Improving Particular Components of the Audio Signal Chain: Optimising Listening in the Control Room

A contextual statement submitted to Middlesex University in partial fulfilment of the requirements for the degree of Doctor of Philosophy by Public Works

Andreas Friesecke
Student ID 9935272

School of Media and Performing Arts
Middlesex University

October 2021

Abstract

"Improving Particular Components of the Audio Signal Chain: Optimising Listening in the Control Room"

Andreas Friesecke

In the field of audio engineering there is a constant need for optimising the listening situation. Listening to, judging and finally optimising the recorded material are essential tasks of audio engineers. The author of this contextual statement has been working in the field of audio engineering since 1993. In addition, various research projects have been undertaken in this field. A selection of three research areas and their published outputs are presented in this contextual statement:

- Research Area 1: Improving acoustic modules to increase efficiency in the acoustical treatment of control rooms
- Research Area 2: Measuring time alignment errors, testing their impact on the listening experience and providing solutions for time alignment of loudspeakers
- Research Area 3: Using equalisation for correcting and shaping a loudspeaker's frequency response

These research areas relate to a consistent listening 'defect' that leads to a blurred and broader sound image. Measures to overcome these defects are presented and proven to be effective by built prototypes and/or products. The results of the research are published in articles and books and can be experienced in the form of hardware systems such as acoustic modules or modified loudspeakers.

Keywords: Audio Engineering, Signal Optimisation, Sound Localisation, Listening Experience, Acoustical Modules, Time Alignment, Equalisation

Acknowledgements

I would like to thank everyone who worked with me in conjunction with my research, especially my former and current colleagues who critically questioned but always supported my ideas. I thank the SAE Institute as my employer who gave me the freedom to undertake the research and facilitated most of my presentations.

I thank Dr. John Dack for being a very supportive supervisor and enduring reader of different drafts of this contextual statement.

I thank Prof. Dr. Benjamin Dwyer assisting me with my academic writing.

I thank Carolina Löhle for giving critical feedback on a document that is not at all in her specialist area.

I thank my wife Susanne for always being patient with me and my research and letting me undertake it in great freedom.

Table of contents:

List of figures:

List of tables:

Statement of original authorship

Declaration

The submitted publications and the contextual statement are all the product of original work by the candidate. None of the submitted publications has been or is being submitted for any other degree of a University or other degree-awarding body.

Hohenbrunn, 10.10.2021 Andreas Friesecke

Wordcount: **25,900**

1 Introduction

The general concept of audio production is to record, store and reproduce audio signals in a way that the original sound experience can be conserved and transported to a different location. The audio signal chain is the route a signal has to travel until it can be heard by the listener (see Figure 1).

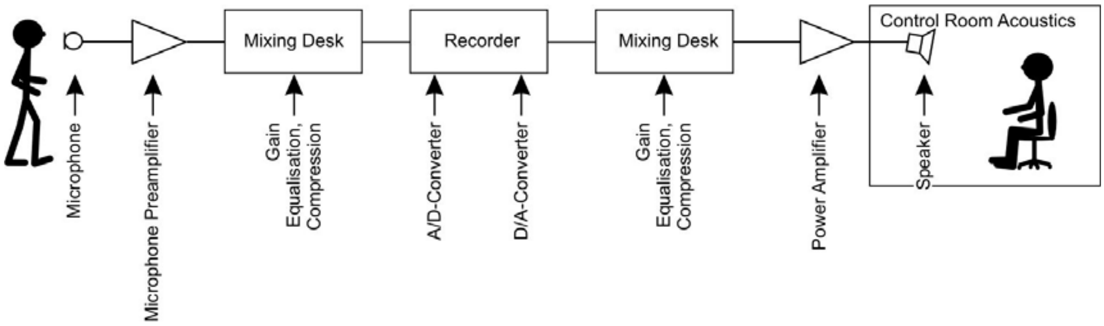


Figure 1: The audio signal chain from the microphone to the listening position (author's translation) (Translated from Friesecke, 2014, p.39).

The audio signal chain starts at the microphone, converting acoustic signals into electrical ones. Different stages of sound manipulation may follow before the signal is stored. The signal which is subsequently reproduced (for example, in a control room) is then played back via loudspeakers that finally transform the electrical signal back into an acoustic one. Under ideal conditions the signal will then reappear in a way that is very similar to the original signal that was captured, for example, in an auditorium of a concert hall. During this process the audio signal may become unintentionally distorted. This might happen for a number of reasons: the introduction of irregularities in the frequency response, adding delayed versions of the signal (such as echoes or room reflections), or deforming the waveform by non-linear transfer functions of certain electronic components.

A critical stage of the signal chain is at the final moment where signals are converted from an electrical signal to an acoustical signal and then being emitted to the listening position. Problematic changes in sound reproduction could come from the loudspeakers and the control room which can all have certain 'defects' contributing to a signal that is then less easily evaluable. But the better the listening situation, the more accurately the signals can be assessed. It is therefore desirable to have as few defects as possible at this final stage.

1.1 Research problem

The research question that arises is: "What measures can be taken to improve the listening situation in the control room?"

The author frequently observed one fact that was always associated with defects in audio reproduction: there was a noticeable reduction in the sharpness of localisation and/or depth of the sound image. Different problems could cause this effect but they are all related to timing issues of the signal. The issues could arise from reflections of the signal in a control room, from timing errors in multi-way loudspeakers or from frequency-dependent delays (that is, phase[1]) of loudspeaker-drivers and loudspeaker-crossovers. Research was undertaken to reduce critical timing errors by improving measures for room acoustic treatment, correcting time-delays of multiple drivers in multi-way loudspeakers and, finally, by equalising loudspeakers. As all these areas refer to the end of the audio signal chain the title of this contextual statement is 'Improving Particular Components of the Audio Signal Chain: Optimising Listening in the Control Room'.

1.2 Research areas

The author of this contextual statement has had a more than 25-year-long professional career in the field of audio engineering; and he has executed extensive research in the field discussed. The outcomes of this research have been published in the author's books and articles, and further presented in a number of conferences.

For the presentation of the author's research in this contextual statement, more substantial outputs have been selected that form a coherent thread. As mentioned, they all focus on improvements at the end of the audio signal chain, particularly loudspeakers and room acoustics. In this context, complex problems that affect acoustics and psychoacoustics are presented. In particular, the research addresses three main areas:

---

[1] Phase is the difference between two sinusoidal signals presented as an angle. It could be converted to a 'group delay' (or simpler 'delay') which is presented in seconds. Shifting signals by a certain phase angle (e.g. 90 degrees) means that the signal is time-delayed differently at different frequencies (according to the example 90 degrees correspond to a quarter of the cycle duration which is 2.5ms at 100Hz, 1.25ms at 200Hz etc.).
A special case is the so-called 'linear phase' which means that the phase shift is proportional to the frequency (e.g. 10 degrees at 100Hz, 20 degrees at 200Hz and 30 degrees at 300Hz etc.). This results in a constant time delay as 10 degrees at 100Hz (278µs) equals 20 degrees at 200Hz (278µs).

- **Research Area 1: Improving acoustic modules to increase efficiency in the acoustical treatment of control rooms:** This research focused on the improvement of room acoustics. New measurement techniques and innovative acoustical modules for controlling room acoustics were developed. The measurement techniques enabled faster and more reliable results for acoustic measurements. This fed into the research question as the newly developed modules helped to improve the listening experience by reducing reflections or scattering soundwaves more efficiently. At the time of this particular research (in 1994) no such modules were available and thus their development made a contribution to the knowledge of that time.
- **Research Area 2: Measuring time alignment errors, testing their impact on the listening experience and providing solutions for time alignment of loudspeakers:** This research investigated timing errors caused by phase shifts in the crossover and by non-aligned positions of the loudspeaker drivers on the mounting plane in relation to the listening position. Measuring these timing errors reliably and testing their impact was the first part of this research task. As these errors led to a blurred sound image, mechanical and electrical solutions for correcting these timing errors were developed. Further, an explanation about the cause of the audible effect was investigated. This research is significant for the research question because it significantly improved the listening situation: time-aligned loudspeakers' phantom sound sources are more precise, and depth imaging increases. At the time of the research (in 1994) there were no published articles concerning this topic in Germany.
- **Research Area 3: Using equalisation for correcting and shaping a loudspeaker's frequency response:** This area investigated the interrelation between mechanical spring-mass systems (i.e., loudspeakers) and electrical resonators (i.e., equalisers) which can correct each other. The investigation focused particularly on the phase shifts of loudspeakers and equalisers, and this led to significantly improved results in the correction and shaping of the frequency responses of loudspeakers. These results directly fed into the research question and provided two options to improve the listening experience: the loudspeaker could either reproduce a more neutral and analytical sound or it is optimised to identify more efficiently problems in the mixing process.

The research in all three areas contributes to answering the research question by optimising acoustic and electroacoustic aspects of the signal chain.

1.3 Overview of published works

This contextual statement provides an insight into the published works and the circumstances in which they were created. For an overview of the published works a timeline is presented in Figure 2. More detailed information about the published works in context with the historical development in the field of research is shown in Appendix A on page 93. The first thorough research started in 1993 and its findings were published in early 1994.

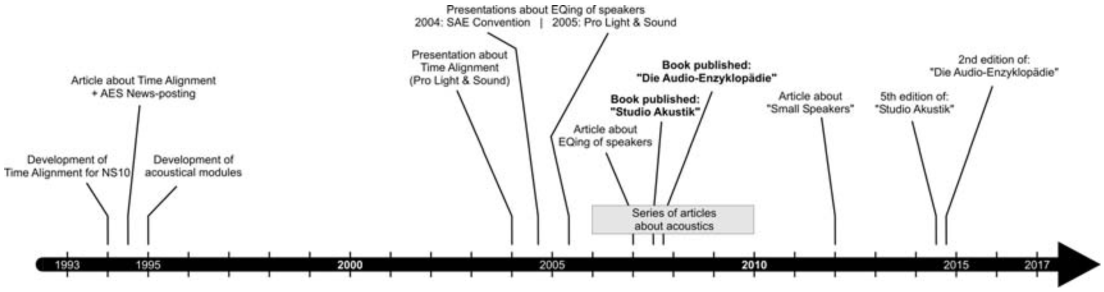


Figure 2: Coarse timeline of the presented research.

There were two different types of publication:

- Five journal articles
- Two books

In addition, conference presentations were given which fed into the published articles and books. These presentations provided the opportunity to present the research findings to the public and obtain feedback before they were published as articles or books. In terms of new knowledge, all these publications had a significant impact in the field at the time they were published because they all revealed information that was not published before. They are listed here in order of their publication date:

[1] ***Article: "Phase Lifting" Magazine: "Studio Magazin", 05/1994; Studio Presse Verlag GmbH***

This article discusses the impact of time-delayed signal arrival from the different drivers of a two-way loudspeaker. It shows the cause, the audible effect, a possible way to time-align the drivers and how to measure the necessary delay for correction. It further presents an explanation about how and why an alignment error becomes audible. Original findings from the author's work about aligning loudspeakers are presented and lead to an original contribution to the field. This article relates to Research Area 2 presented in this contextual statement and demonstrates the correction of a passive studio loudspeaker.

[2] ***Series of articles about acoustics: Magazine: "Recording Magazin", 2006-2010; Publisher PPV Medien***

This series of seven articles focuses on different aspects of studio acoustics throughout seven issues of the magazine. From this series, three are chosen for this contextual statement as they contain findings from the author's research:

1: "Wie trocken darf's denn sein?" (Reverberation) 05/2006
In this article, a new and reliable way of determining the reverberation time is demonstrated in form of measurement diagrams and new possibilities for evaluating measurement data in graphical form.

2: "Wenn der Schall in die Falle geht" (Absorbers) 06/2006
This article presents information about the impact of foil as a covering material for absorbers. It can act as a resonator and can have an impact on the high-frequency efficiency of the absorber.

3: "Schall und Gegenschall“ (Helmholtz resonators) 01/2007
In this article, a new and reliable method for measuring Helmholtz resonators is demonstrated. It further presents the impact of a flow resistance to widen the Q-factor of a Helmholtz resonator.

These articles relate to the Research Area 1 of this contextual statement.

[3] ***Article: "EQ or not EQ - das ist die Frage!" Magazine: "Recording Magazin", 04/2007; Publisher PPV Medien***

This article discusses the equalisation of loudspeakers. Original findings from Research Area 3 are presented in the form of measurement diagrams to verify the written claims. In particular, the usage of equalisers with a non-linear phase response is identified to be superior to linear phase equalisers. At the time of the publication, this finding was a valuable contribution to the field, as linear phase digital equalisers were considered to be the first choice[2] for high fidelity audio.

[2] One example for this is from a book about digital filters. There, "applications requiring linear phase [filters are…] Hi-fidelity audio systems in which phase distortion of recorded music must be minimized to reproduce the original sound with as much fidelity as possible" (Lane and Hillman, 1993, p.7-1)

[4] ***Book: "Studio Akustik"; 190 pages, published 2007, 5th ed. in 2015; Publisher: PPV Medien***

This book is intended for professionals and semi-professionals so they can understand the principles of studio acoustics. In addition to these basic principles, it also provides the author's insights as an acoustic engineer, as well as his findings from all the research areas noted in this contextual statement. However, it focusses mainly on Research Area 1 and presents aspects of studio acoustics more comprehensively than the published material at the time. Furthermore, there are sections in this book that relate to Research Area 2 and Research Area 3, as loudspeakers and studio acoustics interact and both influence the listening experience.

[5] ***Book: "Die Audio-Enzyklopädie" ("The audio encyclopaedia"); 850+ pages, first published 2007, 2nd ed. in 2014; Publisher: DeGruyter***

This book is the largest publication of the author. The title 'encyclopaedia' refers to the comprehensiveness of this book but not to its formal construction, as it is written sequentially in terms of themes. With more than 800 figures and tables, it explains audio technology in a practical and usable way. It provides a broad overview of all relevant topics of audio such as acoustics, psychoacoustics, electroacoustics, electronics, digital audio and more. It contains many practical details about topics including original findings by the author in all three research areas presented in this contextual statement. This book explores the main research question in the chapters dealing with acoustics, psychoacoustics, electroacoustics and measurement techniques. It provides new knowledge to these fields that had not been previously published.

Overview mappings of published works to the research areas are shown in Table 1.

Research Area	Publications
Research Area 1: Improving acoustic modules to increase efficiency in the acoustical treatment of control rooms	[2], [4], [5]
Research Area 2: Measuring time alignment errors, testing their impact on the listening experience and providing solutions for time alignment of loudspeakers	[1], [4], [5]
Research Area 3: Using equalisation for correcting and shaping a loudspeaker's frequency response	[3], [4], [5]

Table 1: Mapping of publications to the research areas.

1.4 Outline of this document

This contextual statement presents work that was undertaken from 1993 to 2007. It presents the context and knowledge of that time in chapter 2. Chapter 3 explains the methodology of the author, discussing different aspects and proceedings to identify best practice. Chapter 4 presents the selected published works in detail, providing insight into the different subjects that were investigated. Results from all research areas are also laid out. Chapter **Fehler! Verweisquelle konnte nicht gefunden werden.** concludes the contextual statement with a summary, a reflection of the proceedings and, finally, a discussion of potential 'next steps'.

2 Context

This chapter places the author's research in its historical context. As all research was undertaken in the past, it is important to view the research in the context of the specific period and its prevailing technological conditions. Many tasks that are quite easy to achieve today were obviously much more problematic in the period under consideration. In order to demonstrate the circumstances and the timeliness of the research, timelines are presented that connect the author's work to the developments in the field at the time of the research. These timelines can be found in Appendix A on page 93.

2.1 Professional background

Early studies in electronic engineering and many years working as a musician formed the basis for the author's decision to work in the field of audio engineering. He started his professional career in this field in 1991 and from late 1992 to early 1995 he worked for the acoustic consulting company ACM in Munich, where he planned and designed studio acoustics and helped to develop new acoustical modules and active loudspeaker systems. At this time, he was also working as a service technician for studios developing specialised electronic equipment. Developing and building this equipment required not only knowledge and skills in electronics, but also a profound knowledge in programming digital signal processors. The latter was gained during the development and manufacturing of complex mixed-signal products mostly in the field of audio engineering. The combination of electronics, acoustics, programming and musicianship led to a skillset that influenced the author's methodology during his research, as he could quickly realise prototypes for tests and observations (see Chapter 3).

The author's first work on Time Alignment Rings was published in 1994 as an article in an audio magazine. In the same year, the Time Alignments Rings were presented and demonstrated at the Tonmeistertagung[3]. Over the years, the author gave numerous papers at the Musikmesse[4] offering new findings in the field of audio technology; as a result these presentations consolidated his reputation as a member of the audio engineering community in Germany. The research results presented at the Musikmesse were also published as journal articles in an audio magazine ("Recording Magazin" published by PPV Medien).

[3] A convention organised by the union of German audio engineers.

[4] The largest fair in the world about music instruments and audio and live sound technology

The idea to write a large book, "Die Audio-Enzyklopädie", emerged in 1996. Finally starting in 2001 it took six years to write the 800-page work, drawing all the figures, getting feedback and finally laying out the final PDF for print. As the book was targeted for a broad readership, no references were used. However, all sources are presented on the corresponding website[5], which also includes additional material. The book was published in 2007 and had a major update with its second edition in 2014.

The book "Studio Akustik" was written on commission. The publisher sought an experienced author in the field of acoustics who would be able to explain this complex subject comprehensibly to a wide audience. There were more than 1000 pre-orders for the first edition, the book is currently in its 7th reprinted edition and is still selling in reasonable numbers[6], which indicates its ongoing significance.

2.2 Contextualisation of the research

This section presents important knowledge regarding the research areas of this contextual statement in order to gain a better understanding of the findings presented in Chapter 4. Much of the following can be found in different chapters of "Die Audio-Enzyklopädie" (Friesecke, 2014) and relates specifically to the three research areas of this contextual statement.

2.2.1 Research Area 1: Improving acoustic modules to increase efficiency in the acoustical treatment of control rooms

For this research area, it is important to look at studio acoustics and at measurement techniques of rooms and acoustic modules.

2.2.1.1 Studio Acoustics

Studio acoustics is an area of research about treating recording studios and control rooms so that room defects are minimised and controlled acoustical conditions are set up. The focus in this contextual statement is on control room acoustics. Reflections, especially in control rooms can be beneficial or detrimental. They are beneficial when they occur in the range of 10 to 20ms after the direct signal. When reflections occur under 10ms of the direct signal they can shift or blur the localisation of

[5] http://www.audio-enzyklopaedie.de/Das_Buch/Quellen.html

[6] In 2020, 298 books were settled by the publisher

phantom sound sources. These reflections should be avoided and an initial time delay gap[7] of 10 to 20ms is recommended for control rooms. If reflections appear later than 20ms they can be heard as echoes, depending on the signal type. Depending on the geometry of the control room, reflections occurring earlier than 10ms or later than 20ms may not be avoidable, but they can be reduced by either absorbing them with absorbers or scattering them with diffusors (Everest, 2001, pp.255–359 and 429–440). Whether an absorber or a diffusor are used is dependent on the approach of the acoustic designer and on the desired reverberation time in the room. While absorbers reduce the reverberation time diffusors do not. Referring to the European Broadcast Union (EBU) recommendation EBU Tech. 3276[8], Maier (2008, p.296) states that the reverberation in control rooms should be uniform across a wide frequency range and should lie in the range of 0.2 to 0.4s, depending on the volume of the room.

One of the most common ways to absorb sound is by using porous materials such as Rockwool or fibreglass wool. These materials have a friction resistance that subtracts energy from a passing sound wave. The absorption coefficient of porous absorbers is dependent on the thickness of the absorber and on the material. The effect of diverse porous materials on sound is well studied and published (Everest, 2001, p.188). Another way to absorb sound is by employing (mass-spring) resonance such as Helmholtz resonators or membrane absorbers. The latter are normally built using a stiff or flexible membrane[9] as a mass in front of a closed air volume that acts as a spring (Everest, 2001, pp.203–209).

Reverberation time can be calculated and measured. For the calculation, the formulas of W. C. Sabine or N. Eyring are used (Everest, 1991, p.109). Normally, a measurement is preferred over calculation, as the absorption coefficients of a room's surfaces are never be known precisely. Both Sabine's and Eyring's formulas can also be used to calculate the absorption coefficient of a material based on the measured reverberation time. In a room with known reverberation time (called a reverberation chamber) an absorptive material is positioned. Then the reverberation time of the room is measured again. The decrease in reverberation time indirectly shows the absorption potential of the material. As a result, the absorption coefficients for the material can be calculated. This method works best with small rooms where the ab-

---

[7] The initial time delay gap (ITD-gap) is the time between the direct signal from a loudspeaker and the earliest reflection from the room.

[8] This recommendation is about the listening conditions in control rooms

[9] Both stiff and flexible materials are used. Common materials are plywood, metal sheets or foils of different weights (up to tarpaulin fabric or even heavier).

sorptive material makes up a significant proportion of the total surface or for rooms with low absorption and, therefore, with long reverberation times so the absorptive material introduces a significant amount of absorption. An example of this process using a simple Rockwool absorber is shown in Appendix C on page 99.

Besides determining absorption coefficients, resonance frequencies are also worth observing. They are relevant for all types of spring-mass absorbers such as Helmholtz resonators and panel or foil absorbers. In this context, not only the resonance frequency itself is of interest but also the Q-factor, which describes how long it takes for a resonance to die out. It is possible to calculate a reverberation time (RT60) from the Q-factor of a resonance (Panzer, 1994, p.86). The higher the Q-factor, the longer the resonance will oscillate. For Helmholtz resonators may not only absorb energy, but can also even prolong the reverberation time if built with a Q-factor that is too high (Everest, 2001, p.229). Helmholtz resonators are not only found in the field of acoustics. Vented boxes are Helmholtz resonators as well (Panzer, 1994, p.59).

Diffusion scatters the sound so that no hard reflections are bounced back to the listening position. Diffusors in control rooms are normally built as QRD[10]. They work in a limited frequency range depending on the width and the depth of the wells. In order to diffuse deep frequencies they need to have deep wells whilst their upper frequency depends on the well-width (D'Antonio and Konnert, 1983, pp.8–10).

2.2.1.2 Acoustic measurements

For measuring room acoustics there is a multitude of possibilities that all lead to the same result. Measurement techniques can be split into measurements in the time domain (impulse, maximum length sequence (MLS)) or in the frequency domain (linear sweep, logarithmic sweep). As seen in the previous subsection, reverberation measurements can be used to test absorbers. Furthermore, reflections in a room and the reverberation of a room can be measured.

Impulse measurements use a very short impulse that is as close as possible to a Dirac pulse. This impulse carries all frequencies with the same level as cosine components (see Appendix B on page 97 and (Randall, 1987, pp.47–55)). The impulse is sent through a loudspeaker into the room and the room response is recorded as

[10] Quadratic Residue Diffusor – a diffusor design based on a quadratic residue sequence developed from a proposal of Manfred Schröder (1975): "Diffuse Sound Reflection by Maximum Length Sequences".

an impulse response. This impulse response can be squared for an Energy Time Curve (ETC) that shows reflections. Alternatively, it can be windowed and then transferred into the frequency domain via the Fourier transformation. Amplitude and phase responses are then available which can be used for frequency-dependent measurements like resonance frequencies or reverberation time. The main disadvantage of the impulse measurement technique is the low amount of energy in the impulse, which leads to a low signal-to-noise ratio of the measurement.

Instead of using an impulse with only a little energy for a measurement, a Maximum Length Sequence (MLS) can be used. An MLS is a predefined pseudo-random sequence that can be depicted as a series of positive and negative impulses of pseudo-random length. An MLS not only sounds like white noise; comparable to the impulse, it also carries all frequencies with the same level. For processing the measurement, the MLS response from the room has to be cross-correlated with the sent MLS (that is, deconvolution) so that an impulse response is derived. As the MLS carries multiple times more energy than the single impulse, the signal-to-noise ratio of this impulse response is significantly better than a single impulse measurement (Everest, 2001, pp.508–509; D'Appolito, 2007, pp.292–294). It has to be noted that the length of the MLS determines the length of the calculated impulse response. For a 65,535-sample-long MLS measured with 48kHz sampling rate an impulse response of 1.37 seconds could be calculated. The impulse response from the MLS measurement can be evaluated comparable to the impulse response from the impulse measurement.

Besides measuring in the time domain, room measurements can also be undertaken in the frequency domain. The simplest measurement technique in the frequency domain is measuring with a sine generator and a level meter. The measurement can be realised by playing single tones one after the other and then measuring their levels. These discrete frequencies are then transferred into a diagram to get the amplitude-frequency response. These measurements can be used for measuring resonators by exciting them with different frequencies while observing the displacement of the resonator's mass, which maximises at the resonance frequency.

Instead of using single frequencies, sweeps can also be used that are then visualised with a level recorder. Sweeps can be linear (e.g. 100 Hz/s) and logarithmic (e.g. two times the frequency per second). At low frequencies, sweeps have to be slow enough to measure a whole period of the signal, but towards higher frequencies the period time decreases and the sweep can be made faster (logarithmic sweep)

(Fehlhaber, 2018). In general, sweeps provide the same measurement results as single tones but reduce the measurement time significantly.

For reverberation time measurements three domains are needed in the measured result: time, energy and frequency. These measurements are also called '3D waterfall diagrams' and they show the frequency-dependent decay of the room. From an impulse response, they are derived by executing multiple Fourier transformations, each starting at a later point in time in the impulse response. For sweep measurements, a delayed sweeping filter can be used to gather the time steps by increasing the delay time step by step. This system is called TDS[11]-measurement. Waterfall diagrams that were virtually cut into slices of octaves or third-octaves can then be used for determining the reverberation time (Everest, 2001, pp.504–508; Crown International, 1997, Appendix C).

2.2.2 Research Area 2: Measuring time alignment errors, testing their impact on the listening experience and providing solutions for time alignment of loudspeakers

This research area focuses on time alignment – an area in which there was only sparse information in 1993. The few sources that existed are presented in this subsection. The term 'time alignment' refers to a process in which different sound sources are adjusted to arrive at the listening position at the same time. In this research, these different sound sources are the different drivers of a multi-way loudspeaker.

At the 54th AES[12] convention in 1976, Edward Long (1976) presented a paper about time alignment. Long's work concentrates on the technique of time alignment with a particular focus on a 'time-align generator'. Using pulse and square wave signals, his measurement setup allowed for easy and real-time measurable results. Although Long's paper is primarily about the process of measuring and optimising time alignment in real-time, it does not explore sufficiently the audible effects of time-aligned systems. However, Long mentions certain effects on stereo material and he refers to other works describing the impact of proper time alignment (especially from R. C. Heyser).

[11] Time Delay Spectrometry

[12] Audio Engineering Society

Richard C. Heyser[13] published a series of articles related to time alignment, although he did not use the wording 'time alignment'. In his two-part article "Determination of Loudspeaker Signal Arrival Times" (Heyser, 1971a; b), he explains measurement systems (especially the TDS system) and how the arrival time of the excitation of a membrane is visible in energy-time measurements. He demonstrated that wavefronts from different parts of the membrane arrive at different times.

Another two-part article "Loudspeaker Phase Characteristics and Time Delay Distortion" (Heyser, 1969a; b) focuses on measuring amplitude responses, phase responses of loudspeakers, and their acoustical position depending on the phase. Heyser explains the relationship between the amplitude and phase of loudspeakers in a comprehensible way. In terms of time delay, he states that "the acoustic position of a loudspeaker should, on average, lie behind its physical position by an amount that is some inverse function of its high-frequency cutoff" (Heyser, 1969a, p.32). This will introduce a larger delay on woofers than expected from their physical appearance. In the second part of this article, time delay is associated with a virtual loudspeaker position. That means variances in time delay (or deviations from linear phase) would result in a frequency-dependent virtual loudspeaker distance (Heyser, 1969b, p.33). If these timing errors happen between multiple drivers (for example low- and high-frequency drivers) then the lobe that is emitted from the loudspeaker is tilted. This is also one of the basic principles of beamforming (Analog Devices, 1993, chap.15). However, most studio loudspeakers have a hemispherical polar pattern and, therefore, lobe tilting would be only relevant if there was a distinct lobe.

In a more maths-based two-part article "The Delay Plane, Objective Analysis of Subjective Properties" (Heyser, 1973a; b), Heyser discusses the representation of audio signals on new domains such as the time-delay domain. This article discusses the subjective effects of time delays (e.g. induced by phase errors). He presents a similar conclusion to those in previous articles stating that a virtual, time-smeared sound source is created by a real loudspeaker with phase errors. Although Heyser's ideas on time delays and their conception are a good foundation for understanding loudspeaker measurement and phase problems, they do not contain explicit information about the time alignment of multi-way loudspeakers or anything about stereo imaging of loudspeakers or, indeed, about localisation errors.

[13] Richard C. Heyser was the inventor of the TDS measurement technique used in the Time-Energy-Frequency (TEF) measurement system. This system laid the foundations of phase measurement for modern loudspeaker design (Davis and Davis, 2015, p.24).

The Finnish loudspeaker manufacturer Genelec built a time-aligned loudspeaker in 1979. It was the model 1024A, a so-called "music monitor" also known by the name "The Dean" (Genelec, 2013, pp.56–57). As Figure 3 shows, only the woofer is aligned to the midrange driver and the tweeter, but not the midrange driver to the tweeter. This differs from the author's findings that aligning the midrange driver with the tweeter is more effective (see subsection 4.2.2.1 at the end).

Figure 3: Genelec 1024A "The Dean" (Genelec 2013: 57).

The Danish manufacturer Klark-Teknik built a two-way loudspeaker with a stepped front baffle named "Jade One" (see Figure 4). In 1994, when the author started researching time alignment, this loudspeaker was still on the market. However, the idea behind this construction was not communicated in a comprehensible manner. The marketing material of the time concentrated on the fact that this was an "integrated speaker" (Klark-Teknik, 1988a, p.21), which meant that crossover, power amplifier and loudspeakers were in the same enclosure. Klark-Teknik did not mention time alignment as a feature in their advertisements. In the manual for that loudspeaker, they state that "… the mounting planes of the two drivers are slightly staggered to optimise the phase characteristics of the system" (Klark-Teknik, 1988b, p.6). There is neither any further explanation around that nor any benefit mentioned. Especially the possible impact of the time alignment on the stereo image of the loudspeaker was not communicated at all.

Figure 4: "Jade One" by Klark-Teknik (Klark-Teknik 1988).

At the Tonmeistertagung 1992 Paul Zwicky presented research on the compensation of timing errors of multi-way studio loudspeakers with the help of digital signal processing. This was published in 1993 in the ‘Tonmeistertagungsbericht’ – a summary of the research output from that convention (Zwicky and Bäder, 1993). Zwicky's article mainly describes sound colouring effects and concentrates on showing that a phase correction is possible but quite often not realised by the manufacturers. In this article Blauert and Laws were cited, who said: "timing errors become marginally audible above 400µs" (Blauert and Laws, 1978).

Since time alignment affects sound sources from one (or nearly one) position to arrive either simultaneously or slightly delayed at the listening position, the impact of non-time-aligned signals arriving from one position with short delays is also relevant. Jens Blauert discusses the impact of reflections from the pinna and that they create frequency-dependent notches that the brain uses as cues for directional hearing (Blauert, 1974, pp.50–52). This part of his research is about single sound sources and how the human ear can detect their direction by the above-mentioned cues. In the section about stereo sound reproduction and phantom sound sources, there is nothing mentioned about the impact of pinna reflections. In the section about delayed sound sources, only localisation of phantom sound sources, the precedence effect and echoes are described. There is no mention of imprecise phantom sound sources in combination with short delays (<1ms) being emitted from one point (Blauert, 1974, chap.3.1). Therefore, Blauert's research does not address the effects of short delays from the same direction on the stereo signal. Two articles from the mid-80s described the effect of natural reflections from the pinna in the range of 200µs that is responsible for sound spatialisation (Burchard, Irrgang and Andresen, 1985, 1987). The authors discovered that the pinna adds a reflection to the direct signal. Depending on the strength of that pinna reflection the sound source is located at different angles. Signals arriving from ±90° (the side of the head) have the strongest level of these reflections. This is the only source providing a finding about the effects of short delays from the same direction.

2.2.3 Research Area 3: Using equalisation for correcting and shaping a loudspeaker's frequency response

For this research area loudspeaker- and filter-theory are particularly relevant. A short section on measurement techniques for loudspeakers, extending subsection 2.2.1.2 from Research Area 1, is added to complete the picture.

2.2.3.1 Signals in the frequency and time domains

The observation of loudspeaker and filter responses can be undertaken in the time and/or in the frequency domain. These domains mathematically relate to each other and this relationship will be briefly explained here.

Audio signals are normally presented in the time domain as a waveform. For example, the waveform shows the output voltage of a microphone plotted over time. It could be argued that the time domain is the domain where the signal is represented in its natural form. In the time domain, the form of an audio signal (e.g. a sine, a triangle or a rectangle function) is easily recognisable. Furthermore, transients, reflections and the time of oscillation can be read directly.

An audio signal can also be represented in the frequency domain. The Fourier transformation is the mathematical formula to convert a signal from the time to the frequency domain and vice versa. In the frequency domain a signal is displayed as a spectrum which either consists of real and imaginary amplitude information plotted over frequency (that is, a complex spectrum) or of magnitude and phase information plotted over frequency. The latter is the most common representation of a spectrum. Quite often only the magnitude is shown. Furthermore, 'amplitude' is often used as a synonym for 'magnitude'. In this contextual statement 'amplitude' is used throughout. In the frequency domain the parameters of equalisers such as centre frequency and quality can be seen easily.

