing negative feedback from the output stage, either globally or locally, (and at least 99.9% of amplifiers in professional use are so designed) the loudspeaker cable passes signal in both directions. The amplifier sends drive voltages to the loudspeaker, which cause currents to flow through the complex impedances which the loudspeakers present as a load.

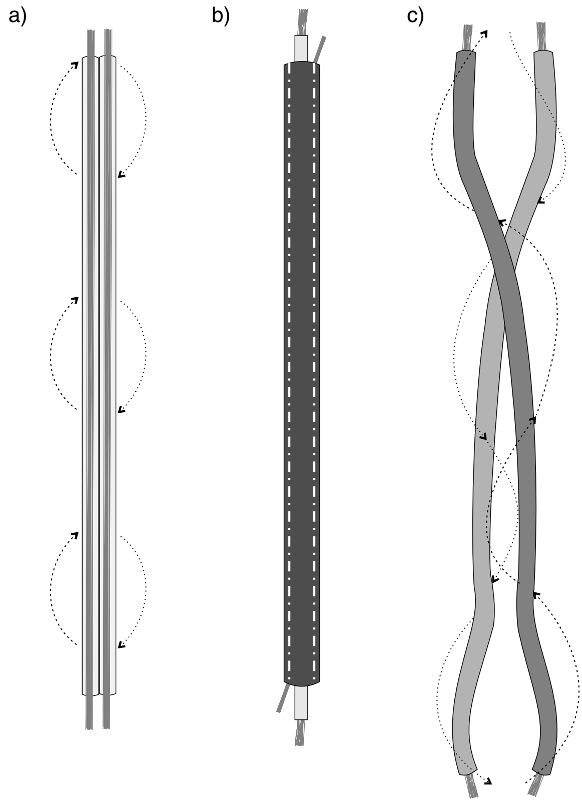


Figure 6.5 Magnetic fields surrounding cables. Better cancellation lowers the inductance. a) parallel conductors. Partial cancellation as parallel conductors exhibit only weak external magnetic fields – the resulting inductance is low. b) Coaxial pair. Almost total cancellation of the magnetic field due to the concentric conductors – the resulting inductance is very low. c) Unrelated pair. Due to the relatively wide and random spacing of the conductors there is little cancellation of their magnetic fields, so the inductance tends to be higher than for the cables shown in (a) and (b)

The reactive components of the impedance, and in particular the moving mass component of the diaphragm/coil assembly, give rise to back-EMFs as the whole assembly resonates in the magnetic field. These EMFs (electro motive forces, or voltages) produced by the resonating loudspeaker acting as an electrical generator rather than as a motor, arrive at the amplifier output terminals via the loudspeaker cables. The circuit of the system is shown in Figure 6.6. The low output impedance of the amplifier cannot effectively damp the back EMFs (reverse voltages) generated by the natural movements of the loudspeaker if an excessive impedance, in the form of a cable, is separating the coil from the output terminals of the amplifier. Cable impedance (or lack of it) is therefore critical in terms of optimising the performance of the amplifier/loudspeaker combinations. It can thus be seen how the cable can control what passes *from* the amplifier *to* the loudspeaker, by virtue of the frequency dependent nature of its impedance,

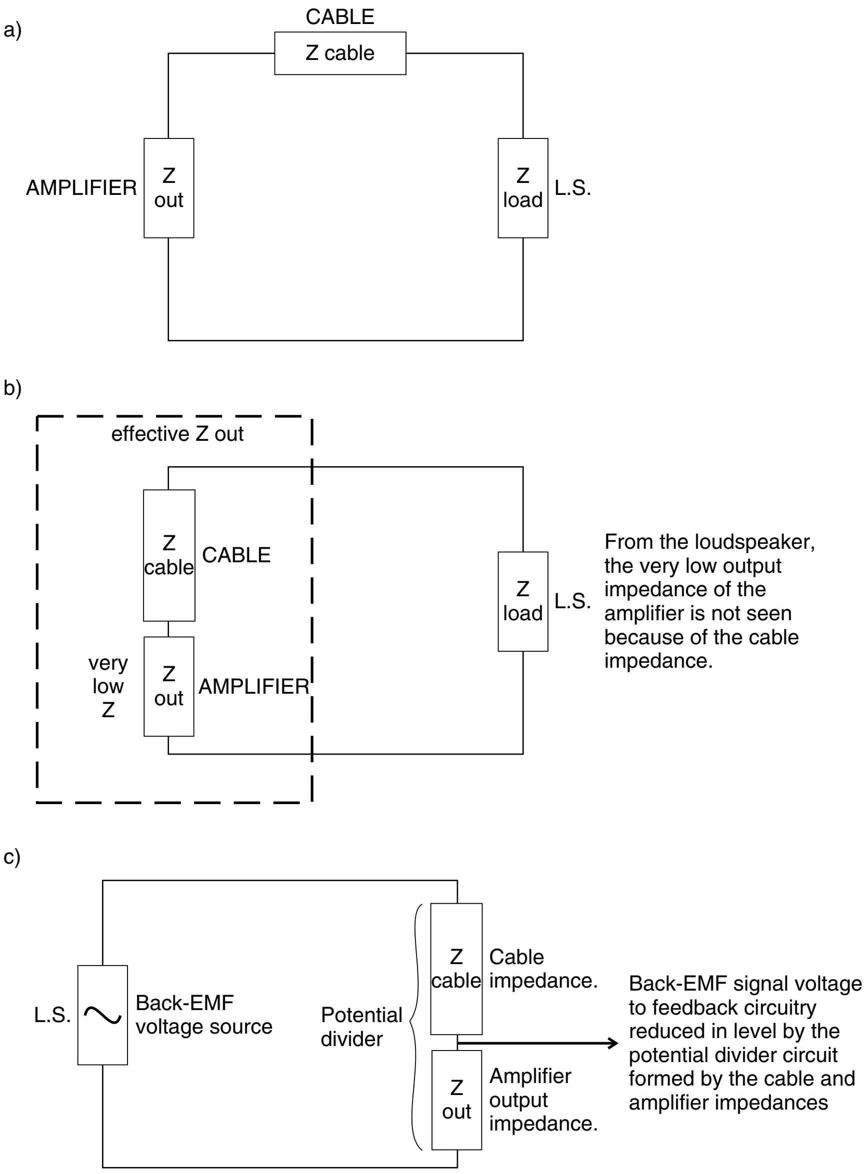


Figure 6.6 Circuit diagram of the amplifier, cable and the loudspeaker impedances. a) Basic circuit. b) Effect on damping. c) Effect on back-emf

and it can also control what passes *to* the amplifier *from* the loudspeaker, and hence affect the damping of the transducer system. The effect of *any* loudspeaker cable therefore must be considered in both directions.

One great problem about generalising about many, or most, of the effects of the performance of loudspeaker cables (once the basic properties of resistance and inductance have been adequately specified) is that their effects can be so system-specific. In other words, what occurs with one combination of amplifier, loudspeaker and location may have very little in common with what occurs with a different combination. The only universal solution for minimising the effect of loudspeaker cables is to minimise their length, by mounting the power amplifiers as close as practically possible to the loudspeakers. A total length of about 2 metres from the amplifier terminals to the motor/driver terminals is a reasonable maximum to aim for. And of course, suitable cable must be used. In the experience of the authors, the differences between cables of appropriate resistance and inductance at lengths below 2 metres are too small to be of any real significance, but exactly what section should be used for what power rating of loudspeaker is something that needs to be worked out case by case. For example, as previously mentioned, an excessively large format cable may be detrimental to the response of a tweeter because the increased cable inductance, due to the cable spacing, may be more of a problem than the increased resistance of a thinner cable.

There exist some large, high powered, passively crossed over, professional monitor systems that are very difficult to drive. When the drive voltages going forward meet the back-EMFs coming in the other direction, all in the highly reactive circuitry of a low impedance, passive, high order crossover, peak currents of up to 100 amps have been measured during some complex musical passages at high studio monitoring levels. It would seem obvious that a cable specified for such a system, using amplifiers capable of driving continuously 3000 watts into half an ohm, would need a higher specification than the cables used in a system of similar power rating and acoustic output but using an active crossover, multiple amplifiers, and where the low frequency drivers presented an almost uniform impedance to the amplifier, (at least in the frequency range over which they were being driven).

When calculating the cross-section of loudspeaker cables, we cannot simply take the approach:

1000 watts into 4 ohms

$$W = I^2R$$

$$I^2 = \frac{W}{R}$$

$$I^2 = \frac{1000}{4}$$

$$I^2 = 250$$

$$I = -\sqrt{250}$$

$$I = \mathit{16\,amps}$$

As 1 mm^2 of cable will safely carry 5 amps, we will therefore use 4 mm^2 cable to give a little margin of security.

In terms of electrical engineering, the above concept is a perfectly safe and viable approach, there is no possibility of the cable overheating, but it in no way takes into account the effect on the sonic perception of musical signals when amplifiers are driving difficult loads. In the case of loudspeaker cables, the emphasis is on the cable impedance rather than thermal/current capacity.

6.10 Some provable characteristics of cable performance

In the autumn of 2002, the authors, together with other collaborators, made a presentation to the Reproduced Sound conference of the Institute of Acoustics[2]. Figures 6.7, 6.8 and 6.9 were taken from that paper. They were from an experiment made at Czerwinski Laboratories in California, by two Russian engineers, Alexander Voishvillo and Alexander Terekhov, accompanied by Eugene Czerwinski, (the founder of Cerwin Vega). [In the 1950s Czerwinski had designed the first of its kind 10,000 watt amplifier using germanium transistors for sonar systems for the US Navy.] They carried out tests on three, 6 metre lengths of different cables, firstly into a resistive 8 ohm load (Figure 6.7), then into a full-range loudspeaker system (Figure 6.8), and finally into a cabinet-mounted low frequency driver (Figure 6.9). The amplifier driving the cables was fed with a multitone test signal[3,4], designed to show up non-linear distortions; in particular intermodulation distortion. In Figure 6.7 it can be seen that all six plots are more or less the same. The left hand plots were measured at the amplifier output terminals, whilst the right hand plots were measured at the other end of the 6 metre cables – at the resistive load. The three very different cables all appeared to perform equally. This was what the experimenters were expecting to occur, independently of the load. They were all very much sceptics with regard to significant loudspeaker cable differences – at least between adequately rated cables. However, when they replaced the 8 ohm resistive load with a passively crossed-over loudspeaker system, they were surprised to see the results as shown in Figure 6.8. With the more complex load, *all* the plots had changed. After changing the load to the low frequency loudspeaker, all the plots changed again, as shown in Figure 6.9. The distortion patterns were noticeably different, not only between the cables, but also between the input and output ends of each cable. The implication here is that the cables change the way that the complex load is seen by the amplifier. Voishvillo reported that upon seeing this, Czerwinski exclaimed “But they’re only short cables!”

Figure 6.10, taken from the same IOA paper[2], shows six twin plots. In each case the upper trace was the measurement at the amplifier end of the cable, and the lower trace was made at the loudspeaker end. The measurements were taken in a city-centre office with typical city EMI (electromagnetic interference). The input signal was a 1.6 kHz square wave, and the load was a TAD 2001 compression driver, mounted on an axisymmetric horn and producing 70 dB SPL at a distance of one metre. This was intended to represent conditions of realistic use. Plots a) and b) are of 5 metres and 50 metres of RG59 coaxial RF cable, respectively. Plots c) and d) are of a typical, transparent insulation, oxygen-free copper loudspeaker cable, again 5 metres and 50 metres respectively; and plots e) and f) are as a) and b), but with the conductors reversed (i.e. screen to hot and core to

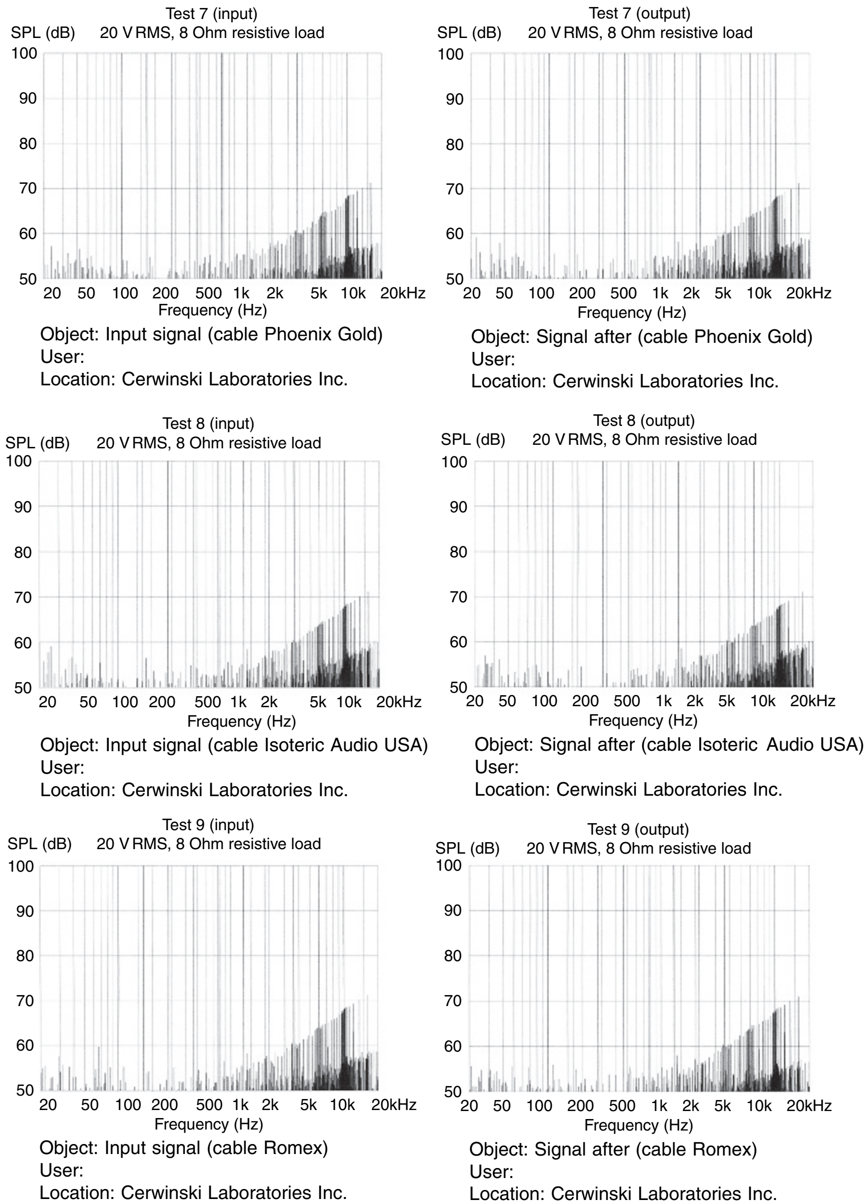


Figure 6.7 Three 6 m cables feeding on 8 ohm resistive load. The left-hand plots are from amplifier output terminals. The right-hand plots were taken from the load

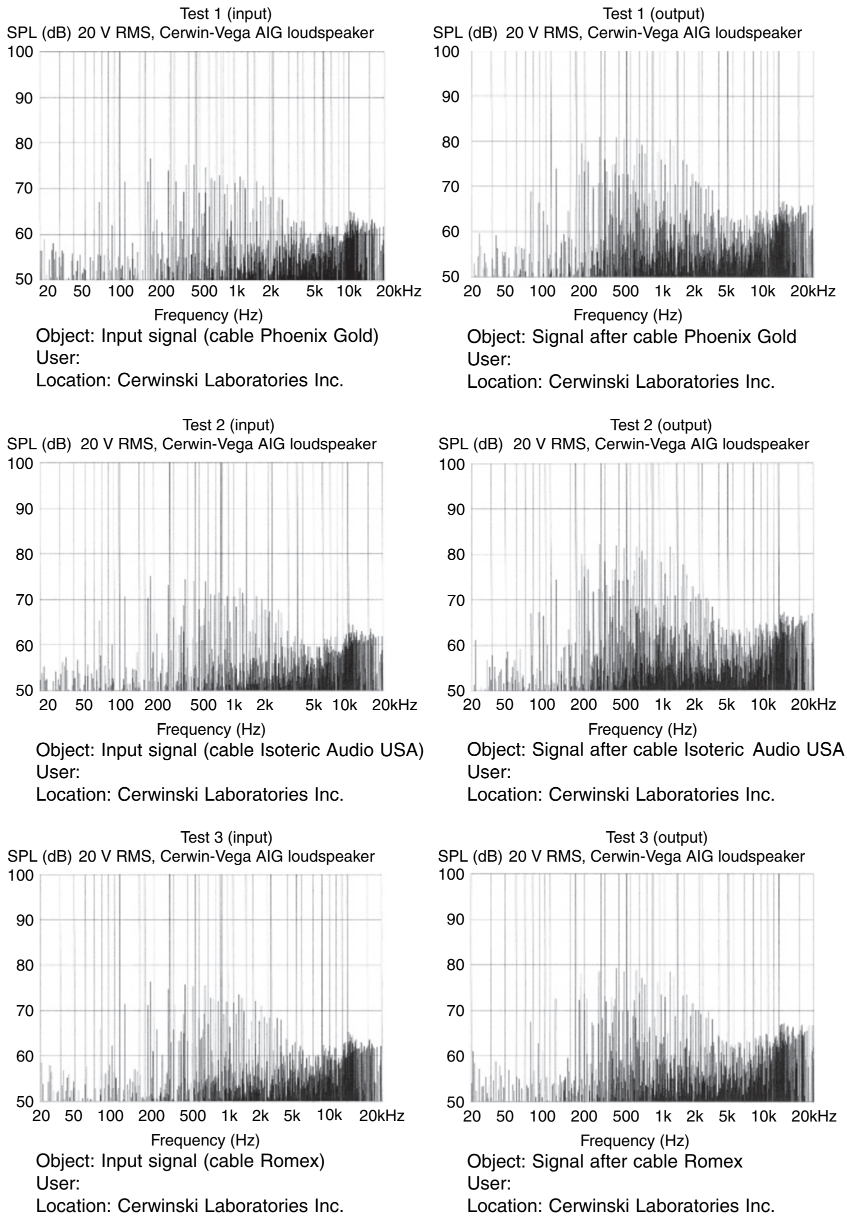


Figure 6.8 Three 6 m cables feeding a full-range loudspeaker with a passive crossover. The left-hand plots are from the amplifier output terminals. The right-hand plots were measurements from the loudspeaker cabinet input terminals

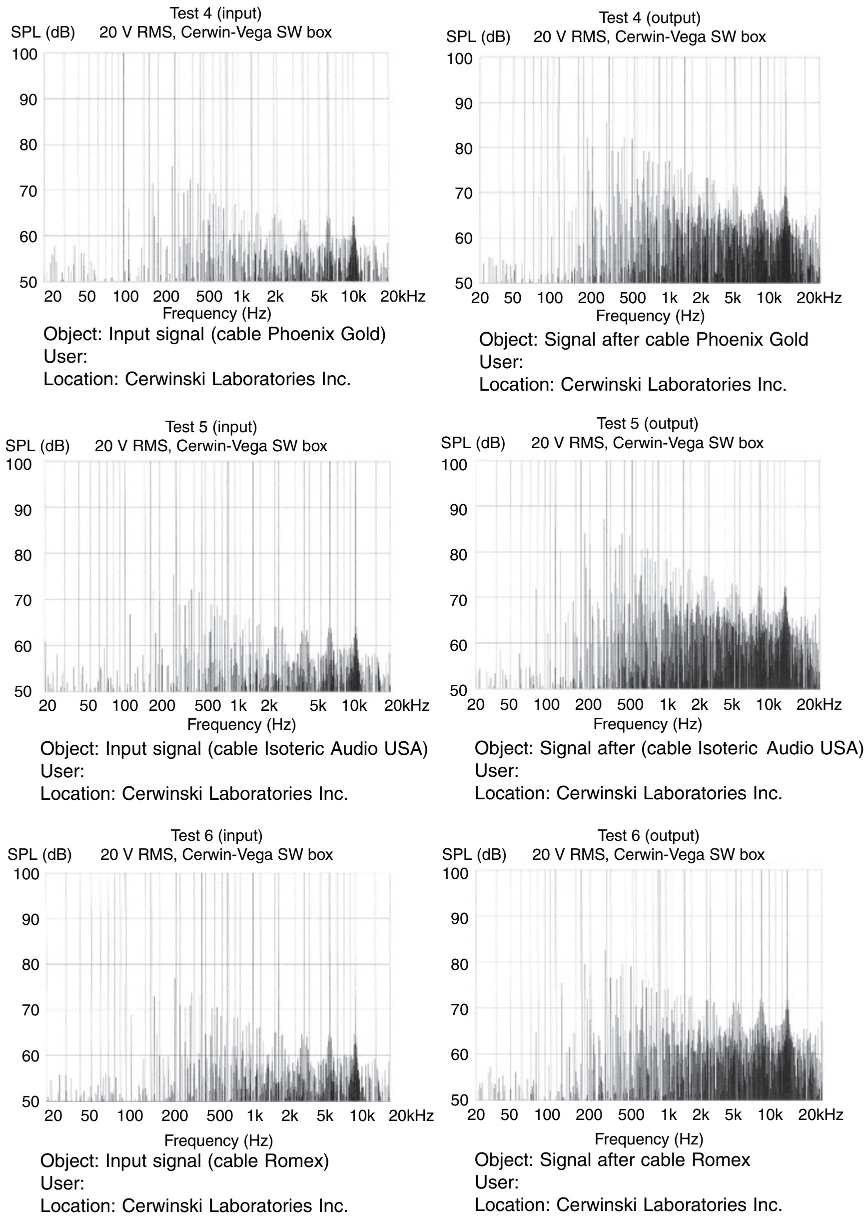


Figure 6.9 Three 6 m cables feeding a sub-woofer box. The left-hand plots are from the amplifier output terminals. The right-hand plots were measurements from the loudspeaker cabinet input terminals

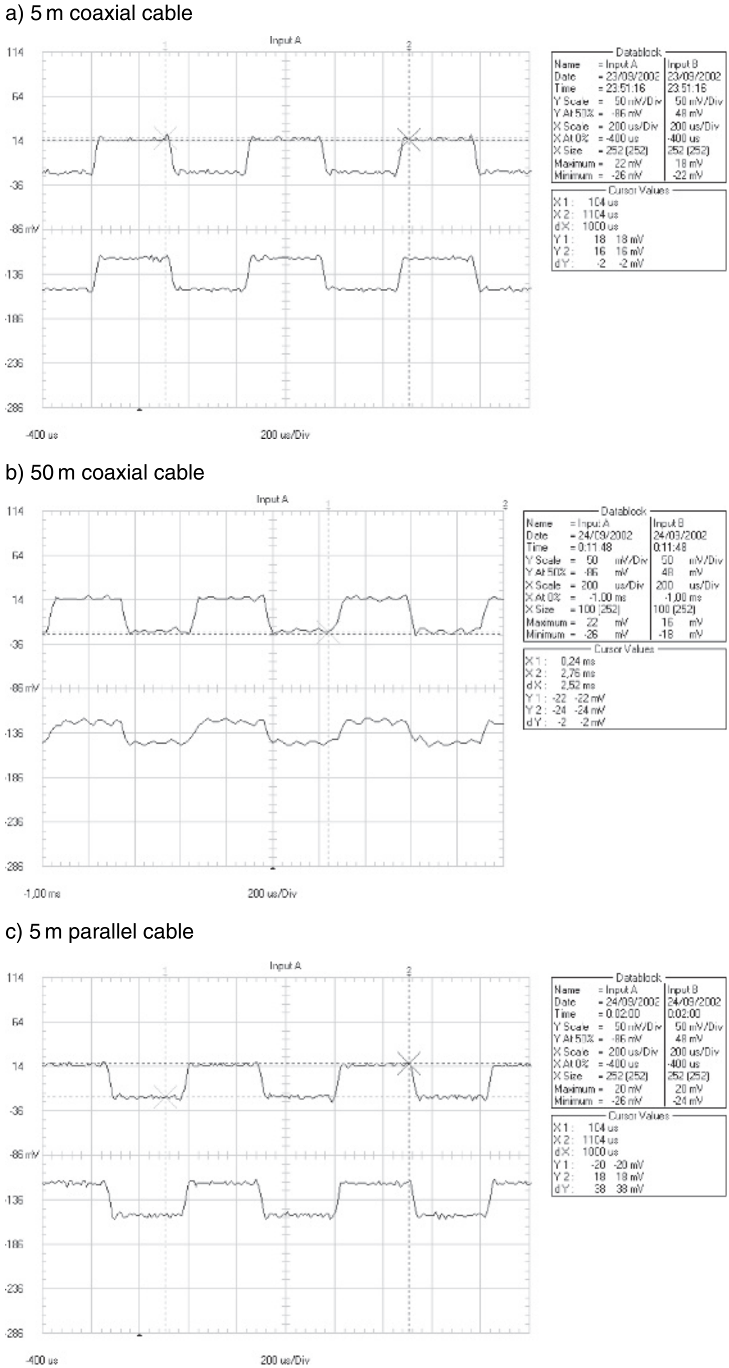
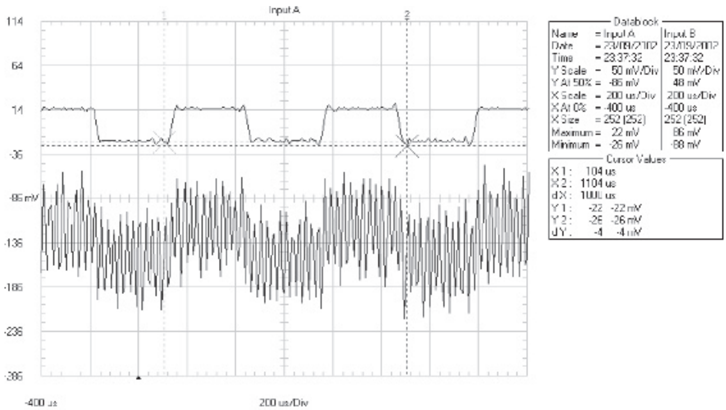
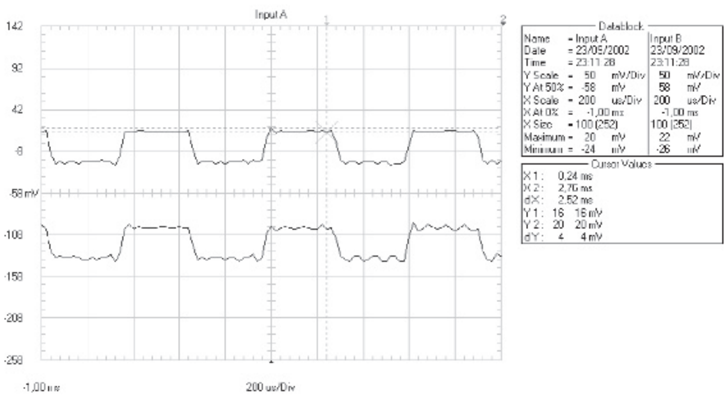


Figure 6.10 Effects of cable on 1.6 kHz square wave. All upper traces are from the amplifier output terminals; all lower traces are from the loudspeaker input terminals

d) 50 m parallel cable



e) 5 m coaxial cable with screen as live



f) 50 m coaxial cable with screen as live

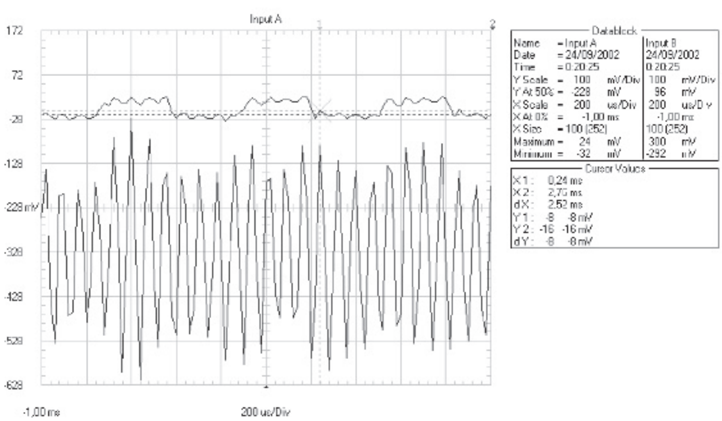


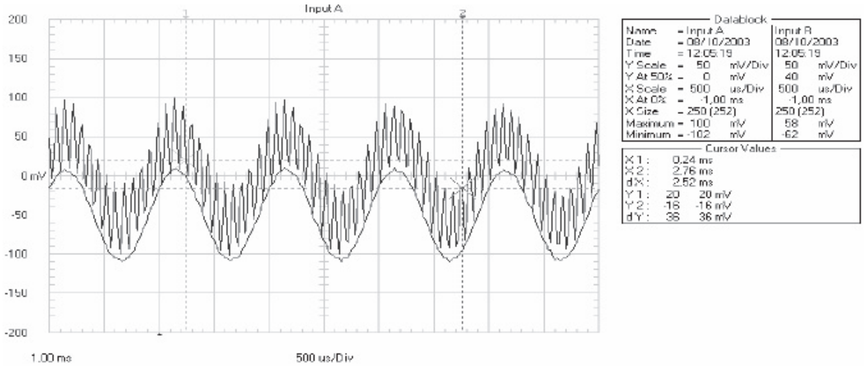
Figure 6.10 Continued

ground). Whilst listening only to the background noise from the amplifier, at very low level, no change could be heard between any of the cables, but the potential for such changes in interference patterns, especially on the longer lengths of cable, does not bode well for the inaudibility of these effects on complex musical signals. It was suspected that some amplifier designs could be more susceptible than others to this type of interference entering their feedback circuits via the output terminals. The gross differences in the interference patterns between b) and f), using the selfsame cable but reversing its polarity [hot down the screen in f)] is suggestive of something untoward occurring.

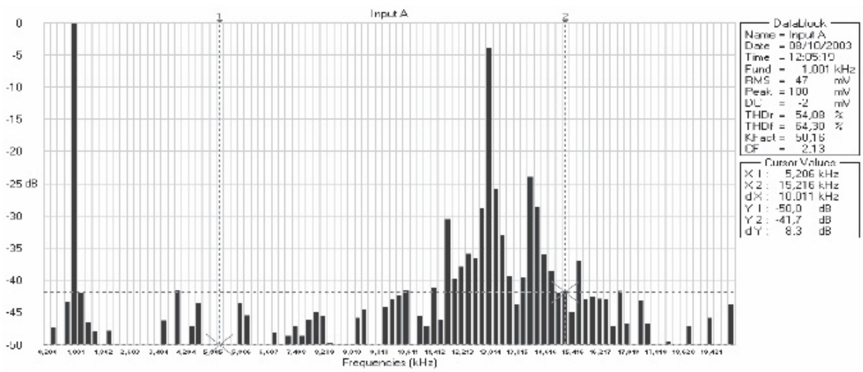
Twelve months later, a further paper was presented, this time to the Reproduced Sound 19 conference, in Oxford[5], largely intended to answer questions raised during the previous year's presentation. A series of five tests were undertaken, interchanging the amplifier, cable, and loudspeaker within the same basic tests set up, to see if any changes were noticed in the interference levels. As had been seen the year before, a considerable amount of hash was being received from the air, including a 13 kHz signal which seemed to disappear around 9 pm each evening, even though nothing in the office or workshop was switched off at that time. In the first set up, a high sensitivity compression driver was used as a load, mounted on a horn and producing a 1 kHz tone at 80 dB SPL at one metre. It was driven by a Class A amplifier, designed for professional use and connected via 28 metres of 4mm^2 OFC (oxygen-free copper), parallel conductor loudspeaker cable. The results are shown in Figure 6.11. The test was then repeated, but with a dome tweeter of similar frequency response but 20 dB less sensitivity. The amplifier was duly increased in gain to once more produce 80 dB SPL at 1 metre, and the resultant plots are shown in Figure 6.12. Inspection of Figures 6.11 and 6.12 clearly shows how, whilst the absolute interference level has remained almost constant, the extra 20 dB of signal has swamped the noise in the second test. The plots were normalised to the level of the 1 kHz signal level, towards the left hand side of each graph. The frequency scales are linear, *not* logarithmic. The implication from this is that if the interference pick up could affect the performance of the amplifier, even if it was not directly audible from the loudspeakers, then the effect would be more noticeable on higher sensitivity drivers than on lower sensitivity drivers.

This was an interesting finding, even though it was not definitive, because the motivation to do the investigations leading to the 2002 paper[4] came after the usual OFC loudspeaker cables were changed to RG59 coaxial cables between the amplifier and compression drivers in a large studio monitoring system. Many people noticed the difference, even though no difference could be measured in the acoustic output. People (professionals) reported the sound as being smoother with the coaxial cable. The studio in question was sited close to emergency service aerials (fire and police) yet nobody had commented about similar problems in similar monitor systems in other locations. The original complaints of an unusual slight harshness in the sound only applied to the high sensitivity loudspeakers, and not to any other loudspeakers in the studio. The change to the RG59 solved the problem.

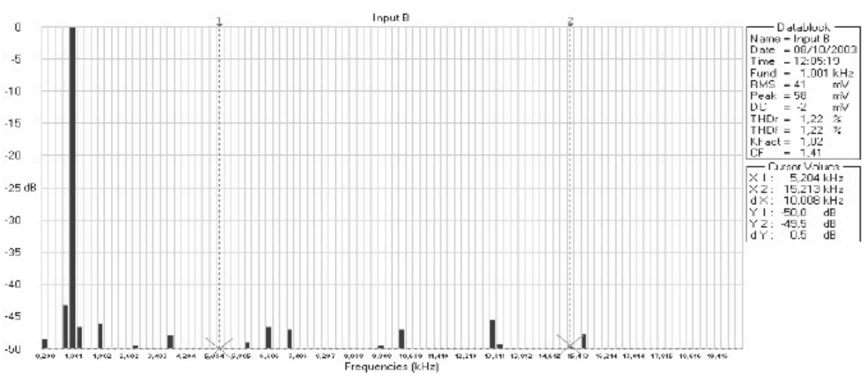
The next test in the series was a repeat of the test whose results were shown in Figure 6.11, except that the parallel conductor, 4 mm^2 OFC cable



a) Input A voltage at loudspeaker terminals
Input B voltage at amplifier output terminals

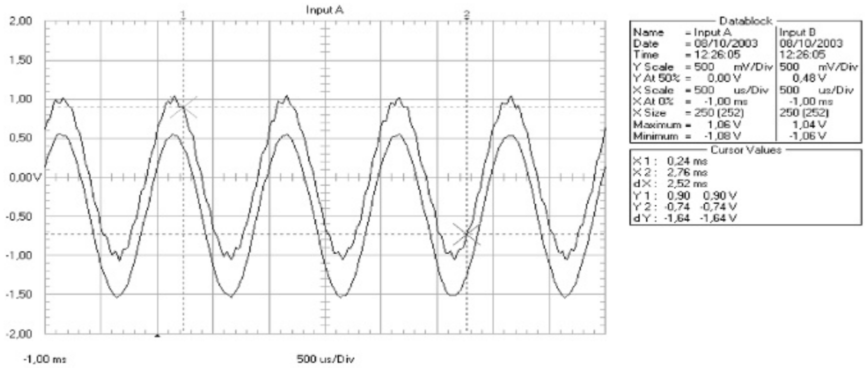


b) Noise spectrum at loudspeaker terminals

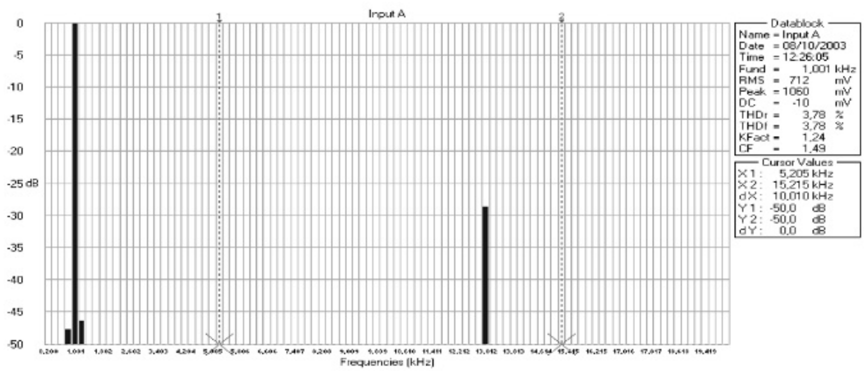


c) Noise spectrum at amplifier output terminals

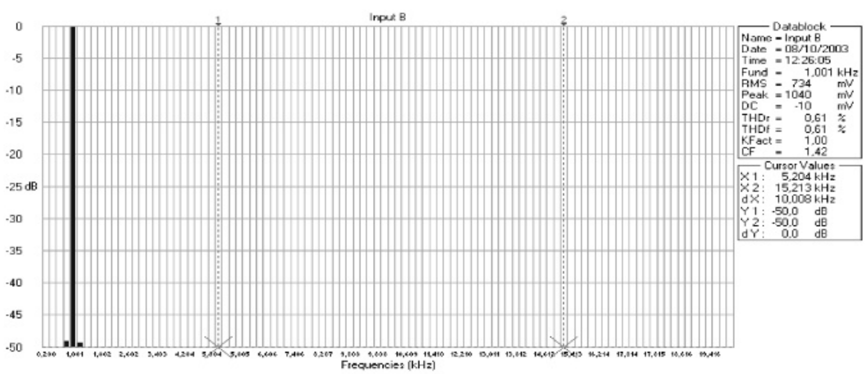
Figure 6.11 Class A amplifier/parallel cable/compression driver



a) Input A voltage at loudspeaker terminals
Input B voltage at the amplifier output terminals



b) Noise spectrum at loudspeaker terminals



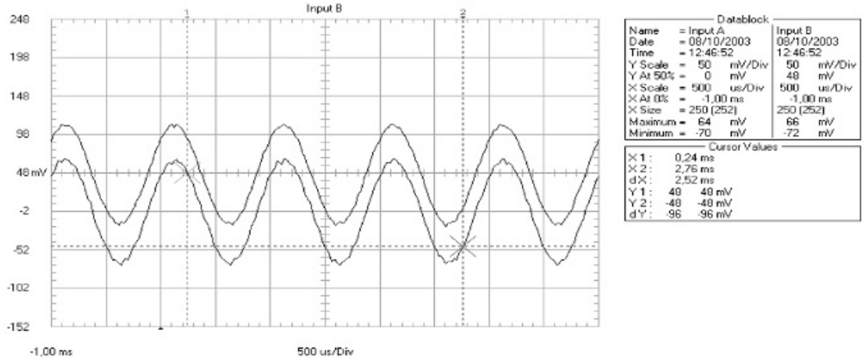
c) Noise spectrum at amplifier output terminals

Figure 6.12 Class A amplifier/parallel cable/dome tweeter was replaced by a 2.5 mm^2 coaxial *loudspeaker* cable (not a co-axial *RF* cable, as used in Figure 6.10). The results are shown in Figure 6.13, which, when compared with Figure 6.11 show much less interference at the loudspeaker end (which may, or may not, be consequential) but a *greater* level of the 13 kHz interference at the amplifier end of the cable. The voltage sensitivity of the measurement scales is approximately equal in each case.

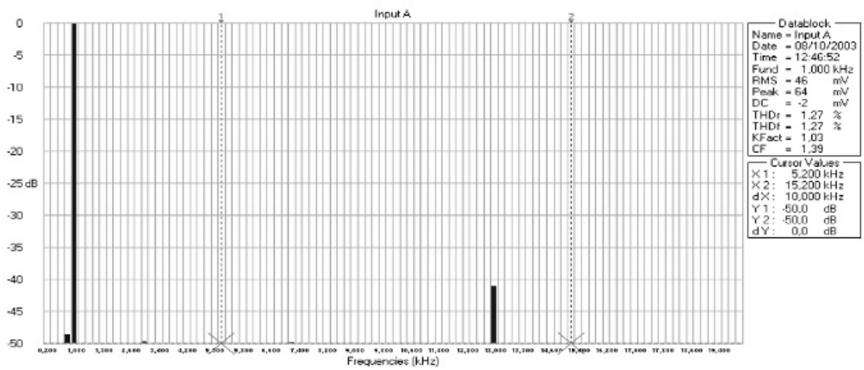
The fourth measurement in the series used the same loudspeaker and cable as in the previous measurement, but the Class A amplifier was substituted for a Class AB amplifier. The amplifiers differed considerably in their quantity and application of negative feedback. The results of this test, shown in Figure 6.14, differ greatly from those shown in Figure 6.13. Repeating the last test with parallel cable in place of coaxial cable gave the results as shown in Figure 6.15. In the paper it was stated "Comparison [of the plots in the figures] shows that the interference signal measured at the loudspeaker driver terminals is affected differently by each amplifier, and that the difference in the interference patterns from cable to cable is also influenced by the amplifiers to which they are connected. Given the obvious interdependence of all the parts of the circuit, the findings help to explain why such a lack of consistency exists between reports of the beneficial effects of using certain loudspeaker cables. It would appear to be the case that certain cable benefits can only be claimed for certain amplifier/loudspeaker combinations, and that any perceived audible improvement heard on any one combination may not necessarily be able to be expected when the cable is used on any other combination".

Perhaps it is also worth looking at some other findings from the same paper, which attempted to measure the losses actually occurring *within* different cables of different construction and with different loads. Figure 6.16 shows the results of some measurements made by connecting the two inputs of a differential amplifier across the amplifier and loudspeaker ends of some two-metre sections of cable[5]. The rise in the plots above 2 kHz is a result of the less effective rejection at higher frequencies, so the graphs are comparative rather than absolute: the decibel (vertical) scale is not calibrated. Nevertheless, the differences are real enough. Figure 6.16(a) shows the plot resulting from the measurement of a 2.5 mm^2, screened, twin twisted-conductor loudspeaker cable. Figure 6.16(b) shows the results for a cat-5 data cable, and Figure 6.16(c) shows the plot for a 6 mm^2 twin, parallel conductor loudspeaker cable. All were driving an 8 ohm woofer, but in the latter case, (c), the cone was blocked to prevent any movement. With the cone blocked there is no mechanical resonance, so the dip at around 80 Hz in the first two cases is absent in the case of Figure 6.16(c). The plots clearly show how the losses within different cables are quite different, and how the changes in impedance at the loudspeaker end can further affect the losses in the cables. All the measurements of Figure 6.16 were on lengths of only *two* metres of cable. Despite the fact that the scaling of the amplitude is only relative, it can still be seen that differences in the losses do exist, not only in terms of overall level but also in frequency balance.

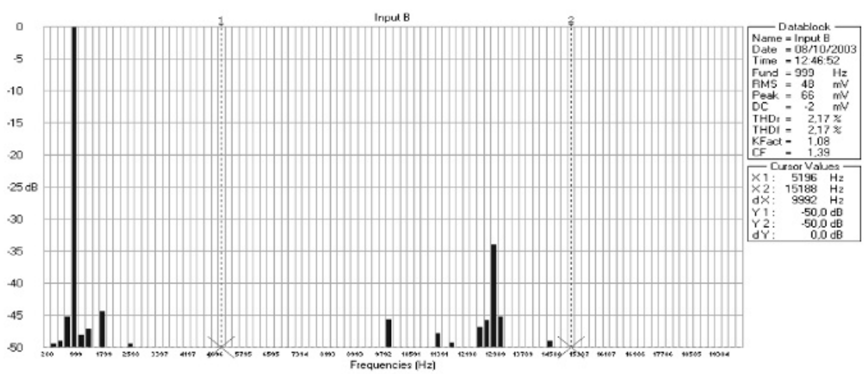
As the evidence presented in this chapter has shown, loudspeaker cables seem to be sensitive to the equipment to which they are connected, and vice versa. What is more, the entire systems seem to be sensitive to their environment, at least in electromagnetic terms. Nevertheless, the concept



a) Input A voltage at loudspeaker terminals
Input B voltage at the amplifier output terminals

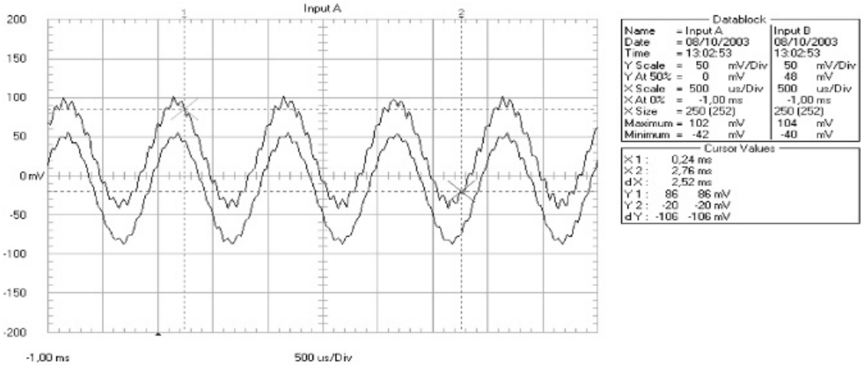


b) Noise spectrum at loudspeaker terminals

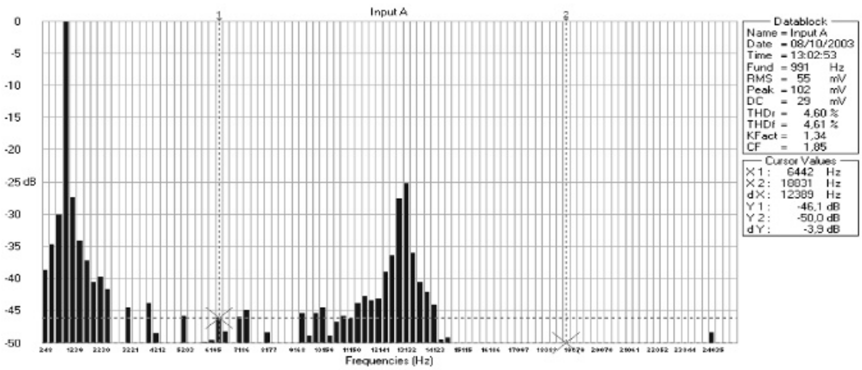


c) Noise spectrum at amplifier output terminals

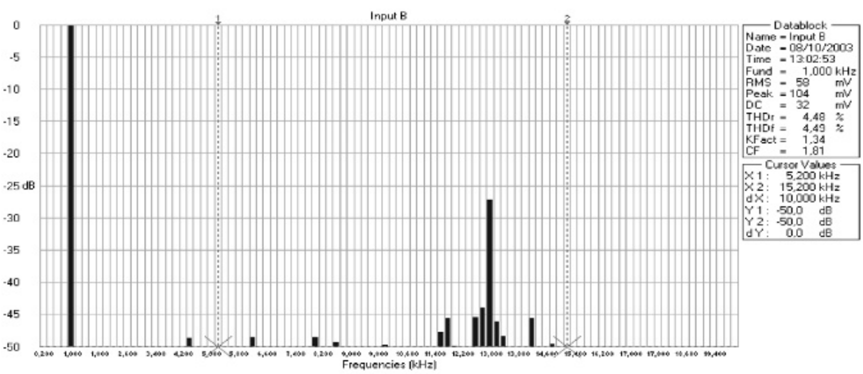
Figure 6.13 Class A amplifier/coaxial cable/compression driver



a) Input A voltage at loudspeaker terminals
Input B voltage at the amplifier output terminals



b) Noise spectrum at loudspeaker terminals



c) Noise spectrum at amplifier output terminals

Figure 6.14 Class AB amplifier/coaxial cable/compression driver

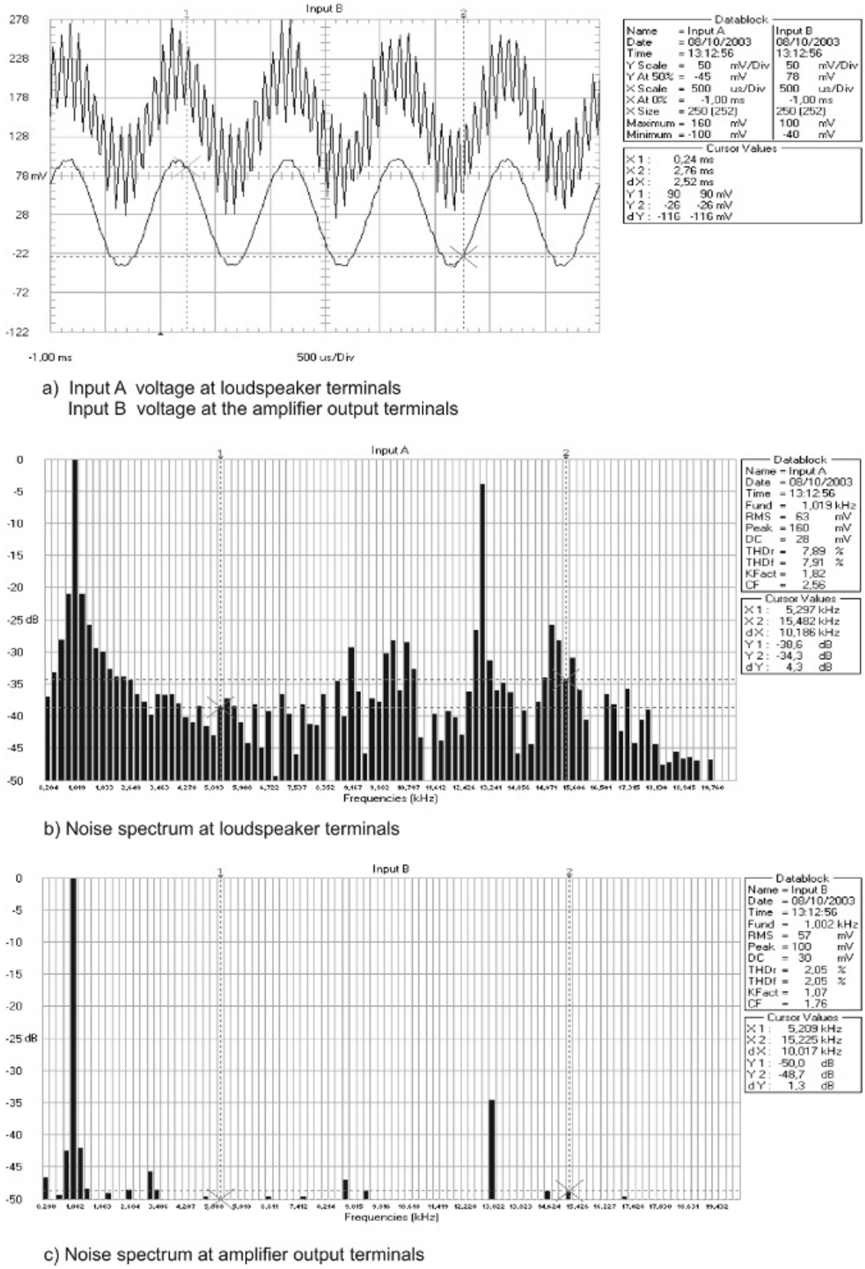


Figure 6.15 Class AB amplifier/parallel cable/compression driver

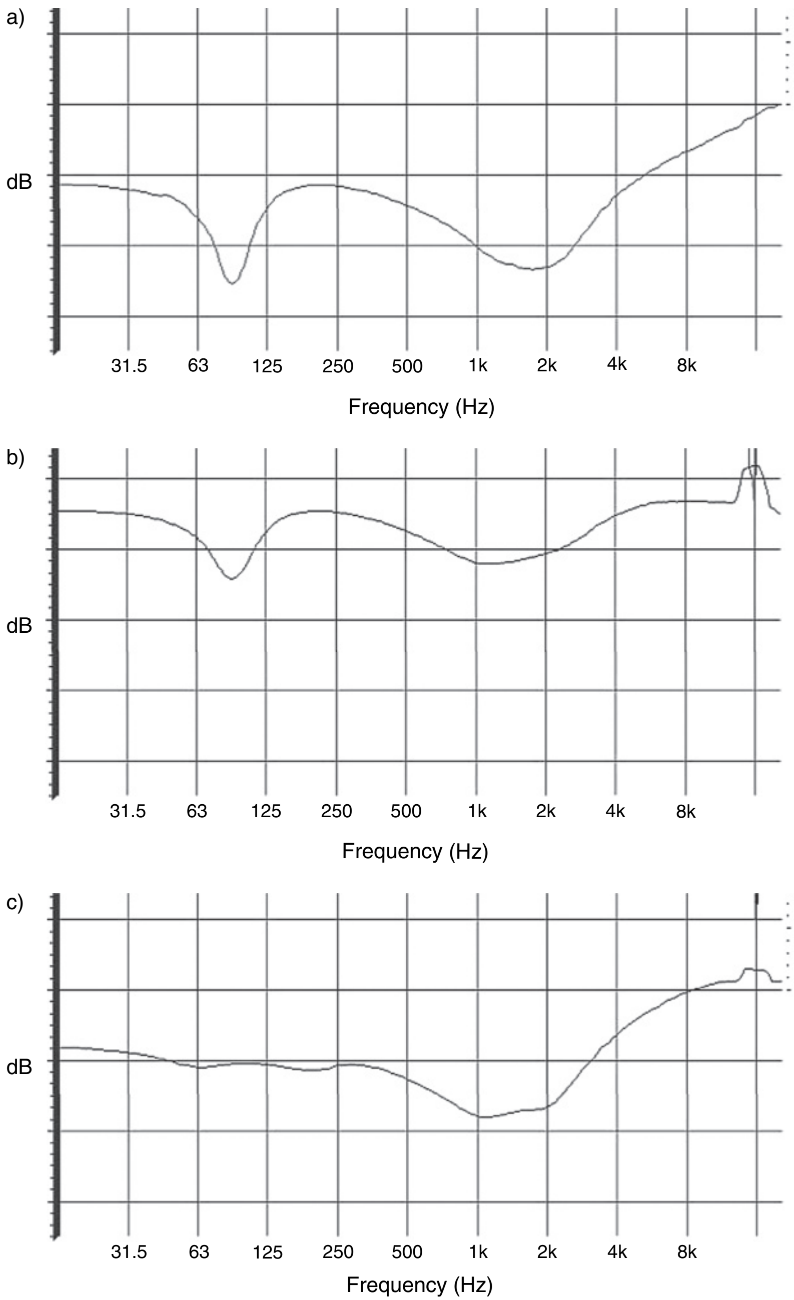


Figure 6.16 a) Spectrum of the difference between the voltage signals at either end of a 2.5 mm^2 screened cable. b) Spectrum of the difference between the voltage at either end of a CAT5 data cable. c) Spectrum of the difference between the voltage at either end of a figure-of-eight cable with the cone of the low frequency loudspeaker blocked (i.e. prevented from moving). Vertical scale not calibrated

of minimising loudspeaker cable lengths seems to be well founded, but the practice of separating the frequency range into narrower bandwidths also seems to reduce the demands made of the amplifiers and cables, alike. This is a concept which it is now worth looking at in a little more detail, before expanding the concept to polyamplification and multiamplificaiton.

6.11 Some passing comments

Unfortunately for people who have not had a very great degree of experience of the subject, sorting out what is solid from what is nebulous in terms of so much that has been published in the hi-fi press about loudspeaker cables is not an easy task. Nevertheless, despite some critics claiming that it is *all* a load of nonsense, it is the opinion of the authors that far too many people of good repute have claimed that sonic differences do exist to be able to dismiss the subject of loudspeaker cables out of hand. One of the main obstacles to verification has been the cost of double-blind, controlled tests, with sufficiently large groups of subjects, coupled with the extraordinarily large number of possible equipment combinations and test locations, *and* the fact that a cable which shows benefits in one of those combinations may show no benefit in another. One component in isolation may not exhibit any noticeable difference over another component. However, when they are used in some combinations with other components, the differences may be readily apparent.

Whilst writing this book, a short, written note was received from Martin Colloms, an eminent authority on the subject of high fidelity loudspeakers. There is nothing to suggest, from either his professional conference papers *or* his well-respected text books, that he is either cloth-eared or a charlatan, and neither can the authors of this book refute his opinions. An extract from his communication will serve as a good summary.

> A fair percentage of my reviewing life has been spent listening to and dealing with cables. Most evaluations I still do blind. I think that I have correlated with reasonable precision the key aspects of cable design from some 400 examples.
>
> Metallurgy: the actual metal or alloy and its state/annealed/crystalline composition/extrusion direction and subsequent build orientation.
>
> Dielectrics alter the sound, as do different plastics in film capacitors, the issues including dielectric factors, DF with frequency, piezo effects, self-damping including semiconducting insulation, and dielectric constant.
>
> Geometry: spacing, stranding, Litz or bunched.
>
> Mechanico-acoustic: physical strength, rigidity, self-damping, microphony issues.
>
> RFI: can be very important; if you spectrum analyse (I work up to 1.5 GHz) the dominant RFI (radio frequency interference) is about 1 MHz for many loudspeaker cables.
>
> Many amplifiers are no longer anything you would recognise by 100 kHz, never mind the 1 MHz, and the RF gets in the output

terminals and intermodulates around the feedback loop. A speaker cable is often a good medium-wave aerial.

Cables vary greatly in how they dump RF interference into the amplifier output port. The Zobel filter has little effect as it is generally buffered by 10 ohms. If the RFI doesn't get in the positive line, it common modes into the ground line.

At 1 MHz or more, the RFI hardly cares what kind of amplifier it is.

Contacts and connectors also matter, as do their tightness and vibration resistance. Acoustic energy ends up being mechanical watts and everything shakes about.

It is amazing how mild that vibration can be and still affect the sound of an electrical component, including a cable. Simple tests on vibration isolation and suppression show this clearly. Acoustic/mechanical coupling remains one of the most insidious modes of sound quality loss.

In a recent conversation, Paul Frindle, one of the designers of the Sony Oxford R3 digital mixing console and many of the Sony professional plug-ins, recalled an occurrence during his time as a designer at Solid State Logic. There had been complaints from many studios around the world that mixes done on the small faders, which were simple potentiometers, sounded better than mixes done on the large faders, which operated VCAs (voltage-controlled amplifiers). They proceeded to make countless measurements at the factory but could find no apparent cause for the described difference in the sound. Some of Paul's comments were very reminiscent of loudspeaker cable anomalies.

P.F. "You could end up in a situation where a single signal sounded fine, but its relationship with everything else was ill-defined. You could set up a mix on the pots (the small faders), with no VCAs, and it would sound great; but going through the main faders, with the VCAs, it would sound oddly wrong. Yet, if you *soloed* any single VCA channel you could hear no difference compared to the small fader. It was fascinating! I found that this was caused by very small amounts of signal dependent delay variance through the VCAs, which was a complex kind of distortion that was unfamiliar within the design context. In my opinion this was one reason that led to the perception that the VCAs sounded bad, and many people went in the direction of consoles with motor driven faders instead.

This problem has some similarity to issues resulting from analogue to digital converters. For instance, if you have different clock-jitter, between channels – this is a lovely one! If your system has timing errors, when you use an ADC and a DAC back to back in a monitoring capacity, the clock is common and simultaneous to both the encoding and decoding stages. Since the timing errors are synchronous they partially null out. However, if you delay the signal – or store it then play it back later – even though the system is the same, the jitter is no longer synchronous, and the resulting sound quality is compromised. In fact, the sound is reminiscent of the mixes done via the VCAs on the *old* SSL consoles. Of course, this is one of many

> good illustrations of how the sound can change through a system, even though the recorded bits are unaffected and no numerical errors have actually occurred.
>
> Another similar problem can exist when you dither signals, because of relative correlation. When you dither a number of channels, the dither noise of any one channel should be unrelated and uncorrelated to any other, or you risk the relative correlation of the dither noise starting to become evident. You can end up with statistically partially 'mono' dither-noise, which can close in the stereo effect, especially on fade-outs. Again, a single channel works fine, a stereo channel apparently sounds fine whilst the music is playing, but as channels build up with the same dither, the stereo begins to close in. We spent a fortune on the R3 finding 256 independent noise sources, because it was a strange problem to solve. It is hard to measure, and the sound problems are something which you often need to be working on day after day before they become clearly apparent. You feel a sort of unease, that something is wrong, but you can't quite explain what it is. At one time, before we found an easier solution, 30% of the processor of the R3 was dealing with this problem. [And the R3 had about 3000 times the processing power of a current (2005) Pro Tools system. P.N.]

The illusiveness of these types of problems are not dissimilar to the illusiveness of definitive, irrefutable evidence about the differences between loudspeaker cables, but whereas Sony and Solid Stage Logic have been able to throw vast amounts of money at the problems, low volume producers of special loudspeaker cables have not enjoyed such luxury. Two of Paul Frindle's sentences in the previous paragraph are very relevant to the issue. "It is hard to measure, and the sound problems are something which you often need to be working on day after day before they become clearly apparent. You feel a sort of unease, that something is wrong, but you can't quite explain what it is". There are the hardnosed objectivists who would say that if you cannot measure it, then it cannot be important, but, once Solid State Logic solved their VCA problem, the complaints stopped from the users around the world. Surely this is evidence that the problem had been real enough. There is no logic in claiming that if Paul and his colleagues had *not* found the problem, then that would have proved that the problem was imaginary. Similarly, if people cannot prove the reasons why cables should sound different, the implication is that not enough effort has been put into finding the problem, and *not* that the problem does not exist. And, of course, a subtle difference in the sound, given rise to by two loudspeaker cables, will almost certainly *not* be audible if the resolution of the loudspeakers being used to audition them is not, in itself, high enough to show up the difference.

Nevertheless, the evidence is now overwhelming that cables *can* give rise to sonic differences. The situation is somewhat paralleled by the relationship between stress and the functioning of the human immune system. Believe it or not, despite the fact that most people 'know' that when people are stressed they seem to be more prone to illness and infection, *no* clinically proven, scientific evidence currently exists to make a definite link between stress and susceptibility to illness. However, in 2002, Ronald Glaser, et al, at Ohio State University, published in the Journal of Consulting and Clinical Psychology a paper stating that whilst no proof existed, the circumstantial evidence was too overwhelming to be ignored. The conclusion was that stress *did* lead to more illness and infection, and that the fact that they could not *prove* how it could do so did not mean that it could not be said that it did do so.

6.12 Multi-cabling

It is widely accepted that bi-wiring with conventional loudspeaker cable will almost universally impart a greater sonic improvement to any multi-way loudspeaker system than a change from a single conventional cable to a single esoteric cable. Bi-wiring takes separate cables from the output terminals of the amplifier to the separated inputs of the high and low frequency sections of the passive crossover filters. Figure 6.17 shows the principles of bi-wiring. Tri-wiring simply extends the concept to three filter sections and three cables. Although each cable still receives the same *voltage* drive from the amplifier, the *current* passed by each cable is only that which relates to the frequency band that it is handling. As it is the *current* which gives rise to the linear *and* non-linear processes which are

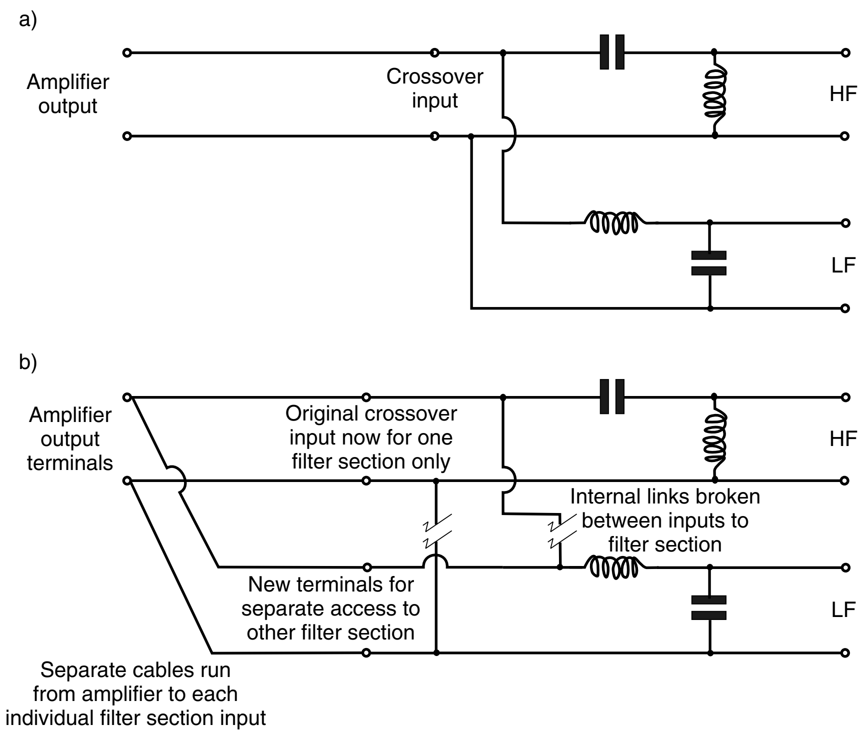


Figure 6.17 a) Conventional amplifier-to-crossover wiring. b) Bi-wiring

attributed to magnetic effects, the separation of the currents into two or more frequency bands can be beneficial. The high frequency signals are therefore unaffected by the heavy low frequency currents.

The materials used for the conductors can influence the sound quality performance of loudspeaker cables, as can the materials used for the insulators. The physical construction of a cable is also considered by many to be a significant factor in terms of sound quality, with some complex plaiting arrangements of the conductors being highly regarded by many specialists. Given these differences, appropriate cable designs and dimensions can be more ideally matched to the frequency and current requirements of the individual drivers in a system if multi-cabling is employed.

6.13 Polyamplification and multiamplification

What has just been discussed is the way in which the separation of the audio spectrum into different frequency bands can make much fewer demands on the loudspeaker cables, and that in separating the bands, intermodulation distortion can be reduced. This idea can be extended one step further by using separate amplifiers, as well as separate cables, a practice known as polyamplification as shown in Figure 6.18. However, once one has progressed to this stage, there can be little justification in using high level, passive filters when low level, active filters could do the job in a more precise and less lossy way, although the design of audiophile-standard active filters is not an easy task. Nonetheless, the polyamplification system *has* seen use, and the amplifiers, like the cables in the previous case, are driven

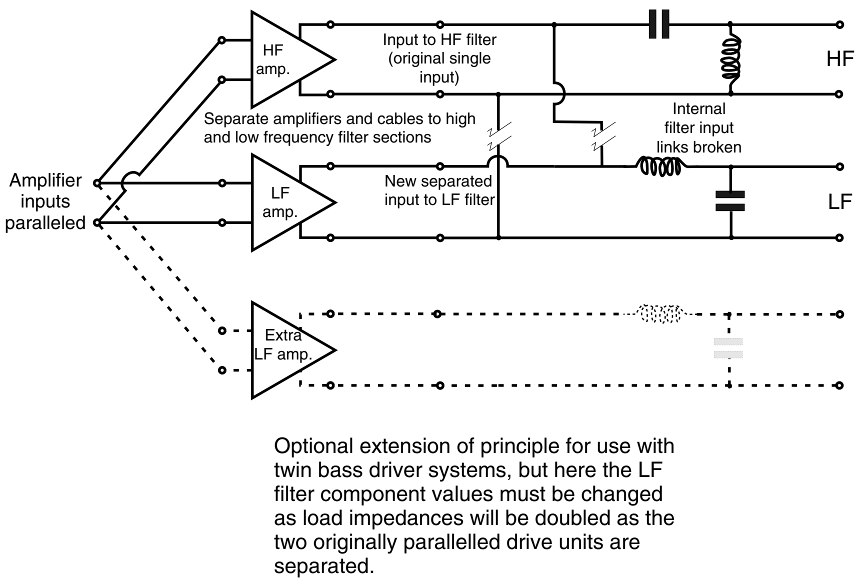


Figure 6.18 Extension of bi-wiring to bi-amplification whilst still using passive, high-level filters

by the full, composite signal voltage, although they only need to supply current over their specific frequency bands, reducing intermodulation distortion to a greater degree than bi-wiring or tri-wiring alone.

When the filters are placed *ahead* of the amplifiers, and the amplifier outputs are connected *directly* to the loudspeaker drive units, with *no* components other than the cable between them, the practice is known as multi-amplification. Multi-wiring (bi-wiring, tri-wiring etc), polyamplification and multiamplification are three distinct steps forward, respectively, in terms of sonic quality, and it usually does not require trained ears to notice the benefits.

There is no doubt that it is asking a lot of any amplifier, or loudspeaker cable, to faithfully pass up to 11 octaves of musical signal with a dynamic range of 90 dB or more. Considering the fact that no loudspeaker *driver* can do this, it seems perfectly reasonable to split the frequency bands ahead of the amplifiers and drive each frequency range independently. The benefits of doing this were amply described in Chapter 5, but further repercussions of this practice on amplifiers and cables can be expanded on, here.

It does not take too much imagination to realise how a 20 or 30 amp low frequency current can modulate high frequency signals passing along the same cable at levels of 40 dB below. The bass currents cause the conductors to flex, mechanically, under the mutual magnetic field. Piezo and structural modulation then carries interference into the low level high frequency system. In terms of *power* ratios, say between tympani and triangles, or bass guitar and mandolin, that 40 dB represents 10,000:1, or a *current* ratio of 100:1. The more that the frequency bands are separated, the easier they are to handle, as with the drive units themselves. It has been the experience of the authors that as the frequency bands become narrower, the need for specially selected cables reduces considerably. The Litz wire referred to in Section 6.8 may be beneficial when used in 10 metre lengths and handling the full audio frequency range, but those benefits when used in 2 metre lengths when handling only frequencies below 1 kHz would be minimal.

In a similar way, the separation of the different frequency bands can allow the selection of amplifiers which are less heavy, expensive, heat producing or physically large, without significantly compromising performance. What is more, if *any* non-linearity exists in the circuitry, which it must to some degree, the potential for the production of intermodulation distortion will be greatly reduced. The potential for the low frequencies to modulate the more delicate high frequencies will not occur, and, as discussed in Chapter 5, the generally prevailing opinion is that multi-amplified systems sound ‘cleaner’ than equivalent systems using single, full-range amplifiers. And of course, with multiamplification, multi-cabling is an automatic result.

6.14 System design

Before selecting amplifiers or cables for any loudspeaker system, it is necessary to have a clear view of the purposes for which the entire system will be used. There simply does not exist any ideal amplifier or ideal cable to serve for all purposes. One cable which improves the sonic quality of given combinations of amplifier and loudspeaker may show absolutely no benefit whatsoever with a different combination. One amplifier which sounds stressed at high levels when driving a loudspeaker exhibiting a 'difficult' input impedance may perform excellently in combination with a loudspeaker with a more uniform impedance curve.

Perhaps nothing highlights the concept of component matching quite so easily as carefully listening to the performance of digital-to-analogue converters (DACs). The subtle, or even not-so-subtle, differences between better quality DACs will be revealed by a high resolution monitor system in a good listening room. However, it is possible to substitute loudspeakers of progressively lesser quality of reproduction, and to continue to repeat the tests until no difference can be heard between the different DACs. To a lesser degree, the same effect may be noticed when changing amplifiers or loudspeaker cables. When considering the components for a system it is not usually considered to be commercially viable to over-specify any component part. For this reason, in general, an amplifier will be used in a powered loudspeaker which is 'only' of a quality up to which it is deemed that no further improvements in amplifier performance would give rise to any significantly noticeable improvement in the sound of the complete system. Even in professional systems, it is rare in the 21st century for manufacturers to over specify anything. However, in terms of professional engineering, as opposed to engineering for the market, it can still be beneficial to design systems with an extra margin of quality to allow for minor deterioration over time, or to allow for minor degradations elsewhere in the chain. A 1000 euro, star quad, polarised, screened cable may possibly make a difference in geographical areas of high electromagnetic interference, and when used with amplifiers which are sensitive to having long 'aerials' hung on to their outputs, but there would be no justification using the same cable in interference-free areas with amplifiers which were immune to the problem.

What equipment is truly professional and what is not can be difficult for many inexperienced people to decide in days when marketing is so aggressive, and when 'professional' is a word often used more as a sales device than an accurate representation. In fact, as so much recording is carried out in places which could hardly be described as professional in the traditional sense, and yet which do much work commercially, the whole line between professional and domestic equipment has become very ill-defined.

