Loudspeaker and Room Transfer Functions Correction using DSP Techniques

Lars G. Johansen

Aarhus School of Engineering
Aarhus University
2014

Part 1

Previous Approaches to Loudspeaker and Room Correction

1.1 Previous scientific reporting on equalisation research

For several years now electronic equalisation has been a topic occupying scien-tists, and while the breakthroughs in room acoustics (and to some extent also psychoacoustics) were made decades ago, recent developments in signal proces-sing techniques and digital hardware now constitute the basis for realtime implementation of equalisation systems. The main areas treated in scientific pa-pers the past fifteen to twenty years relating to equalisation are listed below.

- Correcting for bad room acoustics
- Improving sound quality in cars
- Developing robust inverse filtering techniques
- Developing adaptive equalisation systems
- Modelling the acoustic properties of rooms
- Realising simple and robust realtime equalisation systems
- Creating virtual acoustics, e.g. by phantom sources

Below, the content of the ten most essential papers on room equalisation are presented as the original abstracts in publishing order followed by discussions.

Matti Karjalainen et al.: ***Comparison of Loudspeaker Equalization Methods Based on DSP Techniques,*** *J. Audio Eng. Soc.,* ***47****(1/2), 1999 January/February*

Methods of loudspeaker response equalization using digital filters are compared. In addition to generally known methods and techniques a recently introduced new principle, based on warped filters, is described. Different equalization methods are compared from the points of view of equalization error both objectively and subjectively, computational efficiency, as well as robustness and precision requirements of each method. The study is limited to the linear (magnitude and phase) equalization of loudspeaker free-field responses.

R. Walker: ***Equalisation of Room Acoustics and Adaptive Systems in the Equalisation of Small Room Acoustics,***
proc. of the Audio Eng. Soc. 15th International Conference, Copenhagen 1998

The basic properties of acoustics in small rooms and the effects leading to uneven frequency responses are described. The principles underlying possible equalisation schemes are outlined. This paper concentrates on the underlying causes of response irregularities rather than discussing proprietary solutions. It is shown that additional low-frequency loudspeakers can help to optimize the low-frequency responses without detriment to the higher frequency image localization. By using several low-frequency sources, some response control can be achieved over a limited region. At high frequen-cies, the objective responses of rooms do not match the subjective impressions, and equalisation based on measure parameters is likely to be unsuccessful. Equalisation schemes based on relatively wide-band or short-term response parameters are likely to match the auditory perception mechanisms better and, thus, to be more successful. Automatic or active control of equalisation is essential for complex schemes. However, it adds little to the fundamental principles and the inherent limitations of either low- or high-frequency equalisation.

John N. Mourjopoulos: ***Digital Equalization of Room Acoustics,***
J. Audio Eng. Soc., Vol. 42, No. 11, 1994 November

Signal-processing methods such as digital equalization can in theory achieve a reduction in acoustic reverberation. In practice, however, the realization of these methods is only partially successful for a number of objective and subjective (perceptual) reasons. Two of these problems, the dependence of the equalizer performance on the source and receiver positions and the requirement for extremely lengthy filters, are addressed. It is proposed that all-pole modelling of room responses can relax the equalizer filter length requirement, and the use of vector quantization can optimally classify such responses, obtained at different source and receiver positions. Such classification can be used as a spatial equalization library, achieving reduction in reverberation over a wide range of positions within an enclosure, as was confirmed by a number of tests.

Stephen J. Elliott et al.: ***Practical Implementation of Low-Frequency Equalization Using Adaptive Digital Filters,*** *J. Audio. Eng. Soc.,* ***42****(12), 1994 December*

Adaptive digital filters have been used in an experimental investigation of low-frequency equalization of a single-channel sound reproduction system in a car. The problems are discussed of adapting the digital filter so that a smooth transition is achieved between the equalized low-frequency response (below 400 Hz) and the unequalized high-frequency response. Equalization of the response at only one point in the car is found to cause degradation in the response at others. Multiple-point equalization, in which the response at four positions is best equalized in a least-squares sense, was found to give only modest overall improvements in this case. The best strategy for a single filter appears to be weighted multiple-point equalization, in which the error at the most important listening position in the car is more heavily weighted in the adaptation algorithm. This gave worthwhile improvements in the response at the selected location, without significant degradations at other points. A very similar effect can also be achieved with the single-point equalization systems either by using a leak in the adaptive algorithm or by using an adaptive filter with a smaller number of coefficients.

P. G. Craven and M. A. Gerzon: ***Practical Adaptive Room and Loudspeaker Equaliser for Hi-Fi Use,*** *a preprint of the Audio Eng. Soc. 92nd Conv., Vienna 1992*

This paper describes a soon to be available system for stereo hi-fi use for equalising room and loudspeaker impulse responses across a listening area, using a chirp measure-ment and decimated digital equaliser implemented using a single DSP chip per channel. Filter impulse response lengths up to one second are achievable at low frequencies. Psychoacoustic requirements are discussed for subjectively satisfactory results and naive strategies, such as mean-square optimisation or minimum-phase equalisation are found to be inadequate.

Ronald Genereux: ***Adaptive Loudspeaker Systems: Correcting for the Acoustic Environment,*** *Audio Eng. Soc. 8th Int. Conference, Washington D.C. 1990 May*

Loudspeaker designers have long recognized the influence of the acoustic environment on the perceived quality of an audio system. The availability of powerful digital signal processor (DSP) integrated circuits has created interest in the application of adaptive digital filtering techniques to the equalization of loudspeakers in rooms. A review of some recently proposed implementations is followed by a discussion of the issues which must be considered in order to address the problem successfully. Results from an experimental system are presented.

S. J. Elliott and P. A. Nelson: ***Multiple-Point Equalization in a Room Using Adaptive Digital Filters,*** *J. Audio Eng. Soc.,* ***37****(11), 1989 November*

A method is presented for designing an equalization filter for a sound-reproduction system by adjusting the filter coefficients to minimize the sum of squares of the errors between the equalized responses at multiple points in the room and delayed versions of the original electrical signal. Such an equalization filter can give a more uniform frequency response over a greater volume of the enclosure than a filter designed to equalise at one point only. The results of computer simulations are presented for equalisation in a "room" with dimensions and acoustic damping typical of a car interior, using various algorithms to adapt automatically the coefficients of a digital equalisation filter.

C. Bean and P. Craven: ***Loudspeaker and Room Correction Using Digital Signal Processing,*** *a preprint of the Audio Eng. Soc. 86th Conv., Hamburg 1989*

This paper concerns digital signal processing techniques used to equalise loudspeakers and correction of room effects well away from the loudspeaker itself. Measured results for the linear and minimum phase corrections are presented. Various algorithms for calculating digital equalisation filters are discussed, including techniques to limit the length of the digital filter. Some of the problems inherent in the correction of room acoustics are highlighted and some of the perceived effects after correction are outlined.

M. Miyoshi and Y. Kaneda: ***Inverse Filtering of Room Acoustics,***
IEEE Trans. on Acoustics, Speech and Signal Proc., Vol. 36, No. 2, 1988 February

A novel method is proposed for realizing exact inverse filtering of acoustic impulse responses in a room. This method is based on the principle called *m*ultiple-input/output *in*verse *t*heorem (MINT). Because a room impulse response generally has non-minimum phases, it has been impossible to realize exact inverse filtering of room acoustics using previously reported methods. However, the exact inverse of room acoustics can be realized using the proposed method. With this method, the inverse is constructed from multiple FIR filters (transversal filters) by adding some extra acoustic signal-transmission channels produced by multiple loudspeakers or microphones. The coefficients of these FIR filters can be computed by the well-known rules of matrix algebra. Inverse filtering in a sound field is investigated experimentally. It is shown that the proposed method is greatly superior to previous methods that use only one acoustic signal-trans-mission channel. The results in this paper prove the possibility of sound reproduction and sound receiving without any distortion caused by reflected sounds in a room.

