2015
ISEAT
International Symposium
on ElectroAcoustic
Technologies

Controlling the impulse responses and the spatial variability in digital loudspeaker-room correction.

Lars-Johan Brännmark[1] Mikael Sternad[2]
（1. Dirac Research AB, Uppsala, Sweden;
2. Dirac Research AB and Uppsala University, Uppsala, Sweden）

ABSTRACT

This paper illustrates the main principles for loudspeaker compensation and compensation of the room acoustics that are used by Dirac Research and that have resulted in the technologies Dirac Live® for single-channel compensation and Dirac Unison for joint multi-channel compensation. A first main aim is to control not only the frequency domain properties of the system but also the time domain properties: The impulse responses as measured at different listening positions. In particular, we strive to reduce the "pre-ringings" (pre-echoes) that would otherwise result in an un-natural sound experience. Secondly, we use dynamic models of the sound system that are based on measurements at multiple listening positions. This is important for obtaining a robust design that works over an extended region to provide a large spatial area with good sound quality. Third, we may jointly optimize multiple loudspeakers to better control the sound pressures at different listening positions. This is done by precise phase control of the individual loudspeaker transfer functions at low frequencies. Joint optimization of a set of loudspeakers results in more distinct bass performance, better robustness of the compensation and better control of the impulse responses at different listening positions.

Keywords: Room compensation, robust audio precompensation, sound field control.
(ISEAT 2015, Shenzhen, Nov 14-15. English version, revised October 2, 2015)

[1] Email address: lars-johan.brannmark@dirac.se
[2] Email address: mikael.sternad@signal.uu.se

INTRODUCTION

A system that uses loudspeakers for sound reproduction will affect the quality of the reproduced sound in quite complicated ways. It is time-consuming, expensive and sometimes impossible to reduce undesired effects of listening rooms by changing the loudspeaker placement or by doing a physical re-design (room treatment).

Our aim is to counteract acoustic problems by model-based digital compensation, within the physical limits that are set by the sound reproduction system. We perform sound field control for audio reproduction by high-performance robust room correction algorithms. A compensation system in the form of a set of digital filters is placed between the sound source and the loudspeakers. These discrete time filters are to be designed and adjusted to reduce undesired effects. First, we have to obtain measurements of the acoustic properties of the sound reproduction system at a set of measurement positions in a desired listening space. We then construct a mathematical model of the reproduction system, based on the measured acoustic properties. The filters of the compensation system are then adjusted, based on the model, so that they provide an improved audio performance.

It is far from obvious how such compensation filters should be adjusted. What can realistically be done, and what should be done? At least three aspects have to be taken into account: Properties of the set of *frequency responses* (coloring of the sound) at the measurement positions, properties of the corresponding *time-domain impulse responses* (pre-ringings and echoes) and the variations of the perceived sound with respect to the *spatial location*.

First, a sound reproduction system amplifies different frequencies differently. This is partly due to interference between sound travelling via many different paths through the listening room. Therefore, the frequency response will depend on the listening position in the room. The variability of the amplification with position increases for increasing frequencies. Furthermore, music and speech are nonstationary signals. They do not only consist of sums of tones with constant amplitudes, but contain fast changes and often impulse-like variations. Therefore, time-domain properties of the sound reproduction system are important. Different sound paths have different lengths and therefore different delays due to the finite velocity of sound. A short impulse that is transmitted from a loudspeaker will therefore arrive at a listening position in the form of many copies (echoes) that arrive with different delays. Such a set of delayed impulses represents the *impulse response* of the sound reproduction system and the listening room. The time-domain properties represented by the impulse response will change when the listening positions in space is moved, because the time delays change with position. We will therefore have to work with a set of impulse responses, one for each measurement position.

We also have to take the properties of the human auditory system into account. The perception by individual listeners generates many subtle and quite complicated effects: Some sound properties are obvious to most listeners while others are perceived by only the most experienced listeners. Some properties are noticeable the first time we listen to a reproduction system, but then quickly fade into the background as we grow accustomed. Others are hardly noticeable at the beginning, but grow irritating as the time passes. Some frequency-domain and time-domain properties are masked (made non-perceivable) by the signal processing performed by our brain-stems and brains. For example, frequencies with high amplitude reduce our sensitivity to nearby frequencies (frequency masking). Sound impulses reduce our sensitivity to sound appearing just before the impulse (pre-masking) and just after the impulse (post-masking). Some time-domain properties may have small effects on single-channel sound but may be crucial for the perception of spatial properties of a multi-channel sound stage reproduction.

An ideally adjusted compensation system would give the sound reproduction system perfect frequency-domain and time-domain properties: All frequencies would be amplified according to a desired target frequency response, and this would be accomplished at all specified listening positions in the room. Furthermore, all delayed echoes would be precisely compensated so that they would arrive with equal delay. The impulse response of the compensated system would then become a Dirac Delta function (or Kronecker Delta function in discrete time), at all listening positions. This ideal case has inspired the name of our company, and we can come quite close to it in some interesting scenarios. However, we also need strategies for what to do and what to strive for when the ideal situation cannot be attained.

These questions about non-ideal situations have been crucial in the development of the designs that will be discussed here. We have developed design algorithms over the last decade that use a polynomial equations approach to feedforward controller design, tailored for audio systems [1],[2]. They represent a systematic way to design control algorithms, where physical and acoustical insight can be built into the models, and where both time- and frequency domain aspects are taken into account. The designs are based on extensive previous research on robust control and robust estimation and their properties can briefly be summarized as follows:

- We use linear mixed-phase filters that are designed to invert the dynamics of the sound reproduction system in a controlled and safe way. Specifically, they minimize weighted mean square errors at a specified set of listening positions (also referred to as control points or measurement points), under several constraints.
- The spatial distribution, extent and number of control points is a design aspect that may vary between different use cases. For example, they can be placed in a small region, a larger region or in several separated regions, forming one or multiple extended “sweet spots”.
- The properties of transfer functions in-between measurement points are taken into account in a statistical sense. This increases the robustness of the design and prevents over-compensation.
- The desired system (the target system) is specified as a set of impulse responses at the measurement positions. We thus strive to control the time-domain properties of the compensated system. This is important as precise time-domain control is crucial for a good spatial sound stage in stereo and other multi-channel reproduction.
- For a given target system defined at the measurement positions, the mean square errors to be minimized are obtained as squared differences between the target system and the actual system. The design can be performed under constraints of causality of the controller, constraints on loudspeaker signal amplifications, and constraints on the stability of the compensation filters. These filters may have arbitrarily long impulse responses and may be implemented as infinite impulse response (IIR) filters. The use of long impulse responses is important for obtaining good low-frequency properties.