References

1 Newell, P. R., 'Studio Monitoring Design', Chapter 6, Focal Press, Oxford, UK (1995)
2 Newell. P., Castro, C., Ruiz, M., Holland, K., Newell. J., 'The Effect of Various Types of Cables on the Performance of High Frequency Loudspeakers', Proceedings of the Institute of Acoustics, Vol 24, Part 8, Reproduced Sound 18 conference, Stratford-upon-Avon, UK (Nov 2002)

3 Czerwinski, E., Voishvillo, A., Alexandrov, A., Terekhov, A., 'Multitone Testing of Sound System Components – Some Results and Conclusions, Part 1: History and Theory', Journal of the Audio Engineering Society, Vol 49, No 11, pp 1011 – 1048 (Nov 2001)
4 Czerwinski, E., Voishvillo, A., Alexandrov, A., Terekhov, A., 'Multitone Testing of Sound System Components – Some Results and Conclusions, Part 2: Modeling and Application, Journal of the Audio Engineering Society, Vol 49, No 12, pp 1181 – 1192 (Dec 2001)
5 Castro. S.V., Newell, J. P., Ruiz, M., Holland, K. R., Newell, P. R., 'Loudspeaker Cables for High Frequency Transducers – A Further Assessment', Proceedings of the Institute of Acoustics, Vol 24, Part 8, Reproduced Sound 19 conference, Oxford, UK (Nov 2003)

Bibliography

1 Duncan, B., 'High Performance Audio Power Amplifiers', Newnes, Oxford, UK (1996 – Reprinted with revisions 1997)
2 Newell, P. R., 'Studio Monitoring Design', Focal Press, Oxford, UK (1995)
3 Newell, P. R., 'Recording Studio Design', Focal Press, Oxford, UK (2003)
4 Harris, S., 'Class D Audio Power Amplifiers' Audio Engineering Society, 18th UK Conference Proceedings, London (2003)
5 Colloms, M., 'High Performance Loudspeakers', 6th Edition, John Wiley & Sons, Chichester, UK (2005)

Chapter 7

Loudspeaker behaviour in rooms

7.1 The anechoic and reverberation chambers

It is customary to measure loudspeakers in anechoic chambers, which represent as closely as possible a free acoustic field with no boundaries. The wedges which line the surfaces of most chambers serve to avoid abrupt changes in acoustic impedance and absorb from all angles to a uniform degree. The very high degree of absorption achievable with such systems functions down to a frequency where the length of the wedges exceeds one quarter of the wavelength. Below this frequency the absorption begins to reduce, and reflexions begin to occur. A room with one metre wedges would therefore be anechoic down to a frequency which has a four metre wavelength, which corresponds with a frequency of around 85 Hz. What happens below this frequency depends on the mounting conditions, the shell structure, and the size of the chamber – the larger the chamber, the weaker will be the reflexions at the measuring positions.

Loudspeaker measurements are made in such rooms because they provide an invariable reference which enables like to be compared with like. Real rooms tend to be like human fingerprints, with no two being exactly alike, so measurements made in real rooms would not be representative of the loudspeaker, alone. Although gating techniques exist which can allow computer-based measurement systems to eliminate the room effects, the gating systems themselves tend to introduce anomalies of their own. Large anechoic chambers, although very expensive to build (upwards of 1000 euros per square metre of total surface area for a 70 Hz chamber), really have no equal for measurement accuracy and repeatability.

Where absorption is very high, but incomplete, a room can be considered to be *semi-anechoic*. In such circumstances, and for given positions in the room, a response equalisation curve can be generated which allows compensation of the measurements for the small, smooth irregularities of the room, allowing an anechoic response plot to be realised in less than true anechoic conditions. A *hemi*-anechoic chamber, on the other hand, has one of its surfaces highly reflective and very rigid. If either the source of sound or the measuring microphone is set into the rigid boundary, no reflexions will be evident, although at higher frequencies, where the equipment in the room (either loudspeaker or microphone support) is comparable in size to the wavelength, a weak reflexion may return after travelling between the equipment, the hard-wall surface, and back again; or vice versa. The measured response in such chambers will show a rise at low frequencies due to the mounting in the boundary, the effect of which will be discussed later in the chapter. Fortunately, this can easily be compensated for by electrical means to yield a similar response to a fully anechoic measurement. The hemi-anechoic chamber represents radiation into a hemisphere, whereas in a fully anechoic chamber the waves are free to radiate spherically. The formula for the surface area of a sphere is $4\pi r^2$ ($4\pi\times$ the square of the radius), and for a hemisphere $2\pi r^2$. Consequently, the spaces are commonly referred to as 4π and 2π spaces. An anechoic chamber is shown in Figure 7.1.

At the other extreme, we have the reverberation chambers. These are built with massive, rigid, non-resonant walls and a non-parallel construction. The rooms are typically made from concrete, with the inner surfaces plastered and painted to provide a very smooth surface which is highly reflective up to the limits of audibility. The aim is to provide a uniformly diffuse sound field at all points in the chamber, although when the dimensions become small by comparison to the wavelength some modal, spacially dependent responses begin to appear. By means of the diffuse mixing of the sound, the total energy radiated by the source can be measured at any point within the volume of the chamber that is not too close to a boundary or the direct field of the source itself. Figure 7.2 shows a reverberation chamber with a decay time of around 8 seconds. It is useful in a chapter about room effects to begin with the two extremes of anechoic and diffusely reverberant spaces because all practical rooms lie somewhere in-between, and to a greater or lesser degree exhibit the properties of both of the above extremes.

The response of a loudspeaker measured on-axis and at various angles off-axis in an anechoic chamber is shown in Figure 7.3. The response of the same loudspeaker measured in a reverberation chamber is shown in Figure 7.4. If one can imagine measuring the loudspeaker in an anechoic chamber at all angles and in all planes, and then integrating the results, the

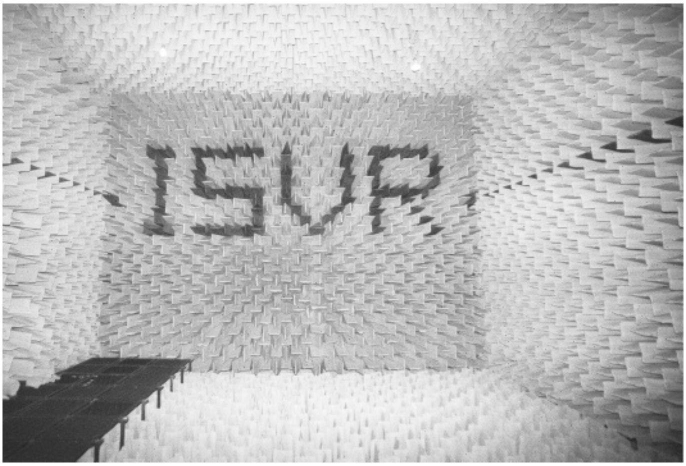


Figure 7.1 The large anechoic chamber at the ISVR, Southampton University, UK. A section of the removable floor grids can be seen at the left-hand side. The wedges are made of glass-fibre, covered in muslin, and are one metre long. Above 70 Hz, the surfaces are almost totally absorbent

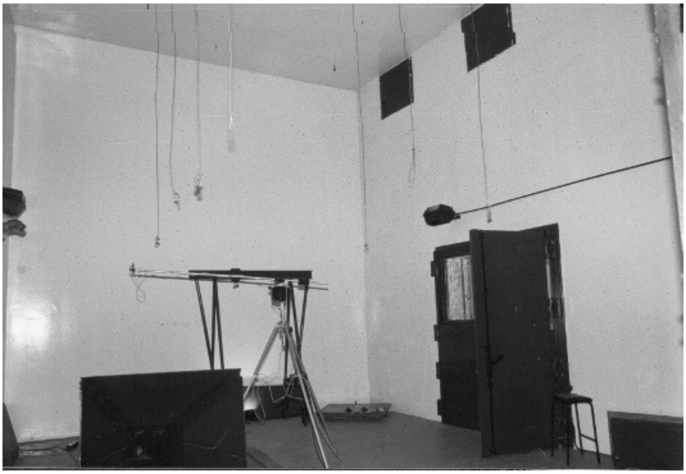


Figure 7.2 The reverberation chamber, adjacent to the anechoic chamber shown in Figure 7.1. The reverberation time at low frequencies is around 8 seconds. At the left of the photograph, the flash from the camera can be seen reflecting from the shiney wall surface

mean response of all of the plots such as the ones shown in Figure 7.3 would look like the response shown in Figure 7.4. The reverberation chambers therefore act as acoustic power integrators, and provide a much quicker means of determining the total power response than the thousands of individual measurements that would need to be taken in an anechoic chamber.

Most cabinet loudspeakers are omnidirectional at low frequencies and rather directional at high frequencies, therefore the extra low frequency power radiating in all directions is the reason for the bass rise in Figure 7.4. To maintain a flat response on axis, more total power must be radiated by the loudspeaker at the frequencies which spread out over a greater volume of the room. There are two common means of describing the directivity, the directivity index (DI) and the directivity factor. The DI is given by $10 \log_{10} Q$, where Q is the directivity factor. In plain English, the DI is the ratio in decibels of the on-axis sound pressure compared to what it would be if the same radiated energy was distributed omnidirectionally. The directivity *factor* is the ratio of the actual directivity to the omnidirectional directivity, and has no units. For example, a source radiating into a hemisphere would have twice the directivity factor of a source radiating into a free field. Radiating into quarter space from a floor/wall junction would yield a directivity factor of 4. Q is perhaps the most commonly used symbol for directivity factor, but one also sees D_F, λ and even R with a subscript Greek letter theta (θ).

7.2 Boundary loading and room gain

When a loudspeaker radiates in unrestricted space, the radiated waves are free to expand in all directions. Below about 250 or 300 Hz, the radiation pattern from a 'monopole' source (such as a loudspeaker cabinet) tends

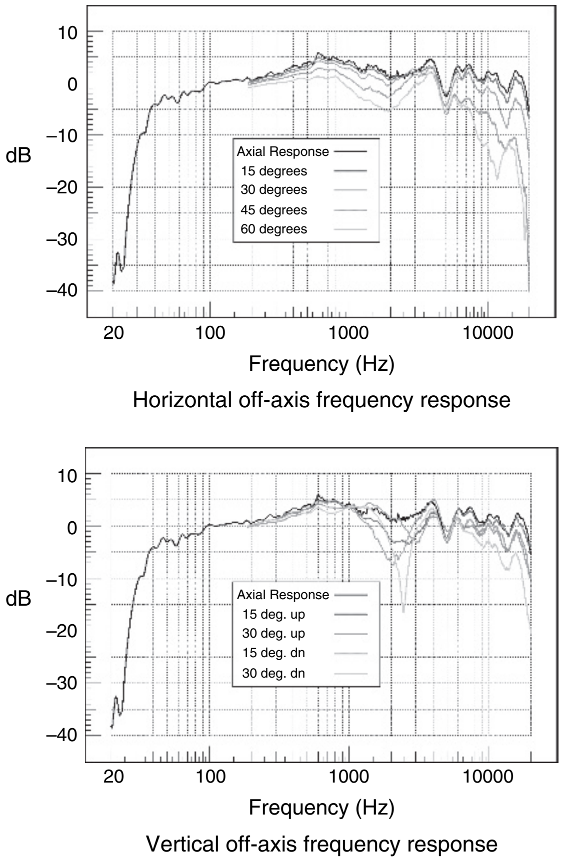


Figure 7.3 Axial and off-axis pressure amplitude responses as measured in an anechoic chamber

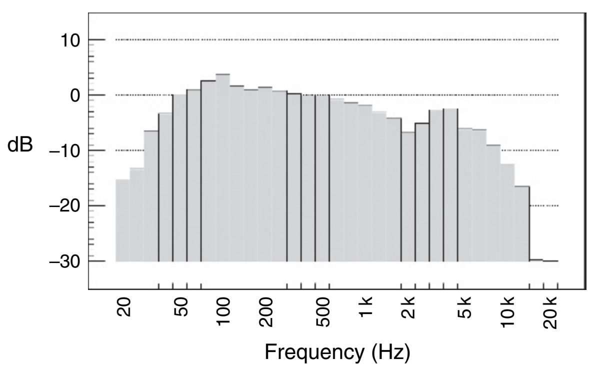


Figure 7.4 Total power response of the loudspeaker whose anechoic responses are shown in Figure 7.3. The reverberation chamber sums the responses from all directions

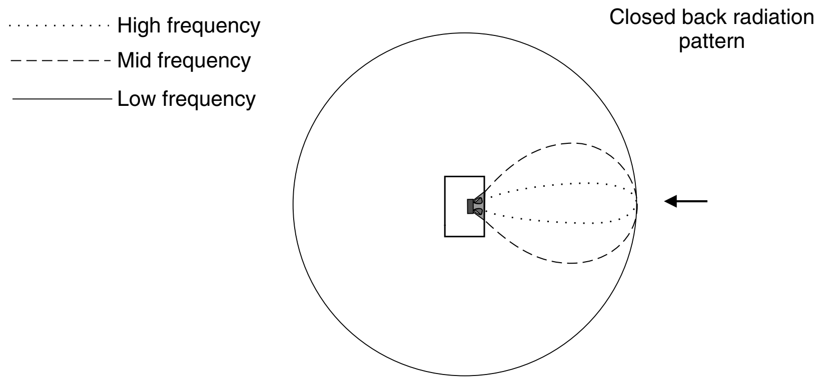


Figure 7.5 Typical radiation pattern from a closed back loudspeaker cabinet
On the axis (arrowed) all frequencies arrive with equal pressure, but the frequencies which radiate over a wider area also radiate more total power in the room, giving rise to the type of power response shown in Figure 7.4

to be omnidirectional, but at higher frequencies, for conventional loudspeakers with the drivers mounted on a single front baffle, the radiation becomes more directional, as shown diagrammatically in Figure 7.5. For dipole sources, which radiate from both sides of their diaphragms, a figure-of-eight radiation pattern is more typical, as shown in Figure 7.6. Most cabinet loudspeakers, whether sealed boxes, reflex cabinets or transmission lines, act as monopole sources at low frequencies. Examples of common dipole sources are flat panel loudspeakers, such as some electrostatics, and open-backed guitar amplifier loudspeakers.

When a wave expands spherically, the sound *intensity* (which is measured not in sound pressure level but in *watts per square metre*) is distributed over

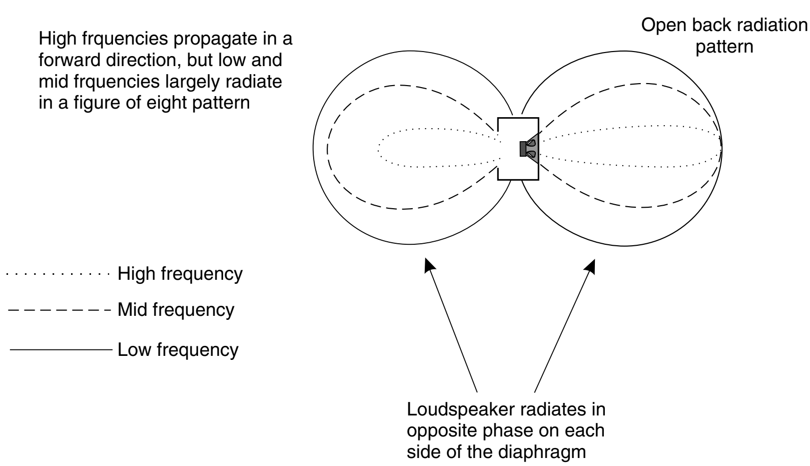


Figure 7.6 Typical radiation pattern from an open-backed loudspeaker cabinet
It is typical for less mid and high frequencies to radiate behind the loudspeaker cabinet due to mechanical obstructions

the surface of the expanding wavefront. Given a sphere of one metre radius, the area of its surface calculated from the formula given in Section 7.1 ($4\pi r^2$) would be:

$$4\pi \times 1^2 = 4\pi \times 1 = 4\pi\,m^2 \qquad 12.6\,m^2$$

A sphere of *two* metres radius would have a surface area of:

$$4\pi \times 2^2 = 4\pi \times 4 = 16\pi\,m^2 \qquad 50.3\,m^2$$

which is four times the area of a sphere of one metre radius

A sphere of four metres radius would have surface area of:

$$4\pi \times 4^2 = 4\pi \times 16 = 64\pi\,m^2 \qquad 201\,m^2$$

π, of course, being approximately 3.142

Sixty-four π square metres is four times $16\pi\,m^2$ which is four times $4\pi\,m^2$. Therefore, each time that we have doubled the radius we have increased the surface area by four times:

$1\,m$ radius $= 4\pi\,m^2$ surface area

$2\,m$ radius $= 16\pi\,m^2$ surface area

$4\,m$ radius $= 64\pi\,m^2$ surface area

The sound intensity is a measurement of power distribution, so for any point on the surface of the expanding wave the intensity is reduced by a quarter each time that the distance from the source is doubled, because the same radiated power is spread over four times the surface area. Each time that *intensity* (or power) is halved, the level reduces by 3 dB. Each time that it is doubled, it is increased by 3 dB, so when the intensity is reduced to one quarter (one half of one half) it reduces by 6 dB (3 dB + 3 dB). Therefore, as the radius of a sphere doubles, and its surface area increases by a factor of four, the intensity at each point on the surface of the sphere is also reduced by a factor of four, and thus falls by 6 dB. This is the principle behind the well-known 'double distance rule', which states that each time the distance from the source is doubled, the sound pressure level (SPL) drops by 6 dB. The reduction is nothing to do with absorption in the air, but is purely a result of spreading the power over a greater area. Air absorption losses at short distances are negligible.

If the wave propagation were to be restricted to travelling along a tube, as shown in Figure 1.1, no expansion would take place, and therefore the only losses would be those incurred due to the non-perfectly smooth-and-rigid walls of the tube acting as absorbers. This is why it was possible in old ships to provide speaking tubes from the bridge to the engine room, and to speak and to be heard clearly over great distances and in noisy environments. In the 19th century, experiments were conducted with speaking tubes which were usable over distances of three kilometers, and conversations (by Biot, in Paris) in a low voice were possible at 1 km[1].

Note, however, that although the sound *power* reduces by 3 dB when it is halved, the halving of sound *pressure* requires a 6 dB reduction. This is because the sound power is proportional to the *square* of the sound pressure. In electrical terms, voltage can be considered to be the electrical pressure, or electro-motive force. In Spanish and Portuguese, for example, the word 'voltage' translates as 'tension', just as in English we can refer to the very high voltages on a cathode ray tube as EHT, or extra high tension. In Spanish and Portuguese, blood pressure also translates as 'blood *tension*', so this concept of using the same word 'tension' highlights the link between voltage and pressure. The well-known equation for calculating the output power of an amplifier from the signal voltage and the load resistance is

$$W = \frac{V^2}{R} \tag{7.1}$$

Where W is the power, V the voltage and R the resistance.

This clearly shows the voltage (pressure) *squared* relationship to the power. Power, on the other hand, is power, whether mechanical, electrical, or heat, and a decibel is always a *power* ratio. A decibel is also always a decibel: there are no separate decibels for power, voltage, pressure or intensity. The relations are fixed and the decibel never changes. If a reduction of 3 dB takes place, the power is halved and the pressure falls by $\sqrt{2}$. If a reduction of 6 dB takes place, the power is quartered and the pressure halves. This relationship needs to be well established here, because when dealing with rooms, we need to deal with radiated *power* from the sources but we measure sound *pressures* within the rooms. This is principally because human ears are pressure detectors, so hearing relates better to pressure changes than to power changes, but loudspeakers radiate acoustic power and heat. Loudspeakers do not radiate pressure, because when all of the positive and negative pressure half-cycles are summed, no net pressure change takes place. A bomb-blast, on the other hand, radiates a unidirectional pressure wave.

7.2.1 Restriction of radiating space

If we mount a loudspeaker cabinet in an infinite, plane boundary, such as is simulated by the hard wall in a hemi-anechoic chamber, two things occur relating to the radiation of the low frequencies which are different to what happens when the loudspeaker is mounted in free space. Firstly, as no sound can radiate behind the source, the radiation can only radiate in a forward direction. As all of the power is radiated forwards, the pressure on-axis will rise by 3 dB, as the half of the radiated power which would have travelled behind the loudspeaker is driven forwards. Secondly, the increased pressure on the face of the diaphragm, resulting from this restriction of the expansion of the wave, tends to resist the movement of the diaphragm. In effect, the radiation impedance has been increased, which in turn gives the diaphragm something more substantial to act upon. The mounting surface can be thought of as a special case of a 180 degree horn of infinite flare rate (see Chapter 4), so it provides a more resistive

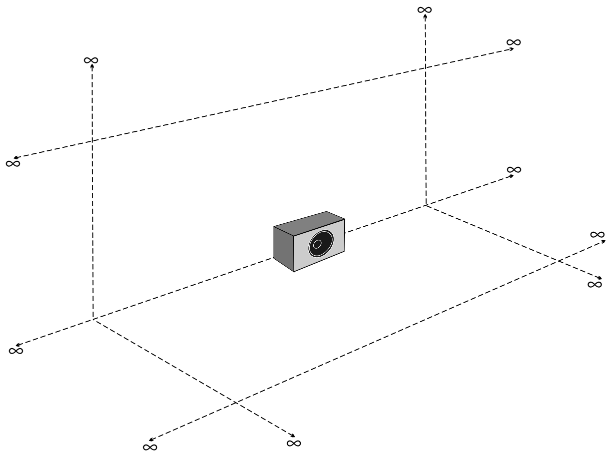


Figure 7.7 Radiation into quarter space (π space)

A loudspeaker bounded by two infinite planes at right angles to each other could only radiate into one quarter of the space compared to that of a free-field. The space restrictions would both constrain the expansion of a propagating wave and give rise to a greater acoustic loading on the diaphragm

termination. This extra loading gives rise to approximately another 3 dB of increased axial sound pressure, resulting from the increase in radiated power, although the precise increase is related to the loudspeaker sensitivity.

If another boundary is introduced, just below the loudspeaker, as shown in Figure 7.7, the same effects occur. This is known as 'quarter space' loading, or π space, because the free-field (4π space) has been halved by the introduction of the first plane (2π space) then halved again by the introduction of the surface immediately below the loudspeaker (π space). If a further plane, rigid boundary surface is introduced at 90 degrees to each of the other surfaces, as shown in Figure 7.8, these effects will be repeated once more. This is akin to mounting a loudspeaker on the floor (or ceiling) in the corner of a room. The potential 6 dB increase which is added to the axial response as each boundary surface is added means that when placed in a three surface corner of a room, the axial radiation at low frequencies can be up to 18 dB (6 dB + 6 dB + 6 dB) higher than when the loudspeaker is receiving the same electrical input power and is radiating at the same distance from the measuring point in free space. This boost is a minimum phase effect, and incurs no penalty in the time or phase responses of the loudspeaker. There is no resonance associated with this effect, and no reflexions exist so long as the radiated wavelengths are substantially larger than the largest physical distances between the radiating surface and the walls enclosing it. This will be discussed further in Section 7.8. The loading can be thought of as mounting the radiating surface at the throat of a horn of triangular cross-section, as shown in Figure 7.9.

A clear significance of this boundary loading effect is that a domestic loudspeaker which is considered to be light on bass may have its LF response reinforced by placing it close to a corner. Conversely, if a corner

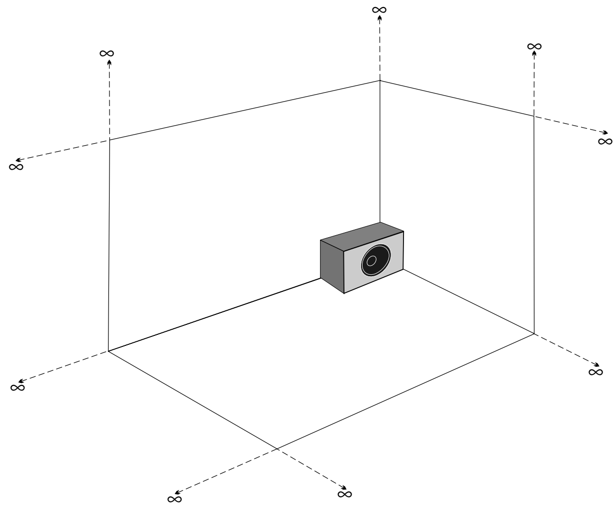


Figure 7.8 Radiation into eighth-space (1/2 π space)
Radiating from the corner of three infinite, rigid boundaries, the loudspeaker would be constrained to radiating into one-eighth of the space of a free-field. In ideal conditions, the low frequency axial sensitivity of the loudspeaker could be up to 18 dB higher than for the same power input into free space

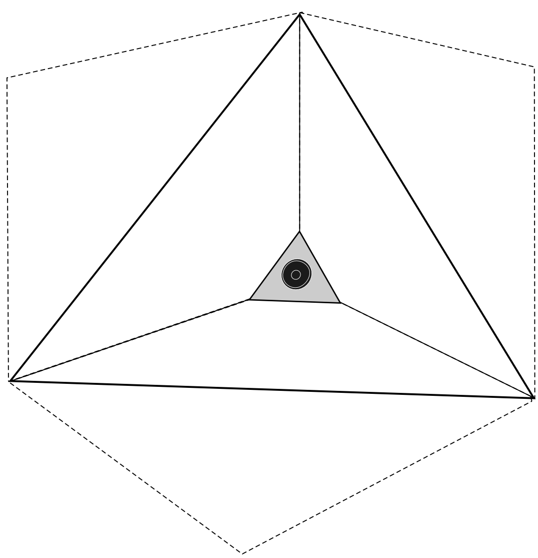


Figure 7.9 A triangular horn
In this representation, Figure 7.8 has been re-drawn to show how, in effect, the loudspeaker constrained by three mutually perpendicular planes is mounted in a triangular horn

is the only position suitable for the placement of a loudspeaker in a room, then a loudspeaker should be chosen which is somewhat bass light in its free-field response if a flat response in the room is the goal. In fact, one of the principal uses of tone controls on hi-fi systems is to compensate for such changes of overall responses due to loudspeaker positioning. The additional radiating efficiency also means that a loudspeaker which is placed in a position where its output is augmented 6 dB with respect to its input will need only one quarter of the power in order to achieve the same SPL in the room. Reliability will be increased and distortion will be reduced. However, in real rooms, other complications may arise such as the driving of extra room modes, and colouration due to the response boost not exactly matching the inverse of the natural roll-off of the loudspeaker.

7.2.2 The mirrored room and mutual coupling

Another way of looking at the above concept is to imagine that the surfaces constraining the loudspeaker were all mirrored, and that in every place where one saw a reflexion of the loudspeaker a real loudspeaker existed which was radiating exactly the same signal. This idea is shown in Figure 7.10, and indeed, if one removed all of the reflective surfaces and put real loudspeakers in place of the reflexions, then the response at low frequencies would be exactly the same as that with one loudspeaker and the boundaries in place. In the case of a single boundary, the low frequency response rose by 6 dB. If two loudspeakers were placed back to back, with no boundary between them and fed with the same signal at the same power level, the result at low frequencies, radiating omnidirectionally, would also be a 6 dB increase. Now, if we drove each loudspeaker with one watt, the total input power would be double that of one loudspeaker, which would be a 3 dB increase. The additional 3 dB needed to yield the 6 dB output increase experienced in real situations is a result of a phenomenon known as mutual coupling.

When one loudspeaker radiates into free air, the local pressure on the diaphragm is that due to its own motion. When an identical second loudspeaker is added, placed very close to the first loudspeaker and driven with the same signal and with the same phase relationships, the local pressure into which each diaphragm radiates is its own pressure *plus* the pressure radiated by the other, adjacent diaphragm. This additional pressure exhibits itself as an increased radiation impedance, which gives rise to more sound being radiated.

Large diaphragms, in general, radiate low frequencies more efficiently than smaller diaphragms because there is a mutual coupling effect between all the individual sections of the one diaphragm. One can think of a large diaphragm as being made up of many small, individual surface areas. The bigger the diaphragm, the more difficult it is for the air in contact with it to move out of the way of the diaphragm motion. The 'congestion' of the air particles presents a more resistive load to the vibrating surface of the driver.

So far, we have only been discussing mutual coupling at low frequencies, but it is not just a low frequency phenomenon. In fact, it occurs at all

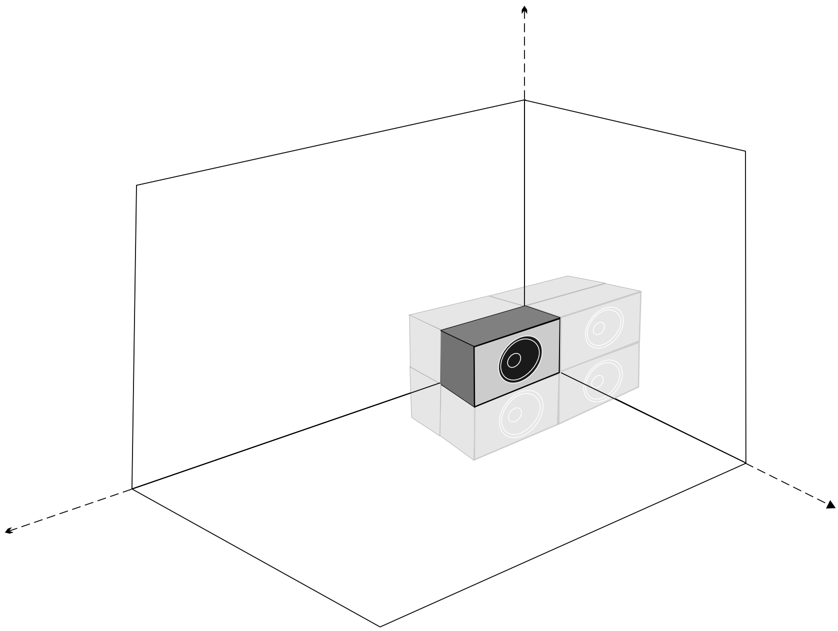


Figure 7.10 The mirrored room

If we imagine a loudspeaker radiating low frequencies from the corner of a room with mirrors on all surfaces, what we would see would be cluster of eight loudspeakers. The power radiated by the single loudspeaker constrained by the walls and the floor is exactly the same as would be radiated in that space if the walls and floor were removed and the cluster of eight loudspeakers, each receiving the same power input as the one loudspeaker in the constrained space, were radiating into a free-field. Within the frequency range of mutual coupling, each doubling of the quantity of loudspeakers, all receiving the same input power, will raise the radiated output by 6 dB. Therefore, compared to one loudspeaker, two will radiate 6 dB more power, four 12 dB more, and eight 18 dB more power, hence the acoustic power increase referred to in Figure 7.8. However, this theoretical maximum is difficult to achieve, and may reduce with increasing loudspeaker sensitivities.

frequencies, but it is only at low frequencies where it is entirely constructive. It can be considered to be 100% constructive up to frequencies where the distance between the radiating surfaces is no more than one eight of a wavelength, which means that the radiation from one loudspeaker reaches the other substantially in-phase with its own radiation. As the distance (or frequency) increases, the mutual coupling boost diminishes and actually becomes slightly negative at some wavelengths as the phase relationship of the two pressure components drift further apart. At still higher frequencies the mutual coupling effect becomes negligible. The zones of coupling are shown diagrammatically in Figure 7.11.

7.3 Room reflexions

When a vibration in the air reaches any boundary, three things will occur. Part of the energy will be *absorbed*, and converted into heat. Another

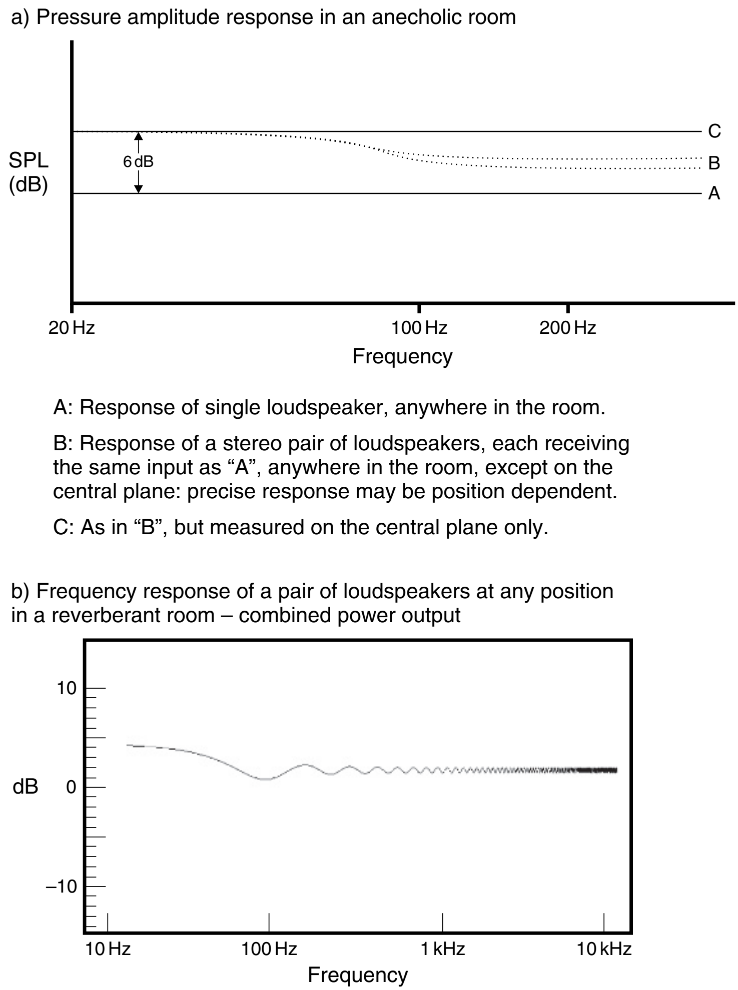
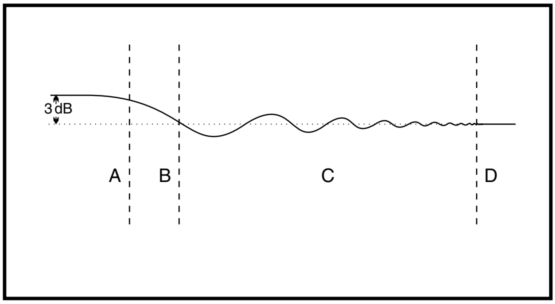


Figure 7.11 Mutual coupling effects – omnidirectional sources (continued on next page)

part of the energy will pass through the boundary, and will be *transmitted* beyond it. The remaining energy will be *reflected* back into the room. As nothing on this planet is either perfectly absorbent, or perfectly reflective, all the above things must occur on each and every contact of a sound wave with a surface. Reflexions from room boundaries will combine at any given point with the direct radiation from the loudspeaker, and the pressure at that point will be the sum of the direct sound and all the individual reflexions which pass through it at the same time. Note, however, that the direct sound will still fall off at a rate of 6 dB per doubling of distance from the source, even though the overall direct/reflected sound level may fall off at a much lower rate.

It is also important to understand that when pressures cancel to zero, the waves do not disappear. They merely pass through each other. Sound

c) Zones of loading/coupling – general response as in b)



Zone A: Region where the separation distance between the loudspeakers is less than the quarter wavelength distance, and where wholly constructive mutual coupling is effective.

Zone B: Region where the separation distance between the loudspeakers is less than half a wavelength, but where the mutual coupling is becoming less effective as the frequency rises.

Zone C: Region where the separation distance is greater than half a wavelength, and where the mutual coupling alternates, as the frequency rises, between being constructive or destructive.

Zone D: Region where the mutual coupling has ceased.

Figure 7.11 Continued

waves have a pressure and a velocity component, and when either one is at a maximum, the other is at a minimum. The concept is similar to that of a swinging pendulum – when the *height* is at a *maximum*, the pendulum stops, before falling back, so the velocity is at a *minimum*. As it swings through the bottom of its trajectory, the height is at its *lowest* point (*minimum* height), but the *velocity* is at a *maximum*. Likewise, when a sound wave hits a wall, the particle velocity must be zero, because the wave motion must stop and change direction. The energy, which can neither be created nor destroyed, must therefore all reside in the pressure. In the pendulum, the energy constantly cycles through potential (height) and kinetic (velocity), and in an acoustic wave it constantly cycles between pressure (potential energy) and velocity of particle motion (kinetic energy). These motions are 90 degrees out of phase. If one considers a sine wave, the energy cannot be at zero when the pressure passes through zero, because it if had disappeared, where would the energy come from to produce the next half cycle? It comes, of course from the velocity component, which was actually at a maximum (it had all the shared energy) when the pressure wave passed through zero.

The pressure components of reflected waves can also load loudspeaker diaphragms, but the fact that they may have travelled considerable distances means that their sound pressures will have fallen significantly by the time that they return to the diaphragm of their origin, so the effect is small. When two identical signals (either electrical or pressure waves) are added together, if either one is more than around 6 dB lower than the other, the total signal will not be significantly greater than the larger of the two, alone. Therefore, reflected energy in the far-field of a loudspeaker will have more effect in terms of the way in which it combines by superposition with the direct signal rather than any effect that it has by directly impinging on the diaphragm.

7.3.1 Resonant modes

A special case arises when a reflexion can, through multiple reflexions, retrace its own path, and arrive at a boundary with the same phase relationship and in the same direction as when it left, as shown in Figure 7.12. In this case, known as a resonant mode, each subsequent reflexion superimposes itself on the previous one, or the sum of the previous ones, and the energy in the mode builds up. Once the source of the energy is switched off, the energy stored in the mode will take some time to decay, dependent upon the reflectivity of the surfaces between which it is resonating. There is much more energy trapped in resonant modes than is normally encountered in single reflexions, and it is the modes which are largely responsible for the colouration which untreated rooms impart to the responses of nominally clear sounding loudspeakers. The energy in a modal resonance can be *higher* than the direct energy received from the source, whereas simple, reflected energy is always lower.

One sometimes hears the term ‘standing wave’ being used for resonant mode. Whilst it is true that all resonant modes *are* standing waves, all standing waves are not resonant modes, so the term ‘room modes’ is preferable to ‘standing waves’. Likewise, whilst all cows are mammals, all mammals are not cows. One can therefore call a cow a mammal, but one cannot call a lion a cow! [And whilst both Scots and English are British, it is wisest not to call a Scotsman ‘English’ or serious violence is likely to ensue!]

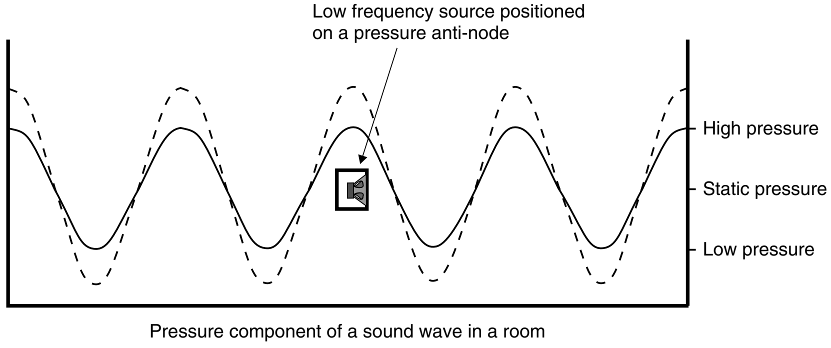


Figure 7.12 The driving of a resonant mode
At resonance, where an exact number of half-cycles can fit between the walls, the direct and reflected waves will superimpose constructively to create a resonant build-up, shown dashed. When the source driving the mode is switched off, the energy trapped in the mode will continue to resonate until it is dissipated by absorption

In Figure 7.13, modal pressure and velocity distribution are shown superimposed. If a conventional monopole loudspeaker, which radiates omnidirectionally at low frequencies, is placed on the node of the pressure wave, where the pressure is at a minimum, it will be unable to radiate a steady tone at that frequency which coincides with the mode, because the modal energy will constantly cancel the direct energy. Conversely, if the loudspeaker is placed at an anti-node, where the pressure is maximum, it will strongly reinforce the mode, just like a correctly timed push of a child on a swing, at the peak of the travel, will add impetus to the motion. On the other hand, a dipole loudspeaker, which radiates reverse polarity pressures equally from each face of its diaphragm, acts as a pressure source, and not a volume/velocity source like a monopole, and couples best at the *pressure node* of a mode, which is, of course, a *velocity anti-node* because of the 90 degree phase shift between the pressure and velocity peaks, as shown in Figure 7.13.

Therefore, if one were to compare a figure-of-eight radiating, electrostatic loudspeaker and a normal cabinet loudspeaker with an almost identical on-axis frequency response when measured in an anechoic chamber, the responses could be expected to vary widely when placed in the same place in a room with reflective boundaries. Part of the difference would be due to the difference in radiation patterns, and hence the different ways in which each one drove the reverberant field. The other great source of variation, particularly at low frequencies, would be due to the different ways that each loudspeaker coupled to the room modes. Consequently, when comparing the sound of open-backed and closed-backed guitar amplifiers it is useless to do so without finding room positions for each of them individually which support their tone. Similarly when assessing domestic high fidelity loudspeakers of monopole or dipole nature, it is essential to consider how each one will couple to the room modes. Clearly, one position will be unlikely to suit both, and it should also be remembered that because of the way that a dipole radiates in a figure-of-eight pattern, it will only drive the modes which occur between the parallel room surfaces which face the front and rear of the loudspeaker, whereas the omnidirectional

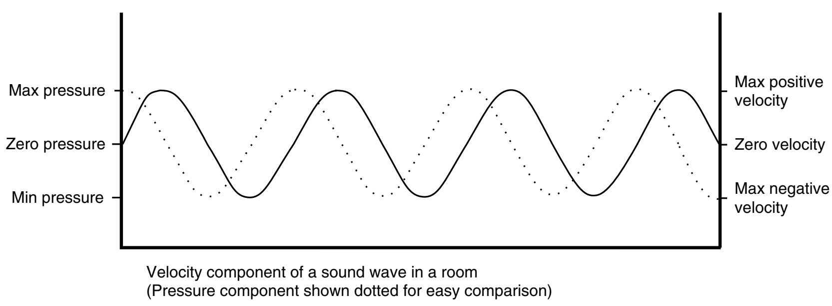


Figure 7.13 The relationship between pressure and particle velocity in a resonant mode

The particle velocity component of the wave must always be at a maximum when the pressure component is at zero, in order to conserve the energy contained in the wave. The energy cannot simply disappear when the pressure is zero, or there would be no energy left to continue the cycle low frequency radiation from a monopole source will drive modes in all three dimensions.

7.4 Flush-mounting

The response of a loudspeaker, except in a free acoustic field (i.e. without boundaries) is absolutely inseparable from the modifications which will be imposed by the room within which it is placed. Rooms form *part* of a loudspeaker system because they provide loading on the diaphragms which actually affect the sound radiation directly from the source. Rooms are not simply environments in which loudspeakers with fixed responses are used. If every position in every room will give rise to its own characteristic response for any loudspeaker placed there, it introduces a great variable in music monitoring conditions if the loudspeaker placement is not *very* carefully chosen. In many music recording studios, and in almost all film dubbing (mixing) theatres, the principal loudspeakers are flush mounted in the front wall.

By this means, questions of room positioning are eliminated, and as all rigid walls are pressure anti-nodes for *all* the modes which they support, these modes will all be driven equally. Conversely, no position *within* the room can drive all of those modes, because the node and anti-node positions vary with frequency. Flush-mounting loudspeakers in a boundary also means that no rear radiation can take place, so none can bounce off the wall behind the loudspeaker and return to the listening position with varying phase relationships to create peaks and dips in the pressure amplitude. What is more, flush mounted loudspeakers will effectively experience 2π radiation, as explained in Section 7.1, and will benefit from the greater sensitivity afforded by the increase in the radiation resistance and the constrained angle of radiation. In a room, which already constrains the wave expansion to some degree, as opposed to in a free-field, the sensitivity increase due to flush-mounting may be around 3 dB to 6 dB. This can be useful at high power levels because the halving of the heat in the voice coil even due to the 3 dB power reduction can be important in reducing thermal compression and increasing long-term reliability. Distortion can also be reduced due to the smaller cone excursions necessary to generate the same SPL. Frequency dependent cabinet edge diffraction effects are also eliminated. When all is considered, flush-mounting enjoys many advantages over free-standing, and no obvious disadvantages; at least in mono and stereo rooms. For these, flush-mounting is the choice of the great majority of large, professional studios. Of course, there also exists a further, non-electro-acoustic advantage for flush-mounting – it leaves the room much less obstructed, especially when very large loudspeaker cabinets are used.

One problem which *can* arise from flush-mounting if it is not well executed is that the front wall can vibrate, thus radiating an unwanted extra, resonant sound energy into the room. Some designers favour the resilient mounting of the cabinets, to avoid transmitting the cabinet vibrations into the structure. Other designers favour the rigid mounting of the cabinets into a heavy, rigid structure, too heavy to significantly vibrate with the available energy. Resilient mounting in a heavy structure is another option, but the advantage of rigid mounting in a heavy wall is that the cabinets are kept absolutely stable under all drive conditions. Cabinets that are intended for mounting in this way usually have no decorative finish except on the front face, because they will be fixed into the frame of the structure during construction.

7.5 Multichannel considerations and phantom imaging

When loudspeakers act in multiples, the response is not necessarily merely the sum of the individual components. When two loudspeaker cabinets are brought close together when reproducing different, uncorrelated signals, at the same sound pressure level, the overall pressure at the same distance from the loudspeakers will rise by 3 dB, which is the simple sum of the two radiated power levels. On the other hand, if the loudspeakers are radiating the *same* signal, the *on-axis* pressure will increase by 6 dB, due to the in-phase pressure superposition, although elsewhere in the room, at places where the phases cancel, lower levels will be evident. Figure 7.14 shows the

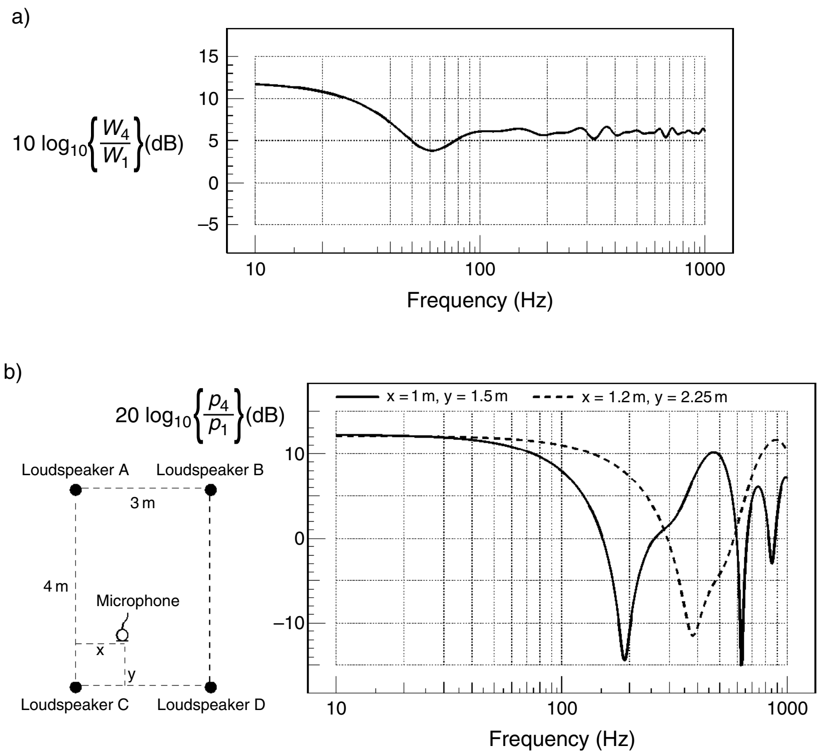


Figure 7.14 Frequency responses for four 'perfect' loudspeakers radiating the same signal. a) Total power response in a highly reverberant room. b) Combined frequency response of four omnidirectional loudspeakers at two different positions in an anechoic chamber

pressure amplitude of the frequency response of four loudspeakers radiating the same signal, measured in free space at two different positions. The only point where a flat response could be received would be at the single point equidistant from the four sources. The loudspeakers represented in the figure are theoretically perfectly flat sources. The significance of this figure is that it shows that no matter how perfect the loudspeakers, or how flat the room, a single signal fed into four loudspeakers will not be able to deliver a flat response except in one spot, due to the phase cancellations given rise to by the different path lengths from the sources. For each frequency, and hence each wavelength, the spacial distribution will be different.

A wavelength is simply the distance in metres which it takes for a wave, travelling at the local speed of sound, to pass through a full cycle. The frequency, in turn, can be defined as the rate of change of phase with time, so the number of cycles per second is also the number of wavelengths that a sound wave will travel through in one second. The equation that relates frequency to wavelength, normally represented by the Greek letter λ (lambda), is:

$$\lambda = \frac{c}{f} \tag{7.2}$$

where

λ = wavelength in metres
c = speed of sound in metres per second
f = frequency in hertz

For example, for 20 Hz and a sound speed of 340 metres per second:

$$\lambda = \frac{340}{20} \quad 17 \text{ metres}$$

For 100 Hz:

$$\lambda = \frac{340}{100} \quad 3.4 \text{ metres}$$

For 500 Hz:

$$\lambda = \frac{340}{500} \quad 0.68 \text{ metres} = 68 \text{ centimetres}$$

For 1 kHz:

$$\lambda = \frac{340}{1000} \quad 0.34 \text{ metres} = 34 \text{ centimetres}$$

For 5 kHz:

$$\lambda = \frac{340}{5000} \quad 0.068 \text{ metres} = 6.8 \text{ centimetres}$$

For 20 kHz:

$$\lambda = \frac{340}{20000} \quad 0.017 \text{ metres} = 1.7 \text{ centimetres}$$

This relationship ensures that the phase relationship between the acoustic pressures radiated from each loudspeaker will vary in all parts of the area between them. In fact, in free space, each quadrant shown in Figure 7.15 will be a mirror image of the adjacent quadrants, but when we take these loudspeakers into a room, the situation changes dramatically. The reflected energy from the room boundaries will complicate the sound field greatly. Unless all the wall surfaces and structures were absolutely identical, the symmetry of the quadrants would also be lost, because the absorption and reflexion properties would be asymmetrical. As the power radiated by the ideal loudspeakers that we are discussing is considered to be flat with frequency, (and let us presume that they are omnidirectional radiators) then the resulting sound field in a room with perfectly rigid walls would be uniform, with a flat response at all places, because the reflexion density would be so great as to ensure that all places receive all frequencies with all phase relationships. This is the integration process that was referred to at the beginning of this chapter when discussing reverberation chambers. Unfortunately, as neither anechoic chambers nor reverberation chambers make good control rooms, it means that all practical rooms exhibit frequency responses that are frequency and position dependent. The difference

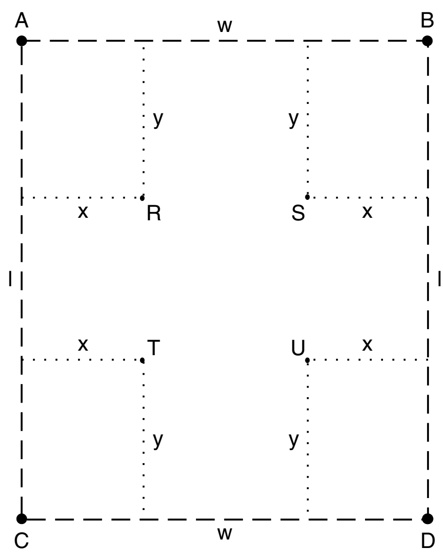


Figure 7.15 Symmetry of radiation patterns of multiple drivers

If loudspeakers at positions A, B, C and D are all radiating the same signal, the microphones at the symmetrical positions R, S, T and U would all receive the same signal. The response at each position would be identical between the response of one loudspeaker in two different places in one room can often be greater than between two rather different loudspeakers in the same place in the same room. Room positioning is therefore a very critical subject, and sound control room designers go to great lengths to try to ensure consistent responses over the critical listening areas.

If we now go back to our four loudspeaker array, and consider its performance in a normal room, the effect on the low frequency performance can be very dependent on the musical signal. With a different instrument in each loudspeaker, the outputs are not correlated, so each loudspeaker will act as an individual source. If the loudspeakers are not symmetrically placed in relation to all three axes of the room, and if the room, itself, is not perfectly symmetrical in shape and construction, then each loudspeaker will drive the room in a different way, and some listening positions in the room may be better for some loudspeakers than for others. Conversely, from the listening position, some loudspeaker positions may be better than others in terms of the received flatness of response.

In the case where the loudspeakers are all receiving the *same* input signal, such as a centrally panned instrument, the situation changes. The correlated signal now means that there are four physically displaced sources, with a very fixed phase relationship between them. As previously stated, the central position would be the only one where the flat response could be received, even *without* the room reflexion complications. Furthermore, mutual coupling between the sources would also take place at frequencies where the distance between the loudspeakers was less than about a quarter of a wavelength, or where a boundary was within about an eighth of a wavelength of a source. The overall frequency response would thus change as the distance between the loudspeaker was changed, or as the loudspeakers were moved in relation to the walls, floor, or ceiling of a room. Not even *one* loudspeaker can produce a flat frequency response in a non-anechoic room, but when *four* loudspeakers radiate the same signal, the situation can become much more complicated. This is a subject that will be discussed further in Chapter 12.

The abovementioned effects have implications for signal panning, also. If a signal in one loudspeaker is panned into the centre of a pair of loudspeakers, the central image, when emanating from two loudspeakers, cannot have the same, in-room response as when only emanating from one loudspeaker. In a totally dead room, when each loudspeaker in a pair is radiating the same signal, the central phantom signal, on the central plane, will be 6 dB higher than the signal radiations from either the left or right loudspeakers alone. It will be 3 dB higher than the sum of the outputs of the left and right loudspeakers radiating individually. This is due to pressure summation on-axis, because when the *pressure* doubles, the response is 6 dB higher. In a reverberant room, no extra level will be detected at the listening position between a single, central source and the phantom sum of a stereo, two sided source, (except at low frequencies where mutual coupling occurs), because the room reflexions will tend to integrate the overall *power* response.

Most mixing console manufacturers use pan-pots which are about $4\frac{1}{2}$ dB down in the centre position, not only because most rooms lie somewhere in-between the anechoic and reverberation chamber responses, but also because mono electrical summing also needs pots which are 6 dB down in the centre (doubling the voltage gives a 6 dB increase). As was made plain in Chapter 1, even single loudspeakers are not simple devices. When we use them in multiples, the way that they behave is complex, and is anything *but* obvious to the majority of users. The acoustic summation of two loudspeakers, except on an extremely thin plane which divides them in an anechoic environment, does *not* sum like an electrical mix of the two signals. Few people using loudspeakers, even professionally, seem to realise that pan-pot laws are related to room acoustics.

To recapitulate, loudspeakers in normal rooms tend to sum power (double power = +3 dB), whereas pressures sum like voltages (double voltage = +6 dB),but the true pressure sum only exists on the central plane in an anechoic room. As 20 kHz has a wavelength of only 1.7 centimetres, and because only within a quarter of a wavelength can true summing be expected, then the full frequency range of the +6 dB summation plane is only around 4 millimetres wide.

7.6 Stereo perception in rooms

The man who conceived and patented the two channel stereo concept was Alan Dower Blumlein, working for the EMI company in England in the mid 1930s. In his patent he referred to a listener in a stereo *seat*. That seat was to form an equilateral triangle with two loudspeakers, which therefore subtended an angle of 60 degrees at the listening position. There is nothing about a pair of loudspeakers that creates stereo. The image that we perceive of a sound stage laid out before us is an illusion created within our brains. This is why we refer to *phantom* images, because there is no sound actually emanating from the directions from which we hear the images arriving unless they are coming from the extreme sides of the sound stage, i.e. out of one loudspeaker only.

Many loudspeakers are said to have good stereo imaging, but in reality it is not the loudspeakers themselves which have good stereo imaging. Good stereo imaging is perceived from loudspeakers that can supply the appropriate information to the ears which the brain can process as a phantom sound stage, but between the loudspeakers and the ears the signal has to cross the room, and rooms can do a good job of scrambling the information. The positioning of the loudspeakers in many studios in confined spaces, or where horizontal reflexions are likely, is guaranteed to diminish the stereo imaging perception from *any* pair of loudspeakers. Loudspeakers of only moderate imaging can easily out-perform potentially better loudspeakers if they are better located. There is nothing inherent in a loudspeaker's performance which give it the ability to create precise stereo imaging sensations independently of its mounting conditions.

In the open space of a domestic lounge, with large expanses of reflective walls, a loudspeaker with a wide, relatively flat off-axis response may tend to give the best stereo imaging. On the other hand, in the clutter of a cramped control room, a narrower directivity with a rolled-off bass power response, but still reasonably flat on-axis, may give the best stereo. It is just simply impossible to say which loudspeaker responses give which best results without knowing the acoustic conditions in which they will be used. Many commercial loudspeakers are designed to give the best results in the majority of the rooms in which their designers have expected that they will be used. Nonetheless, no matter how successful the design may be in the majority of cases, it still does not mean that that design is *generally* better than a design which takes a different approach. The positions of nearby reflective surfaces may strongly influence the choice of loudspeakers for a specific purpose.

Any fast reflexions of less than about 1 millisecond delay with respect to the direct sound (less than about 30 cm path length difference) will pull the image in the direction of the reflexion. Reflexions with more than one millisecond of delay will not do this, as the ear will lock the direction to the first arriving wavefront (known as the Haas effect, the precedence effect, or the law of the first wavefront) but colouration will result from the way in which the signals re-combine with different phase relationships, and transient signals will be smeared. If these things occur asymmetrically, say with a reflecting surface close to one loudspeaker of the stereo pair and not to the other, then the stereo imaging will surely suffer.

Floor reflexions are almost an inevitability, but as these will always arrive at the ear from the same *horizontal* angle as the direct sound, their effect on the stereo imaging is less noticeable. What is more, the ear is much less sensitive in the vertical plane than it is in the horizontal plane, presumably because humans and their ancestors have not had to worry about either predators or prey from either under ground or in the air. Ceiling reflexions tend to behave similarly if the loudspeakers are mounted high up. Wall reflexions *can* contribute to the spacial stability of stereo imaging, especially if the loudspeakers have a wide and flat off-axis response. It seems that the ear can detect reflexion patterns which are uniquely related to certain source positions, but it must be remembered that a *phantom* image has more than one source. The tendency is for stereo images to be more precise in highly absorbent rooms, but more spacially stable in rooms with some lateral reflexions. That is to say, the image in the stereo 'hot seat' will be better in absorbent rooms, but the images will tend to collapse towards one loudspeaker or the other as one moves off centre. Conversely, in more reflective environments, a wider area of stereo perception may be available, but nowhere will it be perceived with the precision afforded by the absorbent rooms. This situation has led to some varied approaches to the design of critical listening rooms, with different designers having different priorities.

7.7 Rooms for critical listening

The rooms in which musical instruments and their amplifiers will be used are usually designed to have some acoustic life which will enrich the sounds and help to inspire good performances. Stereo imaging is not a relevant concept in the design or use of such spaces because the images are almost invariably real, and not phantom. In the control rooms and listening rooms, on the other hand, stereo imaging is a great priority. There is a generally accepted philosophy in the design of such rooms that no reflexions should return to the listening area within 15 milliseconds of the direct sound or with a level above 15 dB below the level of the direct sound. This has led to the development of room geometries which deflect early reflexions away from the listeners, and only allow reflected energy to return via diffusely reflected surfaces. Philosophies such as the Live-End, Dead-End rooms use such principles[2]. Other approaches, such as the Non-Environment rooms, seek to maximally absorb all but the floor reflexions (although in *some* rooms, these too are absorbed), and provide life for the speech and actions of people working within the room by means of a highly reflective front wall[3]. As in the case with the hemi-anechoic chambers, if the source is set into the hard surface, it can only radiate away from it. If there are no reflective surfaces in the room to return the sound waves *to* the front wall, then no reflexions can bounce off the front wall in any way that could create either tonal colouration or image smearing. Readers wishing to study more about control room designs should refer to the Bibliography at the end of this chapter. Figure 12.19 also shows some design concepts for stereo listening rooms.

Room acoustics is a big and complex subject, but even when the room is 'right', and even when the loudspeakers exhibit exemplary performances, it only requires the introduction of furniture and equipment into the listening room, arranged in inappropriate places, to severely affect the perceived, overall response. In recording studio control rooms, the equipment needs to be readily accessible for practical reasons, but the positioning of the equipment needs careful thought. In domestic circumstances it almost goes without saying that no reasonable hi-fi enthusiast would put the dining room table and a few cupboards between themselves and the loudspeakers before listening, yet people, many times out of necessity, do very similar things in control rooms with the mixing consoles and equipment racks. In general, loudspeakers should be mounted in positions which are as unobstructed as possible, and careful thought should be given to the siting of nearby video monitors if sound-colouring reflexions are to be avoided.

Mounting loudspeakers high up is also not a recommended procedure, and Figure 7.16 shows how the tendency is then to listen from above one's head, which will not produce the same audible sensations as listening with the sound in front of one's nose. The high mounting in studio control rooms also tends to produce strong reflexions from the upper surface of the mixing console, which introduces time-smearing of transient sounds and colouration of more steady sounds. And whilst on the subject of mounting, Figure 7.17 shows how multi-driver, non-coaxial loudspeakers should be mounted. Many loudspeaker manufacturers now show such drawings in their instruction leaflets. The drivers should be kept in the same vertical line so that lateral movement of the listeners will not give rise to arrival time differences from the separate drivers in a cabinet; hence the stereo imaging and timbral colouration do not suffer so much. This point is easily demonstrated by feeding a pink noise signal to a loudspeaker and moving one's head laterally from side to side. The tonal change with the loudspeaker drivers displaced horizontally will be much more noticeable than with the drivers mounted vertically. Conversely, the listener could remain

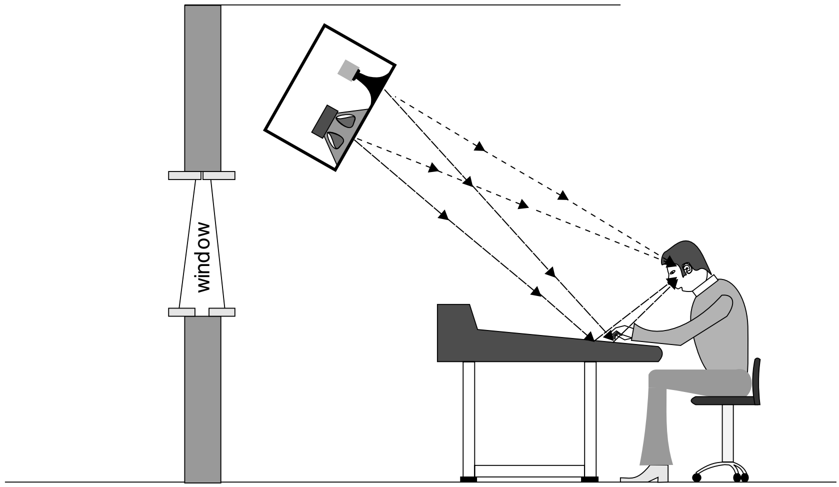


Figure 7.16 When monitor loudspeakers are mounted at a steep angle, the high frequencies, in particular, arrive at the ears from an angle totally inappropriate for the perception of an accurate frequency balance. High frequencies will tend to be under-perceived when a listener is looking at the equalisation controls of the mixing console. What is more, unless the ceiling is highly absorbent, the low frequencies will suffer augmentation due to the proximity of multiple room boundaries. The off-axis radiation, shown by the continues lines, will also tend to reflect from the top surface of the console and smear the transient response

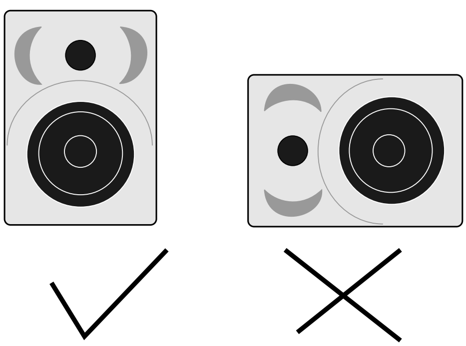


Figure 7.17 Orientation of loudspeakers.

It is preferable to mount loudspeakers with the drive units aligned vertically. When the listener moves to the left or right, the relative distances to the drive units will not change. If the loudspeakers are mounted horizontally, sideways movements will change the relative distance to the high and low frequency drivers, and give rise to phase shifts which will affect the perceived response, especially in the region of the crossover frequency

in the same place whilst an associate swivelled the loudspeaker; the effect would be substantially the same. Mounting reflective surfaces behind the listener as shown in Figure 7.18 should also be avoided if colouration is to be minimised.

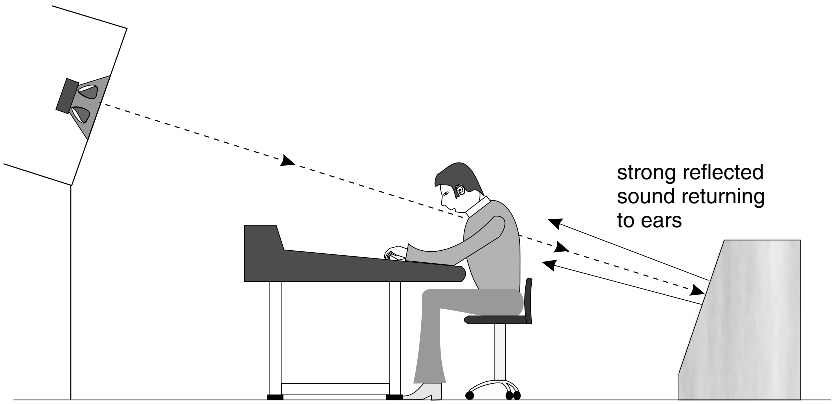


Figure 7.18 In many cases, an effects rack placed behind the engineer for easy access also performs the function of an acoustic mirror, returning strong reflexions to the primary listening positions. This is especially problematical when the main monitor loudspeakers are mounted high up

7.8 Electronic, digitally adaptive response correction

Electrical response-equalisation has long been a feature of loudspeaker design. The earliest forms were just simple potentiometers to adjust the tweeter levels to better subjectively suit their surroundings. Modern monitor loudspeakers with built-in amplifiers offer much more flexibility. Figure 7.19 shows the block diagram and the response flexibility of a Genelec 1030A monitor loudspeaker, designed to provide limited compensation for mounting conditions and boundary proximity. If the loudspeaker were to be placed close to a wall, the radiating space would be reduced, as described in Section 7.2, so the ensuing bass boost would be compensated for by suitable adjustment of the bass tilt and roll-off controls. As previously stated, such loading changes give rise to response changes which are of a largely minimum phase nature (which we shall look at in more detail in the following section), but many response variations which involve signal delays, such as reflexions, resonances and group delays due to filters and loudspeaker driver physical displacements, give rise to *non*-minimum phase responses. Such responses can often only be corrected by the use of acausal filters. ‘Acausal’ means effect before cause, and so these problems can only be dealt with by the insertion of signal delays which allow a filter to act on the signal before being incorporated into the output. There are no analogue solutions to such problems, so inevitably these methods reside in the world of digital signal processing.

The combined response of a loudspeaker and a conventional room is extremely complex. In fact it is absolutely, absurdly complex; far beyond what common sense would suggest. In the 1980s there was great excitement in some circles about the future of digital response correction, and it being the end of the line for room acousticians, but these expectations have not come to pass. It was widely believed that, despite the size of the problem, signal processing technology would develop apace, which indeed it has

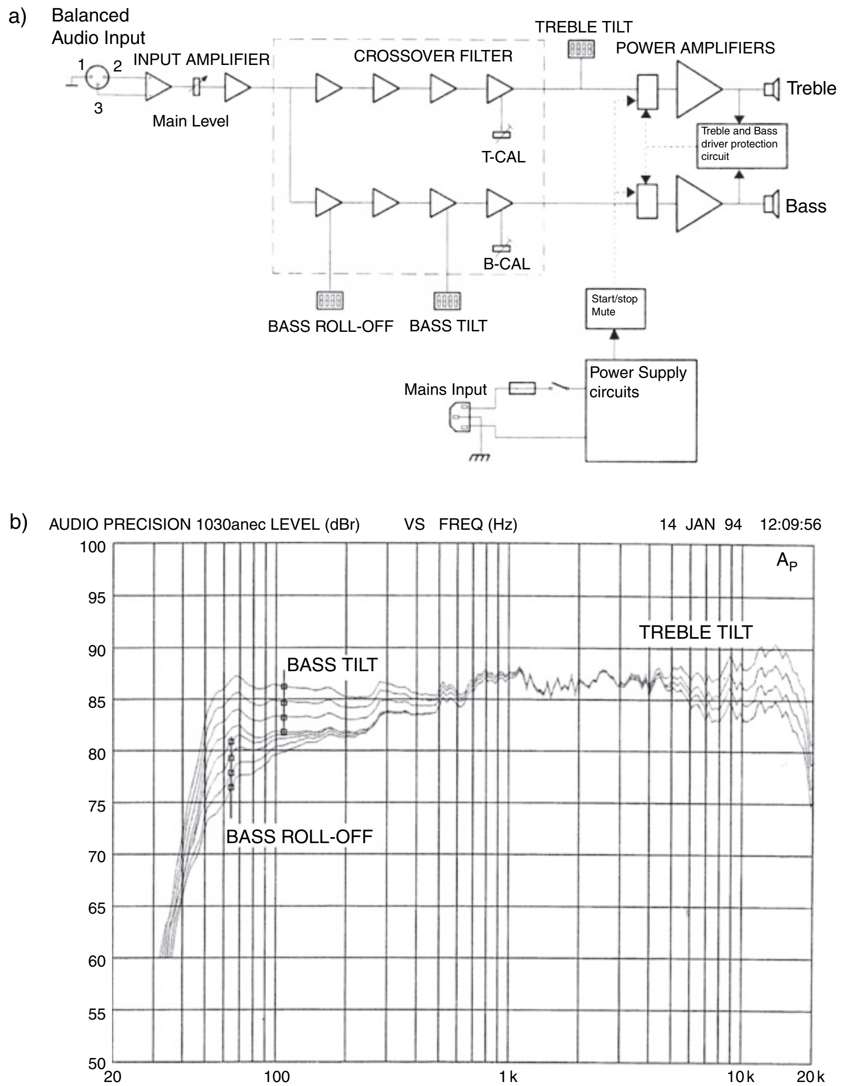


Figure 7.19 Controls and response flexibility of a Genelec 1030A loudspeaker. a) Block diagram showing active crossover filters, power amplifiers and drive units. b) The curves show the effect of the 'bass tilt', 'treble tilt' and 'bass roll-off' controls on the free-field response

done, but complication after complication have arisen to confound many of the hopes. In fact, some excellent room/loudspeaker correction can indeed be achieved with modern technology, but there are prices to be paid for many of the improvements. As will be described further in Chapter 12, if one considers responses below about 100 Hz, and restricts oneself to dealing with largely minimum-phase problems, then signal processing can be put to some very good use. Loudspeaker manufacturers such as JBL and Genelec have adopted the philosophy that whilst a great many people *will* make decisions about professional sound recordings in less than ideal surroundings, largely due to the ever greater financial pressures, then active signal processing can make beneficial contributions to such environments. Using relatively inexpensive technology they can make improvements to conditions where neither space nor money allows for acoustic treatment. However, such responsible manufacturers also acknowledge that there is no real substitute for good room acoustics if the very highest levels of reproduction quality are required. In *very* good rooms, it can actually be the case that signal processing can introduce more problems than it solves, so one should always be aware of this.

It has also been shown that many systems of loudspeaker/room correction are only significantly beneficial at short distances[4]. Figure 7.20 shows the modulation transfer function (MTF) plots of three loudspeakers in a relatively neutral room at distances of one metre and four metres. The MTF scale (vertical) is really a measure of response accuracy in terms of information content; '1' being perfect reproduction and '0' being total inaccuracy. As can be seen from the one metre and four metre plots, the response accuracy generally tends to fall off with distance. (In an anechoic chamber the plots would be identical at both distances). Figure 7.21 shows the same loudspeaker responses after digital response equalisation. It is clear to see that the responses at one metre have been significantly improved, but no such improvement is evident at the four metre distance.