Impulse responses that are needed for measurements and digital filters can also be transferred into the frequency domain. They result in a set of complex gain factors that describe the modification of a complex spectrum. What is known as convolution in the time domain correlates to a multiplication of complex numbers in the frequency domain. An engineer has the choice to observe or modify signals in either the time domain or the frequency domain. The result will be the same (Randall, 1987, p.53).

2.2.3.2 Reading linear phase diagrams

When looking at frequency responses, normally a logarithmic scale at the frequency axis is used. This is because human hearing behaves similarly, and so the diagram and the signal perception match each other. But when it comes to evaluating phase – especially linear phase – a logarithmic diagram is not preferable. As its name suggests, linear phase should be visible as a straight line in the phase response plot. But this is only true if the frequency axis is shown on a linear scale. Figure 5 shows linear phase on a logarithmic scale (left) and on a linear scale (right). From the left diagram it cannot be observed that this phase plot shows linear phase. It is therefore important to evaluate linear phase in diagrams with a linear scale.

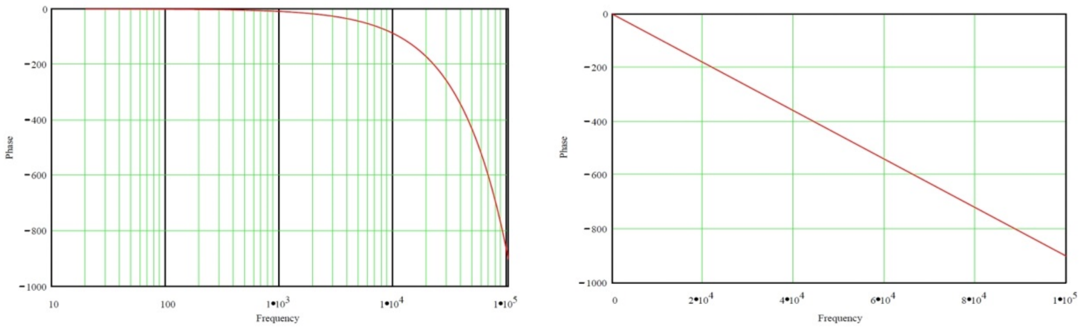


Figure 5: Linear phase on a logarithmic frequency axis (left) and on a linear frequency axis (right).

2.2.3.3 Filter theory

In the context of the published works under discussion, only second-order (that is, biquad[14]) or higher even-order filters are relevant. As every even-order filter can be assembled by a number of second-order filters (that is, staged biquads), only these second-order filters are described here (Lancaster, 1994, p.4-2 and p.5-1; Smith, 1999, p.603). The following subsections present both analogue and digital filters, and outline in particular the restrictions of digital filters.

[14] A biquad is a filter section providing up to two poles and two zeros in the complex transfer function (Smith, 1999, p.600).

Analogue filters

In order to have maximum flexibility, filters are preferably designed as active filters. Especially with state-variable filters (normally used in fully parametric equalisers) the parameters f_0 (frequency), G (gain) and Q or d (quality or damping) can be used independently of each other (Lancaster, 1994, p.7-11). Excellent resources for analogue filters can be found in "Das Aktiv-Filter Kochbuch" by Don Lancaster (1994) and in Chapter 13 of "Halbleiterschaltungstechnik" by Tietze and Schenk (1980). Filters can be described by a complex transfer function where f_0, G and Q (or d) are used as parameters. This transfer function also acts as the basis for the transfer into the digital domain for digital filters. There are many types of filters but for the published works the main focus is on second-order band-pass filters, as these filters are integral parts for fully parametric equalisers. The transfer function and frequency response for the second-order band-pass filter and its resulting equaliser responses are shown in Appendix D on page 105.

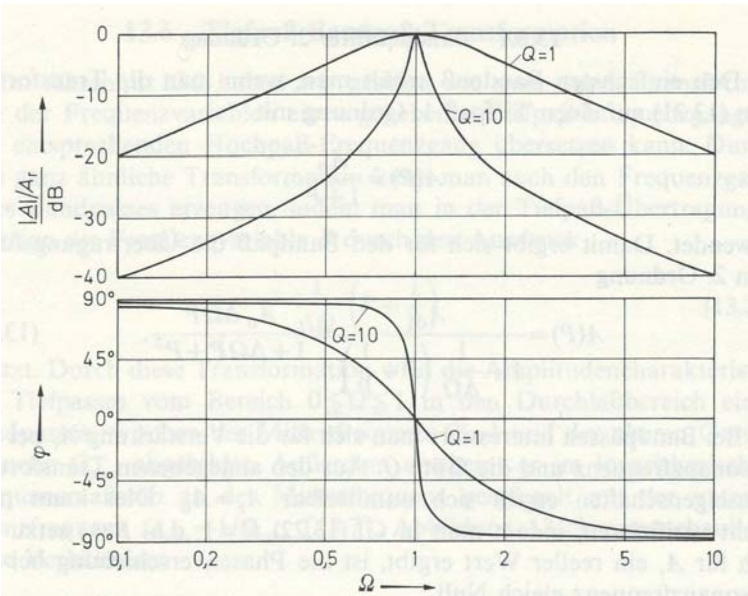


Figure 6: Amplitude and phase response for second-order band-pass filters (Tietze and Schenk, 1980, p.304).

As can be seen in Figure 6, besides the changes in amplitude there is also a phase shift introduced, which maximises in the stop-band of these filters and approaches zero in the pass-band[15]. This phase shift is tightly joined to the amplitude response and cannot be prevented.

[15] The stop-band is the area where the filter attenuates the amplitude of the input signal by more than 3dB. The pass-band is the area where the signal loss is less than 3dB.

Digital filters

Digital filters are commonly known as IIR[16] and FIR[17] filters. Both are calculated by summing up weighted, delayed versions of the input and output signal. One delay stage is called 'tap' and it delays the signal by one sample. IIR filters are also called recursive filters and consist of a feed-forward and a feedback section. So, besides the input signal, parts of the output signal are also used for calculating the impulse response of an IIR filter. IIR filters can easily be implemented as "direct form IIR" filters (Lane and Hillman, 1993, sec.2). For higher-order filters multiple second-order filters can be calculated in series, e.g. in the cascaded direct form (Lane and Hillman, 1993, sec.5). The IIR biquads are comparable to the analogue biquads except that their transfer function is in the digital domain. More about this can be found in Appendix D on page 105.

The time and value quantisation in the digital domain leads to some limitations of IIR filters:

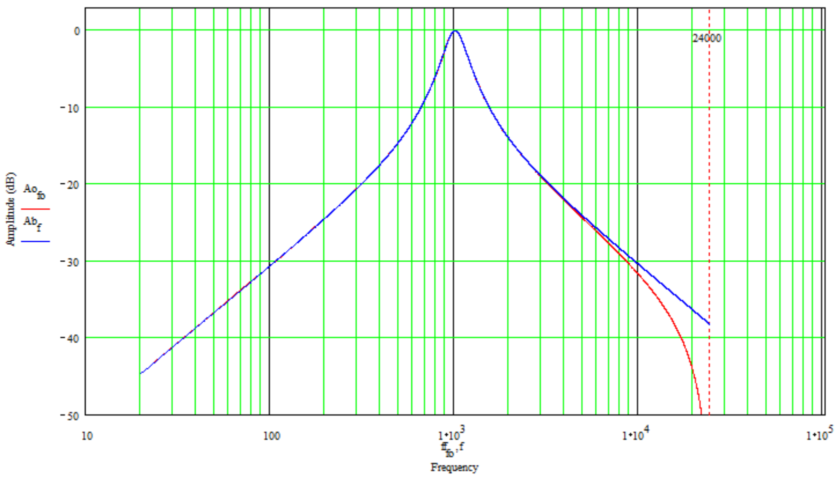
- At the Nyquist frequency, IIR band-pass filters will always have zero output and therefore the band-pass is no longer symmetrical on a logarithmic frequency scale (Lane and Hillman, 1993, p.5-8). For sampling rates of 44.1kHz (Nyquist frequency 22.05kHz) this effect becomes very noticeable at 20kHz. A solution would be to use a higher sampling rate and thus a Nyquist frequency, which is further from 20kHz.
- In all IIR filters the coefficient alpha gets very small for low frequencies. For small numbers the resolution of the processor's word width could lead to rounding errors. The effect is as if the cut-off/centre frequency gets quantised against the word width of the calculation. This could be solved by calculating filters with a greater word width (e.g. double accuracy), which, as a disadvantage, needs more processor cycles.

A comparison of an analogue and a digital band-pass filter with the same f_0 and Q is shown in Figure 7. It can be seen that these filters lead to almost identical frequency responses. Only approaching the Nyquist frequency does the IIR filter deviate from the analogue one.

---

[16] Infinite Impulse Response.

[17] Finite Impulse Response.



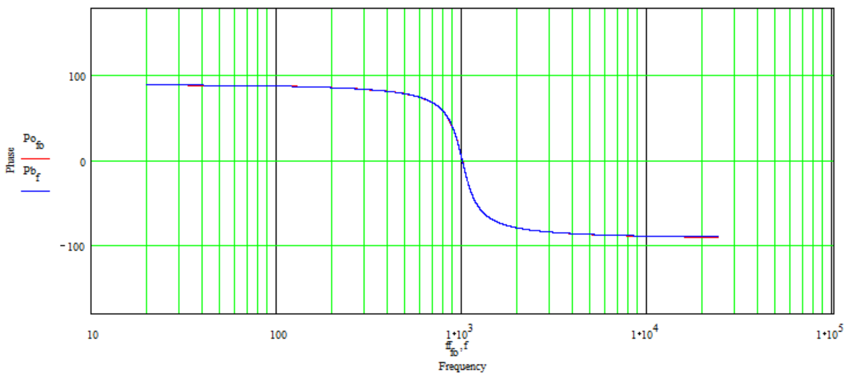


Figure 7: Comparison of an analogue (blue) and digital IIR band-pass (red) with f_0 = 1000Hz and Q = 3.41 (one third octave). The sampling frequency for the IIR is 48kHz.

FIR filters are the other type of digital filters widely used in digital audio. FIR filters only have a feed-forward section, but they are in fact IIR filters with all recursive coefficients set to zero. Therefore, all recursive parts can be omitted. The mathematical operation that is carried out is a "convolution (also called finite impulse response or FIR)" (Smith, 1999, p.276). In the field of audio engineering the term 'convolution' is often heard in conjunction with a 'convolution reverb'. From a mathematics perspective, a convolution reverb and an FIR filter are the same—compare Smith (1999, p.120) and Analog Devices (1993, p.68).

In FIR filters the amplitude response and the phase response are largely independent of each other. It is possible to design filters with (almost) freely designed amplitude and phase responses (Analog Devices, 1993, p.67; Smith, 1999, p.297). This provides the possibility to create a special form of FIR filter: the 'linear phase filter'. This type of filter changes the amplitude response whilst only delaying the input signal by a certain amount. A delay on its own does not affect the hearing experience of a signal. Only when delayed signals are added with non-delayed signals will the delay have an audible impact such as a comb filter. For this reason, the linear phase filters are often claimed to be the best possible filters for correcting amplitude errors. Lane and Hillman (1993, p.7-1) state that linear phase is good for "Hi-fidelity audio systems in which phase distortion of recorded music must be minimized to reproduce the original sound with as much fidelity as possible".

Linear phase FIR filters have a symmetrical impulse response (Lane and Hillman, 1993, p.7-8). This means that there is a pre-ringing[18] before the actual impulse appears that eventually dies out. This pre-ringing is something that can be considered to be unnatural because in nature there is normally an event that triggers something that then dies out, but there is no 'ring-in' and 'ring-out' in this scenario. This pre-ringing is therefore also claimed to be problematic, as Manley state in their manual for the 'Slam!' stereo limiter: "… be careful when assuming that 'linear-phase' is a Good Thing [sic] in filters. It is, in fact, a pretty unnatural sound" (Manley Laboratories, 2004, p.20).

Another aspect of linear phase filters is the introduced latency. As the main impulse is in the middle of the symmetrical impulse response FIR filters, the signal delay is half the length of the filter. For a 100-tap filter it would be 50 samples. Depending on the field of application this might be too long, as audio becomes asynchronous to video, or musicians are disturbed by audible comb-filter effects[19], or even perceptible delays.

Limitations of FIR filters are mainly given by the sampling rate (f_S) in combination with the number of taps. One (linear) frequency band is f_S/taps wide. For a 100-tap FIR filter sampled at 44.1kHz that is 441Hz. The filter, therefore, can be controlled by frequency points starting at 0Hz and then being set 441Hz apart from each other. For filters with finer control (which is especially reasonable for the low-frequency area), the number of taps has to be increased. This leads to very long FIR filters if an accurate control of the bass area is desired (Smith, 1999, pp.297–300).

---

[18] Pre-ringing (or pre-echo) is a side effect in linear phase digital filters. The impulse response of these filters is not only trailing out past the actual signal, it is also fading in in advance of the actual signal. This is an unnatural behaviour and it is still not fully understood how the human brain interprets this effect. Some useful information about this can be found in (Manley Laboratories, 2004, p.20).

[19] A singer that monitors his voice over a headphone will receive two signals in his ear: His own voice through his own head plus the headphone signal through his ears. If the headphone signal is delayed then comb-filter effects will be a result of the addition of these two signals. This leads to frequencies that are cut and others that are boosted. This way the voice of the singer might be heard thinner, confusing the singer and therefore reducing his performance.

Long FIR filters become quite process-intensive. For each tap a multiplication and an addition must be calculated per sample. For example, a 1,000-tap FIR calculated at 96kHz sampling rate would need 1000·96000 = 96 million multiplications and additions per second. This value doubles for a stereo signal. DSPs can handle data retrieval, multiplication, addition and data storage in one processor cycle but for the above-mentioned filter a 200MHz DSP would be necessary for a filter that only has a frequency resolution of 96Hz.

An alternative to calculating the convolution is to do the calculus in the frequency domain. There, it is only a multiplication of the real and imaginary components of the spectrum. For this procedure it is necessary to first transfer the signal from the time into the frequency domain, then do the multiplication and finally transfer it back into the time domain. This transfer could be accomplished by fast Fourier transformations (FFTs), which execute very efficiently on modern processors (Smith, 1999, pp.311–318). For very large FIR filters, this so-called 'fast convolution' or 'high-speed convolution' is the only way to execute the calculation in real-time.

2.2.3.4 Formulas for applying filters to a signal

When filters are applied to a signal (in the time domain), the signal is **multiplied** with the transfer function of the filter. When two or more filters are in series, a series of multiplications take place. When signals are filtered with a crossover for sending low frequencies to the woofer and high frequencies to the tweeter, these filtered signals are **added** at the listening position. When multiplying or adding the transfer functions, both amplitude and phase have to be considered. This will automatically be the case if filters are calculated with complex numbers. If only the amplitude is considered (and phase information is omitted), the result will be erroneous, as the phase shifts introduced by filters could lead to signal cancellations as well as to signal additions.

A very good example of this problem is a loudspeaker crossover. Many passive two-way loudspeakers use second-order high- and low-pass filters for the crossover. The crossover splits the audio signal into a high-frequency portion for the tweeter and a low-frequency portion for the woofer. When listening, tweeter and woofer will be added and therefore the sum of both filter functions will affect the overall frequency response. Adding a second-order high-pass filter to a second order low-pass filter (both Butterworth characteristics -> Q=0.7) with the same cut-off frequency will lead to a notch at the cut-off frequency, as the phase shift between these two filtered signals is 180°. Due to the phase shift, the added signals are in fact subtracted from each other at the cut-off frequency (Dickason, 2002, pp.192–193). To avoid this notch at the cut-off frequency, the tweeter can be connected with inverse polarity. This leads to a subtraction of high- and low-pass filtered signals and due to the phase shift finally to an addition at the cut-off frequency. The result is a +3dB boost at the cut-off frequency (Dickason, 2002, pp.192–193). However, an inverted polarity of the tweeter is not always necessary but in practice dependent on the phase response of the drivers (Stark, 2003, p.160).

2.2.3.5 Loudspeaker theory

As a minimum, loudspeakers consist of a driver surrounded by a box. For passive multi-way loudspeakers there is an additional crossover. The box is needed to avoid the acoustical short circuit of the driver, which happens when low frequencies diffract around the membrane and the pressure in front and behind the membrane cancels out. Most of the boxes are either built as closed boxes or vented boxes. Other models like transmission line boxes or boxes with passive radiators are rarely found.

Using a single driver in a loudspeaker presents several technical challenges. The membrane has to be light to have a good impulse response, but stiff so as not to produce partial oscillations. It has to be large to move sufficient air at low frequencies and at the same time small to avoid too much directivity towards higher frequencies (Dickason, 2002, pp.19–20). These conflicts are solved by building multi-way loudspeakers. They distribute certain parts of the frequency spectrum to dedicated drivers that are optimised for the respective frequency range. This results in a more even directivity of the loudspeaker, and the energy of the signal is distributed to the drivers that can handle it best. Attention should be paid to the crossover network that introduces phase shifts and thus possible delays between the drivers. Ideally, the drivers are mounted close to each other or coaxially to provide a point sound source (Goertz, 2008, pp.425 & 474).

Small boxes are normally used as so-called 'near-field monitors'. They are placed close to the listening position so that the directly radiated sound dominates for the listener. Reflections from the room are only slightly excited and therefore the room has only little influence on the overall sound. However, small boxes normally have a limited low-frequency reproduction because they cannot contain a large woofer. Large loudspeakers are normally placed in the far-field, as their dimensions would prevent them from acting properly as a point sound source if placed too close. However, loudspeakers placed further away interact more with the room and controlled room acoustics are becoming increasingly important for a good listening experience (Moylan, 2015, pp.368–369).

Closed box

Closed boxes act like a second-order high-pass filter (Panzer, 1994, p.33). This filter is directly comparable with an electrical second-order high-pass filter, as its complex transfer function is the same (Panzer, 1994, p.88). The filter is largely influenced by the resonance frequency of the spring-mass-system built from box volume (that is, spring) and membrane mass. The larger the box, the larger the compliance[20] of the enclosed air will be. A larger compliance leads to a lower resonance frequency of the spring-mass system. It also leads to a smaller Q-factor of that resonance. For smaller volumes of the box, the compliance will also get smaller and the resonance frequency and Q-factor will rise. Typical designs for closed boxes aim for a flat frequency response of the second-order high-pass. This is the case for only one specific volume of the box where the filter reaches a Q of 0.7 (Butterworth design). Building the box larger leads to a $Q < 0.7$ (Bessel design), whilst building the box smaller leads to a peak in the frequency response and a $Q > 0.7$ (Chebychev design) (Panzer, 1994, pp.34 and 89).

Vented box

A vented box makes the energy radiated into the box partially utilisable by adding a further resonance that supports the directly radiated sound. As the most energy is needed at low frequencies, the centre-frequency of this resonance is usually at the low end of the loudspeaker. The additional resonance is realised by a port with a defined length and a defined cross-sectional area. The air in that port (plus some air near to the openings) acts as mass, and together with the air-spring from the volume of the box, an additional spring-mass resonance emerges. For frequencies above the port-resonance the vented box behaves like a closed box. Below the port-resonance the air in the port can travel freely and an acoustical short-circuit happens. As a result, the vented box cuts off like a high-pass fourth-order filter (24dB/oct.).

[20] Compliance is the reciprocal of the spring rate. Calculating with compliance instead of spring rate makes the spring-mass resonance formula directly comparable to the electrical LC-resonator formula.

Adding the port-resonance to the output of a closed box leads to a 3dB boost, so the low-frequency response can be extended by about one octave. It is crucial that the port-resonance is calculated properly so that it lies at the falling slope of the closed box and therefore helps to seamlessly extend the low-end frequency response of the loudspeaker. It is recommended the port-resonance should have a Q of about 5 (Panzer, 1994, p.60).

Filtered (vented) box

Closed and vented boxes can be complemented with an additional electrical second-order high-pass filter. This filter can be designed with a Q greater than 0.7 so that its resonance will lead to a peak before the slope appears. Setting the peak frequency at the lower end of the box further extends the low end of the box. For a closed box the additional filter then leads to a 24dB/oct. cut-off slope and for vented boxes it is 36dB/oct. For the latter, the filter is a benefit, as frequencies in the range of the acoustical short-circuit will be attenuated and therefore the membrane excursion will not get too large (Panzer, 1994, chap.8.2 and 8.4).

Amplitude and Phase Errors

Filters and loudspeakers are devices that introduce phase shifts into the signal. Phase shifts are frequency-dependent delays, meaning that certain parts of the signal spectrum are emitted at a different time to others. A linear phase (see footnote 1 on page 2) only results in a frequency-independent delay such as that caused by the time of flight of sound when a loudspeaker is at a certain distance from the listening position. Since the phase and amplitude response (the frequency domain) of a signal is directly related to the waveform (the time domain) of a signal, each non-linear phase shift (even with an unaltered amplitude response) will lead to a change in the waveform of a signal. This can deform, for example, transients of a signal, which are important for signal recognition and localisation. Discussions about the phase sensitivity of our hearing system are found in "Musik im Kopf " (Spitzer, 2013, pp.69–71) and "Psychoacoustics" (Zwicker and Fastl, 1999, pp.187–191)

Heyser (1969a, pp.39–40) states that for minimum-phase loudspeakers, it would be sufficient to present the characteristics of a loudspeaker by amplitude, as the phase will follow the amplitude. He also presents a set of figures that help to identify which phase errors are introduced by different amplitude errors in a minimum-phase loudspeaker (see Figure 8).

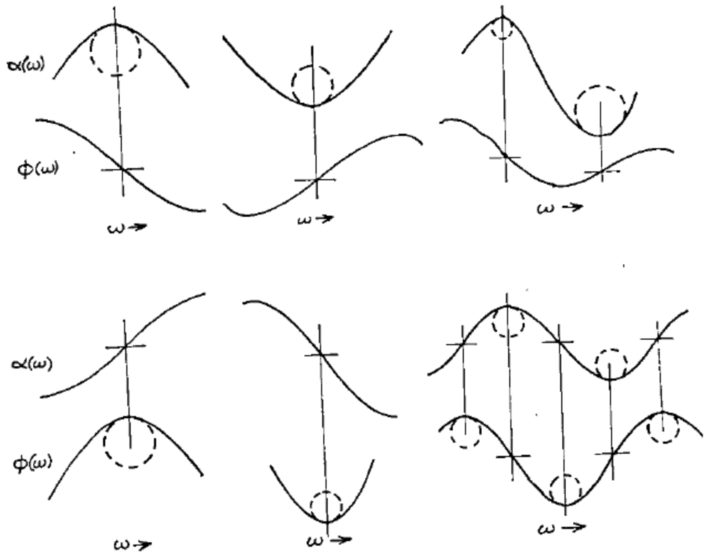


Figure 8: Six examples of amplitude- (a) and corresponding phase (Φ) errors for a minimum-phase loudspeaker (Heyser, 1969a, p.40).

2.2.3.6 Loudspeaker and filter measurements

Measuring loudspeakers and filters is similar to the room measurement techniques shown in subsection 2.2.1.2. However, there are some additional conditions and options that are shown here.

Loudspeakers should be measured in an anechoic chamber to avoid room reflections that lead to comb-filter effects. That said, quite often the measurement is only possible in the control room, as the loudspeaker is already installed and cannot be moved into an anechoic chamber. To attenuate these room responses different options are available. With measurements in the time domain, the impulse response could be overlaid with a window that closes before the first room reflections arrive. In well-treated rooms with a typical initial time delay gap of 10-20ms, this leads to reflection-free measurements down to 50-100Hz (reciprocal of 10-20ms).
For sweep measurements there are two options to attenuate the room reflections:

- Sweep responses could be transferred from the frequency to the time domain and the resulting impulse response is treated as described above (possible for linear and logarithmic sweeps).
- A linear sweep can be filtered with a linear sweeping band-pass that attenuates room reflections. As all reflections always arrive later than the direct signal, they are lower in frequency than the direct signal and therefore the band-pass sweeps to the signal generator with a delay (Time Delay Spectrometry, TDS). Practical implementations use a fixed filter and multiply the

sent and the received sweep in order to get a constant difference frequency that can then be filtered with a non-sweeping filter. This constant difference frequency is only obtainable with linear sweeps.

(Stanley, 1982; Crown International, 1997).

A loudspeaker measurement technique common until the 1990s used a square wave signal. A square wave contains all odd harmonics with decreasing levels as sine components (D'Appolito, 2007, p.257). When a square wave signal is passed through a device under test (DUT) and it comes out as a square wave signal again, the amplitude and the phase response must be flat. If there are any deviations from flat, the square wave will have been distorted and would therefore no longer look like a square wave. Although this way of measuring loudspeakers is very simple to execute, its results are difficult to interpret. Seeing a distorted square wave indicates irregularities in amplitude and/or phase response, but it is nearly impossible to evaluate amplitude and phase errors at specific frequencies.

3 Methodology

This chapter is an account on the author's methodology used in his research projects. Although much research could have been realised in a purely theoretical manner, a key methodological approach is to combine theory and practice. Building prototypes, testing and measuring them, and conducting listening tests are methods that complement the theory that forms the basis for each research project.

3.1 Overarching methodology of the personal research

All studies start with a problem that needs to be investigated and hopefully solved. But not all problems can be identified immediately if they are only studied theoretically. Problems that are not yet captured in theory only become apparent in a practical application. For this reason, the author's methodology always takes into account practical effects or is based on effects that can be observed in practical situations.

In the field of audio engineering, the author observed two approaches: one, which is more technically and theory-oriented and the other, which is more intuitive and practice-oriented. In practice, both approaches blend into each other but most of the time one element is more dominant than the other. In the more technical oriented approach, physical explanations are given while in the more intuitively oriented approach the perception is more important than a technical explanation.

The difficult task when researching in this field is to combine these two methodological approaches, as otherwise a theorist and a practitioner will frequently talk at cross purposes. An example of this can be seen in the '1dB issue': a task could be to try to hear a 1dB[21] change in level. In the audio educational books, the statement is written that "1dB is the average barely perceivable change in sound pressure" (Webers, 1989, p.103). Audio engineers trusting the theory will work in accordance with this statement and believe that there is no perceivable level change below 1dB. But in reality, it depends greatly on the type of signal (e.g. music or noise), the type of transition (sudden or soft change) and the frequency (full range change or only in a limited frequency band) whether a 1dB change can be heard or not. Zwicker and Fastl (1999, pp.175–182) agree on the value of approximately 1dB but they also look closer at different ways of level changes. In practical listening tests they showed that there are signals and conditions when a subject can barely identify a 2dB change and there are other conditions when 0.3dB will be noticed.

[21] dB: Short for decibel. This is a logarithmic ratio to describe differences in power on a logarithmic scale.

Ultimately, only practical tests will resolve what can or cannot be heard. And in the field of audio engineering, it is always the listeners' ears that have the final say. For that reason, any serious research should take into account both the theoretical background as well as the practical impact on listeners. Results gathered from theory and from measurements should therefore always be tested with listeners.

> "There is much to be gained by adopting a dual approach: generating contextualized knowledge on the basis of careful, systematic inquiry and evaluating this through action oriented towards improvement; while at the same time maintaining a critical scepticism and openness to different interpretations that iteratively challenge the action research 'findings' in terms of both the appropriateness of the action and any claims to improvement." (Somekh, 2006, p. 27)

This statement from the book "Action Research" by Bridget Somekh, brings the theoretical research and the practical feedback together. It acted as a methodological framework for the author. Action research "... is applied research, carried out by practitioners who have themselves identified a need for change or improvement..." (Bell, 2006, p.8). Although action research is originally applied to changes in social groups (Somekh, 2006, p.6), it also offers approaches that are appropriate for the author's research; especially when it comes to the listening experience, as there is a need for subjects to participate in the research given that the listening experience is subjective. Action research investigates a "…real-word problem [...] typically at work..." (Denscombe, 2010, p.126) where the subject acts as a "…feedback loop in which initial findings generate possibilities for change..." (Denscombe, 2010, p.126). In the field of audio engineering, these subjects are the ones striving for further improvement and therefore trigger more research regarding issues identified in their professional practice.

Similar to action research, the author's methodology is a cyclical one that approaches a problem step by step striving for improvements. However, it differs from action research in that there is not necessarily a social group engaged in the practical feedback but rather a practical element, as, for example, a prototype, that is used for verification. The phases of the author's research cycle are shown in Figure 9.

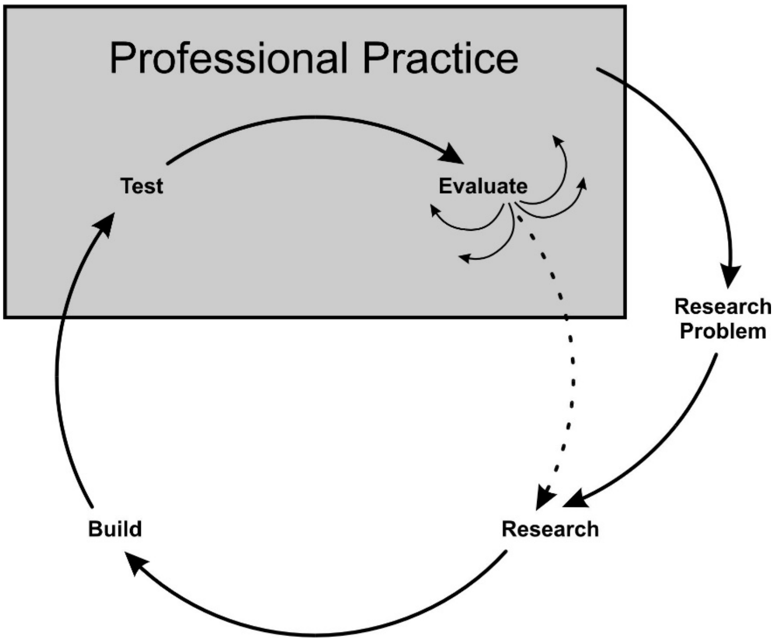


Figure 9: The author's research cycle.

As can be seen, the research problem arises directly from professional practice. A research question or hypothesis has to be formulated and then the research starts. Unique to this cycle is the next step, where some kind of prototype is always built. Testing and evaluating the prototype normally takes place in the professional practice together with other professionals. If the evaluated result is not satisfactory, another research cycle has to be started (dashed line in Figure 9). If the result is satisfactory, it is published and disseminated in the professional community.

This approach has proved to be applicable in all research that the author conducted. The built-phase was particularly necessary, as during the theoretical research many discrepancies were identified. To find the real results, it was necessary to test different theories with the support of measurements or listening tests under practical circumstances. Only this course of action led to trustworthy results that were relevant to the field.

The following subsections will discuss and justify the different phases of the research cycle in greater detail.

3.1.1 Professional practice

A research problem is often identified by practitioners, and in the field of audio engineering it is mostly the audio engineer that strives for a certain auditory experience. Being part of the audio engineering community helps the author to distinguish between a personal problem and a problem that is relevant to the whole community –

for example, a conceptual problem (Booth, Colomb and Williams, 2008, p.56). Such a conceptual problem could be: 'why do certain loudspeakers have large image depth and preciseness in presenting phantom sound sources while others do not?' Although this might have started as an individual acoustic capacity such as, 'why is it not possible for me to judge depth imaging on my loudspeakers?', it is a conceptual problem that is affecting many audio engineers, as it is not related to the listener but to the construction of the loudspeakers.

Professional practice has an additional advantage that is considered to be important: if a problem arises in a professional surrounding it is likely to be a significant problem. In academia, problems are frequently constructed which have only limited (if any) relevance to the industry. Booth et al., state that "Too many researchers at all levels write as if their only task is to answer a question that interests them alone" (Booth, Colomb and Williams, 2008, p.52). Only when others benefit from the research will it have relevance to the industry. Therefore, interplay between the researcher and the end-user is important; and ultimately, it is only the end user who decides whether the auditory experience has improved or not. The real challenge relating to problems constituted by practitioners is to shape them into researchable topics.

3.1.2 Research

Research in the field of audio engineering combines different areas that are interdependent with each other. These areas are electronics, digital signal processing, acoustics, psychoacoustics, and to a certain extent also music. Some of the audio engineering research is more than 100 years old[22]; and in the 1990s was not very accessible. Although the internet already existed at that time, it contained only sparse information, and few or no archives or scans of printed publications. Currently, extensive archives are available.

In the early 1990s, when the author's research began, essential information in the field of audio engineering was contained in a few books focusing on different areas of audio production. Examples include Dickreiter (1987, 1990), which was strongly influenced by broadcast stations, Webers (1989) with a film sound production background and Henle (1990) guided by music production. Ballou (1991) was one of the

[22] Psychoacoustic research for example has a very long history. For example, Fletcher and Wegel (1922) have started their article about "The Frequency-Sensitivity of Normal Ears" with the statement "A large amount of work has been done during the last fifty years in an endeavour to determine in absolute terms the minimum amount of sound that the human ear can perceive"

first books that addressed multiple areas of audio engineering. It was also a publication that the author took as an orientation for his own work "Die Audio-Enzyklopädie".

Valuable information about electronics and digital data processing was gathered from the datasheets and whitepapers that were published from the manufacturers of electronic components and microchips in the form of reference books. These books not only contain information about the devices but also have fully developed application examples using these devices. In the early ’90s, Analog Devices (1992) and Crystal Semiconductor (1994) were market leaders and published extensive datasheet and application-note collections. For the most current research, the papers from the Audio Engineering Society (AES) were important sources. The AES publishes scientific research from the industry and universities. However, in the 1990s there were no searchable archives of these publications.

A very valuable part of the research stage was in the form of discussions with peer experts in the field, through which research issues were shared. As the author was working in the field of studio acoustics, he became acquainted with many fellow experts in that field. Knowing these experts had another benefit that was important for a planned publication: they frequently acted as informal peer reviewers for the author's publications.