- The design is also performed under constraints on the magnitudes of the impulse response coefficients before the main impulse (the "*pre-ringings*" of the impulse responses of the compensated system, see Section 3.1), at all listening positions [1]. Large pre-ringings might otherwise generate quite severe audible distortions.
- The multiple-input multiple-output (MIMO) transfer functions from multiple loudspeakers to multiple listening positions can be pre-compensated jointly by a filter bank. In such a joint design, the signals from different loudspeakers can be made to add in phase at the listening positions at low frequencies. This significantly reduces the variability of the compensated response over the listening regions, as compared to compensating each loudspeaker separately [2].
- A MIMO design can also optimally up-mix stereo or 5.1 coded material to the loudspeaker setting used, for arbitrary loudspeaker numbers and positions. For example, for each channel, one loudspeaker may act as main speaker, while all the others act as support loudspeakers [2]. See the right-hand part of Figure 1.
- The MIMO design can be used for many applications such as car audio systems, beamforming, design of multiple acoustic zones and active noise control [3].

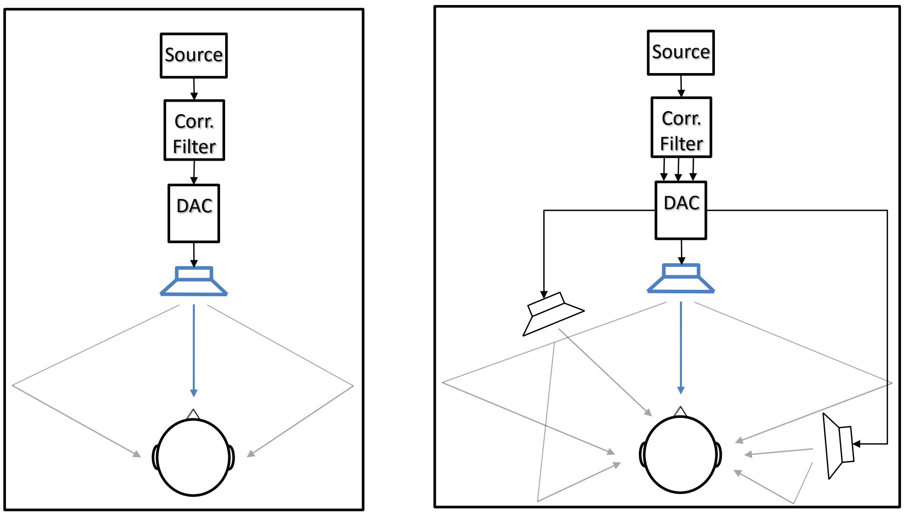

Figure 1: Digital single-channel compensation (left) and multi-channel (MIMO) compensation (right).

Designs for single-channel equalization and room correction (single-input multiple output systems, or SIMO systems) based on the theory in [1] will be discussed in Section 3 below. Joint compensation of transfer functions from multiple loudspeakers (MIMO systems), based on [2] will be discussed in Section 4. We first set these designs in context by briefly discussing some alternative approaches to loudspeaker and room compensation and to sound field control.

1. PREVIOUS WORK

1.1 Single-channel equalization and room correction

Single-channel loudspeaker equalization by the use of digital filters has been extensively studied since the mid-1980s. In a broad sense, the aim of all audio channel equalization schemes is to remove undesired linear (i.e., convolutional) distortions in the signal path of a sound system [4]. Previous work on equalization, and related robustness issues, essentially falls into three categories. In the first category, the filter design aims at a complete signal dereverberation (an inversion of the dynamic system) at a single position in a room. A subsequent robustness analysis then investigates equalizer performance at other spatial positions, or under slightly modified acoustical circumstances. It is well known that this kind of filter design is highly non-robust, causing severe signal degradation when the receiver position changes [5], and even for fixed receivers, due to the “weak nonstationarity” of the acoustical paths in the room [6].

In the second category, the design objective is not a full dereverberation, but rather a reduction of linear distortions, under the constraint that audio performance should not be too much degraded by changes of listener position. The standard approach in this category is to design a filter based on averaging, smoothing or fuzzy clustering of one or several transfer functions and then perform a robustness analysis of the filter [7], [8]. Although based on somewhat ad-hoc principles, some of these methods seem to perform reasonably well and many of them have been implemented in commercial products.

The third category imposes spatial robustness directly on the design by employing a multipoint error criterion to optimize sound reproduction in a number of spatial positions, either by using measured transfer functions at various spatial positions [9] or by direct adaptation of an inverse filter [10]. Our proposed single-channel approach in Section 3 belongs to this category.

1.2 Multi-channel sound field control

Three main approaches to sound field control by joint optimization of inputs to multiple loudspeakers have been explored during the last decades, all of which however suffer from important restrictions: Wave Field Synthesis [11, 12, 13] and High Order Ambisonics [14] are based on theoretical approximations of reality, for example circular loudspeaker positions and echo-free rooms. This is a major disadvantage, since such theoretical approximations are rarely fulfilled in practice. These designs are furthermore performed in the frequency domain and ignore the time-domain properties of the transfer functions.

Finally, we have multipoint Mean Square Error (MSE) designs, which are also in general performed as per-frequency designs [15], with some approaches using spherical harmonics [16], 17]. Our multichannel approach that will be illustrated in Section 4 belongs to the category of multipoint MSE methods.

2. MODELLING OF THE LOUDSPEAKER-ROOM SYSTEM

The steady-state response to sinusoidal input signals of a scalar, linear and time-invariant dynamic system is fully described by its *frequency response*. It is represented by two functions of the frequency, the *frequency magnitude response* and the *phase response*. The frequency magnitude response, or spectrum, is the main characteristic of the behavior of a scalar linear system. It describes the amplification from the source signal to the received signal as a function of frequency. The phase response describes the time delay of each frequency component.

However, as mentioned in Section 1, the time domain properties, represented by the impulse responses of the discrete-time transfer functions is also important. A set of impulse responses from a loudspeaker-room system is illustrated in Figure 2. Our model-based precompensator design is based on such sets of measured impulse responses, one for each loudspeaker-measurement point pair. They may be represented as finite impulse response (FIR) filters or as infinite impulse response (IIR) filters. The measurement points should be placed in the regions of space where compensation is desired. They should at least cover the space where a listeners head is typically placed, but we may use measurements from several regions or from one large extended region. If the spatial region is larger, then the differences between the measured impulse responses at different measurement points will be larger.

There is a unique correspondence between the frequency response (magnitude plus phase response) and the corresponding impulse response in the time domain, via the inverse discrete Fourier transform. There is no unique correspondence between the magnitude response only and the impulse response: The phase response also has to be taken into account.