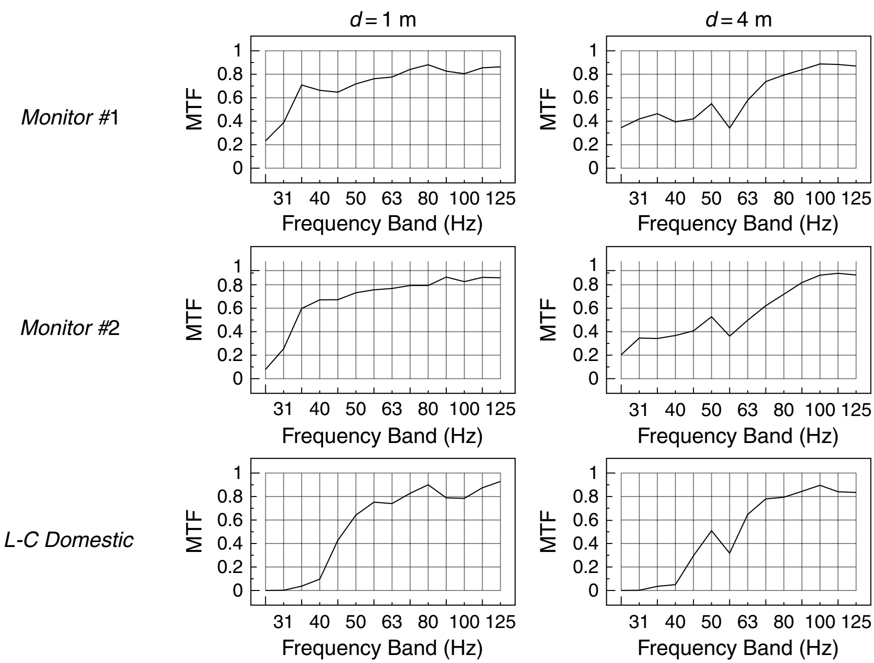


Figure 7.20 MTFs of three loudspeakers in a reasonably well-damped studio recording room at different distances (d) from the loudspeakers

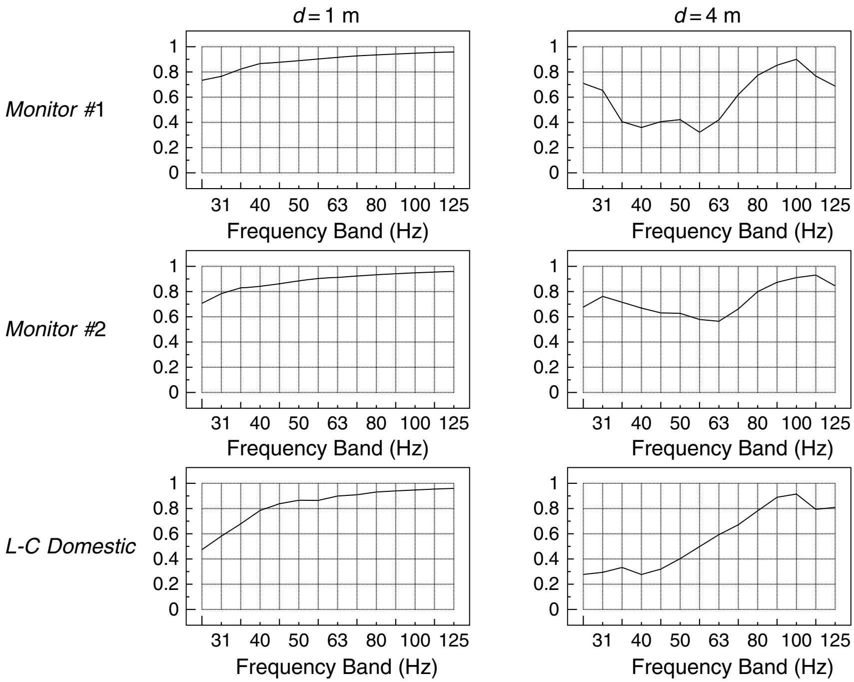


Figure 7.21 MTFs of the loudspeakers in a reasonably well-damped studio recording room at different distances (d) from the loudspeakers after equalisation

The implication is that where the loudspeaker response dominates the overall response, in the close-field, the digital correction can be greatly beneficial, but in the far-field, where the complex room responses tend to dominate the overall response, the correction processes lose control over the response. It thus becomes evident that whilst such equalisation may be effective in the close-field monitoring situations of a poorly treated post-production room, it is no answer when a spacially uniform response is required in a large sound control room, where the non-minimum-phase room responses dominate, *unless* the room is well acoustically controlled. The MTFs are discussed in further detail in Chapter 9.

In 2004, Norcross, et al, published an overview of the situation in an Audio Engineering Society paper[5]. In the abstract to the paper, they stated "When the [response] is *non*-minimum phase, the artefacts [of the correction process] tend to become more severe and become distinctly audible. The artefacts produced by the inverse-filtering process can actually degrade the overall signal quality rather than improve it." They recognised that without doubt, in many cases, the inverse filtering of a room/loudspeaker combination could improve perceived responses, but they were warning of the necessity to carry out very careful subjective listening tests before committing to the use of any signal processing system for professional purposes. Also, as discussed in Chapter 5, whenever digital signal processing is used in a way that needs immediate re-conversion to analogue, the converter quality is also critical if it is not to limit the resolution of a system. Digital processing is better employed within a digital signal chain and then passing the whole chain through only one re-conversion process. Once again though, when we are faced with a chain of processors, converters, loudspeakers and rooms, if only one link limits the resolution then other deficiencies in the chain may well not be noticed. Low resolution loudspeakers may not show the differences between good and very good processors or converters, and the decisions which are made on low resolution monitoring loudspeakers during the recording/mixing process may negatively affect the quality of a recording by not making evident the benefits of superior equipment.

The responses of most loudspeakers in rooms have a mix of minimum and non-minimum phase components, and the mathematically most correct inverse response may not always yield the most subjectively improved response. One loudspeaker may also prove to be more subjectively correctable than another, even where no obvious evidence for this can be seen from their uncorrected responses. Another factor which must be taken into consideration when applying complex inverse filtering is that the correction may be evident only in one small region of a room, and that in most other places the response may be significantly worse than before correction. A further related problem is where correction is used to flatten the axial response of a loudspeaker with an irregular off-axis response. As it is the same driver radiating the signal *and* its correction, it obviously cannot change the axial response by itself without also affecting the off-axis response. The reflexions returning from nearby surfaces with the modified off-axis response may be subjectively undesirable. It is thus very easy to publish the improved axial results, perfectly legitimately, but which do not reflect the true sonic situation in normal use. The significance of the off-axis response could also be very room dependent, in which case so could be the overall effects of the correction.

Furthermore, inverse filtering can produce pre-echoes, which can give rise to audible artefacts. These are inevitable by-products of certain types of inverse, acausal filtering that cannot be entirely eliminated. Sometimes, the processing artefacts can be as low as 60 dB below the signal, yet listening tests have shown them to be audible[5]. Every year, progress is made in the domain of digital signal processing. In live sound technology, the adoption of digital processing has been widespread, rapid and deep. Great strides forward have been made in terms of intelligibility and subjective sound quality. Obviously though, in live performance venues, audience noise, ventilation noise, and even the unavoidable air-related distortions resulting from the very high sound pressure levels close to the loudspeakers tend to mask any low level processing artefacts, so the benefits can clearly outweigh the penalties. Conversely, in recording studios, where the perception of small details is much greater than during live concerts, and where no visual performance is distracting the brain from concentrating on the sound, there have only been limited applications for the digital signal processing of monitor systems because of the great sensitivity to low level artefacts resulting from the correction algorithms. However, when people work on music recordings in rooms where computer discs and ventilation fans are whirring away, producing up to 40 or even 50 dBA at the principal listening positions, many subtleties of the recorded sound will also inevitably not be heard. It is imperative if high quality audio monitoring is to be achieved that such mechanical noises be banished from control rooms. There are no technical reasons why the noisy equipment cannot be located outside of rooms where high resolution loudspeaker monitoring is required. The introduction of so much noisy equipment into the lower level of studios has been a very retrograde step.

7.9 Minimum and non-minimum phase responses

In the previous section we mentioned non-minimum phase responses and acausal filters. A minimum phase system is one in which the phase shift associated with the amplitude response is the minimum that can be allowed whilst still exhibiting the properties of a causal system. A causal system is one in which the output *never* arrives before the input. In minimum phase systems there is a very strict relationship between amplitude and phase, and correcting either one will always tend to correct the other. The response boost at low frequencies given rise to by flush mounting a loudspeaker in a wall is an example of a minimum phase response change, which can be equalised with normal analogue filters to restore the free-field axial response in terms of both frequency response (amplitude and phase) and time response. The essential factor in a minimum phase system is that there is no appreciable delay between the generation of the signal and the effect of whatever is influencing it. If there is no appreciable delay, then there can be no appreciable phase shift, hence only *minimal phase shifts* will be evident. This is the origin of the term 'minimum phase'.

In the case of non-minimum phase responses, amplitude correction, alone, cannot correct the phase responses. The Fourier transform is a mathematical means of linking the time domain representation of a signal to its frequency domain representations of amplitude and phase. The application of the Fourier transform to a signal waveform (time response) reveals the frequency components in terms of their magnitude and relative phase (i.e., the 'spectrum'). The application of the inverse Fourier transform to the spectrum yields the original waveform. This unbreakable connection between the amplitude and phase on one hand, and the time response on the other, means that if the correction of a response in terms of its amplitude, alone, cannot correct the phase response, then the time response will not be correct. Transient sounds can be very dependent upon their waveforms in terms of their sonic characteristics, so non-minimum phase systems tend to have distorted time responses.

The far-field response of a loudspeaker system in a reflective room (but *not* in an anechoic chamber) is an example of a non-minimum phase effect. Here, there is a delay between the signal generation by the loudspeaker and the addition of the boundary reflexions to the composite signal at the listening position. The arrival time differences of the reflected waves give rise to phase irregularities which are frequency and distance dependent, so no simple manipulation of the amplitude response of the source can adequately compensate for the complex disturbances. This is why one-third octave-band equalisation of loudspeaker systems in rooms is only, at best, a very rough approximation of the application of the true inverse of the response, and also why many equalised systems sound no better, or even worse, than the unequalised responses. The equalisers may only, in effect, be moving the response bumps and dips around, and may actually be worsening the transient responses as they try to correct the non-minimum phase amplitude responses.

Another example of a non-minimum phase response is in the combination of the various outputs from crossovers, as discussed in Chapter 5. In any filter circuit, either mechanical or electrical, there are inherent group delays for any signals passing through them. The amount of group delay increases as the filter slope increases and as the frequency lowers. A crossover will consequently have a different group delay associated with each section. When the outputs are recombined they will therefore *not* produce an exact replica of the input signal. For this reason, conventional equalisation cannot be used to correct response errors at crossover points, and physical differences in the driver mounting positions may give rise to further non-minimum phase responses of the same general nature. Amplitude correction of the response irregularities given rise to by either the mechanical or electrical misalignments will lead to further phase distortions and hence further time response errors which may be noticeable on transient signals. The degree of deviation of a response from the minimum-phase response is known as the excess phase. Whenever time-shifted signals are mixed, the tendency is for the excess phase to build up.

Adaptive digital signal processing can deal with these problems, but it can realistically only be achieved to a very high degree at one point in space, or to a lesser degree over a wider area. In all cases, for every part of a room that benefits, another part of the room will suffer a deterioration in the response. In effect, the correction systems are redirecting the acoustic waves. It is a little like having a room with some finite layers of sugar cubes on the floor. The more that one builds up a pile in one place, the level must go down elsewhere. Very high quality control rooms and monitor systems tend to be expensive because there is really no substitute for high quality drive units in big boxes working in heavily acoustically treated rooms if flat, low distortion, high level, wide frequency band, spacially uniform monitoring is required. The problems must be eliminated at their sources, because electronic correction systems all have their compromises and drawbacks. However as mentioned in the previous section, if people will persist in trying to do professional recording in poorly controlled rooms on inexpensive loudspeaker systems, then digital correction may offer some overall performance benefits at affordable prices. Its application to sub-woofer processing will be discussed in Chapter 12.

In general, non-minimum phase response irregularities are difficult to deal with, and so are best avoided by the use of mechanical and acoustical means. The irony is that digital correction is best applied to rooms and components that are so good that they barely need correcting. For example, a related technology is the motional feedback of loudspeakers, where a sensor is placed on the woofer cone and is used to detect the actual cone movement. In Chapter 1, Section 1.5, it was noted how the entire surface of a diaphragm does not always move in unison. Consequently, a sensor on a cone only senses the movement at that part of the cone where it is placed. On transient signals, delayed, non-minimum phase responses can occur which risk wild instability in the system loop, which can be controlled by careful filtering, but the amplifiers tend to need to be much bigger than would normally be necessary in order to handle the superimposed audio signals and correction signals. The costs soon spiral upwards to a point where in many cases it would perhaps be better all-round just to build a better quality loudspeaker system in the first place.

No matter whether digital correction is being applied to the loudspeaker alone, or the loudspeaker/room combination, the need for this extra headroom always needs to be taken into account. This is one reason why large loudspeaker systems are rarely processed, because the amplifiers may need to be unreasonably large; and anyhow, large systems can be made with good transient responses and flat, low distortion frequency responses without the aid of signal processing. Given that the idea of motional feedback has been around since the 1930s [6], the fact that it is still such a rarity suggests that it is not an easily realisable solution for electro-mechanico-acoustic transducer inadequacies. One of the greatest successes in active control has been in extending the lower frequency responses in relatively small boxes. Nevertheless, by whatever means that it is achieved, a given SPL requires a given volume of air to be moved at a given rate, so small drivers, no matter how they are processed, can still only achieve low SPLs at low frequencies.

References

1 Tyndall, J., 'On Sound', Sixth Edition, p 13, Longmans Green and Co., London, UK (1895)
2 Davis, Don and Davis, Chips., 'The LEDE Concept for the Control of Acoustic and Psychoacoustic Parameters in Recording Control Rooms', Journal of the Audio Engineering Society, Vol 28, No 9, pp 585–595, (September 1980)
3 Newell, P. R., Holland, K. R., 'A Proposal for a More Perceptually Uniform Control Room for Stereophonic Music Recording Studios', 103rd AES Convention, Preprint No 4580, New York, USA (1997)
4 Holland, K. R., Newell, P. R., Castro, S. V., Fazenda, B., 'Excess Phase Effects and Modulation Transfer Function Degradation in Relation to Loudspeakers and Rooms Intended for the Quality Control Monitoring of Music', Proceedings of the Institute of Acoustics, Vol 27, Part 8, Reproduced Sound 21 conference, Oxford, UK (2005)
5 Norcross, S. G., Soulodre, G. A., and Lavoie, M. C., 'Subjective Investigations of Inverse Filtering', Journal of the Audio Engineering Society, Vol 52, No 10, pp 1003–1028, (October 2004)
6 Colloms, M., 'High Performance Loudspeakers', 6th Edition, John Wiley & Sons, Chichester, UK (2005)

Bibliography

1 Newell, Philip., 'Recording Studio Design', Focal Press, Oxford, UK (2003)
2 Cooper, Jeff., 'Building a Recording Studio', Fourth Edition, Synergy Group Inc, Los Angeles, USA, (1984)

3 Davis, Don; Davis, Carolyn, ‘Sound System Engineering’, Second Edition, Focal Press, Oxford, UK (1997). NB: originally published by Howard W. Sams, USA (1987)
4 Walker, Robert., ‘A New Approach to the Design of Control Room Acoustics for Stereophony’, 94th AES convention, Berlin (1993)
5 Walker, Robert, ‘The Control of Early Reflections in Studio Control Rooms’, Proceedings of the Institute of Acoustics, Vol 16, Part 4, pp 299–311, UK (1994)
6 Walker, Robert, ‘A Controlled-Reflection Listening Room for Multichannel Sound’, Acoustics Bulletin (Journal of the UK Institute of Acoustics), Vol 24, No 2, pp 13–19, St Albans, UK (March-April 1999)
7 Newell, Philip, ‘Project Studios’, Focal Press, Oxford, UK (2000)

Chapter 8

Form follows function

8.1 The chain

The recording chain under discussion here begins with the musicians and ends in the rooms in which the people who buy the recordings choose to place their music systems. Domestically, we must really limit this discussion to reasonably high fidelity systems, because once we introduce in-car listening and ghetto-blasters on the kitchen table, or portable radios in the bathroom, we begin to enter a realm of variability which, firstly, become unable to be qualified and secondly, in the majority of cases, cannot really be considered capable of truly representing the music producers' wishes. Manufacturers of such equipment may go to great lengths to produce pleasant-sounding equipment, and record producers may go to equally great lengths to ensure the compatibility of their mixes with such systems because they represent the majority of the market for recorded music, but this book is essentially dealing with the concept of high fidelity reproduction. This is not to say that many in-car systems are not capable of remarkably high fidelity in many aspects of their performance, but their reproduction quality is still idiosyncratic in a way that generally sets it apart from what we expect to hear from a 'high-end' system in the home.

The loudspeakers used in the recording chain can be separated into five basic groups:

1) loudspeakers for musical instrument amplification
2) recording monitors
3) mixing monitors
4) mastering monitors
5) domestic high fidelity loudspeakers

There are many people who will argue against this concept, citing that a good, professional, well conceived, well-engineered loudspeaker should be suitable for all of the above purposes. However, the recording chain is a very varied chain, and what will be discussed in this chapter are the specific requirements at each stage of the process which tilt certain designs or concepts towards being advantageous for the different needs of those specific requirements. In fact, the reality of the current situation is that different loudspeakers do tend to be used in different stages of the recording/mixing/ mastering process where financial restraints do not limit the choice of equipment, and there are many good reasons why this state of affairs exists.

In their own ways, the last four of the loudspeakers on the above list all try to achieve the closest approach to the original sound. They are all reproducers, trying to emulate as accurately as possible within their design criteria a faithful acoustical output of the electrical waveform being fed to them. All of them will generally be required to show a wide, flat frequency response, a well-damped time response, low levels of non-linear distortion and a directivity pattern appropriate to the rooms in which they are each expected to be used. The balance of priorities will vary with the intended applications, as we will discuss later, but the above characteristics are common to all of them.

Conversely, loudspeakers for use with musical instruments are part of the instruments which they are amplifying. They are sound *producers*, not reproducers, and what they produce is, in itself, definitive. They are unlikely to have flat frequency responses, and the range of those responses may well be defined by the harmonic range of the instruments with which they will be used. Colouration of the sound is usually a desirable asset – something which is anathema to hi-fi or monitoring – and 'musical' forms of distortion may also be considered to be beneficial. Time responses which contain resonances can impart warmth and character to the musical timbre, and directivity control may be something which is given very little consideration whatsoever. In fact, the design of loudspeakers for the production or reproduction of music have very little in common. Music reproduced via musical instrument loudspeakers may be very far from what the record producers intended, and electric instruments played via high fidelity loudspeakers may tend to sound lifeless and uninspiring.

It may be more appropriate to begin this chapter by looking at the *reproducers* before discussing the producers, because the idiosyncrasies of the latter will be better understood after the rigours of the *reproduction* of music have been better appreciated. Nevertheless, there is no hierarchy of superiority in the concepts, because unless an interesting sound is there to be recorded, there cannot be much enjoyment from reproducing it. Indeed, the factories which produce the better musical instrument loudspeakers spend just as much time and attention on the design, manufacture and quality control of the instrument loudspeakers as they do on the best monitoring loudspeakers. In some ways, designing for flatness and low distortion can be easier than trying to design something to make that elusive, 'magic' sound.

8.2 Recording monitors

Until the early 1970s, recording monitors, mixing monitors and mastering monitors – in those days, mastering being essentially disc cutting – were largely one and the same thing. It was not uncommon for the same type of loudspeakers to be used throughout the principal rooms of a recording studio complex, although in the disc cutting rooms, which were usually smaller in size than the recording control rooms, the sound was rarely the same as in the usually better acoustics of the control rooms and mixing rooms. In those days, also, the musicians tended to stay in the recording rooms, and only ventured into the control rooms when invited to hear a 'playback'.

As time progressed, and the musicians began to become more involved in the whole process, they began to spend much more time in the control rooms, and they began to expect to feel the same sensations as they ventured from the performing studios to the control rooms, in order not to lose the 'vibe'. Volume levels in the control rooms began to rise in order to avoid the deflationary sensation of playing in the studio at 100 dB SPL then listening to the 'take' in the control room at 85 dB SPL. If the levels changed, the perception changed, and with it the 'buzz' of excitement could change, leaving the musicians in doubt about whether they had achieved their aims, or not. Within a very few years, and especially after the advent of synthesisers and other portable keyboard instruments, it began to become commonplace to actually perform in the control room.

It can be argued that if what is perceived at 100 dB will not translate to 85 dB or less, then there is an inconsistency that suggests that the recording may disappoint when reproduced domestically. However, it should be well understood that, at least for multitrack recordings, the *recording* phase is about capturing a *performance*. The experience of the recording staff will be important in deciding if the *sounds* can be optimised at lower levels, but the achievement of the maximum impact of the *performance* can only effectively be captured during the recordings.

Around 100 dB SPL, music begins to affect perception and emotions in a different way to what we perceive at lower levels. Chemical changes take place in the brain, which are not unlike those caused by sex and certain drugs. This explains why music at discothèques needs to be loud, or the sensation to dance will not be stimulated and the tendency towards exhibitionism will not be aroused. The exhibitionist tendency is also important in the recording process, because musicians are performers, and a performance is an exhibition. Therefore, if musicians are to perform at their best, they may well need a stimulus, very similar to the disco dancers needing a stimulus. Obviously, the type of stimulus depends on musical style, but under almost all circumstances, the correct stimulus will not be achieved unless the musicians are performing in a control room at reasonably similar levels of sound pressure to those which they would normally be receiving during a concert performance.

Perhaps coincidentally, the tendency to perform in control rooms began around the same time that some 'mini-PA systems' were beginning to appear as studio monitors, which could easily produce over 120 dB SPL at the mixing console. In all fairness, much more is known about psychoacoustics in 2005 than was known in 1975. However, in those early days, the mini-PA system with a pair of high power 15 inch drivers, a compression driver straight from sound reinforcement technology on a barely modified horn, and a high power compression tweeter seemed to be a reasonable means to achieve the sort of high SPL, wide bandwidth, relatively low distortion sound that was being called for. Two such systems are shown in Figure 8.1, and even at low SPLs the performance of these systems was better than many of the previously used monitors, at least for many types of music, but they did have many design faults by modern standards. Time has refined these concepts, and room acoustics have taken great steps

a)

b)

Figure 8.1 Large studio systems from the mid-1980s. a) Giant Eastlake Audio systems at Marcus Music, London, UK. b) Urei 815s at Jacobs 'Court' studio, Farnham, UK

forward, guided by a much greater understanding of psychoacoustics, but in the conditions of the mid 1970s the failings of the loudspeaker/room combinations in many cases led to problems when using these main monitors for mixing, which in turn led to the use of small reference loudspeakers such as the Auratones, and later the Yamaha NS10s, to name two popular examples. The concept of separate recording and mixing monitors had thus begun to establish itself, and it began to become clear that each had their place in the music recording process. Nonetheless, whilst it is by no means *obligatory* to use different monitors at the recording and mixing stages, it tends to be rather expensive to achieve conditions with a single system which are optimal for both purposes. Cost-cutting has had a big impact on the limitation of monitoring acoustics to something which is often well below what is achievable.

8.2.1 Basic requirements

Recording monitors need to do two jobs at the same time. The musicians may be considering them to be stage monitors – performance monitors – whilst the recording engineers will also be using them to assess recording quality and instrumental timbre. They therefore need to simultaneously achieve high SPLs and great subtlety. The musicians will also be expecting them to go as deep as their lowest bass instrument would achieve in a live performance, or their sense of the performance may be diminished. The musicians may also be distributed around the room, and each of them will be expecting to receive their fair share of the sound. The latter fact tends to mean that the loudspeakers need to insonify the entire room to a reasonably equal degree. If this is to be achieved in a way in which the recording engineer can still hear all the necessary detail in the sound, the room will need to be extremely well controlled acoustically, and the loudspeakers would need to be flush-mounted in the wall. Not surprisingly, these are precisely the conditions which are to be found in most of the control rooms of the world's most famous studios. Some rooms complying with these requirements are shown in Figure 8.2.

A typical frequency response requirement is shown in Figure 8.3. The low frequency response needs to go to around an octave below the lowest note on a conventional bass guitar, at 41 Hz. The extra octave is needed both to minimise colouration due to phase shifts associated with the roll-off, which can extend well above the roll-off turnover frequency, and also to accommodate instruments such as the less common five-string bass guitars, and the sub-bass from synthesisers. The perceived colouration due to the roll-off at around 20 Hz is considerably less in this lowest octave of the audio frequency range.

The roll-off at the high frequency end of the spectrum is something which has developed over time. In the early to mid 1970s, when the recording monitors were also the mixing monitors, it was customary to use multi-band equalisers to achieve a flat response up to 20 kHz. However, it soon became apparent that this was leading to dull mixes in the homes of the record buyers. There were several schools of though about why this should be so. One idea was that the higher monitoring levels in the studios meant that the mixes were being done at levels where the ear was more sensitive to high treble than would be perceived at the levels of typical domestic reproduction. Consequently, what would seem to be a balanced frequency response in the studio would seem top-light in the home. Another reason frequently discussed was that the 'bass trapping' which was employed in many control rooms, to flatten the low frequency response of the rooms, was not a part of most domestic constructions. Therefore, the increased bass build-up in many domestic living rooms required a corresponding treble boost if a balanced low/high frequency relationship was to be achieved. Nevertheless, whatever the reasons actually were, the treble roll-off became a normal alignment for the large monitors, and it seemed to lead to better results. Nowadays, even if the large monitors are not used so frequently for mixing, the generally higher SPLs at which they *are* used seems to be less fatiguing with the roll-off of the high frequencies, and a more natural, representative balance is perceived. A rather similar process occurred in the film industry,

a)

b)

c)

Figure 8.2 a) Kinoshita monitor system at Capri Digital, Capri, Italy. b) Blackwing, London, UK, with its unusual Yamaha NS40 close-field monitors and a pair of 4-way amplified Reflexion Arts 235 monitors, mounted above the soffit. c) Eurosonic, Madrid, Spain

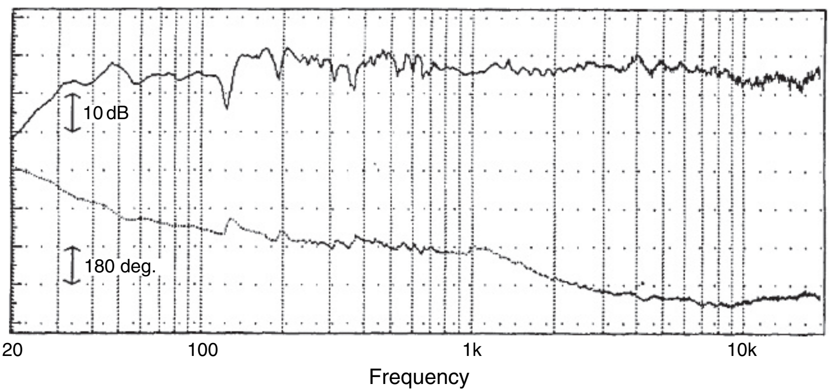


Figure 8.3 Two-way monitor measured on-axis at 2 metres distance in a control room at LMH

where an empirically derived 'X-curve', with a significant high frequency roll-off, became the industry standard, because it works!

By the same token, it would be reasonable to expect that the low frequencies should also be rolled-off. The curves of equal loudness are shown in Figure 8.4, and from them it can be seen that the ear becomes much more sensitive to the low frequency as the level increases. Nevertheless, as we have just discussed, in many domestic rooms the bass is much less controlled than in the most recording studio control rooms, so it would seem to be reasonable that the low frequency increase due to the listening

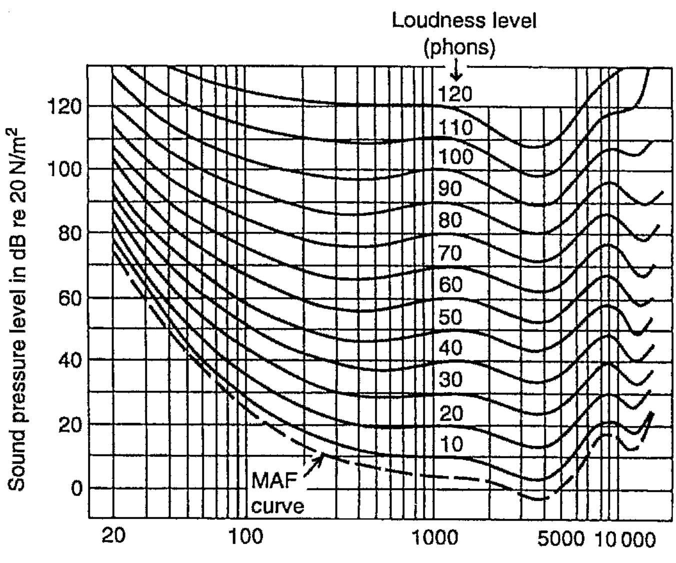


Figure 8.4 The Robinson-Dadson curves of equal loudness

level in the control rooms will very often be matched by the bass build-up in domestic rooms when listening at 10 to 20 dB less SPL. But whatever the reason, at the *recording* stage of the process, the performance is paramount. Who wants to hear a great recording of a poor performance? The frequency response target of Figure 8.3 certainly seems to work for recording monitors, and experience has shown that it also works well for mixing in the far-field.

Another important aspect of recording monitors is a fast decay time across the entire frequency band. A typical decay plot is shown in Figure 8.5. A fast decay is important because any resonant overhang will tend to lift the response at the resonant frequencies, just as a resonant mode in a room will cause a lump in the room response close to its antinode(s). The resonances can also mask detail in the sound, as will be discussed in more detail in Chapter 11. Unfortunately, many small monitor systems purposely employ resonances to extend their low frequency responses (see 'Reflex enclosures' in Chapter 3), but this technique is bound to blur detail and create the potential for misjudgements about the musical balance between instruments containing mixtures of low frequencies of both transient and steady-state nature, such as the combination of bass drums and

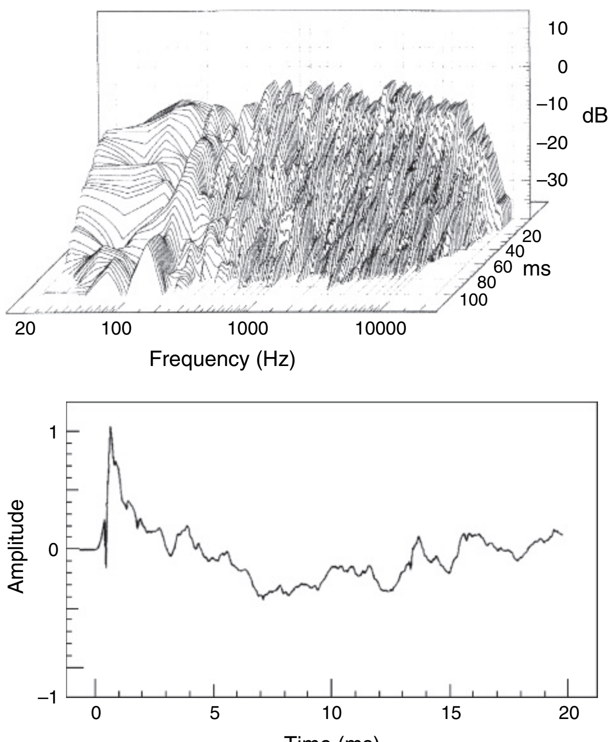


Figure 8.5 In-room decay response of a large monitor system in a well-controlled control room. The resonance evident in the waterfall plot at 120 Hz was the resonance of an empty cable tube in the floor

bass guitars. Their timbral balance will, to different degrees, be coloured by the resonance, so the recording personnel will be left unsure about which part of the sound is contributed by the recording, and which part is contributed by the loudspeakers and the room. It is therefore important to use physically large monitor loudspeakers because at peak levels of 110 dB SPL at reasonable listening distances, it is simply impossible to achieve fast decaying, low distortion, low frequency responses from small loudspeakers. There is no technology or trickery to overcome this problem because it is deeply rooted in the laws of physics.

Mixing is made much easier if the recording stage has been well monitored and controlled. There are far too many cases of problems being heard at the mastering stages due to the mixes having been done on small monitors of recordings which were also done via small loudspeakers, and thus which were not adequately monitored in the first instance. Commercial pressures, and the general decline and de-professionalisation of the recording industry in the late 20th, century has given rise to so many studios which use small monitors for the entire recording process. Many excellent recordings have been made in such studios, partly due to the adaptation of musical styles, such as by using instruments with pre-programmed, well-balanced sounds, and also by the degree of familiarity with the systems which has been developed by the recording personnel. Nevertheless, a whole industry of mastering has since grown to unforeseen levels, partly fed by the uncertainly which many people feel when working in conditions of inadequate monitoring when only small loudspeakers are available.

8.2.2 Proportional costs

Since the 1980s, the cost of multi-track recording systems has plummeted. When inflation is taken into account, the degree to which the prices have fallen is enormous. On the other hand, the price of good recording monitors and excellent acoustics has remained at similar, real, proportionate levels to what they were in the 1980s. That is to say, if a large, stereo monitor system and acoustic treatment cost the same as a new Mercedes car in 1985, then the cost of the same, top-line monitor system and acoustic treatment in 2005 would still cost the same as a new Mercedes car, whereas the price of a recording system has fallen from the price of three entire cars to the price of a replacement engine. Monitor loudspeakers and acoustic control systems are works of engineering, which require skilled labour and careful planning. They do not follow the trend in electronic development at the signal processing level, and micro-miniaturisation and mass production techniques cannot be applied. Even the power amplifiers which are used are physically large devices, which take considerable labour to construct. They also require expensive chassis and bulky components to handle the power levels involved and to dissipate the waste heat. In fact, as time progresses, skilled labour has tended to increase in proportional costs, but mechanisation does not easily lend itself to specialised, low quantity production processes, so high quality amplifier prices also remain high.

This disproportionality in the cost of the recording systems to the monitor systems has strongly militated against the purchase of large recording monitors. In cases where they *have* been bought, but used in rooms which were not suitably treated for financial reasons, they have often unjustly been criticised for being difficult to use. This however, is down to their inappropriate circumstances of use rather than the failings of the monitor systems, themselves, because such systems cannot be expected to work well in rooms which are not appropriately designed. Much of the criticism which modern, large monitors receive are based on this type of misapplication, although it must be said that some manufacturers have clouded the issue by attempting, for marketing reasons, to make their large monitors mimic the rounder sound of the wider-range smaller monitors. The real job of the large monitors is to do well some things that the small monitors *cannot* do. Their job is not simply to be a louder version of the small monitors.

8.2.3 Different approaches

Designing a large loudspeaker system to respond in a delicate and subtle manner whilst producing high SPLs is not an easy task. Unfortunately, as the solutions to some problems become easier as size goes up, the solutions to other problems become more difficult to achieve. The diversity in the design of many of the large systems which will now be discussed reflect the different orders of priority which their designers have given to the points where compromises must be made. A loudspeaker system *cannot* recreate the three-dimensional sound-field produced by an acoustic instrument, but ears are very sensitive to changes in sound-fields. Room acoustics also become more relevant as the loudspeaker to listener distance becomes greater, so the design of the loudspeakers and rooms becomes inextricably linked. The combined sound-field is what the listeners perceive, so if some compromises in loudspeaker system design can be mitigated in their effect by corresponding room acoustic changes, better overall results can be achieved.

The above statement also implies that a loudspeaker system which is designed for one room-acoustic concept may not be appropriate when used in acoustically different rooms, and vice versa. Therefore, to put *any* large monitor system into a relatively untreated room will be asking for problems – small systems at close distances tend to be a better solution. Nevertheless, a good large system in an appropriate room may be able to achieve a level of overall response accuracy which would simply be unattainable by *any* monitor system in a poorly treated room, or any small loudspeaker system in *any* room. However, as previously stated, the option of a large monitor system in an appropriate room is not likely to be a cheap solution to realise.

Figure 8.6 shows four large monitor systems of very high quality, but none of them can achieve their potential without a considerable amount of acoustic design in the control room. And by ‘acoustic design’ we are not referring to sticking some foam panels on the wall. Acoustic control systems which work at 20 or 30 Hz are necessarily large, and their size is dependent upon wavelength, and wavelength alone. They will *not* scale with room size or budget. This means that rooms with well controlled low frequencies need to be considerably larger than the working space which will be required after the room is finished. Two hundred cubic metres would be a good volume to begin with – a room of around 7 m × 7 m × 4 m high (the fact that it is square may be of no account when the necessary degree of treatment is installed[1]) – but such spaces are often considered to be uneconomical in the post 2000 recording world. Nonetheless, economics has nothing to do with physics, so the fact remains that in the top professional studios, where things are built to a quality, rather than to a cost, the control rooms tend to be of over 200 m^3 in their basic shell sizes.

All the systems shown in Figure 8.6 use multiple bass drivers. The production of low frequencies essentially involves moving a quantity of air which is the product of the moving area and the velocity. In other words, a small radiating area can be moved with a high velocity, or a large radiating area can be moved with a low velocity to achieve the same high SPL. However, the low-velocity, high radiating area approach produces significantly less non-linear distortion. The larger surface area also increases the radiation efficiency, as the large area better matches the characteristic impedance of the air in contact with it. Quite obviously, this cannot be achieved in a small box. The use of multiple low-frequency drivers also mean multiple voice coils, and this tends to lead to better dissipation of the waste heat, so problems of thermal compression are reduced.

Another reason for using multiple drivers as opposed to simply using larger, single drivers is because it tends to become difficult to maintain the rigidity of the piston (the cone) much beyond diameters of 15 inches (380 mm). Beyond 18 inches (460 mm) adequate cone rigidity only tends to be possible by employing means that significantly increase the weight of the moving assembly, which leads to reduced efficiency. Conversely, as the resonant frequency is related to both the weight of the moving system and the stiffness of the suspension, a small, light cone can only achieve low resonance with a very loose, low stiffness suspension. That is, as the weight reduces, the stiffness must also reduce if the resonant frequency is to remain the same. This can lead to excessive fragility for professional use, so cones of less than 12 inches (300 mm) tend not be used in large monitor systems, because they tend to be either inefficient or fragile. Where responses are desired down to 20 or 25 Hz, 15 or 16 inch (380 – 400 mm) drivers seem to be an optimum choice.

The choice of whether to put multiple drivers in the same enclosure or to provide them with individual enclosures is another option. Each of two similar drivers in a 500L enclosure behave theoretically exactly as one loudspeaker in a 250L enclosure. Therefore, whether a cabinet with two drivers has a single 500L volume or is divided into two separate volumes of 250L will in no way affect the theoretical performance of the system. If the loudspeakers are reflex loaded, then obviously the two sections would need their individual tuning ports in a divided enclosure, and the sizes of those ports would be different from the ports needed for the larger, single enclosure. However, if the two enclosures were tuned to the same frequency as the single enclosure, the loading on the drivers would be identical to the case of both drivers being situated in one, larger enclosure.

Well, at least that is the case in theory, but in practice the drivers are rarely, if ever, identical in their performance or resonant frequency. Some designers feel that drivers sharing a single cabinet run the risk of de-tuning the systems by one driver dominating the port response, but this problem appears to be negligible with large enclosures tuned to very low frequencies. It is a characteristic of more consequence in small, domestic loudspeakers. What is more, as the resonant frequency goes down, the tuning of big boxes tends to become much broader, much less precisely tuned than smaller boxes with higher resonant frequencies. The only real advantage of using separate enclosures for the double woofers in large cabinets is the ability to modify the overall response by using different tuning frequencies for each cabinet, which can be useful in adjusting the response contour to a desired target function.

a)

b)

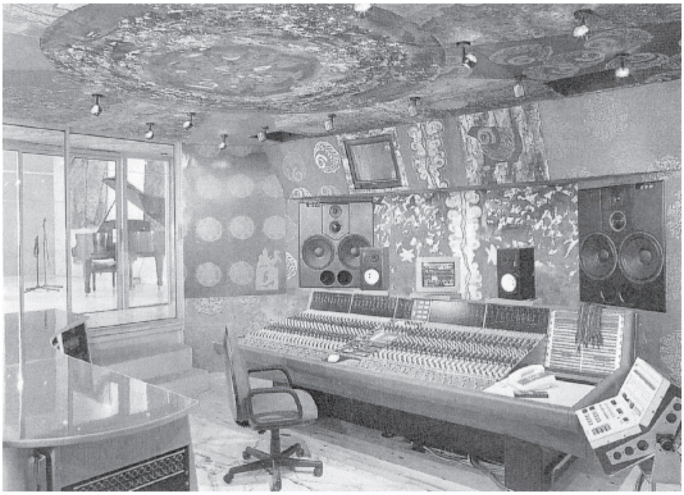


Figure 8.6 a) Garate Studios, Andoain, in the Basque region of Spain, with a Reflexion Arts 234 monitor system. b) Strongroom studios, London, UK, with a Quested Q215 loudspeaker system. c) Genelec 1035A monitor system in JVC Aydama Studios, Japan. d) Olympic Studios, London, UK, with 4-way Westlake Audio HR1 monitors

c)

d)

Figure 8.6 Continued

8.2.4 Crossover points

As described in Chapter 1, the maximum frequency to which a cone driver can operate with reasonable directivity control is when the wavelength is equal to the diameter of the cone. When four drivers are used in a square pattern, the group behaves at low frequencies as one large driver. When two drivers are used side-by-side, the horizontal directivity becomes that of a driver with the same width as the pair; that is, it would be narrower at higher frequencies than that of a single driver. With a vertical arrangement, the horizontal directivity may remain as it would be for one driver, whilst the vertical directivity would narrow. Depending upon the nature of the room acoustics, and whether the surfaces are absorbent or diffusive, designers must decide upon the physical distribution of the drivers, how many crossover points to use, and at which frequencies to make the transitions. The off-axis radiation will need to be as flat as possible if any significant reflected energy is likely to be returned to the listener. Wideband diffusers will return the frequency balance which impinges upon them, so a tonally coloured, non-flat off-axis response reflecting back to the listening position will tend to colour the overall perception in the room.

The general tendency is for manufacturers who sell large monitor systems for incorporation into a wide range of room designs to make multi-way systems. By splitting the frequency range into three or four bands, directivity control can be well maintained and off-axis energy can be kept relatively smooth in frequency balance. Designers who know that their loudspeakers are going to be installed in relatively absorbent rooms can concentrate on the axial response and a region of around 60 degrees in the horizontal and 30 degrees in the vertical, taking the advantage of using two-way systems which minimise the problems normally associated with crossover alignment and acoustic reconstruction. Essentially, in absorbent rooms, the off-axis energy which may be directed outside of the designated working area, and which may *not* have a flat frequency response due to directivity problems, is of little concern because nobody will be there to hear it and it will be absorbed at the boundaries. It therefore cannot return to the room or colour the axial response.

These concepts tend to be very poorly understood by the majority of people now working in the recording industry, and who, in ignorance, may make entirely inappropriate changes to, or of, studio systems. In so many cases, people choose their supposedly favourite loudspeakers and try to use them without any understanding of what they are really doing. It is imperative to understand that a loudspeaker system drives vibrations across a room, and that it is the *combined* response that is perceived. Nobody would buy a Formula One racing engine and expect it to work optimally when mounted in a tractor chassis, yet this is a close analogy of what many people do with their loudspeakers. If an axial response perception is desired, the room must be relatively absorbent. If a more lively sound is required, then the loudspeaker must be engineered to give good off-axis performance. This, as stated earlier, can have a great bearing on the choice of the number of crossover points. Fewer points tend, in general, to lead to better axial responses, whilst more crossover points tend, in general, to facilitate the engineering of a wider, more even off-axis response. However, nothing is absolute here, so designers work to find the most practical compromises for different circumstances of use.

Once that the number of ways has been decided upon, the crossover points need to be chosen. The decisions about where to cross over can be determined either by the usable frequency range of the chosen drivers, or the need to maintain a smooth off-axis response. Clearly though, where two-way systems are concerned, the chosen drivers must deal with greater frequency ranges than the drivers in systems with more crossover points. In the case of the monitor system shown in Figure 8.6(a), the crossover is placed at 1 kHz. It is unusual to operate 15 inch drivers up to such a high frequency, but the d'Appolito layout (in this case both vertically and horizontally symmetrical) and fourth order Linkwitz-Riley filters help to form a line source around the crossover frequency which is well-behaved both on-axis and over the working area. However, the low frequency drivers used in these systems are more expensive to produce than most low frequency drivers of similar size. They need to work smoothly from 1 kHz all the way down to 20 Hz – five and a half octaves. Due to the gradual decoupling of the outer regions of the cones as the frequency rises, the radiating area is progressively reduced, allowing response flatness and directivity to be maintained up to the crossover frequency. Nevertheless, the directivity change will inevitably lead to a non-flat response off-axis, so such loudspeaker designs tend to be used in rooms with absorbent side-walls.

Low frequency alignment and box size will dictate the sensitivity of the low frequency drivers. Some alignments cannot be achieved with high sensitivity drivers, but a reduction in driver sensitivity would require more power from the amplifier in order to achieve the same SPL. This fact has repercussions on amplifier choice, total power consumption, and thermal compression considerations, but, at least with large, in-built monitors, box size – and hence alignment – is not of such a critical nature as when designing smaller monitor systems.

The choice of mid-range radiators for high level monitor systems is largely dictated by function. In the low frequency part of a two-way system it is almost impossible to go from 20 Hz to beyond 1 kHz. Even to achieve 1 kHz is only possible with difficulty in physically large systems. This means that the upper frequency range from 1 kHz, or less, up to 20 kHz, or more, must be handled by one driver. If levels of 100 dB-plus are to be heard at a mixing console, 3 metres from the source, then levels of 110 dB-plus must be produced at one metre from the loudspeaker, with peaks perhaps up to 120 dB SPL. Only compression driver/horn combinations can be expected to work reliably over this frequency range as single drivers at such high sound pressure levels. The radiation pattern from horns is unlike that from pistonic radiators. Horns radiate a pattern much more like a section of a spherical expanding wave. For this reason, they can cover many octaves without the lobing which occurs when a piston begins to radiate at wavelengths which are less than its diameter. For a piston to work at 1 kHz and at levels of 110 dB SPL, it would need to be at least about 4 inches (100 mm) in diameter, if only for reasons of having a voice coil large enough to dissipate the waste heat. Given that at 20 kHz, the wavelength is only about 17 mm, a 100 mm diaphragm would be much too large for controlled radiation over a wide angle. Conversely, a 20 mm diaphragm would have no possibility of being able to dissipate 200 watts of heat from its small coil, so the requirements for radiation at such high SPLs at 1 kHz and 20 kHz are not compatible in a direct radiator.

Horns, on the other hand, exhibit much higher sensitivity. With a sensitivity of 108 dB for 1 watt at one metre, a horn loudspeaker could radiate 120 dB SPL with a total power input of only 16 watts. Given that perhaps 25 per cent of that power would actually be radiated as acoustic power, the dissipation of 12 watts of heat from a 2 inch (50 mm) coil is quite reasonable, especially given the large amount of metal in the magnet system surrounding it. Once the output from this diaphragm is squeezed through a one inch (25 mm) throat and matched to a horn of suitable size, the frequency range of 1 kHz to 20 kHz can be achieved with comfort, even at such high SPLs. A monitor system designer is therefore not totally free to choose whatever type of driver for a given system. Physical and engineering restraints limit the freedom of choice in many cases.

Although much is said about the unpopularity of horn loaded monitors in Europe, despite their wide use in the Americas, as previously mentioned in Chapter 4 many people seem to forget that the very widely used Tannoy 15 inch Dual Concentric monitors were exactly compression driver/horn loudspeakers from 1 kHz upwards. A cross-section of such a driver is shown in Figure 8.7. The monitor system shown in Figure 8.6(a) uses a horn which is relatively similar in geometry, but uses a fixed axisymmetric horn and a much more advanced compression driver with a vapour deposited beryllium diaphragm, the TAD TD 2001. Much of the criticism levelled against horn monitors was due to the misapplication of the technology and the legacy that was left from the ‘mini PA system’ monitors of the 1970s. Unfortunately it must also be added that both ignorance of what *is* possible, and unscrupulous comments from people marketing non-horn systems, have led many people to incorrectly believe that horn loaded upper sections in monitors cannot achieve the highest fidelity, but this is simply not true. The colouration and distortion levels of the monitor system shown in Figure 8.6(a) are *extremely* low, and their directivity in the type of room in which they are intended to be used is not an issue.

The system shown in Figure 8.6(b) is a three-way system, using a 3 inch (75 mm) soft-dome mid-range driver and a 1¼ inch (34 mm) dome tweeter. The crossover frequencies are set at 450 Hz and 4.5 kHz. The lower crossover point is necessary with this cone driver arrangement because the horizontally wide distribution of the total radiating area of the bass drivers would be too large to radiate wavelengths of much less than one metre, equivalent to 340 Hz, without severe directivity problems. The interference patterns would be exhibiting an excessive number of lobes, so an

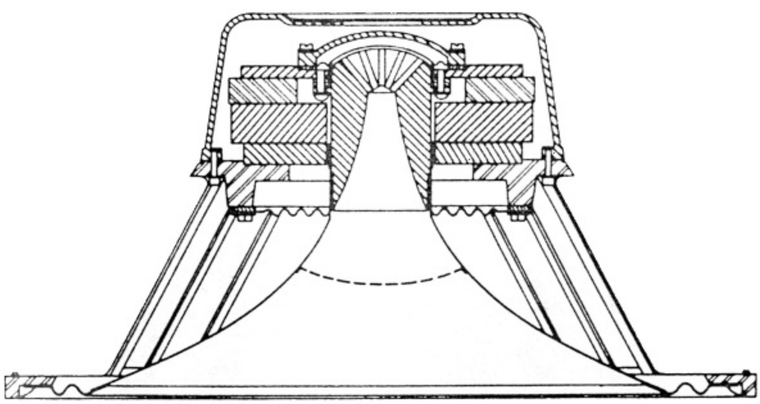


Figure 8.7 A Tannoy Dual concentric with single ferrite magnet serving for both the high and low frequency coils (see also Figure 5.10)

even directivity could not be maintained. As direct radiators can rarely be expected to span more than a decade of frequencies (or 3 octaves) the mid-range driver would begin to suffer from the same problems above 5 kHz, so the $1^1/_4$ inch tweeter takes over the response for the top two octaves.

Figure 8.6(c) shows another three-way design, but this model can be mounted with the bass drivers either side-by-side or vertically. The sculpted panel containing the mid and high frequency drivers can be re-oriented through 90 degrees. The panel takes the form of three shallow horns, or waveguides, which serve to control the directivity of the radiation at mid and high frequencies and reduce the effect of cabinet diffraction. The mid-range drivers are a pair of vertically mounted 5 inch (125 mm) cones, whilst the high frequency driver is a 1 inch (25 mm) compression driver. In this system, the crossover frequencies are 400 Hz and 3.5 kHz.

A four-way system with side-by-side bass drivers is shown in Figure 8.6(d). In this design the bass drivers only operate up to 250 Hz, from where a 10 inch (250 mm) cone driver takes over until 1000 Hz. From here on, a 2 inch (50 mm) throat, 4 inch (100 mm) diaphragm, compression driver, connected to the large wooden horn, takes the response to 4 kHz, from where a small wooden horn, fitted with a 1 inch (25 mm) throat compression driver with a 2 inch (50 mm) diaphragm continues the response to beyond 20 kHz.

From time to time, 5-way systems are also to be found, and all of these concepts from 2-way to 5-way have their applications. It should also be added that they are the results of different design philosophies which may not only reflect their intended application, but which may also reflect the order of priorities which the different designers gave to various aspects of their overall responses.

8.2.5 Power consideration

All the loudspeakers shown in Figure 8.6 are very fine systems, capable of flat responses from 30 Hz to over 20 kHz within ±2.5 dB when flush-mounted, although in many installations a high-frequency roll-off is employed for the reasons described earlier. In each case the designers have opted for different approaches to achieve what is essentially the same goal – a clean, undistorted, full-range sound, with a dynamic capability of supplying sound pressure levels at 3 or 4 metres distance far beyond what the ear can itself perceive in an undistorted manner. This excess of dynamic capability serves to allow transient headroom, damage tolerance, and reliability in long-term daily use.

However, there are big differences in system efficiency. The amplifiers normally supplied with the system in Figure 8.6(a) are specially made two-channel amplifiers, supplying 300 watts Class AG to the bass drivers and 50 watts, Dynamic Class A, to the horn. The extreme sensitivity of the horn means that at levels of 80 dB at 3 metres it would be receiving an input level of only around 10 *milliwatts*, so the low-level performance of the amplifier is extremely important. The Class A design was chosen to ensure the absence of low-level crossover distortion. Even at 100 dB at 3 metres, only about 1 watt is consumed by the mid/high driver, with the low frequency driver taking around 20 watts. This is a very high efficiency system, in which thermal compression is almost non-existent due to the low power levels and the high thermal capacity of the large magnet systems. A *stereo pair*, without signal, draws about 0.4 amps from a 230 volt supply (92 watts) and at full power 4 amps (920 watts).

Some designers feel that by using drivers of lower sensitivity they can better achieve their design goals. The systems shown in Figures 8.6(b) and 8.6(c) use amplifiers with output capabilities of over 2.5 kilowatts per side, a stereo pair requiring at least a 30 amp supply from 230 volts mains. The four-way system shown in Figure 8.6(d) uses medium and high sensitivity drive units, and a pair can operate comfortably from a 10 amp supply.

In all cases, though, oversized power cables should be used in order to keep the supply impedance down. Some amplifiers can draw current in surges, which can instantaneously drop the voltage of supplies without an adequately low impedance. Not only can this rob the bass of transient punch, but the harmonics created by the change in waveform as the voltage sags can cause problems in associated equipment, and can even crash computers. It is best to supply the amplifiers with their own, dedicated supply cable or cables, fed straight from the incoming mains supply at the main circuit breaker board. The breakers feeding the amplifiers may need to be of the delayed action type, because some amplifiers draw high surge currents on switch-on.

The question of total power consumption can be important. In hot countries, such as in southern Europe, it can be difficult to dissipate heat when it is 40°C in the shade. In many places, high current supplies to buildings are hard to organise, and when an *extra* 2 kVA of air conditioning is needed, just to cool the amplifiers, it can put a great strain on the whole studio electrical installation. Many people have said that it can be cheaper to make low efficiency systems because magnets are expensive and amplifier power is cheap. Well, it *is* cheap to *buy* the amplifiers, but the running costs over years can be very expensive indeed when system efficiency is low. Depending upon circumstances, things such as this may or may not be important to the users, but they nonetheless need to be considered because the electricity bills and the heat production cannot be ignored.

There is also a tendency for higher sensitivity systems to exhibit faster transient responses due to the more stable static magnetic fields which are a part of their general character. Computer-aided loudspeaker design programs can often call for lower sensitivity drivers in order to achieve certain design aims, but ears can sometimes ask for something different. One has to be very careful when balancing design parameters if one is not to gain in one department but lose in another. This all tends to be a question of there being too many design parameters which may affect audible responses but which are not available for incorporation into the computer-aided design processes. So, if the whole story does not go in, then it is unlikely that the whole story will come out.

The monitors shown in Figure 8.2(a) are interesting in that they are two-way but use high level passive crossovers. The pair of 16 inch (400 mm) bass drivers operate up to about 300 Hz, with the compression driver handling the upper six octaves. The dynamic impedance can dip to as low as 0.8 ohms at some frequencies and under some drive conditions, and can demand as much as 100 amps (peak) from the amplifiers. Very few amplifiers can deliver this amount of output current, and those that can tend to be very expensive indeed. These monitors are frequently used with FM Acoustics, or JDF amplifiers, which can deliver 3000 watts into half an ohm. In 2000, the price was around 100,000 dollars per pair for the loudspeakers, amplifiers and cables, the latter of which cost around 2500 dollars per side. The demands on the cables are obviously great, and bi-wiring is standards. With the cable to the horn carrying six octaves and the low frequency cable up to 100 amps, the prospects for intermodulation in a single cable would be considerable. The crossover components are also large and expensive, with oxygen-free copper inductor coils.

The engineering of such systems is not a simple task. In fact, despite initially appearing to be the simpler solution, the high-level passive crossover option can be extremely difficult to implement if the highest achievable sound quality is the goal. Multi-amplification actually simplifies many things, even though it perhaps initially seems to be a more complicated approach. However, in the above case, the designer Shozo Kinoshita, in his experienced view, chose the passive crossover option.

One of the largest commercial monitor systems is shown in Figure 8.8. The Quested HM 415 loudspeaker weighs about a quarter of a ton, and the maximum output is claimed to be 130 dB SPL at one metre. The bass units are four 15 inch drivers with external chassis, which help to improve the

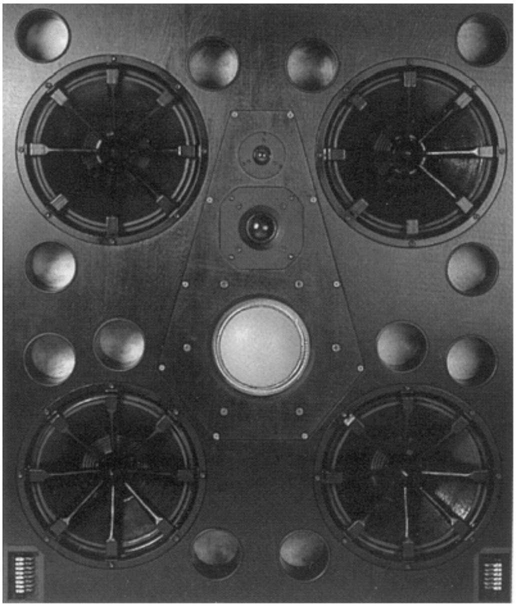


Figure 8.8 The enormous Quested HM415, 1 m 26 in height, 1 m 06 in width, and weighing over 260 kg, using a combination of rigid-dome and soft-dome radiators in the mid-range. Each 15 inch (380 mm) low frequency driver is in its own triple-ported chamber. The external chassis of the LF drivers aid the cooling of the voice coils. The systems are rated to deliver 130 dB SPL at one metre distance

cooling of the voice coils. The low mid-range driver is a 7 inch (175 mm), rigid, polyurethane foam dome. The high mids are radiated via a 2 inch (50 mm) soft dome, and the high frequencies radiate from a 28 mm dome. The amplifier system uses 5 channels, the bass being split into two parallel channels, and the full power consumption is 3.5 kVA *per channel* from the electrical supply for 2400 watts of output power.

8.2.6 Interfacing with the rooms

The 'high-end' recording monitor systems such as the ones described above are the 'Formula One' of monitoring. However, just as the thoroughbred racing cars need a good track to race on, the high-end monitor systems need to be mounted in well designed rooms. Formula One racing cars cannot perform on rough roads, and neither can the top monitor systems perform in poor rooms. As we get further into the discussion of loudspeakers for different applications we will begin to deal with loudspeaker designs which are intended to be more room-tolerant, but the larger systems, partly due to their physical size and widely spaced distributions of the drivers, cannot be considered to be particularly room tolerant. It is therefore entirely unreasonable to make judgements about the sound quality of such systems in rooms which cannot do them justice. It is sad to say that many comments which are heard in the recording industry, relating to large monitors, are based on pure ignorance, or bad experiences of misapplication.

Unfortunately, the requirement for the necessary degree of room control in order to be able to *mix* on the large monitors means that the total cost of providing such a facility is not something that everybody can afford. Large monitor systems in small rooms can be overpowering, because it can be impossible to get them far enough away from the listening position to avoid geometrical near-field effects, where the sound is heard from individual drivers, and not an integrated source. Figure 8.9 shows a scaled down system in a small control room, albeit still in large, 200L cabinets (6 cubic feet). To use any of the systems shown in Figure 8.6 in such a small room would be absurd. The use of such large systems only begins to become viable in rooms of 40 to 50 m^2, which would perhaps leave 30 to 35 m^2 of useable space after treatment. When the cost of the monitor system and acoustic control for such a room would perhaps *begin* at around 40,000 euros, many studio owners decide that it is difficult to get a return on the investment. Nevertheless, the larger studios understand the importance of the recording monitors where engineering excellence is concerned. *Marketing* only really seeks to achieve the maximum profit for a given investment, whereas professionals and artistes seek to earn a living from doing what they do to the best of their ability, and the maximisation of the financial returns are not their sole concern. The pressures to find less expensive solutions are nowadays very great, but it still needs to be understood that just because something cannot be afforded does not mean that it is not necessary!

The tendency to use near-field, or, more properly *close-field* monitors is often simply an attempt to take the room out of the monitoring equation. One problem with this approach is that there is not much room for more than one or two people to hear the optimum sound, and what is more, the

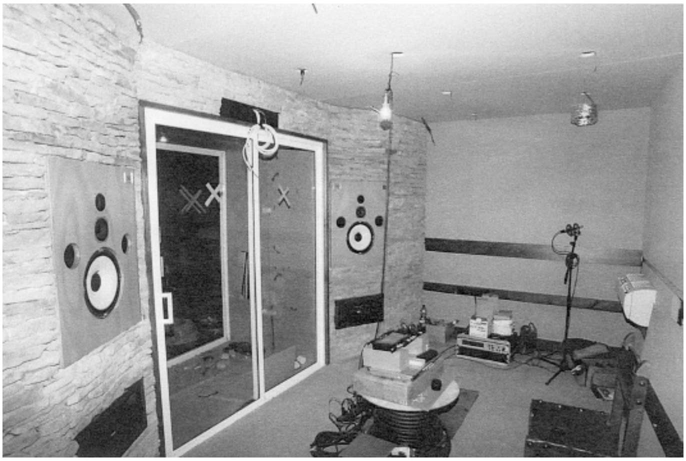


Figure 8.9 A small control room, under construction in Ubeda, Spain, with a miniaturised Reflexion Arts 240c system. The cabinets are large, to allow a generous and fast low frequency response, but the drivers form a compact group to minimise the geometrical near-field problems at short listening distances. The rear wall, side walls and ceiling are highly absorbent behind their fabric coverings

frequency range of the small loudspeakers is necessarily curtailed at the lower end of the frequency spectrum. (Sub-woofers rarely yield accurate bass, as will be discussed in later chapters.) The close-field is the space within the critical distance – the critical distance being that at which the energy which is contributed to the overall sound is equally supplied by the direct and reflected sound fields. It thus follows that as the room decay time is reduced, the close-field increases in size. The high degrees of absorption in some control room designs thus seeks to extend the close-field of the large monitors all the way to the listening position. This effectively yields a close-field response but with a greater optimum listening area. [The acoustic *near* fields, strictly speaking, both geometric and hydrodynamic, are regions very close to the loudspeakers where a highly complex, unintegrated sound-field exists, and where the composite sound that we hear in the far-field has not had time or space to jel into an integrated wavefront. Listening to the monitors shown in Figure 8.6(d) at a distance of one metre would be a typical example – the physical distribution of the individual sources would audibly be very evident.]

There is also a lot of nonsense spoken about control rooms, especially by people who are just repeating hearsay; and it has to be said that the marketing pressures lead to some very partisan passing of opinions which are little more than attempts to gain a commercial advantage. Clearly, the majority of studio owners who cannot afford good control room acoustics and large monitors are rarely going to admit that their studios would be much better if they *could* afford them, so the received wisdom about these things often has little to do with fact. Comments that well-controlled rooms are not necessary should be treated with suspicion. If such things were *not* necessary, then why would almost all the top studios have them?

8.2.7 A word about listening levels

Some people may be shocked reading about listening levels of 100 dB, or more, when so much safety information now restricts industrial levels to 95 dBA or less. However, as the levels within symphony orchestras can easily exceed 95 dBA or 100 dBC (the C-weighting measuring more of the bass) it would seem to suggest that all orchestral musicians should be deaf after a few years of work. This is patently not in accordance with the reality, as most experienced classical musicians exhibit excellent hearing acuity. It also follows that if the performance levels do not damage the ears, (except in some cases where musicians play directly in front of the trumpets), then the loudspeaker reproduction of such ‘natural’ levels should also lead to similar results.

Hammer blows, on the other hand, despite only reading 95 dBA on a sound level meter can produce rapid peaks of 135 dBA or more, but they are too short for the meters to read. It is these peaks which damage the hair cells in the ears, but the more rounded waveforms of music rarely contain any peaks so far above the measured levels. Therefore, whilst *mixing* is not recommended to be carried out at such high levels, for many perceptual reasons, it is still not unreasonable to work during the recordings at the realistic acoustic levels experienced by the musicians. Drum kits would have already been banned in many countries, for health and safety reasons, if a short career as a drummer automatically led to deafness. Classical soprano singers can also produce over 120 dBA one metre from their mouths, and this is why many good recording engineers never place a recording microphone closer than one metre from a soprano – to save the microphones from overload – yet there are few reports of people being deafened by sopranos. Nevertheless, it is still wise to only monitor loudly when deemed necessary, and not to make a habit of doing so if not required. Perceptually, also, mixing is better carried out at levels more close to end user reproduction levels, but that will be dealt with in the next section.

8.3 Mixing monitors

Once we arrive at the stage of mixing the multitrack recording down to stereo, or surround-sound, we are no longer in an environment where we can affect the *performance* of a piece of music. Stimulating the musical performance is no longer either necessary or possible, although ‘vibing’ the mixing personnel, who are still very much a part of the creative process, is still a possibility. A quick blast at 105 dB can, at times, be very satisfying, but mixing at such levels is rarely a good idea. The equal loudness contours shown in Figure 8.4 demonstrate clearly how at different listening levels we perceive different frequency balances. In general, when people are listening to music at home, they tend to listen between 75 and 85 dB SPL. The dBC weighting curve which is used on many sound level meters represents something very similar to the inverse of the 80 phon curve of Figure 8.4. (The more common dBA curve being similar to the inverse of the 40 phon curve – used for the assessment of background noise nuisance.)

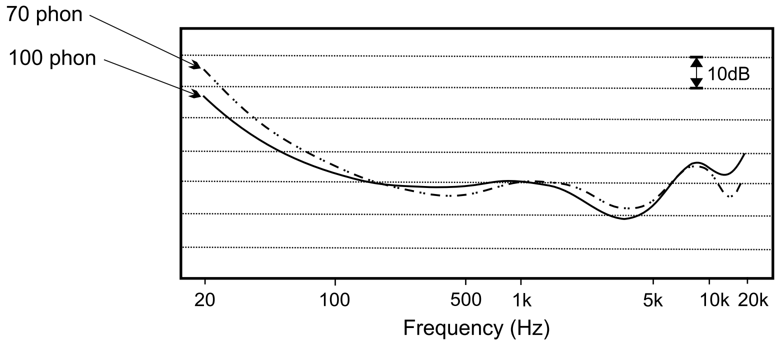


Figure 8.10 The 70 and 100 phon curves overlayed to coincide at 1 kHz. Note that they coincide elsewhere only at around 200 Hz and 6 kHz. At the frequency extremes they differ by as much as 10 dB, with the ear being *more* sensitive low frequencies at the higher SPL but *less* sensitive to the 15 kHz region. The way that the curve intertwine does not make the prospects of automatic correction a very practicable proposition

In Figure 8.10, the 100 phon curve has been superimposed on the 70 phon curve. If we mix at 100 dB SPL, then listen at 70 dB SPL, the perception of the low and high frequencies relative to the mid-frequencies will change by the difference between the curves. A mix which seems balanced in frequency at 100 dB SPL will sound at 70 dB SPL like the bass should have been mixed 3 or 4 decibels higher, and it may sound somewhat dull due to the ear's lower sensitivity to the treble frequencies. Mixing at high levels is therefore unlikely to lead to well-balanced mixes at normal listening levels. Furthermore, mixing at 100 dB SPL plus for day after day will almost surely lead to serious hearing fatigue. It will also lead to temporary shifts in the hearing threshold as the ear's protection mechanisms come into play, and normal perception will not be possible. In fact, from the hearing sensitivity contours shown in Figure 8.4, it can be seen that the contours tend close up at low levels of low frequencies. Despite the fact that 10 dB is generally considered to double loudness, it can be seen that at 70 dB SPL and at 40 Hz there is only about 4 dB between the adjacent loudness doubling/halving contour lines. This implies that a 4 dB difference in overall level at 70 dB SPL may subjectively double or halve the relative loudness of low frequencies, whilst the mid frequency loudness would change by a much lesser amount. This is another reason to avoid mixing at much higher levels than the expected reproduction levels, because the low-frequency/mid-frequency relative balance will be difficult to judge in relation to how it will generally be perceived domestically.

Mixing also tends to be a much more continuous process than recording, so sustained high level mixing can lead to more problems than high level monitoring during recording. Even if recording engineers do animate musicians or look for noises at 105 dB SPL, they probably would not be exposed to such levels for more than about an hour a day, and in short bursts at that. And, despite 'common knowledge' about recording engineers being deaf, there is no evidence to support that idea. Most of them have very acute hearing, which they probably take care of more than most other people do. [DJs, – well, that is another matter!]

In the world of mixing cinema soundtracks, the monitor volume level is fixed, to guarantee that the audiences in the cinemas hear the same level, and hence frequency balance, as the mixing personnel in the studio dubbing theatres. This is a luxury which cannot be enjoyed by music mixers, but, as few people will listen seriously in their homes at either 60 dB SPL or 100 dB SPL, an 80 to 90 dB SPL mixing level seems quite reasonable, or a little lower or higher if desired.

Mixing monitors therefore do not need to be capable of the same output levels as the recording monitors, and this fact can be significant. Ten decibels less in peak output capability means ten times less power handling capacity if the same sensitivity of drive units are used. This could make the engineering much simpler and the area occupied by the drivers more compact. However, the tendency is to use physically smaller loudspeaker cabinets at closer listening ranges. Essentially, what the mixing personnel are trying to do, subconsciously or otherwise, is to remove the room response from the listening chain. To move the loudspeaker close is a much cheaper alternative to treating the room. However, the smaller cabinets require less sensitive drive units if the bass response is to be maintained, (as was described in Chapter 3) so the amplifier power still needs to be quite considerable in many cases, even to work at a maximum average level of 100 dB SPL at 1.5 metres distance. It also means that some of the potential advantages of using much lower powers are not realised, so thermal problems arising from needing to lose the same amount of heat from smaller drive units can have its repercussions on design priorities. In some cases, sensitivities as low as 81 or 82 dB for one watt at one metre are encountered, whereas it is rare indeed to find recording monitors with sensitivities below the 90 dB sensitivity level. The widely used UREI 815s of the 1980s had sensitivities of 103 dB for one watt at one metre. Figure 8.11 shows the two extremes. Obviously, when the smaller loudspeaker needs almost 200 times the power (22 dB) that the large ones need to produce

Figure 8.11 The Urei 815 and ATC SCM10 loudspeaker systems. The smaller loudspeaker needs an input of almost 200 watts to develop 103 dB SPL at one metre distance. The UREI 815 will achieve the same SPL with an input of only 1 watt

the same SPL, the engineering considerations can be very different in the two cases. Such are the problems that face loudspeaker designers.

For many people, mixing is an insecure process. Some of the reasons why will be discussed in Chapter 10, but some of the insecurity is reduced by following fashions in the choice of mixing loudspeakers. In the rock music world, the Yamaha NS10M served as an unofficial reference standard for twenty years or more. Of course, to become an international standard, even if unofficially, the loudspeakers need to be internationally available, so the offerings of large companies tend to be favoured rather than the use of locally manufactured brands – even if their performances are similar. Nevertheless, it is not *all* down to fashion and marketing. There is usually some reason why so many people in so many places gravitate towards certain monitor loudspeakers out of the plethora available to them. Some manufacturers have made studies of hundreds of loudspeaker responses, both professional and domestic, and concluded that a loudspeaker which exhibits something close to the mean response will be likely to succeed as a mixing monitor. However, that work applies only to the frequency response amplitude, but there are many more qualities which people look for during music mixing.