S. T. Neely and J. B. Allen: ***Invertibility of a Room Impulse Response,***
J. Acoust. Soc. Am. ***66****(1), 1979 July*

When a conversation takes place inside a room, the acoustic speech signal is distorted by wall reflections. The room's effect on this signal can be characterized by a room impulse response. If the impulse response happens to be minimum phase, it can easily be inverted. Synthetic room impulse responses were generated using a point image method to solve for wall reflections. A Nyquist plot was used to determined whether a given impulse response was minimum phase. Certain synthetic room impulse responses were found to be minimum phase when the initial delay was removed. For these cases a minimum phase inverse filter was successfully used to remove the effect of a room impulse response on a speech signal.

1.2 Papers on room modelling and subjective evaluation

The following five papers address the room modelling issue in order to establish a basis for equalisation which is more feasible than the blind point-to-point inverse filtering scenario and subjective tests done on equalised room transfer functions.

Y. Haneda et al.: ***Multiple-Point Equalization of Room Transfer Functions by Using Common Acoustical Poles,***
IEEE Trans. on Speech and Audio Proc., Vol. 5, No. 4, 1997 July

A multiple-point equalization filter using the common acoustical poles of a room transfer function is proposed. The common acoustical poles correspond to the resonance frequencies, which are independent of source and receiver positions. They are estimated as common autoregressive (AR) coefficients from multiple room transfer functions. The equalization is achieved with a finite impulse response (FIR) filter, which has the inverse characteristics of the common acoustical pole function. Although the proposed filter cannot recover the frequency response dips of the multiple room transfer functions, it can suppress their common peaks due to resonance; it is also less sensitive to changes in receiver position. Evaluation of the proposed equalization filter using measured room transfer functions shows that it can reduce the deviations in the frequency characteristics of multiple-point room transfer functions better than a conventional multiple-point inverse filter. Experiments show that the proposed filter enables 1-5 dB additional amplifier gain in a public address system without acoustic feedback at multiple receiver positions. Furthermore, the proposed filter reduces the reflected sound in room impulse responses without the pre-echo that occurs with a multiple-point inverse filter. A multiple-point equalization filter using common acoustical poles can thus equalize multiple room transfer functions by suppressing their common peaks.

J. Mourjopoulos et al.:

Pole and Zero Modelling of Room Transfer Functions,
Journal of Sound and Vibration (1991) ***146****(2), 281-302*

The feasibility of using all-pole or all-zero model approximation of room transfer functions was examined especially in respect to the degree that such approximations are suitable for removing room reverberation from signals. Two aspects of the above models were assessed: Their success in reducing room transfer function order and their insensitivity to measurements taken for different source and receiver position inside a room. Both these aspects are crucial to the success of practical dereverberation methods. The tests were carried out on simulated and measured data and it was found that all-pole models are more suitable than all-zero models when both of the above two aspects must be satisfied by the approximation model. It was also shown that, for a given room, an optimum all-pole model exists which approximates the room transfer functions.

P. L. Schuck et al.:

Perception of Perceived Sound in Rooms: Some Results of the Athena Project, *Audio Eng. Soc. 12th Int. Conference*

Experiments comparing the use of multiple versus paired comparisons for loudspeaker evaluation, the use of digital band splitting and equalization, and assessing the scope of the loudspeaker/room interaction effect on loudspeaker preference ratings, are dis-closed. Finally, results of an experiment which utilizes loudspeaker/room equalization to lower the variability of preference scores, as well as increase to them, are presented.

Søren Bech:

The Influence of the Room and of Loudspeaker Position on the Timbre of Reproduced Sound in Domestic Rooms,
Audio Eng. Soc. 12th Int. Conference

A round robin test has been conducted to examine the interaction between a loudspeaker and the room with respect to fidelity of timbre of reproduced sound. Three rooms, three loudspeaker positions, four loudspeakers, four programs, and six subjects were used for the experiment. The statistical analyses show that the main factors which have a significant influence on the assessment of fidelity of timbre are room, loudspeaker position, loudspeaker, and program. Several of the interactions between the main factors are significant, however the most important is that between loudspeaker and positions. The results show that the room will influence a) the overall fidelity of timbre of reproduced sound in all positions; b) the perceived differences between loudspeakers in similar positions; and c) the perceived differences between loudspeakers in different positions in the same room. The loudspeaker position will also have a significant effect on the level of fidelity of timbre. The degree of influence of the room and of the positions will depend on the directivity characteristics of the loudspeaker.

1.3 Discussions on the material

Basically, the topics addressed in these fifteen papers can be condensed into the issues below:

Physical limitations

Why will correction/equalisation never work all over the room, and what artifacts are observed when designing optimal (in some sense) equalising to one part of the room and in fact experience the reproduced sound in another. No magic solutions are given. Essentially, the more optimal equalising in confined parts (or even points) of the room, the more prices are paid in other parts in the sense of deteriorating reproduced sound compared to the one without a correction system employed. For low frequencies, however, it seems like more loudspeakers spread around the room can ease up the problem. Also, appropriate modelling of a room response, eg. by an all-pole function, may prove beneficial. Some report that equalisation based on multiple receivers (microphones) improves overall performance - there is no general agreement though.

Mathematical limitations

Equalisation is in most of the papers related to inverting the frequency spectrum of a measured impulse response. If the response is not representing a minimum-phase system (and it rarely is in ordinary listening rooms) the inversion is not possible without introducing artefacts of some kind. Desired and relatively easy correction must be established then within the set of minimum-phase transfer functions.

Psychoacoustic issues

More of the papers express their concern to what is in fact audible by the human hearing and transform that knowledge into relaxed criteria for the equaliser/-correction system. There seems to be a common agreement on not to attempt too detailed equalisation in both time- and frequency domains as this is likely to cause undesired artifacts even in the sweet spot.

Signal processing issues

In relation to equalisation systems, the late eighties and early nineties were dominated by the fact that it suddenly became possible to implement real-time digital correction electronic hardware. Mostly, however, the reports remain at the discussion level. In fact in this context only the paper by Craven and Gerzon reveals a practically functional system. Maximum equalisation filter length is also discussed as a necessary concern due to limits in available hard-ware. Although more papers have been presented on the practical issues, few, if any, report on entire stand-alone correction systems, ie. a system which by itself can do the acoustic environment measurements (impulse response acquisitions), calculate the equalisation filters/algorithms, and finally perform the real-time handling of the electrical signals.

Room modelling issues

Modelling the room acoustics seems to be a good way of eliminating the effects in a measured room transfer function that usually mess up equalisation attempts. By the use of discrete-time models based on polynomial ratios (ARMA models), the perhaps most annoying acoustical phenomena are captured (the modal resonances), and the beautiful thing is that they immediately relate to digital correction filter design.

1.4 Some practical room correction systems

In 1991 **B&W** launched a room correction system called the **B&W Digital Control Unit**. Essentially it was based on the seventies *Teledyne* experimental system and to a large extent Colin Bean and Peter Craven's earlier theoretical work (see above). A technical presentation can be found in *Hi-Fi News & Record Review 1991 December*.

Marantz AX-1000, a combined room correction system and room simulation system, was accessible in 1992. An impressive piece of equipment for that time representing a power of 13.3 mips.

TacT Audio introduced a room correction system in 1997 called **RCS 2.2**. The equipment only works together with a PC for calculating the equalisation filters, impulse response acquisition and real-time processing are both done independently. Splitting the measured impulses into three frequency bands, the RCS 2.2 uses different strategies for designing the correction filters. In the lowest frequency region, the correction resolution is 0.6 Hz, up to 1,500 Hz the resolution is 5 Hz, and above it is 300 Hz. Processing is powered by four Motorola DSP56002 processors. The RCS 2.2 was received with some excitement in the professional world, and although really expensive it has been reviewed (with good marks) in the home theatre / Hi-Fi literature. Many prejudges pointing towards digital Hi-Fi equipment was put to death, and the reviews reported on noticeable improvements in the subjective quality of reproduced sound. Later, TacT has launched simpler and cheaper versions producing al-most the same subjective quality improvement.

Other recent pieces of room correction equipment are listed below. Also Philips launched a room equaliser in the early nineties (and a not too expensive set of active speakers in fact) **-** neither of those obtained much commercial success.