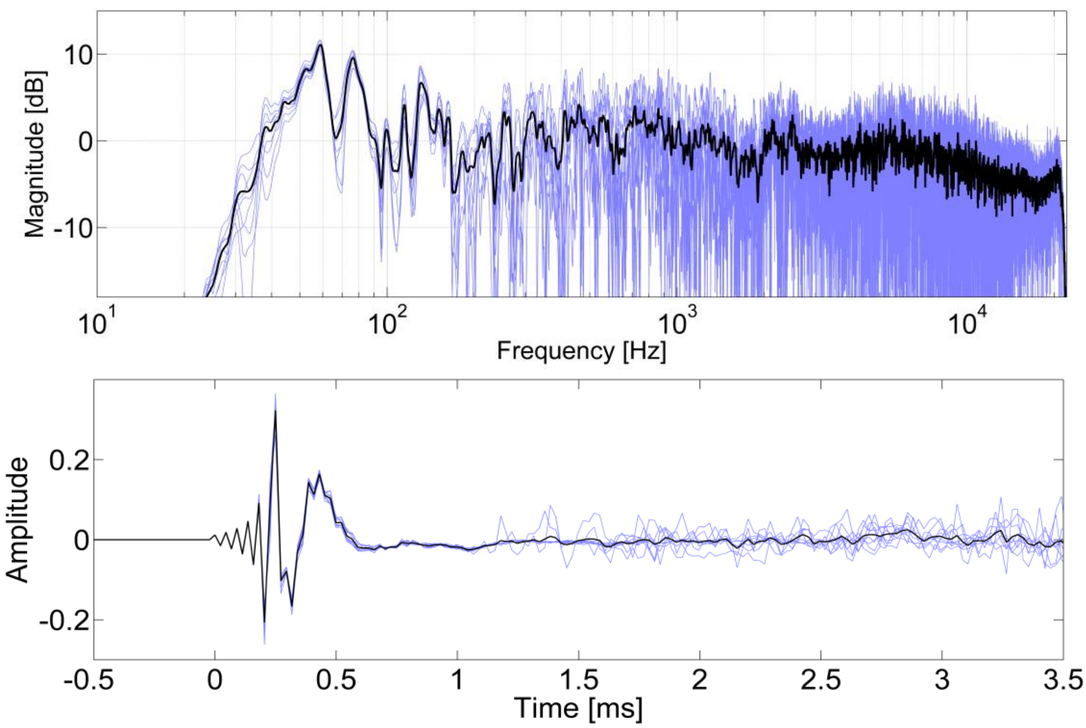

Figure 2: Upper figure: Set of frequency magnitude responses (blue) and the RMS average frequency magnitude response (black) from one loudspeaker to a set of measurement positions of a typical loudspeaker-room system. Note the large variability at higher frequencies. Lower figure: The corresponding impulse responses, with the average impulse response (dark blue). Note the large variability in the tails of the impulse responses, representing diffuse room reflections.

We use such linear time-invariant models to describe the properties of the sound reproduction system as well as the desired properties of the compensated system. The aim of room correction is represented by a target behavior of the compensated system. As mentioned in Section 1, it is typically given as a smooth frequency magnitude response and an impulse response that is concentrated in time, ideally a unit pulse. See Figure 3.

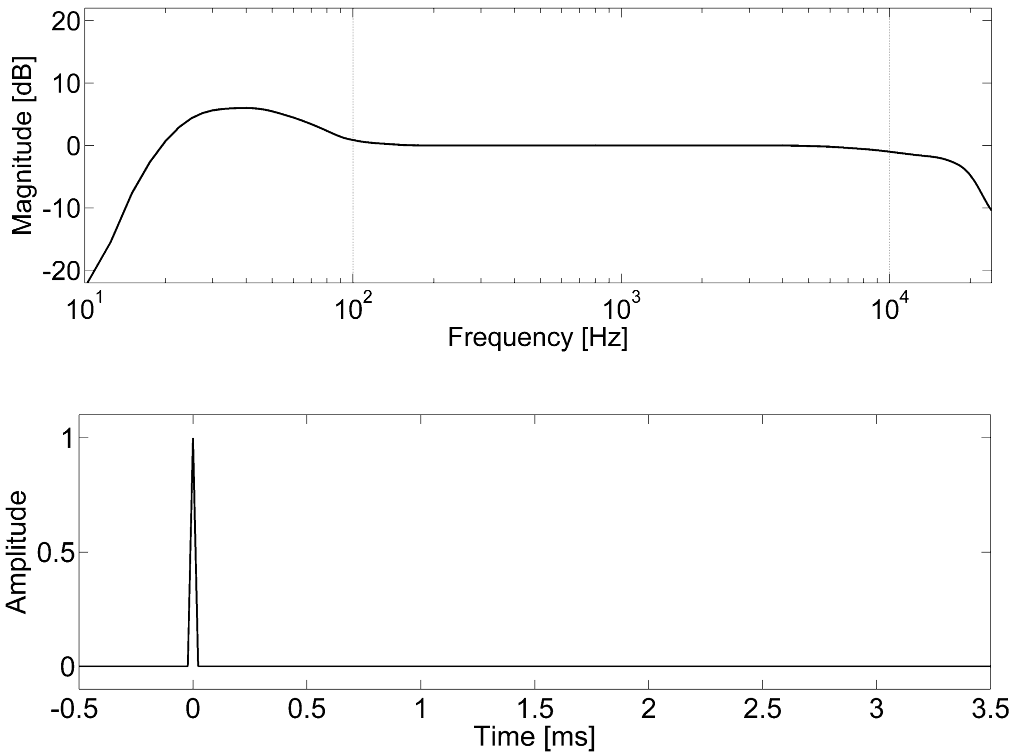

Figure 3: Example of desired target response in frequency and in time. The desired time response is often delayed (shifted to the right), since this makes it easier to approximately attain the target. The delay selected for the target is a design variable, see Section 3.2.

3. COMPENSATION OF SINGLE-LOUDSPEAKER SYSTEMS

3.1 Mixed phase systems and minimum phase representations

For a given spectrum, there exists a unique phase response that results in an impulse response with minimum energy delay for all frequencies. A system with this particular phase response is said to be a *minimum phase system*. All other impulse responses having the same spectrum (same frequency magnitude response) are of *mixed-phase* type.

The impulse response of a minimum phase system has a time envelope that is sharp and distinct at the beginning, and it then decays with increasing time as illustrated by the lowest plot in Figure 4. A minimum phase impulse response has no large bursts of energy (e.g., no strong echoes) in the late parts of the impulse response.