3.1.3 Build

All research areas required the building of a prototype to make the problem palpable and to undertake measurements or listening tests. The build phase is therefore a crucial element in the research cycle. One of the most important aims of this phase is to control all aspects and to get a full understanding of the research topic. A prototype would not work if the topic was not fully understood. Another advantage is that nothing will be hidden when building one's own prototypes, as long as no prefabricated modules were used. Furthermore, a prototype could be exposed to real-world situations and not only subject to theoretical constructs.

Building prototypes in the field of audio engineering requires diverse skills. Knowledge regarding the construction of wooden enclosures for loudspeakers and acoustical modules is needed; electronic circuits have to be designed and soldered; and the programming of digital signal processors or computer programmes is necessary. Prototypes can be built for research purposes only. That implies that it is neither the goal to have a market-ready product nor a well-designed artefact. As an example, a digital filter for audio signals could in fact be programmed in Excel as a spreadsheet. Although it will never be possible to render audio signals in real-time with Excel, this approach might be sufficient for looking at the impact of the filter on the audio waveform and it clearly shows every step of the calculation as if the user were watching audio processing in super-slow motion. As an example, diverse routes were taken to simulate digital filters in Excel, Basic, Mathcad or on DSP-Platforms.

The author's previous knowledge and skills gave him the ability to build prototypes rather quickly. Therefore, the approach of building prototypes could be applied consistently for all research. A great advantage that evolved over time was the fact that knowledge and skills in building prototypes became more and more sophisticated. Advancing older prototype, new ones could be built faster, perhaps even faster than attempts to implement a complex theory. In the author's view, measurements and tests with prototypes reflect real-world situations and supersede theoretical approaches. Good examples of this are the mouth corrections[23] of Helmholtz resonators, which are very difficult to calculate for irregular port shapes but easy to measure. In these cases, the experimental approach might be the only way to get a result in a reasonable amount of time.

3.1.4 Test

The test phase is about gathering data in a systematic and structured way. This can happen during listening sessions or by measuring the results. Both will lead to a better understanding of the circumstances as "observation, reason and experiment make up what we call the scientific method" (Feynman et al., 1997, p.32). In the presented research areas, results were tested by measuring or listening; or by combinations of both. These two variants will be discussed in this subsection.

3.1.4.1 Measurements

Measurements have the advantage that they are objective. If executed properly a measurement is not liable to personal effects such as hearing fatigue or being biased or influenced by a third party. Therefore, it is always feasible to use a measurement as either a basis for further tests or as a verification for mathematical approaches. Although measurements might have disturbing factors, a properly execut-

[23] The mouth-correction describes the amount of air that adds to the air-mass in a duct. This additional air is dragged along with the moving air in the duct. Its amount is highly dependable on the cross-section-shape of the duct.

ed measurement is more valuable than a mathematical simulation, as the measurement shows the real-world result and not only a simulated and, therefore, only a probable one. Thus, the measurements transfer the theoretical findings into practical findings. Starting with measurements is also a feasible way to get a project going and capture initial results for developing a research question.

When executing measurements there are many pitfalls and problems, especially when measuring loudspeakers in rooms. The room response will add to the loudspeaker's response and distort the measurement. It is therefore important to either damp the room or use measurement techniques that can attenuate the room (see subsection 2.2.3.6). One of the most important rules about measurements is that they have to be reproducible (Crown International, 1997, pp.C-2–C-3), so one can trust the results. Within a series of structured measurements, it is important that there is only one parameter that is changed between two measurements so that the cause for certain effects can be isolated. Only by combining many tests with small changes and eliminating certain possibilities is it possible to deduct an outcome from these measurements.

In measurements only issues show up that are measured correctly and looked at with a properly formatted output. Potential problems can be overlooked or declared to be irrelevant when, for example, the measurement does not fully account for them or if a wrong scale is used when looking at the results. Furthermore, simply executing measurements is not enough at this point, as it is the subjective human listening impression that counts. As the presented research areas are about improving the listening experience, listening tests are considered to be an important supplement to consolidate the results of the research.

3.1.4.2 Listening tests

In listening tests, subjective perception is the primary focus. Following a filter model from Bech and Zacharov (2006, pp.41–42), there are three stages where a signal can be measured. Following the first stage of the physical measurement (see previous subsection) listening tests can add perceptual and affective measurements, and complement the objective data with subjective results. In the perceptual measurement stage, subjects can validate specific attributes that can bring improvement, and in the affective measurement stage, subjects can communicate their preference for a particular outcome.

For the participants of listening tests, it is crucial that they trust their ears and remain true to their sensory perception. A prominent publication from Toole and Olive (1994) shows that having a visual impression (such as the size or the material of a loudspeaker) strongly influences the listening experience. These effects, therefore, need to be minimised or eliminated in listening sessions.

One observation the author made during participations in listening tests was that reporting about a certain perception can exaggerate the facts each time a report is presented. This seems to be comparable to the fish that grows each time the fisherman exaggerates his story of catching it. Also, the target group could influence the way of reporting, so effects might be presented at a larger scale to an inexperienced target group than to an experienced one. In tests with people everyone tries to be well received by others and therefore facts could be reported in an exaggerated or biased way towards a desired or supposed result (Leary, 2007, pp.55–57). A proper selection of subjects for a listening test is therefore crucial for valid results. Over the years it was identified that for the investigated topics experienced participants were needed. In the field, they are called 'golden ears' because their way of listening is trained and experienced. Doing tests with inexperienced people leads to poor results; in addition, they lack the special vocabulary required to describe results. However, this group can be useful in gaining facts about 'average' listening habits, which can be important for certain research.

The so-called 'golden ears' are hard to find but a few subjects could be sufficient to get valid results from a listening test. Bech (1990, p.109) discovered in his research about "Listening Tests on Loudspeakers" that "… a trained subject can replace up to 7 untrained subjects if only statistical aspects are considered."

The author's listening tests are normally used to validate a result that was initially observed by the author and then measured physically. For this, the author sought experienced 'golden ears', as they were able to express their impressions in perceptual measurements with the necessary vocabulary. The feedback from the subjects was given verbally based on the 'thinking aloud' methodology that was found to be "… a valuable source of data about the sequence of events that occur while a human subject is solving a problem or performing some other cognitive task" (Ericsson and Simon, 1981, p.10). The author noted that feedback and acted accordingly. As an example, when looking for confirmation of the proper delay time for time alignment the author adjusted the delay time of a pair of loudspeaker drivers while listening to the subjects' feedback. The subjects were not able to see the set delay but could only hear its possible impact on a stereo system with music played on it. According to their verbal feedback, the author determined the delay time corresponding to the auditory attribute that was sought. This process was repeated with multiple subjects to get confirmation.

3.1.5 Evaluate

The evaluation stage is at the very end of each research cycle. Its goal is to provide a recommended course of action for the next cycle or the next researcher. At least the outcome should be an idea about the successive possibilities. At the end of this phase the researcher could either present the results to the public or start over with the next cycle.

This is a quite critical phase, as it clearly shows whether the preceding steps resulted in valuable data for evaluation. In technical terms this could be compared with the problems of a measurement. Only if the measurement sequences are executed properly is a usable evaluation possible. At this stage it all comes together: have the research, build and test phases been successful or not?

Sometimes modifications of the built prototypes will be necessary in order to get more usable results. This seems to be quite normal; research can be planned as accurately as possible – at the end there still remain some uncertainties the researcher has to address. This can happen either through modification or through the limitation of results. These problems present themselves during the evaluation phase and it is important that the researcher observes them and assesses their impact. Modifications and changes can be made as long as they are reasonable and reported to the reader.

If results come out clearly in the evaluation phase, the preliminary stages have been executed properly. Furthermore, if all steps have been reported in a transparent and comprehensible way, all conditions have been met for generating primary literature: "The most original source for scientific papers is a self-conducted investigation or examination" (author's translation) (Theisen, 2008, p.89).

3.2 Handling unexpected results

Every element of research undertaken not only delivers expected results but also some unexpected ones, which creates more questions. The deeper the investigation is, the more likely is the appearance of new problems and questions. The same applies when extremes are used; for example, to bring systems to their limit. The fact that this happens is not new to the experienced researcher but it is important that when new facts are revealed, the researcher should draft new questions and research proposals. These new questions should be discussed with other researchers in the field for verification before publication. Perhaps someone else is already researching in this area or could help to complete the research questions. A single researcher will not be able to resolve all problems alone, and as soon as this is clear, it is time to find someone to conduct further research. Ultimately, it is the result that matters but not the name of the researcher who undertakes it.

3.3 Summary

As noted above, the presented methodology has been used as it is agile and relates to real-world situations; and it can therefore be applied to all of the author's research projects. It has some commonalities with action research, as practitioners participate in the research cycles and there is always some kind of practical result to be tested and evaluated. Unique to the author's methodology is the built phase that is needed to test the validity of the results in practice and verify them with the support of practitioners in the field. It is an elaborate methodology that only becomes feasible in combination with expert knowledge and a great deal of experience in manufacturing and producing hardware and software prototypes that could be used for the practical tests.

The last step in this chapter is to decide how the research projects could be presented. It was decided to present all research in the form of a research story, as this gives the reader the possibility to experience the research best:

> "This is a research story, that is, it describes what you did - the actions you took - and why you did them. The element of explanation transforms the story from storytelling into research. The explanatory elements include saying what the issue was that you wished to investigate; why this was an issue; how you formulated the research question; how you monitored practices; how you gathered data and analysed and interpreted the data; how you came to provisional conclusions; and how you tested the validity of your emergent knowledge claims." (McNiff, 2016, p.64)

The following stories will contain specific aims, additional context, the research-specific application of the methodology, its implementation and the findings.

4 Review of published works

4.1 Research Area 1: Improving acoustic modules to increase efficiency in the acoustical treatment of control rooms

The first research area concerns room and studio acoustics and the development of new acoustical modules. This work is a significant point of origin for further research and it contributes to original findings in the field. The research was undertaken between the years 1993 and 1995. It was carried out at the company ACM (Acoustic Consulting Munich). Under the supervision of one of the CEOs, it was possible to test and try new concepts, propose new designs, combine different mechanisms into multifunctional modules, and build and evaluate new prototypes.

4.1.1 Aims

The aim of this research was the development of new, innovative ways to measure and evaluate acoustical parameters of rooms and acoustical modules. Furthermore, the research focussed on the combination of existing mechanisms for room treatment such as porous absorbers, resonators and diffusors. Finally, prototypes for more efficient acoustical modules were developed. These modules are designed to either absorb or scatter soundwaves over a wider frequency range and therefore control the room acoustics more efficiently than traditional acoustic modules. A better control of the room acoustics improves the localisation of stereo signals and phantom sound sources and thus improves the listening experience.

4.1.2 Process

4.1.2.1 Development of a method for measuring resonators

For a planned series of tests on new acoustical modules some new methods for measuring resonators were developed. At the time of the research, personal computers were not feasible for measurement purposes, as they had neither an acceptable (if any) audio interface nor enough computing power for complex real-time calculations. One option the company chose was the measurement system 'TEF[24]' from Crown (Crown International, 1997) that used TDS measurements (see subsec-

[24] TEF stands for "Time-Energy-Frequency" – a measurement system initially developed by Richard Heyser (Davis and Davis, 2015, p.24).

tion 2.2.1.2), but this system was rarely available as it was used very often in the field for on-site reverberation measurements. An alternative was using a digital storage oscilloscope in combination with mathematical software on a computer. Available was a 50MHz digital storage oscilloscope with averaging functionality and a serial PC interface (Fluke, 1993). Windows 3.11 was the current operating system in 1993 (Markowski, 2016) and Mathcad was available for calculation tasks. With this combination impulse and step response measurements were possible.

One task was to develop a way to accurately measure the resonance of Helmholtz resonators[25]. A set of small test resonators (about 10–30 litres) was built in order to compare the theory with the practice. From the formulas for these resonators, it was clear that there was an uncertainty: the so-called 'mouth correction' that adds to the measurable length of the port in the resonator. For measuring the resonance frequency, the author decided to measure the resonance inside the resonator. Thus, a small hole was drilled in the upper board of the resonator and a microphone was inserted in an airtight arrangement. The microphone then captured the alternating pressure in the resonator's cavity. For the excitation of the resonator a signal had to be found that provided enough bandwidth and energy to excite the resonator. For this purpose, an impulse/step generator was built (see Figure 10). It contained a logic chip together with a resistor and a capacitor generating a repeatable stimulus. Either a short impulse with 10µs or a longer one with 50ms could be generated. The latter provided the step for measurements where more energy was needed (see Appendix B for the difference in energy between an impulse and a step).

Figure 10: The author's measurement set from 1993: Top: A small Sennheiser electret microphone built into a thick pencil. Left: Mic preamp with DIP-switch-programmable gain. Right: Impulse/step-generator with trigger output.

---

[25] A Helmholtz resonator is a cavity with a port (e.g. a bottle). The volume in the cavity together with the air mass in the port results in a mass-spring system called a Helmholtz resonator after its inventor Hermann von Helmholtz.

Although a short impulse would theoretically provide an impulse response where the resonance should be clearly visible, this signal proved to be unusable in practice. The problem was that the low amount of energy in an impulse together with the 8-bit resolution of the digital storage oscilloscope did not lead to any reasonable results.
It was therefore decided to employ a step stimulus. In fact, this is also an impulse but the rising and the falling edge are far apart from each other so that each step can be monitored independently. Another benefit of the step was its frequency distribution, which was suitable for the low-frequency Helmholtz measurements (see Appendix B on page 97). The step was applied to a loudspeaker in front of the resonator and the step response was captured by the microphone. This signal provided a well readable response which could then be further analysed (see Figure 11).

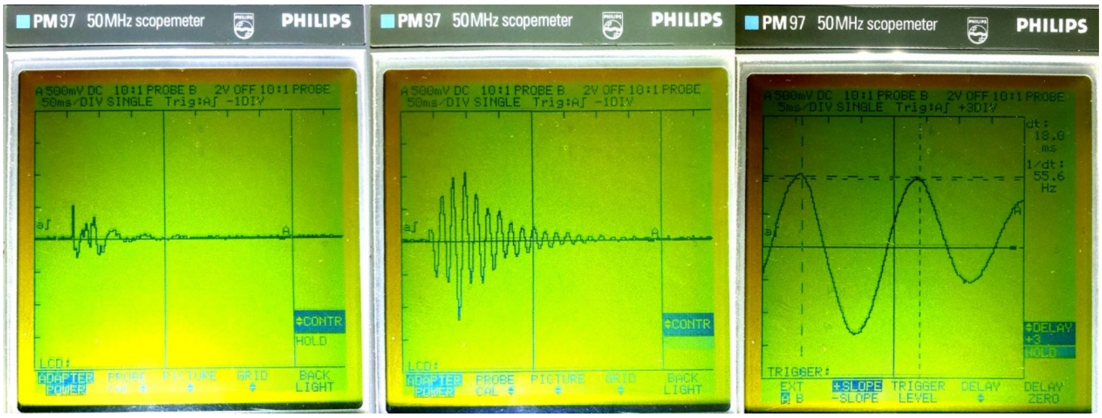


Figure 11: Measurement of a Helmholtz resonator with a step generator. Left: Microphone in front of the loudspeaker without Helmholtz resonator. Middle: Microphone plugged in the measurement hole of the Helmholtz resonator. The resonance and its decay can be seen. Right: Zoom into the resonance to determine the resonance frequency (56Hz) with the oscilloscope cursors.

These measurements were a foundation for all Helmholtz resonators the company had to plan. They did not only show the resonance frequency, but also the decay of the resonance, which indicates the quality (or Q-factor) of the resonance. It was found that the more accurate a resonator was built the better the Q-factor and the longer the decay. As these resonators were intended to absorb low frequencies and reduce the reverberation time it was crucial that they did not have too long a decay, as this would produce the inverse effect: prolonging the reverb (see subsection 2.2.1.1 at the end). This happens especially with large resonators that are built very accurately. A series of measurements with the step generator led to the solution that a flow resistance[26] could be attached over the holes of the resonator in order to decrease the Q-factor and shorten the decay.

[26] In this case one or two sheets of cloth.

A side effect of this research was the finding that capturing resonance frequencies of other materials is possible by putting the microphone close to the resonating surface and then tapping on the surface. Although the decay of these resonances could not be seen accurately, the resonance frequency was easy to read from the recordings of the storage oscilloscope. This led to a way to accurately determine the resonance frequencies of unknown spring-mass systems.

4.1.2.2 Development of software for transferring and evaluating TDS waterfalls

One of the regular tasks for an acoustic consultant is to measure and evaluate the reverberation time of rooms, either for ascertaining the status quo of a room that is to be treated acoustically or to determine if any treatment was successful. Using a sequence of TDS-measurements, a waterfall diagram can be plotted that shows the frequency-dependent decay of the room. By cutting the waterfall in slices of equal frequency bands, averaging these bands and then virtually flipping the averaged slices to the side the reverberation time can be measured. Data for these waterfall diagrams are stored in the form of numbers that represent the energy at a certain time and frequency. The available measurement system (TEF) could either export these numbers as raw data on 5¼" discs or perform an automated pre-evaluation that produces a sequence of levels on a printout. The latter was preferred by the CEOs but it involved additional estimations and a manual transfer from the printout to an Excel spreadsheet which was time-consuming and introduced uncertainties.

The author decided to develop a more user-friendly version for reverberation analysis. For that, the format of the exported raw data was investigated. It was in the form of complex numbers so it could represent both frequency and phase of the measurement. A BASIC program was written that could read the raw data and convert it for a graphical presentation into the virtually flipped sliced as described above. It was now possible to see the room response much more intuitively than in the form of a list of numbers. For evaluating the measurement, a straight line was implemented that could be altered with the keyboard in height and gradient. It then was possible to align this line on the measurement and accurately determine the reverberation time[27]. A set of measurements in third-octave bands was the outcome of this evaluation tool (see Figure 12 and the measurements in Appendix C on page 99).

Subsequently, all reverberation measurements were made with this software as it was a fast and user-friendly tool for evaluation purposes. Two years later, the brother of the CEO ported this software to a Pascal-based Windows application and modified the communication protocol to work with the next generation of this measurement system (Veith, 1995).

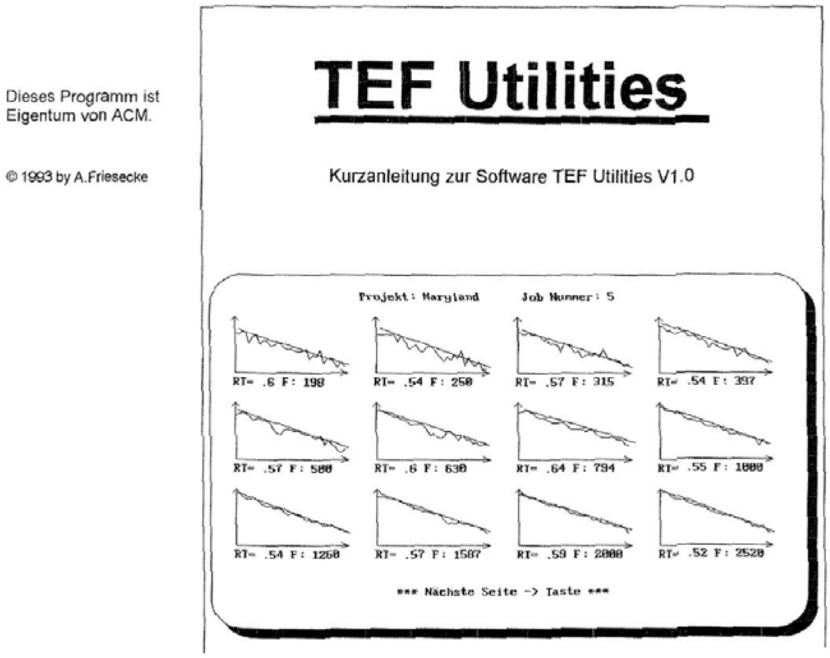


Figure 12: Copy of the first page from the manual for the evaluation software "TEF Utilities". The third-octave filtered reverberation times with the aligned lines can be clearly seen.

4.1.2.3 Testing and designing a new full-range sound absorber

Normally, porous absorbers are used for attenuating reflections. Their efficiency is largely dependent on their thickness (see subsection 2.2.1.1). In order to increase the efficiency of a porous absorber it was planned to combine it with a resonance absorber. It is envisaged that this new hybrid absorber would have better absorption rates at lower frequencies while keeping the thickness of the absorber unaltered. The resonance measurements from subsection 4.1.2.1 were an important basis for this research.

[27] The reverberation time is the gradient of the measured level versus time. It is normally determined in octave bands but for studio purposes in third-octave bands, which leads to about 20-25 results in the frequency range of interest.

The idea was to glue a thin fibreglass mat on the room-facing side of a block of fibreglass wool in a porous absorber. The fibreglass mat would act as a mass but in contrast to the regular membranes still be permeable for sound waves to a certain extent. The fibreglass wool would act as a spring, but in contrast to an air cavity it would absorb the sound passing through it.
In a series of tests different fibreglass mats and fibreglass wools were tested. Ways of glueing them together in a reproducible way were examined, as the glue changed the mass of the mat. The spring-constant of the fibreglass wool could only be estimated and therefore it was necessary to measure the resonance frequency. With a microphone close to the sheet and an oscilloscope it was possible to measure the resonance. The resonance was excited with a finger tap onto the mat (see subsection 4.1.2.1 at the end). The final result was to apply spray-on glue to the fibreglass wool and then attach the fibreglass mat to it[28].

The new absorber modules were then tested in an echo chamber (see subsection 2.2.1.1) and their efficiency was evaluated with the method shown in Appendix C and using the new evaluation method from 4.1.2.2. The result was better than expected and provided new knowledge in the field: with modules of the same thickness an effective absorption became possible that was about one octave lower than with traditional modules. These modules were able to attenuate reflections over a wider frequency range and therefore extend the reflection-free listening zone towards lower frequencies and thus improve the listening experience. An internal guideline with instructions for building the compound absorption material was written and provided to the carpenters that built these modules.

As the last step, a design for these modules was proposed. Not wanting to have the standard flat rectangular shape a variation that was more expensive to manufacture with large bevels at the front was proposed. This design had the advantage that the fibreglass mat was not dampened by the front lining of the absorber. Examples of the manufactured absorbers can be seen in Figure 13.

[28] More details about the materials cannot be given here as they are the intellectual property of the acoustic company.

Figure 13: Bevelled broadband absorbers in the back and Helmfusors (see next section) on the right of the picture at a TV studio in the Sendezentrum München. Although these absorbers are only 12cm deep they absorb sound as efficiently as 20cm deep absorbers. Photo: Sendezentrum München (SZM).

4.1.2.4 Testing and designing a new diffusor-absorber type

The fundamental work on measuring the resonance of Helmholtz resonators (see 4.1.2.1) enabled the creation of a new type of diffusor[29]. The CEO of the company came up with the idea that a diffusor could be used in an additional way. Since it has unused cavities behind its wells one might use them as Helmholtz resonators to absorb and scatter lower frequencies.

Investigating the possibilities of such a diffusor was planned in two stages. In the first stage the cavity-resonator part of the diffusor was built to test individual resonance frequencies and possible frequency ranges. It was important to extend the range of the diffusor towards its lower end and to design the resonators in a way that they resonate evenly spread in a range of about one octave below the lower corner frequency of the diffusor.

Once the resonator design was ready it was necessary to build a prototype for doing measurements with it. A large-format diffusor (1 by 2 metres) was built and all the special requirements in the construction were noted for writing a guide to prospective carpenters that might have to build these elements later on. Most important was the airtightness of the whole construction, which is not important in the standard diffusor construction. Holes for the ports were drilled and metal ducts were inserted in order to get high Q-factors for the resonators. Further, a set of pegs was prepared

[29] A diffusor is an acoustical module that scatters a sound wave. Today's standard is a 'quadratic residue diffusor' that is built from many wells of different depths. Depending on the width and depth of these wells a diffusor has a lower and an upper frequency that limits the effectiveness of the diffusor (Everest, 2001, pp.293–296).

for closing the ducts for comparative measurements. Pieces of cloth were at hand for covering the ports in order to lower the Q-factor of the resonators.

The diffusor was measured in an echo chamber with pegs attached, with open ducts and cloth-covered ducts. The result was that the diffusor absorbs frequencies in the range of the resonators quite well. It also scatters sound down to a lower frequency than without the ducts. For optimal efficiency both in absorption and diffusion the ducts had to be covered with cloth. These covers were later attached to the inside of the resonators on the models built in series. Production examples of these new modules can be seen in Figure 13 and Figure 14).

As a result, another hybrid module – the 'Helmfusor' – was created that was then built as the standard replacement for the traditional diffusors as it worked extremely efficiently without being any larger. As it partly absorbs and partly scatters low frequencies, it reduces hard reflections towards lower frequencies more efficiently than a traditional diffusor and therefore improves the listening experience.

Figure 14: Helmfusor implementation at the back of the large studio of SAE Berlin (opened mid-2016). Photo: SAE Institute GmbH.

Unfortunately, all measurements regarding these new modules were lost during the upgrade of the measurement system from TEF 10 to TEF 20 in 1994. As the TEF 10 used CP/M formatted 5.1/4" disks and the TEF 20 was running on a Windows laptop with DOS-formatted 3.5" disks, it was too elaborate to transfer the data; therefore, no data from that time exists in electronic form. Printouts from that time are also not accessible, as the company ACM filed for bankruptcy in 1996. The author kept some printouts from this time in personal files but these were only about loudspeaker measurements and not about the acoustical modules.

4.1.3 Publications

This subsection expands section 1.3 where a brief description of the published works has already been given. It lists the specific sections of each publication that contain the research findings.

- ***Book:** "Die Audio-Enzyklopädie" ("The audio encyclopaedia"); 850+ pages, first published 2007, 2nd ed. in 2014; Publisher:* ***DeGruyter***
    - **Subsection 1.6.2.2, pp. 72–73 - Measuring Helmholtz resonators:** These two pages describe how Helmholtz resonators can be measured. The method with the step signal from 4.1.2.1 is presented. Also, the measurement setup with the microphone in the resonator is shown in a picture.
    - **Subsection 1.6.6, page 81 - Combined absorbers:** The research outcome from combining traditional porous absorbers with resonance absorbers from the development in 4.1.2.3 is presented and extended to a resonant absorber that can be supplemented with a porous absorber.
    - **Subsection 1.6.7.2, pp. 82–86 - Helmfusor:** These pages present the functionality of a QRD-Diffusor and extend on pages 84 and 85 the widely known basics with the research output from 4.1.2.4. A large Helmfusor is depicted on page 86.

- ***Book:** "Studio Akustik"; 190 pages, published 2007, 5th ed. in 2015; Publisher: PPV Medien*
    - **Chapter 12, pp.139–140 - Measuring RT60:** Different ways of determining RT60 are shown on these pages. The research output from 4.1.2.2 is presented in writing and as a figure on page 140.
    - **Chapter 16, pp. 173–175: Combined absorbers:** These three pages are dedicated to combined absorbers and present the research output from 4.1.2.3 and 4.1.2.4. complemented with some further ideas.
    - **Chapter 17, especially page 188 – Helmfusor:** This page is dedicated to present the research output from 4.1.2.4 in more detail.

- ***Series of articles** about acoustics: Magazine: "Recording Magazin", 2006-2010; Publisher PPV Medien*

  Articles about: Reverberation, Absorbers and Helmholtz resonators: The article "Wie trocken darf's den sein" presents the research output from 4.1.2.2 and describes the need for slicing waterfall diagrams for third-octave reverberation analysis. Detailed results were presented on the website of the magazine where the printed article directly refers to.

  In the article "Wenn der Schall in die Falle geht" there is a section describing porous absorbers being covered by foil to suppress high-frequency absorption. Further, it is pointed out that this foil could act as a resonator. Both are related to the research in 4.1.2.3.

  In the article "Schall und Gegenschall" the measurement setup for Helmholtz resonators as an outcome from the research in 4.1.2.1 is shown and the finding about lowering the Q-factor of resonators by applying sheets of cloth also from the research in 4.1.2.1 is presented.

4.2 Research Area 2: Measuring time alignment errors, testing their impact on the listening experience and providing solutions for time alignment of loudspeakers

The second research area is of particular importance to the contextual statement, as the research started in mid-1993 and continued until 2004 when a new method for electronic time alignment was presented to the public. A modification kit – 'Time Alignment Rings' – for the famous studio loudspeaker NS10 from Yamaha was developed as a result of the research findings. Multiple drivers in a loudspeaker may emit the signal at a different time. This can happen due to the mechanical arrangement of the drivers when drivers are mounted on a flat plane but the acoustical centres of the drivers are shifted from that plane. For coaxial[30] drivers the tweeter might be either in front of the woofer, on the same plane as the woofer or behind the woofer, depending on the mechanical construction. Additional time delays can occur from phase shifts of the crossover networks in the loudspeakers.

4.2.1 Aims

The first aim of this research was to find a reliable way to measure the drivers of multi-way loudspeakers in order to identify possible time delays. A further aim was the development of solutions for time-aligning drivers in both active and passive loudspeakers. Results were verified through listening tests by compensating the measured delays and observing the impact of the correction on the listening experience.

4.2.2 Process

4.2.2.1 Aligning an active coaxial system

The first time alignment was developed for a set of large active loudspeakers that were designed and built by the acoustic consultancy company ACM and installed at the Mastering Studio Munich. These loudspeakers were active three-way loudspeakers with a coaxial mid-high-frequency driver. Together with 12" or 15" woofers in a filtered vented box these loudspeakers provided a flat frequency response from 20Hz to 20,000Hz. One benefit of the coaxial mid-high-frequency driver construction was its point source for mid and high frequencies that should have led to excellent

[30] A coaxial driver is a driver where for example the tweeter and the midrange driver are behind each other and therefore emit their sound from the same location.

localisation and depth imaging of phantom sound sources. The coaxial driver had an 8" midrange driver and a 0.5" tweeter mounted behind the magnet of the midrange driver. The tweeter was configured to drive a small pressure chamber horn that guided the sound through a hole bored in the centre of the mid-frequency driver and emitted it with a controlled polar pattern (see Figure 15). The crossover frequency was 5,000Hz. It was assumed that the high-frequency driver would emit its sound later than the mid-frequency driver. The potential time delay, as well as the audible impact of that delay had to be investigated.

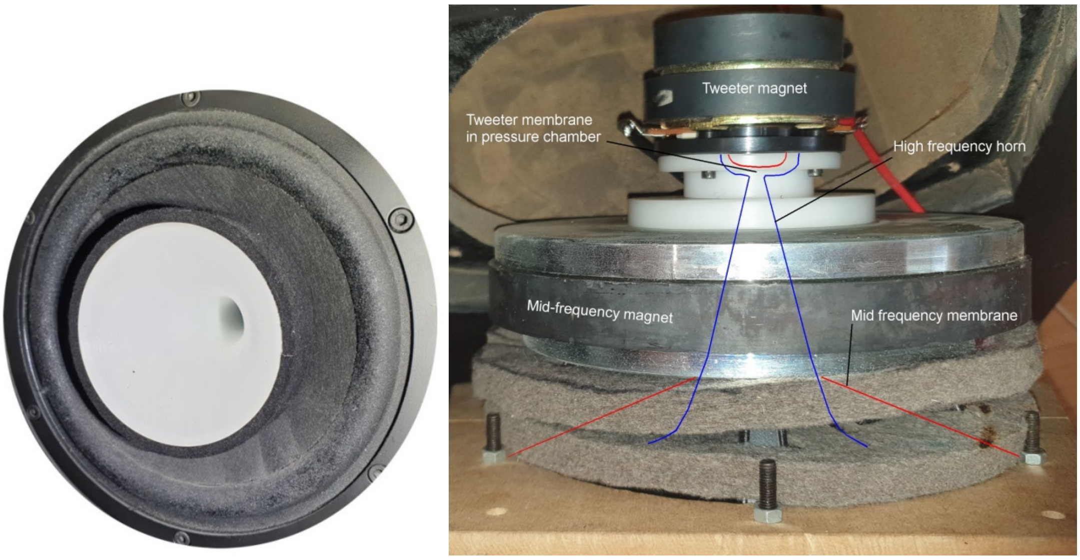


Figure 15: Coaxial mid-high-driver. Left: The mid-frequency driver with a white horn tweeter in the middle. The horn has a black felt on the outside to absorb reflections from the mid-frequency membrane. Right: The horn driver is behind the magnet of the midrange driver. High- and mid-frequency membrane positions are shown.

As a first step the time delay between the midrange driver and the high-frequency driver had to be determined. The ETC[31] measurements from the TEF system only had a resolution of one sample. With 48kHz sampling rate this leads to 20.8µs steps. Phase- and group-delay measurements from the TEF, therefore, seemed to be too inaccurate. Therefore, a step generator and a storage oscilloscope (see subsection 4.1.2.1) were used for the measurements. The step generator had a trigger output for providing a time reference. The oscilloscope was triggered on the rising edge of the step from the generator and it recorded the sound received by a microphone placed at the listening position with 100kHz sampling rate (10µs resolution). This measurement was repeated for the midrange driver and the tweeter and both

[31] ETC stands for Energy Time Curve. It is a measurement where the energy emitted from a sound source is plotted against time (see subsection 2.2.1.2).

measurements were overlaid on the screen. Using 'distance cursors'[32] of the oscilloscope it was possible to accurately determine the time difference between these two drivers. A time difference of 100µs was determined. As expected, the signal from the tweeter arrived later as the tweeter was behind the midrange driver in this construction.