Furthermore, Roister has a quite new set of room equalisers (digital compensators) on the market

- **SigTech AEC-1000/2000**. Can handle 2,200 FIR filter taps in 48 kHz sampling and works adaptively, i.e. the correction filters are changed on-the-fly,
- **Snell Acoustics RCS 1000**. Room correction system with 1.5 Hz resolution,
- **Behringer Ultracurve Pro 8024**. A 1/3 octave equaliser that produces the band corrections automatically based on measurements.

Why not more systems?

Not many systems of the kind have emerged yet, and the world still waits for a simple, cheap, and completely automatic room correction system. Perhaps that is why the concept of room correction has not yet gained widespread popularity. Another explanation is that just ten years ago, digital Hi-Fi equipment was still approached with much scepticism (at least among purists). Perhaps still not completely vanished, the "fear-of-digital" has greatly reduced here at the beginning of the new century. Thus, a breakthrough for the digital room correction may be just around the corner.

Part 2

REALISTIC TARGETS AND FEASIBLE GOALS FOR ROOM ACOUSTICS CORRECTION

2.1 Brief summary and initial recommendations

We know now that sound in both time and frequency domains is influenced quite differently according to the positions of source and receiver. Different transfer paths evoke different patterns of separable reflections as well as different patterns of excited modal frequencies. Operating with the revealed transfer functions, equalisation essentially means creating an exact inverse function, and thus being able to entirely remove the room acoustics by deconvolution (in this context de-reverberation is synonymous). For the reasons discussed earlier, that scenario only works in infinitely small points in the room, hence it is a mathematically beautiful technique but indeed not a feasible approach to practical room acoustics correction system design.

Is anechoic sound preferable?

Also, which sometimes is neglected, human beings do in fact not prefer sound reproduced in total anechoic surroundings. Some acoustic information must be present in order to create a comfortable listening situation, and it is not sufficient to include that piece of information in the recording (neither as a live acoustic event nor as artificial reverberation). So, total de-reverberation (exact equalisation, inverse filtering) is not favourable from a qualitative point of view. It is relevant here to remember the inherent loudspeaker deficiencies namely, apart from non-ideal on-axis impulse characteristics, the usual lack of low frequency reproduction in the sub-octaves typically below 50 Hz and the non-ideal and sometimes unsmooth off-axis characteristics.

The more complex the better?

The room acoustical facts and the way human hearing works fortunately both speak against trying to build up a very complex correction system. At least in the sense of employing a very accurate correction scheme. Although it may seem tempting from a mathematical point of view, a very detailed correction would always only apply to very limited parts of the room which cannot be satisfying from a practical point of view. An optimised listening space of at least 1 cubic metre should be required.

Is accuracy really positive?

Additionally, we tend to define the term *accurate* only from a technical view-point. Maybe such accurate correction does not correlate very well with perceived sound quality, and maybe we fool ourselves if striving for such accuracy. Accuracy may not be a positive goal in this case! One must be aware that sound quality (and sound quality improvement) will always be a rather diffuse and subjective measure, depending on a subtle and not fully understood combined time/frequency behaviour. The human hearing does not comply with the way technical equipment measures room acoustics and performs analysis on impulse responses. As accurate they well may be, in room correction design we are dealing with humans evaluating the improvements - and presumably not being able to discard personal preferences. The really tricky thing is to deal with “non-accuracy” at an appropriate level! However, within that framework the following issues must (and can) be dealt with. Notice that the issues are closely coupled to fig. 2.16 in part A.

End user expectations?

From the user’s point of view, it must be considered realistic to require system operation based on only one initial microphone measurement in the optimised listening space. As literature shows, reasonable performance can be accomplished doing so, and when carefully designed perhaps only with minor drawbacks compared to a multi-microphone system. Also the system design should be aiming at stand-alone operation, thus pointing towards a fairly simple system not involving vast processing resources.

2.2 Targets in the time domain

In the time domain it was shown in part A to be appropriate to separate early from late parts of the impulse response using the statistical time t_{stat} as limit.

Targets below the statistical time

In the early part, the separable reflections (or maybe the combined pattern of the first 5 to 8 reflections) should be considered. At least the most predominant ones (usually, the first floor reflection is among those), must be reduced in magnitude below the limit of audibility. For average listening room these first 5-8 reflections lie within a time span of approximately 10 ms from the direct sound. Preferably, the first 5-8 reflections should be reduced by 6-10 dB. We would secure then that no (or very small) audible effects are retained.

Targets beyond the statistical time

In the late part, statistically modelled, not much can be done. Reverberation time for average listening rooms usually is 400-500 ms, and techniques should reduce T_{60} if considerably larger than that in order to bring up the sensation of a more controlled listening room. Techniques for whitening the reverberation tail should also be applied so that no single frequency region is excessively represented, - with due respect to the fact that the further out in the tail the less high frequency content should be present.

2.3 Targets in the frequency domain

In the frequency domain it seems appropriate to separate low frequencies from high frequencies putting the limit around the Schroeder frequency, f_{schr}, which for average listening rooms amounts to 100-200 Hz.

Targets below Schroeder frequency

In the low frequency region modal resonances are dominant in creating severe peaks and dips in the transfer function spectrum. Peaks which are audible as disturbing or even unpleasant resonances must be removed or reduced. As receiver and loudspeaker positions change so do the peaks and dips. Hence, it may be hazardous to "fill up" the dips. When moving to other positions than the one equalised, the dip compensation is probably not (or less) needed and the resulting severe excess amplification is highly undesirable. For average listening rooms, the bandwidth of even the narrowest modal resonance peak is 3-6 Hz, so compensation with a resolution of 2 Hz will be sufficient. Also, under f_{schr} the equaliser should aim at broadband energy compensation, say in one third octave bands. The properties can be summarised as a desirable target band around 0 dB with a tolerance of ±2 dB. Quite large deviations should be allowed in the way that magnitude changes of order 0-10 dB could be tolerated - but in narrow bands only, eg. 3-5 Hz. A lot of energy can be removed though when attenuating the resonances so maybe, in order to retain the overall sense of bass "power", the target band should be lifted a few decibels.

Targets in the sub frequency region

Embedded in the measured room impulse response, the loudspeaker highpass characteristics show off, usually revealing only little excitation of the room below approximately 50 Hz. As part of the correction, the low frequency reproduction should be extended down to approximately 25 Hz, or as far down as the loudspeaker is capable of handling the more power without introducing audible distortion. Below some f_{low}, eg. 25 Hz, there is no reason for further compensation. The human hearing is only little sensitive in this region and the equaliser might end up taking amplifiers and loudspeakers to the very edge of their performance capacity (most not too expensive loudspeaker's natural roll off frequency lies around 40-50 Hz). Hence, the equaliser target also includes a highpass filter at f_{low}.

Targets beyond Schroeder frequency

In the high frequency region not much can be done without introducing new unpleasant phenomena, but timbre should be considered, ie. the spectral energy should be equalised - presumably in no more detail than what can be done in one third octave bands. This very modest criterium complies well with the fact that considerable position sensitivity is present already at a few times f_{schr} and grows larger with frequency. Hence, there is no physical reason for narrow band compensation at higher frequencies. Psychoacoustically it is difficult to detect a difference of 2 dB in two successive one third octave bands, so a reasonable target band is 0±1.5 dB. Again we suggest a target roll off at high frequencies f_{high}, around 25 kHz, beyond which we may presume to have no interest in compensation. From approximately 1 kHz it may be beneficial to let the equaliser follow a slightly decaying target instead of a completely flat target, say 4-6 dB of total decay up to the upper limit f_{high}. Due to the larger absorption at high frequencies a room response will usually show a decay behaviour, and a subjective evaluation may presumably show a preference towards an equalised response that does not compensate (entirely) for such decay - introducing more high frequency energy.

Targets for phase characteristics

Although doubt still rules concerning audibility of transfer function phase, it must be recommended to strive for linear phase systems. Not an easy task however when exact equalisation is not allowed. Smooth and not to excessive group delay will then be the second best goal.