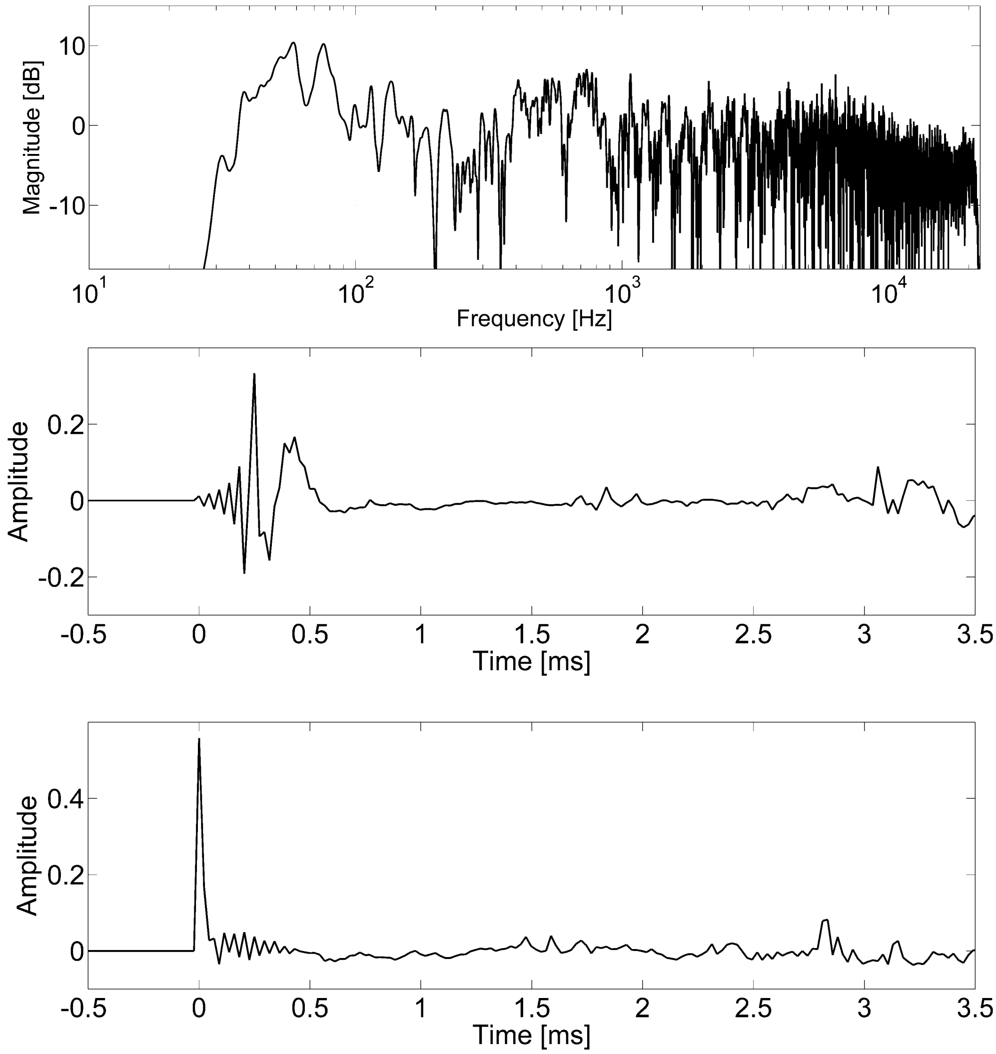

Figure 4: A frequency magnitude response (upper) together with its original measured impulse response (middle), and the impulse response of the minimum phase system that would correspond to this frequency magnitude response (lower).

Typically, loudspeaker–room responses are not minimum phase, they are of mixed phase type. This is the case e.g. for the original measured impulse response in Figure 4 (middle plot).

One particular type of mixed-phase responses has distinct pre- and post-ringing parts. Such responses have a main peak somewhere in the middle, with build-up before and decay after the main peak. Such impulse responses are often generated when a less than perfect pre-compensation filter is convolved with a loudspeaker-room impulse response. See Figure 5 for an example.

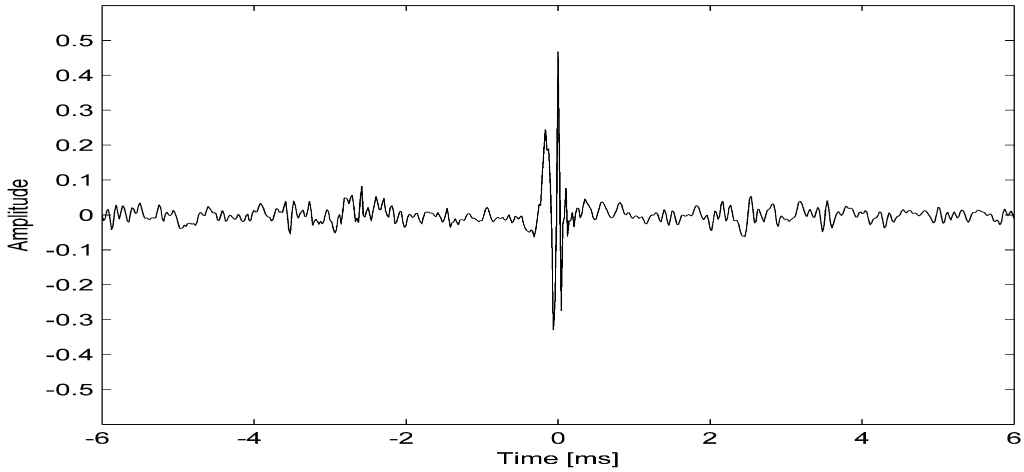

Figure 5: Example of impulse response with pre-ringings (to the left) and post-ringings (to the right).

Post-ringings are impulse response components after the main peak. They are ubiquitous in all sorts of acoustic systems, e.g., a loudspeaker in a room. They can severely distort the reproduced sound.

Pre-ringings are impulse response components before the main peak. Already at fairly low levels and of short duration, pre-ringings give an unnatural, artificial character to the sound and this is very easily detected by human hearing. Typically, pre-ringings of high level and long duration do not occur naturally in the impulse responses of physical systems. They are mostly artifacts introduced by digital processing, e.g., by lossy coders, linear-phase filters or by faulty room correction filters.

3.2 Correction filters for minimum phase and for mixed phase systems

The aim of digital loudspeaker- and room correction is to find a pre-compensation filter which, in some sense, inverts the dynamic response of a linear audio reproduction system. In the ideal case, if an impulse is transmitted through such a pre-compensator and the resulting output signal from the pre-compensator is transmitted through the sound reproduction system, then the end result would just become a reproduction of the original impulse at the measurement positions. An arbitrary signal, which can be modeled as a sequence of scaled impulses, would then be recreated error-free at the measurement positions.

The general theory for the attainable performance and the optimal design of such linear precompensation filters is closely connected to the theory of Wiener filters [18]. A linear pre-compensating inverse filter constitutes a linear feedforward servo controller in a control-theoretic framework. It can be designed as a Wiener filter (also called a linear quadratic optimal feedforward compensator) that minimizes the sums of mean square deviations between the sound pressures and the desired sound pressures at the measurement positions [1, 2, 3]. From Wiener theory, it follows that we may obtain better approximations of an ideal system inverse if we allow the desired output signal to be somewhat delayed. We therefore allow the desired target to be delayed by a specified number of sampling times. This “modeling delay” is specified as a property of the desired response and it is a design variable. Typically, it is selected so large that a further increase gives small improvements. If there are strict latency requirements on the application, then a large delay cannot be used.

For a given compensator, the resulting impulse response at one measurement position could perhaps look like the one in Figure 5, where the target delay has been set to the time zero in the plot. Some components of the signal will go through the system faster, and arrive earlier than the main pulse (to the left of it in the plot). They generate the precursors, or pre-ringings of the impulse response. Other signal components arrive later, and contribute to post-ringings in the later part of the impulse response. It is desirable to reduce the overall distortion in the system (which in an uncompensated system is represented by large post-ringings). But this must not be done at the expense of introducing any audible pre-ringings in the corrected system.