Domestic loudspeakers are usually inadequate for music mixing because they lack the robustness to withstand such operations as the soloing of bass drums, which are essential during the mixing process. Characteristics such as absolute sound quality and the ability to give a realistic response under the conditions of use found in mixing rooms are aspects of mixing monitor performance which may differ very significantly from the requirements of use found in most domestic environments. It is rare, for example, for people to listen seriously to music with the loudspeakers placed on top of the far side of the dining-room table, but this would not be too unlike the placing of loudspeakers on top of a mixing console. Therefore, for professional mixing purposes, the desired anechoic response must take into account the mounting conditions under which the loudspeakers will actually be used if reasonably accurate mixes are to be created.

The aforementioned dominance of the NS10M as a mixing monitor for pop/rock/electronic music, over so many years, was surely in part due to the fact that it exhibited a tendency for its overall response to flatten when placed on top of a mixing console – its predominant position of use. Figure 8.12 shows the gradual change in an NS10's response as a mixing console and room are brought into its proximity. From these results it should be apparent that mounting mixing loudspeakers on a meter bridge or on pedestals are not alternative options. A mixing console will augment the low frequencies but a pedestal will not. The decisions must be made according to the designers' intentions, and the local boundary conditions presented by the rooms unless the monitors are self-powered, and fitted with tilt and roll-off controls to compensate for the acoustic loading differences. Far too many people, even professional mixers, fail to realise that the loudspeaker and its mounting conditions cannot be separated. No fixed response loudspeaker is suitable for all mounting conditions.

The NS10's popularity was also probably due to the fast and uniform decay of its time response. Figure 8.13 shows a selection of waterfall plots from nine different mixing monitors. The similarity of the plots of the

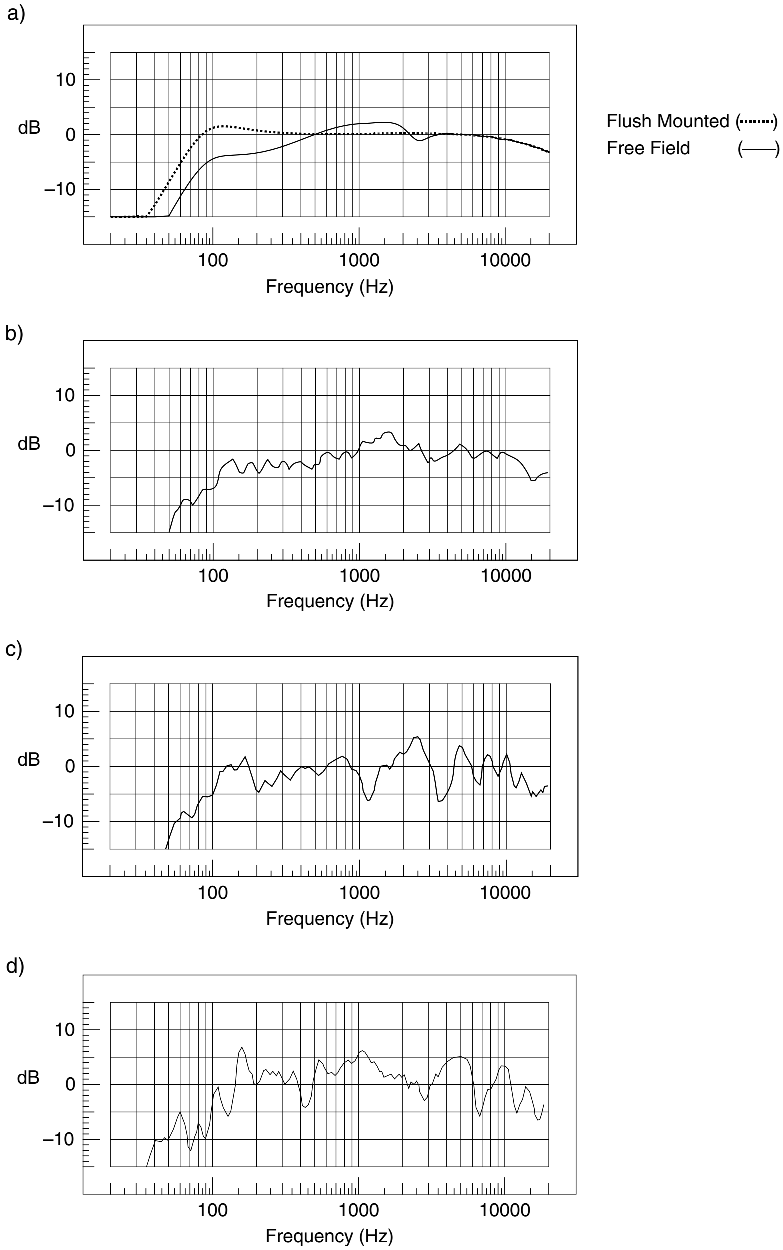


Figure 8.12 a) Response of idealised loudspeaker: flush mounted (dashed line) and free-field (solid line). b) Response of Yamaha NS10M, outside, flown 4 m from floor and wall. c) As b), but mounted on a mixing console metre bridge (all flow 4 m from ground). d) As c), but with the mixing console on the floor in a reasonably controlled room. Note how the overall response gradually changes – (b) shows a response rather similar to the free-field response in (a), whereas (d) shows a response more similar to the flush-mounted response in (a)

NS10M and Auratone 5C are too close to be mere coincidence. The Auratone was one of the first, internationally used 'references' – originally used as a small loudspeaker reference in the days when the large monitors were still the most commonly used mixing loudspeakers. The NS10s very largely displaced the Auratones when they arrived on the scene in the early 1980s. They exhibited a considerably wider frequency response and significantly more output capability, yet still maintained the response characteristics that made the Auratone 5C so popular. The deeper consequences of their responses will be discussed further in Chapter 11, but it can be seen from Figures 8.12 and 8.13 that their responses were tending towards flat and fast when placed on top of mixing consoles, and these two characteristics are conducive to good mixing. Basically, if one could ignore the colouration, and check the low frequencies on other systems, the NS10s and 5Cs told many mixing engineers what they needed to know in order to make a good balance of instruments and reverberation.

The twin demands of lowering production costs and maximising the commercial acceptance of the mixes have led to a situation where by far the majority of music is mixed on relatively small monitor systems, a selection of which is shown in Figure 8.14. Although fashion plays a big role in the choice of which loudspeakers to use, and some commonality of choice is still to be found, there are nonetheless many individual choices made by different people, and it is not by any means unusual to hear discussion between people who cannot understand how each other can use their loudspeakers of choice. Many of the differences come down to different hierarchies of priorities and methods of working – which both

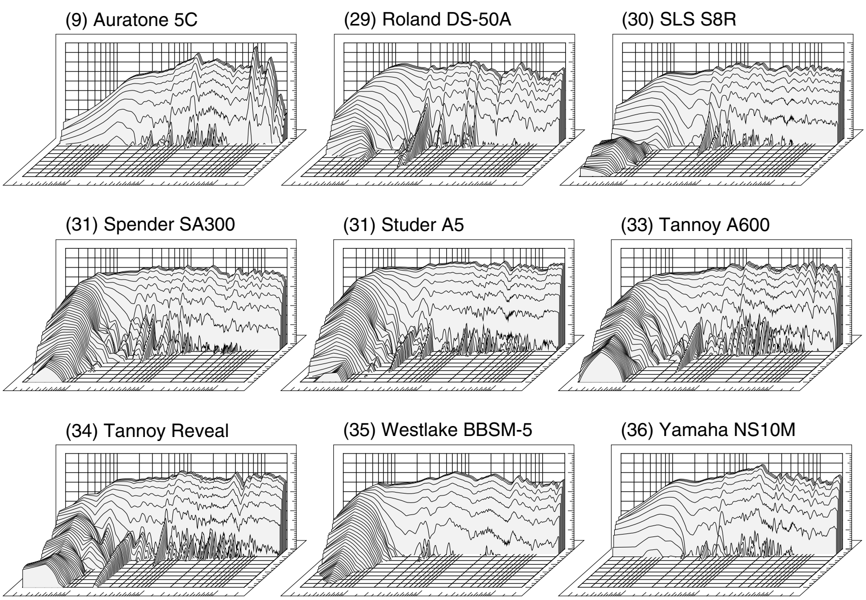


Figure 8.13 Nine waterfall plots – all of close-field monitor loudspeakers in an anechoic chamber

a)
b)
c)

Figure 8.14 Three families of studio monitor loudspeakers. a) Dynaudio Acoustics systems b) Quested Monitoring systems. c) Westlake Audio systems

ultimately translate to individual ways of thinking about the work. Mixing-loudspeaker characteristics can be separated into parameters such as spectral uniformity, dynamic performance (transients), distortion, sound-stage imaging, off-axis performance (very relevant in less controlled rooms), ambience reproduction (transparency), and, of course, the overall impression. However, the order of these characteristics will not be the same for all people. And, what is more, work and leisure requirements may be very different. Many top recording engineers, mixers and producers do not use the same loudspeakers at home that they use in the studios, even though the sizes of the loudspeakers may be similar. At work, they need to hear all the problems; at home, they simply want to enjoy the music. But another reason for the difference is because, in general, almost all serious mixing will be done in rooms with a reasonable degree of acoustic control and a lot of equipment, whereas acoustic control usually does not take much precedence in the designs of their homes, and domestic furniture tends to be rather acoustically different to mixing consoles and equipment racks. The object of the exercise is therefore to do mixes in a professional environment which will translate to a domestic environment. The recording engineers' homes also furnish them with a means of listening as other people listen, which helps them to get another perspective on their work. For these reasons, mixing monitors for use in studios have developed along their own course.

In a report presented in 2004 relating to the selection of a new 'standard' choice of monitors for the BBC [2], the author acknowledged; "Loudspeakers are a sensitive topic on which many people have very different views . . . . ." Despite this, however, in an organisation such as the BBC, with over 600 sound control rooms, some standardisation is necessary because the staff need to move around from room to room, but always need to know what they are listening to. They also use rooms which are rather more dead than most domestic rooms because 'Critical assessment of sound quality cannot be carried out in rooms with large areas of specularly-reflecting surfaces and long and uneven reverberation times, however attractive they might be architecturally[2]'. As we will discuss later, domestic loudspeakers *must* deal with these conditions, but they really cannot replace professional loudspeakers if a good degree of consistency of sound quality is required when used in typical mixing environments.

The BBC carried out their tests in rooms with a mid-band decay time of about 0.2 seconds, which is quite typical of modern control rooms, but it is much lower than is normally to be found in most domestic rooms. The listening tests were carried out during a period of three weeks, by eleven 'democratically chosen' representatives of the staff – all very experienced listeners – on a total of 28 pairs of loudspeakers. In the 'Summary and Conclusions' of the abovementioned report it stated "Inevitably, the final outcome was not entirely clear-cut." The tests were made using large, medium and small loudspeakers, without the listeners knowing the identity of any of them. It is interesting to note that no set of three from any one manufacturer stood out from the rest in either quality or family resemblance. The best family resemblance of a large, medium and small loudspeaker was of mixed manufacturer.

8.3.1 Location dilemmas

The main purpose of the mixing monitors is to reliably enable the mixing personnel to achieve musically and timbrally balanced mixes. Working in the close-field, if the loudspeakers are too small and their low frequency responses are too restricted, then the lower octaves may be leading an uncontrolled life of their own due to the inability to monitor them. Adding shelf equalisation to an instrument at 80 Hz may give rise to unintentional rises at 30 or 40 Hz, leading the mixing personnel to believe that they have obtained the desired tonal balance which subsequently, when played on larger loudspeakers, turns out not to be the case. The use of larger loudspeakers for the mixing may mean pedestal mounting, behind the mixing console, but unless the back of the console is heavily treated with absorbent material, the reflexions from the rear surface can ricochet around the room, affecting the clarity and stereo imaging. This concept of 'mid-field' monitoring also brings more of the room acoustics into the listening equation, and the omnidirectional low frequencies will rarely be as flat as from monitors flush mounted in the wall, which do not suffer from the reflexions of rear radiation and the subsequently response irregularity due to interference at the listening position. Neither will the low frequencies from mid-field monitors be as likely to be as flat as the ones that are available (to a more restricted degree, of course) from loudspeakers mounted in the close-field, where the direct sound dominates the overall response. The whole situation is full of frequently irreconcilable compromises, and great experience may be required in deciding just how the most workable compromises can be achieved.

There is, of course, no fixed response for all mixes. The mixes are, above all, artistic interpretations of the performances which they encapsulate, so if a certain type of loudspeaker leads a certain mixer or producer to achieve their desired results, then those loudspeakers will, for them, be excellent mixing monitors, even though for other people they might be considered to be unusable. Nonetheless, in general, achieving a smooth response at the listening position is still a highly desirable goal.

Mixing monitors are therefore a means to an end, and not an end in themselves. They are tools, and just as with any other tool, such as a tennis racquet or a pair of football boots, different models may suit different individual styles of use *and* personal comfort. However, it is indeed rare to find normal domestic hi-fi loudspeakers in use as mixing monitors, or typical mixing monitors in use as domestic hi-fi loudspeakers, because they are used quite differently and in acoustically different circumstances. But in a growing number of cases, yet another type of loudspeaker has been introduced into the chain to try to interface better the professional and domestic worlds, which we shall now look at in a little more detail.

8.4 Mastering loudspeakers

Historically, the last stage of the quality control of the recording process, both artistic and technical, was the disc cutting, when the tapes were transcribed to cellulose acetate discs. From these discs the stampers would be made, by processes of electroplating, which would in turn mould the vinyl records for sale in the shops. Once all the settings had been decided upon, and a written record had been made for future reference, an acetate test cut could be taken home, or to the record company offices, to assess the sound in known, domestic-type conditions, which were considered to be more representative of end-user conditions than the machinery filled space of the disc cutting room. If all sounded well, the engineers or producers could return to the cutting room, adjust the equipment once more to the settings which had been written down, and a master disc could be cut and sent to the factory. (It was important *not* to play this disc, as the soft acetate was easily damaged.) If it was felt that some adjustment needed to be made to the sound, then the appropriate changes could be made to the settings of the equalisers or compressors before the final disc was cut.

Reference acetates were also frequently cut to assess tracking, because what could be cut on to disc was not always able to be played back on cheaper domestic equipment. The cheaper cartridges could not always keep their styli in the grooves if they were too heavily modulated, and customers would return the discs to the shops if 'the needle jumped', which could be very expensive for the record companies in lost sales. Vinyl discs therefore always had to be made to a realistic lowest common denominator, and it was not only the stylus tracking which made this necessary. In the 1970s and early '80s most domestic loudspeakers were not capable of supplying high levels of low frequencies, so excessive bass levels could either cause distortion on playback or cause the listeners to turn down the bass on the tone controls of their equipment. In neither case would they be hearing what the people recording the disc were intending them to hear, so the artistic intention of the disc could not be fulfilled. Once again, it would be an acetate test disc that would usually be used to confirm the overall domestic compatibility of the musical mix.

With the advent of the compact disc, the greater capabilities of digital recording introduced low frequency response possibilities that were never available from vinyl discs or tape cassettes, and the domestic loudspeaker manufacturers began to respond with more robust loudspeaker systems which could highlight the new advantages of digital recording. Music mixes then needed to be prepared for multi-format release, on vinyl disc, compact disc and tape cassette, and producers still wanted to know, as ever, what the mixes would sound like on the radio. The early 1980s therefore began to see the birth of a new concept within the recording chain; the concept of the mastering studios, where all of these questions could be resolved.

Mastering engineers tend to always work in the same place – their own dedicated mastering room. Their rooms, equipped with their own choices of loudspeakers and equipment become their references. During the course of a year, many more recordings will usually pass through a mastering room than a mixing room, and the recordings which a mastering engineer deals with are likely to vary, in terms of musical styles *and* recorded quality, much more than the recordings worked on by a mixing engineer. Consequently, mastering engineers gain much experience about how a very wide range of recordings sound in their own, personal rooms. In general, mastering rooms are much more sparsely equipped than mixing rooms. There is usually no mixing console between the loudspeakers and the listeners, and what furniture there may be tends to be of an open and not very reflective nature.

Mastering is the last chance to solve problems, or at least to be aware of them before the music goes to the factories and the shops. It is therefore necessary at this stage to be able to listen critically, both artistically and technically. With experience, mastering engineers get to know the relationship between how things sound in their own rooms and how things sound in a wide range of circumstances in the outside world.

Mastering rooms often have acoustic properties somewhere between the typical studio control room acoustics and typical (if such a thing exists) domestic replay circumstances. The monitor systems which they use are frequently of wider frequency range than many mixing monitors, because the mastering personnel need to be able to look for things that have been neglected in the mixing circumstances, and they often do this by using larger loudspeakers which are either *not* of the reflex design, or are reflex boxes with very low tuning frequencies. Mastering monitors can often afford to be larger because they do not need to fit into the spaces left after the mixing console and equipment racks have been installed – both of which are very necessary for mixing purposes. A typical mastering room is shown in Figure 8.15.

The great tendency is to use free-standing loudspeakers, because many mastering engineers feel that these are more typical of the way in which the music will be perceived in the domestic circumstances. The larger boxes and lower low frequency responses facilitate the assessment of the low frequency colouration, and reduce the masking effect which can cover low-level detail when reflex cabinets of a higher tuning frequency are used. Furthermore, the mastering studios need to be able to cope with whatever work comes to them, and if some work is of a very high quality, extended low frequency nature, they need to be able to hear it at its best, because guessing is not what mastering engineers are paid for. Their job, to a very

Figure 8.15 A typical mastering room arrangement. Optimum Mastering, Bristol, UK, with engineer Shawn Joseph and a pair of large PMC transmission lines

great degree, is to give confidence to the rest of the people who have been working on the performances, recording, mixing and editing of the music, and also to the record companies who may be distributing it. For the above reasons, mastering loudspeakers tend to be of an audiophile high fidelity quality, which are capable of being used in rooms with an acoustic which is often less controlled than the studios in which the mixes have been made. Mastering monitors are *not* required to support the stresses of solo'd drums or bass guitars, so they generally do not need to be as robust as recording or mixing monitors. This fact can allow mastering engineer's greater flexibility in their choice of loudspeakers.

From time to time, however, one encounters mastering studios which *are* more like a recording studio control room, but with some notable modifications. The room shown in Figure 8.16 is based on a control room design, but with the loudspeakers mounted in such a way to maximise the accurate listening area, as no compromises need to be made to accommodate a mixing console. The flush-mounted loudspeakers, although of a type used in smaller control rooms, will normally not be used at control room levels, because mastering engineers tend to work at domestic levels, so the high frequency roll-off that would be applied if they were in a recording studio (as discussed in Section 8.2) is left flat. The owner of the mastering room shown in Figure 8.16 is a former recording engineer who feels the need to be able to refer to a more typical, recording-style, full-range monitor when needed, and uses a smaller set of monitors as a secondary reference. As in all matters associated with loudspeaker choice, no absolute, hard and fast rules apply, though strong trends are apparent.

Figure 8.17 shows yet another mastering room. In this case the monitors consist of loudspeakers on stands but with a dedicated sub-woofer below each one. These sub-woofers come into use below about 80 Hz, where the directional characteristics of the auditory system are very poor. This allows the mastering engineer to use full-range monitors, but of smaller physical size than those shown in Figure 8.15. The room is used for much work in

Figure 8.16 A mastering room prepared for a future, centre-front loudspeaker

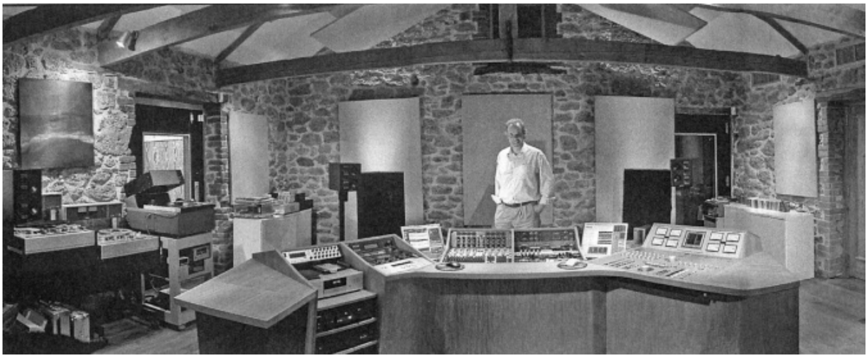


Figure 8.17 Super Audio Mastering – Chagford, Devon, UK

surround sound, and the owner felt that five large, full-range loudspeaker boxes would cause too many acoustic problems if they were surrounding him at close range; especially when only working in 2-channel stereo. However, it is important to note that the sub-woofers are not used with any arbitrary bass management system. They are only used for extending the low frequency responses of the loudspeakers which are mounted above them. There are therefore five sub-woofers for five loudspeaker channels.

There is a significant difference between the responses of small cabinets with sub-woofers below, and larger full-range, integrated cabinets when free-standing in rooms. Although the on-axis responses may be the same in either case, the greater mid/high frequency diffraction of the smaller cabinets gives rise to more reverberant energy in those frequency bands than would arise from generally larger cabinets. In semi-reverberant conditions, typical of many mastering rooms, this gives rise to a better balanced reverberant response compared to the somewhat bass-heavy reverberation which is typical from the larger cabinets, whose greater front baffle areas restrict the mid/high-frequency diffraction. Obviously, though, in a more absorbent general acoustic, the difference would be less noticeable at the listening position, so once again the choice of loudspeaker designs is very closely linked to the rooms in which they will be used.

It is often said by mastering engineers that the loudspeakers which are used for the great majority of music mixing are not sufficiently revealing of low-level details, whereas mixing engineers say that the loudspeakers typically used for mastering are difficult to use for mixing because they do not interface well with the equipment distribution of a mixing studio. They also often speak about the difficulty of mixing on loudspeakers which reveal too much detail, because the information overload becomes an obstacle to concentrating on the mix, and that smaller loudspeakers give a more close approximation to typical end-use conditions. It is interesting to remember that in the 1960s and early 1970s, the people who *mixed* the music were often referred to as *balance* engineers, and that musical balance is the essence of good mixing.

What is very clear therefore is that whilst recording, mixing and mastering *can* all be accomplished on one set of loudspeakers in one room, there are operational and financial reasons why the use of different loud speakers and acoustic conditions can be beneficial at each stage of the process, and that recordings which have passed through those different stages are perhaps better prepared for the very wide range of circumstances in which they will ultimately be heard in peoples' homes.

8.5 Domestic loudspeakers

The scope of domestic loudspeaker design and performance is enormous. Some of the audiophile designs cost more than even the largest recording monitors, whilst the smallest and cheapest are perhaps those in a typical 'ghetto blaster' or portable music centre. All, however, ostensibly serve one purpose – to give enjoyment to the listeners when reproducing music, either recorded or via live broadcast. The range is vast because the degree of importance which people give to listening to music is also very great, from fanatical to inconsequential, and also because the homes in which music will be heard can be architecturally very different. It was recognised a long time ago that many Californian designers of hi-fi loudspeakers aimed for different low frequency alignments than many British designers, because the typical wooden framed houses of California responded very differently to the musical signals than did the solid, stone or brick houses in the UK.

Rooms which are heavily endowed with soft furnishing will be absorbent in the higher frequency ranges. As most domestic listening takes place in the far-field, where the reflected response dominates the direct response from the loudspeakers, a loudspeaker with a brighter anechoic response may sound better balanced in some homes than a loudspeaker with a flatter axial response. Loudspeakers placed against walls or near to corners will exhibit a much reinforced bass response compared to loudspeakers which are placed away from reflective boundaries. A flat loudspeaker may therefore sound bass heavy when close to a reflective boundary, so a choice of loudspeaker with a reduced low frequency response may sound more balanced in such circumstances. In many cases, for reasons of domestic harmony, loudspeakers may have to be placed in conditions which would not suit those having flat anechoic responses, so the range of different loudspeakers available to the public can offer a choice to select a model which best matches the circumstances of use and the affordable price range. It is therefore essential to judge domestic loudspeakers at home for final assessment, unless the acoustic conditions and mounting conditions in the showroom are very close to the conditions of use. And, of course, one should never be reluctant to use any tone controls on the amplifiers to make a final adjustment to taste. The concept that the tone controls should be left flat for a maximally accurate sound to be heard is absurd, unless the listener lives in identical conditions to those in which the recordings were mixed or mastered.

Low colouration, low non-linear distortion, a smooth frequency response which tends towards flat in conditions of use, and an adequate dynamic range are all common goals for either recording, mixing, mastering or domestic hi-fi loudspeakers. However, the fact that in the latter case the acoustic conditions of the rooms in which they will be used will almost never be based on the requirements for the loudspeakers; the loudspeakers must be chosen to suit the rooms, whereas in the former, professional cases the rooms are often treated to accommodate the loudspeakers. For this reason, whereas professional loudspeakers tend towards similarity, the domestic loudspeakers tend towards diversity. The first book mentioned in the Bibliography at the end of this chapter deals with the subject in more depth, and those 500+ pages only deal with high performance loudspeakers, which gives an indication of the scope of domestic loudspeaker design.

Taste is also an allowable parameter in the selection of domestic loudspeakers to a much greater degree than in professional loudspeaker choice. As was mentioned earlier in the chapter, most professionals in the music recording industry do not use at home the loudspeakers that they use at work. Very few people, except some audiophile die-hards, want to hear every error and distortion in a recording when they are listening for pleasure. There is therefore an incentive for the manufacturers of domestic loudspeakers to make products which sound pleasant, even at the expense of accuracy. This is a totally reasonable situation in the design of products which are, after all, intended to be used for enjoyment. Nevertheless, the same basic principles apply to the majority of the four categories of loudspeakers discussed so far. On the other hand, for the final group of loudspeakers to be discussed in this chapter, sounding pleasant is their *only* function, and in *their* design just about everything that we have discussed so far is about to be turned on its head.

8.6 Musical instrument loudspeakers

As mentioned at the beginning of the chapter, loudspeakers for musical instrument amplification are part of a music *production* process, not a *re*production process. Reproduction by loudspeakers can be classified as a situation where the acoustical output is intended to accurately represent the electrical input. Conversely, in music *production* loudspeakers, the controlling factor is the subjectively desirable nature of the output sound, irrespective of its relationship to the electrical input signal.

A bass guitar, for example, has a very dynamic output. The peak-to-mean ratio of the signal is very great, and if a bass guitar is directly injected into a mixing console, without passing through a musical instrument amplifier, it is almost obligatory to compress its dynamic range. If this were not done it would have difficulty blending in a balanced way with the other instruments of the ensemble. Reproduction on domestic equipment would be difficult, because even at reasonable volume levels the peaks could be stressing the entire analogue portion of the playback system. Therefore, when performing live, and not having the advantage of access to electrical compressors, it is often left to the loudspeakers, themselves, to exercise some dynamic range control. Furthermore, the loudspeakers used with amplifiers for musical instruments have often been developed hand-in-hand with the instruments themselves, and hence have also been designed to *enhance* the tonal quality of the instruments. To this end, they have often been designed to introduce linear and non-linear distortions which lead to subjectively desirable colouration, and employ dynamic range compression

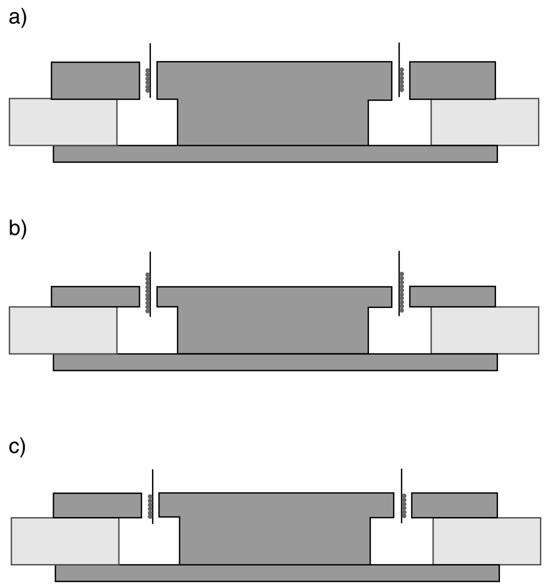


Figure 8.18 Relationship between coil length and magnetic gap length. a) Short-coil/long-gap. b) Long-coil/short-gap. c) Equal coil/gap length

to add punch to the sound. All of these qualities are, of course, anathema to the designers of loudspeakers for high quality music *re*production.

Figure 8.18 shows three types of magnetic gap and coil proportions which can be used for low frequency drivers. Example a) shows a short-coil – long-gap design. This concept is perhaps the best for low distortion reproduction, because the coil remains in the gap and crossing the same number of flux lines throughout its range of travel. It is also apparent that the coil is in good thermal proximity with the large magnet assembly, so it can easily lose the heat which its resistance dissipates. The drawback to this design is predominantly cost, because of the large quantities of magnetic materials which are needed. Example b) shows the opposite approach, the long-coil – short-gap design. In this case, half of the coil is outside of the gap at any one time during its entire range of movement. This is good in terms of linearity of motion with drive signal, but it is bad in terms of efficiency and heat loss from the voice coil. The efficiency is poor because only a proportion of the energy put into the coil is acting on the magnetic flux lines crossing the gap, due to the fact that much of the coil always remains out of the magnetic gap. It is poor in terms of heat dissipation both because only half of the coil is surrounded by metal, to conduct the heat away, and, compared to the design shown in a), because there is less metal to do the conducting. However, less magnetic materials are used, so cost of production can be considerably reduced. In the cases of examples a) and b) the linearity can be excellent, but example a) tends to have the slight advantage.

In example c) the coil and gap are the same length. Production costs can be kept low, and efficiency can be kept high, but it can be appreciated that almost as soon as the coil begins to move, it begins to leave the gap. Once it begins to move out of the gap, less turns of the coil are influenced directly by the magnetic flux lines crossing the gap, although the fringing effect which extends beyond the gap will ameliorate the effect of the abrupt movement out of the gap. The result of this motion is that as more current is applied to the coil, it has less proportional effect as a driving force. The driving force is therefore not linear in its relationship to the drive current, so the acoustic output will be compressed in relation to the electrical input, and harmonic distortion will also be generated. For high quality music reproduction these effects would be disastrous, but they can be highly desirable in their effect on the sound of electric or electronic musical instruments, and even some amplified acoustic instruments, blues harmonica being one obvious example.

It is also essential in the design and construction of loudspeakers for high fidelity sound reproduction that the cones behave as closely as possible to pure pistons, or that any break-up which does exist decouples sections of the cone in order to control the directivity at higher frequencies. Normally, for high fidelity drive units, either strong, heavy, straight sided cones are used (where the weight is also useful for lowering the resonant frequency), or deep profile cone shapes can be applied. If more sensitivity and an extended high frequency response are required, a lighter weight curvilinear cone shape can be employed. These forms are as shown in Figure 8.19. The latter shape facilitates the decoupling of the outer regions of the cone at high frequencies, leaving the central area and the dome as the principle radiating areas, thus improving the control of the directivity whilst still

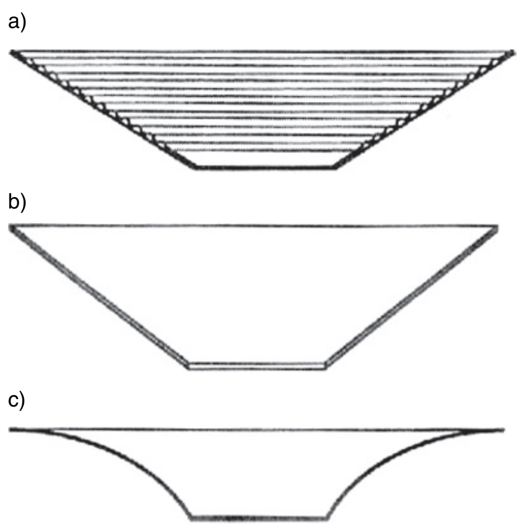


Figure 8.19 Cone profiles – they each serve best for different purposes. a) Ribbed, straight-sided – thickness and ribbing employed to increase rigidity. b) Deep, straight-sided – smaller apex angle employed to increase rigidity. c) Curvilinear – curved profile employed to aid decoupling of outer region of cone

preventing the break-up modes from circulating around the cone. At low frequencies, a straight sided, light-weight, shallow angle cone would not be desirable for high fidelity use, because it would lack the stiffness to act as a good piston. However, a straight sided, light-weight cone may be exactly appropriate for some musical instruments, where the extra sensitivity gives more volume per watt of amplifier power, and the colouration from the reduced stiffness can enhance the tonal response. Cone depth (apex angle) can also be chosen to move resonances into desirable regions. Light, shallow cones can exhibit rising, peaky, mid-range responses, which can be excellent for electric guitar amplification, despite being awful for high fidelity purposes.

What is more, high fidelity loudspeakers are generally intended to be used ‘sensibly’, but musical instrument loudspeakers may be purposely overloaded on many occasions to achieve certain sounds. For this reason, the suspension systems, as well as the voice-coil/magnetic-gap designs, may be used to limit the cone movements and offer some degree of self-protection from overload. Inevitably, this means the use of nonlinear suspensions, which add characteristic sound qualities that would be impossible to accept in a world of high fidelity, yet they may also add further characteristic tone qualities of a desirable nature to a musical instrument loudspeaker. In fact, once all of these ‘defects’ have come together in the right balance of qualities and quantities to produce their own, classic sounds, it can be very difficult for the drive unit designers to ‘improve’ the construction without losing the magic sound. Their sound can be something very difficult to define in technical terms, but the musicians can recognise even the slightest change from the original item. The Celestion company, for example, who produced 12 inch loudspeakers for the classic Vox (blue chassis) and Marshall (silver chassis) guitar amplifiers of the 1960s have had to find sources of exactly the same materials in order to make the same loudspeakers 40 years later. No amount of research and engineering or computer analysis has been able to achieve the same sounds from different designs. Conversely, their current *high fidelity* products are considered to be much superior to their products of the 1960 and 1970s, and none of the older designs remain in production.

8.6.1 Cabinet designs

In the glory days of amateur high fidelity, when almost all recordings were in mono, an enormous quantity of ‘patent’ cabinet designs were written about. Before the advent of stereo gave a whole new dimension to the spaciousness of the reproduction, many strange loudspeaker designs were offered in an attempt to bring more life to the reproduction, and to generally enhance the sensation of ‘being there’. Clearly, by the very fact that they were designed to *enhance* the reproduction, they could not, strictly speaking, be called high fidelity, but many of the designers claimed otherwise. Their point of view was that if the designs gave a greater *sensation* of being at the performance, then that in itself was a greater fidelity to the overall musical experience, even if it was not provable by objective measurement. The fact that many, current, portable music players have 'stereo enhance' and 'super maxi bass' facilities on them would seem to prove their case, at least when listening for pure pleasure is the goal, and when the basic reproduction is not the best that hi-fi stereo can offer. And it must also be said that many, modern, domestic surround systems are more likely to produce a pleasant sensation rather than anything which could be claimed to be a lifelike reality.

As stereophonic domestic music reproduction began to reach much better standards, in the late 1960s, many of the eccentric designs which had previously been offered began to disappear. In stereo, where the spacial enhancement was much less necessary, the strange designs only proved to be sources of annoying colouration, and the fact that they were not capable of producing true high fidelity became very apparent. However, not long after they had begun to fade from the hi-fi scene, some of them, such as the Karlson Coupler, from the 1950s, shown in Figure 8.20, began to be re-discovered by the makers of musical instrument amplifiers. In this new guise, where the concept of fidelity did not exist, the strange cabinets could freely and justifiably contribute their special sound enhancement characteristics to the realm of sound *production*. The original claim for the Karlson enclosure was that it de-tuned the normal resonance of the rear radiation path, making it responsive over a broad band, but objective measurement could largely only show that it produced response irregularities, and little else. It was also recommended for use with full-range

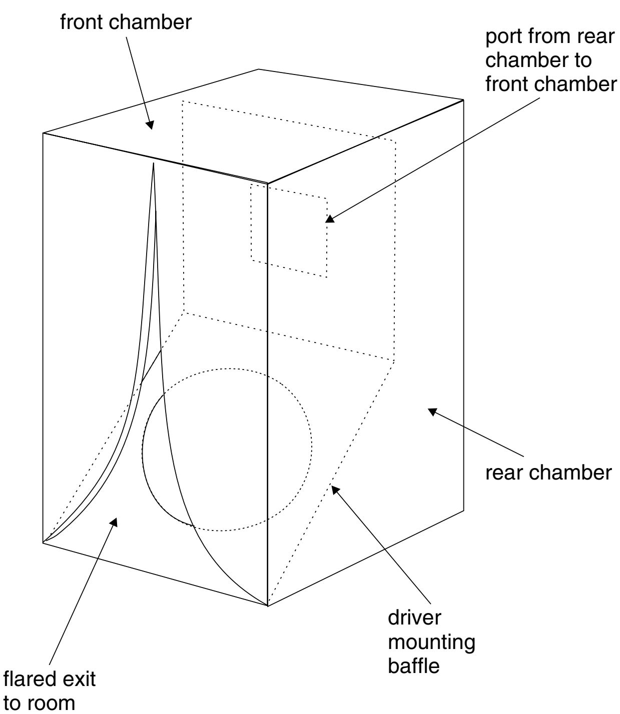


Figure 8.20 The Karlson Coupler, from the early 1950s

loudspeakers, but it is hard to see how severe colouration could be avoided due to resonances and reflexions in the cavity in front of the drive unit. It is unimaginable to be listening to modern day CDs via such an absurd ‘design’, yet in the 1970s it made a big come-back as a popular and successful bass guitar loudspeaker, where its own sound was deemed to be very favourable, and such enclosures continue to be used to this day for instrument amplifiers.

It is therefore very difficult to give specific guidelines for the design of cabinets for musical instrument amplification, because what *sounds* good *is* good; although it might be good for some players and their particular instruments but not good for others. It is also difficult to give target responses for overall performance, because good musical sounds, at their production stage, do not follow many set rules. Designers may well also break the accepted hi-fi standards when it comes to wiring the drive units in the cabinets, such as wiring woofers in series, specifically to *reduce* the damping and enhance the resonances and colouration. This idea may be further extended by using valve amplifiers without any negative feedback on the output stage. This raises the output impedance, and reduces the damping still further. Undersized output transformers may also be employed, to give rise to saturation and ‘overdrive’ effects at high levels. In fact, guitar amplifiers and loudspeakers are crafted more by artistry than by science, but so is the music which they help to create.

Summary

The recording chain from the musicians to the music buying public consists of five very different links. At each stage, the demands on the loudspeakers are different, and even if a perfect loudspeaker did exist, the circumstances of its application would render it non-optimal in many situations where specialised needs exist. Although the fundamental requirements are the same for the reproduction loudspeakers, the specific balance of characteristics can change to quite a large degree—very much so in terms of physical size. However, once we enter the world of musical instrument amplification, many of even the fundamental requirements change drastically, and what can be an absolute taboo for reproduction can be highly desirable for sound production, such as using 12 inch (300 mm) loudspeakers up to 6 or 7 kHz, for example. This not only goes for the loudspeaker drive units themselves, but also for the cabinets, the wiring, and the design of the amplifier output stages.

References

1 Newell, P., ‘Recording Studio Design’, Appendix 1, Focal Press, Oxford, UK (2003)
2 Walker, R., ‘The Selection of Loudspeakers for BBC Radio & Music’, Proceedings of the Institute of Acoustics, Vol 26, Part 8, pp 93-106, Reproduced Sound 20 conference, Oxford, UK (2004)

Bibliography

1 Colloms, M., 'High Performance Loudspeakers', 6th Edition, John Wiley & Sons, Chichester, UK (2005)
2 Fletcher, H., Munson, W.A., 'Loudness: Its Definition, Measurement and Calculation', Journal of the Acoustical Society of America, Vol 5, p 82, (October 1933)
3 Robinson, D.W., Dadson, R.S., 'A Redetermination of the Equal-Loudness Relations for Pure tones', Journal of Applied Physics, Vol 7, p 156, UK (May 1956)
4 ISO226, 'Normal Equal-Loudness Level Contours', International Standards Organisation, (1987)

Chapter 9

Subjective and objective assessment

9.1 The general situation

The human hearing system is quite extraordinarily sensitive and complex. It has a frequency range of around eleven octaves if one considers physical sensation as part of the process, and a dynamic range such that the lowest audible sounds have power levels of only 10^{-12} compared to the loudest sounds *before* the threshold of pain in the ears. That is a power ratio of one million, million times (one trillion in American English). At the lowest detectable sound pressure levels, the lateral movement of the ear drum, (or tympanic membrane) is less than the diameter of a hydrogen molecule, and if the average ear were only 9 or 10 decibels more sensitive, we would experience a permanent hissing sound due to the detection of the Brownian (random) motion of the air molecules. The signal processing of sounds by the brain is also a remarkably refined process. It is thus little wonder that when we reproduce music via the relatively crude devices described in Chapter 2 we are rarely fooled into believing that we are listening to the real instruments.

Nevertheless, back in 1990 David Moulton made the case that loudspeaker reproduction has now reached a stage where, at least with many musical styles, it should be recognised as something in its own right[1]. Music is now being created *on* loudspeakers for reproduction *by* loudspeakers, and much of this music exists in no other form. For many musical creations there was never, at any place or any time, a complete performance of the music as recorded. The late Richard Heyser, the 'father' of the 'Time Delay Spectrometry' measuring system said that in order to really enjoy stereo reproduction, one has to willingly suspend one's belief in reality. We must therefore ask ourselves if we are really trying to reproduce a sense of 'being there' at the original performance, or is there a 'being there' at the reproduction, which may be far more exciting than any live performance could ever be, because it must be accepted that some reproducible music simply could never *be* performed live. Are we now, as David Moulton asked, so accustomed to reproduction via loudspeakers that the loudspeakers, themselves, have become the greatest musical instrument of our time?

We seem now to have two separate outlooks on loudspeakers for music reproduction: 'the closest approach to the original sound', as the Acoustical Manufacturing Company put in their advertisements in the 1940s, or 'the best sound that we can possibly get', with 'best' being highly arbitrary. Nevertheless, it seems that from whichever viewpoint the subject is approached, the general requirements tend to be rather similar: wide bandwidth, low distortion, adequate sound pressure level, fast transient response, low colouration, and so forth. The degrees to which the levels of discrepancies of each aspect of the response are acceptable may vary with the musical styles and the listeners' preferences, but John Watkinson's point of view, that the only criterion we have for the accuracy of a loudspeaker system is the sensitivity of the human hearing system[2], seems to be quite valid. He went on to say that if a reproduction system is more accurate than the human hearing system's error detection threshold, then it needs no further improvement. Nonetheless, that is a difficult goal to achieve when we consider what was discussed in the opening paragraph of this chapter. As shown in Chapter 7, the human hearing system is dealing with wavelengths from as great as around 20 metres to as small as about 1.5 centimetres, a ratio of over 1000 to 1. By contrast, the eye has to deal with almost a one octave range of visible light spectrum, a wavelength ratio of less than 2:1.

Strictly speaking, we would only need one test for a perfect loudspeaker: its ability to perfectly reproduce a delta function supplied electrically to its input terminals. A delta function, otherwise known as a Dirac function, or impulse, contains all frequencies in a very fixed phase relationship, and its waveform is shown in Figure 9.1. Unfortunately, no mechanical system can start and stop instantaneously, so this waveform can never be perfectly reproduced. A delta function reproduced by a good loudspeaker is shown in Figure 9.2, and the degree of reproduction error is patently obvious. Unfortunately, we cannot glean all the information that we need from a visual inspection of the delta function response, so we tend to use a series of individual measurements which highlight particular aspects of a response. Some of them show behaviour in the frequency domain, whilst others show behaviour in the time domain. From them we can build up a picture of how a loudspeaker is responding to its electrical input stimulus, and shortcomings in the response can be assessed.

9.2 Test signals and analysis

The most well known aspect of any loudspeaker performance is the magnitude of the frequency response, which appears in just about every advertising leaflet for loudspeakers. In fact, the *full* frequency response also needs to show the associated phase response, and from the full response every linear aspect of a loudspeaker performance can be derived. A system may be said to be linear if the output contains no frequencies which do not exist in the input signal. A roll-off, or any other deviation from flatness in the frequency response, can be called a *linear* distortion. A system is said to be *non-linear* when frequencies exist in the output which were *not* present in the input signal. For example, if a pure sine wave were to be fed to the input terminals of a loudspeaker, and the measured output showed small amounts of response at twice the frequency and three times the frequency, then those additional frequencies would be the second and third harmonics of the input frequency, and the loudspeaker would thus be generating *non-linear* distortion; in this case harmonic distortion. If the loudspeaker

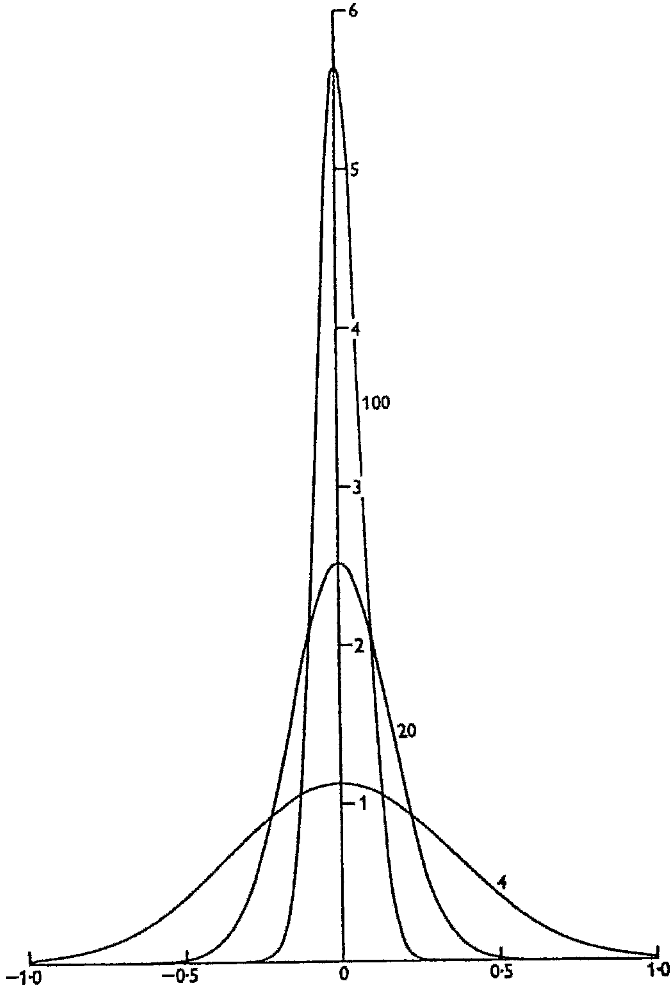


Figure 9.1 Evolution of the Dirac delta function (reproduced from Lighthill, 1964)
In a delta-function one can imagine the energy in a unidirectional signal being gradually narrowed, and each time that it narrows the amplitude increases until, in extremis, it becomes a pulse of infinite height and infinitesimal width

is fed with multiple input frequencies, anywhere from two upwards, then a non-linear system would also produce sum and different tones. In such a case, if the input were to be fed with 1 kHz and 4 kHz, for example, outputs would be noticed also at 1 + 4 kHz, or 5 kHz, and 1 − 4 kHz, or 3 kHz. The products can also further create their own sum and difference tones, and also inter-react with the original tones, producing frequencies such as 3 kHz + 4 kHz (7 kHz), 5 kHz + 1 kHz (6 kHz) and so forth. Whilst harmonic distortions in themselves are not necessarily unpleasant, because all musical sounds are rich in harmonics, the sum and difference products, known as *intermodulation distortion*, can be grossly offensive to the ear.

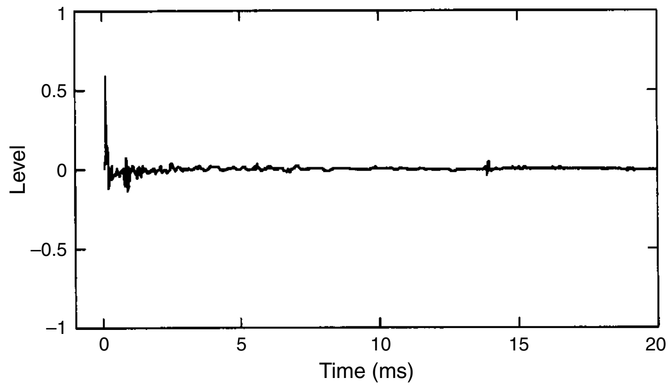


Figure 9.2 Loudspeaker reproduction of an impulse
The Dirac delta-function when reproduced by a loudspeaker inevitably becomes bi-directional and smeared in time, due to the imperfect reproduction

They may or may not coincide with musical harmonics, and they tend to build up into a noise-like signal which accompanies the music. Briggs, in his book Sound Reproduction, published in the 1950s, summed up the situation beautifully in a short quotation from Milton with which he introduced his chapter on intermodulation distortion – "... dire was the noise of conflict."[3] In fact, Gilbert Briggs was so disturbed about the problem of intermodulation distortion that he invited a more knowledgeable specialist, one N.C. Crowhurst, to write the chapter in the Third Edition of the book. In the Second Edition, published in 1950, Briggs had quoted Shakespeare in the chapter on intermodulation which he had written himself (Briggs, that is; not Shakespeare!):

> Find out the cause of this effect;
>
> Or rather say, the cause of this defect,
>
> For this effect defective comes by cause.
>
> Hamlet, Act II, Scene 2.

Over 50 years later, intermodulation distortion is still a significant problem, and we still have no simple way to measure it in an easily interpretable way which intuitively relates to all its audible implications.

Time domain representations of performance are less frequently published, and even less widely understood by loudspeaker users. Nevertheless, they are essential aspects of the analysis of loudspeakers because the phase response, which in concert with the amplitude response is sufficient to define *all* linear aspects of performance, is very non-intuitive. Phase is very abstract; it is a relationship between things, and cannot exist alone. Time domain representations include waveform responses, such as Dirac (delta) and Heaviside function responses (impulse and step-function responses), acoustic source plots which show group delay against frequency, and cepstrum plots, which are useful for finding reflexion and diffraction problems in otherwise complex signals. It is perhaps useful, now, to look at all of these representations and their implications step by step.

9.2.1 Frequency response plots

Figure 9.3 shows a frequency response plot of the axial response of a loudspeaker, derived from a pink noise signal and a dual channel analyser which compared the direct electrical input with the acoustic output, in an anechoic chamber via a measuring microphone. Both the amplitude and phase responses are shown, and it can be seen how there can be no deviation from a straight line in either curve without a corresponding deviation in the other. This is the characteristic of a minimum phase response as discussed in Chapter 7.

Figure 7.19 shows clearly how the amplitude response plots change as equalisation is introduced into a system. Almost everybody reading this book will understand the significance of the amplitude part of the plot, and how ideally, for perfect reproduction the line should be as flat as possible and as wide as possible. From an objective point of view the magnitude plot tells the engineers much about the uniformity, or otherwise, of the pressure amplitude response with respect to different frequencies, and consequently, if those design aims have been achieved.

Subjectively, it has long been considered that the magnitude of the pressure amplitude response (the frequency response in everyday language) is the most significant measure of a loudspeaker's performance, and yet

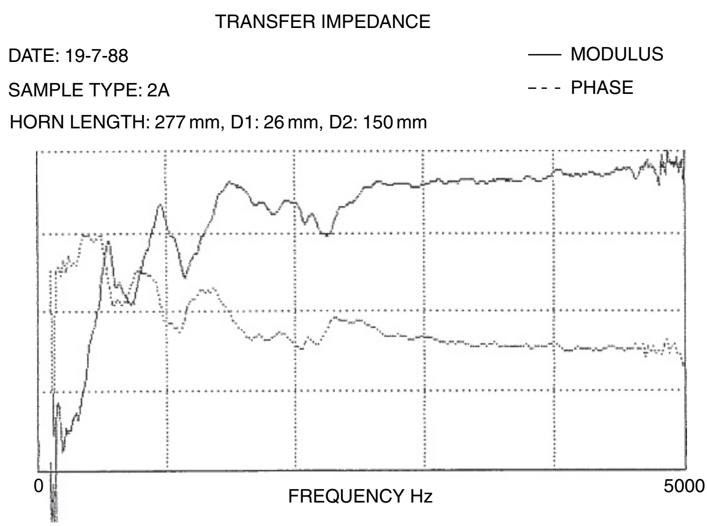


Figure 9.3 A full frequency response

The frequency response of a mid-range horn loudspeaker

The vertical divisions represent 10 dB in amplitude or π radians in phase (360 degrees/ 2π = about 57 degrees). It can be seen how every change in the upper, pressure amplitude plot is accompanied by a corresponding deviation in the lower, phase plot no loudspeakers are truly flat. Smooth deviations from flatness are generally acceptable, and are easily grown accustomed to by listeners who are familiar with the loudspeakers. Many mastering engineers consider extreme flatness to be nice if it can be achieved without other compromises being made, but not essential, because a smooth frequency response deviation is a linear distortion which can be compensated for both mentally and electrically without having to pay the penalty of side effects. On the other hand, *abrupt* changes or irregularities in the frequency response are definitely undesirable. They not only introduce colouration which will only affect music with dominant signal content in the region of the irregularity, but they almost always imply that something else is wrong in the system, and that the abrupt changes are only side-effects of whatever that something else may be. The physics of loudspeaker design really does not permit abrupt changes in frequency response without other consequences, so smoothness is a very desirable characteristic of a curve, perhaps more so than general flatness with an abrupt change somewhere. Abrupt changes are usually accompanied by time response anomalies.

Figure 9.4 shows a series of off-axis plots, both in the vertical and horizontal planes. Their significance is twofold. Firstly, any persons listening off-axis (that is, away from the line which is typically perpendicular to the centre of the face of the loudspeaker cabinet), will hear a response which is characterised by the respective plots in the horizontal and vertical directions. [Note that in a multi-way loudspeaker system it can be very difficult to determine exactly where on the face of the cabinet is the exact acoustic centre of propagation.] The acoustic centre of the loudspeaker shown in Figure 8.6(a) is rather obvious, but the acoustic centre of the loudspeaker shown in Figure 8.6(c) is not obvious at all. The second significance of off-axis frequency responses is that any reflexions which return to the listening position from off-axis radiations will be affected not only by the frequency response of the reflective surface, which for a plastered brick wall would be acceptably flat, but also by the frequency balance radiated in that direction from the loudspeaker. Many, large, multi-way monitor systems do exhibit poor off-axis responses, a price sometimes paid to allow for other design benefits, which is one reason why so many professional control rooms have rather absorbent side-walls that will not return reflexions to the listening position, or geometry that tends to direct this energy into absorbers after the first bounce.

Until now we have been looking at plots of anechoic responses, but once the room response becomes involved in the proceedings, the loudspeakers-only responses tend to get corrupted. Figure 9.5 shows the response of a loudspeaker in an anechoic chamber, and Figure 9.6 shows the response of the same loudspeaker placed on top of a mixing console in a typical small control room. The changes can be seen to be gross, but a measuring microphone is not a pair of ears and a brain. The ability of the ear to know when it is still receiving a smooth direct sound, despite all the chaos surrounding it, is something quite impressive. Moreover, human beings live and work in reflective environments, so if the response corruption is not excessive the irregularities are accepted as what they are – room effects. However, if the room reverberation or response decay time becomes significant, it can mask low-level detail in the sound, so the room decay time is an

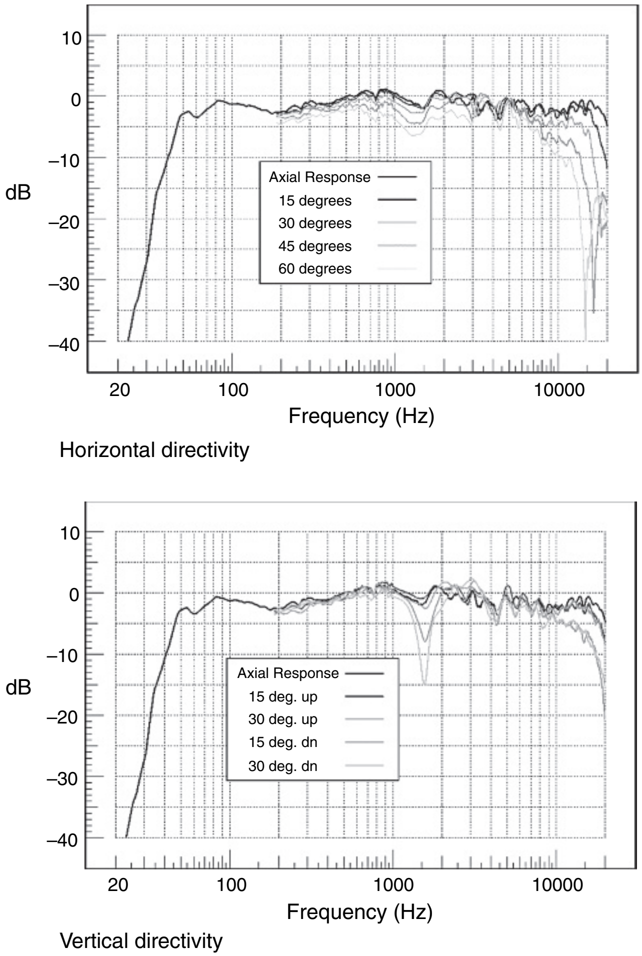


Figure 9.4 Horizontal and vertical directivity plots, showing the responses on-axis and at various angles off-axis

important factor where detailed monitoring is required. The acceptability or otherwise of room effects on loudspeaker responses may depend on the principal reason for listening, such as to the recording quality, or to the performance.

The phase of the frequency response is something which is more useful in engineering processes rather than as something that relates to clearly audible effects, except to say that gross phase errors will be discernible as time response effects.

9.2.2 Waterfall plots

A good general grasp of the performance of loudspeakers can be made from a visual assessment of their waterfall plots. These show a three

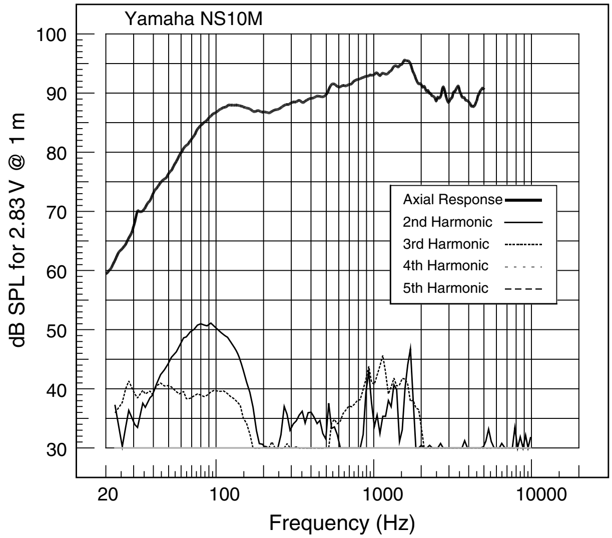


Figure 9.5 Anechoic measurement of frequency response

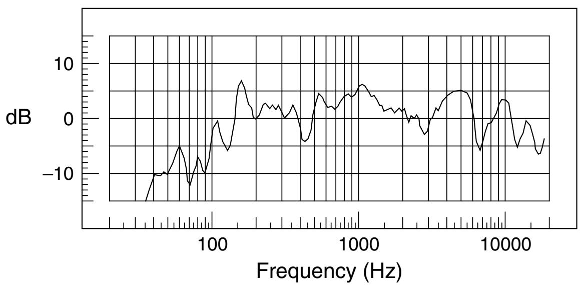


Figure 9.6 Console-top measurement of frequency response

The same loudspeaker as measured in Figure 9.5, but on top of a mixing console in a typical home-studio control room

axis representation of time against frequency against level, as shown in Figure 9.7. In effect, a waterfall plot is a series of pressure amplitude plots taken a few milliseconds apart and displayed by superimposition by computer after the input stimulus has been stopped. An ideal loudspeaker system, which could reproduce an accurate delta function, would show only the top line of the plot – the zero milliseconds line – because the decay would be instantaneous. As explained earlier though, *no* mechanical system can start and stop instantaneously, and the waterfall plots show how the decay takes place, frequency by frequency. A large selection of waterfall plots are shown in Figure 11.1 which are very informative. They show, without any shadow of a doubt, why virtually all loudspeaker systems sound different to one another: none of the responses decay in the same way.

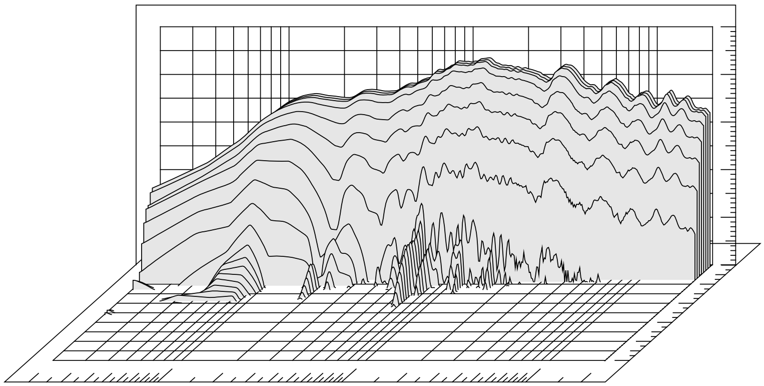


Figure 9.7 A waterfall plot
The cascading lines are the frequency (pressure) responses at 2 millisecond intervals after the cessation of the input stimulus (simulated) at time = 0 milliseconds. Various resonances can be seen continuing to ring until about 40 milliseconds

Subjectively, a long decay tail sounds exactly like what it is; resonance. Figure 9.8 shows the waterfall plots of two different loudspeakers of more or less the same size, a) being a sealed box and b) a reflex enclosure. The box concepts are dealt with in Chapter 3, and the implications are dealt with in Chapter 11, but the general tendency is for the loudspeaker with the faster decays to sound tighter in the bass; the longer decays sound rounder. Balances between bass guitars and bass drums tend to be more reliable and compatible with a range of other loudspeakers when mixed on faster decaying loudspeakers, because the resonances of the individual instruments are heard in a more realistic proportion with each other. Resonant loudspeakers will add their own characteristics to the sound, so it becomes difficult to judge exactly what part of the bass sound is due to the instruments alone, and which part is due to the loudspeakers. Relatively few top mastering engineers use reflex cabinets, and those who do use them tend to use relatively large cabinets with resonances very low down in the audible frequency range. It is believed by the authors that the long-lived and widespread use of the Aurotone 5C and Yamaha NS10 loudspeakers for rock music mixing was largely because of their rapid decays which were uniform with frequency. Even though their frequency responses in anechoic chambers were far from flat, they tended to flatten in the bass region when placed on top of mixing consoles due to the reduction in the radiating angle. Nevertheless, they were still rather bass light, but the accurate time response meant that the relative levels of bass instruments were not confused by resonances, so if the mix was deemed to be bass heavy on larger loudspeakers, it was a relatively simple matter to equalise the mix with a reduction of bass frequencies and still maintain the instrumental balances. This process is often not possible when the balance *between* the instruments has been confused by loudspeaker resonances. This is especially problematical when the resonant frequencies of the loudspeaker reflex port timing is above the 41 Hz fundamental frequency of

a)

b)

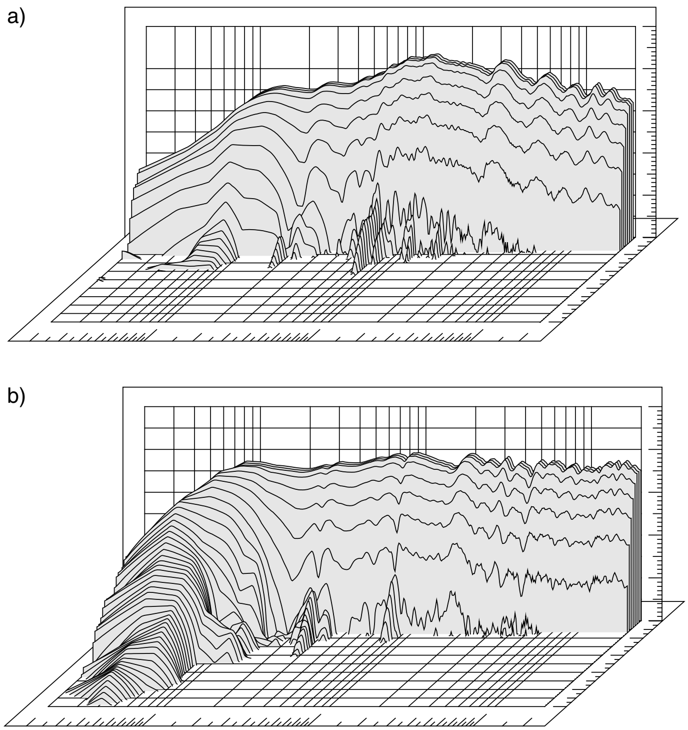


Figure 9.8 Waterfall plots of two loudspeakers. a) A small sealed-box loudspeaker with a relatively rapid delay. b) A small reflex loudspeaker with a considerably longer decay at low frequencies than at mid and high frequencies

the lowest note on a conventional 4-string bass guitar or double bass (the E-string).

Whether such instruments are recorded flat, or stylised by the use of equalisation and signal processing, the final sound must be judged to be suitable for reproduction via a wide range of loudspeakers. This is clearly a difficult task with the situation that exists. The responses shown in Figure 11.1 all represent loudspeakers designed for professional music recording and all were measured in the same anechoic chamber. The situation in domestic reproduction rooms and with non-professional loudspeakers is obviously more diverse. Whilst research has been done both by JBL and Genelec on finding the mean *frequency response* of a wide range of listening conditions, no such work seems to have been carried out to find the mean, most representative *waterfall plot*. It transpires that the mean frequency response (or at least the pressure response) is substantially flat, and many critical listeners also consider that loudspeaker decay times should also be uniform with frequency. Mastering engineers certainly seem to choose predominantly low decay-time loudspeakers – a choice which they have mostly made by ear – as they have found them to aid in the making of more consistent decisions, but the lack of any industry-wide guidelines on response decays is a great pity.

9.2.3 Harmonic distortion

Harmonic distortion plots cannot be generated from noise signals nor complex waveforms. The only practical method of producing a plot such as the one shown in Figure 9.9 is to send to the loudspeaker a sine wave tone which continuously sweeps up in frequency, covering the entire audible frequency range, and use a set of tracking filters which are spaced at one, two, three and four octaves above the principal tone. The outputs of these filters are the second, third, fourth and fifth harmonics of the input tone, and their relative levels can be given in percentages or decibels, which are compared below:

0 dB	100%
−10 dB	30%
−20 dB	10%
−30 dB	3%
−40 dB	1%
−50 dB	0.3%
−60 dB	0.1%
−70 dB	0.03%

Hence, for example, in a loudspeaker which produces 0.3% of harmonic distortion, the distortion products would be 50 dB below the main signal. Another method, perhaps more commonly used, it to measure the level of a single tone, repeated at several different frequencies, and then to measure what remains each time when the frequency of the drive tone is filtered out of the response. This yields a 'total harmonic distortion plus noise' or THD + N figure for each single frequency that is measured, but THD + N

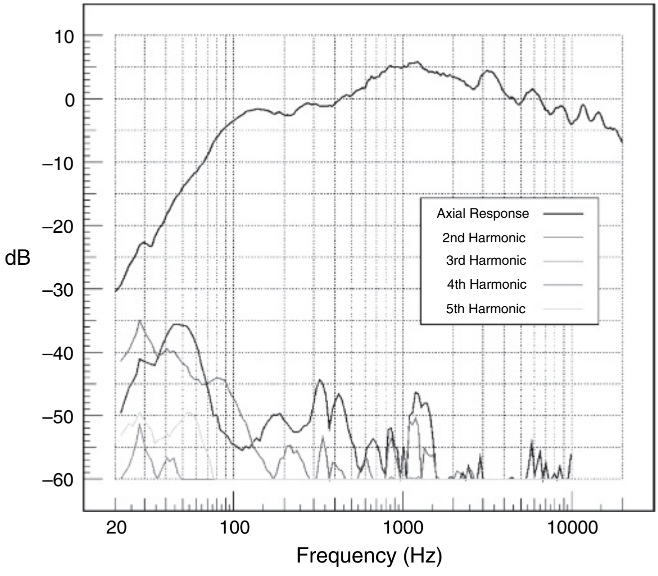


Figure 9.9 On-axis pressure amplitude and harmonic distortion

has consistently failed to relate well, subjectively, to the perceived sound from loudspeakers. Below 50 Hz it seems to be very doubtful that second and third harmonic distortion levels as high as 5% are audible, and it seems questionably whether levels as low as 0.25% are audible at any frequency. The 'maximum operating level' for magnetic flux on analogue tape recorders was set for many years around the 3% distortion level, and some really beautiful sounding recordings were made on those machines. Furthermore, despite the fact that electronic systems, such as amplifiers, with the above levels of distortion could not even be considered for high fidelity use, they may, in some circumstances, make tonally rich sounding guitar amplifiers; so they may not be *accurate*, but they are not necessarily unpleasant sounding. There is also an enormous range of microphone preamplifiers on the market, all sold on the basis of their characteristic sounds, which effectively means their characteristic distortions. These distortions are considered *desirable* by many people during the recording process, although for monitoring purposes it is obviously undesirable to colour the sound or the concept of monitoring the *recording* would not be valid. Nonetheless, and again as mentioned in Chapter 6, such distortions are also considered to be desirable for musical instrument amplification, so we therefore need to face the problem of deciding what level of harmonic distortion is accepted as 'non-intrusive' for each piece of equipment in turn.

The problem with harmonic distortion as a *quality* measure in itself is that harmonics of a low order (2^{nd}, 3^{rd}, 4^{th}) can actually be quite pleasant sounding, and not unmusical at all. All instruments produce large quantities of harmonics, sometimes even more than their fundamental tones, so the question was often asked as to how the ear could detect 0.2% of harmonic distortion from an amplifier reproducing an instrument whose tone was itself perhaps 80% or more of harmonics. The answer must lie in the differences in the *mechanisms which produce the harmonics*, and in what other ways those mechanisms manifest themselves.

A musical instrument produces harmonics from the break-up of its parts into separately resonating sections, which vibrate independently whilst they also vibrate as part of the whole. Figure 9.10 shows a representation of a string breaking into second, third, and fourth harmonic modes, and Figure 9.11 shows the vibrational patterns of a metal plate. Note the perfectly symmetrical behaviour. A harmonic analysis of the sound would be likely to show *only* frequencies which were harmonically related to the fundamental resonance.

When an amplifier produces harmonics, the mechanisms involved are totally different. Harmonics are produced as a result of the transfer function of the amplifier being non-linear, as shown in Figure 9.12, and by other means which are totally alien to natural vibrations, such as the crossover distortion mentioned in Chapter 6. Loudspeakers, also, behave entirely differently to either musical instruments, strings or plates because, at least for sound *reproduction* purposes, they are usually designed *not* to break up into separately moving sections, and they are unlike amplifiers because their moving parts have mass, and hence momentum when in motion. They also have non-linear stiffness in their suspension systems, and, amongst other things, they may have non-linear Bl products, where the drive force is not uniform because of the inconsistent relationship between the static

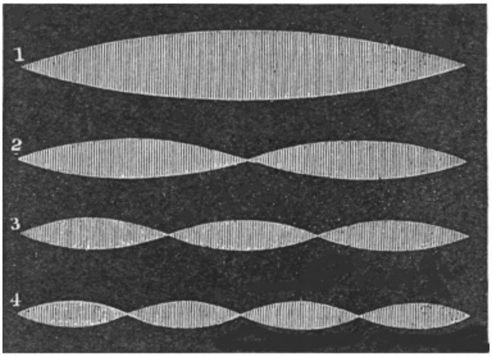


Figure 9.10 Vibrational modes in strings

Representation of a string vibrating, and showing how it can break up into multiple segments at harmonic intervals in response to stimuli at those frequencies

magnetic field (B) and the length of coil (l) when being driven by the voice coil.