Excess phase correction?

Another fundamental issue is: Can we ignore equalisation of the excess phase part in loudspeaker/room transfer functions? At the moment there is no clear answer, in an earlier investigation we showed that the excess phase, under certain circumstances is audible, see [36]. Equalisation of non-minimum transfer functions is generally problematic. We need a delay to obtain a causal impulse response, and therefore we can easily run into problems with pre-responses showing audible (and very annoying) pre-echoes if we move the head to a position where the “compensation of the pre-response” is inaccurate, and such positions do exist. Also, it is not clear how important it would be to separate the two contributing elements of the compound loudspeaker/room impulse response. Craven & Gerzon [79] propose an equalisation of the loudspeaker including non-minimum phase correction. They believe it is important to achieve linear phase in the woofer highpass response. A recent review concerning equalisation of loudspeakers is given by Karjalainen et al. in [74].

Equalising excess phase is a matter of accuracy. A high degree of excess phase correction is possible only when dealing with point-to-point transfer functions. These are generally not desirable, and as shown in fig. 2.1 and table 2.1 the transfer function excess phase increases with frequency and with time. The slight decrease beyond 300 ms is because the high frequency content reduces as time increases, see fig. 2.2). So fortunately one could say, the excess phase is not dominant in the region where correction is feasible.

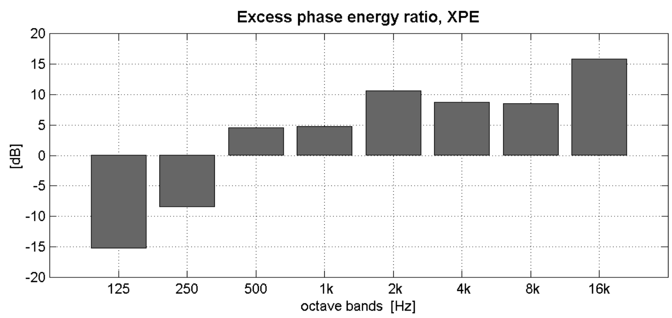

Figure 2.1. *Excess phase as a function of frequency.*

section [ms]	XPE	section [ms]	XPE
0 - 75	-0.5	200 - 300	3.2
0 - 100	-0.4	250 - 350	6.4
50 - 150	-0.3	300 - 400	4.1
100 - 200	-0.4	350 - 450	2.6
150 - 250	0	400 - 500	3.9

Table 2.1. *Excess phase as function of response segments cut out of the impulse response at different times.*

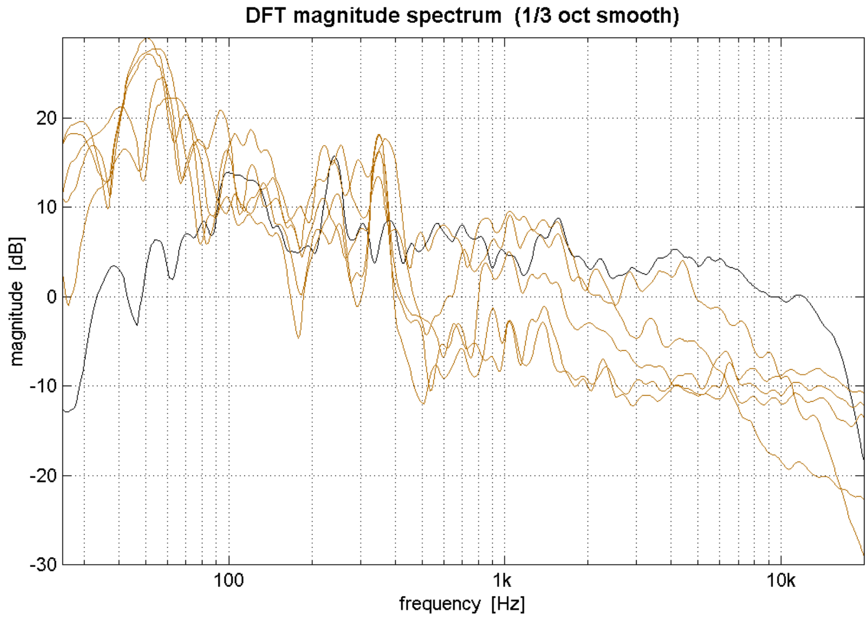

Figure 2.2. *Spectral behaviour of the impulse response segments given in table 1.2. The black spectrum is based on the entire response.*

2.4 Targets for energy relations and other parameters

Early/late energy relations

Controlling the time and frequency domains as suggested above usually also results in a smooth and non-transient behaving system regarding the energy relations measured by the room acoustical parameters like DR, C80, D50 etc. No direct action should be taken to improve in detail these parameters. Using them in objective evaluation however is a powerful tool in a first hand judgement of the success of the correction. If one or more of the parameters show transient behaviour, eg. a big jump in magnitude from one octave band to another, it is most likely that subjective evaluation will reveal characterisations such as *annoying*, *disturbing*, *unpleasant*. Clarity and the DR signal energy relation in particular seem to play an important role for the perception of a high quality room. As a rule-of-thumb the experience tells us that Clarity (for small rooms C35 is applied instead of C80) should exceed 10- 12 dB and DR should be 3-6 dB, see [88] and [90].

Reverberation parameters

Assuming a reasonably low initial value of T_{60}, instead of using a lot of effort to lower the reverberation time further (implying actions close to dereverberation which we do not recommend), it may be beneficial to go for a reduction in early decay time. Even a slight EDT decrease will contribute to the sense of a more damped room with a high subjective clarity.

Temporal repetition

Repetition of events (whole parts of the impulse response) is measured by the temporal diffusion M (also denoted TD), and some diffusion technique may have to be considered if initially M is too small, ie. below approximately 10 dB. Otherwise, this parameter should just serve as a control measure.

2.5 Qualitative specifications and the IMOLE goal

Evaluating the equalised responses, the impulse response should be expected to possess as much “delta impulse” like behaviour as possible with no visible long lasting resonances and exponentially decaying noise-like effects. Similarly, in the magnitude spectrum no sudden jumps should be expected. Rather, it should look smooth with a slight decaying slope. It is also natural (but from a signal processing point of view not at all trivial) to introduce a criteria, call it IMOLE (IMprove Or LEave) saying that equalisers must be designed to generally improve the sound reproduction or at least not deteriorate further the sound reproduction even in spaces away from the “sweet spot” set for correction.

2.6 General issues to address in correction algorithms

Which transfer functions to correct?

It will be assumed in the following that correction always is applied to the combined loudspeaker/room transfer function. Although, by some decomposition technique, it could be possible to separate the effects of loudspeaker equalisation and room equalisation, there are really no reasons for doing so - apart from mere curiosity. For very esoteric loudspeaker sets one could perhaps appreciate a pure room correction. A related issue concerns how to pre-process responses before the equalisation actions take place. Pre-smoothing or averaging of more responses may serve to let the equaliser fulfill the subjective demands more easily.

Positioning of source and receiver

Just like the acoustic properties of the room in which we apply correction most definitely play a role for the final result, so will the initial (or assumed) positions of the loudspeakers and the listener. Different transmission paths correspond to different transfer functions, these being input to a correction algorithm. From an intuitive point of view it is reasonable to assume that in order to reach to correction goals some transfer functions serve as more difficult inputs than others.

The preferable excitation

In a listening room with no electronic correction of sound reproduction, when source and receiver positions fall together with many anti-nodes, we may expe-rience annoying audible phenomena - the modal resonance effects become prevalent in the sound field. On the other hand we completely loose low frequency reproduction if nodes of many resonances coincide with source and receiver positions. The best compromise is positioning of source and receiver where most strong and separable modal resonances are excited to some extent - let us say between 40% and 70%.

Positioning recommendations

From a correction point of view, the more energy already present by inherent room resonance excitation the easier the correction algorithm design becomes. In the extreme case loudspeakers are placed in the corners and all modes and their combinations produce appreciable resonances. The correction algorithm can then concentrate on simply reducing energy in order to meet the goals. Otherwise, it may be forced to put in energy to make up for a possible set of poorly excited resonances. As long as the positions involved are totally fixed we can live with that scenario, but it is not hard to imagine what happens when for example the listener moves to another position in the room causing stronger coupling to the modal resonances. The equalised transfer function for this position might end up showing 10-15 dB peaks. So putting in too much energy compensating for poor receiver position coupling may not be desirable.