Consider a scalar, linear and stable discrete-time system, which describes the transfer function from the loudspeaker input to the sound pressure at one measurement position. If this dynamic system is *minimum phase*, then there exists a unique and causal exact inverse system. See the upper row of figure 6. This inverse will be stable, and it will also be minimum phase. In this case there exists an exact and causal inverse filter, so we do not need to introduce any modeling delay. Therefore, no pre-ringings will appear in the impulse response of the compensated system. If, furthermore, the model of the system is perfect, then the inverse is perfect so there will also appear no post-ringings.

If the system is *mixed phase*, as is mostly the case, there exists no inverse that is both causal and stable. But in that case, an approximate Wiener inverse can be constructed if we allow a modeling delay, see [3]. The attainable quality of the compensation improves with an increasing allowed delay, up to a range where further increased delays provide increasingly small gains.

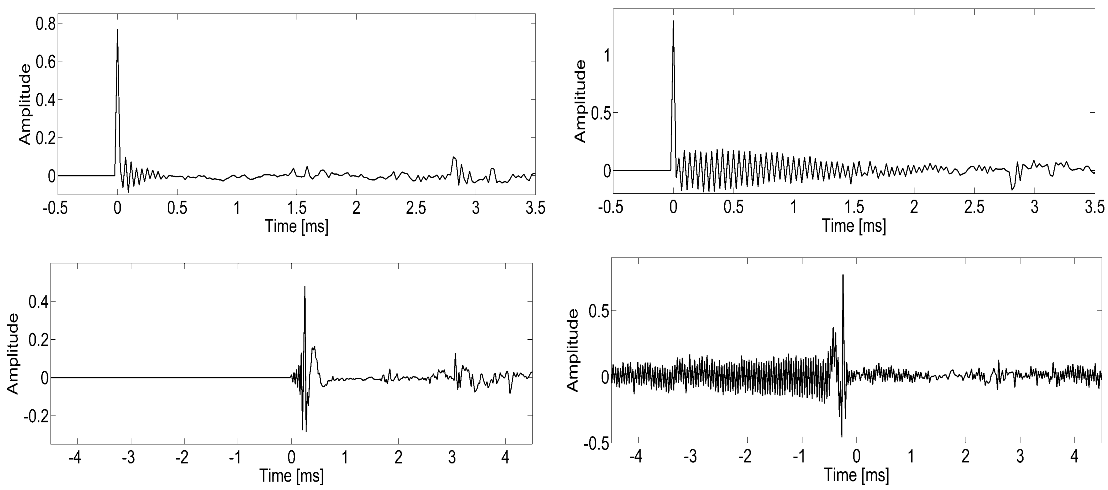

Figure 6: Upper row: Impulse response of a scalar minimum phase system (left) and the impulse response of the corresponding (exact) inverse filter (right). Lower row: Impulse response of a mixed phase system (left) and the impulse response of an approximate inverse filter designed for a modeling delay of 4.5 ms. The plots of all impulse responses have been truncated to the right.

3.3 Single point designs, multipoint designs and robust mixed phase compensation

As illustrated by the lower right-hand plot in Figure 6, it is possible to design a stable Wiener feedforward compensator that would approximately invert a mixed phase scalar system if we allow a modeling delay in the target impulse response. However, such a design will suffer from several limitations:

- Perfect inversion in general requires unrealistically high gains in the inverse filters at some frequencies. This is evident by the large high-frequency oscillations in the impulse responses of the compensation filters (right-hand plots) in both examples in Figure 6.
- The system response varies with spatial position, so an exact inverse designed for one position will not be valid at other positions. Far away from the design position, the compensation may actually destroy the sound properties rather than improve them. This phenomenon is called *over-fitting* to a limited set of measurements.

Rather than perfect inversion in one point in space, we should therefore strive for as good a behavior as possible over an extended spatial region (i.e. at multiple listener positions), under a set of physical and psycho-acoustical constraints on the filter design.

The solution we use is based on multipoint design, using several measurement positions (control points). A criterion to be optimized is defined over a spatial region instead of in one point only. For example, we may minimize the sum over several measurement positions of the mean squared deviations (the MSEs) of the actual system response from a defined ideal response.

A straightforward unconstrained MSE minimization leads to the so-called noncausal Wiener filter as the solution. This solution is obtained under the assumption that an infinitely large modeling delay can be used. Then, the total MSE can be reduced substantially, but the remaining errors appear not only as post-ringings but also as sometimes large pre-ringings.

A better solution is to compute the optimal precompensation under a causality constraint, i.e. to allow a given fixed modeling delay. By holding this modeling delay low, the pre-ringings of the impulse response can be made short in time, so that they are masked by the pre-masking effects of the human auditory system. However, this forces us into compromises with the modeling delay that can sometimes reduce the attainable MSE performance.

We instead use a robust MSE minimization: A design criterion is used that is affected by the sound at all measurement positions. It is minimized under a constraint on the pre-ringings that are allowed at any of the control points.

3.4 Summary and a performance comparison

The discussion above leads to several possible alternatives for the design of precompensation filters. Let us summarize their properties and compare the performances of some of them.

- Single-point methods, which take the response at only one control point in space into account, are non-robust. They will not be considered further below.
- **Minimum phase multipoint design**: We may represent the average response at a set of control points by an average frequency domain function, such as the black line in the upper part of Figure 2. We can then construct the corresponding minimum phase impulse response (as in the lower part of Figure 4 and the upper left part of Figure 6).

This impulse response in general differs from the measured impulse response at any of the control points, since the systems are mostly mixed phase. Still, we may use the inverse of the minimum phase model, and utilize this causal and stable inverse filter (illustrated by the upper right-hand part of figure 6). Such a design will be good in an average sense in the frequency domain. The resulting design uses no modeling delay so the compensated impulse response will have no pre-ringings. However, other time-domain properties are not targeted and remaining post-ringings in the impulse responses may be substantial.

- **Unconstrained mixed phase multipoint design (a noncausal Wiener solution).** This solution takes the MSEs at all control points into account and minimizes a weighed average of them. This generates a noncausal precompensator that therefore requires the use of a very large modeling delay. The resulting performance is often good in an average sense in the frequency domain, but this design does not care about how the remaining errors are distributed over time in the compensated impulse responses.
- **Constrained, robust mixed phase multipoint design.** This solution takes the MSEs at all control points into account and minimizes a weighted average of them under constraints. The constraints may include constraints on loudspeaker powers at different frequencies, a constrained modeling delay and constraints on the allowable pre-ringings of the compensated impulse responses at all control points in the targeted area. Please see [1] for a detailed discussion. The resulting compensators are stable and causal IIR filters, with as long impulse responses as needed. The performance of this strategy is good in an average sense in the frequency domain, and it also corrects time-domain properties. It provides a slightly worse MSE than an unconstrained Wiener solution, but the sound quality is much better from a psychoacoustic perspective.