In fact, harmonic distortion is not a good measure of such subjective subtleties. Some loudspeakers can have 10 dB of difference in harmonic distortion levels yet they may sound very similar, or vice versa. During listening tests carried out by the authors in 1989[4], in an attempt to group according to sonic similarity a selection of twenty mid-range drive units, harmonic distortion performance failed to show any relationship to the pattern of similarity groupings. The fact is that harmonic distortion is really the benign face of non-linear distortion, it is usually the intermodulation distortion which really offends the ear.

9.2.4 Intermodulation distortion

This subject has been investigated in depth by Czerwinski, Voishvillo and their co-investigators who have tried to make representations of analyses which relate measured intermodulation products with sonic perceptions[5,6]. The problem with measuring these distortions is that they are so dependent upon circumstances that it has always been difficult to define them. Intermodulation distortion can change dramatically with level, with the frequency range of the music, with the crest factor of the music (the peak to average relationship) and with many other parameters. Generally, the more complex the musical signal, the more offensive is the intermodulation, as every frequency interacts with every other frequency and with the products of the intermodulation, which in turn intermodulate with themselves. For this reason, a loudspeaker system may sound totally acceptable on relatively simple musical signals, but may sound rather unpleasant with an orchestral crescendo or a heavy concentration of guitars. At the same time, it may well be the case that the perception of the purely harmonic distortion, even up to the higher harmonics, *if* the intermodulation distortion could be separated out, could be totally inoffensive. However, the harmonic and intermodulation products cannot be separated, because

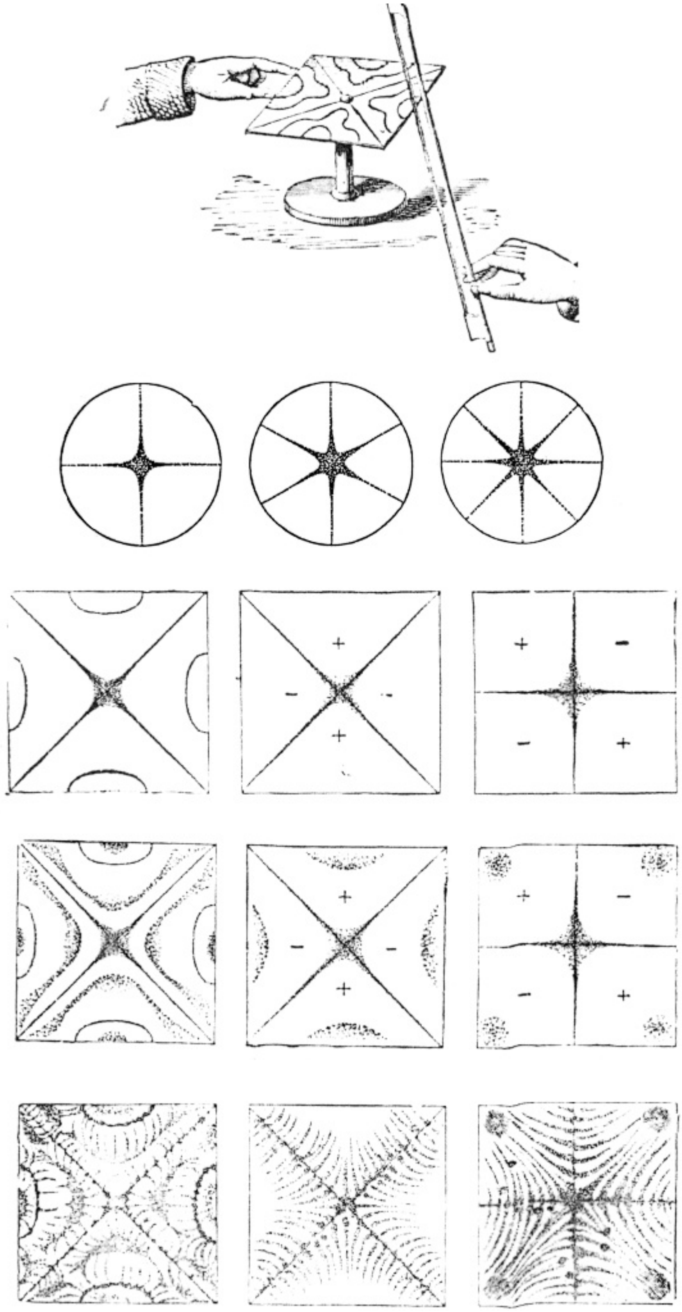
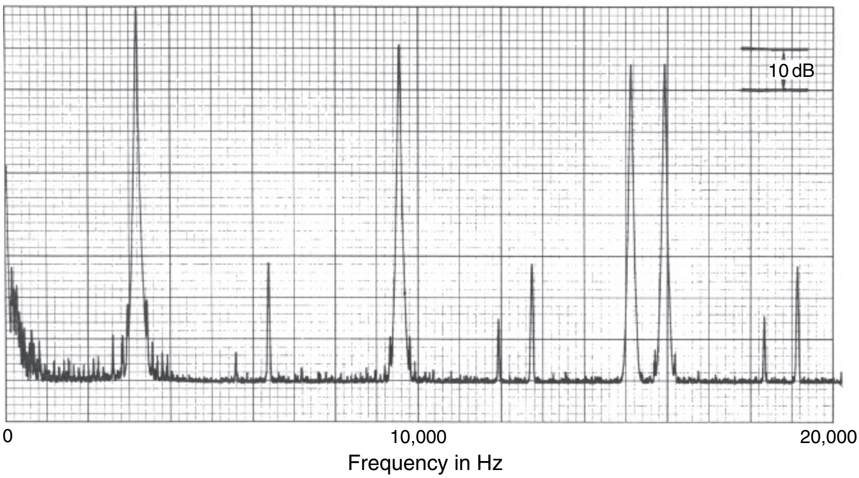


Figure 9.11 Vibrations in plates
An insight into diaphragm break-up. Nineteenth century experiments on vibrations in metal plates [From 'On Sound' by John Tyndall, Longmans Green & Co, London, (1895). Reprinted by Dover Press – highly recommended reading, and still in print]



IM Spectrum, industry standard amplifier using quasi-complementary circuitry

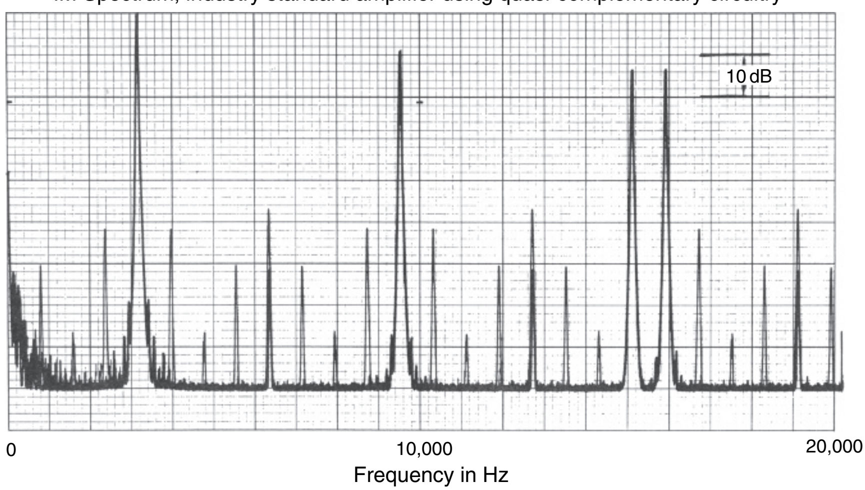


Figure 9.12 Non-linear transfer functions
The two transfer functions show very significant differences in intermodulation distortion in two amplifiers whose harmonic distortion figures are very similar to each other. The four highest peaks are the four frequencies of the drive signal. All the other spikes are distortion products of intermodulation

they are products of the same non-linear processes. Nevertheless, it is unfortunately misleading that the harmonic distortion, which can easily be measured, should so frequently get blamed for the undesirable sounds which are really a result of the intermodulation distortion, which cannot be so easily quantified. As no simple relationship exists between harmonic and intermodulation distortions, neither one can be extrapolated from the other.

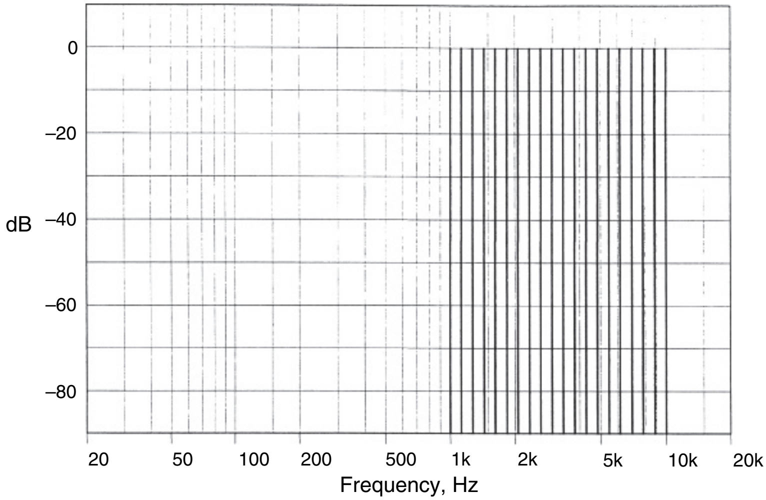


Figure 9.13 Spectrum of a 20-component logarithmic multi-tone signal

Figure 9.12 showed the greatly different levels of intermodulation *and* harmonic products of two amplifiers whose harmonic distortion products, alone, were measured to be very similar. Traditionally, intermodulation has been measured by pairs of tones, say 1 kHz and 5 kHz, but the results have never been particularly representative of any sonic characteristics of the systems under test. The tests shown in Figure 9.12 used four input tones, however, multi-tone signals using ten or twenty frequencies, specially chosen give rise to the widest spread of intermodulation products, can give a visual display of results which intuitively relate much better to what is heard. A twenty tone spectrum is shown in Figure 9.13, and its corresponding waveform is shown in Figure 9.14, which looks quite typical of a musical signal. In fact, statistically, it is also very representative of a real musical signal[7]. The response of a bass driver to a ten tone signal is shown in Figure 9.15[7]. The reason why intermodulation distortion is so audibly offensive can clearly be seen from this graphical presentation. Bear in mind that a distortion-free system would exhibit only the ten vertical lines of the stimulus signal, similar to the twenty, clean lines shown in Figure 9.13. The mass of sum and difference tones shown in Figure 9.15 are like a noise signal which changes dynamically and spectrally according to the stimulus. On a music signal it would be heard as a loss of transparency and openness in the sound, and a loss of low level detail. If the density of intermodulation products from only ten sine-wave tones is as high as shown in Figure 9.15, then it is easy to appreciate that a complex musical signal would produce an underlying, signal-related hash that could take the sweetness out of the music and mask room-sounds.

The fact that intermodulation distortion is the real enemy of both musicality *and* fidelity has been known since the very early days of loudspeakers[8], but it is still so hard to quantify it in any meaningful way simply because it is dependent upon so many dynamic factors, and its

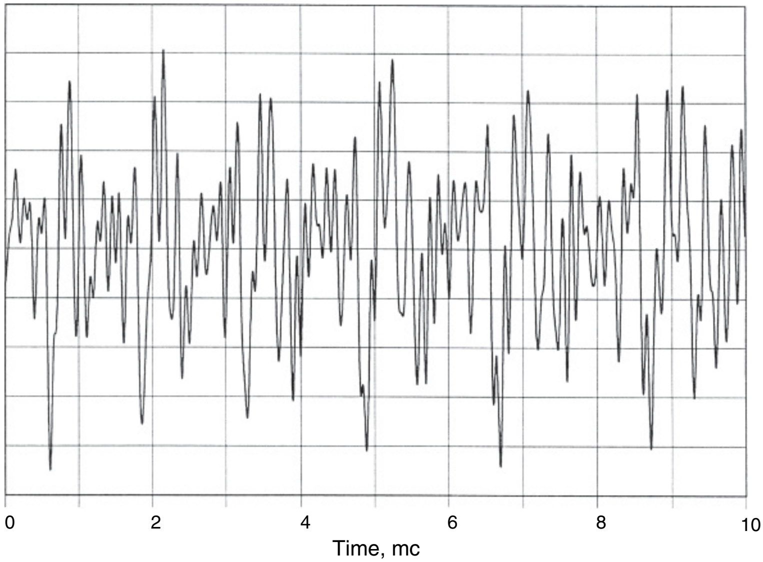


Figure 9.14 Waveform of a 20-component logarithmic multi-tone signal

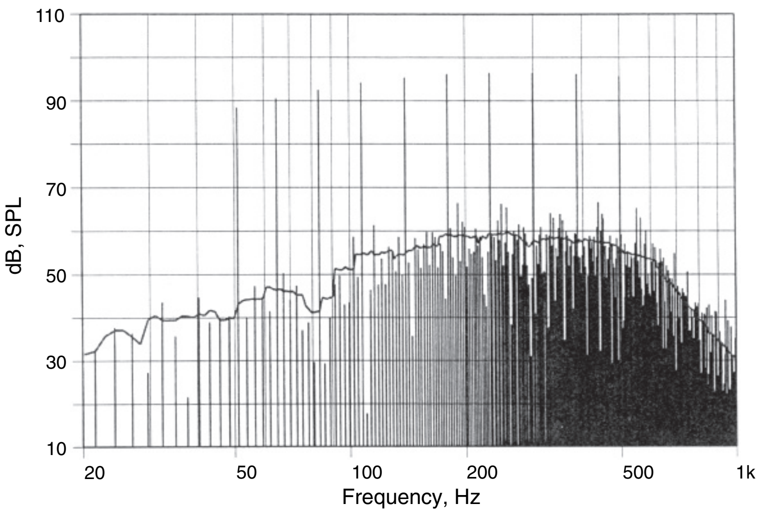


Figure 9.15 Sound pressure reaction of a loudspeaker (long coil, short gap) to multi-tone stimulus. The peak level of the input signal corresponds to X max = 4 mm. The solid curve shows the level of distortion products averaged in a one-third-octave-wide rectangular sweeping window

subjective offensiveness is signal dependent to a much greater degree than is the case for harmonic distortion. The fact that the two types of distortion share the same origins is easily demonstrated by the use of two tones, one fixed in frequency and the other variable. If the two tones are different, the spectral lines on a frequency analysis would show the harmonics, *plus* the sum and difference tones. If the variable frequency source were to be swept to the same frequency as the fixed tone, the pattern of modulation products would change until with the two tones at the same frequency, only the harmonics would be evident. However, where intermodulation is concerned the products are dependent upon so many factors that no truly meaningful figure of 'merit' has been devised to unequivocally define intermodulation distortion performance. To quote from Czerwinski et al, "High-order nonlinearity is very sensitive to the level of the input signal. An increase in input signal which produces a negligible effect on low-order [harmonic] products can wake up the 'evil forces' of nonlinearity, releasing an unfathomable number of high-order intermodulation product 'piranhas' to tear the flesh of the reproduced sound to pieces"[6].

It has puzzled many people for many years why relatively low levels of high-order *harmonic* distortion – 5^{th}, 6^{th}, 7^{th} etc – have been associated with poor sound quality. The principal explanation has been that the higher harmonics are not musically related to the signal, but the very low levels of these distortions have not corroborated this idea when they have been added artificially to sine waves, where they have tended to be *inaudible* at levels which prove to be offensive on music. Czerwinski et al offer the explanation that the low levels of high-order harmonics are, in fact, just tips of high-order intermodulation distortion 'icebergs'. It obviously does not bode well to be metaphorically sailing amongst icebergs in a sea full of piranhas! But that has been the reality of intermodulation distortion – a largely hidden, unquantifiable, yet dangerous enemy.

Distortion mechanisms which give rise to similar levels of harmonic distortion may yield greatly differing levels of intermodulation distortion, and this fact is surely at the root of the long acknowledged lack of any robust correlation between harmonic distortion measurements and subjective audio quality. Having said that, it is obvious that a loudspeaker producing 75% of total harmonic distortion (THD) at 1 kHz would not be considered to be high fidelity, but once we get into the low single figures at low frequencies, or below 0.5% at higher frequencies, then a loudspeaker with 10 dB less THD than another may well not guarantee that it would sound more musically accurate, even when their linear parameters were relatively similar. On the other hand, multi-tone intermodulation distortion presentations *have* begun to reveal good correlation between pure-sounding loudspeakers and clean-looking displays.

9.2.5 Delta-functions and step-functions

A delta-function is shown in Figure 9.16. It is a unidirectional impulse containing all frequencies, and of infinitesimal duration. The response to a delta function defines the full frequency response of any linear system. Mathematically speaking, the delta function is the derivative of the Heaviside function, also known as the step-function. The main problem with using a delta function (also known as a Dirac function, or impulse) in *acoustic* measurements is that it has very little low frequency content; having a spectrum rising by 3 dB per octave, like white noise. This results in a tendency towards poor signal to noise ratios at low frequencies, where

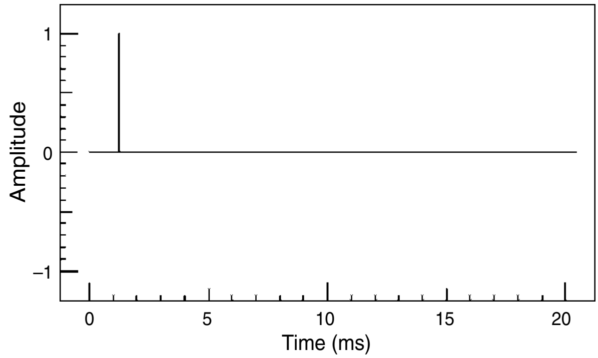


Figure 9.16 A Dirac delta-function, or impulse

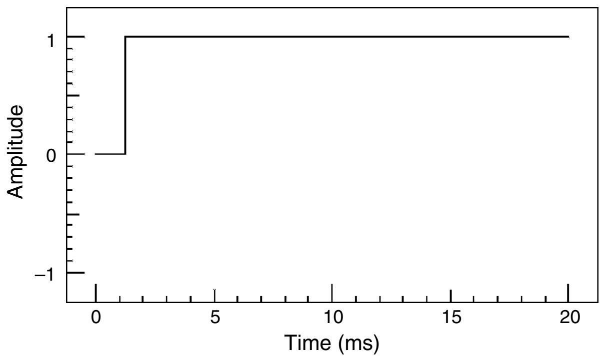


Figure 9.17 A Heaviside step-function

air conditioning noise, ventilation noise and traffic rumble can prejudice the low frequency response accuracy of the acoustic measurement. The step-function (or Heaviside function), shown in Figure 9.17, contains much more low frequency energy. The fact that by processes of either integration or differentiation, either one can be transformed into the other makes the step-function a better option for acoustic measurements, even if it is the impulse response that is ultimately required. Numerous step-function responses are shown in Chapters 10 and 11.

A simple circuit for a rudimentary step-function generator is shown in Figure 9.18. If this signal is applied 15 or 20 times to the input of an amplifier (sufficient to allow for an averaging process to disregard extraneous noises) with about 5 seconds each of 'on' and 'off' time, then by means of the FFT processing of a dual channel recording of the direct electrical output of the box and the loudspeaker output via a measuring microphone, the full frequency response of the transfer function of the system can be obtained. The pressure amplitude response, phase response, waterfall plots, acoustic source plots, cepstrum plots, impulse response and waveform response can all be derived from the step-function. Listening to

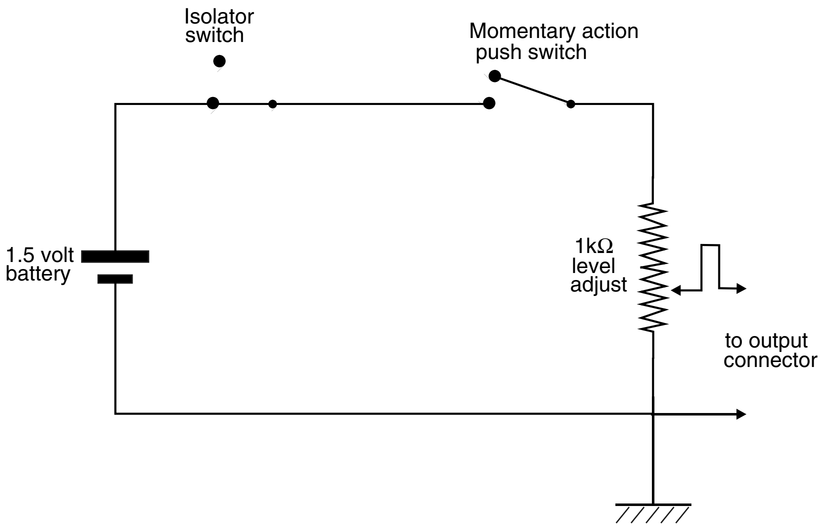


Figure 9.18 A step-function generator
This circuit will emulate quite well the waveform shown in Figure 9.17. For low impedance loads, such as direct connection to loudspeaker drivers, the potentiometer should be set to maximum. A good quality potentiometer should be used

the DC ‘thuds’ can also be quite revealing. A highly damped loudspeaker in an absorbent room will produce a very ‘tight’ impact. In many small control rooms, using small, reflex loudspeakers, the step function tends to sound like an ‘ideal’ bass drum – round and warm, yet solid. This exposes a dangerous situation for making decisions about bass drum/bass guitar/bass synthesiser sounds, because it suggests that much of the perceived sound is likely to be that of the loudspeaker and/or room, and that the ‘great’ sound is not on the recording. To get a better idea of what the step function really sounds like, it can be listened to on a pair of good quality headphones if a reference in anechoic conditions and via fast loudspeakers is not available.

The step function, used in this way, will also expose resonances from things such as open ended cable tubes in the control room, tubular steel mixing console frames, fire extinguishers, furniture resonances, window pane resonances and many other problems that ideally should not be in a critical listening environment. Figure 9.19 shows the response of a control room with open cable tubes in the floor. Although not obviously audible on a pink noise signal, the step source rendered their presence plainly audible to anybody in the room.

If a battery is used as a step source by coupling it and decoupling it directly to a passive loudspeaker system the effects will not be symmetrical, because during the ‘on’ phase, the battery will be connected across the loudspeaker terminals, and its low internal resistance will damp the loudspeaker resonance. On the other hand, when the battery is disconnected, the loudspeaker input terminals will be left open circuit, so the voice coil(s) of the low frequency driver(s) will be left unterminated and free to resonate. The sound of the on and off cycles may therefore sound,

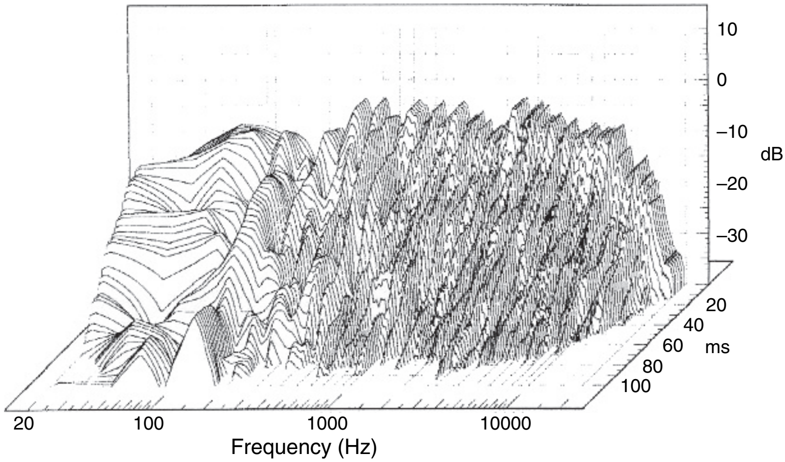


Figure 9.19 Cable tube resonances exposed by a switch box as described in Figure 9.18. Resonances are clearly visible at about 70 Hz and 110 Hz due to open cable tubes in the control room floor. The resonances were clearly audible on step-function excitation

and measure, quite different. However, when the signal is supplied to the loudspeaker via a power amplifier, the low impedance of the output is permanently connected across the loudspeaker terminals, so the positive and negative signals should be much more similar. Bear in mind though that if a $1\frac{1}{2}$ volt battery is used as a step source and connected to the input of a power amplifier for loudspeaker testing purposes, the amplifier should have a flat response down to DC, or the amplifier's roll-off would affect the loudspeaker's true low frequency response.

It should also be borne in mind that the human ear does not always hear the positive and negative pulses in the same way due to its own polarity asymmetry. For this reason a positive-going output from a mixing console or other music source should produce a forward (towards the listener) movement of the loudspeakers diaphragm(s). Although the effect is subtle, if this 'absolute phase' connection is not respected it can give rise to altered perception of the music.

And beware! Not all loudspeakers or drive units *or* amplifiers give a positive-going output from a positive-going input at their red terminals or signal 'hot' input connectors. Many JBL loudspeakers still follow an older standard where the application of a positive voltage to the red terminal causes a movement of the diaphragm *inwards*, towards the magnet. It is difficult for long-established manufacturers to change protocols without creating havoc in their replacement parts markets. Quad and Tannoy are other famous brands who have used this reversed standard, and some eastern manufacturers have copied the lead of such exalted names. As a general rule it is important to either check the manuals and data sheets or physically test any unfamiliar equipment.

The original reason for this old standard of absolute phase was to maintain the phase of the source. For example, a voice pronouncing a 'p' would expel air from the mouth, which would push the microphone diaphragm *inwards*. It therefore followed that in order to maintain the positive pressure in the listening room, the loudspeaker should go *outwards* (i.e. in anti-phase to the microphone), which is perfectly logical! Some older designs of amplifiers also reverse polarity from input to output, and sometimes this was done to 'correct' older loudspeaker standards, so it is always best to check any unknown device for its relative polarity of input and output.

Back on the subject of delta functions and step functions, it is important to note that they should be applied conservatively in terms of level, because subsequent FFT (Fast Fourier Transform) analysis will break down in the presence of distorted signals. The peak of a delta function is so narrow that it can clip without apparently affecting its shape. A clipped spike may look very similar to an un-clipped spike on an oscilloscope, but their frequency contents would be very different. Delta functions are also not very easy to generate in a pure form, but the response can be calculated via the inverse FFT from a white or pink noise signal. White noise also suffers from the poor signal to noise ratio at low frequencies, because of its 3 dB octave (10 dB per decade) rising response. That is, the level at 20 kHz would be down by 3 dB at 10 kHz (or 10 dB down by 2 kHz). At 20 Hz; which is 10 octaves (or 3 decades) below 20 kHz, the response would be 30 dB down. Pink noise, with a flat power spectrum, is a more practical alternative, and from a dual channel recording (one channel straight from the source and the other via a measuring microphone) exact replicas of the step-function and delta-function waveforms can be derived via the inverse FFT. About 2 minutes of noise should be recorded to allow the averaging-out of any extraneous noises.

It is remarkable to think that Fourier, the French mathematician, calculated this relationship in the early 19th century, around 1807. In fact, he was *so* far ahead of his time, and the means of proving the concept practically were still over a century away, that his teacher and mentor, the renowned mathematician Laplace, until his death refused to believe that Fourier's work on these transforms could be correct. Laplace even went so far as to try to discredit Fourier over this issue, but powerful computers have proved his concept beyond question.

Dirac and Heaviside also derived their functions long before the days of transistors or digital computers. One cannot help but wonder at the power of such brains – or whatever it was that they were taking! A selection of step-function responses are shown on short time scales in Figure 9.20. The variability of transient responses should be evident from inspection of the plots. The more similar the waveform is to the electrical input waveform of Figure 9.17, the better will be the transient response of the system.

9.2.6 Acoustic source plots

Whereas the delta and step function responses look at representation of time against amplitude, another way of looking at the time response of a signal is to look at it in terms of time against *frequency*. This, of course is what is displayed on the horizontal plane of a waterfall plot, which shows the response *decay* of a system. We can look at the system *attack* in this

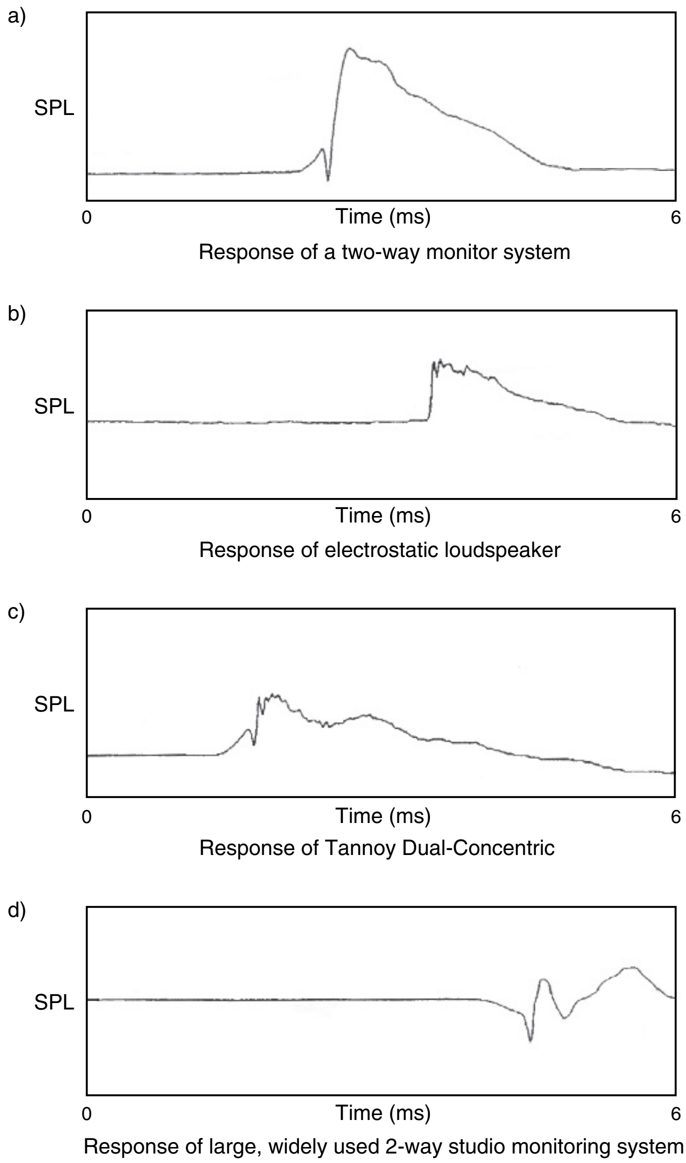


Figure 9.20 Step function responses on short time-scales

domain via an *acoustic source* plot. As the speed of sound at any comfortable listening temperature is around 340 metres per second, time can therefore be converted into equivalent distance. The acoustic source plots shown in Figure 9.21 show the responses of two systems, one a sealed box and the other a reflex enclosure of roughly similar dimensions. It was discussed in Chapter 5 how any filter or resonant system must suffer a 'group delay', where not all the frequencies pass through the system with the same

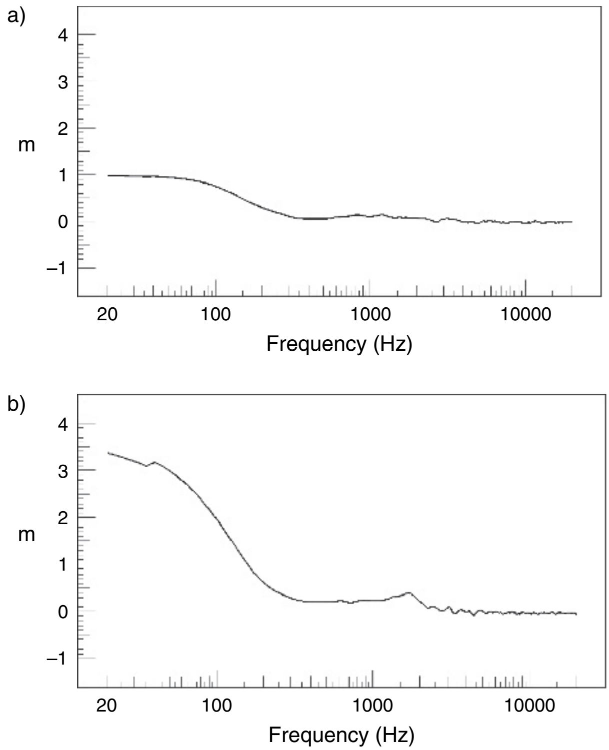


Figure 9.21 Acoustic source plots

Acoustic source plots of the same two loudspeakers whose waterfall plots were shown in Figure 9.8, showing that not only does the reflex enclosure (b) exhibit a longer decay than (a), but also that the low frequencies from loudspeaker (b) effectively emanate from a point over 3 metres behind the physical position of the box. The low frequencies emanate from an apparent point only one metre behind the sealed box (a)

speed. The acoustic source plots show how the different frequency delays give rise to the effect of some frequencies apparently emanating from points somewhere behind the physical location of the loudspeaker cabinets. Some frequencies in a complex sound, after passing through an entire system, actually emanate from the loudspeaker *later* than other frequencies, despite them all having entered the electrical input simultaneously. The low frequencies from the sealed box shown in Figure 9.21(a) can be seen to apparently arrive from a metre behind the face of the cabinet, which corresponds to a delay of about 3 milliseconds, which equates to effectively emanating from about 1 metre behind the force of the loudspeaker. The reflex cabinet shown in Figure 9.21(b) shows a much greater signal delay at low frequencies due to the resonant nature of the box. Here, a 50 Hz signal appears to emanate from a source about 3 metres behind the actual location of the cabinet, which shows the delay in the transient attack with respect to a similar sized sealed box. Obviously, the low frequencies do not really arrive from behind the loudspeaker, but the concept is a useful way of visualising the effect of the group delay on the low frequency components of a transient signal.

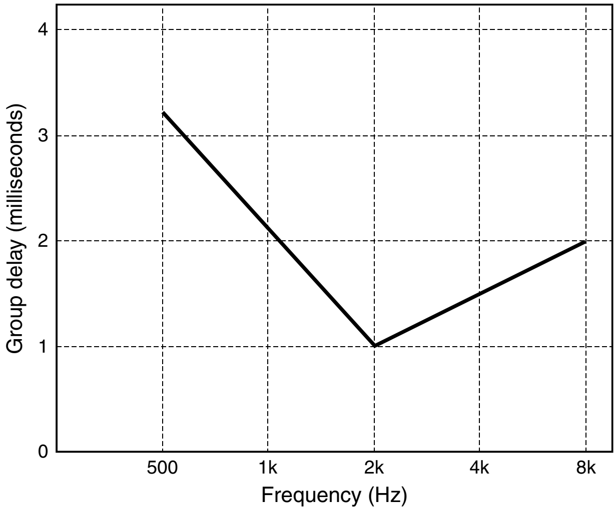


Figure 9.22 The Blauert and Laws criteria for the perception of group delays

Many manufacturers refer to the Blauert and Laws criteria, shown in Figure 9.22. Blauert and Laws determined their results from listening tests, and concluded that any acoustic source plots falling below the line would not be audibly distinguishable in terms of group delay, alone. However, group delays never occur alone, so assumptions about such things should be made with great caution. Effectively the steeper the low frequency roll-off for any given 3 dB down point, the greater will be the group delay and the further behind the physical source will be the apparent source of the low frequency content of a sound.

In some publications the acoustic source has been referred to as the 'acoustic centre', but that term is now generally agreed to refer to the point on a loudspeaker front baffle which most closely corresponds to the measurement axis on which the arrivals from the single or multiple drivers would arrive with the most coherent phase relationship. (See sub-Section 9.2.1.)

9.2.7 Cepstrum analysis

The word cepstrum is an anagram of spectrum. Cepstrum analysis results in plots shown in terms of time against non-dimensional decibels which quantify the gamnitude (an anagram of magnitude). In the world of the cepstrum, phase becomes saphe, high pass filters become long pass lifters, harmonics become rahmonics. Power cepstra were developed in the early 1960s for the enhancement of the detection of echoes from earthquakes in vibrationally noisy environments[9]. The repeated signals become more evident after the inverse Fourier transform of the logarithmic power spectrum, which effectively treats the spectrum as though it were a waveform. Figure 9.23 shows a series of pressure amplitude responses of a loudspeaker with a discrepancy between the on and off-axis responses in the region

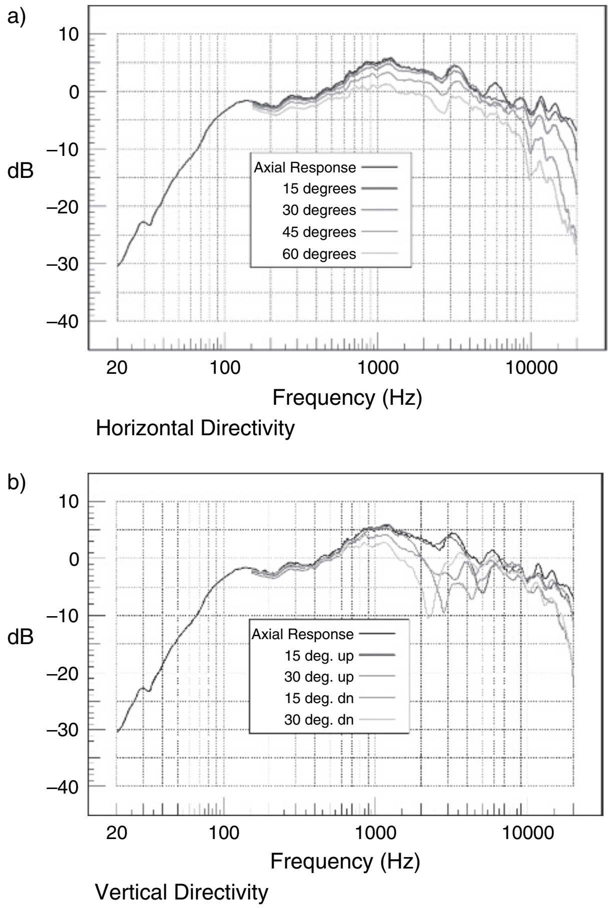


Figure 9.23 On and off-axis pressure responses
Above 4 kHz there can be seen response irregularities in the on-axis response which are not evident in the off-axis responses

around 5-8 kHz, where the on-axis irregularities are not present in the off-axis responses. The cepstrum analysis shown in Figure 9.24 shows a reflexion around 0.4 milliseconds (400 microseconds), which suggests that the problem is one of diffraction from the cabinet edges at a distance of about 7 cm from the centre of the tweeter. One millisecond represents 34 cm at the speed of sound. Four hundred microseconds therefore represents 34×0.4 cm, or 13.6 cm. The half distance, there and back would be $13.6 \div 2$, or 6.8 cm, and the centre of the tweeter was, in fact, about 7 cm from the top and one side of the cabinet.

Cepstrum analysis is not therefore something which directly relates to what we hear (which is not surprising considering its abstract nature) but it can be a powerful tool for diagnosing the sources of complex problems (which is also not surprising considering its original application).[10]

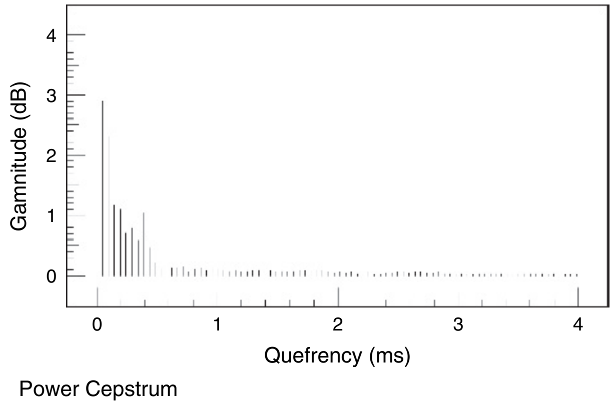


Figure 9.24 The power cepstrum
A strong, single reflexion is evident at about 400 ms, indicating a diffraction problem with the loudspeaker represented in Figure 9.23

9.2.8 Modulation transfer functions

The ubiquitous 'frequency response' plots show the pressure amplitude which a loudspeaker generates at each frequency, or in each defined frequency band, in response to a flat input signal. In many cases, as will be shown in Chapter 11, this simple pressure measurement may fail to reveal many other measured response and sonic differences between loudspeakers. Taken to an extreme, we could scramble the overall phase response by measuring a loudspeaker in a reverberation chamber, and adjust it to give a flat response, but the intelligibility of speech or the resolution of detail in music would be hopelessly lost.

For this reason, in reverberant spaces such as railway stations and airport terminals, a measurement of intelligibility known as a speech transmission index (STI) is often employed which is used to indicate the clarity with which the spoken word would be likely to be heard amongst the background noise and reverberation. Using similar techniques, a system of modulation transfer function (MTF) measurement can be used to indicate the degree of accuracy with which a loudspeaker is reproducing the information content in a musical signal.

Thought of another way, imagine a frequency response like a letter-count in this paragraph. We could individually count all the numbers of the letters a, b, c, d etc, and end up with a table such as a = 28, b = 9, c = 13 etc. If we then shuffled the letters around into one giant anagram, a subsequent letter count would still provide the same result as before; a = 28, b = 9, c = 13 etc, but depending on the degree to which we mixed up the letters, the information content of the paragraph would gradually be lost.

When resonances or group delays within loudspeaker systems or their crossovers and amplifiers smear the time response of a signal, a flat pressure amplitude response may still be achievable, but the onsets of all the components of the music will not arrive in the correct temporal order. They *will* all arrive, but out of sequence due to phase response errors, so an information content loss would be experienced which would equate to shuffling letters around in a paragraph. Fine detail in the musical sounds would be lost, and the jumbled signals would produce other artefacts which were not a part of the original signal. In Chapter 11 is a discussion of the application of this concept to loudspeaker box tunings and port resonances, but here it may be interesting to see how this MTF concept can be applied to room acoustics.[11]

Figure 9.25 shows the comparison of results from a high resolution, full range, flush-mounted monitor system, at a distance of one metre and four metres in the highly damped control room of a music recording studio. The MTF measures the accuracy of response, frequency by frequency, in terms of its fidelity to the input waveform – '1' being perfect and '0' representing no similarity between input and output. It is evident from Figure 9.25 that the control room is not giving rise to any significant loss of information content as the sound waves cross the room, because there is very little difference between plots (a) and (b). [And *no*; despite the oft heard criticisms about absorbent rooms being oppressive, the room is *not* oppressive to be in because there is plenty of reflective surface area sited where the loudspeakers cannot 'see' it, but where it can add adequate life to the speech and movements of people within the room.]

The tests were then repeated using a pair of small loudspeakers in a studio recording/performing room having a relatively neutral acoustic character. Figure 9.26 shows the results, and it can be seen how the MTF drop (information loss) from 1 metre to 4 metres is clearly apparent. Figure 9.27 shows the results of moving the tests into a granite-walled, acoustically live room, using the same loudspeaker and microphone as for Figure 9.26. It is plainly apparent that even at a distance of only one metre, the response has already been significantly degraded with respect to the one metre measurements in the more neutral room. It is also apparent from Figure 9.25 how the full-range, high-resolution, flush-mounted (and expensive) professional monitor system shows more generally detailed information content (a higher MTF at all frequencies) than the small, inexpensive, yet popular 'studio monitor' used for Figures 9.26 and 9.27.

9.2.8.1 Application of room equalisation

A number of companies are now offering monitor systems which purport to deal with the room problems by means of active or adaptive equalisation, to restore a flat frequency response even in relatively uncontrolled rooms.

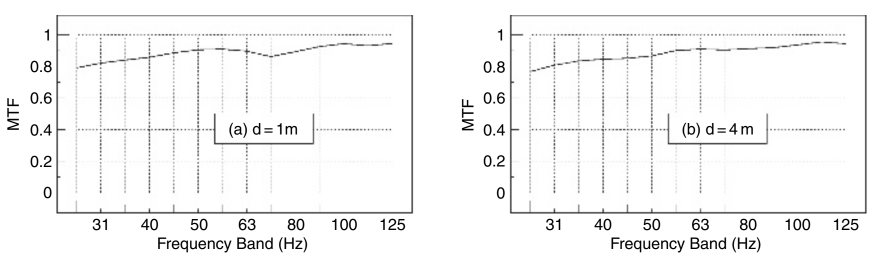


Figure 9.25 Flush-mounted, full-range monitors in a control room

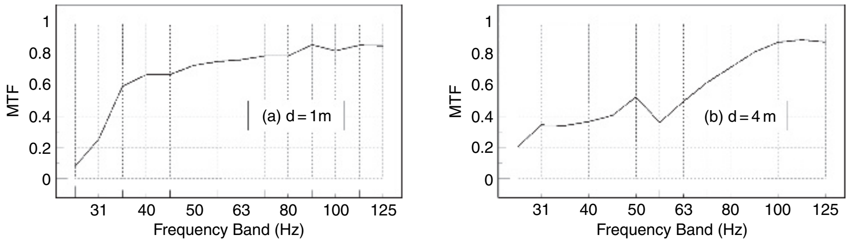


Figure 9.26 Small loudspeakers in a studio performing room

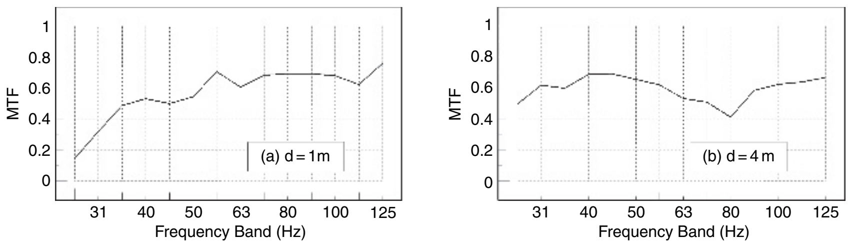


Figure 9.27 Small loudspeakers in a granite-walled, live room

The implication from the publicity often seems to be that room acoustic problems can now be dealt with by signal processing, and also that the highest standards of monitoring clarity can be achieved in less than well-acoustically-designed rooms.

In general, the phase response of a room/loudspeaker system can be separated into minimum-phase (-shift) and excess-phase (-shift) components, as discussed in Section 7.9. The *minimum-phase* components of the response are given rise to by anything which affects the response in a more or less instantaneous way – such as the extra loading on the diaphragm, and the consequent bass boost, when a loudspeaker is placed in a corner. *Excess phase* effects result from time-shifted events, such as group delays in crossover outputs (where the high frequency and low frequency outputs of the filters suffer different signal delays) or reflexions which interfere with a loudspeaker response after returning from a distant surface. In the case of any minimum-phase response modification, the amplitude equalisation will automatically tend to correct the phase errors, and hence the time response (transient response) will also be improved. On the other hand, an excess-phase response will often not have its phase response improved as the amplitude response is flattened, and so its transient response may even be made worse due to time smearing as the amplitude component of the frequency response is flattened.

Figure 9.28 shows the MTFs for the wide-range, high resolution monitor in the highly-damped control room (as shown in Figure 9.25). In this case, its response has been flattened in a computer by the application of a 'perfect' real-time filter, which also employed a 12 dB/octave filter below 20 Hz to prevent wild, out-of-band correction responses. In terms of the MTF, little has changed between Figures 9.25 and 9.28, either at one metre or at four metres. The average MTF has not changed. In the

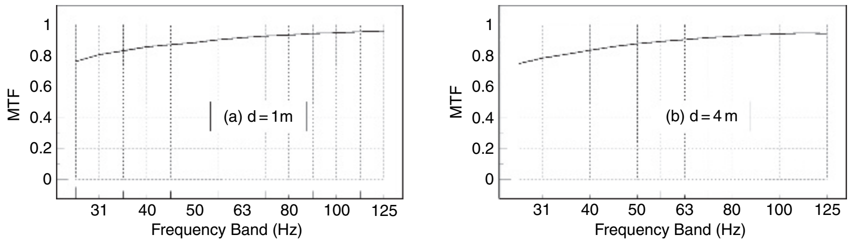


Figure 9.28 Large loudspeakers in a control room – after correction

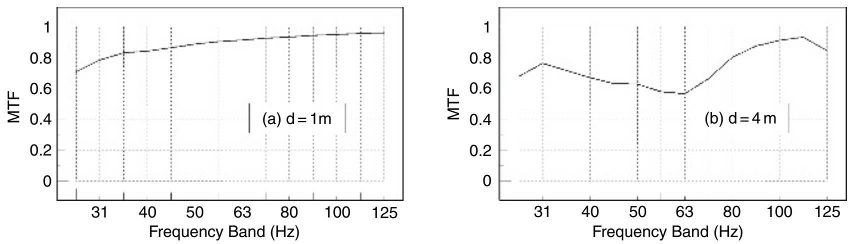


Figure 9.29 Small loudspeakers in a performing room – after correction

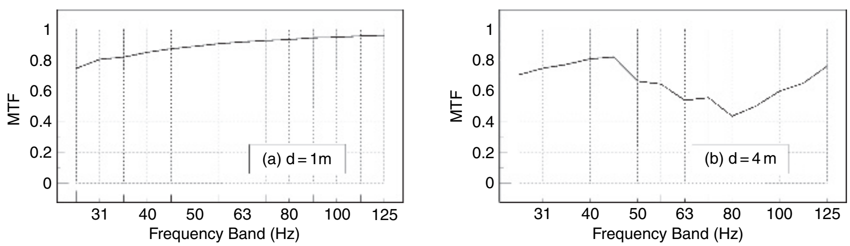


Figure 9.30 Small loudspeakers in a granite room – after correction

case of Figure 9.29, however, which shows the result of 'perfect' real-time equalisation to the smaller loudspeaker in the well-controlled studio room, (as shown previously in Figure 9.26) the MTF response at one metre has been significantly improved by the equalisation, but the response at four metres distance has hardly been improved at all. The results for the same loudspeaker in a stone room, after equalisation to flatness, are shown in Figure 9.30, where it can be seen that the MTF response has also been improved at one metre, but the response at four metres has barely been affected.

These results suggest that the new breed of room-equalised loudspeaker systems can work well at short distances, but that the far-field response in the room may/will not benefit in terms of the resolution of detailed information content, even though the frequency response may appear to be quite flat. In other words, such equalisation may improve the sound for the person close to the loudspeakers, but on the sofa a few metres away the MTF response may remain as bad as ever, or worse! It would appear that only well-designed room acoustics can provide and maintain a large, flat, high-resolution listening area.

All rooms, unless highly absorbent, affect the transmission of information from a loudspeaker to a listener, and even at low frequencies the loss of information content (detail) can be significant. In well-designed listening/control rooms of low decay time (which, once again, need *not* be oppressive to be in if reflective surfaces are strategically placed) the loss of information content is minimal. However, the overall responses in less well treated rooms *can* be improved considerably by modern equalisation processes, but only, it would seem, at relatively short distances from the loudspeakers. Room equalisation does not, in general, significantly reduce the loss of signal information at greater distances.

What the evidence presented here is highlighting is that the flattening of the 'frequency response' is *not* necessarily restoring low level detail and low frequency information accuracy. In fact, as the amplitude part of the frequency response is being flattened, the phase response may be suffering degradation. This may make it easier to achieve a correct musical balance for a mix, but it may *not* do anything to improve the assessment of things such as the fine structural detail or the transparency of the room sounds within the recordings.

Once we get into the lower MTF regions at low frequencies, experience has shown it can become more difficult to balance percussive and more continuous sounds, such as bass guitar to bass drum balances. A good MTF and a fast transient response at low frequencies therefore remain essential features of a good mixing environment.

Clearly, this type of insight into loudspeaker and room responses requires techniques such as MTF analysis in order to be able to confirm and see what ears have been telling people for decades, but which many 'conventional' established forms of loudspeaker measurement systems have signally failed to reveal.

9.2.8.2 A D-to-A dilemma

A further point should also be raised about mid-priced loudspeaker systems which incorporate digital equalisation. The quality of the D to A converters should be considered when thinking about using them. A good quality pair of D to A converters for monitoring a recording made via good quality A to D converters costs around 1000 euros/dollars or more. Clearly, on entire loudspeaker systems costing only 1000 euros/dollars, and having digital inputs, the converters used in the electronics will probably cost nearer to tens of dollars. When auditioning different converters using these loudspeakers, this situation could (and in fact does) lead to conclusions such as: "When we made comparisons, the mid-price A to D converters sounded just as good as the super-expensive ones". Such conclusions could easily be drawn when monitoring via mid-price monitor systems which use digital equalisation systems and low-cost D to A converters in acoustically untreated or inadequately treated rooms.

John Watkinson raised this issue when he suggested that the resolution of a loudspeaker system could be tested by reducing the bit rate of a digital signal until the loss became noticeable[12]. The loudspeakers making audible the smallest bit-rate reductions being the ones with the greatest resolution of fine detail. Although some holes can be picked in this argument, the basic concept does seem to hold water. In practice, the problem which this highlights is that if the limitations of the D to A conversion of the monitor system, or poor MTFs due to bad room acoustics, lead to bad decisions about the choice of A to D converters, the deficiency will be forever locked into the recording. Conversely, excellent A to D conversion, even if not revealing itself on all reproduction systems, will be fully enjoyable by those who do listen to the recordings via high quality reproduction chains. However, measurement systems for accurately predicting subtle sonic differences in A-to-D and D-to-A conversion are still not well-defined, and no simple, powerful tool is yet readily available. Nevertheless, it should be rather self-evident that if any D to A converters in the monitor system are not of the highest quality, then they will limit the ability to monitor the quality of any other converters in the recording chain. Cheap, quality-control monitors with digital inputs are therefore, effectively, a contradiction in terms.

9.3 Sound fields and human perception

Once all the objective testing has been completed, the ultimate assessment of the quality of a loudspeaker intended for musical use must be made by the ear. Unfortunately there are some aspects of perceived loudspeaker quality which do not easily lend themselves to objective measurement, yet which are important aspects of perceived quality. No matter how good a recording may be, and no matter how good its reproduction in terms of spacial imaging and definition, one incontrovertible fact is that its reproduced sound-field will bear little resemblance to the sound-field of the original instrument. Of course, if the music is an electronically based creation, where no real instruments ever existed other than the electronic system itself, then the loudspeaker playback on the original monitor system on which it was mixed *is* the definitive sound field. However, a loudspeaker reproduction of a cello recording will inevitably give rise to a huge spacial distortion in terms of the sound field. A bowed cello radiates sound from a large area, and in a very complex manner. Walking round a cello whilst it is being played, a listener may experience a small reduction in the high frequencies when passing immediately behind the cellist, whose body will tend to cast a high frequency shadow, but otherwise the tone would be perceived to be relatively independent of position. The source is very distributed, with many parts of the instrument radiating a wide range of frequencies at the same time. The sound field pattern from the instruments contrasts sharply with that from any loudspeakers radiating a reproduction of a recording of the same instrument.

The recording microphone is a pressure transducer, or a pressure gradient transducer if of figure-of-eight pattern, which responds to the radiated sound received at the small place that the diaphragm occupies, and converts the sound pressure changes into an electrical signal. In stereo, with two microphones, a two-channel signal can be recorded which can convey to the ear of a centrally placed listener, via loudspeakers or headphones, enough information to give a sensation of the source positions in terms of left and right, but not in terms of height. Once these signals are reproduced via multi-driver loudspeakers, the 'no-height' information is also separated into frequency bands. All the high frequencies above a certain crossover point will come from two points (the tweeters), perhaps no more than 3 cm in diameter, spaced at the extreme left and right of the sound-field. The composite high frequency sound pressure is thus beamed at the two ears in two separate rays, from two points in space, with no height information.

Compare this to listening to the instruments of a string quartet or a rock group. The sound sources are multiple and are distributed in width and height. High frequencies arrive from the entire drum kit, with the cymbals above ear height and the snare drum below, or with the violin above and the cello below. If the recordings were made in an anechoic chamber, then played back in a reflective room, comparison to the live performance at the same place in the same room would show many differences. The real instruments would each be located at different places with regard to the nodes and anti-nodes of the room modes. Conversely, the phantom images of the recorded instruments would all originate only from the two loudspeaker positions, and only the nodes and anti-nodes relevant to the two places occupied by the two loudspeakers could affect the overall response. A phantom central image of a guitar would not couple to the room at its phantom position, but at each loudspeaker position. Therefore a pair of loudspeakers would send a signal to the ears from two places, whereas a real group of musicians would send their sound to the ears from many individual positions. In reflective/reverberant conditions, a pair of loudspeakers would couple to the room acoustics at only two places, whereas a group of musicians would each couple to the room in a different way, and even the individual drums in a kit would couple differently to the room modes, and each produce their own unique reflexion patterns and timings. Given the extreme sophistication of the human hearing system it would be stretching the imagination to even hope that such differences between live music and music reproduced via loudspeakers could go unnoticed.

[The pinnae of the ears (the ear flaps) are highly refined devices which collect sound in different ways from different horizontal and vertical directions. The sound pressure differences at the entrances to the ear canals are therefore not merely the differences that would occur if microphones were placed in the same location in the absence of a head, torso and pinnae. A pair of ear canals receives cues about the directions of sounds in the horizontal and vertical planes which are not available to a pair of microphones, and as such the composite responses are not the same. By the use of dummy heads with ear flaps, binaural recordings of great spacial sensitivity can be made for playback over close-fitting earphones, but such recordings will not work via loudspeakers because the head, torso and pinnae of the listener would introduce a second set of processing which would confuse the delicate information in the binaural recording.]

The ears therefore receive a different set of cues from the distributed set of sources than from a phantom sound stage created by two sources, and the resulting sound-field distortion can lead to great perceptual differences. As loudspeakers cannot three dimensionally reproduce the reality of a live performance, we must therefore look at loudspeaker reproduction as a performance in its own right. The question then to be asked is how well the loudspeakers can transmit the emotions which were generated by the original performance. Given that the musicians were likely to be using the tones of their instruments to manipulate the sensations which they were intending to convey, it would seem obvious that their timbral subtleties should be preserved as accurately as possible. However, exactly how that timbral fidelity can be achieved may depend on certain compromise decisions, but the optimum compromise for one set of instruments may not be optimum for a different set of instruments. The compromises for large loudspeakers with great source areas may be different from the compromises which seem most apt for smaller, more compact sources, but even the differences between those sources may be dependent on the musical genre. Furthermore, to create an appropriate rendition of any given piece of music in different room acoustics may also favour one set of loudspeaker design compromises over another.

When music is created by electronic or electric sources, it is itself created *on* loudspeakers for reproduction *by* loudspeakers. One presumes that the loudspeakers on which the music is finally mixed are the reference for the timbre and balance of the instruments, but the question arises as to how to choose those loudspeakers. Are they to be chosen to give the widest range of options for the recording personnel, or are they to be chosen to be the most appropriate to enable the widest range of likely domestic loudspeakers to reproduce the intentions of the producers in the majority of circumstances? This dilemma often leads to the use of different recording and mixing loudspeakers, as was discussed in Chapter 8.

There are in fact some market tendencies which do exist. Leaving aside the audiophiles for now, and also leaving aside the people for whom music is just something to fill the empty air, there is a tendency, for people who choose their loudspeakers by careful listening before buying, to choose different loudspeakers depending upon the type of music that they mostly listen to. People who like orchestral music tend to buy different loudspeakers to those who like rock music. In fact, there is also a tendency for recording engineers who work principally with classical music to use different monitor loudspeakers to those who record rock music. What is more, there are traceable similarities between the recording/mixing and domestic listening loudspeakers used by each group.

The orchestral/acoustic music listeners tend to value loudspeakers with low non-linear distortion, low colouration and a smoothly rolling-off low frequency response, coupled with wide and smooth directivity to give rise to plenty of the lateral reflexions which are necessary to produce a sense of spaciousness. Pinpoint stereo positioning is often low on their agenda, because in the reflective and reverberant acoustics of a concert hall, no precise positional localisation is possible at a live concert; and neither is it usually considered to be desirable, because it would imply an acoustic that could not support the spaciousness – the two things are generally mutually exclusive. On the rock music side, flat low frequency responses are often valued, even if they cut-off quite abruptly. Colouration is to some degree acceptable because instruments are often so heavily equalised in the recording and mixing processes that no real reference exists. Heavy percussion and transient signals are commonplace, so high sound pressure level capabilities are often required, and as the audibility of small amounts, or even not so small amounts, of non-linear distortion is doubtful on such high impact recordings, non-linear distortion levels may be tolerable which would be too high for classical music enthusiasts. Relatively narrow directivity may also be deemed to be desirable for rock music enthusiasts because too many room reflexions can detract from the transient impact of the fast changing music. However, all of the above-above-mentioned items are tendencies, only, and will not apply in all circumstances.

Of course for a price, if money is no great object, very many of the most desirable properties could be reasonably incorporated into one design, but it would not be cheap and it would not be small. At 2006 prices, if one were prepared to pay 2000 euros per pair of 50 litre boxes, one could begin to approach a compatible design if the room acoustics could also be reasonably contoured to requirements. Unfortunately though, it is sad to say that even people in some supposedly professional parts of the music industry consider such costs and sizes to be beyond their circumstances. In such cases, it is really important to understand that when compromises are made when using cheaper and smaller loudspeakers, they may not be as suited to some music or rooms as they are to others, so any reference which they provide may not be as robust or broad-based. Further implications of mix compatibility will be discussed in Chapter 10, but in general, within reason, as loudspeaker costs and sizes increase, and rooms become better controlled, the easier it is to achieve a more universal set of monitors. Small, cheap loudspeakers, if they are good at all, tend only to be good for a limited range of circumstances and uses.

It is therefore difficult to be too rigid in trying to determine threshold levels for 'good' performance when considering the implications of the various characteristics discussed in Section 9.2 because what is optimal for any given size of loudspeaker will depend upon its use. Ultimately, the goal of listening to music is to enjoy it, so what gives pleasure has value, even if it can be technically argued against. However, in general, the closer that one can get to technical excellence, the overall sonic performance of a loudspeaker usually improves, and it is surely incumbent on a professional recording industry to be fully aware of what is on a recording, even if 99.9% of the purchasers of the end products are not going to hear the subtleties. The music buyers can invest in their domestic entertainment systems according to the degree that they are important in their lives, but professional attitudes are more demanding. If nothing else, professional pride requires that the end users should not become aware of recording errors that passed unnoticed through the recording and mixing process.

9.3.1 Further perceptual considerations

The fact that our pinnae, middle ears, inner ears and brains are unique to each of us introduces aspects of physiology and taste into the questions of loudspeaker parameter optimisation. Culture and ethnicity also have a bearing on the subject. When concentrating on listening to music, all human beings tend to react with the side of the brain which relates to them being right or left handed. When listening in a relaxed way, the tendency is for the activity to switch to the opposite hemisphere. Mongoloid races, such as Japanese, tend to process western music with one half of the brain and oriental-style music with the other half of the brain[13].

Dr Diana Deutsch, at the University of San Diego, California, published finding showing that the *place* of our birth, irrespective of being from local descendants or not, can affect our musical perception. The median pitch of the language and accent with which we first learn to speak can fix certain aspects of our musical perception for the rest of our lives, and may even affect the perception of complex pitch sequences in terms of whether we hear them to be rising or falling[14]. Southern English and Californian populations were shown to perceive a tri-tone pitch sequence in opposite ways. During listening tests at the Institute of Sound and Vibration Research, a band-limited, anechoic recording of an acoustic guitar chord was perceived by some listeners to change its notational inversion when played through different loudspeakers, whilst other listeners heard the same chord *notation* but a change in the timbre[15].

During the installation of a monitor system in London in the late 1980s, there arose a situation with two well respected recording engineers who could not agree on the 'correct' amount of high frequencies from a monitor loudspeaker system which gave the most accurate reproduction when compared to a live cello. They disagreed by a full 3 dB at 6 kHz, but this disagreement was clearly not related to their own absolute high frequency sensitivities because they were comparing the sound of the monitors to a live source. The only apparent explanation for this is that because the live instrument and the loudspeakers produced different sound fields, the perception of the *sound-field* was different for each listener. Clearly, all the high frequencies from the loudspeaker came from one very small source, the tweeter, whilst the high frequency distribution from the instrument was from many points on the strings and various parts of the body. The 'highs' from the cello therefore emanated from a distributed source having a much greater area than the tweeter. Of course, the microphone could add its own frequency tailoring and one-dimensionality, but there would seem to be no reason why this should differ in perception from one listener to another.

During research in the late 1970s, Belendiuk and Bulter[16] concluded from their experiments with 45 subjects that "there exists a pattern of spectral cues for median sagittal plane positioned sounds common to all listeners". In order to prove this hypothesis, they conducted an experiment in which sounds were emitted from different, numbered, loudspeakers, and the listeners were asked to say from which loudspeaker the sound was emanating. They then made binaural recordings via moulds of the actual outer ears of four of the listeners, and asked them to repeat the test, via headphones, of the recordings made using their own pinnae. The headphone results were very similar to the direct results, suggesting that the recordings were representative of 'live' listening. Not all the subjects were equally accurate in their correct choices, though, with some, in both their live *and* recorded tests, scoring better than others in terms of identifying the correct source position. Very interestingly, when the tests were repeated with each subject listening via the pinnae recordings of the three other subjects in turn, the experimenters noted, "that some pinnae, in their role of transforming the spectra of the sound-field, provided more adequate (positional) cues . . . than do others”. Some people who scored low in both the live and recorded tests, using their own pinnae, could locate more accurately via other peoples’ pinnae. Conversely, via some pinnae, none of the subjects could locate very accurately. However, for all the subjects, listening through their own pinnae sounded most natural to each of them.

Interestingly, some people do claim to hear height information in two-channel stereo recordings which can carry no such information, but this is an effect of their own pinnae being stimulated in different ways by different loudspeakers and mounting conditions. Such people really do hear height in the stereo images, but *only* as an artefact of their own pinnae, and not of the recordings or the loudspeakers.

Given these differences, and all of the aspects of frequency discrimination, distortion sensitivity, spectral response differences, directional response differences, psychological differences, environmental differences, cultural differences, and so forth, it would be almost absurd to expect that we all perceive the same balance of characteristics from any given sound. True, whatever we each individually hear is *natural* to each one of us, but when any reproduction system creates any imbalance in any of its characteristics, as compared to a natural event, the aforementioned human variables will inevitably dictate that any shortcomings in the reproduction system may elicit different opinions, vis-à-vis the accuracy of reproduction, from different people. So, to the question of whether it is more important to reduce the harmonic distortion in a system by 0.2%, or the phase accuracy by 15 degrees at 15 kHz could well be an entirely personal matter, which could be task related, music related or room related, or even related to a person’s psycho-physical make-up.

In fact, it could also be experience related. Human beings have a great tendency to acclimatise themselves to familiar sounds. If we go into a strange house, we often note the change in the sound of our voices, yet in our own homes we rarely notice a change from going from outside to inside, because we are so used to it. People who have been working for years on a certain genre of loudspeaker also have a tendency to hear music in relation to the way that they are accustomed to hearing it. It is hard to say whether people make some choices about loudspeakers because they suit their concept of what sounds right to them, or because they have a recognisably familiar sound that they feel comfortable with. In the realms of acoustic music, if a listener is also a regular concert-goer, the reference of the general sound of live instruments is always in the back of their mind as a reference, but with heavily processed music, only the people in the control room at the time of the mixing *really* know what the sound should be like. People can and do get accustomed to how a wide range of non-acoustic music sounds on their own loudspeakers, then presume that that is how everything should sound. [Conversely, I had a recording which I liked very much and had played it on many systems, the better ones of which yielded a remarkably rich yet obviously natural drum sound. By chance, several years later, and on a different continent, I met the mixing engineer, and told him what I thought of the recording. “Natural?”, “I had to e.q. the hell out of it to get that sound!" he exclaimed. So we have to be very careful about our supposed references. P.N.]

Despite all of the complications of human perception, objective analysis of loudspeakers is fundamental for keeping the progress on the rails because subjectivism has far too many variables. Toole speaks of a circle of confusion[17], where recordings considered to be 'good' are used to assess loudspeakers, which are used to make recordings, which are made to assess loudspeakers, which are used to make recordings . . . ad infinitum.

If this is done in conjunction with continued reference to objective loudspeaker measurements, it is perhaps one of the only means that we have to refine our assessments, but Toole's implications were that this has often been a 'circle of testing' which has occurred too often *without* the strict objective constraints. Where, for example, domestic loudspeakers of average quality have been the de facto reference, and monitoring loudspeakers have been chosen which give the best compatibility with this arbitrary reference. Then the new sounds recorded via this system become the reference recordings, and so forth. Some marketing people may actually applaud this, but professionally it is a loose practise.

Moulton[1] discusses listening in the living room, the bedroom, the car, on headphones, on the 'ghetto blaster' in the street, then concludes that in order to sound compatible on all these irreconcilably different playback systems, a recording must be:

a) comparatively simple spacially and texturally
b) limited in note-to-note dynamic range
c) exaggerated in spacial characteristics that do not become significantly degraded timbrally when summed into mono and,
d) must not depend on their musical effect for any frequencies outside of a range from 80 Hz to 10 kHz

The latter requirement dispenses with three whole octaves of the audible range – octaves where the 'power' and 'air' are to be experienced. Even people who cannot hear tones above 10 kHz can still usually hear when the 10 to 20 kHz band is cut, so if we were to monitor only to this lowest common denominator, we would be robbed of much of the musical experience when listening on audiophile systems. Working like this does no justice to either the art, the technology or the science. The discerning listeners deserve more.

Whilst the 80 Hz to 10 kHz brigade are unquestionably the bread and butter of the industry, the artistes, the dedicated professionals and the audiophiles are the great driving force that keeps it fresh and interesting. Said Roederer, "Music is not a waveform, but a psychological and spiritual construct within the mind. The waveform and its physical dimensions are simply carriers of musical information"[18]. Yet, that waveform is best defended from corrupting influences if the objective performance of a loudspeaker is designed to best convey it. Despite the subjective assessment being the ultimate arbiter of any loudspeaker's performance, it is only by objective means that we can keep focussed on the engineering targets, without which, and by subjective means only, we would surely become lost.

It may well be that the dynamic ebb and flow of the music, the subtle timing differences and other characteristics are better exhibited by some loudspeakers than by others, but these are things for which we have no reliable, measurable descriptors. However, the overwhelming tendency is for a loudspeaker which scores highly in all the objective measurement regimes detailed in Section 9.2 to show the musical characteristics in a more artistic and exciting light. Nevertheless, it must always be the case that until we have perfection, we will always be faced with different balances of compromises serving some situations better than others.