Part 3

Digital Signal Processing Techniques for Correction Algorithm Design

3.1 Transfer function decomposition and Hilbert transform

The Z-transform H(z) of a measured room impulse response h(n), although non-parameterised, can be modelled by a generalised digital IIR filter as in eq. 3.1. Essentially, the generalised systems modelling encompasses both numerator and denominator polynomials. The roots a_j in the numerator symbolise the zeros in the transfer function inside the unit circle and the b_j are the zeros outside the unit circle. Correspondingly, c_i denote the inside of the unit circle poles of the transfer function and d_i the outside poles.

$$H(z) = \frac{\sum_{j=0}^{M} \beta_j z^{-j}}{1 - \sum_{i=1}^{N} \alpha_i z^{-i}} = \frac{\prod_{j=1}^{M_{in}} (1 - a_j z^{-1})}{\prod_{i=1}^{N_{in}} (1 - c_i z^{-1})} \cdot \frac{\prod_{j=1}^{M_{out}} (1 - b_j z)}{\prod_{i=1}^{N_{out}} (1 - d_i z)} \qquad (3.1)$$

Classification of systems

Using the discrete time systems definition and examining the *Regions Of Convergence* in the Z-domain, it can be shown that the four combinations of *causality* and *stability* fall in two categories as shown in table 3.1.

Stability/Causality systems characterisation		**Equiva-lent to:**	**Stability/Causality systems characterisation**	
stable	causal	$	unstable	non-causal
unstable	causal	$	stable	non-causal

Table 3.1. *Combinations of stability and causality.*

Only stable systems of course are of interest leaving no room for any d_i in H(z). Systems with zeros inside the unit circle only (no b_j) are called *minimum phase* systems, and we refer to the phase contribution from the outside zeros represented by b_j as *excess phase*. The term excess phase is sensible since it can be shown that this is really purely an extension of the phase shift in H(z). It has no relation to the magnitude of H(z). Systems which have zeros both inside and outside the Z-domain unit circle are called *mixed phase* systems $H_{mix}(z)$.

Minimum/excess phase decomposition

Through decomposition, any transfer function H(z) can be split into a product of a minimum phase part, an allpass part, and a pure delay according to eq. 3.2 (sometimes $H_{allpass}(z)$ also contains the delay z^{-n}). The minimum phase part consists of all the poles, the natural “inside” zeros (a_j), and any “outside” zero b_j mapped to the inside with magnitude $1/r(b_j)$, call them b'_j. The allpass part consists of the original “outside” zeros b_j and poles cancelling out the artificially introduced zeros b'_j, these poles are denoted by a'_j. All possible magnitude information of H(z) then is held in $H_{mph}(z)$, whereas the magnitude of $H_{allpass}(z)$ as defined will always be unity. An example decomposition is shown graphically with poles and zeros in the Z-domain in fig. 3.1. It can be shown (described below) that the minimum phase thus defined and the magnitude in a transfer function are unambiguously linked together.

$$H(z) = H_{mph}(z)\,H_{allpass}(z)\,z^{-n} \tag{3.2}$$

The Hilbert transform

Splitting a sequence h(n) into a sum of its *even* and *odd* parts, it can be shown, see e.g. [54], that the Fourier transform H(z) is composed of the Fourier trans-form of $h_{even}(n)$ as the real part and the Fourier transform of $h_{odd}(n)$ as the imaginary part. For a causal sequence h(n), the sequence can be composed in terms of its even part only, and thus the complex Fourier transform H(z) is actually determined entirely from its real part $H_{RE}(z)$. Now using the complex convolution theorem, an analytic relationship between $H_{RE}(z)$ and $H_{IM}(z)$ can be established, - this is called the *Hilbert transform*. Splitting H(z) in terms of the magnitude and the phase angle, and taking the complex logarithm, we have an expression of the form in eq. 3.3.

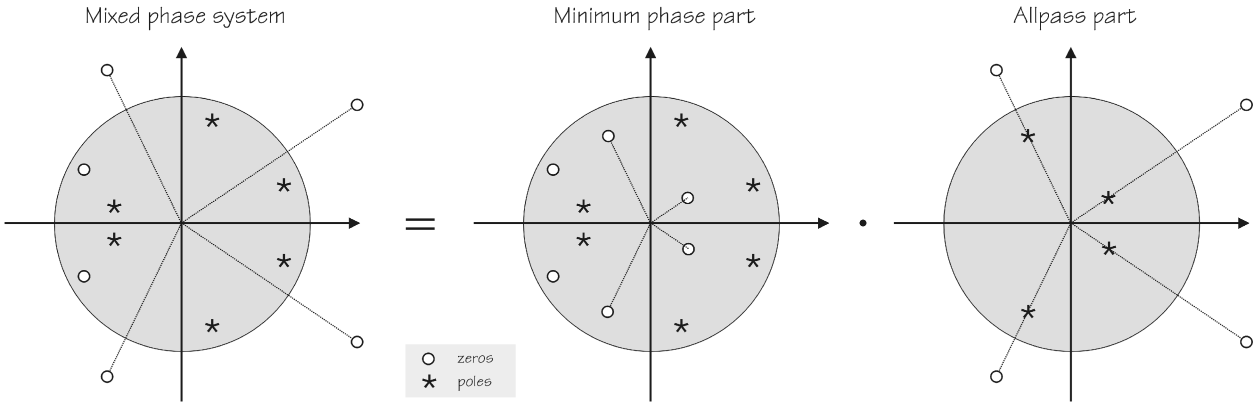

Figure 3.1. *Mixed phase system decomposition into the minimum phase part and the allpass part representing excess phase.*

$$\hat{H}(z) = \log\{H(z)\} = \log|H(z)| + j\arg\{H(z)\} \tag{3.3}$$

In the right-hand side the real part represents originally the magnitude in H(z), and the imaginary part represents originally the phase angle of H(z). Taking the inverse Fourier transform, we arrive with a sequence $\hat{h}(n)$, the *complex cepstrum*, for which we can say that the causal part corresponds to the real part of its Fourier transform $\hat{H}_{RE}(z)$. Performing the logarithmic operation on eq. 3.1, inspecting the poles and zeros, and expressing the inverse Fourier transform as a series expansion, it is seen that the a_j and c_i, representing the minimum phase in h(n), in fact form the causal part of $\hat{h}(n)$.

Hence, since for all causal (and real and stable) sequences $H_{IM}(z)$ is analytically related to $H_{RE}(z)$ through the Hilbert transform, if we now compose $\hat{H}(z)$ entirely by means of $\hat{H}_{RE}(z)$, finding H(z) taking the complex exponential and finally turning back into the time domain to h(n) by the inverse Fourier transform, h(n) will be minimum phase. The minimum phase part $h_{mph}(n)$ of a response h(n) can be found then by the magnitude of H(z) only. Also, the magnitude |H(z)| can be derived unambiguously from the minimum phase. Thus, the Hilbert transform in this setup actually forms a pair of equations, eq. 3.4 and 3.5, where *P* denotes the *Cauchy principal value*.

$$\arg\{H(e^{j\omega})\} = -\frac{1}{2\pi} P \int_{-\pi}^{\pi} \log|H(e^{j\omega})| \cot\left(\frac{\omega-\theta}{2}\right) d\theta \tag{3.4}$$

$$\log|H(e^{j\omega})| = \hat{h}[0] + \frac{1}{2\pi} P \int_{-\pi}^{\pi} \arg\{H(e^{j\omega})\} \cot\left(\frac{\omega-\theta}{2}\right) d\theta \tag{3.5}$$