As the next step, the impact of this delay had to be evaluated. There was no existing literature describing the impact of two delayed sound sources being emitted from exactly the same position with such a short delay time. Therefore, a listening test was set up. The idea was to delay the midrange driver and virtually shift its sound source backwards to match with the one from the tweeter. As the test system was an active loudspeaker system[33] it was possible to insert a delay between the midrange output of the crossover and the amplifier for the midrange driver. Two digital delays (Lexicon 224) were used to align the midrange drivers to the tweeters. The successful alignment was controlled with the oscilloscope. Fortuitously, the professional delays of the early '90s were able to produce such short delay times. More modern ones have quite long latencies, for example, the Yamaha D5000 from 1998 has a minimum propagation delay of 2.8ms (Yamaha, 1998, p.29).

Both midrange drivers of a stereo setup were equipped with variable digital delays. By increasing the delay time, it was now possible to virtually move the midrange drivers step by step backwards in order to align them to the tweeters. By listening to music and changing the delay time back and forth, the author identified a significant improvement of the stereo image at around 100µs delay time. To get confirmation five 'golden ears'[34] were invited to participate in a perceptive listening test. They had to sit in the listening position of the control room and listen to music both in stereo and mono as well as to pink noise in mono. The author adjusted the time delays and the subjects had to immediately report their perception. Without seeing the set time delay, they concordantly mentioned the improved stereo image at 100µs time delay for stereo signals. For mono signals, the subjects reported an improvement in sharpness of the phantom sound source at 100µs time delay.

---

[32] Digital oscilloscopes mostly have two movable cursors and can measure different data in between them. Obviously it is the distance of the cursors (time difference); other data like frequency, voltage etc. vary from system to system (Fluke, 1993, p.6-24).

[33] Passive loudspeaker systems have one power amplifier and the crossover behind the amplifier. Active loudspeakers have the crossover first and then one power amplifier for each driver (Henricksen, 1991, pp.560–565).

[34] There were participants from the mastering studio and from the acoustic consultancy company. All had more than 5 years professional listening experience.

Compensating the timing error between the midrange driver and the tweeter led to an effect where the stereo image seemed to 'snap into place'. Phantom sound sources became more precise and the depth of the image opened up. Comparing the delayed version with the non-delayed version by activating the bypass function on the delay made it clear: without delay the sound was smeared in the range of the crossover frequency[35] of the midrange driver and the tweeter. As the loudspeaker was a custom-made model for the mastering studio, the mastering engineers were asked for their preference using an affective listening test. They unanimously voted for the time-aligned version. It became clear that the 100μs measured with the step generator was correct. The delay of the midrange drivers was finally implemented with analogue all-pass filters in the crossover.

It was further noted that the effect of aligned drivers emerges even when the alignment is close to the perfect value. Having ±10μs offset from the ideal value did not lead to audible differences. In a further test, the timing error between the woofer and the midrange driver was corrected as well and the perceptive listening test was repeated. There was no comparable effect perceived by the subjects, so an alignment of the woofer to the midrange driver was not followed any further.

4.2.2.2 Aligning a passive studio loudspeaker

Inspired by the results from the active loudspeaker, further attempts were made to reproduce the effect with a two-way passive studio loudspeaker. It was planned to observe the impact of time alignment on a loudspeaker where low- and high-frequency drivers are not mounted coaxially but side-by-side. The loudspeakers chosen were a pair of "NS-10M Studio" from Yamaha – one of the most renowned loudspeakers in the world (Ward, 2008). Time aligning these seemed to have the highest relevance to the industry.

Timing errors in these passive loudspeakers could happen for two reasons:

- The woofer is expected to be late as its voice coil and therefore its acoustical centre is mounted deeper inside the box than the voice coil of the tweeter. For the NS-10M there could be a 2.5–3.5cm mechanical difference. The exact distance is not visible without dismantling the drivers.

[35] The crossover frequency of that system was at 5,000Hz. At this frequency midrange driver and tweeter complement each other. Therefore the effect is most audible at the crossover frequency.

- The crossover could introduce additional delays. For the NS-10M there is a crossover with second-order filters. By simulating this crossover its phase-error between tweeter and woofer can be calculated for the crossover frequency. This simulation led to a phase shift of 54° at 2,150Hz, which corresponds to 70µs time delay or 2.4cm distance error (see Appendix E on page 113).

In combination, the mechanical plus the electrical time delay could lead to a distance error of approximately 4.9–5.9cm. To get an accurate result it was decided to measure the prospective timing error. These measurements were far more difficult than for the active system as in passive loudspeakers the power amplifier drives the crossover and behind that, the signal is split into low and high frequencies. For measuring the woofer and the tweeter separately it was necessary to remove the respective other driver. A timing error of 160µs was measured using the same measurement technique as before (see Figure 16). As expected, the woofer was too late. Using the speed of sound in air 160µs correspond to 5.4cm distance.

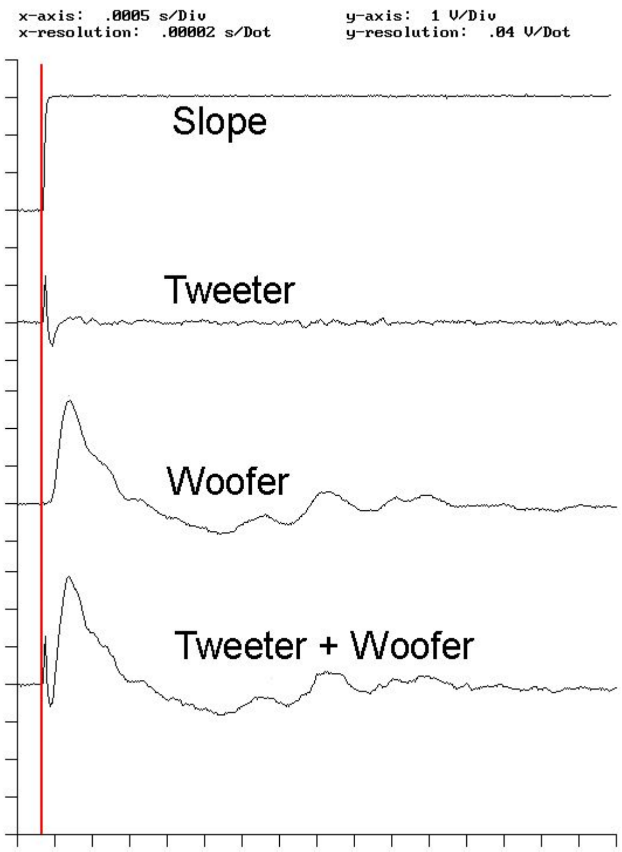


Figure 16: Composition of measurements for the NS-10M studio loudspeaker: From top to bottom: Step signal from the step generator (aligned to red line for reference), signal from high-frequency driver only, signal from low-frequency driver only, signal from the whole loudspeaker.

As the effect of the stereo image 'snapping in place' was very evident with the active coaxial driver, it was planned to have a comparable setup with the passive loudspeakers. An adjustable delay for two loudspeakers (stereo delay) was needed that could be operated in real-time while listening to music. Using the digital delay from the studio was not an option, as the passive loudspeakers had no separate inputs for the woofer and the tweeter. For the test, a second pair of NS-10Ms was borrowed that were built in the same year and had a comparable frequency response. In the first pair the tweeters were disabled; in the second pair the woofers were disabled. The second pair with the still-active tweeter was put straight on a table. The first pair with the still-active woofer was placed upside-down beside the second pair on two round pens so it could roll forward and backwards. In doing so, a 'split loudspeaker' was created that could shift the woofer back and forth in order to make the desired alignment. An elastic strap was attached to the back of the rolling pair that pulled the loudspeaker backwards. A cord was attached to the front that led to the listening position. Now the listener was able to pull on the cord and both woofers came closer to the listening position. Relaxing the cord led to a backwards movement due to the tension of the elastic strap (see Figure 17).

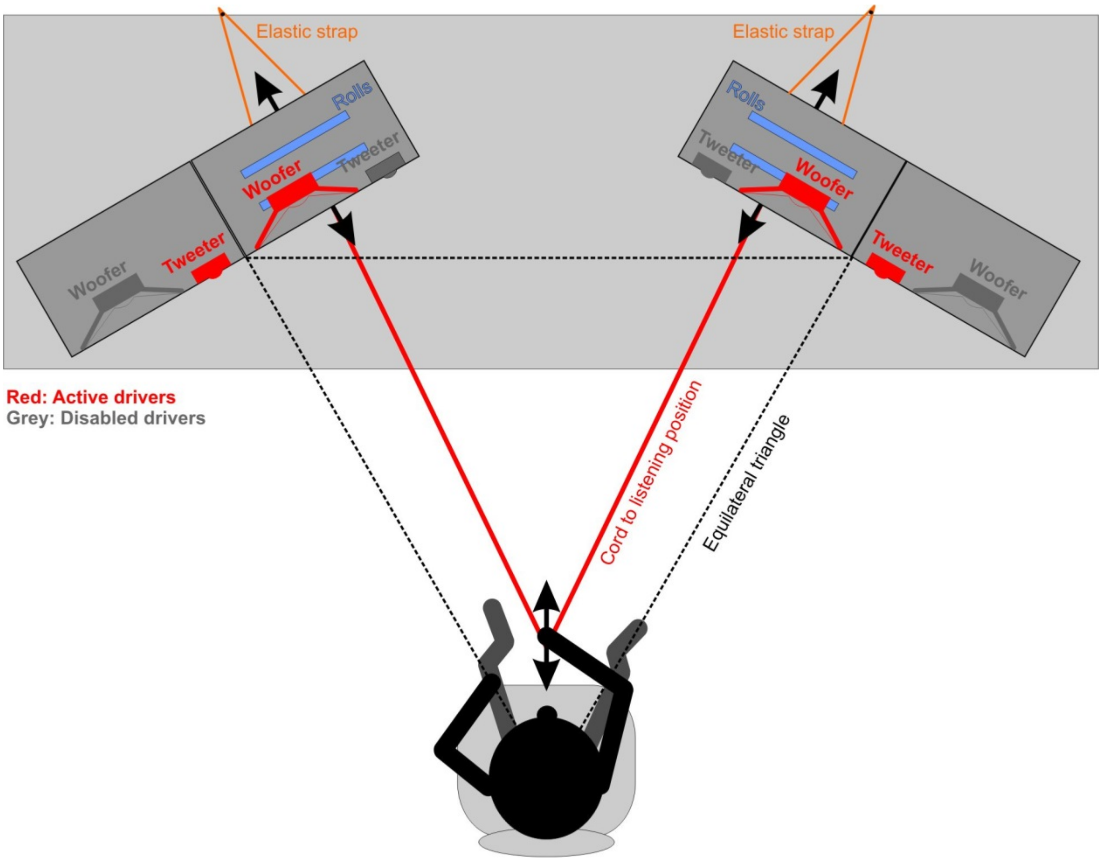


Figure 17: Test setup for time aligning the NS-10 loudspeakers. Two 'half loudspeakers' could be shifted against each other in order to align the position of the woofer to the tweeter.

To confirm the measured delay time, a perceptive listening test with multiple subjects (normal listeners and 'golden ears'[36], about 10 in total) took place. Normal listeners without much listening experience were included as the effect of the time alignment appeared to be more noticeable than with the active speaker from subsection 4.2.2.1. The subjects were placed in the listening position as shown in Figure 17. The test started with mono pink-noise played on both loudspeakers. The subjects could then roll the woofers back and forth and report their auditory sensation. Their task was to find the position where the noise appeared as sharp as possible in the middle between the loudspeakers. All subjects found this position quickly and thereby validated that the measured 5.4cm offset would be the right alignment. Listening to stereo music and repeating the listening test led to the result that the stereo image seemed to 'snap into place' with the correct alignment.

Having found the right offset for the woofer, wooden spacer rings were built to move the woofers of one loudspeaker pair outwards. Both loudspeaker pairs were restored to full working order and stacked one on top of the other. Using a changeover switch, the time-aligned loudspeakers and the original pair could be selected alternatively. With this setup another series of affective listening tests were conducted with more subjects[37]. The subjects who also participated in the test from subsection 4.2.2.1 noticed that the effect was more significant than with the active speaker. This was explained by the larger delay of 160µs that had to be compensated and due to the fact that the crossover frequency was around 2kHz and is therefore closer to the most sensitive frequency range of the human ear (Fletcher and Munson, 1933). The overall result proved the effectiveness of the time alignment. Even listeners without any knowledge of the audio field could immediately hear the difference. What is of even greater significance, they were able to describe the impact. While the non-aligned loudspeakers produced a flat and partly smeary image, the aligned ones had much more depth and showed precise phantom sound images. Furthermore, it was no longer possible to localise the aligned loudspeakers themselves – only their phantom sound images. Asked for their preferences, the subjects all voted for the time-aligned speakers. The loudspeakers were measured in amplitude and phase with and without the spacer rings. The results can be seen in Figure 18 as separate measurements and in Figure 19 in combination. Although the frequency response of the loudspeakers changed, it was decided that the improvement in localisation could outweigh these changes.

---

[36] These were the same as in subsection 4.2.2.1.

[37] These were the 'golden ears' from subsection 4.2.2.1, audio engineers from SAE Institute Munich, musicians and unexperienced listeners.

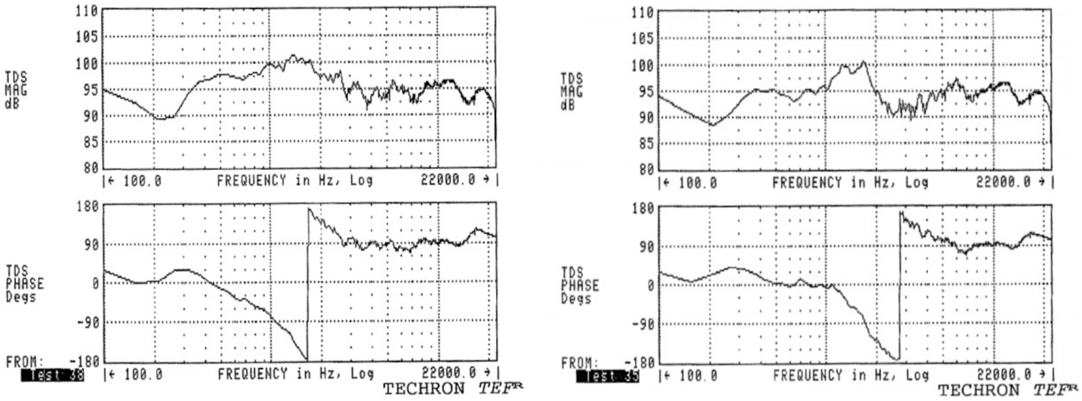


Figure 18: Measurements of the NS-10M without (left) and with (right) Time Alignment Rings.

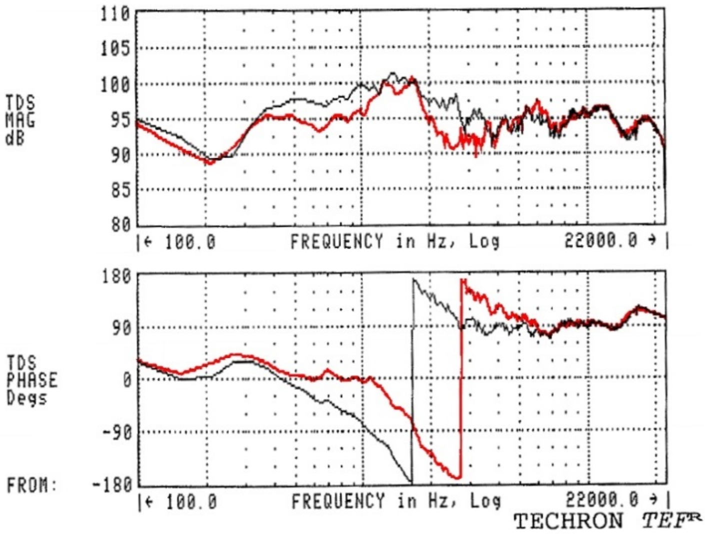


Figure 19: Comparison of the NS-10M without (black) and with (red) Time Alignment Rings.

It was therefore decided to build a modification kit for the NS-10M loudspeakers from Yamaha. A distance ring was designed (see Figure 20) and 1,000 pieces were ordered as an injection moulding. This time alignment set is still being sold by the author and although it alters the total sound of the NS-10M the author received largely positive feedback about the impact of the time alignment modification (see some examples in Appendix H on page 134). The modification kit can be seen in Figure 21 and in Figure 22.

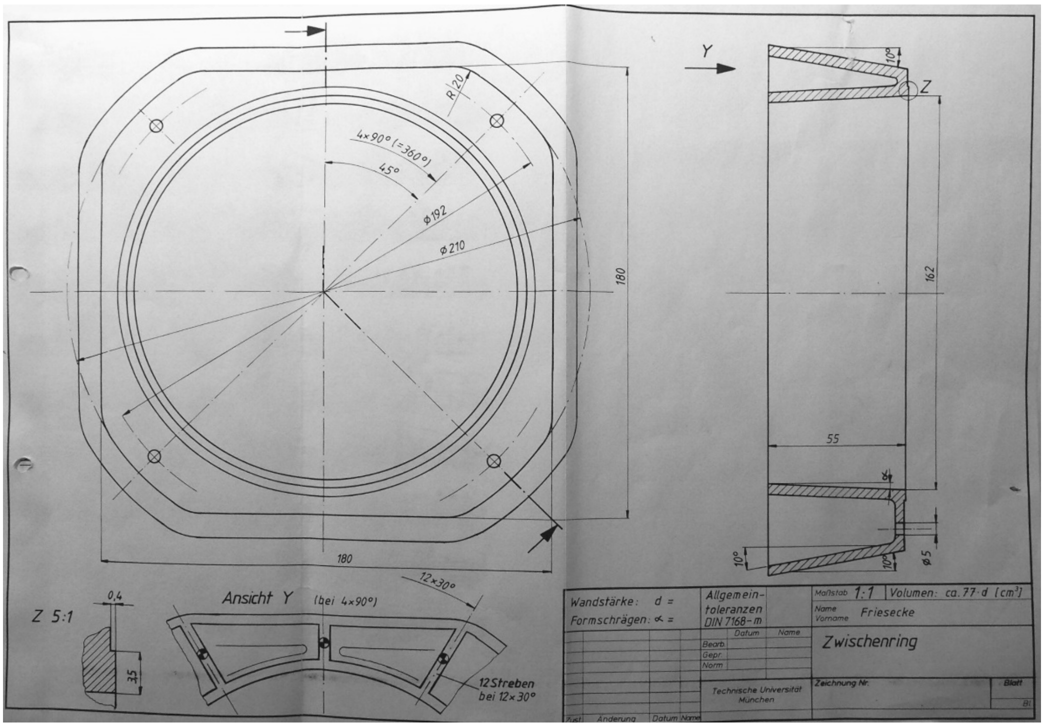


Figure 20: Technical drawing of the "Time Alignment Ring" for the injection moulding company (technical drawing by Falko Friesecke).

Figure 21: NS-10 loudspeakers with the Time Alignment Ring attached.

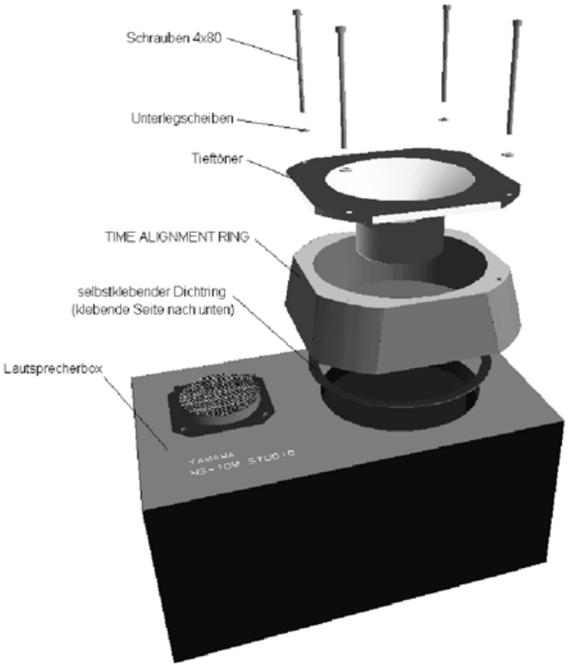


Figure 22: Modification kit as it was printed in the author's article about time alignment.

From the listening tests on pages 51 and 55 it could be seen that time alignment has a large impact on the listening experience. However, it was unclear why time alignment would not only alter the sound colouring of a loudspeaker but essentially alter the localisation of phantom sound sources. At the time of the research only a limited number of sources could be accessed such as the "Tonmeistertagungsberichte", the AES papers from 1990 onwards and some standard books about sound engineering from German, UK and US authors. The research from Long and Heyser (see subsection 2.2.2) were not known at this time. However, they did not contain an explanation of the impact that was observed in the author's research.

The research from Zwicky only looked at frequencies below 500Hz, the loudspeakers from Genelec aligned only the woofer, and for the Klark-Teknik loudspeaker "Jade One" no information was given on the stepped baffle (see subsection 2.2.2).

The explanation for the audible time alignment effect was finally affected by the research of Burchard, Irrgang and Andresen (1987). The author of this contextual statement found that there was a correlation between the signals from non-aligned loudspeakers with the pinna reflections reported by Burchard et al. Therefore, he claimed that a non-aligned loudspeaker produces artificial pinna reflections. If the human ear's localisation mechanism is deceived by a loudspeaker that produces two signals 160µs apart from each other, then the brain identifies this sound source at the side of the head. But in fact, the real source is only 30 degrees off to the centre and the pinna reflections for the true direction appear as well. As a result, the sound sources cannot be clearly positioned by the brain and seem to smear and broaden between the loudspeakers. This conclusion was published in the author's article "Phase Lifting" from 1994 (see 4.2.3).

4.2.2.3 Electronic time alignment

In mid-2003, the author decided to re-examine the time alignment topic again and investigate an electronic time alignment option for the NS-10M. Digital loudspeaker management systems like the Behringer Ultradrive DCX2496 (Behringer, 2013) were equipped with all functions necessary for time aligning loudspeakers. The aim was to realise a purely electronic time alignment that could be applicable to all kinds of loudspeakers – active or passive – *without any modification of the loudspeaker itself*. It should only modify or 'pre-process' the input signal. To apply this to the NS-10, it should delay the tweeter instead of moving the woofer outwards. As the tweeter in this passive loudspeaker is not accessible independently, a frequency-dependent pre-processing was necessary.

The digital crossover was used to split the signal into two frequency ranges that were divided exactly at the same frequency as the loudspeaker's internal crossover. Linkwitz-Riley filters with fourth-order were used, as these have the same phase response for both high- and low-pass filters, and the sum of the split outputs add to a flat amplitude-response. (Lancaster, 1994, chap.8; Team Elektor, 1995, p.43; Tietze and Schenk, 1980, chap.13). Although this leads to a phase shift of the total signal it does not introduce phase shifts between low- and high-frequency outputs.

A delay was inserted into the high-frequency path of the crossover and adjusted to the desired delay time of the tweeter (160µs). In the last stage these two signals had to be added again. The digital crossover did not offer a summing stage at its outputs, so the signals were added via a passive resistor network of two equal resistors integrated in a cable accepting a loss of 6dB in level. In order to correct any errors in the frequency response, equalisers were inserted in both paths. The whole alignment and equalising process was closely monitored with a measurement system. As the equalisers introduce time delays themselves, finding the right parameters for the loudspeaker alignment was an approximation process between the delay and the equalisers. The block diagram for the electronic alignment is shown in Figure 23. It could be realised either in analogue (the delay would then have to be an all-pass filter) or digitally.

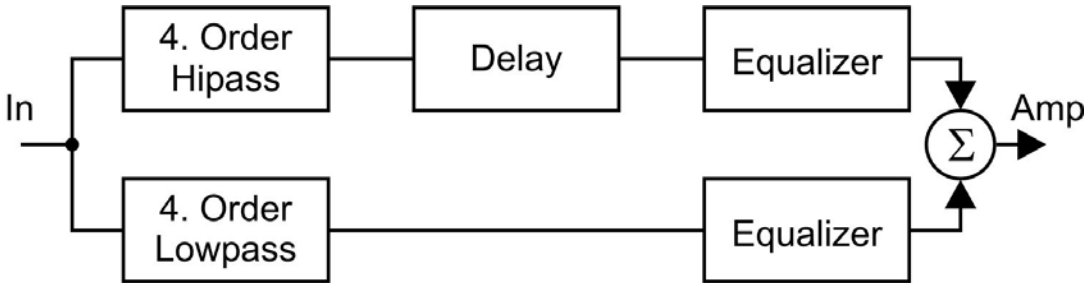


Figure 23: Block schematic of the electronic time alignment. This version is for delaying the tweeter, which is normally necessary.

The first test was conducted with a pair of unmodified NS-10 loudspeakers. The electronic time alignment had the same audible effect as the mechanical alignment. Using the equalisers, the frequency response was even better than with the mechanical alignment kit. This outcome was presented at the Musikmesse Frankfurt in 2004. A second test was carried out in a studio and using two Behringer loudspeaker management systems, a 5.1 system with active loudspeakers was aligned. The result was so effective that playing a mono signal over the left and right loudspeaker at the same time gave the impression that the signal was coming only from the centre-speaker. But in fact, there was no signal applied to the centre-speaker. A perfect phantom sound source appeared.

This research led to the development of a concept for an electronic time alignment that can be applied to all types of loudspeakers without changing any component of the loudspeaker itself. The full-range pre-processing produced results comparable to processing active loudspeakers solely in its high- or low-frequency paths.

4.2.3 Publications

This subsection expands section 1.3 where a brief description of the published works has already been given. It lists the specific sections of each publication that contain the research findings.

- ***Article:** "Phase Lifting" Magazine: "Studio Magazin", 05/1994; Studio Presse Verlag GmbH*
    - **Article about time alignment and its impact on the listening experience:** This article presents the research output from 4.2.2.2 and introduces the Time Alignment Kit to the market. It explains the im-

pact of two delayed signals from the same direction on the stereo image by localisation mechanisms of the pinnae and the brain.

- ***Book:*** *"Die Audio-Enzyklopädie" ("The audio encyclopaedia"); 850+ pages, first published 2007, 2nd ed. in 2014; Publisher:* ***DeGruyter***
    - **Subsection 5.6.1, pp. 138–139: Delayed sound sources from one direction:** This subsection explains principles of localisation and relates them to timing errors of multi-way loudspeakers. It presents the research output from 4.2 from the listener's perspective. This subsection refers to subsection 6.10.2.14 in the same publication (see below) for correcting timing errors.
    - **Subsections 5.5.1-5.5.2, pp. 333–335: Timing errors from crossovers:** The research output from 4.2.2.2 about the timing errors of crossover networks is presented in this section. It is shown that second-order crossovers that lead to a flat overall frequency response will have a delay between high- and low-frequency output.
    - **Subsection 5.8.1, p. 350: All-pass filters for time alignment:** The usage of all-pass filters for delaying loudspeaker drivers for time alignment was a result of the research in 4.2.2.1. In this paragraph, all-pass filters are recommended for time alignment of mid-range and high-frequency drivers.
    - **Subsection 6.1.2.2, p. 411: Phase response of electrodynamic transducers:** This subsection discusses timing issues related to phase shifts from crossovers, which are a research output from 4.2.2.2.
    - **Subsection 6.10.2.14, pp. 470–472: Mechanical and electrical time alignment:** The research output from 4.2 is summarised in this section. All three options for time alignment are presented and supported by diagrams.

- ***Book:*** *"Studio Akustik"; 190 pages, published 2007, 5th ed. in 2015; Publisher: PPV Medien*
    - **Chapter 11, p. 131: Optimising loudspeakers with time alignment:** In this section of the book the research output from 4.2.2.1 and 4.2.2.2 is presented and briefly explained. The Time Alignment Kit from the author is also shown in a picture so the reader can see how time alignment can be implemented.

4.3 Research Area 3: Using equalisation for correcting and shaping a loudspeaker's frequency response

The third research area concerns the equalisation of loudspeakers in different contexts. Research in this area started in 1993 with a comparison between different pre-processing options for loudspeaker correction. This developed to loudspeaker correction with multiband parametric equalisers and building a loudspeaker that was largely dependent on equalising for reaching a useful output. Loudspeakers are the outlets for the signal that the audio engineer listens to. They essentially influence the perception of the mix in the control room and therefore contribute to the result. Usually, a flat frequency response of loudspeakers is desired as this setting is as neutral to the sound as possible. Furthermore, a linear phase response is desired, so all frequency components of the signal arrive simultaneously.

4.3.1 Aims

This research investigated the effectiveness of equalisers on correcting the amplitude and the phase response of loudspeakers, for getting a flatter frequency response and a more linear phase. Digital (both FIR an IIR) and analogue filters were tested and evaluated. To demonstrate the effectiveness of equalisers an intentionally incorrectly designed loudspeaker was corrected with equalisers and the results were evaluated. The correction of the loudspeakers should not only lead to a flatter amplitude response but also to a more linear phase. The equalised loudspeakers should produce an overall neutral sound and have very good localisation in terms of phantom sound sources and depth imaging. Small loudspeakers should be equalised to sound like larger ones, so they could be placed as near-field loudspeakers and therefore have little interaction with the room acoustics. A secondary aim of this area was the investigation of how the frequency response of a loudspeaker could be shaped using equalisers so its response is beneficial for the mixing process.

4.3.2 Process

4.3.2.1 Programming a convolution EQ

Finite impulse response filters (FIR filters) are normally associated with linear phase filters and the literature quite often only explains linear phase applications. In fact, these filters do not necessarily need to be linear phase (see subsection 2.2.3.3). They could have almost any phase response and a non-linear phase should lead to better results than a linear phase. To test this, the impulse response of a loudspeaker was measured with the impulse generator from subsection 4.1.2.1 and a storage oscilloscope. The result was transferred into Mathcad and there it was Fourier-transformed. In the frequency domain the amplitude and the non-linear phase response were mirrored (that is, still non-linear) and the resulting frequency response was transferred back into the time domain. The resulting impulse response was used for the FIR filter. FIR filters can be calculated either in the time domain as convolution[38] or in the frequency domain as multiplication of real and imaginary parts of the spectrum. In this case the measured impulse response was convolved with the calculated correction impulse response.

The result was close to the ideal impulse delayed by the length of the measured original impulse response (see Appendix F from page 116, Figure F1). The remaining artefacts in the impulse response were quite likely a result of the noise that was in the measurement of the impulse response and that could not be corrected due to discontinuous amplitude and phase jumps.

Results that emerged from this research were:

- The correction is only valid for exactly one position in the room where the measurement took place. At a different position the result might become worse as the correction does not only correct loudspeaker errors but also any errors introduced from room reflections.
- Noise in the measurement leads to artefacts that remain in the corrected impulse response.
- There is a latency which is exactly as long as the measured impulse response. This makes sense as the phase was mirrored and therefore the correction impulse response must delay the whole signal. The longer the measured impulse response (e.g. when measuring down to low frequencies), the longer the introduced latency will become.

A further test was made with an artificial impulse response from an equaliser band that was set to boost. Calculating an IIR band-pass and adding it to a simulated flat signal led to this boost band. Its amplitude and phase response were mirrored and then used to calculate a correction impulse response via an inverse Fourier transformation. This (noiseless) impulse response was then fed into the correction programme for convolution. The result can be seen in Appendix F in Figure F2. The

[38] A convolution is a mathematical formula for "combining two signals to form a third signal" (Smith, 1999, p.107). Usually an impulse response is combined with an input signal in order to form the output signal (convolved signal).

correction led to a perfect flat amplitude and a flat phase response. This result confirmed the above-mentioned statement that a correction with a non-linear phase should lead to better results than one with a linear phase.

As one last test the correction of only the amplitude of a loudspeaker was undertaken. This equals the linear phase filter application. The amplitude of the measured loudspeaker was mirrored and the phase for the correction was artificially set to linear (and therefore noiseless) introducing only a delay into the correction signal. The result showed that although the amplitude was corrected very well there were still irregularities in the phase response of the corrected signal. The resulting impulse response was far from an ideal impulse and the overall result was therefore categorised to be worse than the one with the inverted measured phase response (see Appendix F, Figure F3).

4.3.2.2 Comparison of a frequency- and phase-corrected driver according to a patent from Peter Pfleiderer with the same driver corrected with a conventional fully parametric equaliser

In mid-1993, a new device for compensating loudspeaker errors triggered a test that was undertaken by the author. The discussed device provided a pre-processing of the audio signal in order to compensate driver deficits (Pfleiderer, 1987). It promised better phase correction than with other devices, and as a proof of concept a square wave was shown that was distorted by the driver. After applying the correction device, the square wave could be recognised again. As discussed in subsection 2.2.3.6, a square wave can be used for measuring loudspeakers but it is not a user-friendly test signal. The author developed a hypothesis that with a fully parametric EQ[39] it should be possible to do at least the same correction, if not better. This is mainly because an analogue fully parametric EQ should be versatile enough to adapt to all kinds of errors. Further, it should correct both amplitude and phase at the same time[40]. A series of measurements were conducted where the device from Pfleiderer (see Figure 24) was benchmarked against two fully parametric four-band equalisers (see Figure 25).

---

[39] In a fully parametric EQ the parameters centre frequency (f_c), quality (Q) and cut/boost (dB) can be set independently of each other. It is therefore the most versatile equaliser for professional studio use.

[40] This statement is based on the finding that the artificial equaliser band from chapter 4.3.2.1 could be undone by a mirrored equaliser band.