Let us illustrate the performance of the last three designs (minimum phase, unconstrained mixed phase and constrained robust mixed phase) on the acoustic system with 9 measurement positions (control points) that has been illustrated by Figure 2. In all cases, the properties at all measurement positions are equally weighted when performing the designs.

The frequency magnitude responses of the resulting precompensation filters are illustrated in Figure 7. None of the designed precompensation filters require excessive filter gains. The minimum phase (upper) and the constrained mixed phase (middle) compensators do in this case have almost the same frequency magnitude response. They should therefore provide almost equal compensation of the frequency magnitude response. The resulting performances of the compensated systems are illustrated by Figure 8, both in the frequency domain (left) and in the time domain (right). The minimum phase design equalizes the average frequency magnitude response well, but it leaves the time domain properties almost unchanged compared to the original impulse responses (cf. Figure 2). The unconstrained mixed phase design reduces the post-ringings of the impulse responses but generates significant pre-ringings. The robust mixed phase design controls both pre-ringings and post-ringings in the time domain and it also provides a good average performance in the frequency domain.

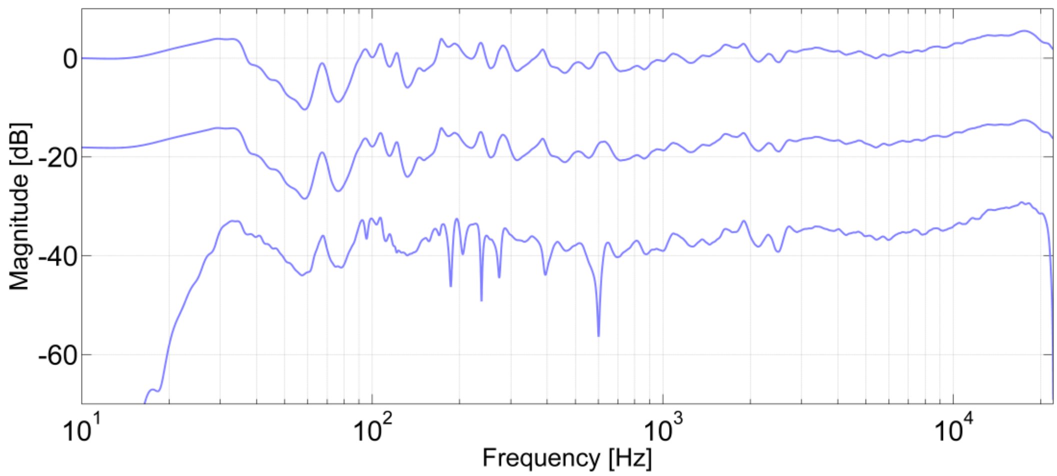

Figure 7: Frequency magnitude responses of scalar linear precompensation filters that are obtained for the sound reproduction system illustrated by Figure 2. Upper: Minimum phase multipoint design. Middle: Constrained robust mixed phase design. Lower: Unconstrained (noncausal Wiener) mixed phase design. Note that the curves have been shifted vertically to improve the clarity of the plot.

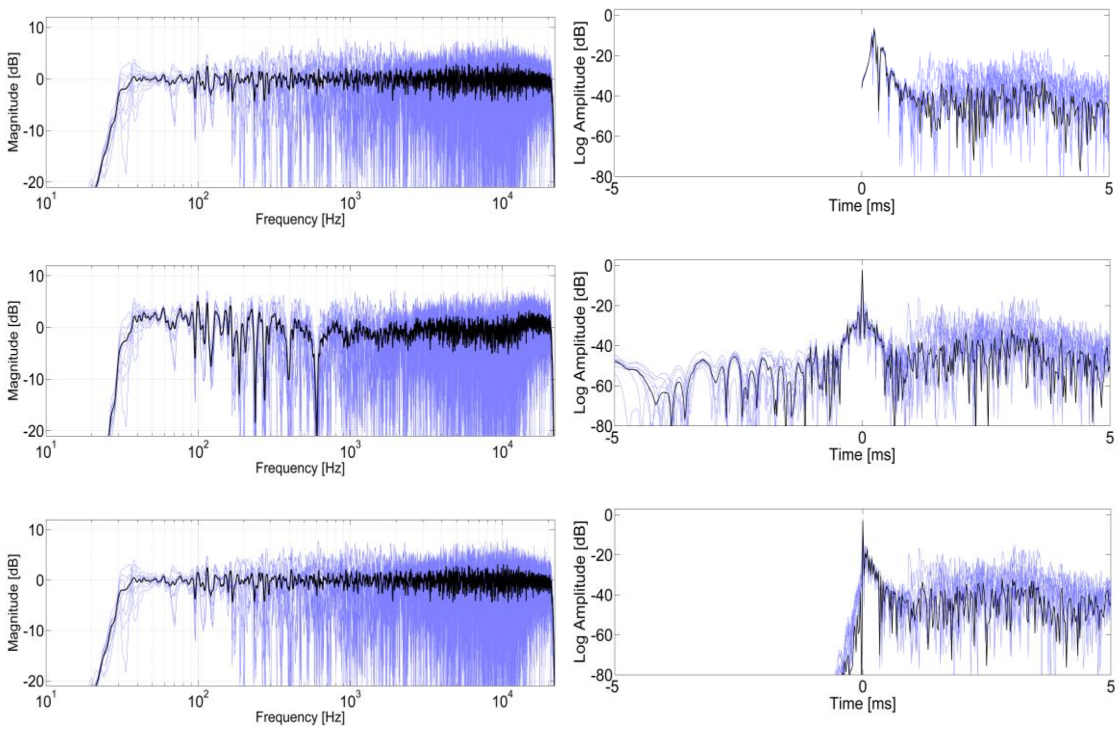

Figure 8: Simulated frequency magnitude responses (left) and impulse response amplitudes in logarithmic scale (right), in the 9 measurement positions of the system illustrated in Figure 2, when it is compensated by the three compensator designs illustrated by Figure 7. Upper: Minimum phase multipoint design. Middle: Unconstrained (noncausal Wiener) mixed phase design. Lower: Constrained robust mixed phase design, performed under pre-ringing constraints. (This design corresponds to the middle precompensator gain plot in Figure 7).

4. MULTICHANNEL COMPENSATION

Single-channel correction can improve performance in an average sense over a spatial domain. In particular, time domain and frequency domain properties that are common to all listener positions can be corrected. For example, broadband direct sound and low-frequency room effects will be fairly common over an extended sweet spot, and can therefore be compensated well. However, the spatial variability is not reduced by single-channel correction. This can be seen by comparing the variability in the frequency domain after compensation in Figure 8 to the variability that was present in the original system in Figure 2.

A sound reproduction system in general uses more than one loudspeaker, and we have so far discussed compensation of each loudspeaker individually. Such a design will improve the system but it leaves several problems unsolved. For example, crossover optimization (time-alignment of spatially separated tweeter, midrange, and woofer drivers) is not straightforward with single-channel methods and it requires separate manual tuning.