Homomorphic deconvolution

Separation of minimum phase systems and allpass systems can be accomplished by employing homomorphic deconvolution. It can be shown that if a sequence x(n) contains minimum phase only, then its complex cepstrum $\hat{x}(n)$ will turn out to be causal. Similarly, given a causal cepstrum $\hat{x}_{caus}(n)$, it is ensured that it in fact represents a time domain sequence $x_{mph}(n)$ with a Z-transform con-taining minimum phase only. Consequently, the minimum phase part of a response h(n) can be extracted by first forming the complex cepstrum, then deleting any non-causal information, and finally by reverse operations turning back to the time domain, see fig. 3.2. The complex cepstrum can be estimated using eq. 3.6 (the larger L the better), and the minimum phase part is found in the cepstral domain by multiplying $\hat{h}(n)$ with a causalisation function $l_{mph}(n)$ as in eq. 3.7 leaving back only the causal part of the cepstrum.

$$\hat{h}(n) = IDFT_L\left(\ln\left|DFT_L\{h(n)\}\right|\right) \quad (3.6)$$

$$l_{mph}(n) = 2\,u(n) - \delta(n) \quad (3.7)$$

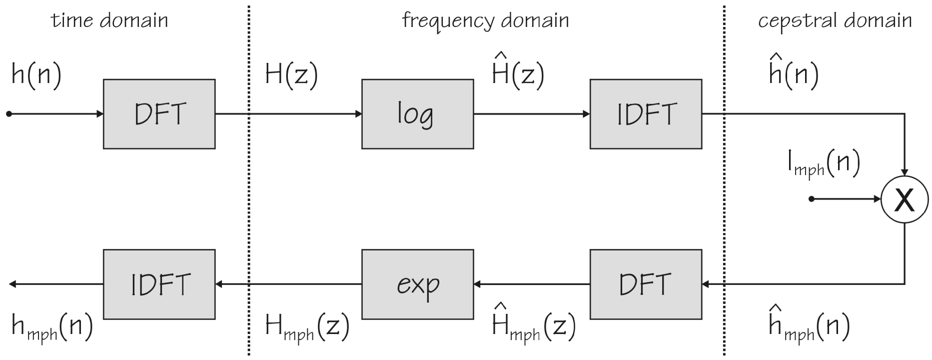

Figure 3.2. *Homomorphic deconvolution of h(n) finding the minimum phase part $h_{mph}(n)$.*

The allpass part $h_{allpass}(n)$ is then determined by dividing the complex frequency spectrum of the minimum phase part into the spectrum of the original response and then transform it back into the time domain, see eq. 3.8. For large values of L, the operations described above are an approximation to using the discrete *Hilbert transform* described above.

$$h_{allpass}(n) = IDFT_L\left(\frac{H(z)}{H_{mph}(z)}\right) \quad (3.8)$$

Non-causal excess phase equalisation

Inverting a mixed phase system $h_{mix}(n)$ leads inherently to instability. The inter-esting thing is however that an unstable but causal system also can take the form of a stable but non-causal system, so by allowing non-causality the correc-tion of maximum phase systems actually does become possible. The excess phase in a room impulse response can then be equalised by introducing a delay. To account for all the excess phase, ideally the non-causality thus imposed should last infinitely long which is of course not possible. From sheer practicality, equalising excess phase is then a compromise between the degree of correction and the amount of delay which can be tolerated.

Optimally, when equalising $h_{max}(n)$ in a point-to-point scenario, no artefacts are present in the correction delay part but the non-causal correction will introduce artefacts whenever the reproduction system is altered even slightly. The arte-facts can be audible, eg. as pre-echoes and/or pre-reverberation, which is extremely annoying. Presumably, these audible phenomena represent a severe sound quality degradation, so excess phase equalisation should be handled with great care and in general it cannot be recommended.

Invertibility of room impulse responses

We can immediately invert $H_{mph}(z)$ but not $H_{allpass}(z)$ since the zeros outside the unit circle in $H_{allpass}(z)$ turn into poles with the same location when inverted and then producing an unstable system. This means essentially that we can compensate for the magnitude and minimum phase, but not for the excess phase. Allowing non-causal systems (inherently by introducing delays), it is possible however also to compensate for at least some of the excess phase. Blind equalisation approaches operating on a mixed phase system $H_{mix}(z)$, room transfer functions always are, can lead to awkward results due to $H_{allpass}(z)$, so in order to guide the equa-liser, it might be beneficial to decompose $H_{mix}(z)$ and feeding only $H_{mph}(z)$ to the equaliser. If then excess phase correction is paramount then, it can be taken care of separately.

3.2 Parametric transfer function modelling

Modelling a transfer H(z) in a parametric way can be useful in equalisation, particularly when the phenomena in H(z) seem to be in good accordance with the technique leading to the parameterised model. In general, taking a starting point in eq. 3.9, parameterised models can be classified in three categories, the MA (moving average) models, the AR (autoregressive) models, and the ARMA (combination of MA and AR) models. A *moving average* model emerges when one or more b_j is different from zero and all a_i are zero saying that no denomi-nator polynomial exists and H(z)=B(z). Hence only modelling by zeroes is possible, and since zeroes represent dips in the frequency magnitude spectrum MA modelling is probably not the best way to model resonances.

$$H(z) = \frac{B(z)}{A(z)} = \frac{\sum_{j=0}^{M} b_j z^{-j}}{1 + \sum_{i=1}^{N} a_i z^{-i}} = \frac{b_0 + b_1 z^{-1} + \cdots + b_M z^{-M}}{1 + a_1 z^{-1} + \cdots + a_N z^{-N}} \tag{3.9}$$

Autoregressive (AR) modelling

When the B(z) polynomial has coefficients $b_j = 0$ (apart from the constant b_0), H(z) is an autoregressive function $H(z) = b_0/A(z)$. Here we have roots in the denominator causing peaks in the magnitude spectrum. This is more like what we are looking for since these peaks could well resemble the modal resonance peaks in the measured transfer function. One way to establish an autoregressive model is through *Linear Prediction*. Linear prediction assumes a H(z) = 1/A(z) model and will attempt to find the A(z) polynomial coefficients a_i so that the error between the model and the measurement is minimised in the least squares (LS) sense. The proce-dure assumes that a particular sample of say an impulse response h(n) can be formed (or predicted) as a linear combination of previous samples like in eq. 3.10. The a_i here are called *Linear Predictive Coefficients*, hence the modelling technique is referred to as the LPC method.

$$\hat{h}(n) = \sum_{i=1}^{N} a_i h(n-i) \tag{3.10}$$

There are more ways to approach estimation of the coefficients a_i by minimi-sing the LS prediction error, a popular one however is the autocorrelation approach (estimation can also be based on covariances). It can be shown [xxx] that a_i can be found solving the set of equations in eq. 3.11, where $r_h(k)$ refer to the autocorrelation coefficient of h(n). Eq. 3.10 can be put more simply as a matrix equation as in eq. 3.12, where **R** is a Toeplitz matrix. In figs. 3.3 through 3.6 are shown the frequency magnitude spectra of AR modelling up to 290 Hz of a measured impulse response, the orders being 12, 24, 48, 72 respectively. Not surprisingly, when N is increased the modelling gets better - N can be chosen too large however, rendering the R matrix singular or near-singular. This is commonly referred to as over-modelling.

$$\begin{bmatrix} r_0 & r_1 & r_2 & \cdots & r_{N-1} \\ r_1 & r_0 & r_1 & \cdots & r_{N-2} \\ r_2 & r_1 & r_0 & \cdots & r_{N-3} \\ . & . & . & \cdots & . \\ r_{N-1} & r_{N-2} & r_{N-3} & \cdots & r_0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ . \\ a_N \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ . \\ r_N \end{bmatrix} \tag{3.11}$$

$$Ra = r \tag{3.12}$$

Naturally, the impulse response of an AR model (which can be considered an IIR filter with no numerator) is infinitely long, but one great thing about the AR approach is that when using the model for straightforward inverse equalisa-tion filter design, the equalisation filter G(z) becomes an FIR filter according to eq. 3.13. FIR filtering is equal to moving averaging, it has finite impulse response, and it is inherently stable. AR modelling is attractive then because of its ability to capture the phenomena in the measured transfer function that we want to address, and because it produces simple and stable and minimum phase inverse filters. As shown in [xxx], AR modelling does not always work satisfactorily, eg. handling a cosine seems difficult. In practice though, model-ling room transfer functions - at least the low frequency parts up to a few hundred herz, seems to cause only little trouble.