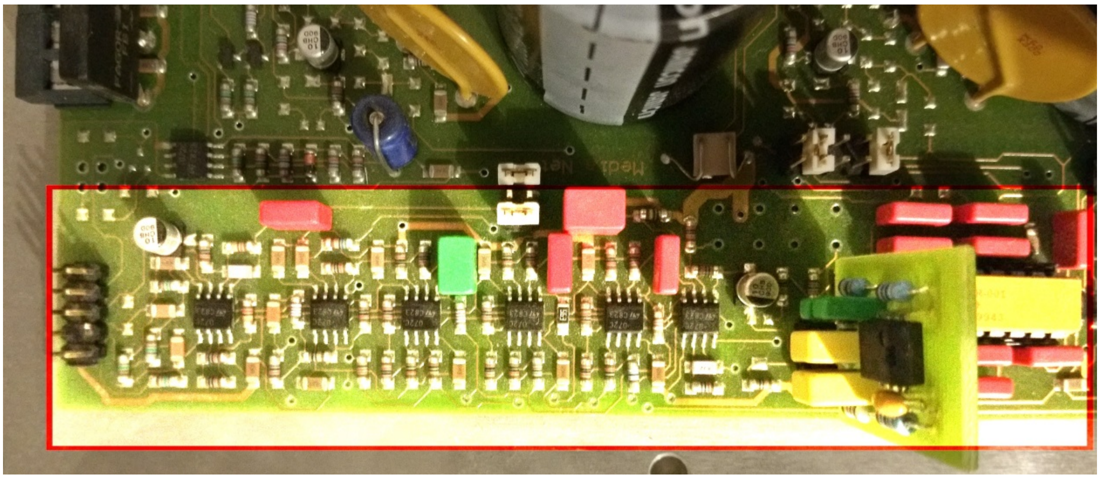


Figure 24: Compensation circuitry from Pfleiderer in an unsealed version (photo of a loud-speaker that appeared on the market in 2002).

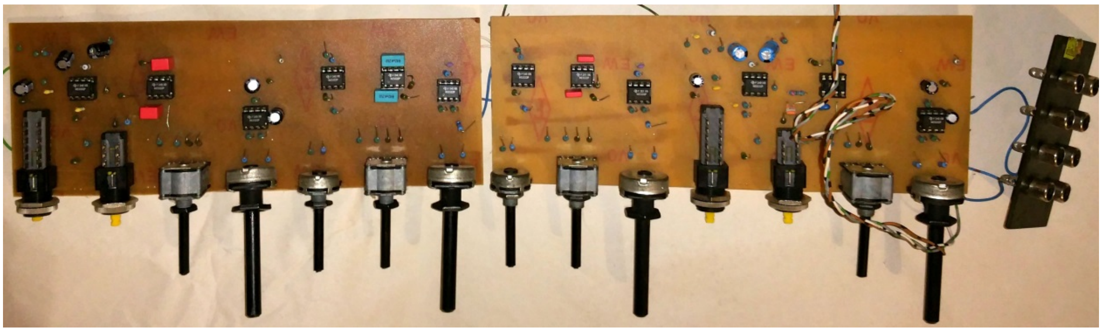


Figure 25: Four-band fully parametric equaliser (built in 1993 from schematics from the Neve VR mixing desk). For the test two of them in series were used.

The driver (a point source driver with a mechanical crossover also from Pfleiderer (1993) – see Figure 26) that was corrected had a large dip in its frequency response towards high frequencies (above 7kHz) and some other irregularities (see Figure 27, left section). All tests were controlled with the measurement system TEF where amplitude and phase plots could be made. With this system a straightforward correction process was possible as the fully parametric equaliser could independently set frequency, gain and quality parameters and the results of the correction could be monitored with the TEF system. For the Pfleiderer circuit the correction was more difficult as its parameters were interdependent and not intuitively useable.

Figure 26: Two-way coaxial driver with a mechanical crossover. The rubber ring decouples the metal tweeter from the paper cone above 7kHz. Below 7kHz cone and tweeter form one large membrane.

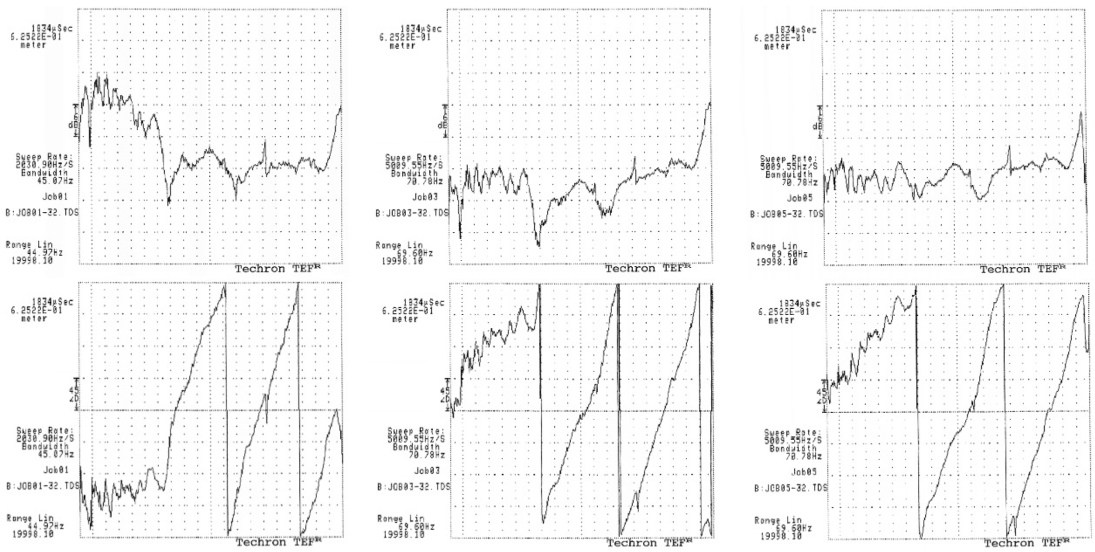


Figure 27: Amplitude and phase diagrams from the test of the Pfleiderer loudspeaker: Left: Uncorrected response. Middle: Correction result by Pfleiderer's circuitry. Right: Correction with two four-band equalisers. The frequency scale for all diagrams is linear with 1kHz per section.

The result was that both devices led to a satisfactory result with an obvious advantage of the equaliser (see Figure 27, middle and right section). While the apparatus cited by Pfleiderer was able to compensate larger, general amplitude errors (such as 20dB jumps in the amplitude response) in one stage, the fully parametric EQ was able to focus more accurately on details in the frequency response. Two EQ bands were needed only to compensate the large jump in the frequency response at 7kHz but as there were eight bands available, there was enough room for correcting further details as well. The Pfleiderer system only had four stages and could therefore not treat minor errors. Furthermore, the correction circuitry from Pfleiderer produced much more noise in its sealed compensation module than the high-class studio EQ from Neve.

Concerning both amplitude and phase the results with the equaliser were better and led to a more linear phase. This can be seen when a ruler is aligned to the measurements in the right section of Figure 28. It can also be observed that below 7kHz the phase is less steep than above 7kHz. This is because of the mechanical crossover of the driver: below 7kHz the whole woofer emits a signal, while above 7kHz only the inner spherical cap, which is at a greater distance from the measurement microphone, emits a signal.

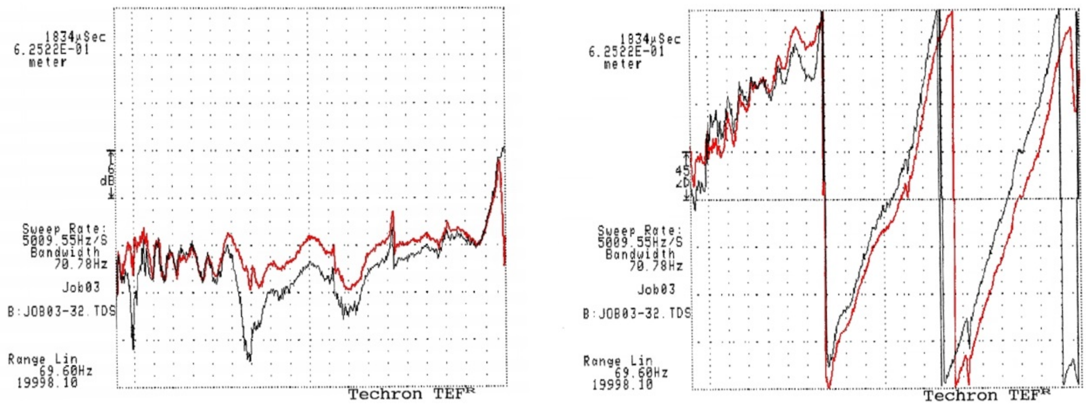


Figure 28: Amplitude (left) and phase (right) diagrams from corrections with Pfleiderer's circuit (black) and with equalisers (red). The frequency scale is linear at 1kHz per section.

An important observation from this series of tests is that a standard equaliser can compensate amplitude and phase errors simultaneously. For accurately addressing the error a fully parametric equaliser in combination with a measurement system is needed. Furthermore, there should be enough bands to correct all possible errors. It was interesting to see how the correction of an amplitude error led to the correction of the corresponding phase error as well. An example is shown in Figure 29, which also demonstrates well the suggestions from Heyser (1969a) (see last paragraph in subsection 2.2.3.5). With one standard equaliser band a bump in the amplitude frequency response at ~2,200Hz (blue) is flattened (red). At the same time the equaliser flattened the phase frequency response as well (orange: before – green: after). Note: both amplitude and phase are considered to be best when flat.

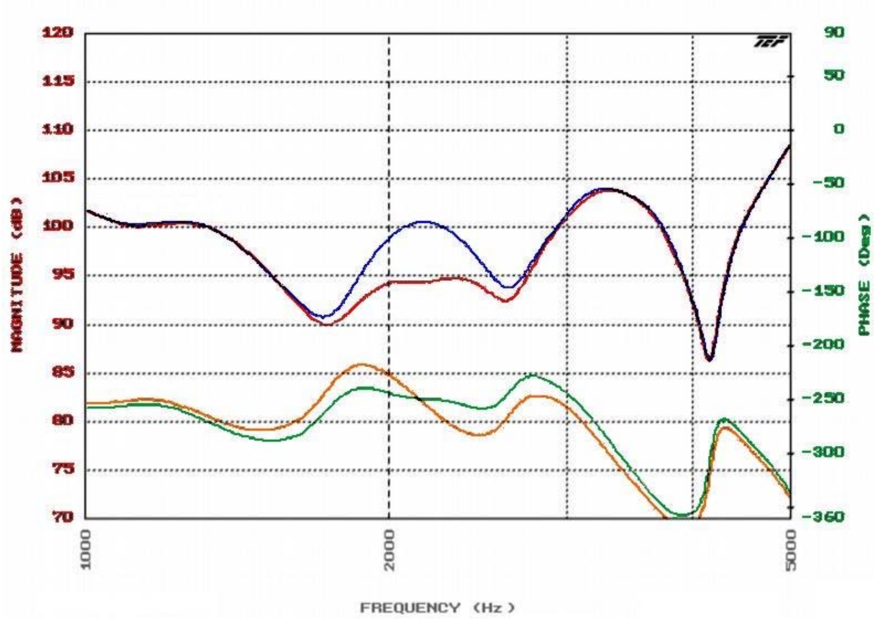


Figure 29: Compensating a loudspeaker error at about 2,200Hz with a fully parametric equaliser. The frequency response (blue) is corrected to the red one. At the same time, the erroneous phase (orange) flattens to the green one.

In discussions about equalisers with or without phase impact (see subsection 2.2.3.3 – linear phase FIR filters) the author prefers filters with phase impact (i.e., non-linear phase) for loudspeaker correction. As already seen in 4.3.2.1, the linear phase filter would keep the phase uncorrected while an analogue filter or digital IIR filter would correct the phase while also correcting the amplitude.

4.3.2.3 Intentionally building an incorrectly designed small-format loudspeaker and equalising the errors

In order to demonstrate the possibilities of equalisation in conjunction with versatile and modern digital crossovers a thought-provoking test was designed. The idea was to build a pair of small two-way loudspeakers with coaxial drivers from a car accessory portfolio. The loudspeaker box was designed arbitrarily but small in size and had a volume of about 1.5 litres. The dimensions were chosen in a way that the box would fit conveniently into a CD shelf (15x15x15cm). Normally loudspeakers of that size would work from about 150Hz and above. Slightly larger in size are, for example, the well-known Auratone 5C Super Sound Cubes, which start at about 140Hz (Auratone Corporation, n.d.).

According to the loudspeaker theory in subsection 2.2.3.5, a closed box (as the simplest example) equals a second-order high-pass filter. The second-order high-pass of the closed box can be envisaged as a first-order high-pass (from the piston of the membrane) in combination with the mass-spring resonance of the membrane, its suspension and the air in the box. This substitution can be seen in Figure 30 for a resonance with Q_{tc}=0.7 (which would normally be considered as a box with the 'right' size[41]).

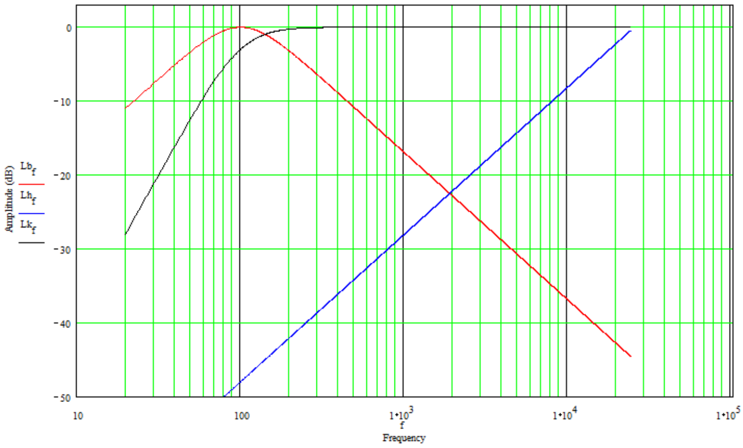


Figure 30: Construction of a second-order high-pass (black) from a band-pass (red) with f_0=100Hz and Q=0.7 and a high-pass (blue) with $f_0 = \infty$. Note: The diagrams are shifted in level for better comprehensibility.

For smaller boxes Q would increase and for larger boxes Q would decrease. Possible frequency responses can be seen in Figure 31.

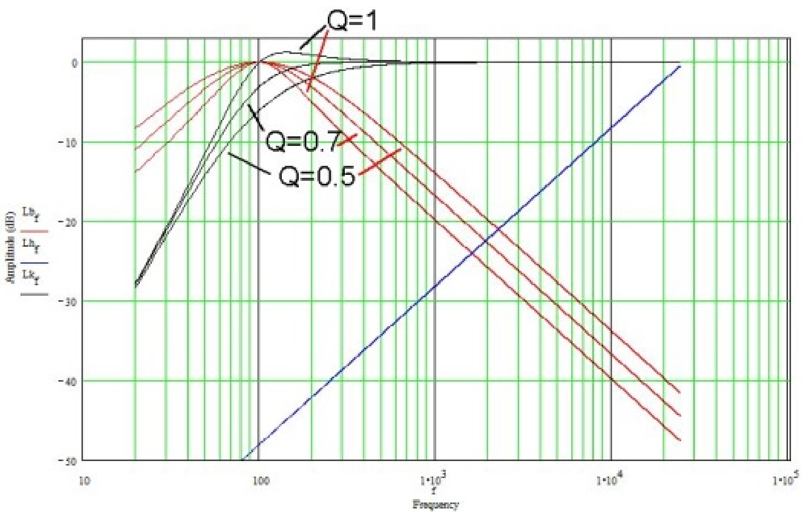


Figure 31: Theoretical amplitude curves for a closed box with Q=0.5 (Bessel), Q=0.7 (Butterworth) and Q=1 (Chebychev).

[41] One of the pioneers of loudspeaker box design, Richard Small (1972, p.801) states that "In audio systems the flatness and extent of the steady-state amplitude-vs-frequency response – or simply frequency response – is usually considered to be of greatest importance." and that "… Q_{tc} = 0.71 […] is a second-order Butterworth (B2) maximally-flat alignment". In Small's design example (1973, p.14) he then calculates the example with Q_{tc}=0.707 as default value.

The drivers were 4" JBL P452 Coax for car hi-fi use. For these drivers, which had a soft suspension, the box can be seen as 'too small' and a resonance should become visible in the low-mid frequencies.

The loudspeaker was designed as a vented box with two ports that were tuned to about 45Hz (see Figure 32). As a result, the Helmholtz resonance of the ports was not aligned to the high-pass function of the driver but well below their lower cut-off frequency so a large gap would occur. The low Helmholtz resonance was chosen intentionally as a frequency response of the whole loudspeaker down to 40Hz was desired to make a small nearfield-loudspeaker sound like a larger one. Further, the box was planned to finally be a filtered vented box that avoids an overly large membrane excursion.

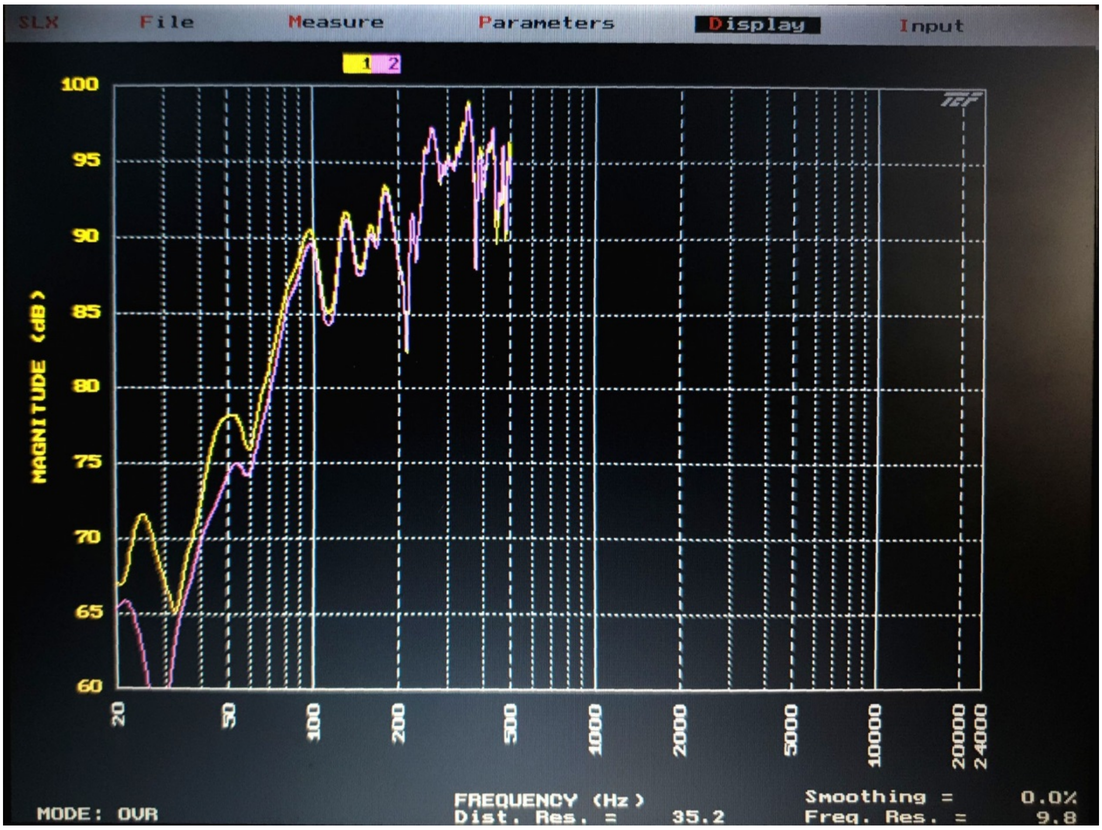


Figure 32: Bass frequency response without equalisers. Violet: Ports closed. Yellow: Ports opened. The additional energy from the ports can be clearly seen around 50Hz. Note: The comb filter on the measurement is from room reflections; the difference below 32Hz is noise in the measurement.

In addition to the large irregularities in the bass that came from the improper volume of the box, there were smaller irregularities expected, which could come from the drivers themselves such as resonances within the membrane or from the construction. The hypothesis of the author was that most irregularities would come from mechanical spring-mass problems and should be correctable with equalisers. For example, the peak due to the 'too small' box should be correctable by a band-pass EQ set to cut. The gap between the output from the ports at the Helmholtz resonance frequency and the directly radiated sound should be correctable by one or more band-pass EQs set to boost.

The coaxial drivers had a very small tweeter (diameter smaller than 1cm) and therefore the crossover frequency was expected to be rather high which might lead to partial oscillations of the woofer's membrane. Some other problems that were not spring-mass related were also expected such as reflections from inside the enclosure or from behind the membrane (see Figure 33). All these problems lead to interferences as multiple signals sum up. This could be either the direct signal and a reflection or signals from differently moving parts of the membrane. These interferences can be seen in the measurements as comb filter effects or as deep notches. Correcting these problems with an equaliser was attempted but proved to be limited in effectiveness. A deep notch could be boosted with an equaliser set to a very high Q-factor but this led to a ringing effect[42] of the equaliser and thus introduced a new problem. It was therefore decided to either correct these problems mechanically (e.g. through additional porous absorptive material) or leave them uncorrected.

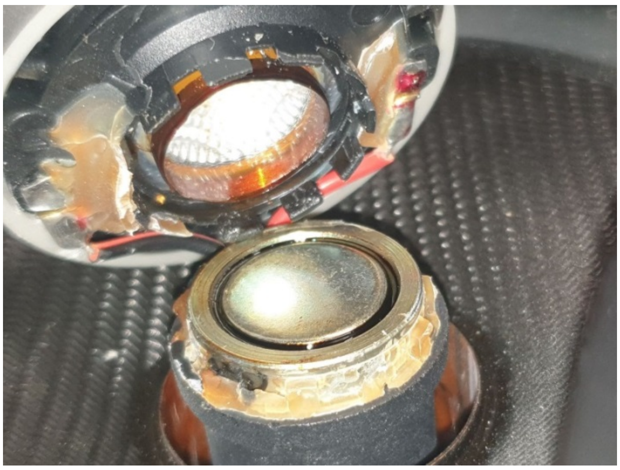


Figure 33: Missing absorptive material behind the membrane can lead to reflections.

The loudspeaker was connected to a Behringer Ultradrive DCX2496 loudspeaker management system and four small car hi-fi power amps. An active two-way system was planned. For this, the passive coaxial drivers were modified in a way that tweeter and woofer could be driven separately. For the correction task the DCX2496 offered up to nine fully parametric equaliser bands per output and again up to nine

[42] Towards higher Q-factors equalisers set to boost act like a ringing bell and produce a decaying tone in addition to the processed audio signal (Dove, 1991, pp.836–837).

equaliser bands per input (Behringer, 2013, p.14). In this way each driver could be equalised with nine bands and the overall system could then be tuned with another nine bands. At a later stage of the project the power amplifiers were integrated into the case of the DCX2496 to get a more compact and portable system (see Figure 34).

Figure 34: Small active loudspeakers with a digital crossover for demonstrating equalisation effects. On the left side of the crossover, four car hi-fi power amplifiers were integrated into the crossover.

All steps were controlled with the TEF measurement system. As a first step, the level of the tweeter and the woofer were adjusted so they matched approximately. Then, a good-sounding crossover frequency was to be determined. Fourth-order Linkwitz-Riley filters were used, as they have a flat-sum response (see subsection 4.2.2.3). The crossover frequency was determined by sweeping through the range of 2kHz to 8kHz while listening to music. This procedure was preferred against any theoretical or measurement approaches, as the loudspeaker should have a good sound at the end, and this can only be determined through listening tests. There was a preferred small frequency range between 5 and 6kHz where the loudspeaker sounded best. The crossover frequency was finally set to 6kHz.

The next step was the time alignment of the drivers. In this coaxial driver the tweeter was in front of the woofer and therefore arrived too early. A delay of 90µs (30mm) was set in the tweeter's path. Following time alignment, the major errors in the frequency response were adjusted: the resonance from the 'too small' box had to be attenuated. The gap between the port-resonance and the low end of the driver had to be filled, the area around the port-resonance had to be adjusted, and the filter for the filtered box had to be installed. For this purpose, the following adjustments were made (see Table 2):

Error	Adjustment
Compensating the overly small box	Bell-EQ 223Hz / -4dB / Q 2.5
Gap between Helmholtz resonance and woofer cut off	Bell-EQ 69Hz / +7.5dB / Q 1.4
Adjusting resonance area of the woofer	Bell-EQ 43Hz / +12dB / Q2.5
Getting a filtered box	High-Pass 40Hz / Butterworth 12dB/Oct

Table 2: Adjustments for compensating the low end of the loudspeaker

The boost at the port-resonance was supporting the tiny ports. This boost can be achieved without larger problems because the membrane excursion at the port-resonance is small and therefore the driver tolerates more power without reaching the maximum membrane excursion. Only airflow noise in the ports might be a problem. The additional high-pass filter at 40Hz cuts off the frequency range below the port-resonance where the acoustical short-circuit occurs (see Figure 35).

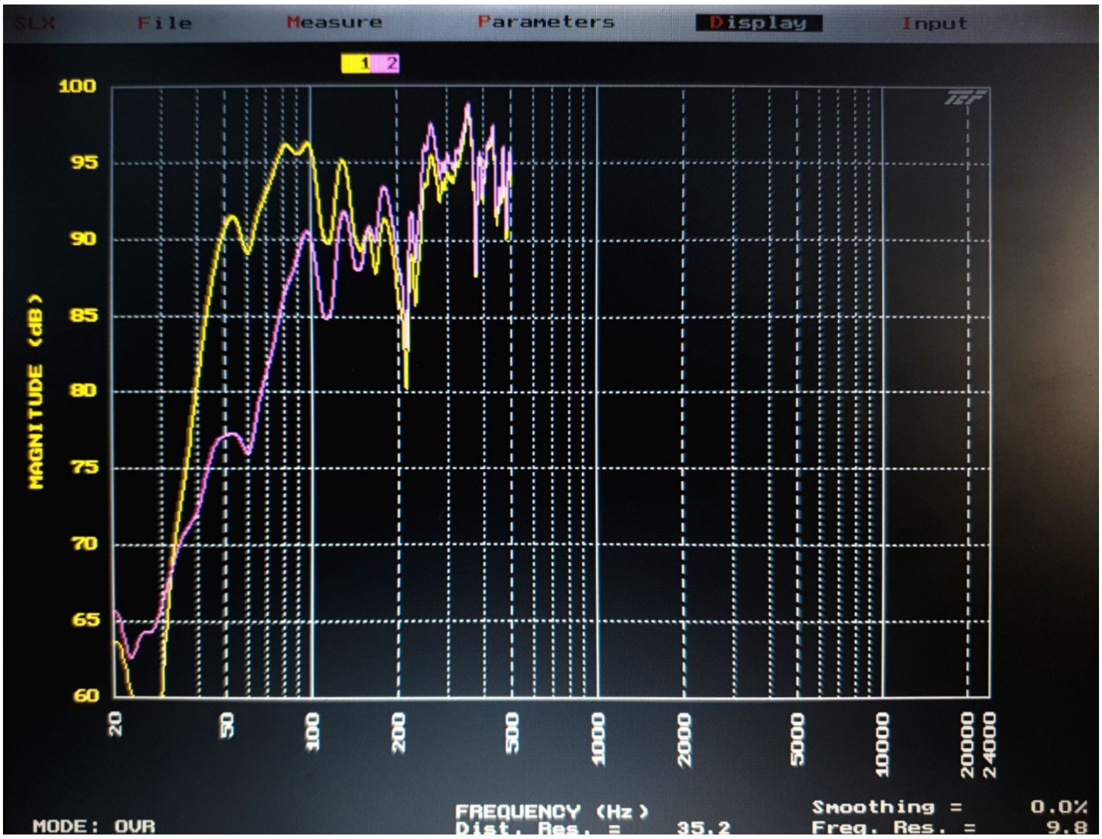


Figure 35: Bass response without (violet) and with an equaliser (yellow). Note: The comb filter on the measurement is from room reflections.

The final step was to look at smaller 'defects' of the driver. These could come from material self-resonance and damping effects. Corrections were made according to the irregularities seen in the measurements and treated in order of severity. All remaining EQ bands of the woofer were used. For the tweeter only two EQ bands were used to correct larger gaps in its frequency response (7.5kHz and 17kHz). The overall result can be seen in Figure 36 and Figure 37. Unfortunately, the drivers had some larger interference and reflection problems that became visible as deep notches and that were not corrected. Although the problems of the driver could not be corrected completely, it can clearly be seen that the phase response smoothens. From a very non-linear response with jumps (cyan line in Figure 37), it gets closer to linear (red line in Figure 37).

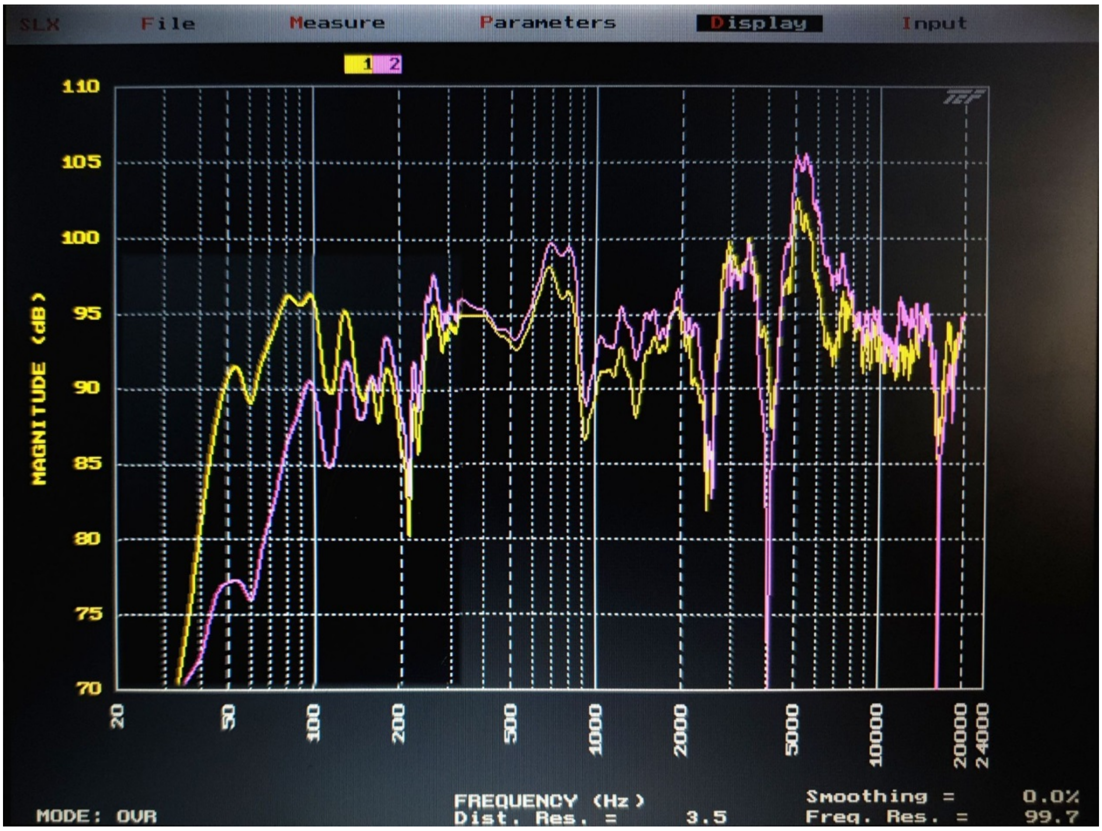


Figure 36: Full-range amplitude response comparing the equalised (yellow) and the unequalised loudspeakers (violet). Note: Low and high frequencies were measured differently and then graphically combined (split frequency = 300Hz).

Figure 37: Phase response comparing the equalised (red) and the unequalised loudspeakers (cyan). The equalised response is getting much closer to linear phase. Note: Linear phase would be a straight line of an arbitrary angle.

Results from this research were:

Using the two-way drivers with active filters brought a large improvement over the passive filters they were delivered with. But being uncorrected, the loudspeakers still sounded small, unbalanced and imprecise. Time alignment together with equaliser-corrections of irregularities in the range of 1kHz to 7kHz improved the precision and overall balance. Extending the useable range towards the low end made the loud-speakers sound larger. Setting the port-resonance at the lowest targeted frequency was advantageous in terms of membrane excursion and made it possible to produce a decent sound level even at low frequencies. The aim of having loudspeakers that are small, can be placed in the nearfield of the listener and sound like larger ones was reached. This outcome was presented at the SAE Convention 2004 and at the Musikmesse Frankfurt in 2005 and 2008.

Further results from this research were:

- Errors from different spring-mass resonances are correctable by an electronic resonating system (equaliser). Amplitude errors and phase errors are corrected simultaneously when using analogue or IIR-based equalisers.
- For optimal results a fully parametric EQ is needed. In order to compensate different kinds of errors the equaliser needs to be flexible.
- Purely mechanical errors cannot be compensated. For example, partial oscillations on the membrane and reflections from the suspension, the magnet and the grill cannot be corrected. Treating these problems requires changes to the construction of the drivers or the box. For example, the car hi-fi driver was not a good choice. It only fulfilled the request for a coaxial construction but had mechanical issues with the membrane material, the grill and the tweeter position.
- The two ports in the small box had an extremely small diameter but still had to be very long. Each port had an inner diameter of 1.3cm which led to 1.33cm² surface for each port (= 2.65cm² in total). The ports were folded once inside the box. For an unfolded version their diameter could have been only 5mm. At higher levels these small ports produced wind noise as there was a rapid air exchange in the ports.

Recommendations for an updated version of this loudspeaker:

- Usage of a better driver with a larger tweeter and fewer mechanical problems.
- Redesigning the ports to have less wind noise.

This updated version was realised in 2019 and can be seen in Figure 38.

Figure 38: Original (left) and updated (right) version of the small active coaxial two-way system. Main differences are the better drivers and the longer and differently formed port.