The stereo image is often greatly improved after performing single-channel correction, but this aspect is not explicitly taken into account in the criterion functions used by the single-channel designs discussed in chapter 3. It would be desirable to extend the single-channel precompensator design to be able to simultaneously optimize the whole multichannel system. This has been done in [2] for the case of robust mixed-phase multipoint design. The cautious, pre-ringing constrained solution of [1] has here been extended to the compensation of MIMO systems. Loudspeaker correction is performed using the same target as in the single-channel method, but all available loudspeakers are used to come closer to the target. As a result, the spatial variations are reduced, and crossover/driver alignment is to a greater extent automatized. For example, if we have two speaker elements that work in different frequency ranges and are also at different positions, then the resulting design will contain an optimized crossover filter.
The use of a criterion of pairwise similarity between left and right stereo channels further helps to improve sound stage imaging [19]. Furthermore, a MIMO design can form the basis of a unified solution to the problems of equalizer design, crossover design, delay and level calibration, sum-response optimization and up-mixing (i.e. routing 2-channel or 5.1 channel source material) to N loudspeaker outputs in a car audio system [21]. It can be used to give the listener the experience to be in another listening space, with different room acoustics and differently placed loudspeakers, as compared to the physical listening space [21].
The MIMO design approach also applies to “personal audio” applications, i.e., acoustic zones, and to active noise control [3], where the use of multiple control loudspeakers can significantly enlarge the zone of silence [20].

Figure 9 below illustrates the result of robust MIMO design in the form of frequency magnitude plots, based on a set of measured channels to 64 control points. Note the decrease of the variability of the compensated transfer functions in the low-frequency region as the number of utilized loudspeakers is increased. Figure 10 illustrates the combined time-domain and frequency domain properties of the same compensated system. The “waterfall plots” illustrate the decay of different frequency components of an impulse. A significant tightening of the bass response can be noted as the number of co-optimized loudspeakers increases.

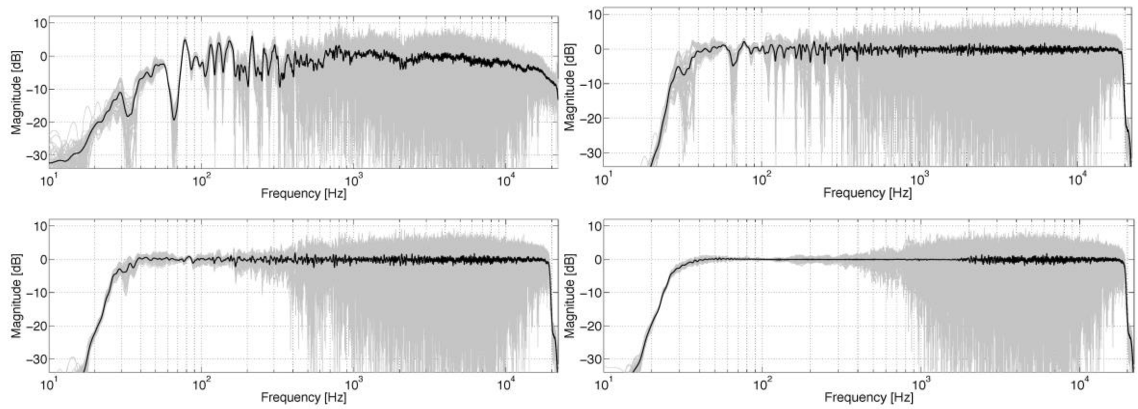

Figure 9: Simulated frequency magnitude plots, for one uncompensated loudspeaker and with robust mixed-phase compensation of this speaker (upper row). Performance of robust mixed-phase MIMO compensation that uses six and 16 loudspeakers (lower row), all from [2]. Gray lines reflect the variations in a cubic grid of 64 measurement points with 3 dm sides. Black lines are the RMS average responses.

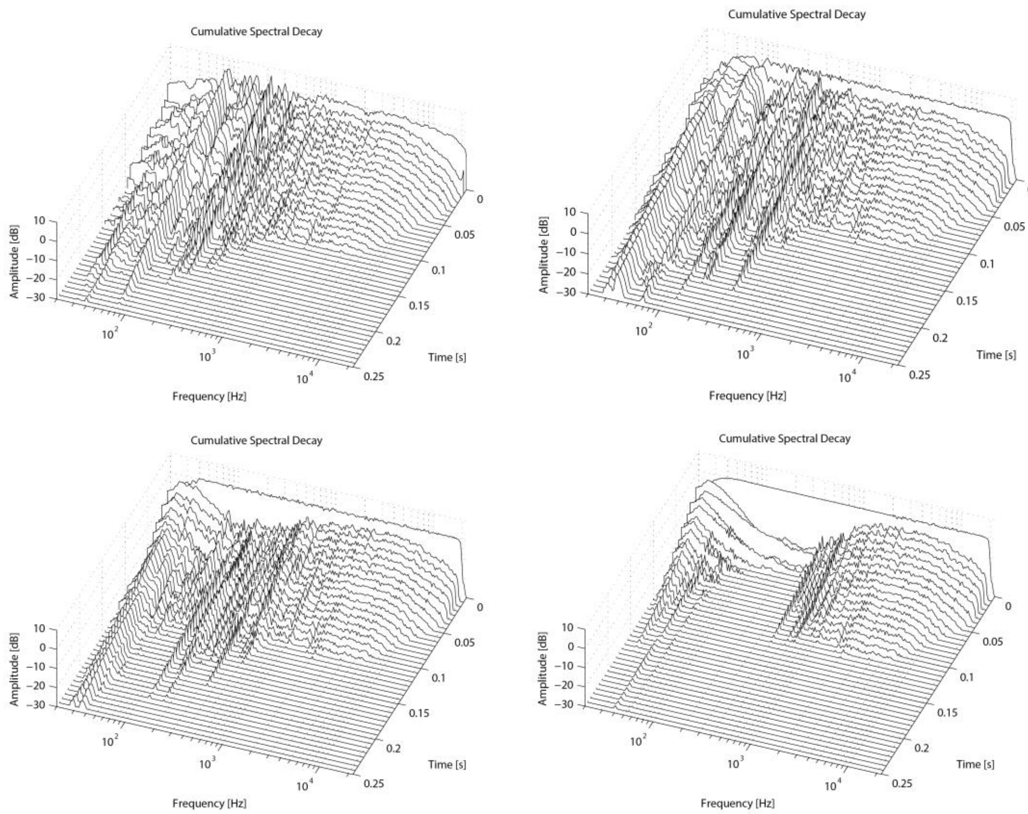

Figure 10: Waterfall plots illustrating the cumulative spectral decay of one loudspeaker (upper left), evaluated in a cubic grid of 64 measurement points with side 3 dm. Upper right; The result after single-loudspeaker compensation. Lower: The results of MIMO compensation with 6 loudspeakers (a 5.1 system, left), and with 16 loudspeakers (a 14.2 system, right). From [2]. Same designs are used as in Figure 9.