$$G(z) = \frac{1}{H(z)} = A(z) \tag{3.13}$$

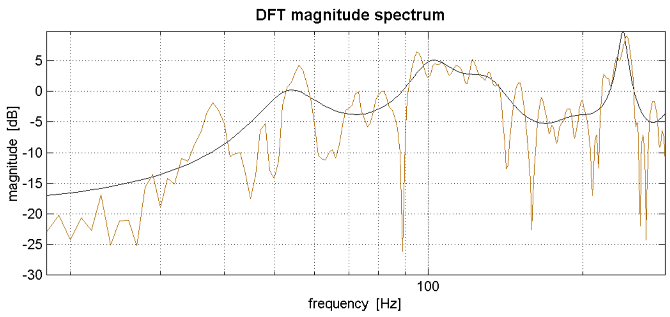

Figure 3.3. *Order 12 AR modelling of room response (grey).*

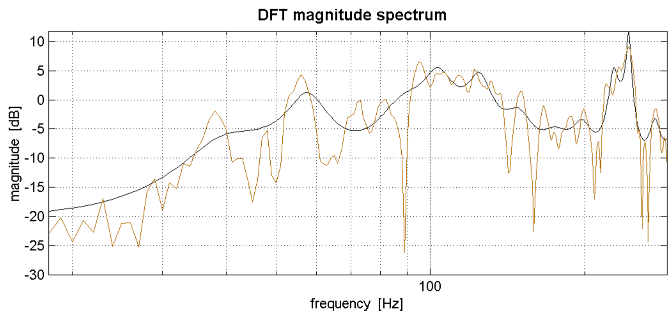

Figure 3.4. *Order 24 AR modelling of room response (grey).*

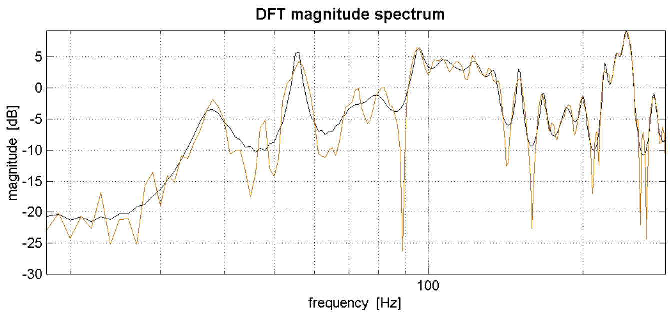

Figure 3.5. *Order 48 AR modelling of room response (grey).*

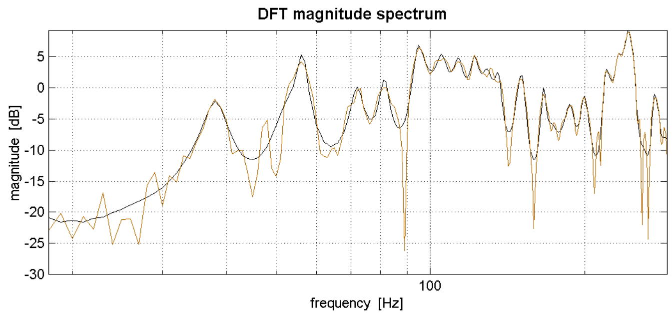

Figure 3.6. *Order 24 AR modelling of room response (grey).*

ARMA modelling

The combination of MA and AR yielding models with both numerator and denominator polynomial may seem attractive since it most obviously should capture more characteristics in the measurements. There are two reasons why this is not quite so obvious. First, the ARMA model will also try to find the dips in the frequency magnitude spectrum of the measurement, but we are actually not interested in those since equalising design based on the ARMA model then will possess peaks which cannot be recommended according to the discussions in section 2. The second reason is that the procedures behind forming the ARMA models, estimating both b and a coefficients, are quite complex and not always robust. Using there procedures would require extensive checking for model validity which also contributes to the complexity. Furthermore, solutions to ARMA modelling are never derived analytically, they always come as iterative improvements of an initial estimate, so a stop criterion must be applied. Among the methods for ARMA modelling are

Prony's method, the Stieglitz-McBride method, and the Yule-Walker method (based on frequency domain specifications), see [xxx] and [xxx].

3.3 Spectral inversion, smoothing, and regularisation

Inverting a smoothed spectrum

Without any modifications, a pure inversion of H(z) is generally not possible without tolerating considerable delays. If equalisation of minimum phase only can be accepted though, we can decompose H(Z) and invert $H_{mph}(z)$. For the reasons discussed previously even this is probably not a good idea in practical correction systems, but a feasible approach could be to *smooth* the spectrum, i.e. perform an averaging in 1/N octave bands. This way, narrow band effects are averaged out and in fact a time domain smearing is imposed also. Now it is no problem finding an inverse spectrum of the smoothed H(z). When such smoothing is done, any phase information is lost initially. However, by using the Hilbert transform, we can derive a completely new phase part and construct a new complex Fourier transform from the smoothed magnitude part. Turning back into the time domain, and allowing a small delay (necessary to account for a slight non-causality due to the smoothing), we have a minimum phase equaliser based on a smoothed transfer function.

Transfer function regularisation

If no smoothing is allowed (or perhaps in a combination), so-called regularisation of a transfer function subject to inversion can be done, see [xxx]. Regularisation, referring to eq. 3.14, will suppress the dip (zeroing) effects with a desired amount determined by the D constants, and hence the inverse transfer function, G(z), will not suffer from equal size peaks relative to the initial dips. This can be advantageous when we want to design low frequency equalisation by spectral inversion instead of using the AR modelling. Still the inversion should be based on a minimum phase decomposed version of H(z).

$$G(z) = \frac{1+\rho_1}{H(z)+\rho_2} \tag{3.14}$$

3.4 Warping the frequency scale

Frequency warping is a way to redistribute the attention on the frequency scale. For example, more focus can be put on the low end of a frequency band at the expence of the high end detail. Actually, frequency warping is a con-formal mapping where the normal delay element z^{-1} in discrete-time systems is replaced by a first order allpass filter D(z) as in eq. 3.15. The Z transform of an impulse response h(n) can the be written introducing the frequency warping as in eq. 3.16, and looking at a single sinosoid with frequency T=2Bf/f_s, its warped counterpart is given by T' in eq. 3.17.

$$D(z) = \frac{z^{-1}-\lambda}{1-\lambda z^{-1}} \tag{3.15}$$

$$H(z) = \sum_{n=0}^{\infty} h(n)\, z^{-n} = \sum_{k=0}^{\infty} w(k)\, D^k(z) \tag{3.16}$$

$$\omega' = \arctan \frac{\left(1-\lambda^2\right)\sin(\omega)}{\left(1+\lambda^2\right)\cos(\omega)-2\lambda} \tag{3.17}$$

Hence, we have a nonuniform-resolution frequency representation of H(z). This can be very advantageous when trying to reflect the mechanisms of the human hearing, where a logarithmic-like frequency dependent frequency reso-lution is observed. Choosing 8 rightly (0.7-0.75), will produce a frequency scale resembling that of the *Bark* scale. Now, impulse responses can be warped, equalisation filters can be determined in the warped domain, and the equali-sation filter response can be dewarped (same procedure just using negative 8). The drawback is however that using D(z) turns FIR filters into IIR filters, so stability is not automatically ensured (particularly for large filter orders), and the equalisation filters have infinite impulse responses which must be truncated (if not in fact the equalisation is carried out in the warped domain). These WFIR filters can represent a more adequate allocation of filtering capacity in acoustical applications, the frequency dependent resolution can be determined through eq. 3.18. More information on frequency warping can be found in [xxx] and .....

$$\Delta f_{WFIR} = \frac{f_s}{N} \frac{1-\lambda^2}{1+\lambda^2-2\lambda\cos(2\pi \frac{f}{f_s})} \tag{3.18}$$

In fig. 3.7 are shown five frequency magnitude versions of a room impulse response h(n) up to 1200 Hz. The upper black spectrum is a straightforward Fourier transform, the grey spectra are Fourier transforms of the impulse re-sponse warped using 8=0.4, 0.6, 0.75, 0.85 respectively. Fig. 3.8 shows from the upper spectrum and down h(n) equalised by LPC derived inverse filters of orders 25, 50, 100, 200, 400 respectively. The same equalisation filter orders are applied in fig. 3.10, only here the filters operate on a warped response with 8=0.75. Fig. 3.9 shows what happens when the same order (N=100) equali-sation filter is applied to the five degrees of warping shown in fig. 3.7. It is seen that when applying the warping, the same equalisation performance is obtained with a considerably lower order of the filter - approximately a factor of three. Also, the equalisation “effort” seem to be distributed more evenly at the logarithmic frequency scale.