The size of the box is now 16x16x16cm. The driver is an SB-Audio SB12PFC25-4-COAX 4" driver. The port was changed to a flat, rectangle port which is folded two times inside the box. The increased length leads to a three times larger surface area (8.1cm²) of the port and therefore minimises wind noise. More details about the construction can be seen in Appendix G on page 121. The updated version was corrected as before but with adapted parameters. The new crossover frequency is now 2.5kHz, as the tweeter is larger. As a result, the woofer had to reproduce less energy from high frequencies and therefore fewer problems with partial oscillations of the membrane. The correction needed six EQ bands for the woofer and six for the tweeter. The overall result is better than with the first model and has no more problems. The loudspeaker can now produce quite loud and deep signals. The actual limitation only comes from the 15W amplifier and no longer from the driver (see Appendix G from page 129).

4.3.2.4 Modification of loudspeakers to shape their sound

The most noteworthy loudspeakers are the NS-10 from Yamaha and the 5C Super Sound Cube from Auratone. Although they date from the 1970s and '80s, these two in combination with a large-format loudspeaker are still used in many studio environments (McChane, Hong and Baccigaluppi, 2016). But in regard to frequency response these loudspeakers have many disadvantages. They do not have a really good low end and their mid-frequency range is overemphasised. Investigating the impact of these drawbacks was the main focus of some smaller research undertaken over the years. The author observed that mixing with loudspeakers that have a good low end and a very flat frequency response (such as the Behringer Truth 2030A or the Mackie HR824) is much more difficult than using loudspeakers like the NS-10M. It was assumed that the good low end was responsible for masking mid frequencies and therefore mid frequencies cannot be judged properly. As a result a mix sounds good if it is reproduced with enough bass.

To verify this, a simulated loudness measurement showing the masking effect was undertaken. Figure 39 shows an analysis with a recorded bass subgroup (bass and bass drum). Based on a software listing to calculate loudness according to DIN45631 (Zwicker et al., 1991) an Excel spreadsheet was developed. The spectrum of a bass subgroup was set as input to monitor the masking effect. As a result, the new hearing threshold from the masking effect (blue) was compared with the input spectrum (pink). It can be seen that some frequencies are being masked (pink lines below the dotted new hearing threshold) and therefore cannot be properly assessed.

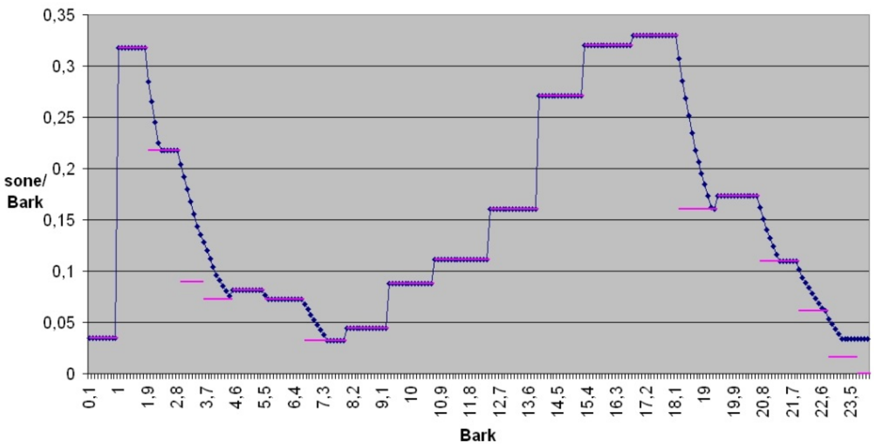


Figure 39: A bass subgroup that is not properly mixed. Bass and bass drum were only added in level, but no equalising or ducking was implemented. It can be seen that the masking effect leads to a new threshold of hearing (blue) which masks some of the frequency bands (pink).

To compensate for the masking effect there are two options: either the low end has to be weaker or the mid frequencies have to be emphasised. Both are the case for the NS-10M loudspeakers. That is why these loudspeakers present the mid frequencies in greater detail, like a camera lens that zooms into a section to focus on a specific detail. Having the frequency range from 1–3kHz (around 8–16 Bark) over-emphasised where the human ear is most sensitive (Fletcher and Munson, 1933) leads to scrutinising this range in greater detail and mixing it very accurately. At the same time the author recommends weakening the low end and therefore reducing its masking effect. Both measures could be reached by equalising a loudspeaker. Although these equalising measures produce an unrewarding listening experience in terms of overall sound, they make the loudspeakers extremely effective for mixing. All above-mentioned measures were demonstrated in lessons using equalisers in the monitor section of a mixing desk.

Comparing the Yamaha HS50 to the Yamaha NS-10M revealed a similar problem. Although the HS50 was claimed to be a successor to the NS-10, it sounded very different and also received some unfavourable reviews when comparing its sound to the NS-10[43]. The HS50 was a vented box with more low end and a quite flat frequency response. As such, it compares favourably with many other loudspeakers, such as the above-mentioned Behringer Truth 2030A. A measurement confirmed this similarity. But the HS50 has a unique feature: with switches on its back, it could

[43] Some reviewers on a loudspeaker review thread stated "…hyped treble […] hyped bass, and subdued mids. […] If you are mixing - these are not NS-10s.", "These are not even close to NS10s, they are a ported speaker…", "I would only consider buying the HS50Ms as second pair for reference, but even then, if you want the NS10 thing buy some NS10s" (Gearspace, n.d.)

be equalised to bring its frequency response close to the one of the NS-10. For that the low end has to be cut and the mid frequencies have to be boosted. The author recommended these setting for a better mixing experience but got the response that most engineers did not apply them as they preferred the sound of the unfiltered box. This led to the conclusion that the first sound impression that a loudspeaker makes is more important than the mix one could make with it.

The following results were observed during this research:

- Overemphasising a certain frequency band is like an optical zoom into this band. Information in the band is heard more prominently and therefore mixed more accurately.
- A weak low end is a large benefit if the mixing focus lies on the mid-frequency range. As low frequencies mask mid and high frequencies (Zwicker and Fastl, 1999, pp.106–110) an overemphasised low end makes a loudspeaker sound very generous. Almost everything on it sounds fine. In order to have a very critical-analytical loudspeaker a weak low end is needed.
- For a mixing engineer it should not be directly relevant if a loudspeaker sounds acceptable or not. It is only important that the mix that was achieved via these loudspeakers sounds good at the final stage. A loudspeaker therefore cannot simply be judged by its frequency response or evaluated through listening tests. A mixing test would be needed where mixes are performed on diverse loudspeakers and the mixing result is then evaluated.

4.3.3 Publications

This subsection expands section 1.3 where a brief description of the published works has already been given. It lists the specific sections of each publication that contain the research findings.

- ***Article:*** *"EQ or not EQ - das ist die Frage!" Magazine: "Recording Magazin", 04/2007; Publisher PPV Medien*
    - **Article about equalising loudspeakers:** This article summarises the research output from 4.3. It shows how corrections affect amplitude and phase response, discusses linear phase filters and shows the limits of an equalised loudspeaker.

- **Book:** *"Die Audio-Enzyklopädie" ("The audio encyclopaedia"); 850+ pages, first published 2007, 2nd ed. in 2014; Publisher:* ***DeGruyter***
    - **Subsection 5.5.4, p. 336: Equalising loudspeakers:** This subsection describes two ways to correct loudspeaker responses either through modification of the crossover or through equalisers. Fully parametric equalisers are recommended as they proved to be suitable (research output from 4.3.2.2).
    - **Subsection 6.10.2.11, pp. 468–469: Membrane excursion when equalised:** This discussion about the limits of the membrane excursion explains the mechanical limits of a loudspeaker towards low frequencies. It also makes clear that an equaliser cannot override these limits as found out in 4.3.2.3.
    - **Subsection 7.12.2.2, p. 577: Linear phase filters and pre-ringing:** This subsection uses the research output from 4.3.2.1 and shows the pre-ringing of linear phase filters as presented in Appendix F, Figure F3 (middle section on the left).
    - **Subsections 10.3.3.3 and 10.3.3.4, pp. 699–700: Equalising loudspeakers – phase correction:** These sections use the research output from 4.3.2.1 and 4.3.2.2 to discuss the usage of different equaliser types. They point to loudspeaker correction and show that traditional equalisers correct both amplitude and phase while linear phase equalisers only correct the amplitude.

- **Book:** *"Studio Akustik"; 190 pages, published 2007, 5th ed. in 2015; Publisher: PPV Medien*
    - **Chapter 11, p. 132: Equalising low frequencies:** This paragraph points the reader towards the possibility of equalizing the low-frequency range of a loudspeaker as an outcome of the research in 4.3.2.4. However, improving the room acoustics is claimed to be superior to equalising as it addresses the cause and not solely the symptoms.

4.4 Summary

Although the various research areas presented in this chapter have different approaches and results, they all have one common aim: the optimisation of the last stage of the signal chain and thus the improvement of the listening experience in the control room. All research led to clearly audible differences and improvements in the listening experience. However, the results from Research Area 2 and 3 should be considered as a recommendation for optimisation, as their impact is perceived differently by individual listeners. Everyone can decide for themselves whether an optimisation is considered valuable and whether it is worthwhile taking any measures for optimising the personal signal chain.

5 Conclusion

In this chapter, the three areas of this contextual statement are briefly summarised and evaluated. From this summary, it could be seen that the results from all research areas led to an improvement of the control room listening situation. At the end of this chapter, a brief outlook at possible future projects is given, followed by a personal assessment of the published works and a reflection of the contextual statement.

5.1 Research Area 1: Summary and critical evaluation

Acoustics was one of the first areas to be researched. It is a wide field and only a fraction – a part of studio acoustics – was investigated. Good acoustics in the control room of a studio is the first step in optimising the listening situation and an essential prerequisite for assessing all other aspects presented in this contextual statement. Assessing a mix or the depth-staging of phantom sound sources is difficult without proper acoustic treatment of the listening situation. As a result of the author's research, prototypes of acoustical modules were built and measured. They can attenuate or scatter reflections more efficiently than traditional modules did before.

When the research started, acoustics was a field where only a few companies were on the market and little was communicated to the public. Measurement equipment was expensive and difficult to use. Today, like many others, this field is democratised. The book "Studio Akustik" made a significant contribution to this and set a foundation for audio engineers who wanted to understand studio acoustics or even build their own acoustic treatment. At the time of its first publication in 2007, this book collected and disseminated much information that was impossible for the general public to access. In this book, the author's experience was shared and valuable information for planning, calculating, measuring and building studio acoustics was provided. Latest developments like the combined absorbers and the Helmfusor were presented as innovative solutions for current acoustical treatment. These new and compact types of acoustic modules were first published by the author. With these new acoustic modules, the acoustic treatment and thus the listening situation can be optimised more efficiently than with traditional modules. These modules are still part of the planning of professional studio designers today.

5.2 Research Area 2: Summary and critical evaluation

Time alignment was one of the most positive outcomes achieved, as it substantially improved the listening situation. Different ways of time alignment were developed and published. From mechanical to different electrical solutions a suitable way for every type of loudspeaker was proposed and demonstrated. On top of this, a reliable way to determine the alignment error with a step response was shown. Although time alignment aspects were already discussed in the 1970s, no publication addressed time alignment in the range of 2000-5000Hz and discussed the impact of short timing errors of 100-200µs. The author's research and his publications were the first to address and explain these issues.

The improvement in localisation is enormous and could easily be sensed by amateurs and inexperienced listeners. The stereo image of time-aligned loudspeakers is excellent, allowing one to hear sharp phantom sound sources and a high image depth. Furthermore, perfectly aligned loudspeakers become inaudible. This means that the listener can only localise the phantom sound sources produced by the loudspeakers but no longer localise the position of the physical loudspeakers. This is also a possible test that could be used as an indicator for time-aligned loudspeakers: playing back the same signal over two loudspeakers simultaneously will only result in a phantom sound source that is extremely sharp in the middle between these loudspeakers and nowhere else.

One disadvantage of the mechanical time-aligned passive loudspeakers is the obvious change in amplitude response. Thus, listeners have to decide whether they would like to have the original frequency response or the improved localisation. In the case of the marketed Time Alignment Rings, a 30-day money-back guarantee was given if the buyer was not satisfied with the result.

The electrical time alignment of passive loudspeakers could not only correct the timing between woofer and tweeter but also keep the frequency response very similar to the original one. It is much more expensive than the mechanical solution and needs sophisticated signal manipulation currently not completely available[44] in a loudspeaker management system. The electrical time alignment is useful for time-aligning loudspeakers that should not be modified mechanically or electrically in any way. Finally, time alignment of active loudspeakers is an easy task today as the crossovers are more and more realised in digital signal processors. Adding a delay

[44] The final summing stage (see Figure 23 on page 60) is missing, so an external summing device is needed.

either in the high- or in the low-pass path is not difficult and should be a designated option.

5.3 Research Area 3: Summary and critical evaluation

Equalisation of loudspeakers has many facets. From a technical correction of the frequency and phase response to a controlled shaping of the loudspeaker's character it can be used in many ways. To use it properly there must be a clear vision of the desired outcome, as loudspeakers can be tuned in many different directions. Having a well-tempered loudspeaker that reproduces most material pleasantly might be the goal for purely listening. In studios, however, loudspeakers must also critically present problematic parts of a mix and ideally point out any weaknesses. To do so the loudspeaker can attenuate low frequencies to reduce the masking effect and emphasise mid frequencies to improve scrutiny in this area. Therefore, equalisation can improve the listening situation for both musical enjoyment and critical listening.

All standard equalisers produce phase shifts. This phase shift is normally beneficial for the practical application of loudspeaker correction as it is the inverse of the phase shift the defect incurred. Linear phase equalisers only compensate amplitude errors and not phase errors. They further introduce a pre-ringing into the signal which might be detrimental. However, with the convolution needed for a linear phase equaliser one could also build a universal equaliser in which amplitude and phase can be corrected largely independent of each other. If this happens the result could reach optimal levels.

The author's publications provided an important insight into loudspeaker correction. There were no previous publications that showed that standard equalisers correct both amplitude and phase. Very often phase shifts are considered to be bad and it is claimed that only linear phase equalisers can provide the most transparent sound correction. However, evidence for that claim is rare. Demonstrating that the standard equaliser's phase shifts are crucial for an appropriate loudspeaker correction whilst linear phase equalisation does not correct all errors is an original contribution to the field of audio engineering. The demonstration of the impact of equalisation to small loudspeakers at several presentations offered a broad range of people the opportunity to think about equalisation and its possibilities.

Equalising loudspeakers is only wise in conjunction with accurate measurements. Performing this task only by listening to the loudspeakers is not recommended. In combination with a reliable and accurate measuring system, precise improvements are possible as long as the equalisers are versatile in use (e.g. fully parametric equalisers). Although an equaliser can correct a vast number of errors, it cannot go beyond the limits of physics. Once the excursion of the membrane gets too large, once the ports of a vented box start to compress a signal or make airflow noise, these limits are reached. Furthermore, reflections inside the loudspeaker cabinet are not correctable by equalisers.

5.4 Outlook

Students regularly study studio optimisation, loudspeaker and room measurement and loudspeaker construction. The results from the research could find a place in these projects, as they relate to specialist knowledge that is not easily found on the internet. Time alignment can be a fixed element in each loudspeaker construction and it is easy to realise when considered right from the beginning of the construction.

As processes like the electronic time alignment and the equalisation of loudspeakers are inserted in between the output of the mixing desk or DAW and the power amplifier, it is obvious that these modifications could be realised in the form of a plugin[45] in the control room section[46] of a DAW. Electronic time alignment could be accomplished via presets of common loudspeakers or derived empirically by listening to the precision of the phantom sound sources. Equalisation could simulate different famous loudspeakers on the physical ones and so reduce the number of loudspeakers that have to be installed.

5.5 Personal findings, reflection of the contextual statement and conclusion

During the creation of this document the author had to look for old notes, remember the circumstances and recall the thoughts that he had at the time of the research. It was realised that it is not self-evident having reached all these goals and that a good amount of innovative and creative work was necessary to undertake all these projects. At a time when research took place mainly in libraries and by talking to other

[45] A plugin is a piece of software that is inserted into the audio signal on a DAW. It can either pass the audio through and add something to it or it can replace the original signal with a recalculated one.

[46] The control room section of a mixing desk normally contains loudspeaker selection; mono, stereo, L<->R swap and external inputs. Having an insert in this section is not common in analogue desks. As of 2020 only very few DAWs have a control room section comparable to a mixing desk. But even if they provide one there is currently no possibility to insert plugins in this section.

experts, it was much harder to get information than it is today. Nowadays researching internationally has become normal with the internet. But at the same time the internet does not contain every single piece of information and there is still a need and a benefit to discuss and exchange ideas with other experts in the field; particularly those whose publications have not appeared online.

During the creation of the contextual statement, the author faced the problem that it is not so easy to retrace previous research in a way that fulfils the requirements of an EQF level-eight work, mainly because the research was not originally undertaken as part of such a work. Luckily, much information survived throughout the years – some in the form of printouts in personal binders, others in electronic form as files on floppy disks and backup drives. Fortunately, most of the material is still accessible as Windows operating systems are highly backwards compatible. Based on this material, procedures and outcomes could be restored and finally presented in this contextual statement. Only data from the early days that were stored on a custom CP/M system were lost when this system was put out of operation in 1994.

Within the contextual statement the professional development of the author and excerpts of his research was presented. Over many years the author has conducted research with his colleagues and supervised undergraduate and postgraduate research in the field of audio technology. Many results were shared with the public in the form of presentations at fairs, as publications in magazines and in the form of subject-specific books. For the author it was always important to reach a broad target group and present his research in a comprehensible and demonstrative way. This was achieved successfully as the author's publications serve as a reference for many audio enthusiasts.

Bibliography

Ahnert, W. and Tennhardt, H.-P., 2008. Raumakustik. In: S. Weinzierl, ed. *Handbuch der Audiotechnik*. Berlin: Springer.pp.181–266.

Analog Devices, 1992. *1992 Audio/Video Reference Manual*. Norwood, MA: Analog Devices.

Analog Devices, 1993. *ADSP-2100 family user's manual*. Analog Devices technical reference books. Englewood Cliffs, NJ: PTR Prentice Hall.

Auratone Corporation, n.d. *5C Super Sound Cube Manual*. Coronado, CA: Auratone Corporation.

Ballou, G. ed., 1991. *Handbook for sound engineers: the new audio cyclopedia*. 2. ed. Carmel, IN: Sams.

Bech, S., 1990. Listening Tests on Loudspeakers: A Discussion of Experimental Procedures and Evaluation of the Respose Data. *Audio Eng. Soc*, Paper 8-14, pp.101–115.

Bech, S. and Zacharov, N., 2006. *Perceptual Audio Evaluation: Theory, Method and Application*. Chichester, England ; Hoboken, NJ: John Wiley & Sons.

Behringer, 2013. *Behringer Ultradrive DCX2496 User Manual*. Bothell, WA: MUSIC Group Services.

Bell, J., 2006. *Doing your research project: a guide for first-time researchers in education, health and social science*. 4. ed. Maidenhead, Berkshire, England: Open University Press.

Blauert, J., 1974. *Räumliches Hören*. Stuttgart: Hirzel.

Blauert, J. and Laws, P., 1978. Group Delay Distortion in Electroacoustical Systems. *Acoust soc Am.*, (5.), pp.1478–1483.

Booth, W.C., Colomb, G.G. and Williams, J.M., 2008. *The craft of research*. 3. ed. Chicago guides to writing, editing, and publishing. Chicago: University of Chicago Press.

Burchard, J.M., Irrgang, E. and Andresen, B., 1985. Das äußere Ohr des Menschen als schallverdoppelndes Organ. *Sprache Stimme Gehör*, (9), pp.100–103.

Burchard, J.M., Irrgang, E. and Andresen, B., 1987. Die Funktion der menschlichen Ohrmuschel. *Spektrum der Wissenschaft 6/87*, pp.66–74.

Clemen, C., 1998. *Lautsprechermodell*. [online] Available at: <https://www.hs-augsburg.de/~clemen/lehre/audio/LS_Modell.PDF> [Accessed 1 Jun. 2020].

Crown International, 1997. *TEF Sound Lab TDS Software User's Manual*. Elkhart, IN: Crown International.

Crystal Semiconductor, 1994. *Crystal Semiconductor Audio Databook*. Austin, TX: Crystal Semiconductor Corporation.

D'Antonio, P. and Konnert, J.H., 1983. The Schroeder Quadratic-Residue Diffusor: Design Theory and Application. In: *Audio Engineering Society Convention 74*.

D'Appolito, J.A., 2007. *Lautsprecher-Meßtechnik: PC-gestützte Analyse analoger Systeme*. 3. ed. Aachen: Elektor-Verlag.

Davis, D. and Davis, C., 2015. Audio and Acoustic DNA - Past and Present. In: G. Ballou, ed. *Handbook for sound engineers*, 5. ed. Burlington, MA: Focal Press. pp.3–36.

Denscombe, M., 2010. *The good research guide: for small-scale social research projects*. 4. ed. Open up study skills. Maidenhead: Open University Press.

Dickason, V., 2002. *Lautsprecherbau: bewährte Rezepte für den perfekten Sound*. 2nd ed. Aachen: Elektor-Verlag.

Dickreiter, M. and Schule für Rundfunktechnik eds., 1987. *Handbuch der Tonstudiotechnik. Bd. 1: Raumakustik, Schallquellen, Schallwahrnehmung, Schallwandler, Beschallungstechnik, Aufnahmetechnik, Klanggestaltung*. 5. ed. München: Saur.

Dickreiter, M. and Schule für Rundfunktechnik eds., 1990. *Handbuch der Tonstudiotechnik. 2: Analoge Schallspeicherung, analoge Tonregieanlagen, Hörfunk-Betriebstechnik, digitale Tontechnik, Tonmeßtechnik*. 5. ed. München: Saur.

Dove, S., 1991. Consoles and Systems. In: G. Ballou, ed. *Handbook for Sound Engineers*, 2. ed. Carmel, IN: Sams. pp.771–906.

Ericsson, K.A. and Simon, H.A., 1981. *Protocol Analysis*. Typescript available at: <http://digitalcollections.library.cmu.edu/awweb/awarchive?type=file&item=39233> [Accessed 1 Jul. 2021].

Everest, F.A., 1991. Acoustical Design of Audio Rooms. In: G. Ballou, ed. *Handbook for Sound Engineers*, 2. ed. Carmel, IN: Sams. pp.103–141.

Everest, F.A., 2001. *The master handbook of acoustics*. 4. ed. New York: McGraw-Hill.

Fehlhaber, W., 2018. *Signal Analysis II: Linear vs Logarithmic Sine Sweep*. [online] Available at: <https://www.linkedin.com/pulse/signal-analysis-ii-linear-vs-logarithmic-sine-sweep-william-fehlhaber> [Accessed 7 Jun. 2020].

Feynman, R.P., Leighton, R.B., Sands, M.L. and Feynman, R.P., 1997. *Vorlesungen über Physik - Band 1: Mechanik, Strahlung und Wärme*. 3. ed. Vorlesungen über Physik. München: Oldenbourg.

Fletcher, H. and Munson, W., 1933. Loudness, Its Definition, Measurement and Calculation. *J.A.S.A.*, 5(Oct 1933), pp.82–108.

Fletcher, H. and Wegel, R., 1922. The Frequency-Sensitivity of Normal Ears. *Proceedings of the National Academy of Sciences of the United States of America*, 8(1), pp.5–6.

Fluke, 1993. *93/95/97 50MHz Scopemeter - Users Manual*. s.l.: Philips.

Friesecke, A., 2014. *Die Audio-Enzyklopädie: Ein Nachschlagewerk für Tontechniker*. 2. ed. Reference. Berlin: De Gruyter Saur.

Gearspace, n.d. *Yamaha HS50M*. [online] Available at: <https://gearspace.com/board/reviews/673329-yamaha-hs50m.html> [Accessed 7 Jul. 2021].

Genelec, 2013. *35 years of sustainable development*. [online] Available at: <https://www.genelec.com/sites/default/files/media/About%20Us/Genelec%20story/35th_anniversary_book_130507.pdf> [Accessed 17 Dec. 2017].

Goertz, A., 2008. Lautsprecher. In: S. Weinzierl, ed. *Handbuch der Audiotechnik*. Berlin: Springer. pp.421–490.

Henle, H., 1990. *Das Tonstudio-Handbuch. Praktische Einführung in die professionelle Aufnahmetechnik*. München: Carstensen.

Henricksen, C.A., 1991. Loudspeakers, Enclosures, and Headphones. In: G. Ballou, ed. *Handbook for Sound Engineers*, 2. ed. Carmel, IN: Sams. pp.497–592.

Herget, T., 2020. *HTH L-Modelle*. [online] L-Modelle - Erklärung zu den Induktivitäts-Parametern! Available at: <http://www.traumboxen.de/theorie/lmodell.htm> [Accessed 21 May 2020].

Heyser, R.C., 1969a. Loudspeaker Phase Characteristics and Time Delay Distortion: Part 1. *J. Audio Eng. Soc*, 17(1), pp.30–41.

Heyser, R.C., 1969b. Loudspeaker Phase Characteristics and Time Delay Distortion: Part 2. *J. Audio Eng. Soc*, 17(2), pp.130–137.

Heyser, R.C., 1971a. Determination of Loudspeaker Signal Arrival Times: Part 1. *J. Audio Eng. Soc*, 19(9), pp.734–743.

Heyser, R.C., 1971b. Determination of Loudspeaker Signal Arrival Times: Part 2. *J. Audio Eng. Soc*, 19(10), pp.829–834.

Heyser, R.C., 1973a. The Delay Plane, Objective Analysis of Subjective Properties-Part 1. *J. Audio Eng. Soc*, 21(9), pp.690–701.

Heyser, R.C., 1973b. The Delay Plane, Objective Analysis of Subjective Properties-Part 2. *J. Audio Eng. Soc*, 21(10), pp.786–791.

Klark-Teknik, 1988a. JADE one MKii. *Sound on Sound*, [online] April 1988. Available at: <http://www.muzines.co.uk/ad/4471> [Accessed 21 Feb. 2019].

Klark-Teknik, 1988b. *Operating and Service information JADE one MKII*. Kidderminster, England: Klark-Teknik.

Lancaster, D., 1994. *Das Aktiv-Filter-Kochbuch*. 8. ed. Vaterstetten: IWT-Verlag.

Lane, J. and Hillman, G., 1993. *ImplementingIIR/FIR Filters with Motorola's DSP56000/SPS/DSP56001 Digital Signal Processors*. 2. ed. s.l.: Motorola.

Leary, M.R., 2007. *The curse of the self: self-awareness, egotism, and the quality of human life*. Oxford: Oxford Univ. Press.

Long, E.M., 1976. A Time-Align Technique for Loudspeaker System Design. In: *Audio Engineering Society Convention 54*. [online] Available at: <http://www.aes.org/e-lib/browse.cfm?elib=2283> [Accessed 25 Jul. 2020].

Maier, P., 2008. Studioakustik. In: S. Weinzierl, ed. *Handbuch der Audiotechnik*. Berlin: Springer.pp.267–311.

Manley Laboratories, 2004. *Manley Slam! Stereo Limiter And Micpre*. Chino, CA: Manley Laboratories.

Markowski, D., 2016. *Windows 3.1*. [online] Available at: <http://www.winhistory.de/more/win31.htm> [Accessed 22 Jan. 2018].

McChane, J., Hong, A. and Baccigaluppi, J., 2016. 5C Super Sound Cube speaker. *Tape Op*, [online] (111). Available at: <https://tapeop.com/reviews/gear/111/5c-super-sound-cube-speaker/> [Accessed 12 Mar. 2019].

McNiff, J., 2016. *Writing up your action research project*. London ; New York, NY: Routledge.

Moylan, W., 2015. *Understanding and crafting the mix: the art of recording*. 3. ed. New York: Focal Press, Taylor & Francis Group.

Panzer, J., 1994. *Die Konstruktion von Baßlautsprechergehäusen: der Entwurf von Baßlautsprechergehäusen mit Hilfe der Filtertheorie*. 5. ed. RPB-Taschenbuch. München: Franzis.

Pfleiderer, P., 1987. *Device for compensating reproduction errors in an electroacoustic transducer*. U.S. Pat. 4675835 A.

Pfleiderer, P., 1993. Pfleid FRS 20 R. *Audio Review*, (131), pp.132–137.

Randall, R.B., 1987. *Frequency analysis*. 3. ed. Naerum: Brüel & Kjaer.

SB-Acoustics, 2020. *4" SB12PFCR25-4-COAX*. [online] Available at: <http://sbacoustics.com/product/4-sb12pfcr25-4-coax/> [Accessed 14 Mar. 2020].

Schröder, M.R., 1975. Diffuse Sound Reflection by Maximum Length Sequences. *J. Acoust. Soc. Am.*, 57(1), pp.149–150.

Small, R.H., 1972. Closed-Box Loudspeaker Systems Part 1: Analysis. *J. Audio Eng. Soc*, 20(10), pp.798–808.

Small, R.H., 1973. Closed-Box Loudspeaker Systems Part 2: Synthesis. *J. Audio Eng. Soc*, 21(1), pp.11–18.

Smith, S.W., 1999. *The scientist and engineer's guide to digital signal processing*. San Diego, CA: California Technical Pub.

Somekh, B., 2006. *Action research: a methodology for change and development*. Doing qualitative research in educational settings. Maidenhead: Open University Press.

Spitzer, M., 2013. *Musik im Kopf: hören, musizieren, verstehen und erleben im neuronalen Netzwerk*. 10. ed. Stuttgart: Schattauer.

Stanley, G.R., 1982. A Microprocessor Based TEF Analyser. In: Audio *Engineering Society Convention 72*. [online] Available at: <http://www.aes.org/e-lib/browse.cfm?elib=11813> [Accessed 25 Jul. 2021]

Stark, B., 2003. *Lautsprecher-Handbuch: Theorie und Praxis des Boxenbauens.* 7. ed. München: Pflaum.

Team Elektor, 1995. Aktive 24-dB-Linkwitz/Riley-Weiche. *Elektor Halbleiterheft 1995*, (7/8).

Theisen, M.R., 2008. *Wissenschaftliches Arbeiten: Technik - Methodik - Form ; [jede Seite ein Muster: Nachschlagen = Erfolg haben!]*. 14. ed. WiSt-Taschenbücher. München: Vahlen.

Tietze, U. and Schenk, C., 1980. *Halbleiter-Schaltungstechnik*. 5. ed. Berlin: Springer.

Toole, F.E. and Olive, S., 1994. Hearing is Believing vs. Believing is Hearing: Blind vs. Sighted Listening Tests, and Other Interesting Things. In: *Audio Engineering Society Convention 97*. [online] Available at: <http://www.aes.org/e-lib/browse.cfm?elib=6338> [Accessed 15 Jun. 2020].

Veith, C., 1995. *TEF Utility*. [computer program]

Ward, P., 2008. *The Yamaha NS10 Story*. [online] Available at: <https://www.soundonsound.com/reviews/yamaha-ns10-story> [Accessed 29 Nov. 2017].

Webers, J., 1989. *Tonstudiotechnik: Schallaufnahme und -wiedergabe bei Rundfunk, Fernsehen, Film und Schallplatte*. 5. ed. Franzis-Handbuch. München: Franzis.

Yamaha, 1998. *Professional Digital Delay D5000 - Operation Manual*. s.l.: Yamaha.

Yamaha, n.d. *NS-10M Studio Monitor Speaker System User Manual*. s.l.: Yamaha.

Zwicker, E. and Fastl, H., 1999. *Psychoacoustics: facts and models*. 2. ed. Springer series in information sciences. Berlin ; New York: Springer.

Zwicker, E., Fastl, H., Widmann, U., Kurakata, K., Kuwano, S. and Namba, S., 1991. Program for calculating loudness according to DIN 45631 (ISO 532B). *Journal of the Acoustical Society of Japan*, 12(1), pp.39–42.

Zwicky, P. and Bäder, K.O., 1993. Kompensation von Zeitfehlern in Mehrband-Studiolautsprechern mit Hilfe digitaler Signalverarbeitung. In: *TMT Proceedings*. München: Saur.

Appendices

Table of contents:

Appendix A: Detailed timelines

In this appendix three timelines are presented. They show the developments of the author in the three research areas presented in this contextual statement. To get a better understanding of the context the developments in the market are also presented in the lower section of each timeline.

The timelines presented are:

Figure A1: Timeline for the Research Area 1.

Figure A2: Timeline for the Research Area 2.

Figure A3: Timeline for the Research Area 3.

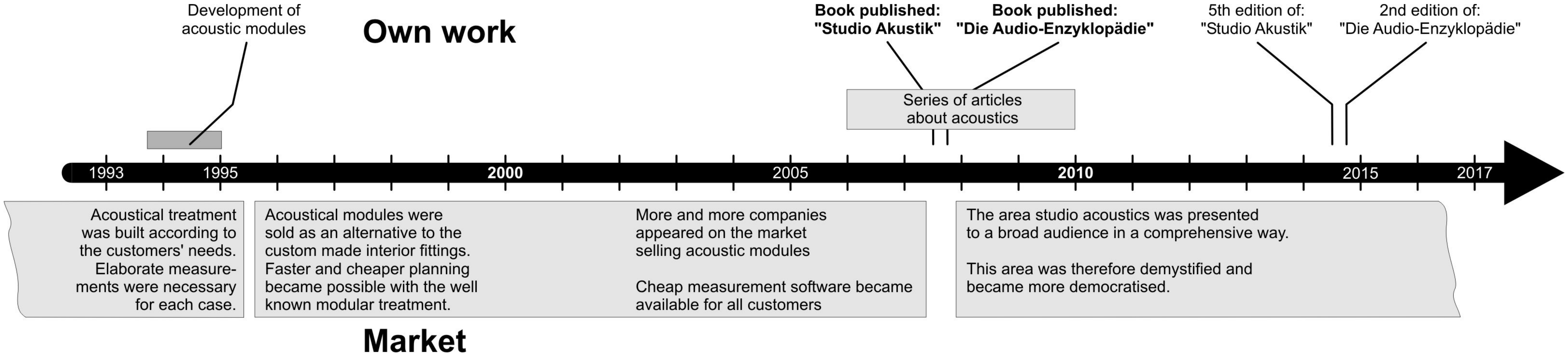


Figure A1: Timeline for the Research Area 1.