5. CONCLUSIONS

We have here illustrated the challenges of loudspeaker equalization and room response correction and have outlined some solutions to these challenges. While the physical properties of the loudspeakers and the listening room will place fundamental limits on what can be done, digital signal processing can provide large improvements within these limits. However, to perform compensation successfully in difficult cases requires the use of a nontrivial set of design aims, methods and algorithms that simultaneously takes multiple aspects into account. In particular, we should simultaneously consider the frequency response, the time domain properties and the variability of these properties over an extended listening area. We have shown how robust mixed phase designs can do this successfully. We have furthermore illustrated that a joint co-design of digital precompensators of all loudspeakers in the reproduction system is quite powerful. It can not only improve the average mean square error performance but also reduce the variability of the acoustic transfer functions within the listening area.

The mixed–phase single-channel designs as well as the multi-channel designs that have been illustrated here have been introduced in successful commercial products.[3] We expect them to appear in many new advanced products and applications in the near future.

REFERENCES

[1] L-J. Brännmark and A. Ahlén, "Spatially robust audio compensation based on SIMO feedforward control," *IEEE Transactions on Signal Processing*, vol. 57, no. 5, pp. 1689-1702, May 2009.

[2] L-J Brännmark, A. Bahne and A. Ahlén, "Compensation of loudspeaker-room responser in a robust MIMO control framework," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 21, no. 6, pp. 1201-1216, June 2013.

[3] A. Barkefors, M. Sternad and L-J. Brännmark, "Design and analysis of linear quadratic Gaussian feedforward controllers for active noise control," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 22, no. 12, pp. 1777-1791, December 2014.

[4] M. Karjalainen, T. Paatero, J. Mourjopoulos and P. Hatziantoniou, "About room response equalization and dereverberation," Proc. of *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, WASPAA´05, pp. 183-186, New Palz, NY, October 2005.

[3] The single-channel design is used in the technology ***Dirac Live***. It can e.g. be used for room compensation by a virtual sound card for those who listen to music via a loudspeaker system with a computer as main sound source. See www.dirac.se, where the software for design and compensation can be obtained. Dirac Live precompensation filters can also be downloaded into hardware processors by miniDSP, see [22] and by Emotiva [25]. More expensive solutions that integrate Dirac Live room compensation are offered by the Casablanca IV processor by Theta Digital [23] and by the RS20i system by Datasat Digital Entertainment [24]. Dirac Live is also used by BMW, Rolls Royce, Bentley, and by Volvo, among others, in high-end audio systems for cars. The MIMO design, ***Dirac Unison***, is used in the Bowers & Wilkins sound system for the newly released Volvo XC90. Earlier versions (***Dirac Dimensions***) are used in sound systems by Bentley and by BMW.

[5] B.D. Radlovic, R.C. Williamson and R.A. Kennedy, "Equalization in an acoustic reverberant environment: Robustness results," *IEEE Transactions on Speech and Audio Processing*, vol. 8, no. 3. pp. 311-319, May 2000.
[6] P. Hatziantoniou and J. Mourjopoulos, "Errors in real-time room acoustic dereverberation," *Journal of the Audio Engineering Society*, vol. 52, no. 9, pp 883-899, Sept. 2004.
[7] S. Bharitkar and C. Kyriakakis, "A cluster centroid method for room response equalization at multiple locations," Proc. of *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, WASPAA´01, pp. 55-58, New Palz, NY, October 2001.
[8] P. Hatziantoniou and J. Mourjopoulos, "Real-time room equalization based on complex smoothing: Robustness results," *Proc. 116 AES Convention*, Berlin, Germany, May 8-11 2004.
[9] R. Wilson, "Equalization of loudspeaker drive units considering both on-and off axis responses," *Journal of the Audio Engineering Society*, vol. 39, no. 3, pp. 127-139, 1991.
[10] S.J. Elliott and P.A. Nelson, "Multiple-point equalization in a room using adaptive digital filters," *Journal of the Audio Engineering Society*, vol. 37, no. 11, pp. 899-907, 1989.
[11] A.J. Berkhout, D. de Vries and P. Vogel, "Acoustic control by wave field synthesis," *Journal of the Acoustic Society of America*, vol. 93, no. 5 pp. 2764-2778, May 1993.
[12] E. Corteel, "Equalization in an extended area using multichannel equalization and wave field synthesis," *Journal of the Audio Engineering Society*, vol. 54, no. 12, pp. 1140-1161, December 2006.
[13] S. Spors, R. Rabenstein and J. Ahrens, "The theory of wave field synthesis revisited," *Proc. AES 124th Convention*, Preprint 7358, Amsterdam, May 2008.
[14] M.A. Poletti, "Three-dimensional surround sound systems based on spherical harmonics," *Journal of the Audio Engineering Society*, vol. 53, no. 11, pp. 1004-1025, November 2005.
[15] O. Kirkeby and P.A. Nelson, "Reproduction of plane wave sound fields," *Journal of the Acoustical Society of America*, vol. 94, no. 5, pp. 2992-3000, November 1993.
[16] Z. Wang and S.F. Wu, "Helmholtz equation – Least-squares method for reconstructing the acoustic pressure field," *Journal of the Acoustical Society of America*, vol. 102, no. 4, pp. 2020-2032, October 1997.
[17] T. Betlehem and T.D. Abhayapala, "Theory and design of sound field reproduction in reverberant rooms," *Journal of the Acoustical Society of America*, vol. 117, no. 4, pt. 1, pp. 2100-2111, April 2005.
[18] H.W Bode and C.E. Shannon, "A simplified derivation of linear least squares smoothing and prediction theory," *Proceedings of the IRE*, vol. 38, pp. 417-425, April 1950.
[19] A. Bahne, L-J. Brännmark and A. Ahlén, "Symmetric loudspeaker-room equalization utilizing a pairwise channel similarity criterion," *IEEE Transactions on Signal Processing*, vol. 61, no. 24, pp. 6276-6290, December 2013.
[20] S. Berthilsson, A. Barkefors and M. Sternad, "MIMO design of active noise controllers for car interiors: Extending the silenced region at higher frequencies," *American Control Conference (ACC)*, Montreal, Canada, June 2012.

[21] M. Johansson, L-J. Brännmark, A. Bahne and M. Sternad, "Sound field control using a limited number of loudspeakers," *AES 36th International Conference*, Dearborn, Michigan, USA, June 2009.

[22] miniDSP Dirac Live Series: http://www.minidsp.com/products/dirac-series

[23] The Casablanca IV: http://www.thetadigital.com/casablanca_iv_controller_info.shtml

[24] Datasat RS20i: http://www.datasatdigital.com/consumer/products/rs20i.php

[25] Emotiva: https://emotiva.com/products/pres-and-pros/accessories/dirac-live-full-emotiva