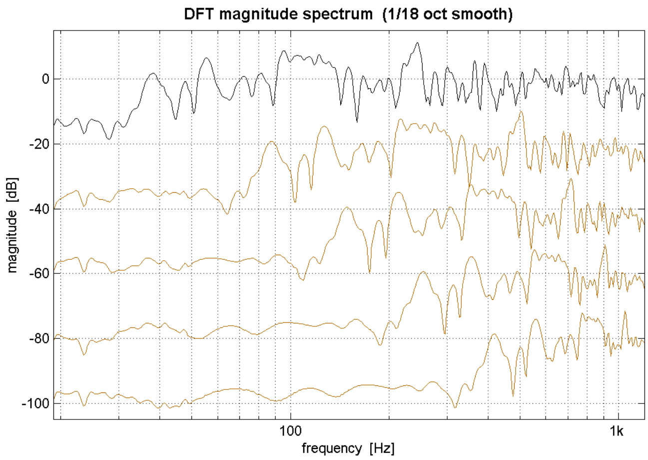

Figure 3.7. *Four warped transfer functions (grey curves) - the upper is not warped. 8 factors are 0; 0.4; 0.6; 0.75; 0.85.*

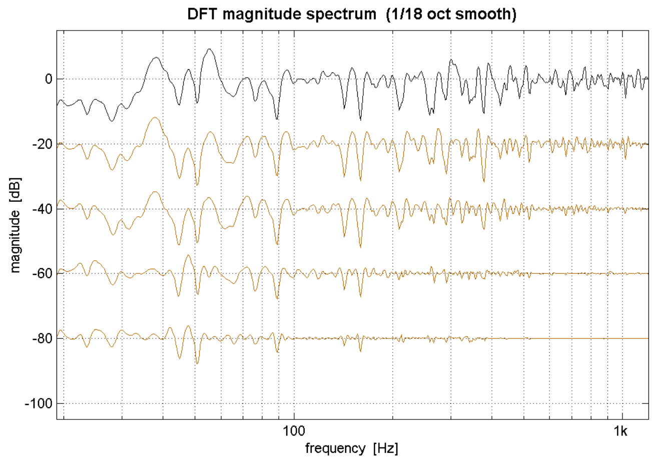

Figure 3.8. *Equalisation of the non-warped transfer function - LPC derived inverse filters of orders 25; 50; 100; 200; 400.*

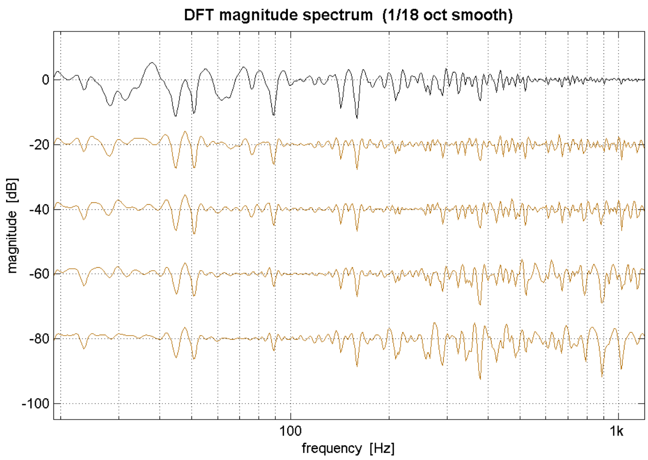

Figure 3.9. *Four warped transfer functions (grey curves) - as in fig. 3.7 equalised by an LPC inverse filter of order 100.*

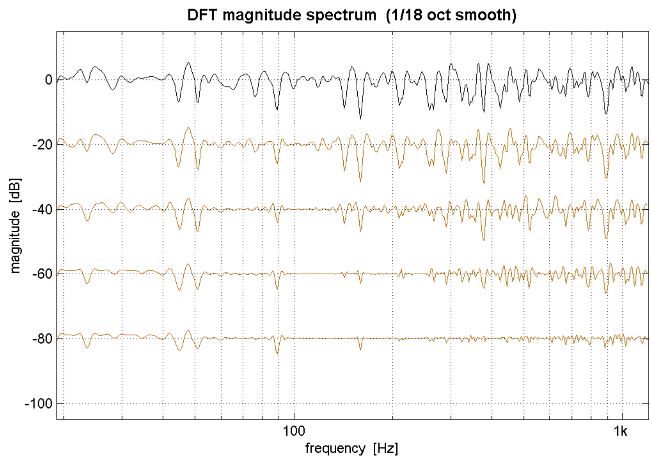

Figure 3.10. *Equalisation of the 8=0.75 warped transfer func-tion - LPC derived inverse filters of orders 25; 50; 100; 200; 400.*

3.5 Separable reflections attenuation and diffusion

Attenuation of early reflections

A technique has been developed for attenuating early strong reflections in a room impulse response h(n). The technique qualifies be the fact that it does not try to deconvolve the reflections, that would be alarming from a position sensitivity point of view. Instead it attenuates each reflection and anything else in a small time span around the reflection. The algorithm is not extremely com-plicated and can easily be incorporated in a room acoustics correction scenario. By the techniques described in the above sections, only frequency domain effects are addressed directly and we can just hope that the actions will also have a positive effect in the time domain. The reflections attenuation algorithm addresses annoying time domain effects. Forming the algorithm involves the steps below, and it is a quite new way to address room acoustics correction from a practical viewpoint.

- a segment c(n) of length t_c covering the early reflection is cut out of h(n)
- the magnitude spectrum of c(n) is smoothed getting G(z)
- G(z) is inverted and reverse transformed to g(n)
- g(n) is causalised into $g_{caus}(n)$ by a delay t_{caus}
- $g_{caus}(n)$ is multiplied with a special window

Early reflections diffusion

As an alternative to the reflections attenuation, in order to render the first strong reflections inaudible as separable phenomena, a diffusion filter could be applied. A small sequence (a few milliseconds in length) of white noise, which is exponentially weighted to decrease in average to 10%, is convolved by the measured impulse response. The early strong reflections are then smeared in time and the early part of the response will contain more energy, so Clarity will increase but DR will probably not since the direct sound is not amplified. The situation would resemble that of having many reflections close to each other. Their amplitude may be fairly high but due to the small spacing their individual contributions are probably rendered inaudible.

3.6 Excess phase equalisation

Equalisation of excess phase

Since $h_{allpass}(n)$ holds no information about the frequency magnitude, we can convolve the initial response by this and only the phase is changed. In fact, it can be shown that performing the convolution as given in eq. 3.19 results in a complete removal of excess phase. So only a minimum phase version of h(n) is left. Of course for infinitely long sequences, eq. 3.19 cannot be determined, so one will have to choose a finite length of the causalisation. Also practical rea-sons can dictate such a restriction, e.g. introducing delays of just a few hundred milliseconds destroys syncronisation in a combined audio/visual reproduction. This reduces the amount of excess phase that can be corrected for. Also to minimise the risc of pre-echo and pre-reverberation effects, causalisation should probably be chosen fairly small.

$$h_m(n) = h(n) \otimes h_{allpass}(-n) \tag{3.19}$$