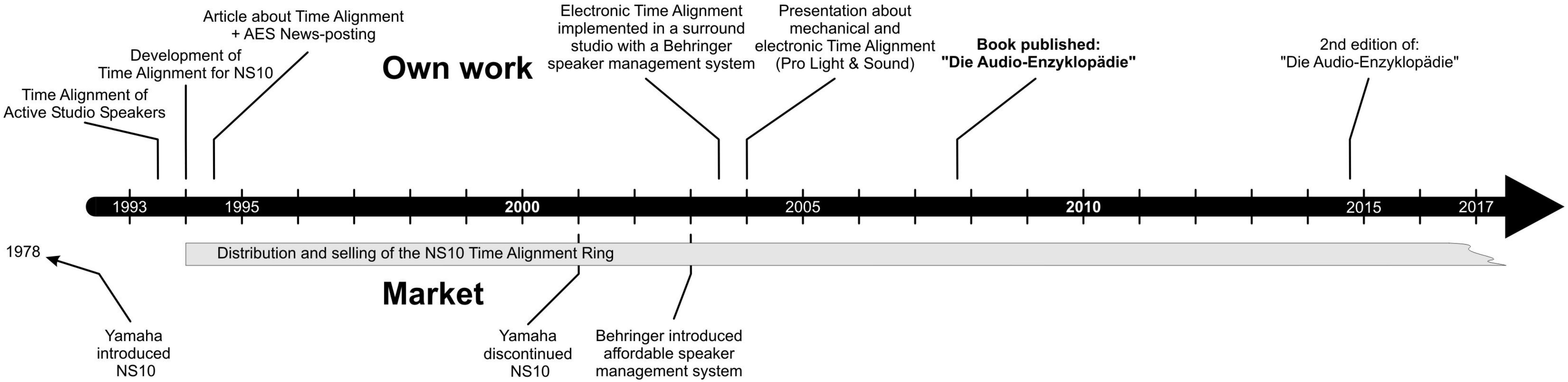


Figure A2: Timeline for the Research Area 2.

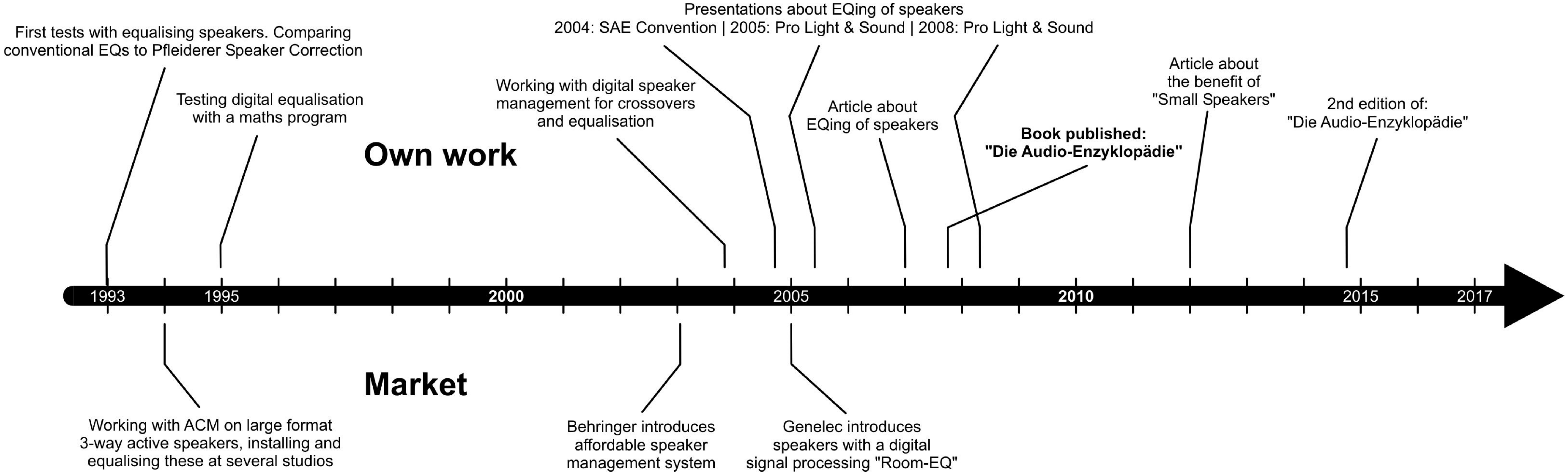


Figure A3: Timeline for the Research Area 3.

Appendix B: Energy distribution of impulse and step

Measuring frequency responses with an 'ideal impulse' is quite easy, as an impulse that is only one sample in length contains the same energy in all harmonics of a spectrum. In Figure B1 the waveform of an impulse with 1 sample length at sample number 100 is shown. In Figure B2 the amplitude spectrum derived by a 1,024-sample-long FFT (fast Fourier transformation) is presented. It can be seen very well that the impulse's energy that averages across 1,024 samples is now shown as 512 harmonics with the same energy for each harmonic ($\frac{1}{1024} = -60.2dB$).

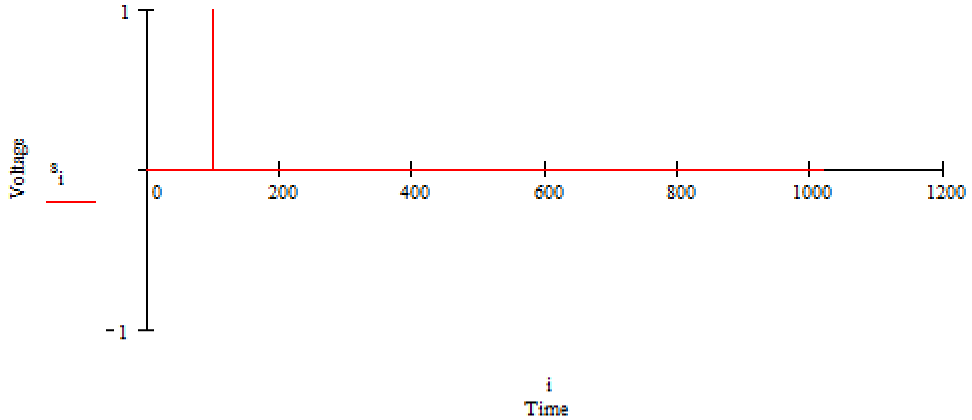


Figure B1: Waveform of an 'ideal impulse' with one sample length.

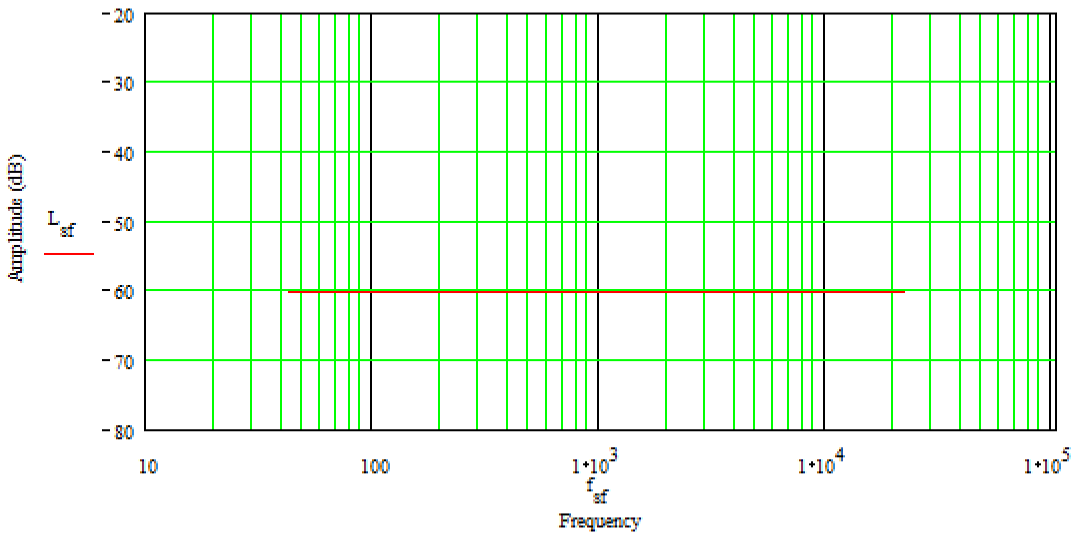


Figure B2: Spectrum of the 'ideal impulse' presented as 512 harmonics from a 1,024-point FFT.

This extraordinary energy distribution is only true for an impulse with only one sample length. For a longer impulse the amplitude spectrum is based on a $\frac{\sin(x)}{x}$ function (Randall, 1987, p.46). A step can be seen as part of a longer pulse and so it has a comparable spectrum. In Figure B3 a step starting at sample number 100 is shown. In Figure B4 the amplitude spectrum of this step is presented. It can be seen that the step not only has more energy in total but also has the most energy in the low-frequency range.

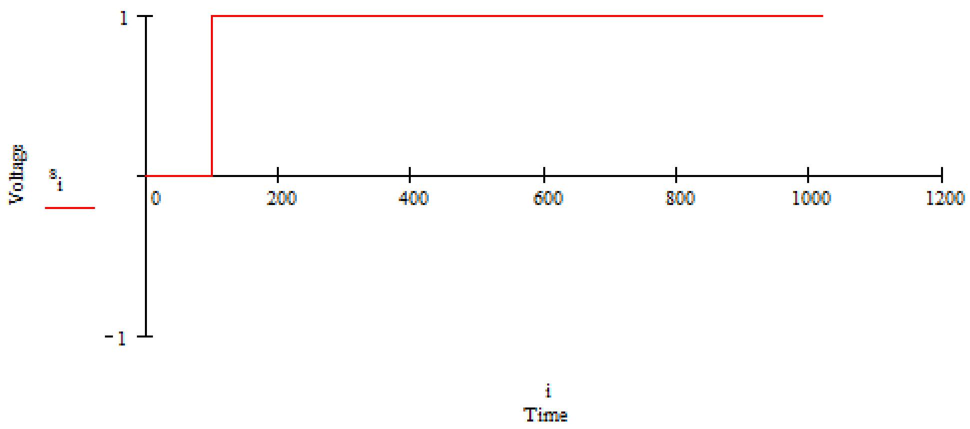


Figure B3: Waveform of a step signal.

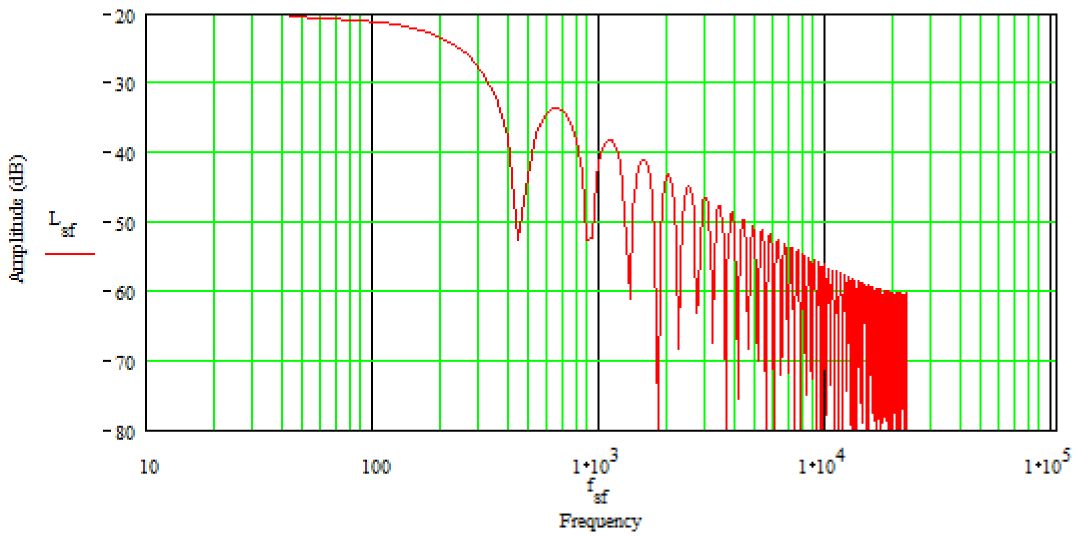


Figure B4: Spectrum of the step presented as 512 harmonics from a 1,024-point FFT.

Appendix C: Measuring absorption coefficients in an echo chamber

If the absorption coefficient of a material is unknown then a measurement of that material in an echo chamber is a possible way to empirically derive the absorption coefficients for that material. An echo chamber is an empty room with hard surfaces, so its reverberation time is long. Ideally, the room has a good diffusivity to avoid standing waves. However, an empty cellar or bathroom could also be used as an echo chamber. The smaller the room, the less absorptive material is necessary to get useful results from the measurement.

Two measurements are needed:

- The frequency-dependent reverberation time of the empty echo chamber.
- The frequency-dependent reverberation time of the echo chamber with the absorptive material installed in the room.

For the following example a small tiled bathroom was used. It has a volume of $13m^3$ and a total surface of $34m^2$. The absorptive material was a $2m^2$ wooden cabinet filled with 20cm thick Rockwool. The cabinet was open at one side. Its closed back was positioned directly against the wall of the room. The waterfall diagram of the empty room is shown in Figure C1. Figure C2 shows the waterfall for the same room with the absorptive material inside. It was mounted to a wall and would therefore replace $2m^2$ of the tiled wall with absorptive material. It could be seen that there is less energy in the room and the waterfall declines more quickly. Figure C3 shows the waterfall of the empty room cut into third-octave slices, calculating the mean of each third-octave band and then aligning a slope for each third-octave. This slope represents the RT60 for each third-octave. The same process for the modified room can be seen in Figure C4. Again, it can be seen that the room with absorptive material has shorter reverberation times.

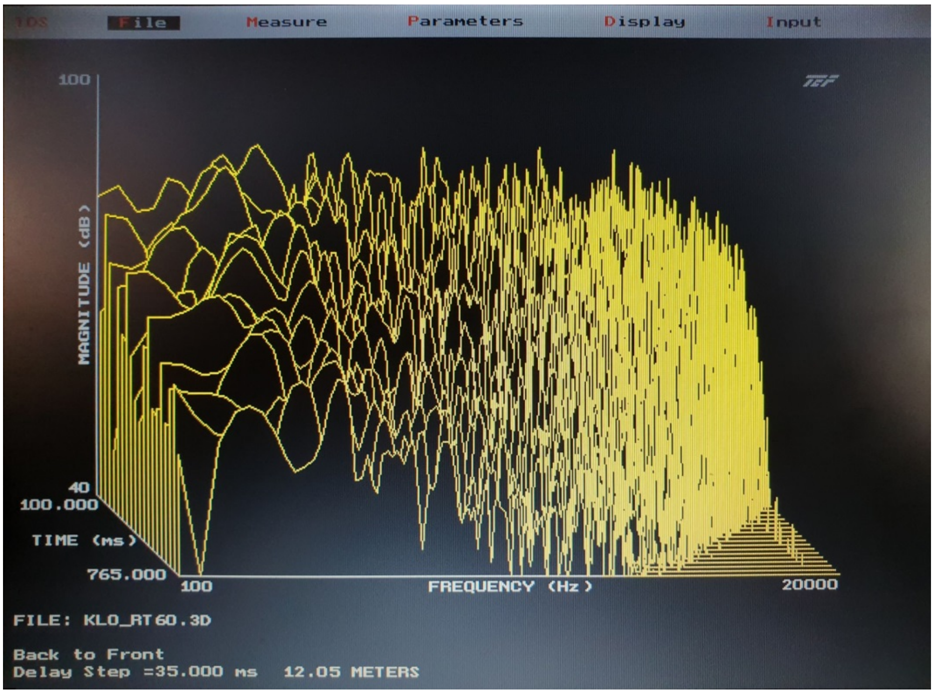


Figure C1: 3D waterfall diagram of the empty echo chamber.

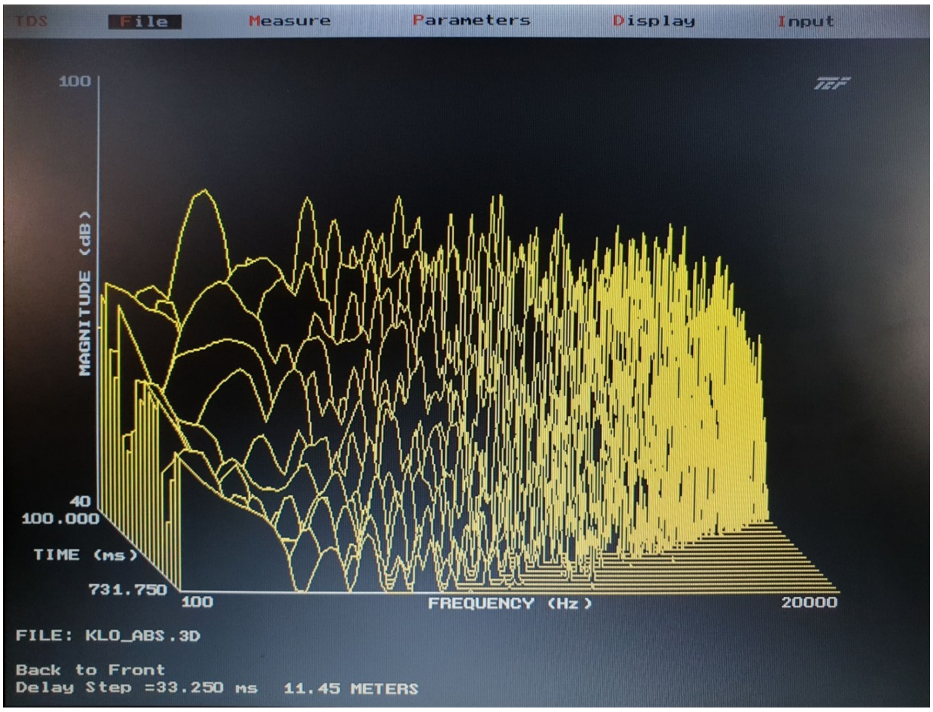


Figure C2: 3D waterfall diagram of the echo chamber with the absorptive material brought inside.

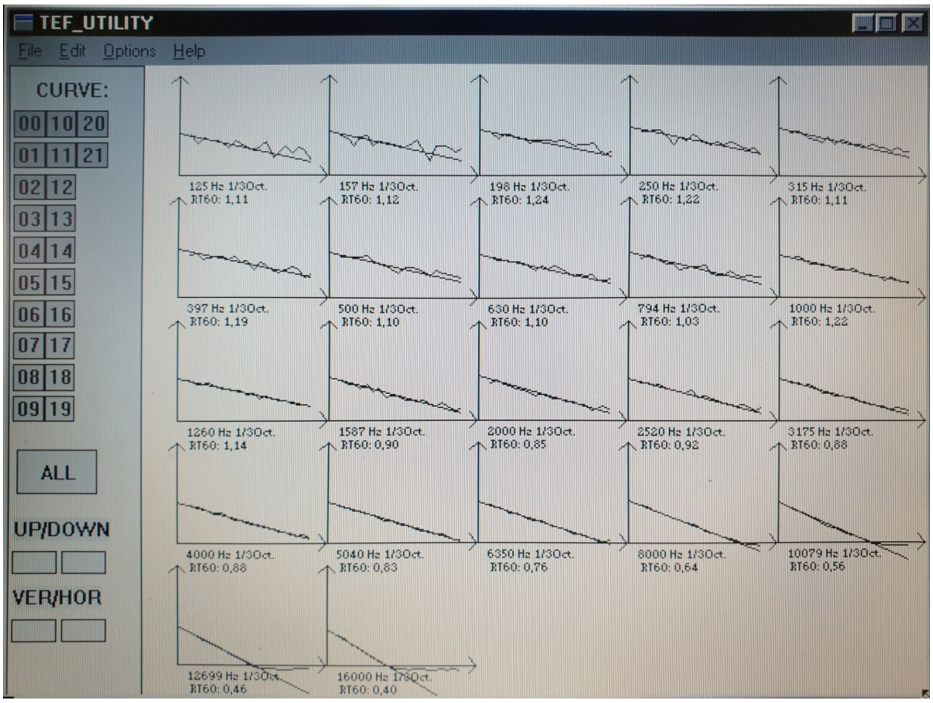


Figure C3: RT60 analysis of the empty echo chamber.

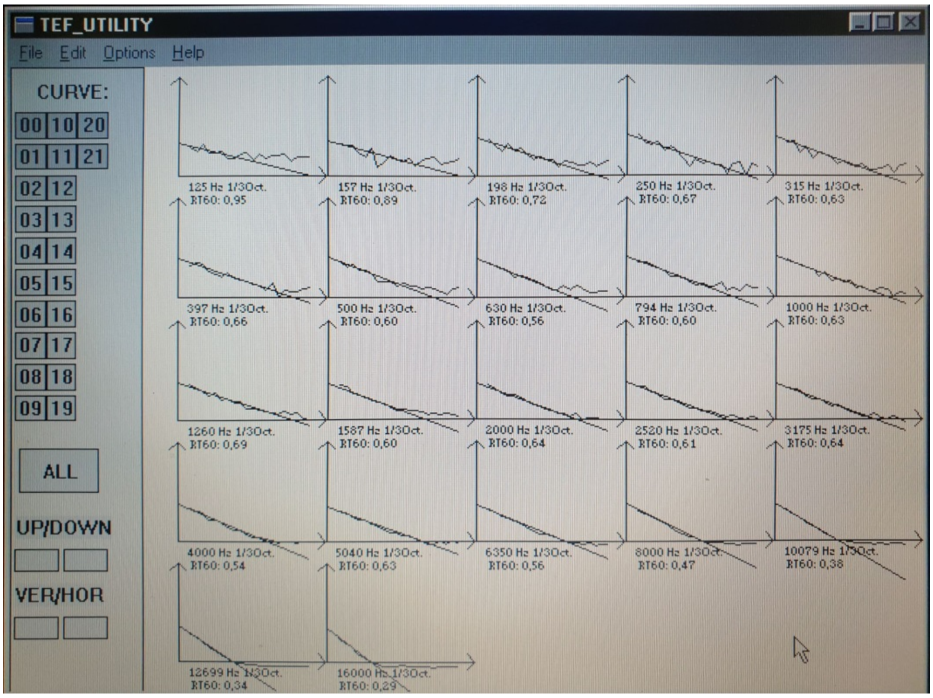


Figure C4: RT60 analysis of the echo chamber with the absorptive material brought inside.

The results from the RT60 analysis are entered into an Excel spreadsheet for calculating the absorption coefficients for each third-octave (see Figure C5). The formula for calculating the reverberation time is[47] (Ahnert and Tennhardt, 2008, p.190):

$$RT60 = 0{,}163 \cdot \frac{V}{-\ln(1-\propto) \cdot S}$$

where:
α: Average absorption coefficient
V: Volume of the room (m³)
S: Total surface of the room (m²)
RT60: Reverberation time of the room (s)

Solved for α the formula is:

$$\propto = 1 - e^{\frac{-0{,}163 \cdot V}{S \cdot RT60}}$$

RT60 will be measured twice and so α can then be calculated for:

- the empty room: α_{empty}
- the room with the absorber inside: α_{test}.

The average absorption coefficient is derived by adding up each absorptive surface multiplied with its absorption coefficient and then dividing the result by the total surface of the room:

$$\propto = \frac{\propto_1 \cdot S_1 + \propto_2 \cdot S_2 \ldots}{S}$$

where:
α: Average absorption coefficient
α_1: Absorption coefficient of first material
α_2: Absorption coefficient of second material
S: Total surface of the room (m²)
S_1: Surface of first material (m²)
S_2: Surface of second material (m²)

For the room with the absorber inside the absorber reduces the surface S, which absorbs with α_{empty} by the surface of the absorber (the new surface will be: S - S_{abs}). It further adds the absorber's surface S_{abs} with the absorber's absorption coefficient α_{abs}. So the following can be stated:

$$\propto_{test} = \frac{\propto_{empty} \cdot (S - S_{abs}) + \propto_{abs} \cdot S_{abs}}{S}$$

[47] Note: This slightly modified formula omits the absorption of air.

This formula can be solved for α_{abs}:

$$\propto_{abs} = \frac{\propto_{test} \cdot S - \propto_{empty} \cdot (S - S_{abs})}{S_{abs}}$$

The final calculation is performed in a spreadsheet (see Figure C5) for all third-octave measurements. The reverberation time RT60 for the empty room and the room with absorber is shown in Figure C6. Finally, the result for α_{abs} is shown in graphical form in Figure C7.

Room height	2,91	m
Room depth	1,65	m
Room length	2,65	m
Room surface	33,771	m²
Room volume	12,724	m³
Absorber Surface	2	m²

Frequency band (Hz)	**125**	**160**	**200**	**250**	**315**	**400**	**500**	**630**	**800**	**1000**	**1250**
RT60: Echo chamber	1,11	1,12	1,24	1,22	1,11	1,19	1,10	1,10	1,03	1,22	1,14
RT60: Echo chamber with absorber	0,95	0,89	0,72	0,67	0,63	0,66	0,60	0,56	0,60	0,63	0,69
Alpha: Echo chamber	0,05	0,05	0,05	0,05	0,05	0,05	0,05	0,05	0,06	0,05	0,05
Alpha: Echo chamber with absorber	0,06	0,07	0,08	0,09	0,09	0,09	0,10	0,10	0,10	0,09	0,09
Alpha: Absorber	0,20	0,28	0,61	0,70	0,71	0,70	0,78	0,89	0,72	0,79	0,60

Frequency band (Hz)	**1600**	**2000**	**2500**	**3150**	**4000**	**5000**	**6300**	**8000**	**10000**	**12500**	**16000**
RT60: Echo chamber	0,90	0,85	0,92	0,88	0,88	0,83	0,76	0,64	0,56	0,46	0,40
RT60: Echo chamber with absorber	0,60	0,64	0,61	0,64	0,54	0,63	0,56	0,47	0,38	0,34	0,29
Alpha: Echo chamber	0,07	0,07	0,06	0,07	0,07	0,07	0,08	0,09	0,10	0,12	0,14
Alpha: Echo chamber with absorber	0,10	0,09	0,10	0,09	0,11	0,09	0,10	0,12	0,15	0,17	0,19
Alpha: Absorber	0,59	0,44	0,59	0,47	0,74	0,44	0,52	0,61	0,87	0,81	0,96

Figure C5: Calculation of the absorption coefficients as a result of the changes in RT60 between the empty room and the same room with the absorptive material mounted.

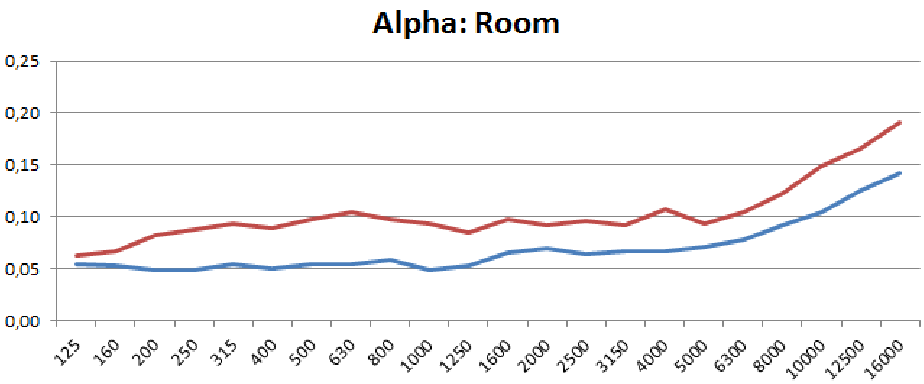


Figure C6: Frequency-dependent absorption coefficients for the whole room. Blue: Empty room. Red: Room with absorber (frequency axis to the right in third-octaves).

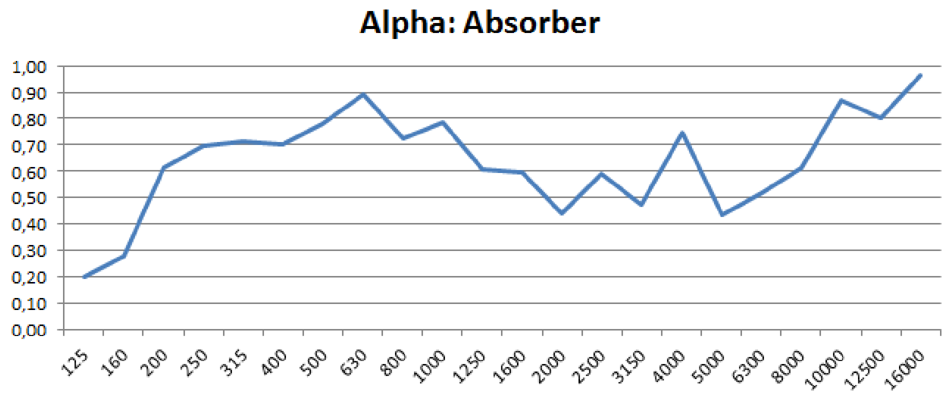


Figure C7: Frequency-dependent absorption coefficients of the test material determined by the echo-chamber measurement (frequency axis to the right in third-octaves).

Appendix D: Filter transfer functions for second-order band-pass filters and corresponding equaliser responses

Analogue band-pass filters

The second-order analogue band-pass filter has the following transfer function (adapted from Tietze and Schenk, 1980, p.303):

$$h(f) = \frac{G}{Q} \cdot \frac{S(f)}{S(f)^2 + \frac{S(f)}{Q} + 1}$$

which leads to

$$h(f) = \frac{G \cdot S(f)}{Q \cdot S(f)^2 + S(f) + Q}$$

where:

h(f): Complex transfer function. Note: h(f) = 1 represents the original signal

S(f): Complex frequency variable: $S(f) = \frac{i \cdot f}{f_0}$

i: Imaginary unit

f: Frequency

f_0: Centre frequency of the band-pass

G: Gain factor for the band-pass: G = 1 leads to unity gain (0dB) at the centre frequency

Q: Quality of the band-pass

The amplitude response and phase response of this band-pass can be seen in Figure D1.

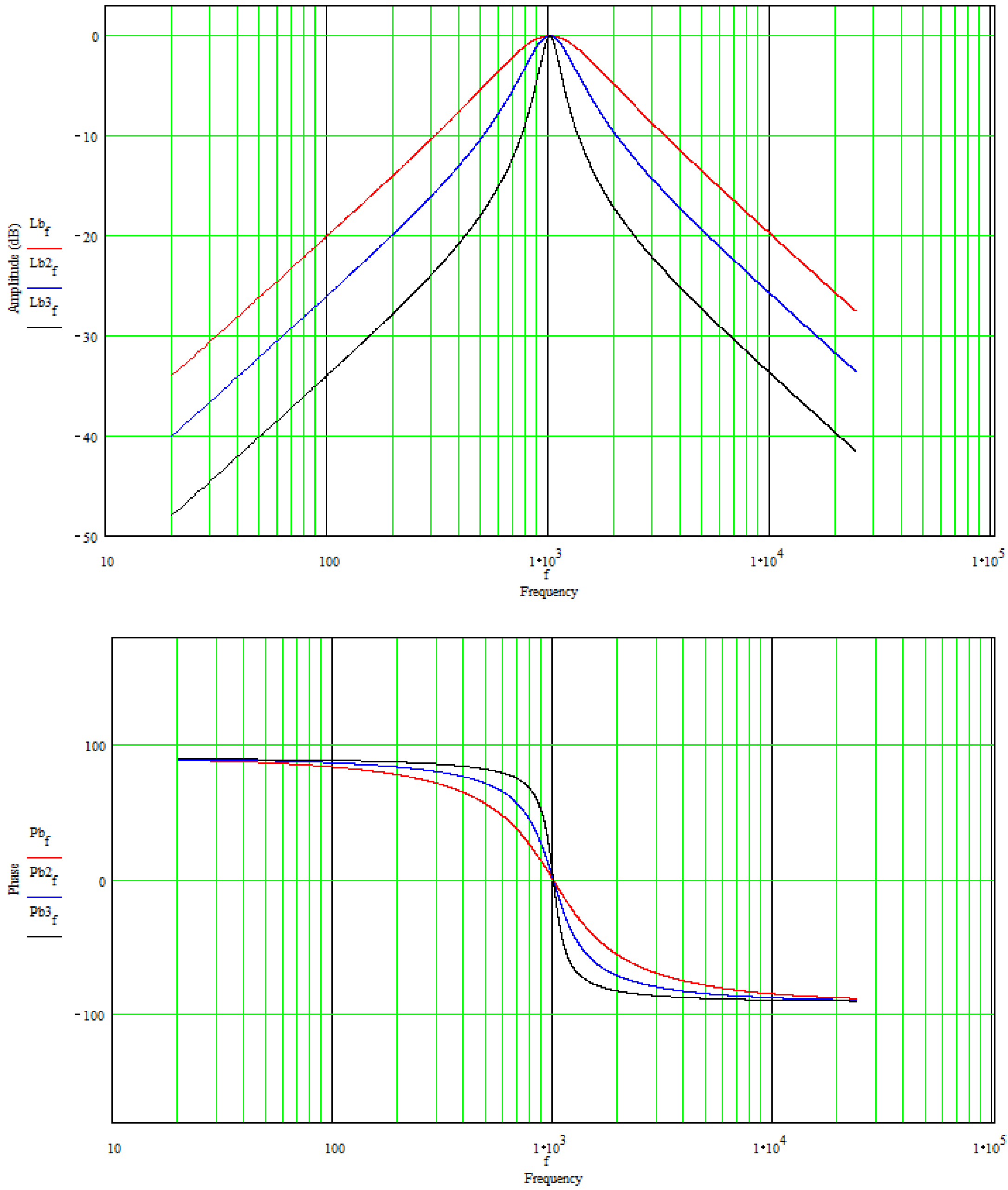


Figure D1: Amplitude response and phase response of an analogue second-order band-pass with f_0 = 1000Hz and Q=1 (red), Q=2 (blue) and Q=5 (black).