References

1 Moulton, D., 'The Creation of Musical Sounds for Playback Through Loudspeakers', AES 8th International Conference, The Sound of Audio, Washington DC, USA (1990)
2 Watkinson, John; 'The Jitter Bug', Resolution, Vol 1, No 4, UK (October 2002)
3 Briggs, G., 'Sound Reproduction', Third Edition, p 174, Wharfedale Wireless Works, Bradford, UK (1953)
4 Newell, P., Holland, K. R., 'Do All Mid-Range Horn Loudspeakers have a Recognisable Characteristic Sound?' Proceedings of the Institute of Acoustics, Vol 12, Part 8, pp 249–258, Reproduced Sound 6 conference, Windermere, UK (1990)
Also in: Newell, P., 'Studio Monitoring Design' Chapter 12, p 200, Focal Press, Oxford UK (1995)
5 Czerwinski, E, Voishvillo, A., Alexandrov, S., Terekhov, A., 'Multitone Testing of Sound System Components – Some Results and Conclusions, Part 1: History and Theory', Journal of the Audio Engineering Society, Vol 49, No 11, pp 1011–1048 (November 2001)
6 Czerwinski, E., Voishvillo, A., Alexandrov, S., Terekhov, A., 'Multitone Testing of Sound System Components – Some Results and Conclusions, Part 2: Modeling and Application', Journal of the Audio Engineering Society, Vol 49, No 12, pp 1181–1192 (December 2001)
7 Voishvillo, A., 'Assessment of Loudspeaker Large Signal Performance – Comparison of Different Testing Methods and Signals', Presented at the 111th AES Convention, New York, USA (November/December 2001)
8 Janovsky, W., 'The Audibility of Distortion' (in German), Elek, Nochr.-Tech, Vol 6, pp 421–430 (Nov 1929)
9 Bogert, B. P., Healy, M. J. R, Tukey, J. W., 'The Quefrency Analysis of Time Series for Echos: Cepstrum, Pseudo-Autocovariance, Cross-Cepsrum and Saphe Cracking', in Rosenblatt, M., (editor), Proceedings of the Symposium on Time Series Analysis, pp 209–243, Wiley, New York, USA (1963)
10 Holland, K. R., 'Use of Cepstral Analysis in the Interpretation of Loudspeaker Frequency Response Measurements' Proceedings of the Institute of Acoustics, Vol 15, Part 7, pp 65–72 (1993)
11 Holland, K., Newell, P., Castro, S. and Fazenda, B., "Excess Phase Effects and Modulation Transfer Function Degradation in Relation to Loudspeakers and Rooms Intended for the Quality Control Monitoring of Music". Proceedings of the Institute of Acoustics, Vol 27, Part 8, Reproduced Sound 21, (2005)
12 Watkinson, J and Salter, R., "Modelling and Measuring the Loudspeaker as an Information Channel", presented to the 'Reproduced Sound 15' conference of the Institute of Acoustics, Stratford-on-Avon, UK (November 1999)
13 Davis, D., Davis, C., 'Sound System Engineering', Second Edition p 9, Focal Press, Oxford, UK (1997)

14 Deutsch, D., 'Paradoxes of Musical Pitch' in Scientific American, Vol 267, No 2, pp 70–75 (August 1992)
15 Newell, P., 'Studio Monitoring Design' Chapter 5, Section 5.6, Focal Press, Oxford, UK (1995)
16 Belendiuk, K., Butler, R. A., 'Directional Hearing under Progressive Impoverishment of Binaural Cues' Sensory Processes, 2, pp 58–70 (1978)
17 Toole, F., 'Art and Science in the Control Room' Proceedings of the Institute of Acoustics, Vol 25, Part 8, Reproduced Sound 19 conference, Oxford, UK (2003)
18 Roederer, J. G., 'Introduction to the Physics and Psychophysics of Music' 2nd Edition, pp 11–12, Springer Verlag, New York, USA (1979)

Bibliography

1 Voishvillo, A., 'Assessment of Nonlinearity in Transducers and Sound Systems —From THD to Perceptual Models', presented at the 121st convention of the Audio Engineering Society, San Francisco, USA (October 2006)

Chapter 10

The mix, the music and the monitors

10.1 Physics or psychology?

The recording personnel in many small studios seem endlessly to be searching for that magic pair of loudspeakers that they can trust for whatever type of music they mix on them. They swear by one pair of loudspeakers as being their ultimate reference, then after the first mix that fails to sound as good on another system – say at home, or in the A&R department office – they lose all faith in them and seek a new 'reference'. Until they find that new reference, they may live in a state of unease, and even panic, as their perceived anchor to 'reality' loses its grip. Why such well-trusted loudspeakers could suddenly lose their authority has long puzzled many people, but the fact is that the loudspeakers, themselves, rarely *are* the sources of the problems. They usually have not changed at all between the days when they were used to make the 'magic' mixes and the days when the mixes were not perceived to travel so well. As often as not, the problem lies not in the loudspeakers, but in the music, as we shall see later. However, many other things can also change along the way, such as the perception of the recording engineers who are using them. They often seem to pass through the following phases during the period of use of any, one type of loudspeaker system:

1) First audition – forming an initial opinion.
2) Getting to know them – evaluating and refining the opinion.
3) Increase in worries about minor aspects of the sound.
4) Boredom with the sound and doubts about its 'accuracy'.

In so many instances, after finding a new, favourite loudspeaker, people say that they have finally found a monitor to which they can accurately refer their recordings and mixes, only to find that six months later they again declare them to be 'wrong' and unusable. If in reality nothing in the control room had changed in six months, and it is very doubtful that a pair of good quality loudspeakers would have changed their characteristics in such a short time of use, then the only thing which could have changed is the way in which the users were *perceiving* the sound.

Sounds exist only in peoples' brains. We all hear differently – we all perceive the sounds differently – so it is difficult to *compare* musical perceptions, and no doubt we all have our individual hierarchies of priorities in terms of what is important in any given musical rendition. Furthermore, almost any experienced recording engineer or producer would freely admit that on any given day, the 'accuracy' of the monitor system or the 'rightness' of a mix can seem to be dependent upon things such as the general mood the day before, or how well they have slept. Less experienced people, however, often believe that *their* hearing system is always accurate and 'highly tuned'. If, one day, a monitor system does not sound 'right' to them, or even when somebody else tells them that they do not sound right, they call people in to carry out a thorough test of the system, and it is often some absurd response change, carried out by the person or persons called in to make the adjustments, which ultimately placates them. But perhaps something is heard with the new settings that were not noticed the day before, so in the minds of the recordists, this signifies that more detail is now being perceived. The confirmation by measurement that all was previously well with the loudspeakers frequently does nothing to deter the requests for something to be changed.

In fact, quite often, rather than the measurements confirming that all was well, they can induce a sense of even greater insecurity, because some recording staff may feel that their comments were not being listened to, and that their integrity and reputation were being undermined. At this point something *must* be found so that honour and credibility can be restored. Therefore, under growing pressure to come up with something, and whilst being influenced by what are often misleading descriptions of the perceived problems, perhaps some illogical change will be made to the monitoring response which nonetheless seems to provide a credible 're-alignment' of the loudspeaker. The change has a novelty value which brings a new 'reality' with it. Honour has also been satisfied, because a problem was found and rectified, yet the probability is that in almost all cases, no physical problem existed and no aspect of the loudspeaker response had actually changed *until* the adjustment was made. The only problem that the adjustment really fixed was the one in the mind of the person reporting the problem, and it possibly, in reality, took the loudspeaker response in a direction *away* from absolute accuracy.

So what underlies this sort of insecurity and variability of perception? There is little doubt that mood can affect perception. The stress resulting from a rejection of a mix by a recording company can certainly deflate the confidence, especially of less experienced sound recordists, and once doubt and insecurity sets in it develops a momentum of its own. Nevertheless, the problem under discussion here is too frequent and widespread to be simply a case of a confidence crisis, because even when good moods are restored for all parties, the problem of the mix being incompatible between a number of loudspeaker systems may still remain, even though other mixes done on the same monitor system may not have suffered from the same problem.

10.2 The musical dependence of compatibility

The symptoms of a very common problem tend to be that a mix which was done on a certain monitor system, in which the mixing personnel used to have great confidence, and which sounded great in the control room, did not sound good on the radio, or in some other important reference place. They then compare this situation to that of a mix which had been done some months earlier, (of a different piece of music – in almost all cases) on the selfsame monitor system, which sounded great in *all* the places in which it had been checked.

The perceived implication is normally that if the earlier mix sounded good in the control room, and then proceeded to sound good in several other 'trusted' locations, then it must be the case that the *later* mix, which did *not* sound good in the other locations, must have been incorrectly mixed. This is usually thought to be due to a changed control room monitor system performance misleading the mixing personnel into making erroneous judgements. However, if a new monitor system is used, and a more compatible mix results, it almost never seems to be the case that anybody then bothers to remix the *earlier* music, which showed good general compatibility, on the new reference system, and then play it back in the other reference places to see if it *still* sounded good on all the systems. This is a crucial test, because the musical source material can be very influential in terms of the compatibility between loudspeaker systems. It could be the case that the musical arrangement was the source of the compatibility problems. This fact can easily be demonstrated. Two monitor systems in a control room can often be adjusted to sound almost identical on one voice or instrument, only to then sound very different from each other on a different musical signal.

It is also common knowledge that some music, mixed by certain people in certain famous studios, seems to sound good wherever it is played, yet many mixes, done by lesser mortals, only sound good on a few types of loudspeaker. Many people seem to think that some special skills or equipment are being used in the recording, mixing or mastering processes by the people producing the 'sound good anywhere' mixes. They believe that they must have some special, 'truthful', monitor systems, and wonder what they can be. But what people often fail to remember is that the top people in the top studios are also often working with very highly skilled and experienced musicians, who also have a lot of experience in how to achieve well-arranged, *musically* balanced recordings, which show a remarkable degree of tolerance to minor level or equalisation changes.

When working in recording studios, up until the early 1970s, an almost universally present member of the recording team was the musical arranger. Things were worked out and rehearsed in advance, and each instrument had its own place in both the time and the frequency domains. It was very much a part of the job of the musical arranger to make sure that instruments did not clash with one another either in terms of time, pitch, or timbre. However, as things have progressed, the recording process had gradually become somewhat more anarchic, and this had led to mixes in which many instruments may be fighting for the same time and frequency spaces. They need to be so delicately balanced in order to sound 'right' that even minor differences in any aspect of the response of another loudspeaker system can be sufficient to render the balance unsatisfactory. Quite arbitrarily, the loudspeakers on which music is mixed can often be declared to be inaccurate, but from the plots of Figure 11.1, we can see quite clearly how *no* loudspeakers are technically accurate.

10.2.1 Sine waves and pink noise

Let us think of the compatibility problem this way. If we put a mid frequency sine wave into five good quality loudspeakers in turn, then adjust all the levels to give the same sound pressure, it is probable, unless serious distortion is present in any of the systems, that the sine wave will sound the same in each case. The sine wave represents just about the simplest signal source that could be used. Now let us go to the other extreme, and use a signal that contains *all* frequencies, such a pink noise. Almost certainly, even after the most careful adjustments, the pink noise would sound noticeably different when reproduced by each of the five different loudspeakers – its tone colour would certainly change – and it may even sound different if played alternately through the two loudspeakers of a matched stereo pair.

The significance of this, in musical terms, is that the more that a musical mix tends towards a noise signal (in many cases this means the more complicated that it becomes) the more it will tend to sound different when switched from one loudspeaker to another. A lightly blown flute will therefore tend to sound more similar when reproduced by a wide range of loudspeakers than would be the case for a fuzz guitar. The latter could noticeably change in timbre, or even in pitch, even when played on relatively similar loudspeakers.

So, the more information that exists in a musical mix, the more of it there is to get upset by minor response differences. Clean and simple mixes, or mixes of music which is well arranged in terms of the temporal and frequency distribution of the different musical parts, will tend to be more robust than mixes which are highly congested and/or heavily processed. Many of the recordings of Dire Straits always seem to sound musically balanced no matter on whatever systems they are played, from an audiophile hi-fi system to a hotel radio alarm clock. Further consideration of their excellent musical arrangements surely goes some way to explaining why this is the case. This situation can puzzle many people who have gone to great lengths to ensure that their monitors are well set up, and they are very respectably flat when analysed with pink noise. However, as discussed in Chapter 9, there is far more to a loudspeaker response than its steady-state pressure amplitude response.

10.3 Real responses vs. preconceived ideas

There is a tendency in much of the music recording world to believe that the responses of all 'good' loudspeakers are more or less the same, and that it is merely the precise application of their proprietary construction techniques which makes the minor differences between them, but the reality is somewhat brutally different, as Figures 10.1 to 10.5 show. The plots show five aspects of the responses of a Yamaha NS10M and two of its 'competitors' in the field of close-range music monitoring. These measurements were all made in one, fixed position in the same, large, anechoic chamber. Even before any room response anomalies could be considered, such as those given rise to by directivity differences, the simple, on-axis, anechoic

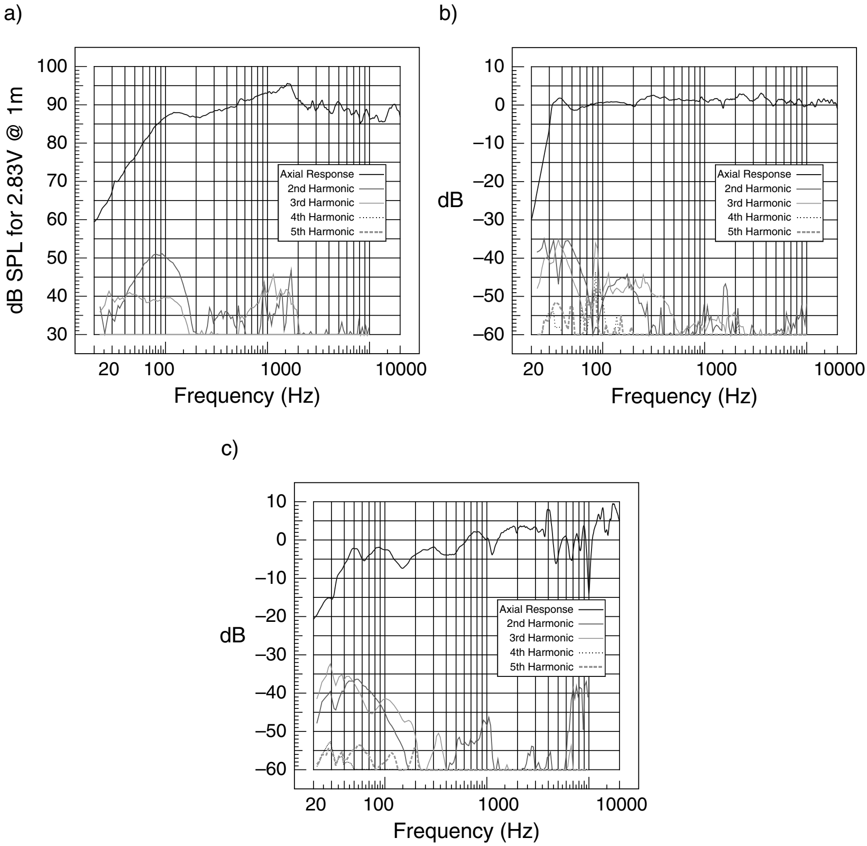


Figure 10.1 Pressure amplitude and harmonic distortion responses of three loudspeakers. a) Yamaha NS10M. b) Genelec S30D. c) SAE TM160A

responses of the three loudspeakers, as shown in Figure 10.1, are different to the point of absurdity if they are all ostensibly for the same sort of professional use. The levels of harmonic distortion are also quite different. Given their disparate responses, there is simply no way that a wide range of musical mixes *could* sound similar when alternately played on each of the three loudspeakers, but the degree to which they do sound different may well be largely a question of instrumentation and arrangement, rather than simply a matter of the balance of the mix.

Figure 10.2 shows the acoustic source plots of the same three loudspeakers referred to in the previous paragraph. As each metre on the vertical scale represents about 3 milliseconds, it can be seen that from the Yamaha the 50 Hz component of the signal is delayed by about 4 ms (about 1.3 metres).From the SAE the delay is around 6ms, and for the Genelec, the delay is in the order of *ten* milliseconds. The plots of Figure 10.2 show the delay in the *attack* of the low frequencies, but by contrast, Figure 10.3 shows the delay in the *decays.* The waterfall plots show how the different frequencies decay in response to any signal which excites

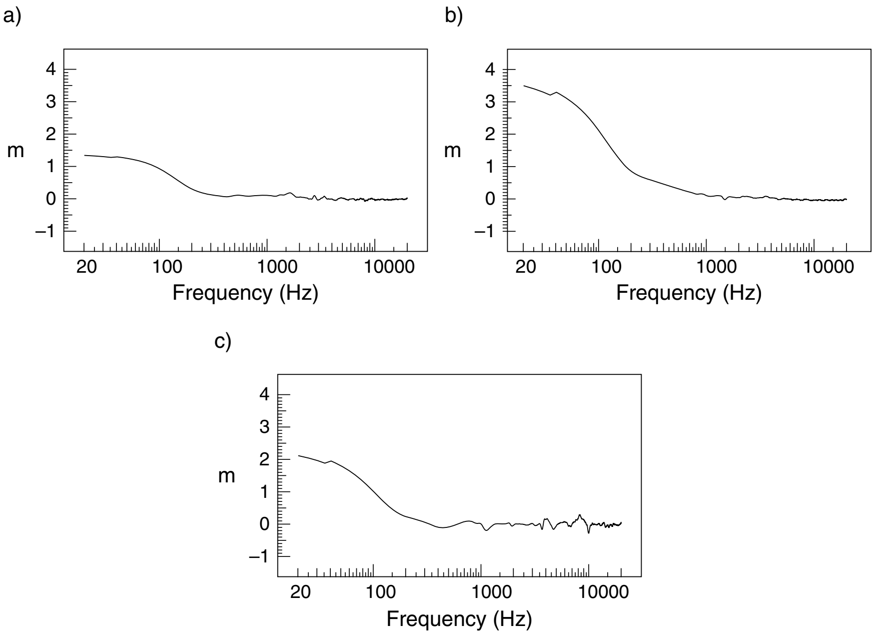


Figure 10.2 Acoustic source plots. a) Yamaha NS10M. b) Genelec S30D. c) SAE TM160A

the loudspeakers. Again, the temporal response variations are enormous. On these plots, the 60 Hz component from the Yamaha is 30 dB down by 20 ms; the Genelec is barely down by 30 dB even after *100 ms*, and the SAE response lies somewhere in-between. How *can* a bass drum sound the same on each loudspeaker? These longer responses can play havoc with the balance between bass drums and bass guitars.

Figure 10.4 shows an electrical step function. An easy way of generating such a function would be to connect a 1½ volt battery to the input terminals of a loudspeaker – active *or* passive. Figure 10.5 shows the time response (which is in fact the waveform response) of the same three loudspeakers as shown in Figures 10.1 to 10.3 when subjected to a step-function stimulus. Once again, the resulting waveforms are very different. They could hardly be expected to sound the same, and indeed they do *not* sound the same. How a mix would 'travel' from one pair of these loudspeakers to another would depend on the musical arrangement, the frequency range which the music covered, and many other factors. The musical style, or genre, may determine on which loudspeakers the mix was deemed to sound most right, and this may, in turn, lead people to false conclusions about which loudspeaker was *generally* most right. Surely, though, such vaguearies can hardly be considered to be a desirable part of any 'reference monitoring' process. From Figures 10.1 to 10.3 it is easy to understand how pink noise would sound very different if played through each pair of loudspeakers in turn. It should be equally easy to appreciate how spectrally complex music would suffer a similar fate.

With such a widely varying range of loudspeakers in use as so-called 'references', there is little wonder that no, one, mix, done on *any* set of

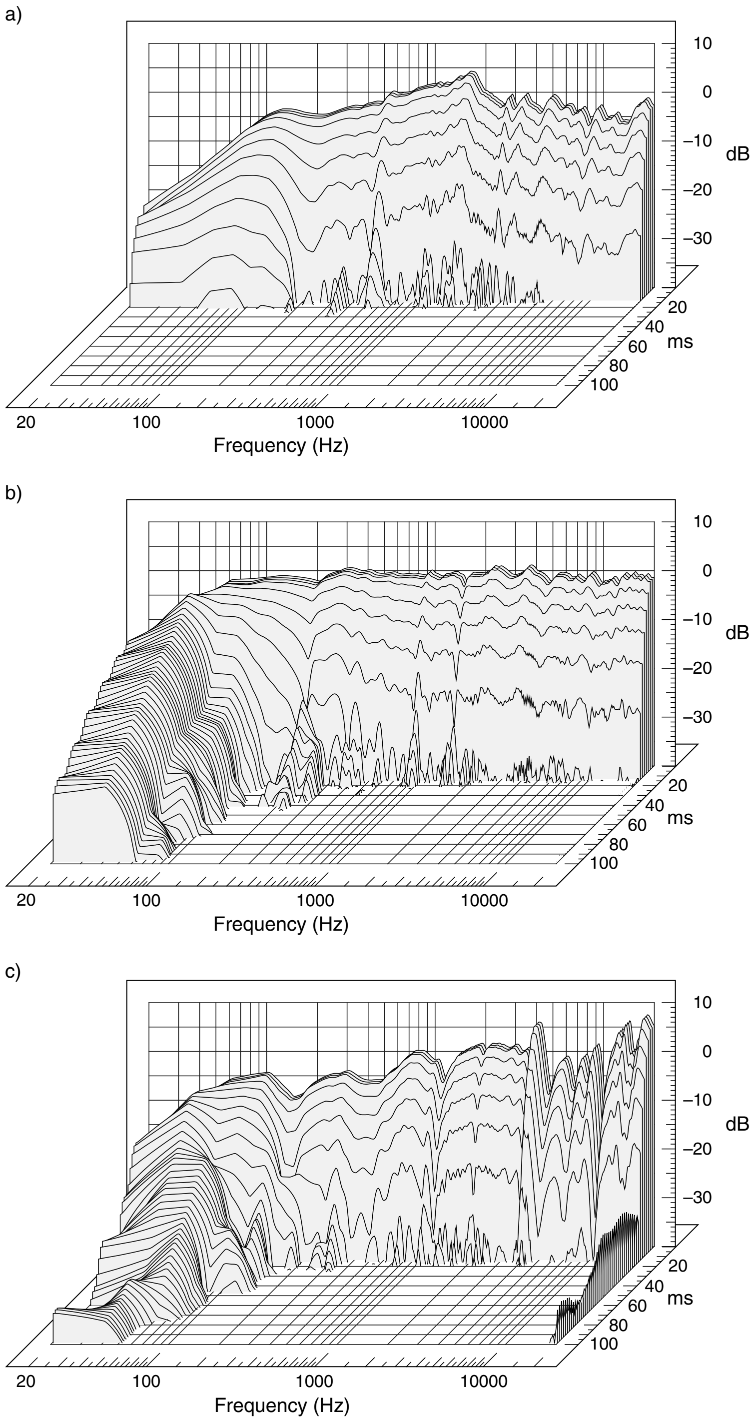


Figure 10.3 Waterfall plots. a) Yamaha NS10M. b) Genelec S30D. c) SAE TM160A

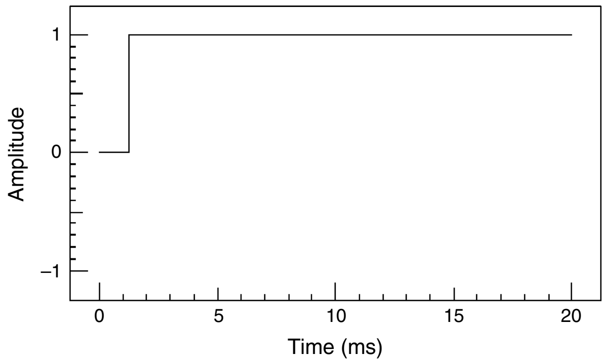


Figure 10.4 An electrical step-function waveform

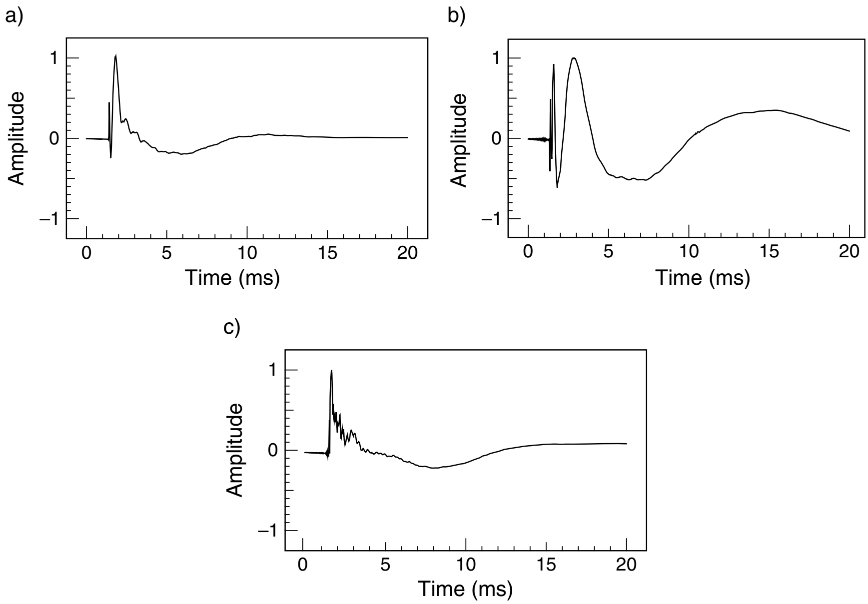


Figure 10.5 Step-function responses of the three loudspeakers shown in Figures 10.1 to 10.3. a) Yamaha NS10M. b) Genelec S30D. c) SAE TM160A

monitor loudspeakers, can be expected to sound equal on all of them. Nevertheless, a well-distributed musical balance, of well-arranged music, can go some considerable way to ameliorate the problems, but much modern music is very complex. Especially when music has been produced entirely on one set of loudspeakers, it is inevitable that it will tend to sound wrong on many other systems in many different rooms. The fact that so many *different* references are available seems to devalue the very concept of their use *as* references. Furthermore, the ear, itself, especially as part of an overall system of musical perception, has also been shown to be a rather variable reference, and it is usually only a combination of experience and confidence which can solve this problem. This is exactly what people are usually paying for when they engage the services of a reputable mastering engineer.

It is also interesting to note that in general, classical recordists tend to use different loudspeakers to rock music recordists. This fact serves to highlight how far we are from perfection in loudspeaker design, because a perfect loudspeaker would be an incontrovertible reference, *unless*, that is, the buyers of one type of music recording had a general trend towards all using a specific genre of imperfect loudspeaker in their homes, and to which it seemed appropriate to make mixing compromises. To some extent this case actually exists, with the vast majority of domestic listeners to rock music using small, bass reflex enclosures, the problems of which will be discussed further in the next chapter.

It seems to be the case that highly experienced and successful recording personnel stick tightly to their chosen monitor systems, and they are *very* reluctant to change them – even when they know them to be lacking. Conversely, many inexperienced recording personnel provide the bread and butter for the mid-priced monitor loudspeaker industry, rushing to change their references every time that somebody complains about the failure of a mix to travel well. The reality is that in the latter case, the problem usually lies in the music and the mix, and not in the loudspeakers. A robust musical arrangement, well played and well mixed, *will* travel. A complex mix of a poor arrangement will probably not fare so well.

Acknowledgement

This chapter was partly inspired by the following paper which, although on a different subject, highlighted many parallel concepts.

Bailey, Mark; 'Perception of Music: The Element of Surprise', Proceedings of the Institute of Acoustics, Vol 24, Part 8, Reproduced Sound 18 conference, Stratford-upon-Avon, UK (November 2002)

Chapter 11

Low frequency and transient response dilemmas

11.1 The great low frequency deception

Recording industry publications often contain advertisements proclaiming the true and lifelike responses of 'accurate' small monitor systems. If only from the sheer quantity of such advertising, it is entirely reasonable that many people would be led to believe that accurate low frequency reproduction from small boxes at quite high monitoring levels was an easily achievable goal, but the reality is rather different. At realistic listening levels for recording studio or mastering use, the low frequency response of small loudspeakers cannot be as accurate in terms of frequency response *and* transient response as that of a good large system, flush mounted in the front wall of a well-controlled room. The laws of physics simply will not allow it.

Various techniques can be used to flatten and extend the low frequency pressure amplitude of loudspeakers in small-to-medium-sized cabinets, but what so many people are unaware of is the degree to which these methods of response extension can distort the time responses of the systems, and mask considerable amounts of the low-level detail which a complex musical signal may contain. The magnitude of the problem is clearly depicted in Figure 11.1, from which it can be seen that of the 38 specimens tested, when the time, frequency and pressure responses are viewed together, no two out of the 38 plots look the same. It can also be added, without fear of contradiction, that no two loudspeakers *sound* the same, either. Using these loudspeakers as monitoring references therefore is more a question of interpretation rather than anything absolute.

The variability between these 'reference' loudspeakers is made all the more alarming when one considers two further points. Firstly, that all the measurements shown in Figure 11.1 were made in the same position in the same, large (611 m^3) anechoic chamber. Obviously, when one uses these loudspeakers in typical control rooms or domestic rooms, the responses will be further adulterated. Secondly, it must be appreciated that amongst the 38 examples are some very fine loudspeakers, in fact all were submitted for test by manufacturers who were proud of what they had achieved, and all were presented as professional music-monitoring loudspeakers. Therefore, as these loudspeakers represent 'the higher end' of loudspeaker production, it is lamentable to think how poor the responses can be at the lower end of the market. Furthermore, it should also be obvious that if these plots represent the best that such famous manufacturers can do, then we are not dealing with any problems that can be easily solved. So, perhaps we should now look at the implication of what the plots represent, and analyse the problems step by step.

11.1.1 The air spring

Moving coil loudspeakers in boxes are volume-velocity sources. The acoustic output is the product of the area and the velocity of the diaphragm, so, for any given output, either a large volume of air can be moved slowly, or a small volume of air can be moved quickly. Let us say that, for a given SPL, the diaphragm of a 15 inch (380 mm) woofer in a 500 litre box moved 2 mm peak to peak. With an effective diaphragm radius of 6½ inches, which is about 160 mm, (we do not count the surrounds as part of the radiating area), the radiation area would be 80,000 mm^2. Moving 2 mm peak to peak means moving 1 mm from rest to the peak in either direction, so the unidirectional displacement would be $80{,}000 \times 1\,mm^3$ or 80,000 mm^3. This is equal to 0.08 litres, so the static pressure in a 500 litre box would be compressed (if the cone went inwards) by 0.08 litres, or by one part in 6250 of the original volume.

For the same SPL, a 6 inch (150mm) loudspeaker in a 10 litre box would still need to move the same amount of air (80,000 m^3). However, with an effective piston radius of only 2½ inches, or 65 mm, the cone travel would need to be about 12 mm peak to peak, or 6 mm in either direction, so it would also need to travel six times faster than the cone of the 15 inch loudspeaker, (i.e. 6 mm instead of 1 mm in the same period of time for any given frequency). What is more, the displacement of 80, 000mm^3 (0.08 litres) in a box of only ten litres would represent a pressure change in the box of one part in 125 of the original volume, compared to one part in 6250 in the 500 litre box. The air compression inside the box would therefore be 50 times greater than that in the 500 litre box, and there are several consequences of these differences.

Anybody who has tried to compress the air in a bicycle pump with their finger over the outlet will realise that air makes an effective spring. They will also realise that the more the air is compressed, the more it resists the applied force, and the bicycle pump can rarely be compressed much more than about half way. The force needed to compress the air by each subsequent cubic centimetre increases with the compression, so the process is not linear. In the case of the 15 inch and 6½ inch cones referred to above, the cone in the small box would have a much harder job to compress the air by 1/125th of its volume than the cone in the large box, which only needs to compress the air by 1/6250th part in our previous example. (Cone *size* is irrelevant, here – only the volume displacement matters.) Large boxes therefore tend to produce lower distortion at low frequencies, because the non-linear air compression is proportionally less. The concept is shown diagrammatically in Figure 11.2.

The non-linearity of the air spring can perhaps also be better understood when one considers that it would take an infinite force to compress 1 litre of air to zero volume, yet it would take only a moderate force in the opposite

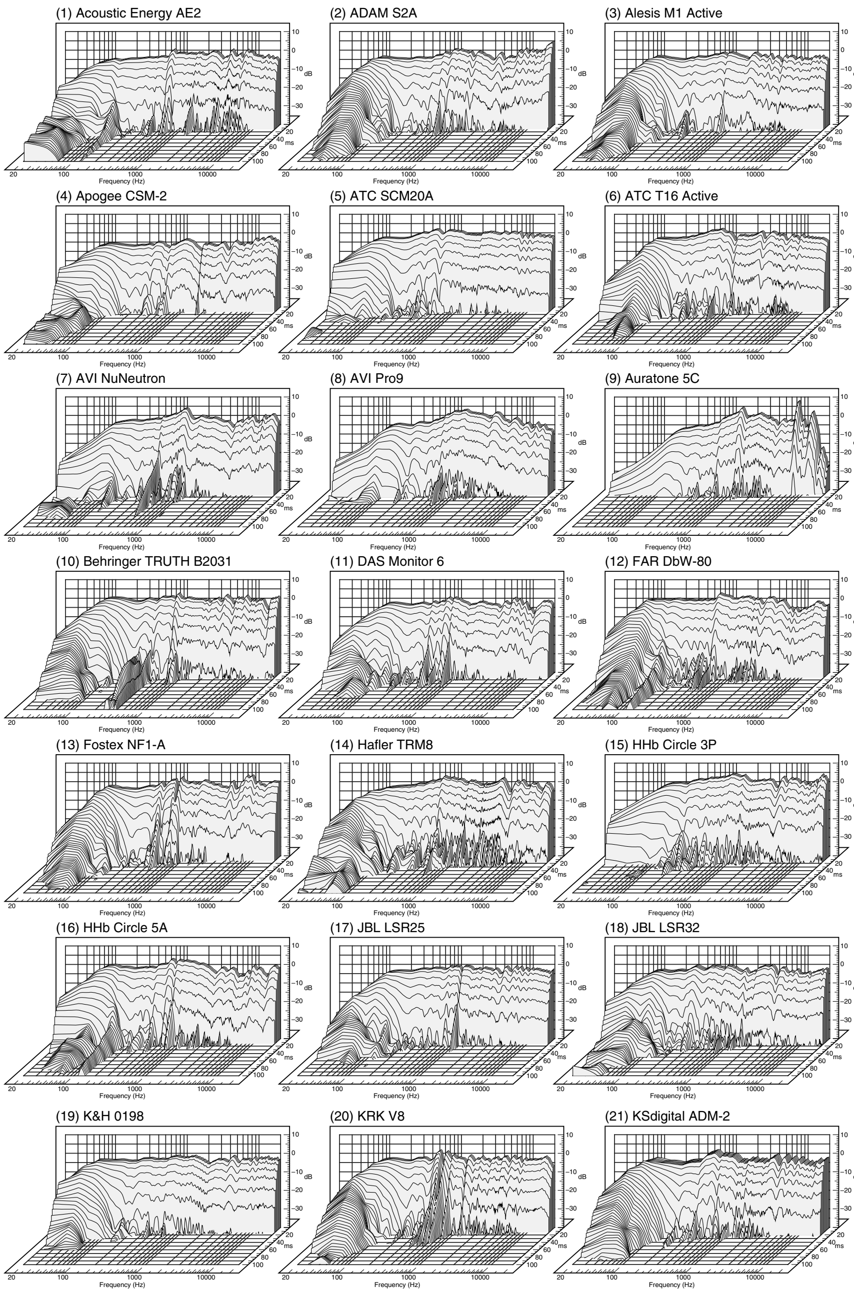


Figure 11.1 Waterfall plots of the anechoic responses of 38 loudspeakers

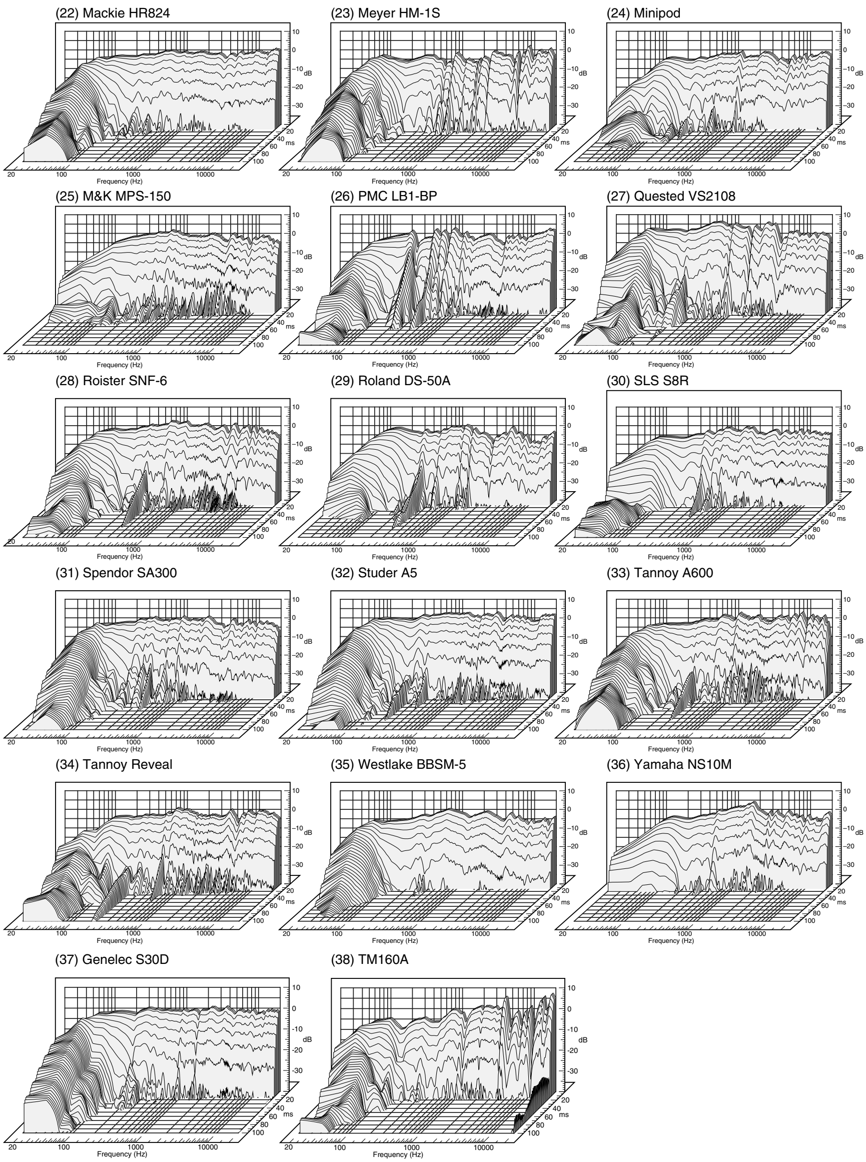


Figure 11.1 Continued

direction to rarefy it to 2 litres. The forces needed for a given change in air volume in each direction (in this case ±1 litre) are thus not equal, so the restoring forces applied by the air on the compression and rarefaction half cycles of the cone movement are also not equal. The non-linear air-spring forces thus vary not only with the *degree* of displacement, but also with the *direction* of the displacement. Changing air temperatures inside the boxes also add more complications of their own, and waste-heat from the voice

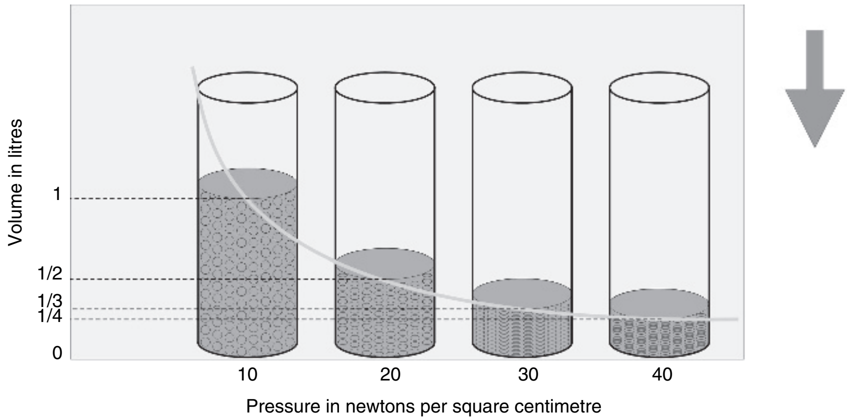


Figure 11.2 Boyle's law
Each pressure change of 10 newtons produces progressively less change in the volume of the gas. The process is therefore not linear, and can give rise to harmonic distortion

From this it can be seen that the more that a gas is compressed, the more it will resist further compression. The line passing through the cylinders is a curve, *not* a straight line, and this gives rise to the non-linear distortion (principally harmonic distortion) due to the back-loading on the diaphragm of a low frequency driver. In a big box, the relative compression is less, so for any given displacement the non-linear distortion due to the effect of the air spring will also be less

coils when the loudspeakers are being driven by a musical signal ensures that the temperature of the air *will* change during use. Therefore, for any given SPL, loudspeakers in small boxes tend to produce more distortion than similar loudspeakers in larger boxes.

11.1.2 Size, weight and sensitivity

The main controlling factor for the extension of the low frequency response of a loudspeaker system is its resonant frequency, because the low frequency response of any conventional loudspeaker system will begin to fall off quite rapidly below the resonant frequency. There are systems which drive the loudspeakers well below resonance, with the application of electronic compensation, but they can run into problems of cone excursion limitations, and are not in widespread use. The resonance is a function of the stiffness of the air spring formed by the air inside the box, coupled with the moving mass of the loudspeaker cone/coil assembly. The fact that the air inside a small box presents a stiffer spring than the air in a larger box, (because it is proportionally compressed more for any given volume displacement) means that it will raise the resonant frequency of any driver mounted in it, compared to the same driver in a larger box (i.e. loaded by a softer spring). The only way to counter this effect, and to lower the resonant frequency to that of the same driver in a larger box, is to increase the mass of the cone/coil assembly. [Imagine a guitar string; if it is tightened, the pitch will increase. Maintaining the same tension, the only way to lower the note is to thicken the string, i.e. make it heavier.]

The problem that one now encounters is that to move the heavier cone, in order to displace it by the same amount as a lighter cone in a larger box, more work must be done, so more power will be needed from the amplifier. The sensitivity of a heavy cone in a small box is therefore less (for the same resonant frequency and bass extension) than for a lighter cone in a larger box. If the sensitivity is to be maintained in a smaller box, by using the same weight of cone and coil, then inevitably the resonant frequency will be forced upwards. So, as the box size decreases, either the bass extension or the sensitivity (or both) will be reduced. The only way to maintain the bass extension *and* the sensitivity is to increase the size of the box or the size of the magnets.

Larger boxes often tend to use larger drive units. A large diaphragm will tend to be heavier than a smaller one, and the large diaphragm may also need to be heavier to maintain its rigidity, as discussed in Section 7.2.2. This would suggest a lower sensitivity in free air, but larger drive units often also have larger magnet systems, which can easily restore the sensitivity, and a larger radiating area, in itself, increased the radiation efficiency, as was discussed in Section 7.2.2. Nevertheless, in small boxes, the greater pressure changes may require stiffer, heavier cones, in order not to deform under high pressure loads, so the efficiency can again be caused to reduce. Once more, a bigger magnet could be an answer, but it may not be a simple task to use a bigger magnet because it could seriously obstruct the free air movement behind the cone, and it could reduce the internal volume of the box, hence further stiffening the spring and again raising the resonant frequency, which could only be offset by making the cone assembly even heavier. If this were the case, the extra weight of the cone would have to be compensated for by using more power to drive it, and because putting in more power means that it would probably need a bigger (and heavier) voice coil to take the extra power, it could therefore need even more power to restore the output. So, now it should be becoming obvious that there are just *so* many things which conspire against the extended low-frequency performance of small boxes. There is no way out – we just keep going round in circles.

Let us now consider two actual loudspeakers of similar frequency range but very different size. A small loudspeaker such a the ATC SCM10 would need to be driven by almost 200 watts in order to give the same SPL at one metre as the double 15 inch (380 mm) woofered UREI 815 receiving *one* watt of input. The two loudspeakers are shown in Figure 8.11. As has just been discussed, reducing the box size demands that either the low frequency response will be reduced, or the sensitivity will be reduced. High sensitivity *and* good low frequency extension can *only* be achieved in large boxes. If the ATC seeks to achieve a good low frequency extension in a small box, then the sensitivity must be low; the air-spring physics dictates that this must be so. The ATC SCM10 has a box volume of about 10 litres; the UREI 815 contains almost 500 litres. Given that they cover the same frequency range, the sensitivity difference of about 22 dB is the result – hence the ATC needing almost 200 watts to sound as loud as the UREI receiving *one* watt.

11.1.3 Further consequences of small size

When small cones move far and fast, they also tend to produce more Doppler distortion (or frequency modulation), and this problem is often exacerbated by the small woofers being used up to higher frequencies than the large woofers, which can make the Doppler distortion more noticeable. The high frequencies are being radiated from a diaphragm which is moving backwards and forwards with the low frequency signals. (For more about Doppler distortion see the Glossary.)

Long cone excursions also mean more movement in the cone suspension systems (the surrounds and the spiders), which also tend to be non-linear in nature as cone excursions increase. That is, the restoring forces are rarely uniform with distance travelled. This tends to give rise to higher levels of intermodulation and harmonic distortion than would be experienced from larger cones of similar quality, moving over shorter distances. The larger movements also require greater movement through the static magnetic field of the magnet system, which tends to give rise to greater flux distortion and, even more audible, non-linear Bl profile distortions.

Furthermore, the reduced sensitivity of the smaller boxes means that more heat is expended in the voice coils compared to that produced in the voice coils of larger loudspeakers for the same output SPLs. This problem is even further aggravated by the fact that the smaller loudspeakers have greater problems in dissipating the heat, because there is less air surrounding them in the smaller boxes. This can lead to thermal compression, as the hotter the voice coil gets, the more its resistance increases, so the less power it can draw from the amplifier for any given output voltage. The resulting power compression produces yet more distortion products, so it can clearly be seen that the distortion mechanisms acting on small loudspeakers are far greater than those acting on similarly engineered large loudspeakers. And even that is not all; small cones rapidly punching thought the air can produce turbulence, which can be a source of strange noises due to the shearing of the air at the edges of the cone. There are thus many reasons why large-coned drivers moving over short distances tend to produce less distortion than smaller ones, of similar quality, moving over larger distances. All of these reasons militate *against* the low frequency performance of small loudspeakers.

11.2 Commercial solutions

The commercial pressures on loudspeaker manufactures tend to come from people who are largely ignorant of these problems. The typical customers demand more output of a wider bandwidth from ever smaller boxes, so loudspeaker manufactures try to rise to the challenge. One example of a technique used to augment the low frequency output is to use a reflex loaded cabinet, (as described in Chapter 3) with one or more tuning ports. In these systems, the mass of air *inside the ports* resonates with the spring which is created by the air trapped within the cabinet. If the resonant frequency is chosen to be just below where the driver response begins to roll-off, then the overall response can be extended. The resonance in the tuning ports begins to radiate sound just where the drivers begins to lose their output, and so the overall response can be extended downwards.

The effective extension of the low frequency response by means of reflex loading also increases the loading on the rear of the driver as resonance is approached. This helps to limit the cone movement and to protect the drivers from overload. Unfortunately though, once the frequencies pass below resonance, the air merely pumps in and out through the ports, and all control of the cone movement is lost. In many active monitor systems, electrical high-pass filters are used to sharply reduce the input power to the drivers at frequencies below the cabinet resonance frequency. This enables higher acoustic output from the loudspeaker systems within their intended bandwidth of use, without the risk of overload and mechanical failure due to high levels of programme below the resonance. By such means, a flat pressure amplitude response can be obtained to a lower frequency than with a sealed box of the same size, and the maximum SPL can be increased (typically 3–4 dB for similar sized boxes) without risking drive unit failure, but there is a price to be paid for these gains. The phase response will be compromised, and hence the uniformity of the time (transient) response will tend to be lost.

11.2.1 The time penalty

It must be understood that a resonant system can neither start nor stop instantly. The time response of reflex loaded loudspeakers therefore tends to be longer than that of similar sealed box versions, which means that transients will be smeared in time: the impulse response will be longer. Moreover, the effect of the electrical high-pass filters is to further extend the impulse response, because the electrical filters are also resonant (tuned) circuits. In general, the steeper the filter slope for any given frequency, the longer it will ring. More effective protection therefore tends to lead to greater transient smearing. Figure 11.3 shows the low frequency decay of a sealed box loudspeaker, with its attendant low frequency roll off. Figure 11.4 shows the low frequency response of an electrically protected reflex cabinet of somewhat similar size. Clearly the response shown in Figure 11.4 is flatter until a lower frequency, but a flat frequency response is not the be-all and end-all of loudspeaker performance. Note how the response between 20 Hz and 100 Hz has been caused to ring on, long after the higher frequencies have decayed.

Figure 11.5 shows the corresponding step-function responses, and Figure 11.6 the acoustic source plots. These plots clearly show the time response of the reflex cabinets to be significantly inferior to the sealed boxes. The low frequencies from the reflex enclosure arrive later (as can be seen from Figure 11.6), and take longer to decay (as can be seen from the decay tails in Figure 11.5), which both compromise the ‘punch’ in the low frequency sound. Figure 11.7 compares the two plots of Figure 11.6 with the acoustic source plot of a large, wide-band, flush-mounted studio loudspeaker system in a well-controlled room. From this comparison it should be obvious why the NS10M (and Auratone 5C before it) earned a reputation as a punchy little ‘big’ monitor.

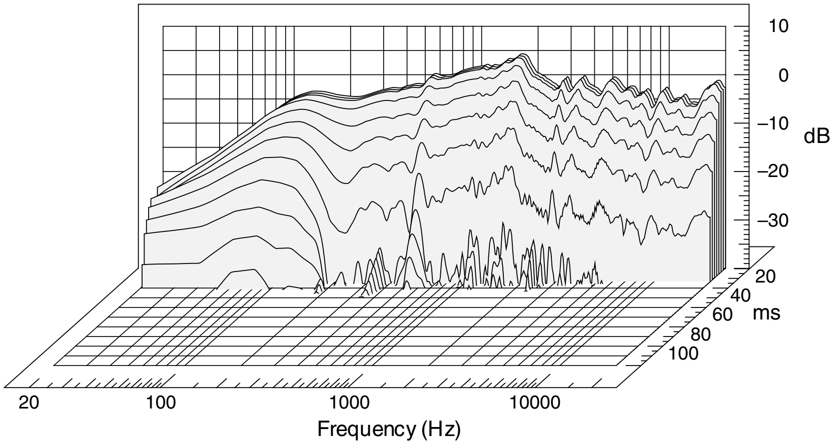


Figure 11.3 Waterfall plot of a small, sealed-box loudspeaker (NS10M)

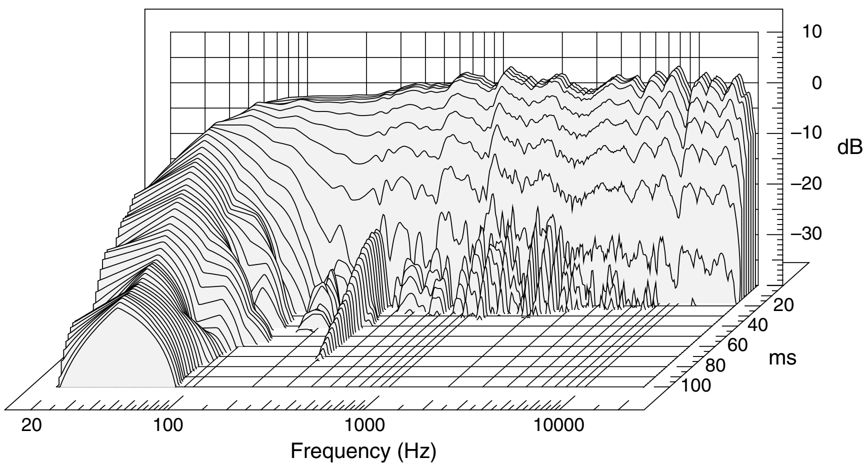


Figure 11.4 Waterfall plot of a small reflex (ported) enclosure of similar size to the loudspeaker whose response is shown in Figure 11.3

A sealed box cabinet will exhibit a 12 dB per octave roll-off below resonance, but a reflex enclosure will exhibit a 24 dB per octave roll-off when the port output becomes out of phase with the driver output. As the system roll-offs are often further steepened by the addition of electrical protection filters below the system resonance, sixth, and even eighth order roll-offs (36 dB and 48 dB per octave, respectively) are quite common. With such protection, some small systems *can* produce high output SPLs at relatively low frequencies, but the time (i.e. transient) accuracy of the responses may be very poor. We will return to this topic in the next sections of this chapter.

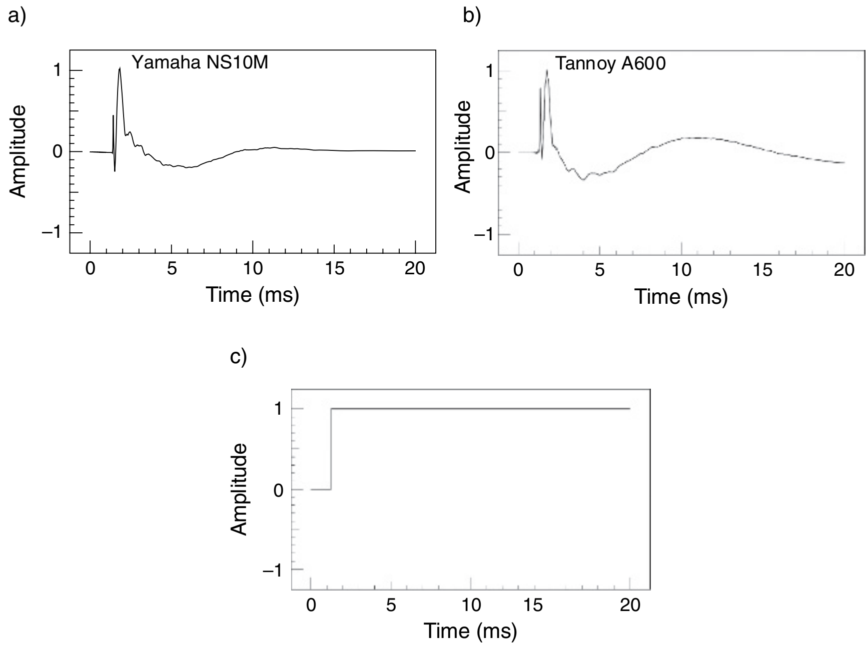


Figure 11.5 Step function responses corresponding to the waterfall plots shown in Figures 11.3 and 11.4, compared to the electrical input signal shown in (c). Note how rapidly the NS10 (a) returns to a flat line on the zero amplitude level

Inevitably, the different resonances of the different systems will produce musical colourations of different characters. This may not be a serious problem for use in domestic listening, but such inconsistency of colouration does little to help the confidence of the users in recording studios. If a mix sounds different when played on each system, then how does one know which loudspeaker is most right, or when the musical balance of the mix is correct? The resonances of the sealed boxes tend to be better controlled, and are usually much more highly damped than their reflex-loaded counterparts. This leaves the magnitude of the frequency response of sealed boxes as their predominating audible characteristic, but it is usually the *time* responses of reflex enclosures (related to the 'phase' part of the frequency response) which give rise to their different sonic characters. There is considerable evidence to suggest that the many years of use of the Auratone 5C and Yamaha NS10Ms as mixing monitors has been due to their rapid response decays. It can be seen from their response plots (Numbers 9 and 36 in Figure 11.1) that all the frequencies decay at an equal rate – none can be seen to be hanging on after the other frequencies have disappeared.

A simple roll-off in the low frequency response of a loudspeaker used for *mixing* is in itself not a great problem, because any wrong decisions about a balance can usually be corrected by equalisation at a later date, such as during mastering. As previously stated in other chapters, an error in

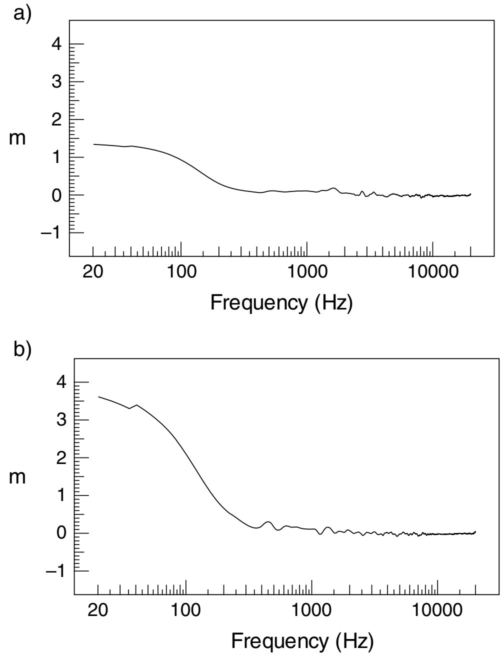


Figure 11.6 Acoustic source plots corresponding to the waterfall plots shown in Figures 11.3 and 11.4

In these plots the low frequency response delay is shown in terms of from how many metres behind the loudspeakers the low frequencies are apparently emanating. As each metre corresponds to about 3 milliseconds, it can be appreciated how the low frequencies from behind the sealed box (a) arrive much more 'tightly' with the rest of the frequencies than they do in the case from the reflex cabinet

the *time* response, such as that added by tuning port and filter resonances, can lead to misjudgements especially between the percussive and tonal low frequency instruments, such as bass drums and bass guitars, whose relative levels *cannot* be adjusted once they have been mixed together. The time response errors of the loudspeakers will have led to erroneous mixing decisions which cannot be unscrambled. The concept will be treated in a more definitive manner in Section 11.7.

11.2.2 The transient trade-off

A problem therefore exits in terms of how we can achieve flat, uncoloured, wide-band listening at relatively high sound pressure levels from ten-litre boxes. At the moment, basically, the answer is that we *cannot* do it. Just as there is a trade-off between low frequency extension, low frequency sensitivity/efficiency (see Glossary) and box size; there is also a trade-off between low frequency SPL, bass extension, and transient accuracy if bass reflex loading and electrical protection are resorted to in an attempt to

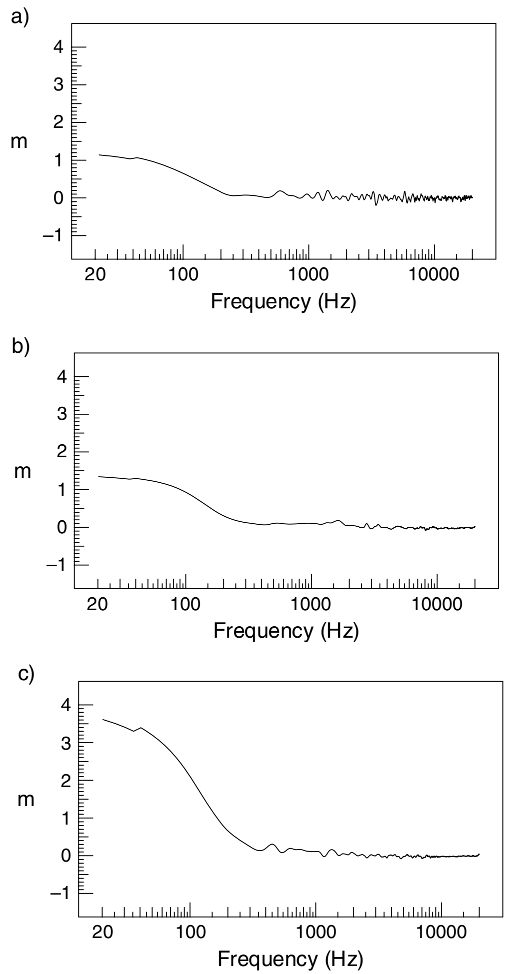


Figure 11.7 Acoustic source plots of Figure 11.6 compared to the acoustic source plot of a large, flush-mounted, studio monitor loudspeaker. a) A 700 L, wide-range, flush-mounted studio monitor loudspeaker. b) A Yamaha NS10M. c) A popular, small, reflex-loaded, close-field loudspeaker system. Note how the response of the NS10M mimics that of the large monitor loudspeaker system. There is therefore little wonder that the NS10M has a reputation for having a 'rock and roll punch'

defeat the box size limitations. In effect, the bass extension is gained at the expense of transient accuracy.

We therefore have a state of affairs whereby, at *low* SPLs, good low frequency extension *can* be achieved from small boxes, but the non-linearity of the internal air-spring can lead to high distortion when the cone excursions, and hence the high degrees of internal pressure changes, become significant. Suspension and magnet system non-linearities can add further problems, and remember also that in a small box there often exists the problem of how to get rid of the heat from the voice coil. Thermal overload and burnout are always a problem at high SPLs due to the high power necessary to overcome the limitations of the poor system efficiency. Larger, higher-sensitivity systems not only produce less heat for any given SPL, but also are much better at dissipating it. They thus win on both counts.

From the waterfall plots of Figures 11.3 and 11.4 it can be seen that, whatever the box type, the decay is never instantaneous. There is always a slope to the time representation, which in these plots is depicted by the time 'slices'. One can imagine the slope of the plot in Figure 11.3 continuing below the 'floor' formed by the frequency and time axes, with the lines of the 'waterfall' continuing to cascade down. The question has often been asked whether the electrical flattening of the low frequency response would inevitably lengthen the time response, even with the sealed boxes – effectively extending the response on the time scale as the level was brought up from below the 'floor' on the amplitude (vertical) scale. In truth, the tendency is for the flattening of the amplitude response to *shorten* the time response (i.e. steepen the slope) by means of its correction of the phase response errors which are associated with the roll-off. This means that a large or small sealed box, equalised or not, would still exhibit a much faster time response than a reflex enclosure. Figure 11.8 shows the comparative effect.

Unfortunately, the type of equalisation shown in Figure 11.8 is not a practical option, because even at very low sound pressure levels, the excursion limits of a small drive unit would be exceeded at the lowest frequencies, and the drive-power levels would double every time that an extra 3 dB of boost was applied. Figure 11.8 only serves to show how, in principle, corrective equalisation would beneficial to the entire time/frequency/phase response.

It seems reasonable that the more extended bass response of reflex enclosures may well be valuable to 'vibe' the musicians during the recording process, where concentration is more on achieving a good performance rather than necessarily looking at the subtleties of each sound. However, during the mixing process, another, more critical view is required, and hence perhaps a different set of loudspeakers. At the mastering stage, the requirements for sonic transparency can become even more critical, and hence we have arrived at the sort of differentiation of uses described at length in Chapter 8.

There are those who say that very fast time responses are not necessary from loudspeakers, because their decay times are considerably shorter than many of the rooms in which most of them will be used. However, what they fail to realise is that the small loudspeakers are usually being used in the close-field, which is normally considered to be within the critical distance where the direct sound and room sound are equal in level. It therefore follows that if one is listening in the close-field, the responses of the loudspeakers will predominate in the total response. Indeed, this is one of the principal reasons for the use of close-field listening. The room decay will therefore *not* totally mask the loudspeaker decay, except in rooms which are so live that the close field extends only a matter of centimetres from the loudspeakers, but such rooms would hardly be appropriate for monitoring.

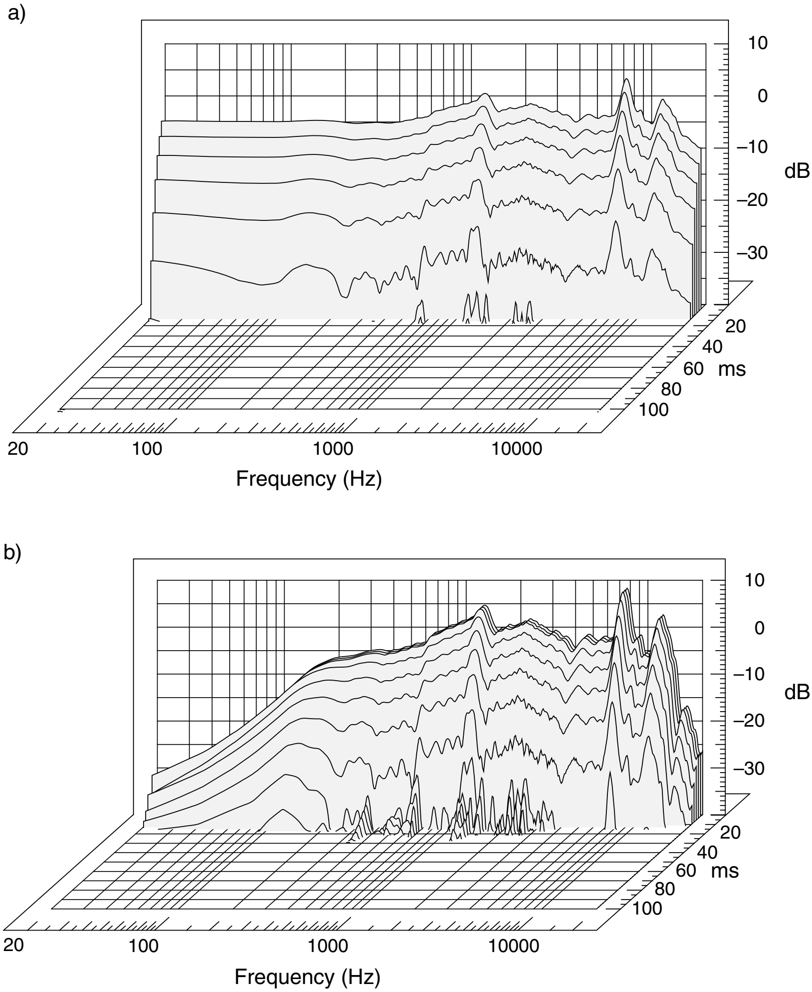


Figure 11.8 Time vs. frequency – the effect of equalisation on a sealed box loudspeaker. a) Waterfall plot showing the effect on the time response of electrically flattening the response of an Auratone 5C. Although the response is now flatter down to a lower frequency than the reflex enclosure shown in Figure 11.4, the time response has not been extended – the decay is much more rapid. Unfortunately, such equalisation is not a practical solution because the loudspeaker would overload even at low SPLs. Compare this plot with the unequalised response shown in (b), from which it can be seen that the time response has not been lengthened in the slightest by the corrective equalisation

11.3 The evolution of the desk-top monitor

It came as a bit of a shock during interviews with some highly respected recording personal to find not only how little they knew about loudspeakers, but also how little they *cared* about knowing what was going on inside the boxes. They seemed only to be interested in whether they worked, or not, for their own decision-making during recording or mixing sessions. In effect, it seems that they have largely given up trying to understand a subject which usually appears to be cloaked in so much mystery. The problem with this is that it provides very little feedback to the loudspeaker designers, and the trend towards aggressive marketing of the products also does little to increase the breadth of understanding of the users.

Somewhat disturbingly, even the designers of many professional loudspeaker system are under the influence of powerful marketing pressures. When they receive instructions from their paymasters to design a new loudspeaker, sound quality can be as low as fourth or fifth on the list of design priorities. Things such as to be 100 euros a pair cheaper than the perceived competition, an eye-catching design that will look good in advertisements, comparable size to a competitor, and to be louder than the nearest competing models are often typically seen to be more important than absolute sound quality by the largely all-powerful marketing/sales people, even in a supposedly *professional* recording industry.

However, despite this, mastering engineers, recording engineers, producers and musicians *have* been able to find some common paths, which, even if they have not been clearly marked by technical guidelines, have nevertheless been leaving some clear sonic footprints. Look, for example at Figure 11.9. How many users of Auratones and NS10s have ever realised that they possessed such similar frequency responses, (at least below 6 kHz) or how different their 'inverted V' anechoic frequency responses were from most other loudspeakers? The waterfall plots shown in Figure 11.10 (which show pressure amplitude against time against frequency) are also very similar, as are the step function responses shown in Figure 11.11. These plots help to explain why such a large number of recording engineers and producers moved to the NS10 as a louder and deeper replacement for their trusted Auratone 5Cs. The NS10 basically gave them more SPL and more bass than the Auratones, but the general characteristics of the sound remained the same. The reasons why only became apparent about 20 years after the switch had taken place by the judgement of ears, alone[1].

The plots in Figure 11.9 are anechoic chamber responses, and it is clear that in the frequency domain they are not flat with frequency. Technically, they seem wrong, but sonically, at least for achieving a musical balance between instruments, many users find them to be right when used in their typical console-top locations, the effect of which on the pressure response was shown in Figure 8.12.

Let us consider the implications of Figure 8.12(a), shown again, here, as Figure 11.12. The solid line shows the predicted response of a loudspeaker such as an NS10M in a free-field (as in an anechoic chamber). The shape is not unlike the measured anechoic responses of the NS10M and the Auratone 5C, as were shown in Figure 11.9. The broken line shows the predicted response when such a loudspeaker is flush mounted in a wall. When one considers that the NS10M was originally designed as a bookshelf loudspeaker, the logic behind its anechoic response should now be clearly apparent. In fact, if sited next to *any* large reflective surface, such as the top surface of a large mixing console when placed on its meter bridge, the response would also tend towards that of the broken line. Figure 11.13 shows the actual response of an NS10M on a meter bridge in

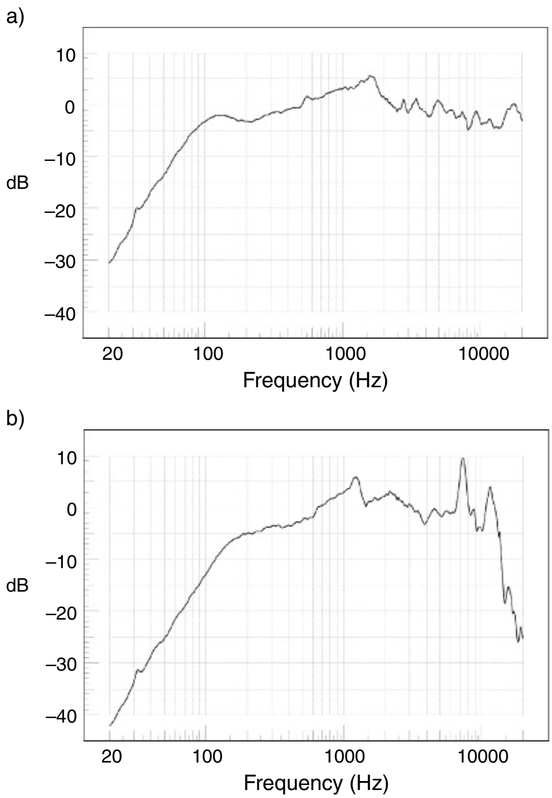


Figure 11.9 Comparisons of pressure amplitude responses of the NS10M and the Auratone 5C loudspeakers. a) NS10 anechoic frequency response plot. b) Auratone 5C anechoic frequency response plot. Note the general similarity of the frequency responses (apart from the peak at 8 kHz in the Auratone's response). The fact that they are *not* flat is also worthy of note

a typical recording environment. Obviously the desk (console) tops cause many response irregularities, but, the *general* tendency towards an overall response flatness is also very apparent in Figure 11.13. The low frequency response has clearly been augmented, and what is more it has been augmented in a non-resonant manner that does not affect the rate of decay of the low frequencies.

The great significance of this is that if a loudspeaker which had a flat *anechoic* response were to be similarly positioned, it would also be subjected to the same type of response modifications, and so would tend to produce an excess of low frequencies. This is the reason why many active loudspeakers have bass frequency adjustment switches, with recommended settings for different mounting conditions. With passively crossed-over loudspeakers, this flexibility in response correction is only usually achievable by the use of external equalisers, but it is seldom a practical solution. Low frequency response controls can be incorporated into an active design for a few cents, but a dedicated equaliser with the sonic transparency necessary for high quality monitoring may be vastly more expensive than the

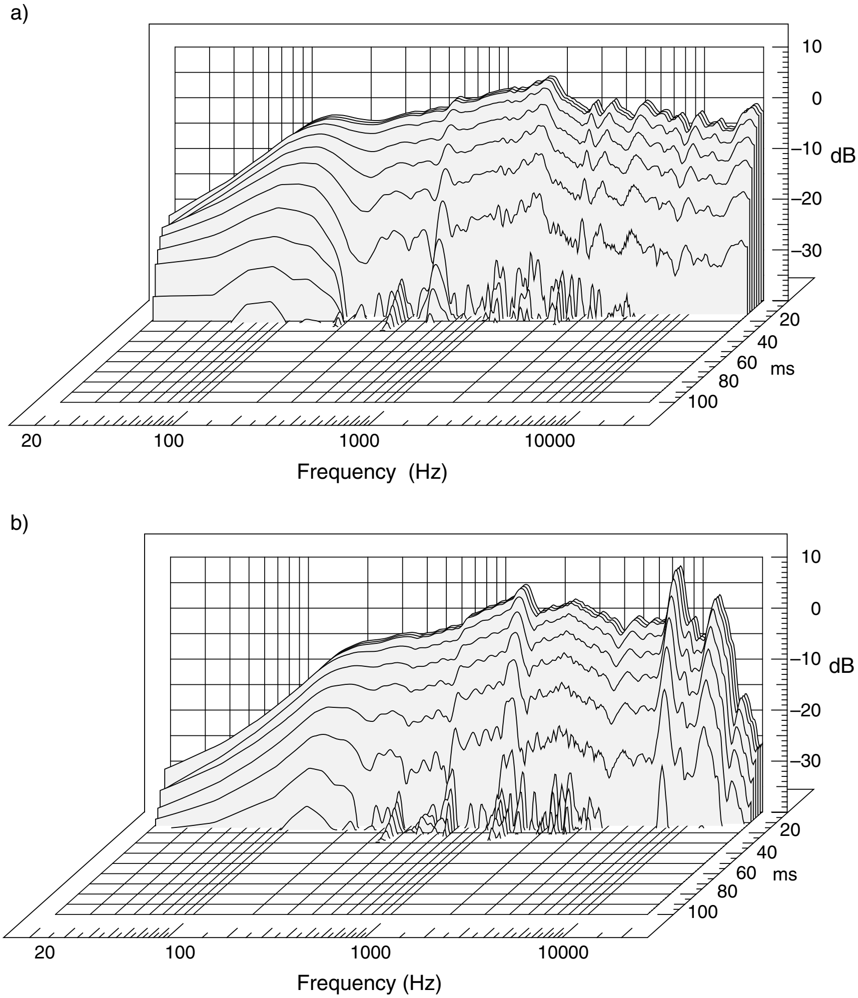


Figure 11.10 Comparison of the waterfall plots of the NS10M and the Auratone 5C loudspeakers. a) Waterfall plot of NS10M. b) Waterfall plot of Auratone 5C. By 20 milliseconds after the wideband excitation has ceased, the response at almost *all* frequencies has decayed below the −40 dB level. This is *not* the case with most small reflex loudspeakers, of which the plot of Figure 11.4 is more typical

loudspeakers with which they are being used. This tends to discourage their use. Obviously, the use of a cheaper equaliser, with its own sonic character, would be an absurd choice if accurate monitoring were the goal. Passively crossed-over loudspeakers therefore tend to be best chosen with responses appropriate for their conditions of mounting, but a large proportion of the people involved in the recording world seem not to understand this rather important point.

The upshot of this is that loudspeakers get located in positions that are physically practical for working, but which may not be conducive to

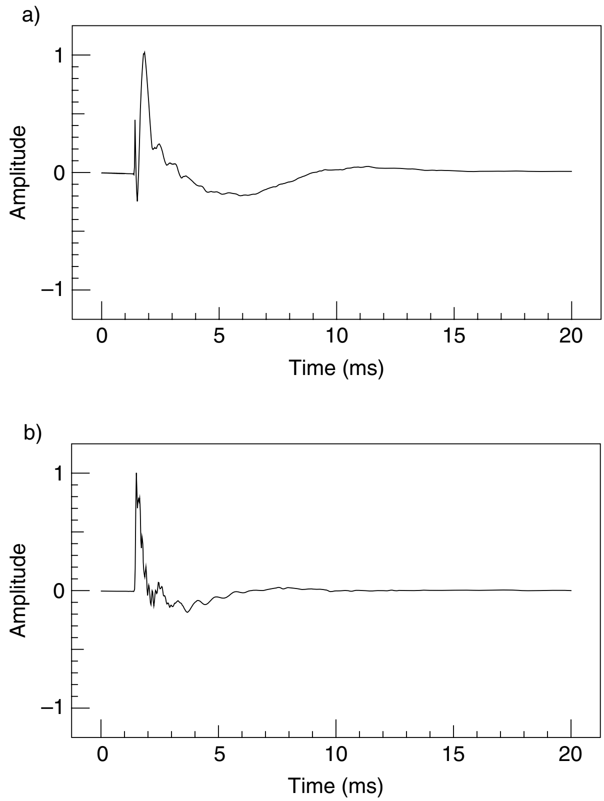


Figure 11.11 Comparison of the step-function responses of the NS10M and Auratone 5C loudspeakers. a) NS10M step function. b) Auratone 5C step function. By 8 milliseconds after the impulsive excitation, the responses have decayed to a flat line. Again, this is not typical of most loudspeakers of which, once again, the step-function response shown in Figure 11.5(b) is perhaps more typical

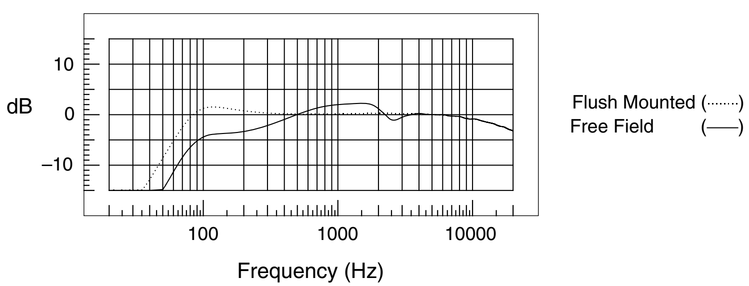


Figure 11.12 Response of an idealised loudspeaker, of similar size to an NS10M, under different conditions of mounting

The solid line shows the free-field response, and the dotted line shows the expected response if flush-mounted

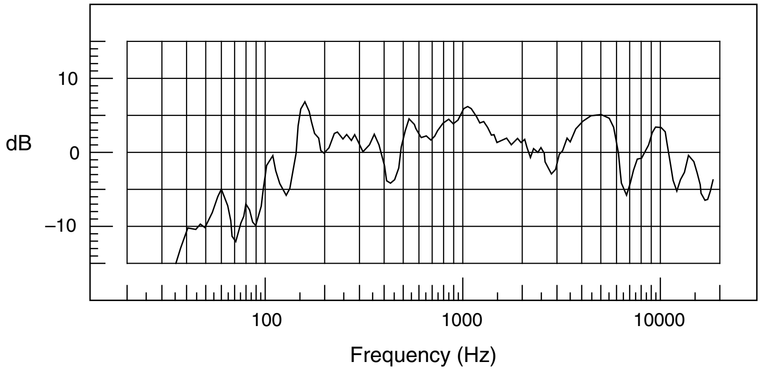


Figure 11.13 Response of a Yamaha NS10M loudspeaker on top of a mixing console meter bridge in a typical, small control room

Note the general tendency towards an overall flat response. The irregularities are desk-top reflexions, which would tend to similarly affect the responses of *any* loudspeakers so mounted

optimising the flatness of the low frequency response, and usually no measures are subsequently taken to correct the overall response. A small bookshelf loudspeaker, for example, will be acceptably flat, in wideband terms, when mounted on a meter bridge of a suitably sized mixing console, but the consequent comb-filtering due to the desk-top reflexions will reduce the sonic openness and transparency. (The effect of the reflexions can be seen in the irregularity of the response shown in Figure 11.13). These latter two aspects could be restored by positioning the loudspeakers behind the console, on pedestals, but then the low frequency reinforcement would be lost and the bass response would fall off more rapidly. Neither situation is therefore ideal. In the case of pedestal mounting, because the low frequency roll-off is a result of the lack of loading on the bass driver, it could be equalised by a suitably sonically neutral equaliser. [Note that this is *not* a room response problem, which could *not* be properly corrected by means of equalisation, but rather it is a loudspeaker air-loading problem, which *can* legitimately be equalised]. The equaliser would, in this case, correct the loudspeaker response in terms of both amplitude and phase, but unfortunately it would also reduce the headroom, so overloads would be likely at lower than normal volume levels. This is therefore not perceived to be a useful solution.

From this discussion it should be apparent that a loudspeaker mounted on the meter bridge of a mixing console, in a control room, could not be expected to perform similarly when placed on pedestals in a mastering room. So, this is one clear example of why mastering engineers may need different loudspeakers to the ones used in recording studios – they use them differently. What is more, it must be fully understood that loudspeakers which are to be mounted on pedestals tend to need a flatter *anechoic* low frequency response than loudspeakers which will be mounted *in* walls, next to walls, or on meter bridges. A loudspeaker with a fixed, passive crossover and with a flat anechoic frequency response is therefore *not* suitable for meter bridge mounting.

11.4 The great time deception

The problems discussed so far in this chapter would appear to be at the root of the frequent observation by mastering engineers that the recordings from the studios with less-experienced personnel frequently display incorrect balances between the bass instruments. If a loudspeaker exhibits a resonant low frequency response, it will add this resonance to the response of the instrumental sounds that it is reproducing. This may be of little significance to a resonant instrument like a bass guitar, but it may very distinctly alter the character of a tight, fast-decaying, percussive bass drum sound. The added resonant energy could mislead the mixing personnel into believing that the bass drum was louder in the mix than was actually the case. They would hence balance the instruments with the bass drum lower than it should be with respect to the bass guitar. When it comes to the mastering stage, little can be done to rectify the situation, because the equalisers or compressors usually cannot act on the bass drum without also acting on the bass guitar. The only solution may be to go back to the studio and do another mix with more bass drum, and incur whatever extra expenses and wasted time that may be involved.

Conversely, if a mix was done on loudspeakers with fast, low-frequency decays, albeit deficient in bass extention, the situation may not be so bad. If the instruments were balanced *between themselves*, a simple equalisation process (taking off whatever low frequency had been added due to the lack of bass on the mixing loudspeakers) could return the overall response to that which the recording staff *thought* that they had at the time of mixing. But it is the errors in the loudspeakers' time responses which are the great deceivers in so many cases, and the mixing errors which *they* lead people to make are often unable to be corrected by any currently known signal processing device.

It still seems to be the case that too many loudspeaker manufactures who make products specifically aimed at the music recording market are paying too much attention to the flattening of the pressure amplitude of the frequency response and too little attention to the shortening of the time responses. This could, at least in part, be down to the fact that the anechoic response flatness sells loudspeakers by looking good in brochures, whereas a short time response does little to sell the loudspeakers because 99% + of the users are unaware that a time response problem even exists. This is a pity, because we now have the evidence available to show how important the fast decay of a time response is to the sonic neutrality of a system[2].

11.5 Resonant tails and one-note bass

Figure 11.14 shows the frequency responses of ten different small loudspeakers, all ostensibly designed for use as small monitors in recording studios[3]. It can be seen how the responses below 200 Hz are all quite different. Numbers 1 and 2 are sealed boxes, which roll off naturally at 12 dB per octave. Number 1 is a larger box than Number 2, so it typically exhibits a flatter response to a lower frequency before the roll-off

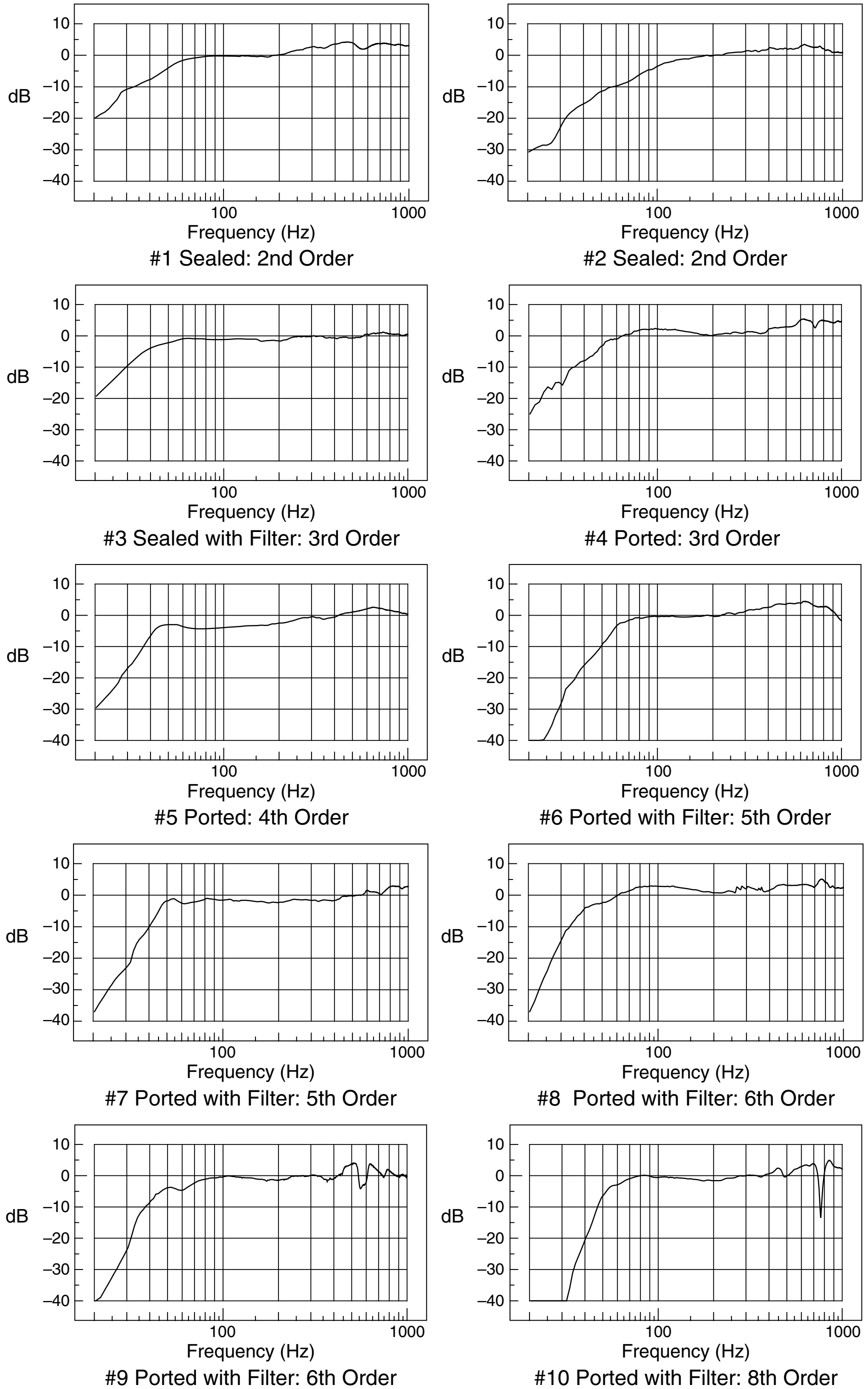


Figure 11.14 Low frequency responses of ten small, monitoring loudspeakers

begins. The eight subsequent plots all show the responses of loudspeakers which have used different acoustic or electro-acoustic means to attempt to maintain the flatness of the response to lower frequencies than would be achievable if the boxes had been sealed. Number 10 exhibits a response not unlike Number 1, albeit with a steeper roll-off, but in a rather smaller box than that of Number 1.

Figure 11.15 shows another series of plots. These show the responses of six of the loudspeakers shown in Figure 11.14, but after excitation with a burst of four cycles of 60 Hz. The ‘order’ referred to in the captions below each of the plots relates to the overall rate of low frequency roll-off: each ‘order’ represents 6 dB per octave. Whilst there is nothing in Figure 11.14 to suggest that there could be any significant problem with the means that have been used to extend the flatness of the on-axis responses,

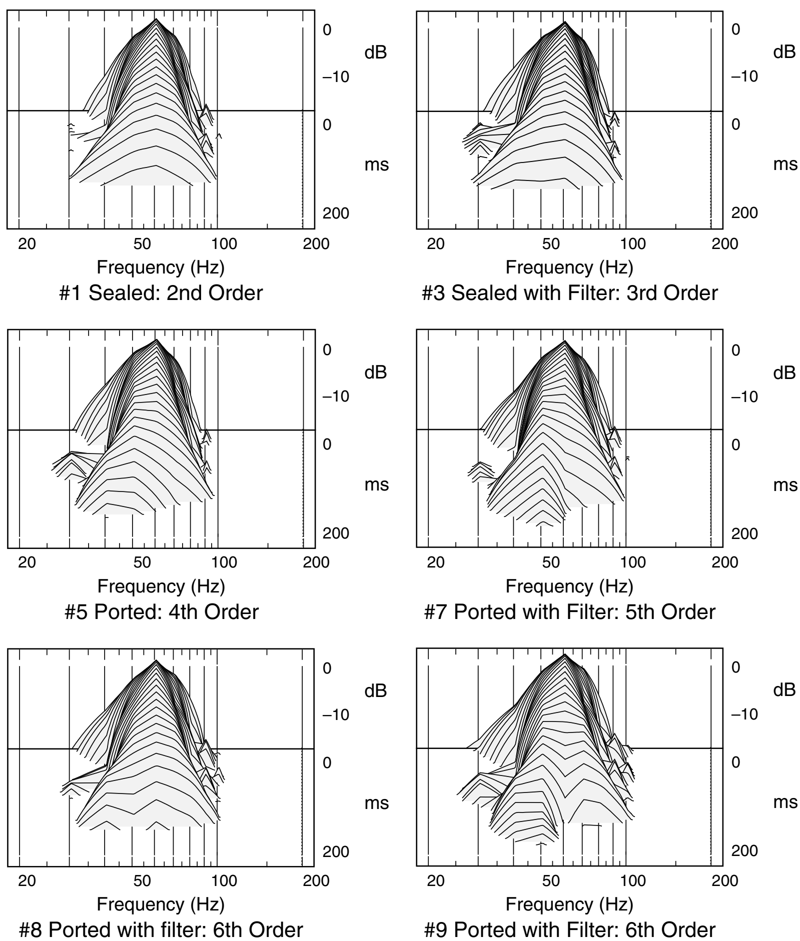


Figure 11.15 Waterfall plots of 4 cycles of a 60 Hz tone reproduced by six different loudspeakers

Figure 11.15 begins to cast doubt on their validity. Note what happens as the box resonances are used to extend the low frequency flatness, and the electrical filters are applied in order to prevent over-excursion of the cones at frequencies *below* those of the box resonances. Number 1, a simple sealed box, shows a relatively straightforward decay at 60 Hz. But look at the ported enclosures, Numbers 5 and 7. The port tunings are not high Q, so their resonances are rather broad. This means that they can be excited by other frequencies which are reasonably close to their own resonant frequencies. The four cycles of 60 Hz are sufficient to excite the nearby port resonances, and it can be seen that as the 60 Hz excitation (at the top of the plots) decays, the electro-acoustical resonances continue to ring on. In effect, the decay of the excitation note shifts in frequency. In the cases of Numbers 8 and 9, double resonance can be seen to continue, with the initial excitation frequency bifurcating into separate resonant decays. If the excitation was a musical signal, then these resonances may not relate to the musical input. They may even be very discordant.

Of course, this can also happen in ‘boomy’ rooms, where room resonances are excited at frequencies nearby the original musical input. Some concert halls or performance stages are notorious for their ‘out of tune’ decays, which can be quite noticeable and undesirable during acoustic performances. In the cases of large, often multi-purpose rooms, the problems can be difficult to treat, due to the sheer, physical scale of the treatments, but in loudspeakers the problems are surely the result of either ignorance or inappropriate engineering design. Clean-sounding bass is hardly likely to be heard from loudspeakers which are generating their own low-frequency signatures.

11.6 The masking of detail

Another aspect of loudspeaker resonances is their tendency to mask detail in the signal. Figure 11.16 shows representations of the reproduction of a modulated noise signal. This was a pseudo-random noise with a bandwidth from 35 Hz to 70 Hz, modulated by a 10 Hz sine wave. The process was not unlike that which is used to measure speech intelligibility in voice evacuation systems (Speech Transmission Indices – STI and RASTI). In these plots, the shallower the modulation depth, the more the loudspeaker is likely to mask low frequency detail – or bass articulation as it may be referred to. The sealed box, with the second order roll-off, shows a modulation depth of around 32 dB (from 50 dB, down to 18 dB). The sealed box with an electrical protection filter (Number 3) reduces this to about 26 dB. By the time we get to the ported enclosure with the second order electrical protection filter, the modulation depth is down to about 14 dBs.

As the boxes become progressively more tuned and protected, the detail in low level signals can become progressively more lost. Once again, larger, sealed boxes perform much better.

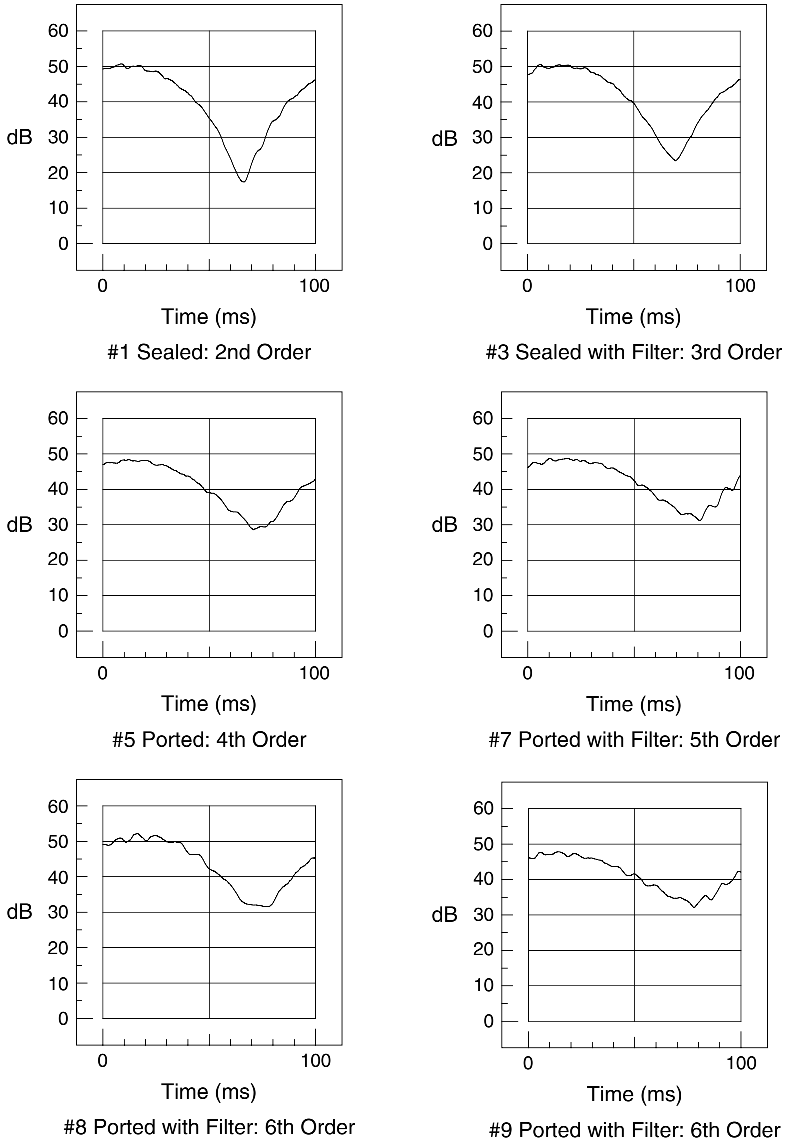


Figure 11.16 Averaged, convolved, modulated noise via six different loudspeakers

11.7 Theoretical equalisation and excess phase

Figure 11.17 shows the *waterfall* plots of the anechoic, on-axis responses of the same ten loudspeakers represented in Figure 11.14. They are all boxes of reasonably similar size. It can be seen that the low frequency resonances tend to increase in their decay time as the order of roll-off increases. Due to the general irregularity of the frequency responses, it

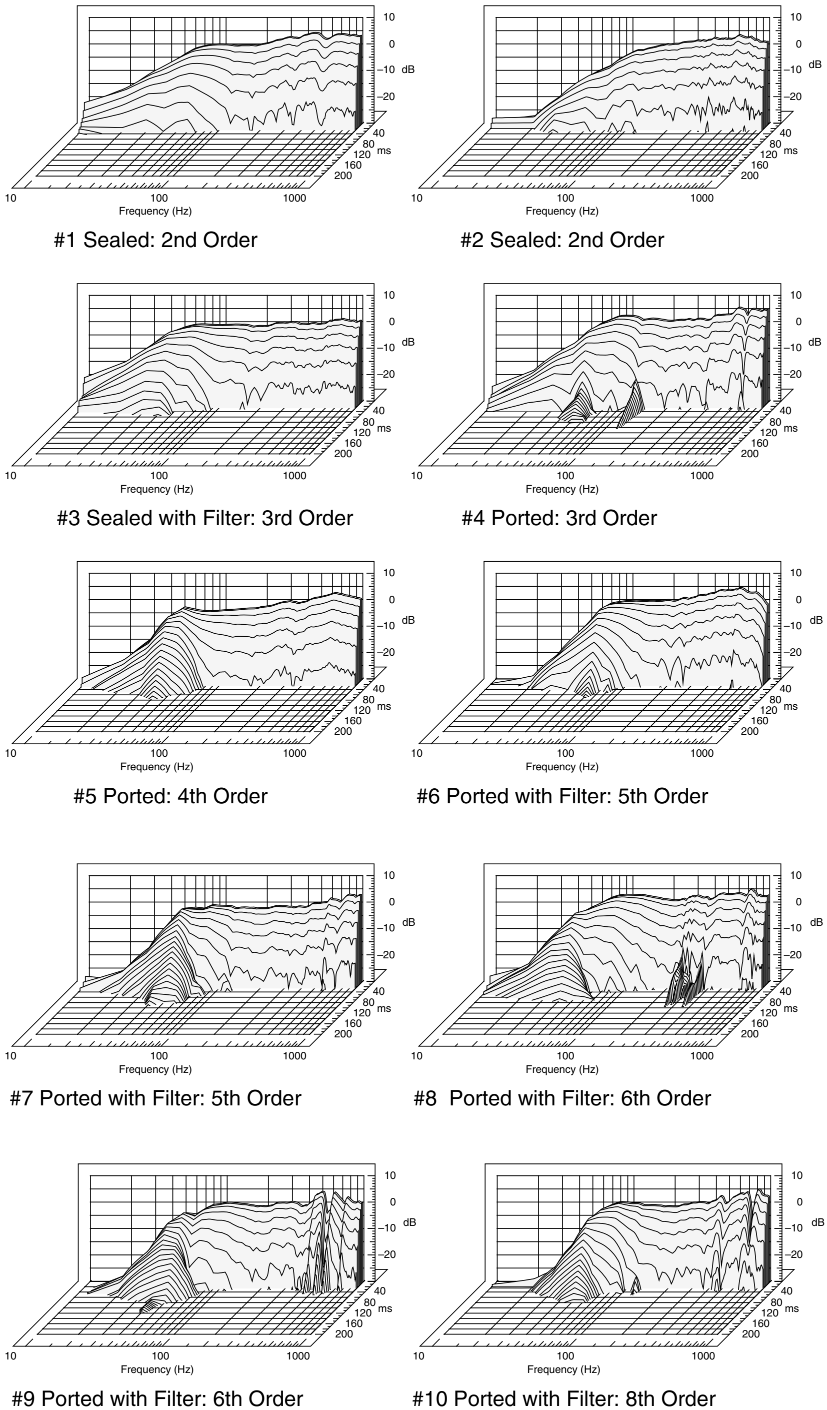


Figure 11.17 Waterfall plots of the same anechoic responses as shown in Figure 11.14

is hard to accurately compare one response with another. Nevertheless, it should be easy to appreciate that each loudspeaker would be likely to lead mixing personnel to make different level and equalisation decisions with respect to the low frequency instruments in their musical mixes. So, in terms of references, they are clearly not compatible. The big question is, to what degree would the mixing misjudgements made on each loudspeaker be correctable?

If people make different equalisation decisions at the mixing stage, then it should be possible to reverse them at the mastering stage if the decisions were global, that is, if extra bass was added to most low frequency instruments, or a little extra top was added. However, as was discussed in Section 11.4, if people have been led into equalisation or level decisions due to time response irregularities in the monitors, then those decisions are idiosyncratic to those monitors only, and may not relate well to the balance as heard in the mastering rooms. For the purposes of a research paper for a conference on acoustics, an attempt was made to 'do what the mastering engineer would do', by trying to equalise the response of the ten loudspeakers represented in Figure 11.17, and to try to see what was left in the response if they were all equalised to a common flatness[3]. Figure 11.18 shows that the results after the minimum phase portions of the frequency responses (the parts which are correctable in their phase responses as they are corrected in their amplitude responses) were 'normalised' by a digital computer. *The idea behind this being that if all of the loudspeaker responses could be equalised to be flat, then the mixes done on all of them should subsequently be equalisable back to flat.*

All of the upper traces of the ten plots shown in Figure 11.18 can be seen to be more or less equal, but the rates of decay have *not* been made uniform. Although the top line of the response of Number 3 is very similar to the top line of Number 8, the continuing resonance of loudspeaker Number 8 would undoubtedly lead to a more smeared, coloured, resonant sound in the bass response, and a probable perception of *more* bass, all due to time domain artefacts which have no solution in the frequency domain (i.e. by equalisers). No matter how flat the frequency responses of these loudspeakers can be made, the fact remains that their time response discrepancies would still lead to different mixing decisions. So, the implication, here, is that the loudspeakers which can be equalised flat in the frequency domain, without any significant hang-over in the time domain, *should* tend to lead mixing personnel into subsequently equalisable mixes. Judgemental errors made on loudspeakers which show resonant hang-over, even after frequency response flattening, would tend to need to be re-mixed. In other words, given non-perfect loudspeakers, the ones which *can* be equalised closer to 'perfection' in both the time *and* frequency domains will tend to lead to more correctable mixes at the mastering stage.

11.8 Modulation transfer-function and a new type of frequency response plot

Fashions and marketing can have a huge influence on the loudspeaker buying public. This is especially worrying when the loudspeakers will not

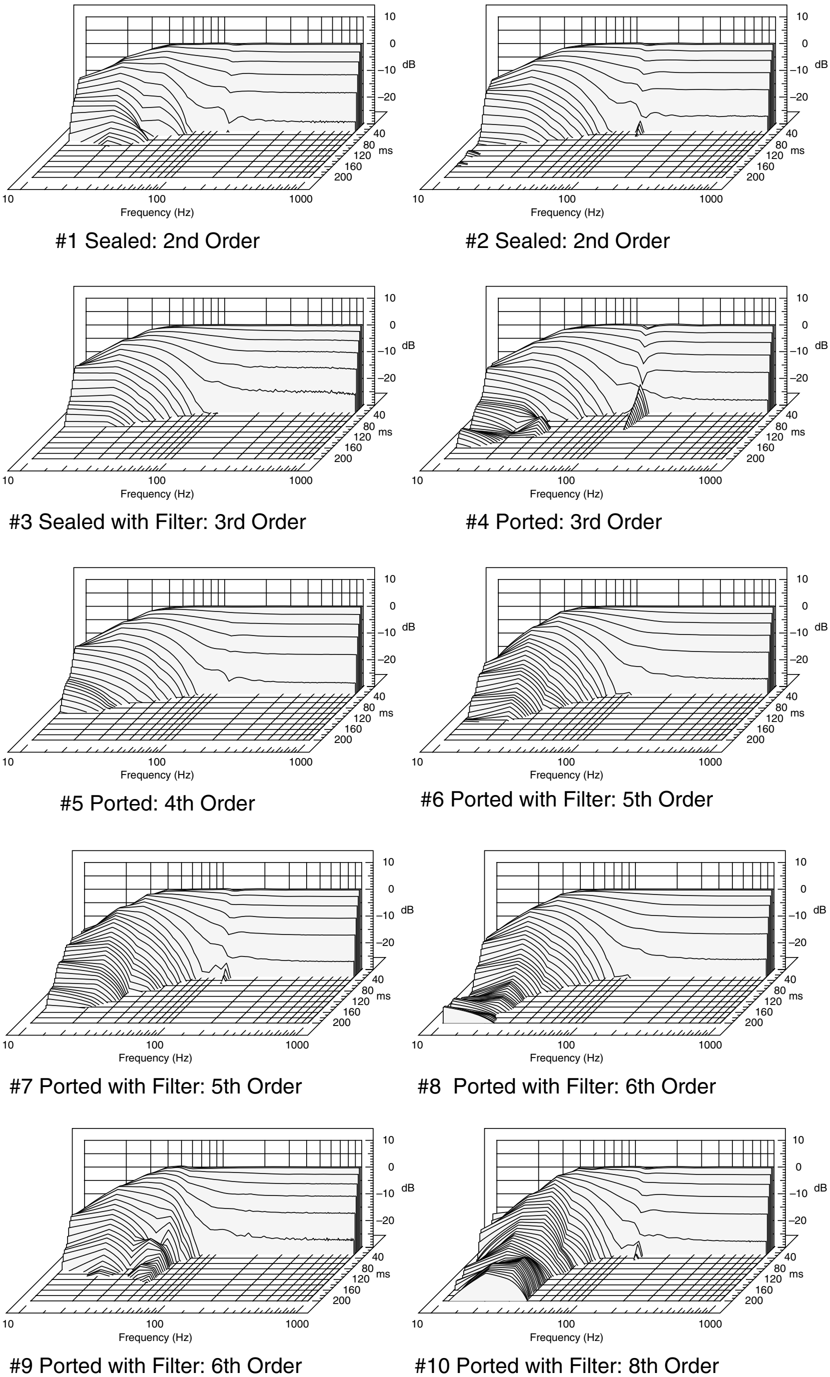


Figure 11.18 Waterfall plots of the excess-phase responses of the loudspeaker responses shown in Figure 11.17

simply be used for the enjoyment of listening to music, but rather will be used, supposedly, to try to mix to some reference standard. If we don't use some at least *reasonably* standard conditions for music mixing, then what is there for a 'high-fidelity' domestic systems to be faithful to? In the Concise Oxford Dictionary, 'fidelity' is defined as 'strict conformity to truth or fact; exact correspondence to the original'. But how can one conform in the home if the original sources of references are so variable? Taste, of course, comes in to this question, but so many of the tonal balance difference between music recordings in the shops are due to the wide variability of monitoring conditions. Something more revealing than the frequency response graphs has long been needed, and the MTF concept, introduced in Chapters 7 and 11, can be used to give much more insight into the potential *sonic* accuracy of different loudspeakers.

Traditionally we have relied on the modulus of the frequency response (commonly referred to as simply the frequency response) as the standard of reference, but it should be clear from Figure 11.1 that even two loudspeakers with similar 'top lines' on the waterfall plots can hardly be expected to sound the same if the remainder of the decaying responses are so difference. Something more was needed to define the low frequency responses, so research was undertaken to try to adopt *S*peech *T*ransmission *I*ndex (STI) techniques to look at the ability of loudspeakers to convey the full information content of the signals[4]. The STI is a means of calculating the voice modulation carrying the message through the general background noise. The idea proposed for the research was to use the concept of a typical modulation transfer function (MTF) calculation to determine how much detail could be carried on a low frequency signal without being blurred by resonances or distortions. To investigate this, the impulse responses of the loudspeakers were convolved with a modulated noise signal; a pseudo-random noise with a bandwidth of about 35 to 70 Hz, modulated by a 10 Hz sine wave. The results are shown in Figure 11.19, three of which are taken from the same group of results shown in Figure 11.16.

A sealed box exhibits a second order roll-off (12 dB/oct) below the resonant frequency of the system. Adding a first order, high-pass electrical protection filter yields a third order roll-off (18 dB/oct). A ported enclosure typically shows a fourth order (24 dB/oct) roll-off below resonance, and adding a second order protection filter results in a sixth order (36 dB/oct) roll-off. As can be seen from Figure 11.19, as the order of roll-off increases, the modulation depth reduces. In the cases shown, the depth varies from about 28 dB for the simple sealed box to only about 14 dB for the sixth order ported system. This indicates that bass detail is being lost in the ported systems, especially the ones with the added electrical protection filters. Nevertheless, although this shows *what* is happening, it still does not provide and easily describable accuracy measure. What was needed was a means by which to demonstrate not only the flatness of the response but also the accuracy with which the detail in the overall signal was being preserved.

Obviously, the modulation depth of the 20 to 30 Hz region of a loudspeaker whose response is already 40 dB down by those frequencies is not going to play much part in the subjective assessment of the sound, so a way had to be found to cross-correlate the modulation accuracy with

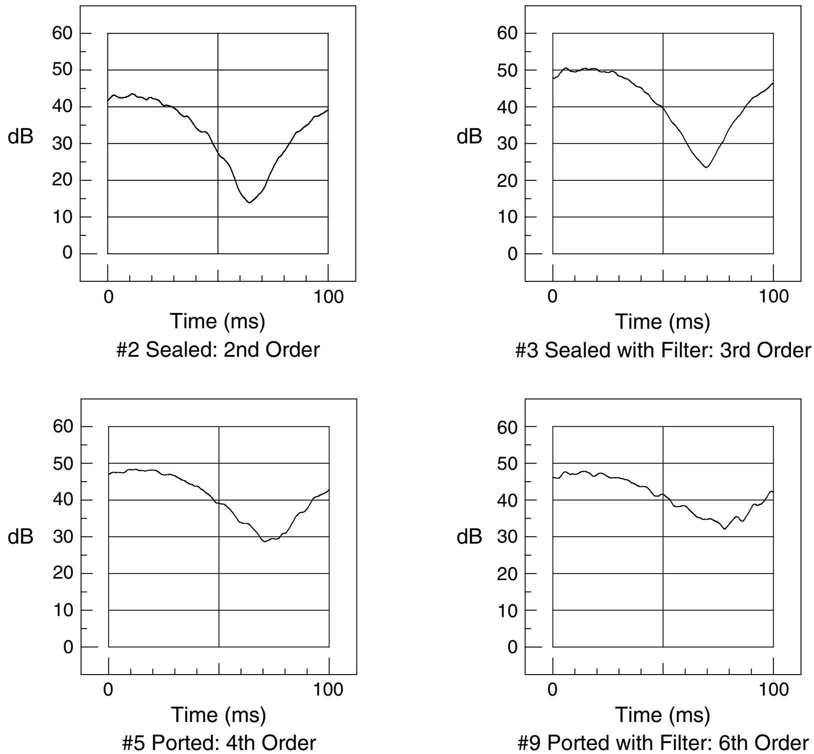


Figure 11.19 Averaged, convolved, modulated noise responses for the four example loudspeakers. It can be seen how as the order of low frequency roll-off increases, the modulation depth decreases. In other words, the *information content* in the signal decreases

the useful response of each system. A level of 85 dB SPL was chosen as a typical listening level, then the loudspeaker responses were convolved with an inverse of the minimal audible field curve (MAF) from the typical Robinson-Dadson curves for the sensitivity of human hearing – as shown in Figure 8.4 – similar to the well-known Fletcher-Munson curves. This yielded a compromise between response flatness and modulation accuracy. The plots of Figure 11.20 show the resultant MTF curves for the same four loudspeakers as in Figure 11.19. The MTF (vertical) scale is marked from 0 to 1. Zero would represent a response with no relation to the input signal, and 1 would represent perfect reproduction. The graphs are therefore plots of reproduction accuracy versus frequency. They are *not* frequency response plots; however, *from* them it can be seen which loudspeakers keep the highest MTF accuracy down to the lowest frequencies, so they are, in effect, low frequency *quality* response plots.

Whilst the ‘frequency response’ plots of Figure 11.14 would suggest that the ported enclosure with the fourth order roll-off would be a much better monitor loudspeaker than the simple sealed box shown in plot Number 2, the plots of Figure 11.20 show that the small sealed box maintains

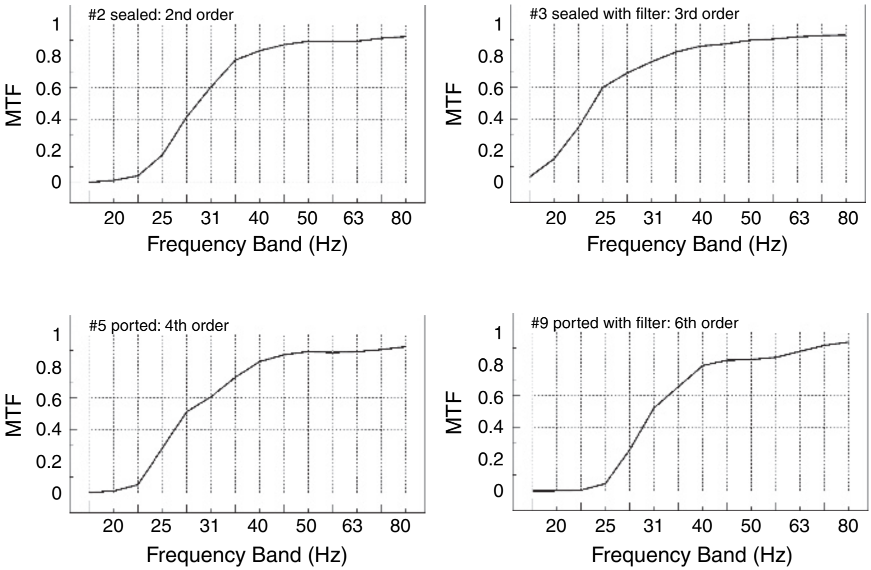


Figure 11.20 MTF results for the four example loudspeakers at a level of 85 dB SPL and incorporating compensation for minimum audible field (MAF)

a response accuracy, in terms of detail in the sound, down to a lower frequency than the fourth order ported box. In these plots, the total area below the curve is effectively a measure of response quality, whereas in Figure 11.14, the area below the curve only represents level; irrespective of whether that level is of anything connected with the music, or not, as shown by the extraneous resonant frequencies in Figure 11.15, which only serve to swamp the detail in the musical signal.

Of course, for anybody buying a loudspeaker from the advertising brochures, the response of the ported fourth order enclosure in Figure 11.14 would seem to greatly improve upon the response of the simple sealed box in the same figure. Even typical in-room response measurements on a spectrum analyser would tend to support that case, yet the results in Figure 11.20 would tell a rather different story, with the accuracy of reproduction, (looked at from a broader perspective than simply the sound pressure level) favouring the simple sealed box. The sixth order ported cabinet, which shows an almost similarly extended 'frequency response' in Figure 11.14, comes off much worse in Figure 11.20, suggesting that it produced 'artificial' bass, which would not be consistent with the concept of accurate reproduction.

This system of low frequency response assessment is the subject of continuing research. The whole point of the exercise is to try to define which loudspeakers are monitoring the musical signals which are being delivered *to* them, and which ones are 'inventing' their output. As previously stated, a 'truthful' loudspeaker, even if it is short on overall bass level, will tend to give rise to mixes which *can* be corrected at the mastering stage, or even by the tone controls on domestic hi-fi systems. Conversely, a loudspeaker which shows a poor MTF response, no matter how flat its 'frequency response' may appear, may well give rise to mixes which are the results of deceptions in the time domain, or the 'smudging' of the fine detail in the sounds being reproduced. Such loudspeakers are likely to give rise to mixes which are *not* correctable by means of equalisation in the later stages of production, and lost detail is something that simply will not be heard. Obviously, no judgements can be made about things which can not be heard, and so artistic opportunities may be lost. Loudspeakers with poor MTFs are scrambling the message, whereas a loudspeaker with a good MTF but a low-frequency roll-off is simply reducing the level of the message at those frequencies, hence the possibility of a correction by subsequent low frequency equalisation of the mix.

11.9 Summing-up

In this chapter it has been demonstrated how the fundamental principles of electro-acoustics dictate that loudspeaker cabinets and low frequency drive units need to be large if extended low frequency responses are to be achieved with good system sensitivity and low distortion. It is possible to trade size for sensitivity, low-frequency extension for transient accuracy, or sensitivity for low frequency extension, but it is not possible to simultaneously maximise the performance in terms of low frequency extension, low distortion, high system sensitivity and fast transient response except in large boxes. Furthermore, it has also been shown how fine detail in the sound may be lost when resonant systems are used to extend the low frequency responses or protect drivers from over-excursion.

Also, in general, loudspeakers with lower order low-frequency roll-offs tend to offer more precise transient accuracy than higher order designs. In terms of response accuracy it is always beneficial to extend the low frequency response as far as possible, but whilst *minimising* the ultimate slope of roll-off. If this can only be done at suitable listening levels from small cabinets by means of employing *resonant* cabinets, it means that they can never achieve the overall low frequency response accuracy of well-designed large cabinets, because only large cabinets can exhibit low order roll-offs and high SPLs at very low frequencies. Furthermore, it has been shown that mixes which have been carried out using loudspeakers with low-order roll-offs and high SPLs are more readily equalisable at a later date if the overall frequency response is deemed to be inappropriate for the music. Conversely, bass drum/bass guitar balances deemed to be incorrect after having been mixed on resonant loudspeaker systems show a tendency towards being less correctable due to simultaneous time and frequency domain errors.

Well-designed larger loudspeaker boxes will win on almost all aspects of reproduction, which is why the more reputable mastering engineers tend to avoid relying on small loudspeakers when an accurate, overall assessment of a recording is required.

Always remember the incontrovertible truth: size *is* important!

References

1 Newell, P. R., Holland, K. R., Newell, J. P., 'The Yamaha NS10M: Twenty Years a Reference Monitor. Why?' Proceedings of the Institute of Acoustics, Vol 23, Part 8, pp 29–40; Reproduced Sound 17, Stratford- upon-Avon, UK (2001)
2 Newell, P. R., Holland, K. R., Mapp, P., 'The Perception of the Reception of a Deception', Proceedings of the Institute of Acoustics, Vol 24, Part 8, Reproduced Sound 18, Stratford- upon-Avon, UK (2002)
3 Holland, K. R., Newell, P. R., Mapp, P., 'Steady State and Transient Loudspeaker Frequency Responses'. Proceedings of the Institute of Acoustics, Vol 25, Part 8, Reproduced Sound 19, Oxford, UK (2003)
4 Holland, K. R., Newell, P. R., Mapp, P., 'Modulation Transfer Functions - a Measure of Loudspeaker Performance'. Proceedings of the Institute of Acoustics, Vol 26, Part 8, pp 107–115; Reproduced Sound 20, Oxford UK, (2004)
Copies of the above papers can be obtained from the Institute of Acoustics, (UK), Tel. +44 1727 848195, Fax, +44 1727 850553 or email *ioa@ioa.org.uk*
The full set of response plots can also be obtained from Reflexion Arts, S.L., Tel. +34 986 481155, Fax +34 986 413412, or by e-mail, *ventas@reflexion-arts.com*

Chapter 12

The challenges of surround sound

The choice of loudspeakers for use with surround-sound systems is never easy. One problem is that the concept of surround means so many different things to so many different people. Except for professional cinema applications, where Dolby and THX, along with a few other companies, have laid down strict guidelines, the rest of the world of surround exists in near chaos. Even in professional, industry magazines, recording engineers and producers, who ought to know better, can be seen promoting the idea that there *are* no rules, so everybody should do it in their own way. Well; if there are no rules in the studios, then where are the references for the domestic reproduction systems to comply with?

The majority of this confusion stems from the fact that modern, music-only surround sound was never conceived and controlled purely by the music business, but has developed as an adjunct to other surround technologies which have developed for their own purposes, and which were never primarily intended to be capable of audiophile quality music reproduction. Exactly how best to deal with the subject of the recording and reproduction of music in surround is something which is still very controversial; even in professional circles.

In domestic applications, we have to accept that in the majority of cases *one* loudspeaker arrangement will be used for home cinema *and* the reproduction of music-only SACDs and DVDs in surround. This is a very unsatisfactory situation, but it would be totally unrealistic to expect people to have ten or more loudspeakers in their lounges. Firstly, therefore, it may be beneficial to look at the reasons for the undesirability of the above circumstances by looking at a range of professional surround mixing circumstances. This will also establish a more global understanding before embarking on a more detailed discussion of the requirements of each individual loudspeaker.

12.1 Surround sound in professional studios

Surround sound, in the guise of quadrophonics, first came to prominence in the early 1970s. The concept of the use of close-field monitors had not yet become widespread, so the first quadrophonic control rooms tended to consist of the two front halves of the then normal stereo control rooms placed face to face. A 1970s-style quadrophonic room is shown in Figure 12.1. There was some pressure at the time to spell the word quadr*a*phonic, with an ‘a’, which was said to be more etymologically correct, but the compatibility

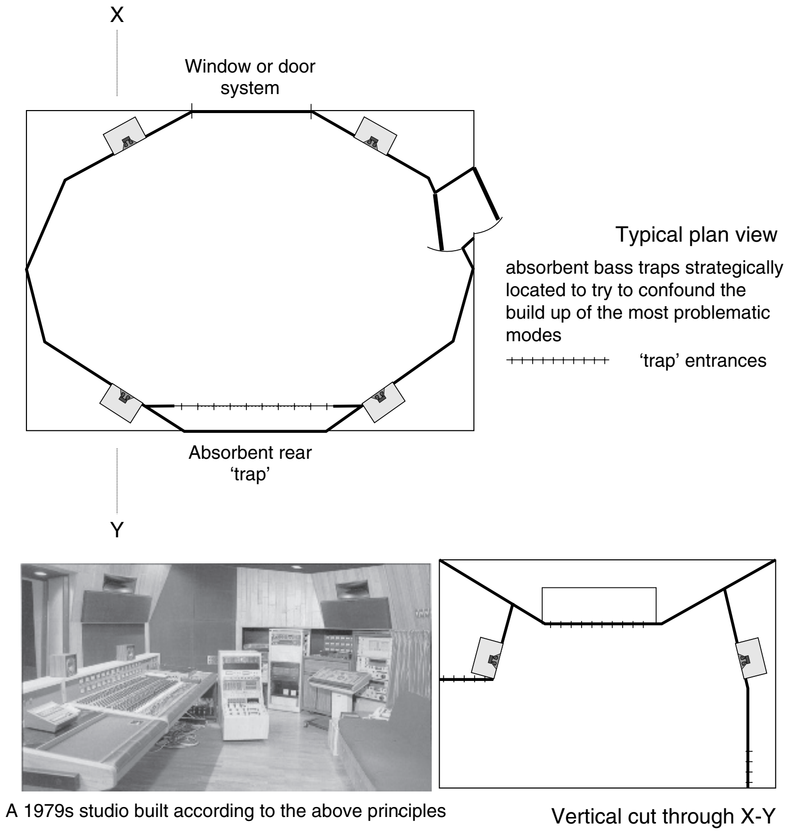


Figure 12.1 A quadrophonic control room from the 1970s
In effect, the room consists of the two front walls of a stereo room placed face to face. Note that the front loudspeakers are flush-mounted, above the windows, but the rear loudspeakers are mounted above the soffits of the machine alcoves – another source of asymmetry

with stere*o*phonic and mon*o*phonic led to a tendency to use the 'o' spelling. However, when The Who released their famous film Quadr*o*phenia, that seemed to kill off the 'a' spelling. It was in such uncertain conditions and the battles of the format wars that quadrophonics sprung into life.

In those days, quite frankly, the sound in most cinemas was bad. Dolby, to their great credit, began to specify that cinemas which used the Dolby processed soundtracks should comply to certain standards. Dolby understood the limitations of the phase-matrixed quadrophonic music systems, (which were not convincing the record-buying public to purchase many quadrophonic music systems) but they saw the possibilities of stabilising the centre-front image in cinema presentations by using one of the four quadrophonic channels in the centre-front position, effectively making a three-channel frontal stereo sound-stage. The remaining channel they fed to multiple rear and side loudspeakers, all supplied with the same signal, to give the effect of a distributed source of ambiental sound. This was the Dolby Stereo system. The quadrophonic and Dolby Stereo layouts are

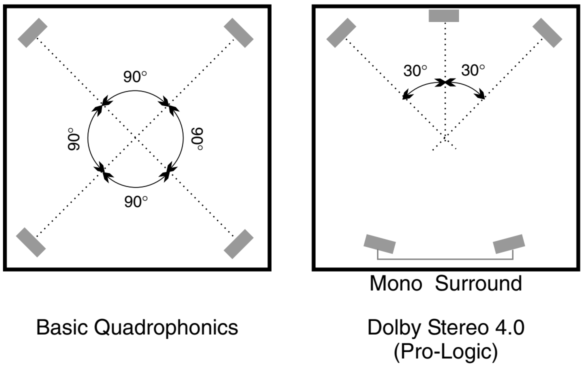


Figure 12.2 Comparison of quadrophonic and Dolby Stereo formats, both using four channel matrix-encoded signals

shown in Figure 12.2. Both systems used phase-matrixed analogue recordings, which exhibited very poor separation between channels, and 'logic' enhancement systems were usually employed to improve the inter-channel discrimination. (Some vinyl disc systems used high frequency carriers, for the rear channels, instead of phase matrixing, but they were fraught with problems.) The quadrophonic music systems of the 1970s used four identical loudspeakers, but the Dolby Stereo system required large loudspeakers of controlled directivity behind the screen, and multiple, small, restricted bandwidth, wide-dispersion loudspeakers for the surround channel, to maximise the audience coverage and the sense of spaciousness. Therefore, at this very first bifurcation, the loudspeaker requirements for four channel surround had already changed.

Around this time, still in the mid 1970s, Tomlinson Holman proposed, for *domestic playback*, the use of dipole (figure of 8) loudspeakers for the surround channels[1]. Even at this very early stage, many people were beginning to realise that the 360 degree distribution of primary instruments was not suitable for a great deal of music, and that the rear channels were often better employed for ambience. The use of dipoles in this position, as shown in Figure 12.3, with their nulls facing the listeners, ensured that

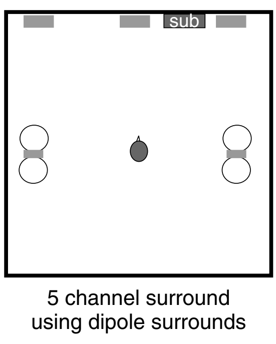


Figure 12.3 Dipole surround loudspeakers

only the sound that reflected from the walls would be heard, thus creating a well diffused ambient surround field.

Figure 12.4 shows a reasonably similar idea, proposed by Fosgate[2], but using bipolar loudspeakers instead of the dipoles proposed by Holman. Again, the loudspeakers are generally facing away from the listener, and in the drawing each side of the surround uses 3 drivers – a pair in the side cabinets and one in the rear cabinet. There was also in the proposal a means of creating four channels from the rear pair, in order to further diffuse the ambience. In the above configuration there are no true nulls, as there are in Figure 12.3, but only the low frequencies will radiate directly to the listener. The middle and high frequencies will tend to be received via reflexions only.

Another attempt at providing a diffuse side/rear surround field is shown in Figure 12.5. This method has been used in small film dubbing and video post-production studios[3]. In this case the conventional box loudspeakers are mounted so as to face mathematically derived diffuser panels which are fixed to the walls. Once again the low frequencies will radiate omnidirectionally, but the middle and high frequencies can only arrive at the listening position via the diffusers. A variation on this theme, using naturally diffuse sources, is shown in Figure 12.6. This system uses distributed mode loudspeakers as single surround sources on each side wall [4]. Figures 12.1 to 12.6 clearly show how professional and domestic systems have both developed around concepts which have used some very different types of

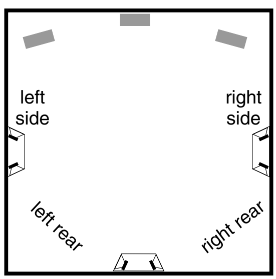


Figure 12.4 Bipolar surround, proposed by Fosgate, using processors to decode seven channels processed from five

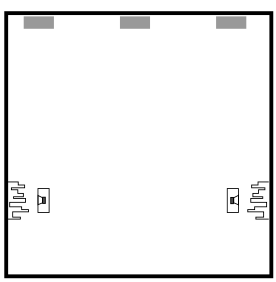


Figure 12.5 A five-channel format using wall-mounted diffusers to disperse the sound from the conventional rear loudspeakers. (After David Bell)

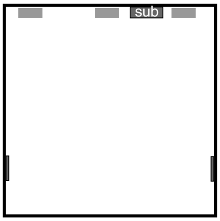


Figure 12.6 Diffuse surround sources – NXT DML panels

loudspeakers to try to create the desired illusion. Unfortunately, with so many concepts in use – and we have barely begun discussing them – it should already be apparent that compatibility problems are to be expected between the mixing and domestic playback environments.

In cinemas, the environments are usually well controlled, in order to conform to standards which ensure reasonable commonality with the dubbing theatres – the studios in which the film soundtracks are mixed. Although this book is not intended to be about cinema and video systems, it is necessary to look at them here so that their influence on the development of domestic entertainment systems can be better understood. For many people, since the beginning of the 21st century, the playback of music recordings in the home has only been via their home theatre systems, which are primarily geared to the reproductions of film and video soundtracks, but it must be clearly established that there is no single type of loudspeaker which can serve for all purposes. The monopoles, dipoles and bipoles represented in Figures 12.2, 12.3 and 12.4 are fundamentally different radiating sources, as are the diffuse sources shown in Figures 12.5 and 12.6, and the T.V./video style systems shown in Figure 12.7. There is no possibility of these diverse sources producing identical sounds. Consequently, if music is not mixed on systems of the same general concept as the reproduction systems, then the reproduction cannot be anything other than compromised. The resulting sound may be pleasing, and may even be desirable, but it will not be 'high-fidelity' in the sense that a good domestic stereo system, well positioned in an appropriate room, will be expected to radiate a sound which is very close in its balance and timbre to that which

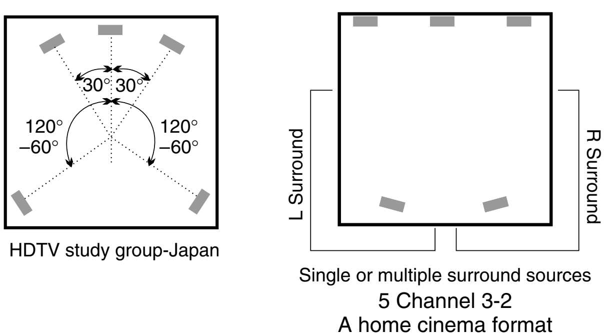


Figure 12.7 Incompatible domestic formats

was being heard by the recording personnel at the time of mixing. With surround sound, the variables are so many that the degree of fidelity which we have come to expect in stereo is rarely achievable.

12.2 Cinema sound

Feature films, in general, are mixed specially for the big screen, where there is an important psychological correlation between picture size and the appropriate sound pressure level. Figure 12.8 shows a behind-the-screen installation of a mixing room (dubbing theatre) equipment for either 5.1, 6.1 (Dolby EX, with left, centre and right surround channels) or 7.1, the Sony Dynamic Digital Sound (SDDS) system. The surround loudspeakers are 12 JBL 8340A cabinets, specially designed for use as cinema surround loudspeakers. A smaller 5.1 room, using JBL behind-screen loudspeakers is shown in Figure 12.9. Rooms such as these are used for the final mix, usually

Figure 12.8 A cinema studio monitor wall for Sony SDDS 7.1 surround

Figure 12.9 Auditel 'Audi A', in Paris, France, using a JBL cinema system

Figure 12.10 The 'Kyoto' control room at Eurosonic, Madrid, Spain, designed by Sam Toyashima. The room is used for mixing stems for film soundtracks. The surround loudspeakers are pedestal mounted at the sides and rear of the room. The large, right-front monitor is the black rectangle at the left of the photograph. The console-mounted loudspeakers are for stereo television mixing

from 'stems' (or pre-mixes) of music, dialogue and effects. In many cases, however, the music is mixed in smaller rooms without screens. Such a room is shown in Figure 12.10. This room has a large, Quested frontal monitor system in a room designed by Sam Toyashima. The surround loudspeakers are pairs of smaller Genelec monitors, on pedestals, in an L-shaped, side and rear configuration. The monitor system is essentially flat, with a small high frequency roll-off in typical music recording studio style.

Once the music stems go to the dubbing theatres to be mixed with the dialogue and effects, the surround/front balance will be readjusted to the cinema environment. What is more, the overall sound will be re-equalised, because after the perforated projection screen has been placed in front of the loudspeakers, the response at the mixing positions is adjusted to the 'X-curve', as shown in Figure 12.11. Mixing will normally take place with peak levels reaching 105 dBC or more at the listening position. A picture of a life-sized tank, firing a shell with a 'bang' at 85 dBC would simply not sound credible, so large sounds go with large screens. Bearing in mind that when played in the home on a small screen, even 85 dBC may be considered to be too loud for the neighbours, a quick glance at the equal loudness curves shown in Figure 8.4 will reveal that the perception at 80 dB and 100 dB is quite different. Film protocol demands that the mix be done at the same level as the cinema presentations, so that the perception of relative frequency balance can be accurately assessed, unlike the music mixing process, where no such control exists. For home cinema release, the soundtrack may have to be re-equalised, not only because of the lower probable level of reproduction, but also because many home systems do not reproduce with an X-curve equalisation. (For more on the X-curve see the Glossary).

A typical domestic playback system is shown diagrammatically in Figure 12.12, conforming to the ITU775 recommendation. When a film is played back on such a system, using loudspeakers of typical domestic

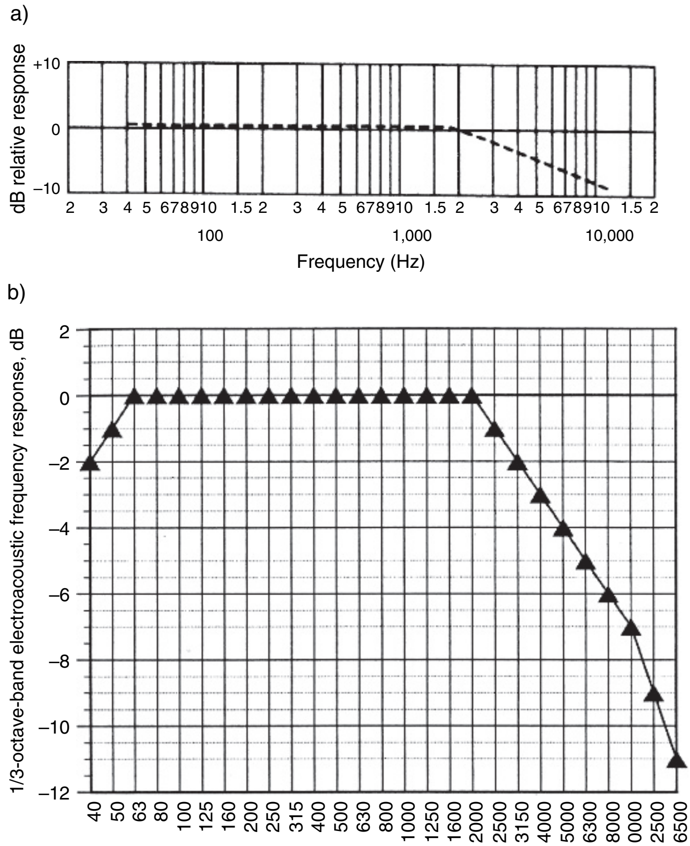


Figure 12.11 The X-curve. a) ISO Bulletin 2969 – recommended response curve for motion picture loudspeaker systems. b) The X-curve for cinema monitoring, to be measured spacially averaged in the far-field of the sound system with quasi-steady state pink noise and low-diffraction (small) measuring microphones. The room volume must be at least 200 m^3 (6000 cubic feet). The curve is additionally adjusted for various room volumes, as per SMPTE 202

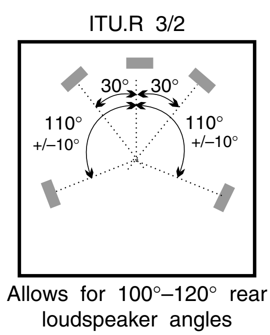


Figure 12.12 ITU.R 775, 3/2

nature, the overall perception is bound to change because the type and distribution of the loudspeakers are different to those on which most films are mixed. Very little source material is therefore specifically mixed for reproduction on home theatre systems. In audio-visual presentations, the changes in the perception of the soundtracks tends to be compensated for by the strong domination of the eyes over the ears. The small screen also tends to require less low frequencies in the sound. However, when we listen to music-only mixes, there are no such distractions, and the lack of fidelity can be conspicuous.

12.3 Music mixing

Unlike in the cinema world, no standard mixing conditions have ever been enforced for music mixing, and recommendations have been ignored in the majority of instances or, at best, have only been respected in certain aspects. There has been a free-for-all which has continued unchecked largely because the domestic surround sound systems have been predominantly cinema/video orientated, and music mixes have been accepted in a compromised form by people who were unwilling to accept a separate, music-only system at home. The lack of any clear domestic trends of high fidelity music orientated surround systems has not encouraged recording studio owners to invest in specialised surround-sound mixing facilities, so the studios, themselves, are therefore tending not to give any clear leads. In many cases there have been few clear-cut decisions about whether music should be mixed four-channel, five-channel, or five-channel plus a discrete low frequency channel. (5.1 as it was dubbed by Tom Holman.) The argument for four channels has been put forward by many people who have felt that the centre-front channel is compromised in many 5.1 domestic systems, because of the conflict between the loudspeaker and the video screen vying for the same spot. Considering the fact that the centre-front in music mixes is the location for many lead vocals and primary instruments, the risk of compromising it has led many music mixers to continue to make phantom centres from the front left and right channels to avoid the risk of the mix being ruined by playback through systems with inadequate centre channels. In most cases, the front left and front right loudspeaker will, indeed, be the best loudspeakers in any set up.

Another body of opinion has been avoiding mixing vocals in the centre channel because of the equalisation changes which are noticed if the recording is 'folded down' to four or two channels due to the asymmetrical or symmetrical reception by the ears of a phantom centre or discrete centre image respectively. The problem is shown in Figure 12.13, from which it can be seen how the signals arriving from 30 degrees either side of centre front suffer a path length difference of about 10 cm between the nearest and furthest ear from each loudspeaker. As discussed in Chapter 5, arrival delays between two drivers will give rise to a cancellation at the ear around the frequencies which arrive in the region of 180 degrees out of phase. Conversely, the sound from *one* loudspeaker will suffer the same cancellation effects if the *ears* are displaced so that the arrival of the wavefront is not simultaneous at the two ears. Consequently, a phantom centre

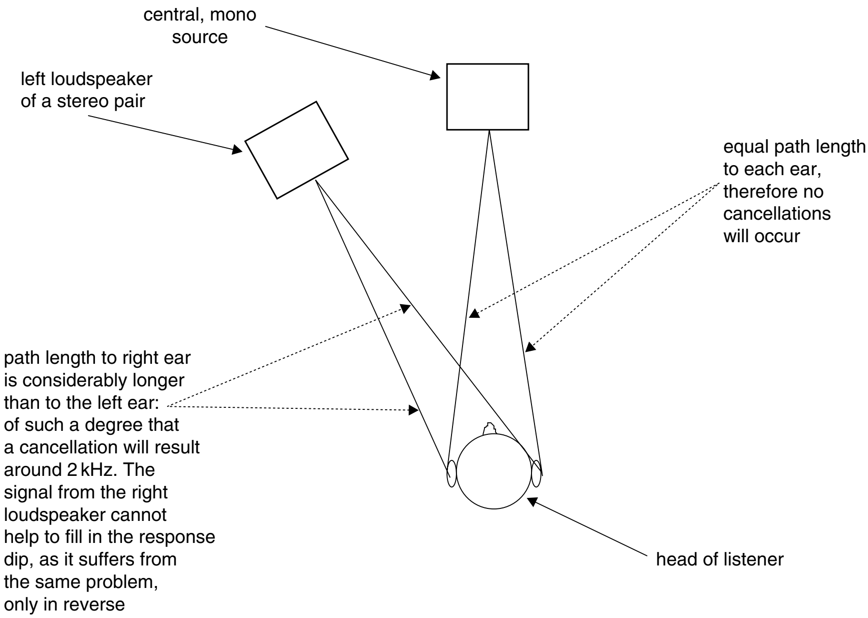


Figure 12.13 Path length anomalies for phantom central images

image will suffer from some cancellation in the region of around 2 kHz. When making decisions about equalisation of the phantom centre-front instruments or voices, this dip is automatically taken into account. When listening to the same sounds through a discrete centre-front channel, no such arrival delay dip occurs, so, for any given sound, the equalisation decisions would be different when mixing with a discrete or a phantom centre. This has led some people to complain that the sounds which they are accustomed to recording and monitoring via a stereo pair of loudspeakers sound too present when monitored through a discrete, centre-front loudspeaker. They therefore opt to use phantom centre-front images in order to try to maintain better equalisation compatibility when the surround mixes are heard in normal, two-channel stereo.

The other big dilemma is often whether to mix to 5 or 5.1 channels – that is, with or without the sub-woofer as a dedicated channel. In reality, the sub-woofer channel has little to do with music mixing. The .1 was a cinema invention to provide extra headroom for low frequency effects, such as during explosions. In music, there is really no equivalent function, so five-channel mixing tends to be the norm. When listening to a 'pseudo 5.1' system which is really a five-channel system with a bass management system feeding a common sub-woofer, there will be perceptual differences as compared to listening on five full-range loudspeakers, due both to the mono low-bass and to the arrival time anomalies that will be experienced in all but the very centre of the array. The perception will also be likely to change significantly with changes in the frequency below which the content of the five channels was filtered and sent to the sub-woofer. There is no standard frequency of cross-over into sub-woofer(s).

12.4 Sub-woofers – discrete and managed

To get the best out of any sub-woofer, the cabinet needs to be large. The sub-woofers shown in Figure 12.8 are real sub-woofers, designed to outperform even the 600 litre cabinets of the five main loudspeaker systems, which easily reach down to 20 Hz themselves. With ten 15 inch loudspeakers in the five 600 litre cabinets, and with reflex ports tuned to 20 Hz, the low-frequency output of this main system, alone, is quite prodigious. Each of the sub-woofers visible in the photograph consists of a McCauley 6174, eighteen inch driver (460 mm), with a free-air resonance of 20 Hz and a sensitivity of 94 dB for one watt (2.83V) at one metre. The frequency range is quoted by the manufacturers as 15 Hz to 800 Hz and the rated power capacity is 800 watts. The boxes are also of 600 litre capacity, with one driver in each box, and are reflex tuned to 16 Hz. They are constructed from a sandwich of 2 layers of 19 mm chipboard with a 5 kg/m^2 plasticised bituminous deadsheet in-between, and lined with a sandwich of 20 mm felt/3.5 kg deadsheet/20 mm felt. The boxes are also internally braced in all three axes, and are surrounded by 10 cm of concrete. The 40 mm-long voice coil allows a cone travel in any one direction of 15 mm under a constant force factor (Bl), and a 25 mm unidirectional travel with reduced Bl before mechanical limitations arise. The total cone excursion can therefore reach 50 mm (2 inches) peak to peak. Each drive unit weighs over 16 kg, and the ferrite magnet alone, of 13000 gauss, weighs almost 12 kg, and has a Bl factor of 15.3 tesla.

This system is being explained in detail because it is a professional sub-woofer, intended to deliver high quality sound at cinema levels, day after day, year after year. The response decay is rapid and uniform in the frequency bandwidth of use. The loudspeakers have no characteristic 'boom' to their sound. Their positioning at the junction of the front wall and the floor effectively mounts them in quarter space (or π space) as explained in Chapter 7. This effectively raises the sensitivity by 6 dB compared to the quoted 94 dB in free space. The use of two drivers, side by side, increases the loading on the cones still further, and the mutual coupling so achieved adds another 3 dB to the low frequency sensitivity, yielding 103 dB for 1 watt at 1 metre for the pair of cabinets. Structural losses are kept low by making both the floor and monitor wall out of thick concrete. The listening distance in this room is about 8 metres, so the 'double distance rule' (stating that the SPL falls by 6 dB every time the distance from the source is doubled, [see Chapter 7]) would mean that 103 dB at 1 metre would fall by 18 dB at 8 metres, to 85 dB. Each multiplication of the power by 10 adds 10 dB to the SPL, so 1000 watts ($10 \times 10 \times 10$) would deliver 30 dB more than the 85 dB for 1 watt. The maximum peak short-term SPL to be expected on rare occasions in cinema would not exceed 115 dBC, so the 1600 watts rms rating of the pair of loudspeakers (or 3200 watts for 10 milliseconds) gives the system adequate damage tolerance. At all normal cinema levels, the loudspeakers are working well within their low distortion range, and, as high level peaks are always short-lived, thermal compression is not a factor. The amplifiers driving the sub-woofers are Class AG, made by Neva Audio in St Petersburg, Russia, and can deliver 2000 watts into 8 ohms. The above description outlines the lengths that one needs to go to outperform the low frequency response of a good, medium to large sized studio monitor system.

Purely in terms of high fidelity, when mixing music in surround on a monitor system using five, full-range loudspeakers, it is best to use their own low frequency capability rather than a separate sub-woofer, because the five discrete sources can give better spacial perception, a seamless response because no low frequency crossover is involved, a tighter and better spacially distributed transient response, and an overall low frequency sound quality that outperforms the majority of commercial sub-woofers. In fact, in many studios which are used to mix music stems for cinema, and where some low frequencies may be mixed to a separate sub-bass (Low Frequency Effects [.1]) channel, the low frequency feed is often distributed across the flush-mounted front loudspeakers, instead of feeding a discrete sub-woofer of commercial design, precisely because the LF response of the main system is likely to be better than that of a commercial sub-woofer. However, as will be discussed later, there are other concerns which may lead to other decisions about the best option for mixing formats.

There can also be some confusion caused by the different meanings of LFE as have been applied to sub-woofer use. In the Dolby Digital (Dolby Surround) language, LFE stands for the *L*ow *F*requency *Effects* channel that is the discrete (.1) channel of low bandwidth, used in addition to the 5 main channels. Conversely on many music monitoring systems, the LFE is the Low Frequency *Extension* provided by the sub-woofer, whose feed is derived from a bass management system which re-directs to the sub-woofer the low frequency content of the five principal channels, below the frequency where the five main loudspeakers begin to roll-off in their frequency response. The two concepts are very different. In the former case, the five main channels are full range, and the LFE channel may at times operate up to 500 or 1000 Hz, although normally the upper limit is more likely to be around 120 Hz. In the latter case, the sub-woofer should ideally only operate below 80 Hz, or the low frequencies will be localised by the ear to the position of the sub-woofer. A further concept of Low Frequency *Enhancement* also exists, in the form of an optionally reproducible low frequency effects channel for digital television, but whose content is not essential to the overall programme. We therefore have LFE standing variously for Low Frequency *E*ffects *E*xtension, or *E*nhancement. There is little or no commonality between the three concepts, and some implementations are optional. So much for standardisation.

Dolby now recommend a sub-woofer distribution for cinema use as shown in Figure 12.14, with one woofer placed 33% of the distance from one side wall, and the other woofer placed 20% of the distance from the other side wall. This gives rise to a more or less central image, without either risking the localisation of the effects channel to one side of the cinema or symmetrically driving the room modes from a central location. However, in home use, one single sub-woofer is normally used whose frequency response is limited by a crossover, the frequency of which is determined by the low frequency response of the five main loudspeakers. Somewhat absurdly, this can be as high as 400 Hz when used with mini-satellite loudspeakers. In such cases, localisation of many low frequency sounds to the 'sub'-woofer in inevitable. For discerning listeners,

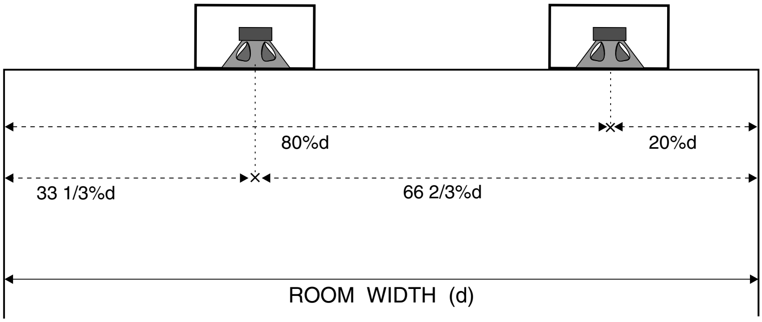


Figure 12.14 Sub-woofer siting for cinema installations

Dolby recommendations for siting two sub-woofers; one of them one fifth of the room width from one side wall, the other one third of the room width from the other side wall. This not only avoids both the localisation of the sound to a single sub-woofer placed off-centre, but also avoids the symmetrical driving of the strongest room modes by the central placement of the sub-woofer(s)

the crossover to a sub-woofer must be kept below 80 Hz, and lower if possible, or spacial information will be lost and localisation of sound to the sub-woofer will occur. [However, see Section 12.6.]

Obviously, when we listen to five discrete channels we localise sounds at their sources, but a system with a bass management system will re-distribute the low frequencies, sometimes with odd consequences. For example, a synthesizer located in the left rear loudspeaker may sound strange if its output *above* 200 Hz comes from the left rear and the output *below* 200 Hz comes from the front-located sub-woofer. Again, this may in many cases sound acceptable, but it is not high fidelity inasmuch as this was probably not what the music producer intended. However, some mixes are actually carried out with bass management systems because some mixing personnel feel that this is more typical of how the majority of people will listen at home. This is a market-led approach, rather than a high fidelity approach, but there will be justification for this idea in Section 12.6, on grounds of sheer practicability.

12.5 Size versus performance compromises

The low frequency decay of a good studio loudspeaker system is shown in Figure 12.15. The response extends quite low, and the decay is fast and uniform with frequency. In general, such a response is only possible with relatively large boxes, but compactness has been a very desirable characteristic of loudspeakers for surround systems, because the domestic practicality or social desirability of placing five or six large cabinets in a dwelling space is not good. Even in professional studios, the lack of universally accepted surround formats for music mixing has made owners reluctant to invest heavily in surround control rooms. For this reason, many stereo rooms have been adapted for surround use by the installation of free-standing systems. Again, for reasons of practicality in rooms already containing much other equipment, many studios have opted to use rather small loudspeakers.

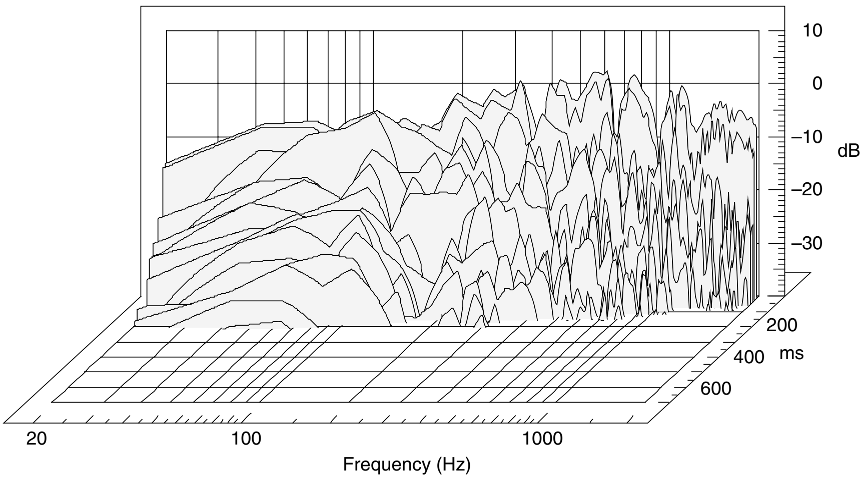


Figure 12.15 Responds decay of a full-range loudspeaker in an acoustically well-controlled room

A single sub-woofer must handle the low frequencies from five channels, and so needs to have the same acoustic output capability as that of all five principal loudspeakers connected in phase and receiving the same signal. By its very nature, a sub-woofer is required to have an extended low frequency response, but to achieve this in a small box with the sensitivity necessary in order not to burn up with the sheer electrical input power needed to handle the combined output of five channels is, as we saw in Chapter 11, a great conflict of requirements. The sensitivity problem has been addressed by some manufacturers by the use of bandpass cabinets and horn loaded cabinets, but each only addresses part of the overall response shortfall.

The bandpass concept was described in Chapter 3. These systems tend to exhibit multiple resonances, as it is the resonant effects that help to augment the sensitivity and flatten the pressure amplitude response over the desired band. The response of one such cabinet is shown in Figure 12.16. The waterfall plot shows a flat response for just over an octave, but the decay shows a degree of resonance that would not be conducive to tight, high-impact bass sounds. The step function response in the same figure shows just how slow the attack is. The response actually builds up with time, rather than immediately starting to decay. Neither the attack nor the decay could really be deemed to be concordant with the concept of high fidelity.

Figure 12.17 shows the response of a horn-loaded cabinet intended for use as a sub-woofer. Typical of the high acoustic loading presented to the diaphragm by the horn, the time response is characterised by a fast attack, as shown in the step function plot, and a fast decay which is clearly observable in the waterfall plot. Unfortunately, the bass extension is rather poor because the horn mouth needs to be proportionate to the wavelength in order to couple the horn to the room, as explained in Chapter 4. Once again, when size is kept small, we have a trade off between bass extension and a clean transient response. Therefore, a small, high-fidelity, high sensitivity sub-woofer is a contradiction in terms.

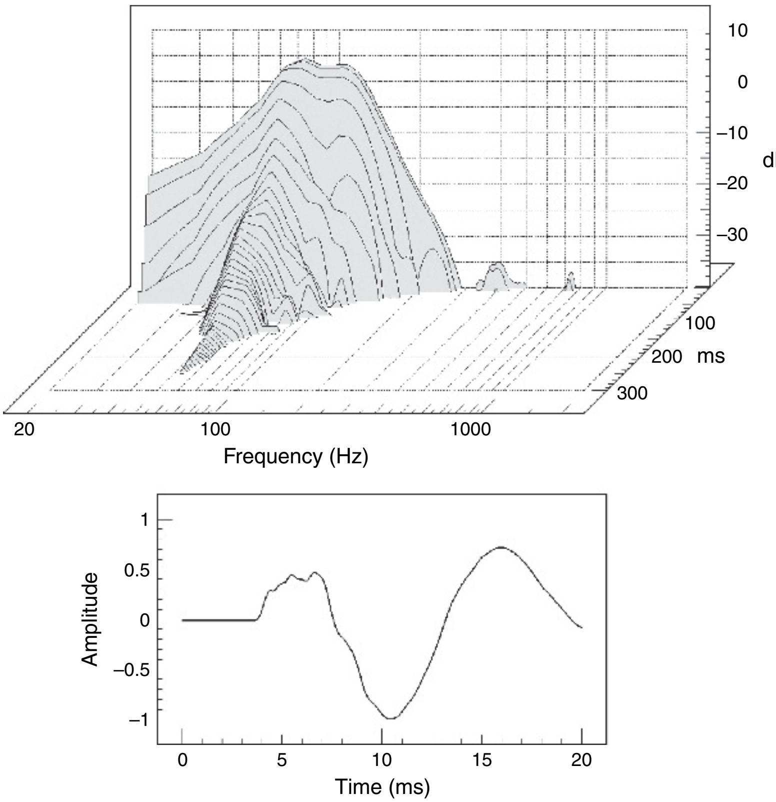
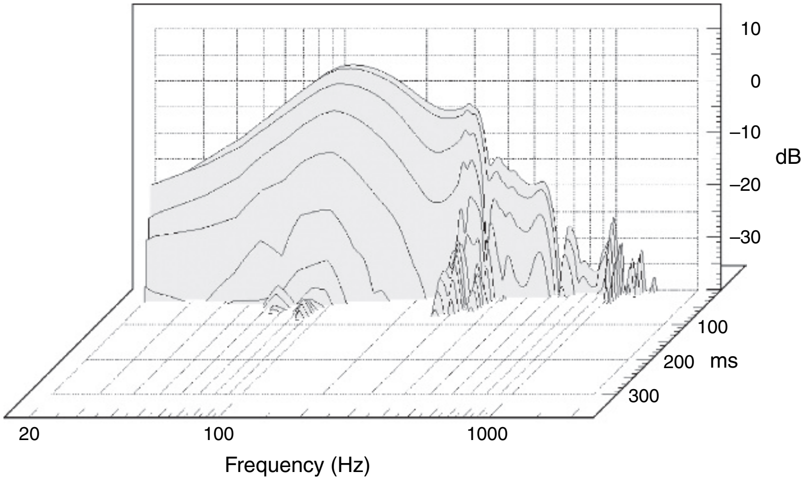


Figure 12.16 Waterfall plot and step-function plot of a typical bandpass sub-woofer in an anechoic chamber

'Woofer', perhaps, but sub-woofer only if a low fidelity transient response can be tolerated. A horn-loaded loudspeaker which is a *true* sub-woofer is shown in Figure 12.18, but small it is not. Nevertheless, in some cases, the corner placement of a horn-loaded sub-woofer may go some way to extending the low-frequency output, but it will exite *all* the possible room modes, and hence may provoke severe room resonances. This would, of course, defeat the object of using the loudspeaker with the better transient response.

The need for small size in order to be marketable has led to an emphasis on the multichannel spacial sensations, and relatively little has been openly discussed about the ways in which many loudspeaker performances have dropped vis-à-vis older loudspeakers intended for stereo. There is a tendency for surround to compromise absolute fidelity for the extra sensation of space, but the excitement of the spaciousness can be short lived once the absolute fidelity loss is recognised. Loudspeaker manufacturers are obviously going to make what they can sell, and if super hi-fi surround systems appeal to only a very small market, then the bigger market will be primarily catered for. There has been a general recognition that the



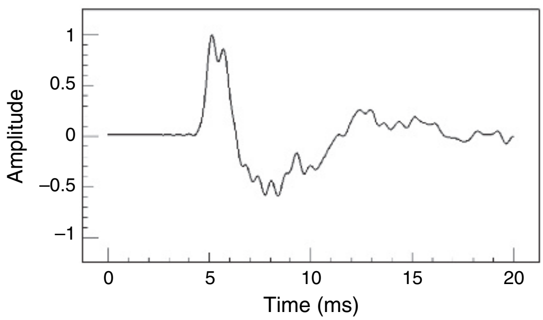


Figure 12.17 Waterfall plot and step-function plot of a horn-loaded sub-woofer in an anechoic chamber

Figure 12.18 A concrete horn, under the stage, loads a cabinet with four 18 inch (450 mm) low frequency drivers. The response is essentially flat down to 20 Hz, and 120 dBC on the dance-floor is achieved with ease and reliability. The mouth size is 4 m × 1 m

vast majority of domestic users will use 'satellite and sub-woofer' systems, and much effort has gone into finding means by which the low frequency cabinets can be positioned and controlled in order to get a real improvement in the LF response, even beyond that which could be expected from good stereo loudspeakers in the same rooms. This concept will be further explored in the next section.

12.6 Compound sub-woofers and electronic control

Whilst no electronic signal processing system can be expected to outperform the excellent acoustic treatment of rooms, it is entirely justifiable to recognise that good rooms are expensive and rare. For this reason, some loudspeaker manufacturers have seen a real need for a means of deriving a decent low frequency performance in acoustically poor rooms, and the use of separate (sub-) woofers can be a real advantage in such circumstances as exist in many domestic listening rooms. Toole gave a good overview of the situation in his 2003 paper, 'Art and Science in the Control Room'[5]. He made the observation that below about 100 Hz, room resonances can be controlled electro-acoustically, although response *dips* cannot be equalised. Between about 100 Hz and 300 Hz, mounting geometry, adjacent boundaries, and strong solitary reflexions can cause problems that defy simple electronic cures, and most probably need to be dealt with acoustically. Above 300 Hz, the loudspeaker dominates, so there is no substitute for using good, well-mounted drivers here.

Below 100 Hz, in small rooms, such as many sound control rooms and domestic listening rooms, the number of resonant modes will be few, and mostly well separated. This gives rise to great spacial variation and amplitude response variation. A listener seated at an antinode of a resonance will hear a greatly augmented response, whereas at a node the response would be diminished, perhaps almost to the point of cancellation. The perceived frequency response of a loudspeaker driving the room from a boundary would therefore be dependent upon the listening/receiving position. Conversely, a loudspeaker placed at a node of a resonance would be incapable of driving the mode at that frequency, whereas when placed at an antinode it would strongly drive the resonance. Hence, any given loudspeaker in any given room being measured or listened to at any given point in the room would exhibit a low-frequency response which was entirely position dependent, and likely to vary wildly from place to place unless the room was heavily damped, i.e. acoustically dead.

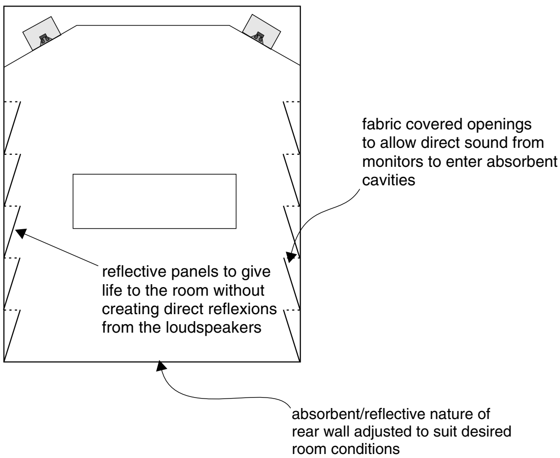
Back in the days of mono recording it was customary to find a loudspeaker position which drove a room in the most uniform manner achievable, and then find a listening position with the flattest overall response. Stereo complicated matters, because the relationship between the driving and the listening positions was fixed by the ideal equilateral triangle which needed to be maintained between the two loudspeakers and the listener(s). Finding *three* relatively flat positions was much more difficult than finding just two, especially when the three needed to maintain the equilateral triangle geometry. The extra degree of difficulty led to the development of highly specialised design concepts for stereo control and listening rooms, some examples of which are shown in Figure 12.19. All of the rooms shown are asymmetrical from front to back: the end that generates the sound is in all cases acoustically different from the end which receives the sound. This works fine for stereo, but for multichannel surround sound, where the rear channels may carry significant programme information, and not just ambience, this approach will not suffice. Dr Toole argues that in anything but very highly damped rooms with enormous bass absorption systems, the ability to obtain a flat low-frequency response from five full-range loudspeakers is beyond us. In fact, the authors of this book came to a similar conclusion in a 2004 conference paper, when discussing the concept of how to make a symmetrical room for surround sound with individual loudspeaker responses as good as can be achieved in a stereo room, especially with flush-mounted monitors[6].

Toole is convinced that the five full-range loudspeaker option, when used below 80 to 100 Hz, can only lead to appalling bass responses in anything other than anechoic chambers, especially when one considers all the possibilities of the combinations of loudspeakers which may be radiating simultaneously whilst supporting some of the many possible phantom sound stages. He maintains that the bass summation is not only a cost-saving measure, but is a superior way to reproduce bass in non-ideal rooms.

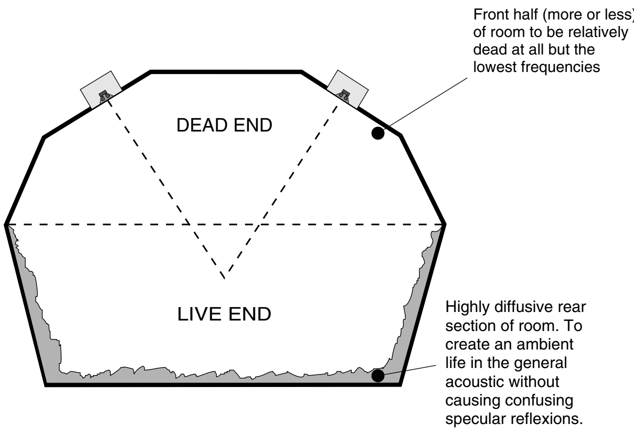
Given this concept of mono bass below 80 Hz, a technique for achieving a flat, in-room response already exists, with the use of a combination of four distributed sub-woofers whose responses are modified by signal processing and parametric equalisation, both in the form of individual equalisation and global equalisation. This can bring an overall loudspeaker/room response to within about 3 dB of flat over a wide listening area. Moreover, the fact that room *modal* responses, as opposed to reflexions, are of a minimum phase nature, this amplitude correction will also tend to correct the phase, and hence the time response. Toole has been most insistent that this is the only practical solution for domestic surround sound listening, and, as such, it only makes sense to monitor the same way in the studio control rooms, because almost nobody will hear the mixes reproduced in a totally discrete way. On the other hand, Holman[7] has argued that stereo sub-woofers can be essential in order to create a greater sense of spaciousness, and advocates the positioning of a sub-woofer half way down each side of the room. The authors of this book agree with both of them, hence the title of their aforementioned 2004 paper – Surround Sound – The Chaos Continues[6]. Spaciousness and flatness cannot both be maximised in the same system; except, of course, in an anechoic chamber!

Shortly before the publication of this book, Floyd Toole published a paper in the Journal of the Audio Engineering Society in which he expanded further on the concepts of low frequency responses for surround systems[8]. Dr. Toole has said that in his opinion, if room modes can be actively suppressed by processed multiple woofers, and a relatively flat, uniform, low frequency sound field can be achieved in less than perfect rooms, then achieving this in mono below 80 Hz may well be significantly more desirable than the extra spacial effects of stereo low frequencies

a)



b)



c)

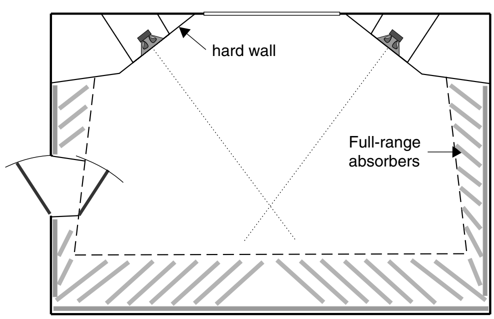


Figure 12.19 a) Typical room in the style of Wolfgang Jensen. b) Live-end, Dead-end control room. c) Non-environment control room. d) BBC-style, controlled reflexion room. e) Ishii/Mizutoni listening room. f) An IEC listening room

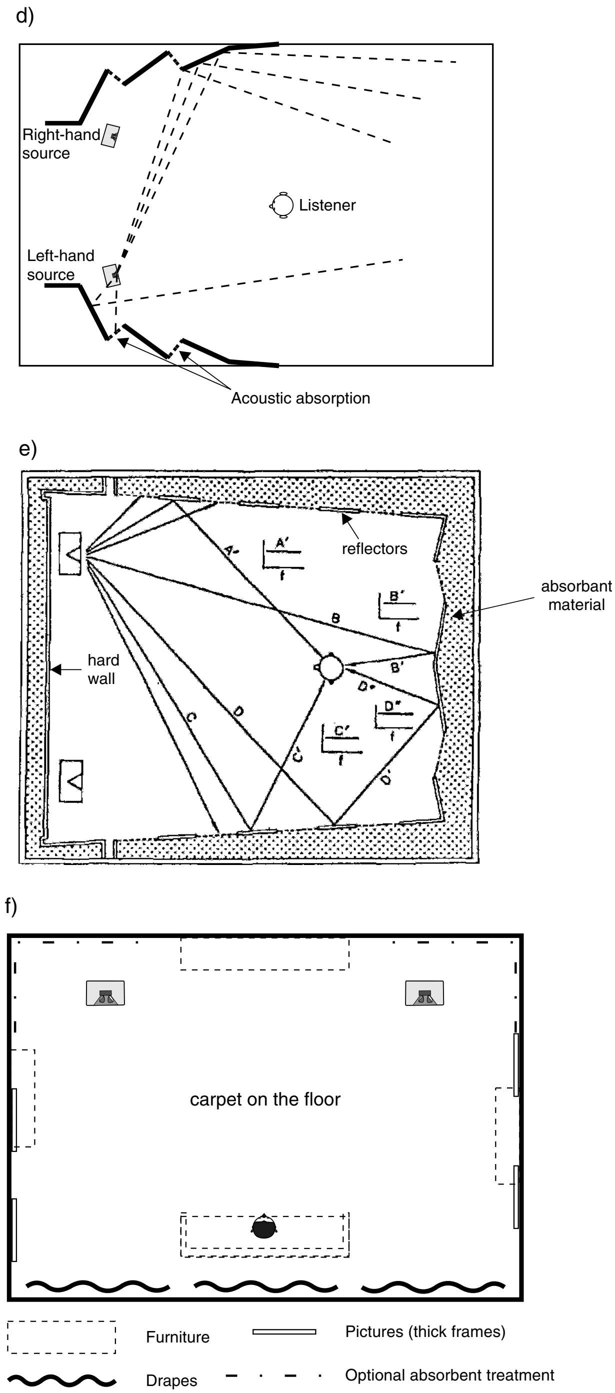


Figure 12.19 Continued

which must suffer modal irregularities. There is a lot of reason in what he says.

12.7 System considerations

What must be understood is that surround sound is not simply stereo plus another dimension. Surround sound is something totally different, and needs to be assessed as such. As we have already seen from Figure 12.19, the stereo rooms will not adapt well to discrete surround sound, although they can work quite well if the surround channels are restricted to ambience. This reappraisal of rooms for surround use, with careful juxtapositioning of diffusive and absorptive surfaces, evenly distributed around the room, also requires a reappraisal of loudspeaker performances. In many professional stereo rooms, the axial response of a loudspeaker tends to be paramount, because off-axis irregularities can, in many cases (although not all applications) be controlled by absorption.

However, in surround rooms, loudspeakers often tend to need to exhibit very smooth off-axis responses, because of the wider distribution of diffusively reflective surfaces in the room. They also need a wider directivity pattern in the horizontal plane because the listening positions tend to be broader than stereo listening areas. Figure 12.20 a) shows how a producer and engineer may experience essentially the same overall balance in a stereo listening room, but in surround, they may have to be side-by-side, otherwise the producer would remain too close to the rear loudspeakers. This is another argument for the use of the centre-front channel, because it can anchor the stereo centre image even for off-centre listeners, thus allowing two people to hear essentially the same mix when sat side-by-side. (However, there is also a good argument for larger mixing rooms, where small positional changes will less affect the perception!)

In Figure 12.20 b) the rear, P_2 position is not valid because the person sitting there would hear too much of the rear channels in proportion to the front channels. The left side, P_1 position would work reasonably well if the front sound stage was anchored with the central images coming from the centre loudspeaker. However, if a phantom central image were to be used, the person at position P_1, would predominantly only hear it coming from the left loudspeaker, due to the precedence effect and the earlier arrival time of the signal from the left loudspeaker. (Unfortunately, this means that the decision whether to use discrete centre or phantom centre may be dictated by something as arbitrary as the size of the mixing room.)

The precedence effect, or the 'law of the first wavefront' means that for the arrival of two similar signals, varying in arrival time between about 1 millisecond and 40 milliseconds, the perceived direction of the source will be from the direction of the first sound to arrive, unless the level of the later arriving sound is 6 or 8 dBs higher. In reality, for a centrally panned image, the later arriving sound will also be *lower* in level, because it will be arriving from further away, so the directional pull to the earlier sound will be reinforced. Similar sounds arriving within less than about 700 microseconds (0.7 ms) will not be so affected, and sounds arriving more than about 60 ms apart will be heard as two separate sounds – a sound

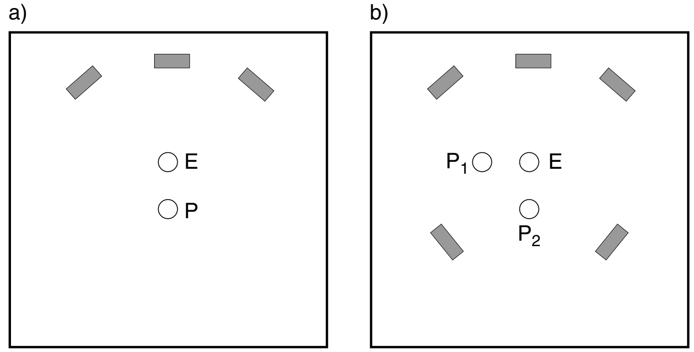


Figure 12.20 a) In a stereo room, given a suitably absorbent rear wall, the engineer and producer at positions E and P would hear essentially the same mix. The only significant difference would be a slightly reduced level at position P. b) In a surround room there are no positions where two people can hear the same balance. Relative to the engineer's position (E) position P_1 would experience a left-heavy mix, and P_2 a rear-heaving mix

Only large rooms can resolve this problem, but, ironically, the current trend is towards smaller rooms

and its echo. There are 'grey areas' between 0.7 and 1 millisecond, and 40 and 60 milliseconds, which can depend on the nature of the signal and individual perceptions. For this reason, Dolby Digital Surround systems in cinemas have a delay, adjusted to circumstances, on the signals fed to the surround channels so that if any sound exists in all loudspeakers, then even for people at the back of the cinema, the behind-the-screen loudspeaker signals will always arrive first. This keeps the action located on the screen. However, the music industry does not apply such delays because the rear channels may be used for discrete instruments, and not just ambience.

What is more, in some cases, the cinema surround loudspeaker arrays are also tapered in level, becoming lower towards the back of the theatre. There will always be a level reduction from the frontal loudspeakers as a listener moves towards the rear, so a corresponding level reduction in the surround loudspeakers can help to prevent the surround sounds from dominating the more important screen sounds. A 4 dB taper from the frontmost to the rearmost surround loudspeakers is about the maximum which can be used before the surround distribution would itself become compromised. In the configuration shown in Figure 12.20 b), with only single loudspeakers on each of the rear channels, no such tapering would be possible, so the front/rear balance would depend upon the listening position. Room reflexions can help to ameliorate the problem, but smooth off-axis responses are essential if the reflexions are to be returned with an evenly distributed energy balance, or substantial colouration of the sound would be experienced.

For the above reasons it can be very difficult in small rooms to find positions where two or more people can enjoy similar listening experiences from surround sound systems, especially when the use of the centre channel has not been fully exploited for reasons of lack of faith in the end-users' choice of adequate centre-front loudspeakers. In larger rooms, which allow greater distances between the loudspeakers, the relative time and level differences are generally less, so a larger area can usually be achieved where a more common listening experience can be realised. However, as always, when the distance from the loudspeaker to the listener is increased, the room acoustics come more into play. If the room acoustics cannot be changed, then loudspeakers may be chosen with more suitable directivity patterns, but this would, of course, also change the relationship of the direct to ambient sound. There are no easy solutions to problems of 'universal' surround sound.

Loudspeakers that are good for stereo are therefore by no means necessarily good for multi-channel systems. The loudspeaker directivities and total power responses need to be considered in relation to room acoustics and the size and orientation of the designated optimal listening areas. The question of whether five surround channels will be used for the free distribution of instruments or as a frontal stereo sound-stage plus ambient rear-channels is another factor which will greatly influence loudspeaker choice. Quite clearly, a system using monopole front loudspeakers with dipole rears whose nulls are facing the listeners will not be capable of giving equal reproduction perception to a guitar, for example, panned from one type of loudspeaker to the other. Conversely, a symmetrical system which *can* do this may be unable to exhibit the same enveloping ambience as the dipole. When considering true high fidelity in surround sound, the choice and the positioning of the loudspeaker is greatly more complicated than the corresponding choices in stereo.

Unfortunately, it is also the case that very few studios are as well prepared for surround as they are for stereo, and the people doing the surround mixes are generally much less sure of what to do, which is not surprising when surround can mean so many things to so many different people. Some, who perhaps should know better, even applaud the free-for-all, no rules approach in the name of artistic freedom, but, as was mentioned in the opening paragraph of this chapter, if there is no reference standard at the time of mixing, then how are the domestic listeners to know what loudspeaker arrangement is most appropriate to reproduce the intended sound of any given mix?

The range of loudspeakers which are currently used in the control rooms where music surround mixing takes place is wildly variable, as are the control room acoustics. There is no degree of standardisation even vaguely approaching that of film dubbing theatres, and no companies such as Dolby or THX to 'enforce' the standards. As previously explained, rooms which are optimised for stereo are not optimal for surround, and loudspeakers which are optimal for use in good surround use may not be optimal for stereo. Mixes done for cinema are not optimally reproducible on domestic surround systems which are optimised for music, and vice versa. Music which has been mixed on five, discrete, full-range loudspeakers will not be optimally reproduced on 'home theatre', bass managed systems. High fidelity music-only surround and home cinema systems are very different. This is both unfortunate and disappointing, because very few people will be prepared to purchase, or live with, two completely separate systems. However, for people who are happy to listen to all their music via MP3 coding perhaps this discussion about surround loudspeaker optimisation is superfluous, yet unless there is a public demand for quality, the industry will stagnate, and the whole artistic process of music production will be demotivated such that creativity will surely suffer. The lack of general awareness about the whole question of loudspeaker suitability for surround sound is, in itself, only adding to the confusion.

References

1 Holman, Tom., 'New Factors in Sound for Cinema and Television', Journal of the Audio Engineering Society, Vol 30, Nos 7/8, pp 529–539 (July/August 1991)

2 Fosgate, James; matrix engineer and designer of the Dolby ProLogic II system

3 Chase, Jason., 'Hi-Fi or Surround, Part Two', Audio Media, European edition, Issue 92, pp 122–6 (July 1998)

4 Newell, P.R., Holland, K.R., and Castro, S.V., 'An Experimental Screening Room for Dolby 5.1', Proceedings of the Institute of Acoustics, Reproduced Sound 15 conference, Vol 21 Part 8, pp 157–66 (1999)

5 Toole, Floyd E., 'Art and Science in the Control Room', Proceedings of the Institute of Acoustics, Vol 25, Part 8, Reproduced Sound 19 conference, Oxford, UK (2003)

6 Newell, Philip R., Holland, Keith R., 'Surround Sound – The Chaos Continues', Proceedings of the Institute of Acoustics, Vol 26, Part 8, pp 135–147, Reproduced Sound 20 conference, Oxford, UK (2004)

7 Holman, Tomlinson., '5.1 Surround Sound – Up and Running', Focal Press, Oxford, UK (2000)

8 Toole, Floyd., 'Loudspeakers and Rooms for Sound Reproduction – A Scientific Review', Journal of the Audio Engineering Society, Vol 64, No 6, pp 461–476 (June 2006)

Glossary of terms

This glossary has been rolled forwards from previous work by the authors, with some new additions specifically relating to the text of this book. Experience has shown that although some of the terms may be familiar to many readers, they have often been misunderstood or misused. This glossary attempts to clarify the definitions of the terms, at least as used in British English.

Acausal filter
A filter applied, usually digitally, which has advanced knowledge of the signal arriving, via delaying the main signal on which it is intended to act – effect before cause through prior knowledge of the cause. See **Causal filter**.

Active systems
Filters or loudspeaker systems, for example, where an external power source needs to be applied as well as the drive signal. A filter based on semiconductors or valves, driven from an external battery or mains supply, and placed ahead of a power amplifier in a loudspeaker system is an example of an active filter. See **Passive systems**.

Anechoic chamber
A room that is designed to simulate free-field acoustic conditions by means of the placement of highly absorbent materials on all surfaces. The absorbent materials are usually in the form of wedges pointing into the room. This arrangement ensures that sound waves arriving at the room edges are maximally absorbed for all angles of incidence. The lower frequency limit of anechoic performance is set by the length of the wedges, which are effective down to a frequency where the wedge length is equal to one-quarter wavelength. However, what little reflexions still exist will return more weakly from the distant walls of the larger anechoic chambers than from the nearer walls of a smaller chamber using similar wedges.

Variously known as anechoic rooms or free-field rooms.

See also **Semi-anechoic chamber** and **Hemi-anechoic chamber**.

A to D
Analogue to **D**igital converter. A device using a highly stable internal clock which samples the audio voltage waveform at a rate higher than at least twice the highest frequency of interest, and puts out a digital binary signal that represents the voltage level of each sample. For example, to sample a maximum frequency of 8 kHz, the sampling rate would need to be in excess of 16,000 samples per second. See also **D to A**.

Audio frequency range
The range of frequencies over which the human ear is sensitive is usually considered to be from 20 Hz to 20 kHz. A number of commonly used frequency ranges are listed below. The span of frequencies quoted for each range should not be treated as exact; they are included as an approximate guide only.

Name	Frequency range
Infrasonic	0–20 Hz
Very low	15–50 Hz
Low	20–250 Hz
Lower mid	200–1000 Hz
Mid	250 Hz–5 kHz
Upper mid	2–6 kHz
High	5–20 kHz
Very high	15–25 kHz
Ultrasonic	20 kHz–10^{13} Hz

Back e.m.f.'s
The electromotive forces (voltages) which are generated in a loudspeaker system by the mechano-magnetic interactions. They superimpose on the drive signal (forward e.m.f.) but are usually largely damped by the energy sinking (absorbing) action of the low output impedance of an amplifier, which thus provides a high damping factor. Excessive impedance in loudspeaker cables, for example, can reduce this damping effect, and hence in such systems the back e.m.f.'s would play a greater part in the overall response of the system.

Bookshelf loudspeaker
A genre of domestic orientated loudspeakers whose low frequency responses are aligned to try to achieve a flat far-field response when the loading provided by a wall is taken into account, as would typically be the case when a loudspeaker was mounted on a bookshelf.

Causal filter
A filter in which the effect takes place after the cause. See **Acausal filter**.

Cellular
Composed of cells, which can be either open or closed. In closed cell foams, for example, each cell acts like a small balloon. When compressed or distorted the air trapped inside cannot escape, and so the cell acts as a good spring but a poor sound absorber. Open cell foams are generally poorer springs but better acoustic absorbers. (For equal densities.)

Close field
The region close to a sound source, such as a loudspeaker, in which the sound field is largely that due to the source, and is little affected by the room reflexions, resonances, or reverberation. See also **Near field**.

Codec (**co**de-**dec**ode)
An algorithm for allowing data compression in digital systems, usually in accordance with psychoacoustic phenomena, which allows maximum data compression with minimum acceptable (perceptually/subjectively dependent) audible degradation of the reconstructed sound.

Damping
Damping refers to any mechanism that causes an oscillating system to lose energy. Damping of acoustic waves can result from the frictional losses associated with the propagation of sound through porous materials, the radiation of sound power, or causing a structure with internal losses to vibrate.

dBA
A weighted (filtered) measurement scale where the filter curves are roughly the inverse of the 40 phon contour of equal loudness. The scale was principally developed to relate to the subjective annoyance of noise around the 30–50 dB SPL ear sensitivity at mid frequencies, by correcting the high and low frequency measured levels to the same subjective level as the mid frequencies. (See Figure 8.4.)

D to A
Digital to **A**nalogue converters receive digitally coded signals, representing voltage waveforms, and by means of clocking and filtering, produce an analogue output voltage that should be as close a representation as possible of the waveform represented by the digits. See also **A** to **D**.

dBC
Similar to dBA but based on the inverse of the 80 phon contour of equal loudness, relating more to the subjective frequency balance at typical music listening levels. The low frequency response of the dBC curve is almost flat.

Deadsheets
Limp membranes having considerable inertia but little stiffness. They are widely used for the mechanical damping of acoustic waves. Those used in loudspeaker construction are normally between 3 and 15 kg/m^2.

Decibel
The standard unit of measure for level, or level difference. One tenth part of a bel. Multiplying a given quantity by 1.26 (to two significant figures) will give an increase in level of one decibel.

For example: $1\,\mathrm{W} \times 1.26 = 1.26\,\mathrm{W}$ (+1 dB relative to 1 W); $1.26\,\mathrm{W} \times 1.26 = 1.59\,\mathrm{W}$ (+2 dB relative to 1 W); $1.59\,\mathrm{W} \times 1.26 = 2.00\,\mathrm{W}$ (+3 dB relative to 1 W). It must be borne in mind that in *all* cases, a decibel represents a power ratio.

Decibels and sound pressure level (SPL)
Many observable physical phenomena cover a truly enormous dynamic range, and sound is no exception. The changes in pressure in the air due to the quietest of audible sounds are of the order of 20 μPa (20 micro-pascals), that is 0.00002 Pa, whereas those that are due to sounds on the threshold of ear-pain are of the order of 20 Pa, a ratio of one to one million. When the very loudest sounds, such as those generated by jet engines and rockets, are considered, this ratio becomes nearer to one to one thousand million! Clearly, the usual, linear number system is inefficient for an everyday description of such a wide dynamic range, so the concept of the bel was introduced to compress wide dynamic ranges into manageable numbers. The bel is simply *the logarithm of the ratio of two powers*; the decibel is one tenth of a bel.

Acoustic pressure is measured in pascals (newtons per square metre), which do not have the units of power. In order to express acoustic pressure in decibels it is therefore necessary to square the pressure and divide it by a squared reference pressure. For convenience, the squaring of the two pressures is usually taken outside the logarithm (a useful property of logarithms); the formula for converting from acoustic pressure to decibels can then be written:

$$\text{decibels} = 10 \times \log_{10} = \left\{\frac{p^2}{p_0^2}\right\} = 20 \times \log_{10}\left\{\frac{p}{p_0}\right\}$$

where p is the acoustic pressure of interest and p_0 is the reference pressure. When 20 μPa is used as the reference pressure, sound pressure expressed in decibels is referred to as *sound pressure level* (SPL). A sound pressure of 3 Pa is therefore equivalent to a sound pressure level of 103.5 dB, thus:

$$\text{SPL} = 20 \times \log_{10}\left\{\frac{3}{20 \times 10^{-6}}\right\} = 103.5\,\text{dB}$$

The acoustic dynamic range above can be expressed in decibels as sound pressure levels of 0 dB for the quietest sounds, through 120 dB for the threshold of pain, to 180 dB for the loudest (severely ear-damaging) sounds.

Decibels are also used to express electrical quantities, such as voltages and currents, in which case the reference quantity will depend upon the application (and should *always* be stated).

When dealing with quantities that already have the units of power, such as sound power or electrical power, the squaring inside the logarithm is unnecessary and the ratio of two powers, W_1 and W_2 expressed in decibels is then

$$10 \times \log_{10} \left\{ \frac{W_1}{W_2} \right\}$$

Distributed mode loudspeaker (DML)
A type of loudspeaker where the motor system excites a plate of high modal density. The source is diffuse and acts like neither a piston nor a point source. Radiation takes place from both sides of the panel, but the response is not typically that of a more conventional dipole, even when the panel is unbaffled. The stereo imaging is not as good as with conventional loudspeakers, but the response tends to be less disturbed by the room acoustics. Such loudspeakers can be advantageous as the ambient rear channels of a surround loudspeaker system, for example.

Doppler distortion
Frequency modulation dependent on the speed with which a source of sound is either approaching or receding from a listening position. The most common example is perhaps that of a train whistle, or horn, which suddenly drops in pitch as the listening position is passed. The whistle exhibits a constant frequency of output, but when it is approaching the listening position, the period between the cycles is compressed as the arrival delay shortens with time. As the source approaches, effectively the wavelength is shortened. This gives rise to the impression that a higher frequency is being emitted. The time of arrival for each subsequent cycle is reduced as the train approaches because it is emitted from a point nearer to the listener than was the previous cycle. Once the listening point is passed, the opposite effect occurs, with each subsequent cycle emanating from a more distant position, lengthening the wavelength and hence suffering a greater arrival delay. The frequency is thus perceived to have lowered. The degree of pitch change is dependent upon the relative speed of the sound source and the listener. (C. J. Doppler, 1842.)

Efficiency of loudspeakers — See Sensitivity and efficiency

Eigenfrequency
Frequency of the **Eigentones (see below)**.

Eigentones
The natural resonant tones of a space (or any other resonant system, in fact), ‘eigen’ being German for ‘own’. A room’s own tones. If a room is driven (excited) by a noise signal containing all frequencies, then all the eigentones will be driven. When the drive signal is terminated, however, only the eigentones will continue to ring-on, the decay rate being a function of size and total absorption. If a room is driven by a *musical* signal, only the eigentones that correspond to frequencies in the musical signal will be driven, and then only those eigentones will ring on when the music stops.

The eigentones thus dictate which frequencies will ring-on when the drive signal is stopped, but the input signal determines which eigentones are driven. Eigentone is another term for resonant mode. (See **Mode**.) If a room is driven at a frequency that does *not* correspond with an eigentone (mode), then the room response will decay more rapidly, once the drive signal has been stopped, than the frequencies which correspond with the eigentones. This is why widely spaced room modes give rise to uneven overall responses in a room. Frequencies corresponding to eigentones (modes) will be reinforced or cancelled, depending on source and receiver position, but other frequencies will be unaffected. See **Standing waves and resonances**.)

Euro (€)
The European currency unit, roughly equivalent to 1.25 US dollars in 2006. During the reading of the first draught of this book for assessment by the publishers, a referee requested that the currencies used in the text should be converted to one standard unit, such as the euro or the U.S. dollar. In reality, this proved difficult to achieve in any meaningful manner because the wild fluctuations of the currencies would lead to distorted senses of values outside of the local territory of that currency. In mid 2001, the euro would buy 87 cents of a U.S. dollar. By mid 2004 it would buy 1 dollar 34 cents, a rise in value of over 50 percent. Consequently, reference in this book to things manufactured in the USA is given in dollars, and to things made in Europe, in euros. Where any other currencies are stated, some means of assessing the relative value at the time referred to will be given in the text.

Far-field
The far-field is the region in which the radiation from a loudspeaker, or other acoustic source, can be considered equivalent to that of a point source, and in which the sound pressure *and* velocity reduce by 6 dB for each doubling of distance from the source.

Feedback

1. The instability that occurs when the output from a microphone is reproduced by a loudspeaker system, the output of which arrives back at the microphone. Regenerative feedback develops when the overall gain is greater than 1. Also known as 'howlround'.
2. The instability that occurs when an output of an electrical amplification system is re-applied to its input.
3. A term used by orchestral musicians to describe the way in which they hear the output from their own instruments reinforced via reflexions in their performance space. When fed to them via loudspeakers or headphones, the usual term is *foldback*.
4. *Negative* feedback (phase reversed) is used in electronic amplifier circuitry to reduce distortion.

Fourier Transform
The mathematical transform linking the time domain representation of a signal to its frequency domain representation. Application of the Fourier

Transform to a signal (waveform) reveals the frequency components in terms of their magnitude and relative phase (the Spectrum). Application of the inverse Fourier Transform to the spectrum yields the original waveform.

Frequency
The rate of change of phase with time. The number of complete cycles per unit interval of time for a sinusoidally time-varying quantity. The repetition rate of any cyclic event. Unit: Hertz: (Hz); previously measured in units of cycles per second (c.p.s.), and even earlier in half-vibrations per second.

Frequency response
The response of a system in terms of its amplitude and phase response. See **(Pressure) amplitude response** and **Phase response**.

Group delay
The frequency dependent response delay through electrical or mechanical systems which are given rise to by phase distortions. The group delay is related to the degree of phase shift.

Haas effect
H. Haas, 1951. See **Precedence effect**.

Hemi-anechoic chamber
An anechoic chamber with one hard boundary in which the sound sources can be embedded. Equates to 2π space – hemispherical radiation.

Impedance
A combination of resistance and reactance. Symbol Z. The ratio of pressure to velocity in acoustic systems, or voltage to current in electrical systems, expressed at a given frequency. (See also **Resistance** and **Reactance**.)

Infrasonic
Relating to frequencies below approximately 20 Hz. See **Subsonic**.

Intensity
'Sound intensity' is a very specific term, and represents the flow of acoustic energy. It is measured in units of watts per square metre, and should not be confused with either sound power level (SWL), sound pressure level (SPL) or loudness, but it *is* associated with the sound intensity level (SIL).

Interference field
See **Standing wave field**.

Kinetic energy
Energy of motion. Energy possessed by a body due to its motion, which can be converted to other forms of energy through the application of a braking force.

Linearity
A system can be said to be linear when the output contains only the frequencies applied to the input. A falling amplitude response with frequency is therefore a *linear* distortion. Such a response is *NOT* non-linear, as may be erroneously stated in many advertisements. See **Non-linearity**.

Loss factor
The reciprocal of the Q-factor. See **Q-factor**.

Magnetostriction
The property of some materials, for example iron, of expanding and contracting due to the influence of an applied magnetic field.

Microphone directivity patterns
Most microphones consist of a small diaphragm that moves in response to changes in the pressure exerted on it by a sound wave; the diaphragm motion is then detected and converted into an electrical signal.

The simplest form of microphone is one that has only one side of the diaphragm exposed to the sound field. If the diaphragm is sufficiently small, such a microphone will respond equally to sounds from all directions, and is termed ‘omni-directional’.

A microphone which has both sides of the diaphragm open to the sound field will only detect the difference between the pressures on the two sides. When a sound wave is incident from a direction normal to the diaphragm, there will be a short delay between the pressure on the incident side and that on the far side, and the microphone responds to the resultant pressure difference. When a sound wave is incident from a direction in line with the diaphragm, the same pressure is exerted on both sides of the diaphragm and the sound is not detected. This arrangement results in a ‘figure-of-eight’ or dipole directivity pattern.

If an omnidirectional microphone element and a figure-of-eight microphone element are mounted close together and their outputs summed, the resultant directivity pattern will lie between the two extremes of omni-directional and figure-of-eight patterns. If the sensitivities of the two elements are the same, the combined directivity pattern is known as cardioid, because of its heart shape. Various other patterns, such as hyper-cardioid and super-cardioid are achieved by varying the relative sensitivities of the omni-directional and figure-of-eight elements. Similar directivity patterns can be realised using only one microphone element. The microphone diaphragm is mounted at one end of a short tube, and the delay introduced to one side of the diaphragm by the tube gives rise to an approximation to a cardioid directivity pattern. More complex directivity patterns can be achieved by using more than two elements; the SoundField microphone, for example, has four elements that can be combined in a variety of ways.

Minimum phase and non-minimum phase

A minimum phase signal/system is one in which the phase shift associated with the amplitude response is the minimum that can be allowed whilst still exhibiting the properties of a *causal* system (one in which the output *never* arrives before the input). As there is a strict relationship between amplitude and phase in such systems, correcting either one will inevitably tend to correct the other. The low frequency response boost given rise to by the flush mounting of a loudspeaker in a wall is an example of a minimum phase response change, which therefore *can* be equalised to restore the free-field response in terms of both amplitude *and* phase (and hence it will also restore the time [transient] response). The essential factor is that no appreciable delay is involved between the generation of the signal and the effect of whatever is influencing it. If there is no appreciable delay, then there can be no appreciable phase-shifts, hence minimum-phase, or more fully, minimum phase-shift.

'Non-minimum phase' responses are those where amplitude correction, alone, cannot correct any phase disturbances. The far-field response of a loudspeaker in a reflective room is an example of a non-minimum phase effect. Here, there is a delay between the signal generation by the loudspeaker and the superimposition of the boundary reflexions on the overall response. Reflexion arrival times create phase irregularities, which are frequency and distance (time) dependent, so no simple manipulation of the amplitude response of the source can adequately compensate for the complex disturbances.

Another example of a non-minimum phase effect is in the combination of the various outputs of crossovers. In any filter circuits, either mechanical or electrical, there are inherent **Group delays** for any signal passing through them. The amount of group delay increases as the filter frequency lowers, and as its order (6, 12, 18, 24, etc. dB/octave) increases. A crossover will thus have different group delays associated with each section, and when the outputs are recombined, they will *not* produce an exact replica of the input signal. For this reason, conventional equalisation cannot be used to correct for response errors at crossover points. Amplitude correction will lead to further phase distortion, and hence time response errors. In practice, most crossovers above first-order are non-minimum phase devices.

Mode

A special pattern of vibration whose position remains invariant, as when a travelling wave and its reflexions superimpose themselves between two or more boundaries such that the peaks and troughs in the waveform coincide, and appear static. See **Standing wave field**.

Modes and resonances

Sound consists of tiny local changes in air density that propagate through the air as a wave motion at the speed of sound. The speed of sound is around 340 metres per second at normal room temperature and, although it is temperature dependent, it is independent of variations in the ambient pressure, and is the same at all frequencies. The frequency of a sound wave is measured in cycles per second (c/s or cps) known more usually these days as hertz (Hz), and is usually represented by the symbol f. The distance that a sound wave travels in one cycle at any frequency is known as the wavelength, represented by the symbol λ (lambda), and has the units of metres. The speed of sound is represented by the symbol c. The relationship between wavelength, frequency and the speed of sound is simple; wavelength is equal to the speed of sound divided by the frequency, or $\lambda = c/f$. Therefore, for example, a sound wave at a frequency of 34 Hz has a wavelength of $340/34 = 10\,\mathrm{m}$.

As a sound wave propagates away from a source in a room, it will expand until it reaches a reflective room boundary, such as a wall, from which it will reflect back into the room. The reflected wave will continue to propagate until it reaches other boundaries from which it will also reflect. If there is nothing in either the room or the boundaries to absorb energy from the wave, the propagation and reflexion will continue indefinitely, but in practice some absorption is always present and the wave will decay with increasing time. The point on the cycle of a sound wave (the *phase* of the wave) when it reaches the boundary depends upon the distance to the boundary and the frequency of the wave.

A rigid boundary will change the direction of propagation of an incident sound wave, but will maintain its phase, so the phase of a reflected wave can be calculated from the *total* distance propagated from the source. If this total distance is equal to a whole number of wavelengths, then the wave will have the same phase as it started with. When two boundaries are parallel to each other, a sound wave will reflect from one boundary towards the other, and then reflect back again to where it started, continuing back and forth until its energy is dissipated. If the distance between the boundaries is such that the 'round trip' from the source to the first boundary, on to the second boundary, and back to the source is a whole number of wavelengths, then the returning wave will have the same phase as the outgoing wave, and will serve to reinforce it. This situation is known as resonance. Resonances can also occur due to reflexions from multiple boundaries; the necessary requirement being that the sound wave eventually returns to a point with the same phase as when it left. One can imagine a whole set of possible combinations of reflexions in a typical room that allows the wave to return to its starting point, and therefore a whole set of frequencies for which resonance will occur. In fact, in theory, every room has an infinite number of possible resonances.

As stated above, if there is nothing in the room or the boundaries to absorb energy from a sound wave, a short duration sound pulse (or transient), emitted from a source will propagate around the room indefinitely. Of the infinite number of possible paths that the wave can take, only those that correspond to resonances at frequencies contained in the pulse will be continually reinforced; all other paths will not be reinforced. After a short time, the resulting sound field can be thought of as simply a sum of all of the resonances that have been excited. These resonant paths are known as the natural modes of the room, and the resonant frequencies are known as the natural frequencies, or '**eigentones**', of the room; both are determined uniquely by the room geometry. 'Eigen' is German for 'own', so the eigentones are a room's own particular, natural, resonance frequencies.

When sound absorption occurs within the room or boundaries, resonant modes still exist, but the wave will decay at a rate determined by the amount of absorption. To maintain a given sound level in a room in the presence of absorption, the source needs to be operated continuously, at a level dependent upon both whether or not resonant modes are being excited and the amount of absorption present. When the sound source emits a transient signal in the presence of absorption (for example, switching off a continuous signal), many different paths – not just resonant modes – will be excited, but after a short time, only the resonant modes will remain (because they tend to have less absorption); the room will 'ring' at the resonance frequencies until the modes decay. The reverberation time of a room is a measure of the average rate of decay of the sound in the room when an otherwise continuous sound source is switched off; it is the time taken for the sound level to fall to minus 60 dB relative to its initial, continuously excited, level. As the amount of absorption is increased, the sound level at the resonant frequencies will reduce, but the bandwidth of each mode (the range of frequencies over which the mode can be excited to a significant degree) will increase. When the boundaries are fully absorbent, the room modes no longer exist (an anechoic chamber).

When sounds such as speech or music are heard in a room, the level of the continuous components of the sound will be determined by whether or not they coincide with any room resonances that are excited. The transient components will 'hang on' at the resonance frequencies after the transient has gone. It should also be understood that the perception of the modal activity is not uniform within the space, because of the spacial distribution of the nodes and antinodes.

Mutual coupling

Mutual coupling is the term used to describe the interaction between two or more sound sources radiating the same signal. If the diaphragms are receiving different, uncorrelated inputs, then the output *power* summation will be simply that of the output of the different diaphragms. However, if the diaphragms are receiving the *same* input, then for frequencies whose wavelengths are greater than eight times the distance between the diaphragms (i.e. the diaphragms are less than one-eighth of a wavelength apart) the outputs will be substantially in-phase at any point in the room. The radiated *pressures* (not powers) will then superimpose, giving rise to a 6 dB increase in SPL for each doubling of the radiating area if the diaphragms are radiating the same power. This implies *four times* the output power, yet, when the radiated powers are equal but the radiated signals are *uncorrelated* (i.e. totally different signals) only a 3 dB SPL increase (due to the simple power summation) would result. Where a close boundary reflects a wave back on to the radiating surface of a diaphragm, the effect is the same as if a second diaphragm were radiating the same signal – the mirrored room analogy – and so the diaphragm radiated twice the power that it would do if moved away from the boundary.

This seemingly $1 + 1 = 4$ situation is due to the fact that, for a given diaphragm velocity, the *power* output is proportional to the diaphragm radius raised to the fourth power. Doubling the diaphragm area therefore yields a fourfold increase in power. The increased low frequency radiating efficiency of large diaphragms can be thought of as being due to all the individual parts of the diaphragm mutually coupling. The pressure radiated by one part of the diaphragm resists the movement of the adjacent parts. The increase in radiation resistance on a mass-controlled diaphragm, typical of a heavy woofer (whose movement can be considered independent of the local air pressure), gives rise to increased work being done, by having a greater pressure of air to push against, so more power is radiated.

As the frequency of radiation rises, or the loudspeakers (or the loudspeaker and a reflective surface) are sited further apart (i.e. the diaphragms are separated by *more* than one-eighth of a wavelength), the coupling becomes less in-phase so the radiation boost reduces. As the frequency or separation distance continues to rise, regions of in and out of phase interference will result, giving rise to a combined output power response as shown in Figure 7.11. Only for a listener on the central plane between the loudspeakers will the 6 dB pressure summation be maintained.

In a reverberant room, the output from a pair of stereo loudspeakers reproducing a central mono signal will sum by 6 dB on axis, at all frequencies, but the reverberant field will be driven by the combined power response, as shown in Figure 7.11b). The *overall* sound, therefore, will be darker (with more low frequencies) than that from a central mono source radiating the same signal and receiving the same input power as the mutually coupling pair.

Much more on this subject can be found in Reference [1]. at the end of the Glossary.

Near field

There are two quite distinct and separate definitions of the near field of a source of sound; one is related to the geometry of the source whilst the other has to do with the rate of expansion of radiating waves. The region beyond both near fields is known as the far field.

The **geometric near field** is defined as that region close to a source where the sound pressure does not vary as the inverse of the distance from a source. The extent of the geometric near field is dependent upon the detailed geometry of the sound source and is finite only at frequencies where the wavelength is shorter than a typical source dimension (for a circular piston this is when the wavelength is equal to the piston diameter); there is no geometric near field at lower frequencies. A point source does not have a geometric near field at any frequency.

The **acoustic near field** (or **hydrodynamic** near field) is defined as that region close to a source where the air motion (velocity field) does not vary as the inverse of the distance from a source, although the acoustic pressure may. The extent of the acoustic near field is inversely proportional to frequency: it is large at low frequencies and small at high frequencies. The sound field radiated by a point source has a sound pressure that varies as the inverse of distance from the source at all frequencies and distances. The air motion only varies as the inverse of distance in the far field. For practical sources, the extent of the acoustic near field is affected also by source geometry.

Decisions relating to the proximity of a listener to a close-field monitor should be made by considering the geometric near field only; the ear, being essentially a pressure sensitive organ, is insensitive to the presence of the acoustic near field.

Newton

A standard unit of force; symbol N. Easy to remember examples are that one newton is roughly equal to a *weight* (on the earth's surface) of 100 g, and that an apple of medium size is attracted to the earth by a force of around one newton. A force of one newton bearing down on a spring would be applied by a 100 g weight on the surface of the earth.

Non-linearity

A system is said to be non-linear when the output contains frequencies which were *not* present in the input signal, and are not due to system noise. Harmonic distortion, intermodulation distortion and rattles are sources of non-linearities. A system exhibiting any of these is said to be non-linear. See **Linearity**.

Noise weighting curves (dBA, etc.)

The human ear does not have a flat frequency response; a low frequency noise will generally sound quieter than a higher frequency noise having the same sound pressure level. A measurement of sound pressure level therefore does not yield an accurate measurement of perceived loudness unless the frequency content of the noise is taken into account. Noise weighting curves are used to convert sound pressure level measurements into an approximation of perceived loudness, by discriminating against low and high frequency noises. The most commonly used noise-weighting curve is known as A-weighting. An A-weighting curve is simply a filter with a response that rises with increasing frequency up to 2kHz, above which it falls off gently.

The frequency response of the human ear changes with changes in sound pressure level (see Figure 8.4), so different weighting curves are required for different levels. The dBA curve was developed for signals having loudness below 40 phon, the dBB curve was intended for somewhat higher levels. At levels over about 80 phon, the dBC curve should be used. Other curves are also in use, such as dBD, which can be used for high level industrial noise, and dBG, which is used for infrasonic and very low frequency noise assessments.

The widespread use of the dBA curve for the assessment of noise can give rise to poor results in situations when another weighting curve is more appropriate. Only at about 1 kHz and 6 kHz do all the curves agree. Between 3 and 4 kHz, errors of up to 10 dB can be found, and at low frequencies, the A-weighting curve can over-assess or under-assess noise nuisance levels by up to 20 dB, depending upon level. The dBA curve is often used at relatively high levels; a purpose that it was never intended for, and is not suited to, but sometimes this needs to be done for comparison purposes.

In any case, noise weighting should only be applied when one requires an approximation to the perceived loudness of a sound; it is therefore of most use in noise assessment. Noise weighting should never be applied when absolute values of sound pressure are required; in the measurement of loudspeaker frequency response, for example. Here, a flat measurement (unweighted) should be used.

Objective and subjective assessment
In acoustics in general, and in audio in particular, there is often some disagreement between that which our measurements tell us and that which we hear. In audio, objective assessment involves measuring the performance of a piece of equipment using instruments, and comparing this performance with a desired specification. Subjective assessment, however, involves auditioning the equipment under carefully controlled conditions and assessing particular aspects of the sounds that are heard. The successful assessment of the quality or suitability of a piece of audio equipment therefore, ideally, needs both approaches. Objective assessment is more easily carried out in the laboratory, or during production runs, than subjective assessment. To make a reliable and repeatable subjective assessment usually requires the ears of a number of subjects, and hence, often, a large amount of time.

Particulate
Relating to particles. Particulate motion = motion of the particles.

Pascal
A pressure of one newton per square metre. See **Newton**.

Passive systems
Systems, such as filters, without any source of external power other than the signal energy itself. An inductor/capacitor filter immediately before a loudspeaker drive unit is an example of a passive filter. See **Active systems**.

Phase response
The relative phase of the input and output signals as a function of frequency.

Phon
A unit of perceived loudness, such that a given change in the phon level would always produce an equal, subjective loudness change, irrespective of the actual SPL change. The contours in Figure 8.4 represent phon levels, which it can be seen do not relate directly to the physical sound pressure levels.

Pink noise
Filtered white noise (reducing 3 dB/octave with frequency) which yields equal energy per octave.

Pinna
Plural: pinnae. The outer part of the ear that projects outside the head. The ear flap.

Potential energy
The energy of position; such as imparted on a body by raising it in a gravity field. The energy concentrated in a spring is also potential energy.

Precedence effect
Also referred to as the **Haas effect**, and the law of the first wavefront. When two short-duration sounds are heard in rapid succession, the tendency is for the second sound to be psychoacoustically suppressed. The pair of sounds is perceived as one sound, coming from the direction of the first arriving source. The precedence effect operates when the second sound arrives within approximately 0.7–40 ms after the first sound.

The affect can often be overridden if the second sound is 7–15 dB higher in level than the first sound.

(Pressure) amplitude response
The ratio of the output amplitude of a system divided by the amplitude of the input as a function of frequency. When sound pressure is involved, the term 'pressure' is prefixed: for electrical or mechanical systems, the term 'amplitude response' suffices.

Psychoacoustics
Unlike the related discipline of acoustics, which is concerned with the physics of sound, psychoacoustics is the science of the perception of sound, particularly by humans. The stereo illusion, the cocktail party effect and the perception of pitch are all examples of psycho-acoustic phenomena.

PVA
Poly-vinyl acetate. A water-based adhesive that is water resistant once dry.

Q-factor
A measure of the sharpness of the peak in a resonant system. It is defined as

$$Q = \frac{f\mathrm{res}}{\mathrm{Bh}}$$

where

$f\mathrm{res}$ = resonant frequency

Bh = the half-power bandwidth of the resonance

Reactance
A reactive system is one which stores and releases energy without loss. Inductors and capacitors are reactive. Symbol X. The phase quadrature part of **Impedance**, q.v. Reactance can be thought of as the resistance to the flow of AC, as exhibited by inductors and capacitors, and is highly frequency dependent.

Resistance
Anything which impedes the flow of energy in such a way that the losses are dissipated (more usually into heat, but also into work). Symbol R. The in-phase part of an **impedance** q.v. Resistance acts equally on AC and DC currents, independent of frequency.

Semi-anechoic chamber
A room in which the absorption is incomplete, and contains a residual reflected component that can be corrected for during measurement analysis. See **Anechoic chamber**.

Sensitivity and efficiency
The efficiency of a loudspeaker is the proportion of total radiated sound power relative to the total electrical input power. The sensitivity is usually measured over a limited frequency band — usually in the centre of the loudspeakers' flat range of operation — at one metre distance and with an input of 2.83 volts. [Some early systems of measurement used 1 mW of input power at 30 feet distance.]
Efficiency is given as a percentage, whereas sensitivity is measured in dB SPL for one watt at one metre (2.83 volts into 8 ohms). The input signal is usually pink noise, suitably band-limited to suit the operating range of the loudspeaker or driver in question.
A complication arises in the specification of the sensitivity of compression drivers because the horns to which they are connected, by virtue of their different directivity patterns, distribute the radiated output power over different areas. It is therefore better to consider a combined driver/horn system as one unit before specifying the sensitivity, because the axial sensitivity will fall as the total area over which the sound is distributed is increased.

Shuffler
A type of circuit used in headphone reproduction to try to create inter-aural cross-correlation to simulate the effect of loudspeaker listening. This is done in an attempt to produce a frontal sound stage, because the stereo sound stage is generally inside or above the head of a listener when using headphones. Pre-World War II, the word was also used for other purposes, relating to middle/side (M & S) microphone matrixing.

Sine wave (and its frequency content)
A sine wave is a graph of the value of a single frequency signal against time. Strictly speaking, for a signal to consist to a single frequency, the sine wave must have existed for all time, as any change to the amplitude of the signal, such as during a switch-on or switch-off, gives rise to the generation of other frequencies: this has important implications for audio. Most audio signals

contain pseudo-steady-state sounds, such as notes played on an instrument. When these sounds are reproduced by an imperfect audio system, the excitation of any resonances in the reproduction chain will depend upon the frequency content of the signal. During a long note, the signal may be dominated by a few discrete frequencies, such as a set of harmonics, and the chances of resonances being excited are slim. However, during the start and stop of the note, a range of frequencies is produced, above and below those of the steady-state signal, and the chances of resonances being excited are increased. This phenomenon leads to the apparent pitch of the note being 'pulled' towards the frequency of any nearby resonance during the start, and particularly the end, of the note.

Sound Power Level (SWL)
The level of sound power, expressed in decibels, relative to a stated reference value. The unit is the decibel referenced to 1 picowatt (1 pW).

Sound Pressure Level (SPL)
The unit is the decibel, referenced to 0 dB SPL at 20 micro-pascals (20 μPa). It is defined by $20 \log_{10} (p_{rms}/p_{ref})$. **S**ound **P**ressure **L**evel, or SPL, doubles or halves with every 6 dB change, unlike the sound power, which doubles and halves with 3 dB change, because the power relates to the square of the pressure. In the acoustic and electrical domains, sound power equates to electrical power and SPL to voltage. Subjective loudness tends to double or halve with 10 dB changes: 10 dB higher being twice as loud, and 10 dB lower being half as loud. See also **Intensity**. Ten decibels relates to a ten times *power* change.

Squawker
Term of American origin for a mid-range loudspeaker – onomatopoeically relating to the sound of a large bird such as a seagull or macaw.

Standing wave field
The pattern of wave superposition that occurs in a reflective environment, whereby the distribution of peaks and troughs in the response throughout the space appear to be stationary. See **Mode**.

Standing waves and resonances
Standing waves occur whenever two or more waves having the same frequency and type pass through the same point. The resultant spatial interference pattern, which consists of regions of high and low amplitude, is 'fixed' in space, even though the waves themselves are travelling.

Resonant standing waves only occur when a standing wave pattern is set up by interference between a wave and its reflexions from two or more surfaces. *And*, when the wave travels from a point, via the surfaces, back to that point, it is travelling in the original direction. *And*, when the distance travelled by this wave is equal to an exact number of wavelengths. The returning wave then reinforces itself, and if losses are low, the standing wave field becomes resonant.

The simplest resonant standing wave to visualise is that set up between two parallel walls spaced half a wavelength apart. A wave travelling from a point towards one of the walls is reflected back towards the other wall, from which it is reflected back again in the original direction. As the distance between the walls is one half of a wavelength, the total distance travelled by the wave on return to the point is one wavelength; the wave then travels away from the point with exactly the right phase to reinforce the next cycle of the wave. If the frequency of the wave or the distance between the walls is changed, a standing wave pattern will still exist between the walls, but resonance will not occur.

It should be stressed that standing waves *always* exist when like waves interfere, whether a resonance situation occurs or not, and that the common usage of the term 'standing wave' to describe only resonant conditions is both erroneous and misleading. That is, all resonant modes are standing waves, but all standing waves are not resonant modes. See also, **Eigentones**.

Step-function
Alternative names are 'unit step function' 'Heaviside step function' and 'Heaviside function'. (O. Heaviside, 1892.)

$$\mathrm{H}(x) = 0 \text{ for } x < 0$$
$$\mathrm{H}(x) = 1 \text{ for } x > 0$$

Its value at $x = 0$ is not defined. The alternative notation $\mathrm{u}(x)$ is more common in signal processing.

Subsonic
This, in British English at least, is an aerodynamic term meaning below the speed of sound (as opposed to **Supersonic**). Its use implying below 20 Hz is incorrect.
Compare:
Infrared with ultraviolet
Latin: sub – under, super – on top of, above
Infra – below, ultra – beyond
Conventionally, 'sub' usually pairs with 'super' and 'infra' with 'ultra'. (See **Infrasonic**.)

Supersonic
An aerodynamic term meaning *above* the *speed* of sound. Its use relating to beyond the frequency range of hearing is archaic. **Ultrasonic** is the term now used for frequencies above the range of human hearing.

Transfer function
Alternative term used (at least in electro-acoustics, although not in all subjects) for the frequency response. What you get out relative to what you put in. A flat frequency response implies a flat transfer function.

Transient response
The response of a system to an impulsive input signal. An accurate time (transient) response requires an extended frequency response and a smooth

phase response. A low frequency amplitude response roll-off, for example, will give rise to the lengthened time (transient) responses, as shown in Figures 6.3 and 6.4. The more the frequency response is curtailed, either in terms of frequency of turnover or steepness of slope, the more the transient response will be smeared in time.

Tweeter
Term of American origin for a high-frequency loudspeaker, onomatopoeically imitating the high-pitched 'tweet-tweet' sound made by small birds.

Ultrasonic
Relating to frequencies above approximately 20 kHz. Some authorities limit the term to a maximum of 10^{13} Hz, beyond which the term 'hypersonic' is used. Hypersonic is also used in aerodynamics, relating to *speeds* beyond five times the speed of sound.

Weighting
The pre-multiplication of data by a set of weighting factors. A bias applied to improve measurement compatibility with subjective assessment.

White noise
A random noise signal containing all frequencies. Statistically the response has equal energy per bandwidth in hertz. For example, 20 Hz to 25 Hz (5 Hz bandwidth) would have equal energy to the band from 1000 Hz to 1005 Hz (also a 5 Hz bandwidth), and hence on a spectrum analyser shows a response rising 3 dB per octave as the frequency rises.

Woofer
Term of American origin for a low frequency loudspeaker – onomatopoeically relating to the deep bark, or 'woof' of a large dog.

X-curve
An empirically derived curve for the equalisation of monitor systems in the cinema industry. The curve is shown in Figure 12.11, and is used, somewhat flexibly, according to room size and decay time to improve the sonic compatibility of the perceived frequency response. Due to many reasons, of which not all the mechanisms or implications are fully understood, the frequency balance of the soundtracks tend to become brighter sounding as the room size increases, and/or the decay time increases. Thus, a large room with a longer decay time, using a flat monitor response, would be perceived to be over-bright when compared to the sound in a smaller room with a shorter decay time.

Reference

1 Borwick, John, *Loudspeaker and Headphone Handbook*, 3rd Edn, Chapters 1 and 9, Focal Press, Oxford, UK and Boston, USA (2001)

Thanks to Professor James Angus for verifying the accuracy of this glossary.

Index