It can be seen that the band-pass has a gain of 0dB at the centre frequency and an attenuation towards higher and lower frequencies. The slope of the high- and low-frequency attenuation is approximating to 6dB/octave. Its phase response is 0° at the centre frequency and approximates +90° towards low and -90° towards high frequencies.

Building a parametric equaliser band from this band-pass is in principle an addition or a subtraction of the band-pass filtered signal from the unfiltered signal. However, in practice it is desired that cutting a frequency leads to symmetrical frequency responses compared to boosting a frequency. This leads to a slightly modified (red) formula for the cut process:

Boost:

$$hb(f) = 1 + h(f)$$

$$hb(f) = 1 + \frac{G \cdot S(f)}{Q \cdot S(f)^2 + S(f) + Q}$$

Cut:

$$hc(f) = 1 - \frac{G \cdot S(f)}{Q \cdot S(f)^2 + S(f) + G \cdot S(f) + Q}$$

where:
hb(f): Complex transfer function of the boosting equaliser band
hc(f): Complex transfer function of the cutting equaliser band
Having: G = 1 leads to ±6dB cut/boost
G = 2.16 leads to ±10dB cut/boost
G = 3 leads to ± 12dB cut/boost

Operating the boost band and the cut band in series leads to a mathematical multiplication of their transfer functions (for detailed calculations see end of this appendix). The result of this multiplication is:

$$ht(f) = hb(f) \cdot hc(f) = 1$$

So, the cut band can exactly undo what the boost band has added to the signal (see Figure D2).

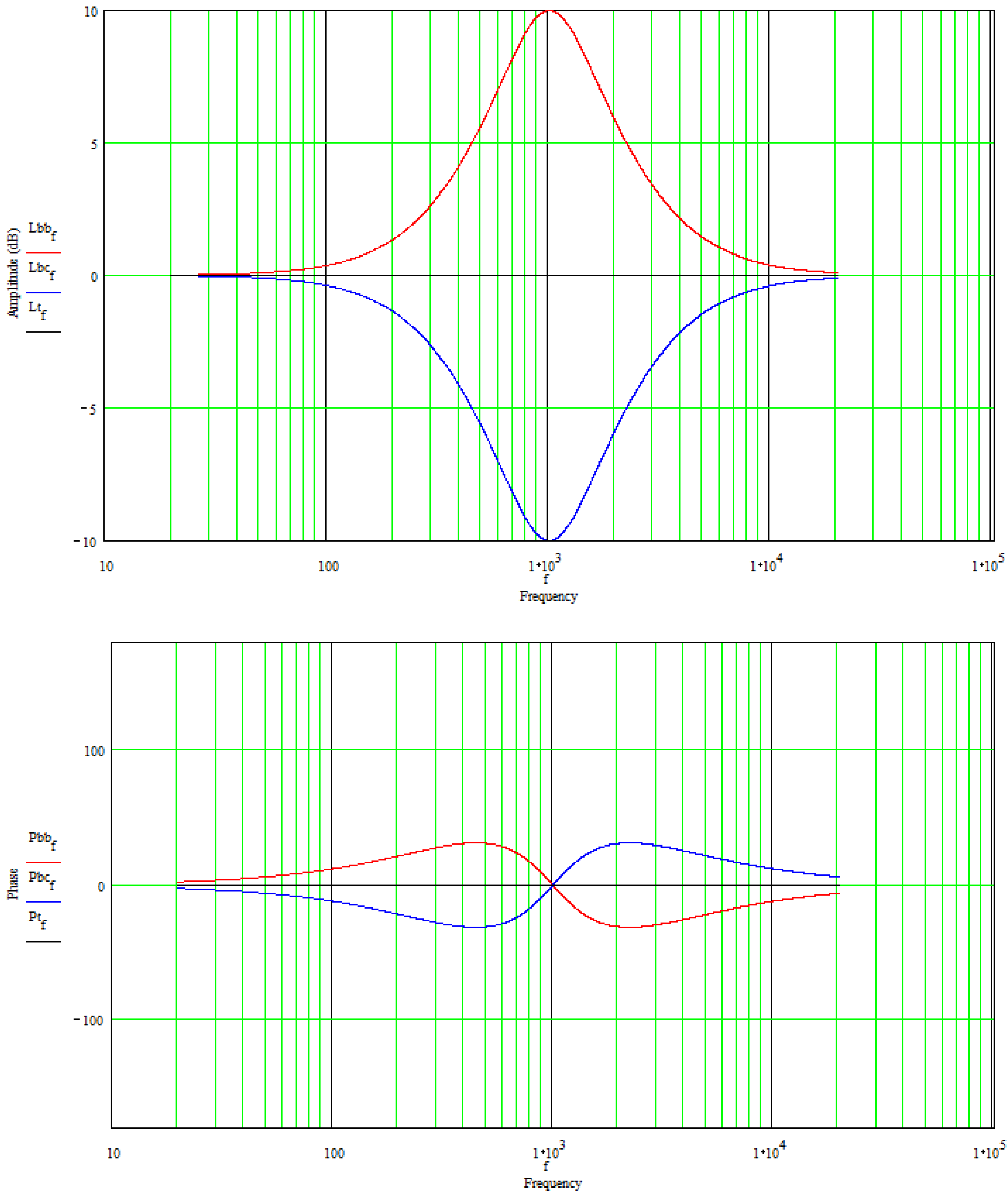


Figure D2: Amplitude and phase for an equaliser set to boost (red) and cut (blue) with Q=1 and G=2.16 (±10dB). Boosting and cutting in series leads to a flat amplitude and a flat phase (black).

Digital IIR filters

Digital band-pass filters can be calculated with IIR filters. IIR filters are filters with an infinite impulse response – hence their name IIR. Their impulse response is calculated using sample delays (z^{-1}) and multipliers in a feed-forward and a feedback section. A typical implementation of a digital second-order band-pass is the 'direct form' shown in Figure D3.

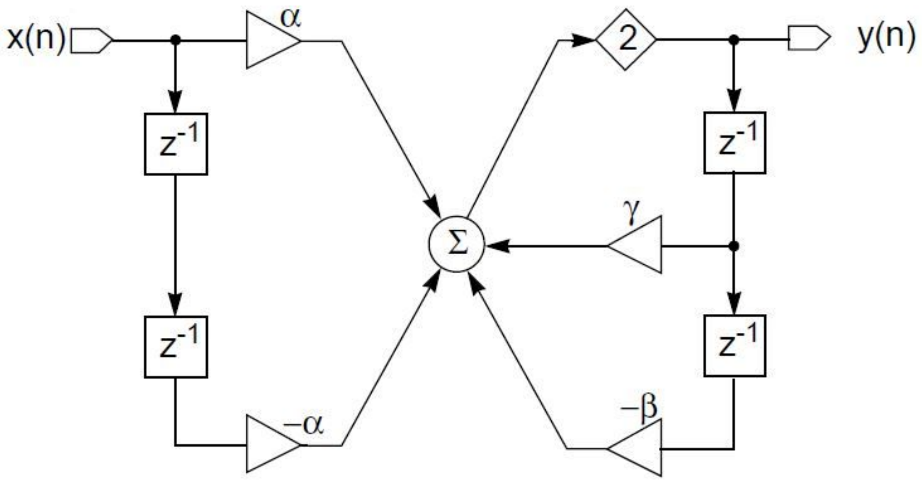


Figure D3: Network diagram of a direct-form second-order IIR band-pass filter (Lane and Hillman, 1993, p.2-24).

The transfer function for IIR filters is in the time-discrete Z-domain. The transfer from the analogue transfer function in the S-domain to the digital transfer function in the Z-domain for an IIR band-pass is well described in Lane and Hillman (1993) on pages 2-1 to 2-9. The coefficients α, β and γ have to be calculated using the sampling frequency, f_0 and Q. Formulas for this can be found in Lane and Hillman (1993) on page 2-24. Stimulating the filter with a 1-sample pulse in a maths program leads to an impulse response of the filter. This can be Fourier-transformed to analyse the amplitude and phase response (see Figure D4).

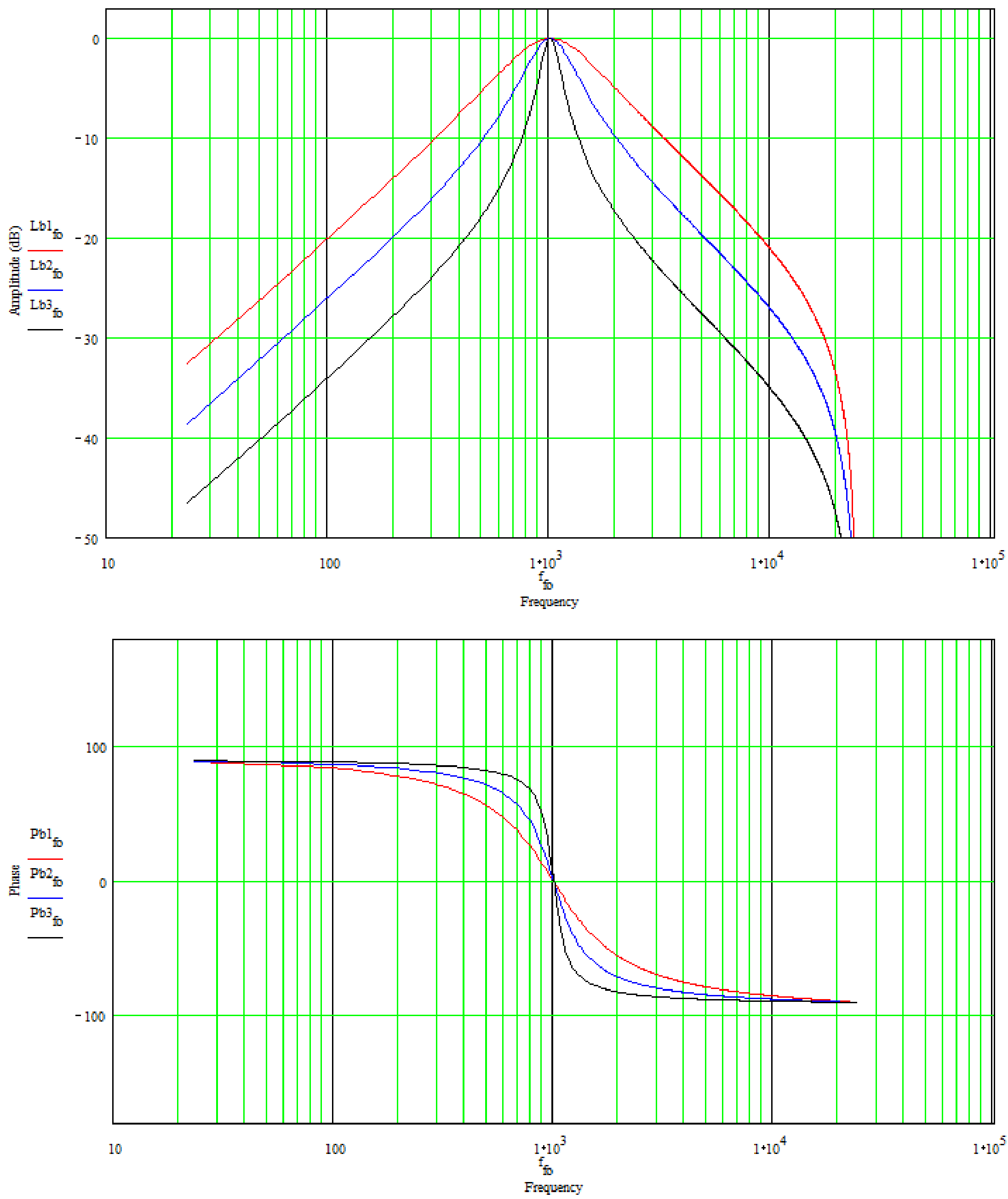


Figure D4: Amplitude response and phase response of a digital second-order direct-form IIR band-pass with f_0 = 1000Hz and Q=1 (red), Q=2 (blue) and Q=5 (black).

It can be seen that the frequency responses are comparable to the analogue ones except for the deviation near the Nyquist frequency[48]. At the Nyquist frequency the output of the IIR band-pass is null, which leads to minus infinity decibels in the amplitude plot. As Lane and Hillman (1993, p.5-8) state, "Imagine the zero at plus infinity in the analog response mapping into the zero at fs/2 in the digital case".

As the digital band-pass is comparable to the analogue one, the equalisers that are built with that band-pass are comparable as well. So both analogue and digital

[48] The Nyquist frequency is half of the sampling frequency.

equalisers can be used alternatively. Attention has to be paid at high frequencies where the IIR band-pass response differs from the analogue one if the Nyquist frequency is too close to 20kHz. Using sampling rates of 96kHz or higher will reduce this problem.

Digital FIR filters

FIR filters can be seen as a special form of IIR filters. They only have a feed-forward section and no feedback section. As there is no feedback they have a finite impulse response (FIR). Their implementation can be seen in Figure D5.

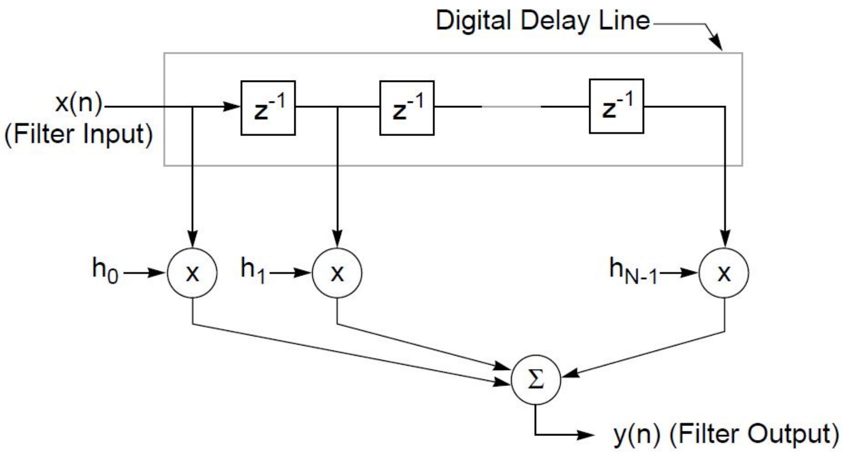


Figure D5: Network diagram of an FIR filter with N taps (Lane and Hillman, 1993, p.7-3)

In order to make the filter effective at least one period of the signal waveform must fit into the filter. That leads to very long (many taps) filters for low frequencies. For high frequencies these filters can have fewer taps. FIR filters are the same as a convolution as they sum up delayed and weighted samples as in convolution. They can convolve any impulse response to an audio signal regardless if it is an equaliser curve a room response or a reverb.

A special form of FIR filter (but by far not the only possible one) is a linear filter that has a symmetrical impulse response. This filter only changes the amplitude response of an audio signal and introduces a latency (that is, a delay) of half the filter length. So, it has a linear phase as a frequency-proportional phase shift is introduced (that is, a constant group delay).

In order to get the impulse response loaded into the coefficients of an FIR filter one could draw the desired amplitude response and phase response into a frequency plot and then Fourier-transfer this plot to an impulse response.

FIR filters can form band-passes or cut/boost equaliser bands comparable to the analogue ones. This can easily be achieved by loading the (truncated) impulse response of an IIR band-pass as coefficients into the FIR filter. However, the FIR filter would use far more processor cycles than an IIR filter (up to 1,000 times as many for very low audio frequencies or high Q-factors) without any benefit. Using a linear phase filter for calculating the band-pass filter is a questionable task as this setup could only correct the amplitude error but not the phase error introduced by an analogue equaliser. It would further introduce an unnecessary latency, as linear phase FIRs always have a symmetrical impulse response and therefore delay a signal by half the number of taps. For a better understanding of FIR filters, especially linear phase FIR filters, Lane and Hillman (1993) provide a good explanation in chapter 7.

Calculation of boost and cut in series:

For simplification: S = S(f)

$$\left(1+\frac{G\cdot S}{Q\cdot S^2+S+Q}\right)\cdot\left(1-\frac{G\cdot S}{Q\cdot S^2+S+G\cdot S+Q}\right)=$$

multiplication

$$1+\frac{G\cdot S}{Q\cdot S^2+S+Q}-\frac{G\cdot S}{Q\cdot S^2+S+G\cdot S+Q}-\frac{G^2\cdot S^2}{(Q\cdot S^2+S+Q)\cdot(Q\cdot S^2+S+G\cdot S+Q)}=$$

common denominator

$$\frac{(Q\cdot S^2+S+Q)\cdot(Q\cdot S^2+S+G\cdot S+Q)+G\cdot S\cdot(Q\cdot S^2+S+G\cdot S+Q)-G\cdot S\cdot(Q\cdot S^2+S+Q)-G^2\cdot S^2}{(Q\cdot S^2+S+Q)\cdot(Q\cdot S^2+S+G\cdot S+Q)}=$$

factoring out G·S in the red section

$$\frac{(Q\cdot S^2+S+Q)\cdot(Q\cdot S^2+S+G\cdot S+Q)+G\cdot S\cdot(Q\cdot S^2+S+G\cdot S+Q-Q\cdot S^2-S-Q)-G^2\cdot S^2}{(Q\cdot S^2+S+Q)\cdot(Q\cdot S^2+S+G\cdot S+Q)}=$$

$$\frac{(Q\cdot S^2+S+Q)\cdot(Q\cdot S^2+S+G\cdot S+Q)+G^2\cdot S^2-G^2\cdot S^2}{(Q\cdot S^2+S+Q)\cdot(Q\cdot S^2+S+G\cdot S+Q)}=$$

$$\frac{(Q\cdot S^2+S+Q)\cdot(Q\cdot S^2+S+G\cdot S+Q)}{(Q\cdot S^2+S+Q)\cdot(Q\cdot S^2+S+G\cdot S+Q)}=1$$

Appendix E: The possible impact of the crossover of the NS-10M Studio on the timing of low- and high-frequency drivers

The Yamaha NS-10M Studio uses a passive crossover with second-order high- and low-pass filters (Yamaha, n.d., p.4). The specifications claim that the crossover frequency is at 2kHz (Yamaha, n.d., p.3). The schematic of the crossover is shown in Figure E1. It can be seen that the drivers are connected with the same polarity to the output of the crossover. So, they are driven in phase, which means their signals can be added.

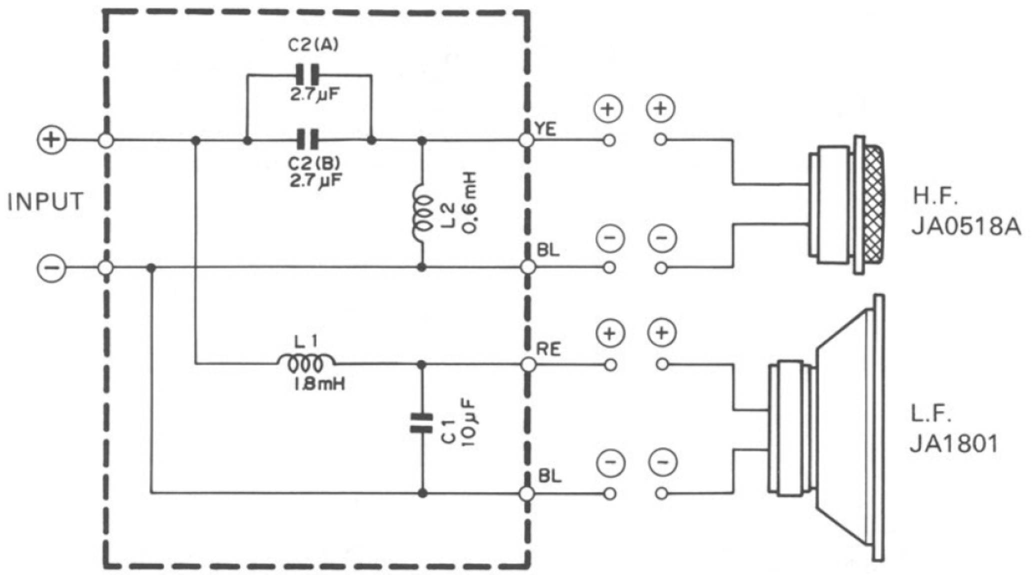


Figure E1: Schematic of the crossover of the NS-10M Studio (Yamaha, n.d., p.4).

The crossover should not be calculated on its own as the impedance of the drivers has to be considered as it puts a complex load on the output of the filters. According to Herget (n.d.) the complex load of each driver can be calculated as a network of resistors, inductors and one capacitor, as shown in Figure E2.

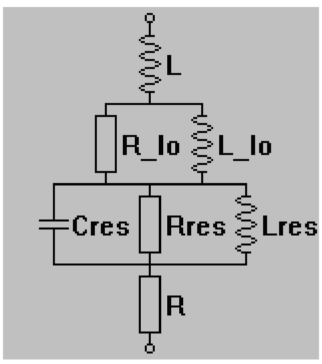


Figure E2: Schematic for the simulated impedance of a driver (Herget, n.d.).

For the calculation of this load, complex numbers have to be used. Resistors appear as real numbers, whereas inductors and capacitors appear as complex and frequency-dependent numbers:

$$R = R$$

$$X_L = i \cdot 2 \cdot \pi \cdot f \cdot L$$

$$X_C = \frac{1}{i \cdot 2 \cdot \pi \cdot f \cdot C}$$

where:
X: Impedance of the inductor or capacitor
i: Imaginary unit
f: Frequency
R: Resistance
L: Inductance
C: Capacity

Using these formulas the whole crossover with its connected drivers can be calculated as a network of complex resistors connected in series and in parallel. Measuring R and empirically determining L, R_lo, L_lo, Cres, Rres and Lres the impedance of the loudspeaker given in its datasheet can be approximated. Figure E3 shows the result of this calculation, where the simulated impedance is laid over the impedance plot from the datasheet. The simulation does not match exactly but it is the closest result that was possible with the simulation network from Figure E2.

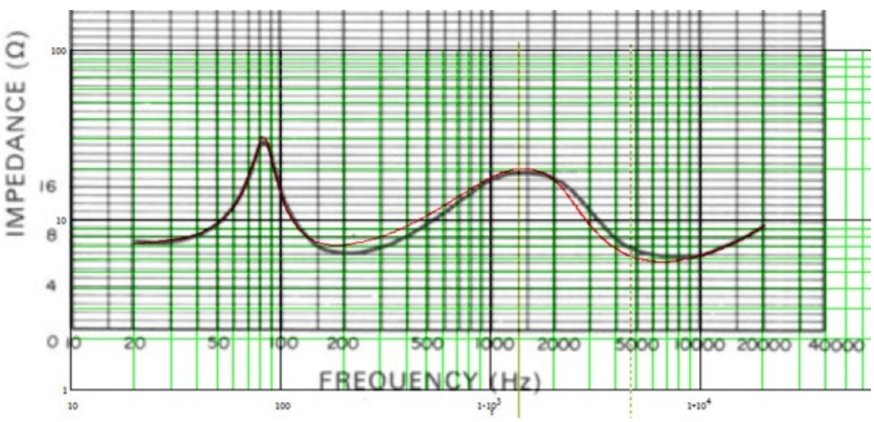


Figure E3: Simulated impedance (red) laid over the impedance plot from the NS-10M datasheet (Yamaha, n.d., p.3).

Using the empirically derived load data the frequency response of the crossover could be calculated. The result is shown in Figure E4.

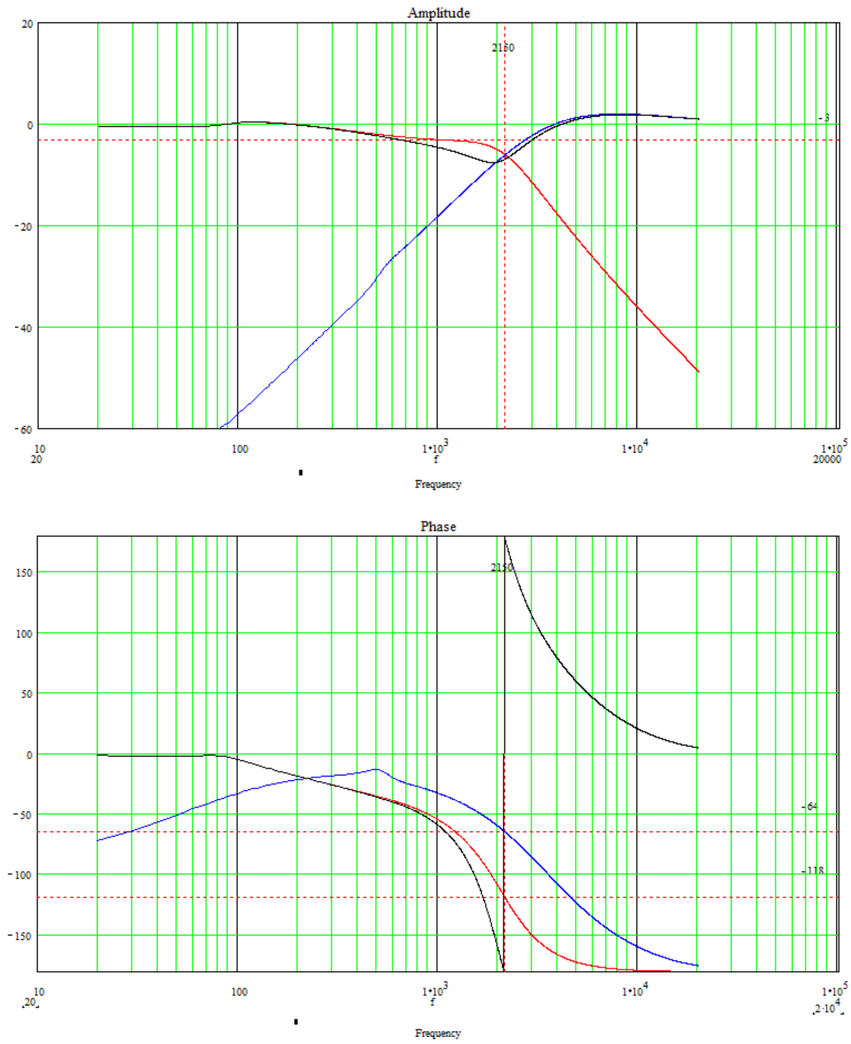


Figure D4: Amplitude response and phase response of the NS-10M crossover (with simulated driver loads). Red: woofer. Blue: Tweeter. Black: Sum of woofer and tweeter.

It can be seen that at the crossover frequency 2.15kHz there is a large phase shift between the high- and the low-pass output. The high-pass output (blue) has less delay (-64°) than the low-pass output (red, -118°). The phase difference is 54° and this would lead to a time delay of:

$$t = \frac{\varphi}{f \cdot 360°} = \frac{54°}{2150Hz \cdot 360°} = 70\mu s$$

This delay would correspond to 2.4cm distance of the woofer behind the tweeter.

Appendix F: Testing the equalisation of loudspeakers with FIR filters

A possible and useful application for an FIR filter is the convolution of custom impulse responses from for example measured and inverted loudspeaker responses. This way both amplitude errors and phase errors of the loudspeaker can be corrected at once. An example of this procedure is shown in Figure F1. On the left diagrams, there is the impulse response, in the middle ones the amplitude response and on the right ones the phase response of each stage. In the top section, the measured signal from an arbitrary loudspeaker is shown. Its amplitude and phase responses are inverted (middle section). The resulting impulse response is shown in the middle left diagram. Next, a convolution (Smith, 1999, p.120) of the upper left with the middle left impulse response is calculated. The result is in the lower section on the left. It is a near-ideal impulse, delayed by the impulse length of the input impulse response. The amplitude response is quite flat and the phase response is quite linear (bottom section). The remaining irregularities come from the quantisation noise that occurs in the original measurement. Please bear in mind that the lower section phase diagram is on a linear frequency scale. The phase is hard to evaluate as it alters between 0° and close to 180° or −180°. It would be perfectly linear if it always went to 180° (or −180°, which is the same).

One problem with this way of correcting loudspeakers is the latency that is introduced into the signal in order to correct all phase errors. This method of correction would therefore not be useable for any applications needing low-latency or latency-free loudspeaker responses. Another problem is the pre-ringing that can be seen in the impulse response of the correction filter.

A similar test was undertaken with an artificial and therefore noiseless impulse response from a mathematically calculated equaliser band. This can be seen in Figure F2. The calculated impulse response of an equaliser boost band (IIR band-pass added to the direct signal) was exported (top section). Again, amplitude and phase were inverted (middle section). After the convolution of the input impulse response with the correction impulse response the ideal impulse appeared (bottom section). This corresponds to the result in Appendix D, where boost and cut of analogue or digital band-pass filters can undo each other. It further shows that without any noise the correction works perfectly.

In the last test of this appendix the loudspeaker measurement from the first test was taken again (see Figure F3, upper section). This time only the amplitude was inverted but the phase was generated artificially and noiseless in order to simulate a linear phase correction filter. After the convolution of the measured impulse response with the correction impulse response the result is a near-flat amplitude response. It is even steadier than the one in the first test. This is a positive impact from the noiseless artificial phase response of the correction stage. However, the phase in the corrected output is not linear (not even close to it). This can also be seen in the output impulse response, which is far from an ideal impulse.

So when the task is to do an ideal correction of a frequency response then:

- A low noise measurement has to be undertaken.
- The correction filter has to be fed with an individual amplitude response and an individual phase response.

A linear phase filter would not bring the desired perfect signal correction that could be reached with inverted amplitude and phase responses.

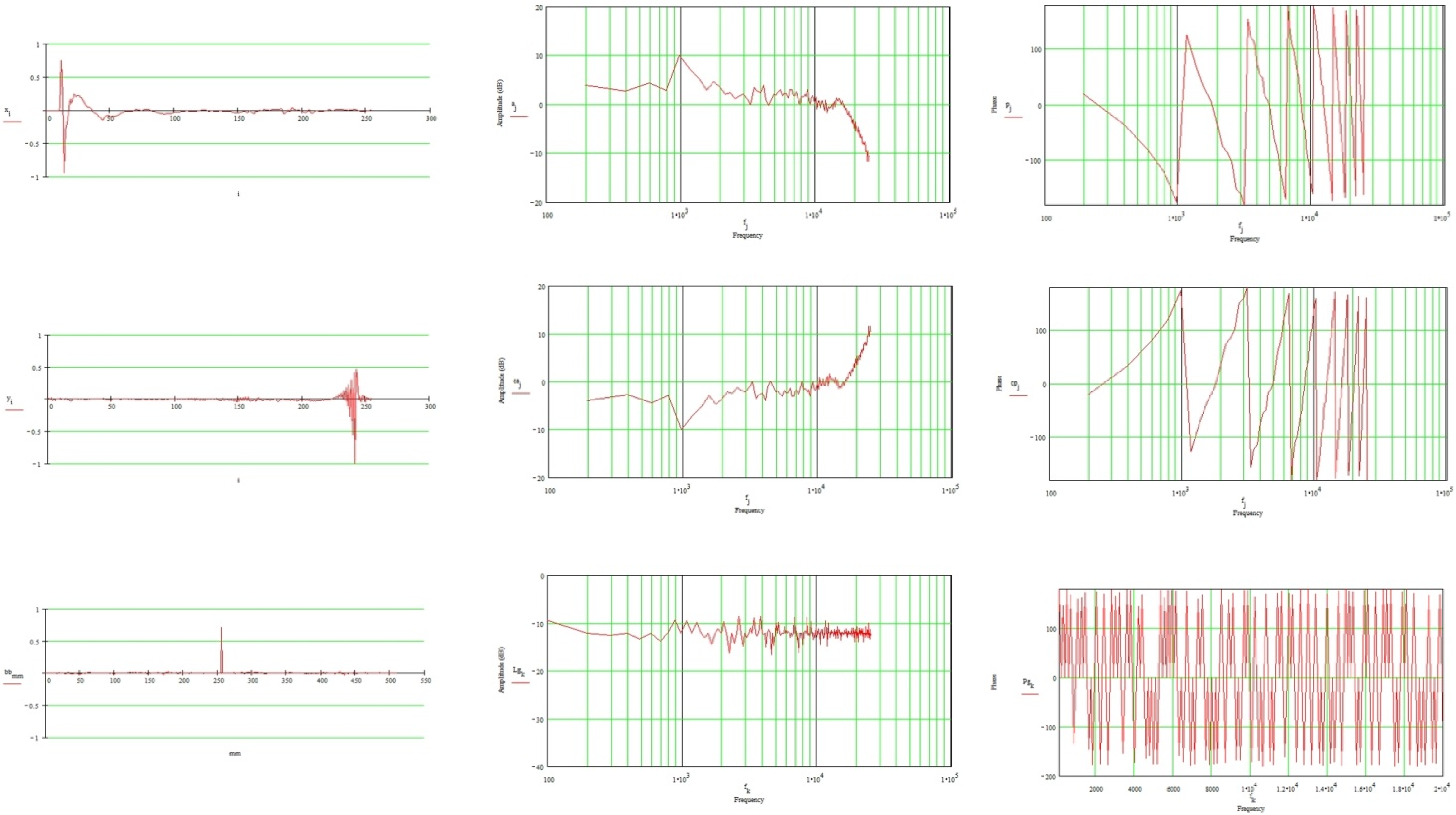


Figure F1: A mathematically simulated loudspeaker correction with a 256-tap FIR filter. The top section shows the original loudspeaker measurement. The middle section shows the filter derived from an inverted frequency response of the measurement. The lower section shows the result of the filtering. On the left: Impulse response. Middle: Corresponding frequency response. Right: Phase response. Please bear in mind the linear frequency scale on the lower right phase response.

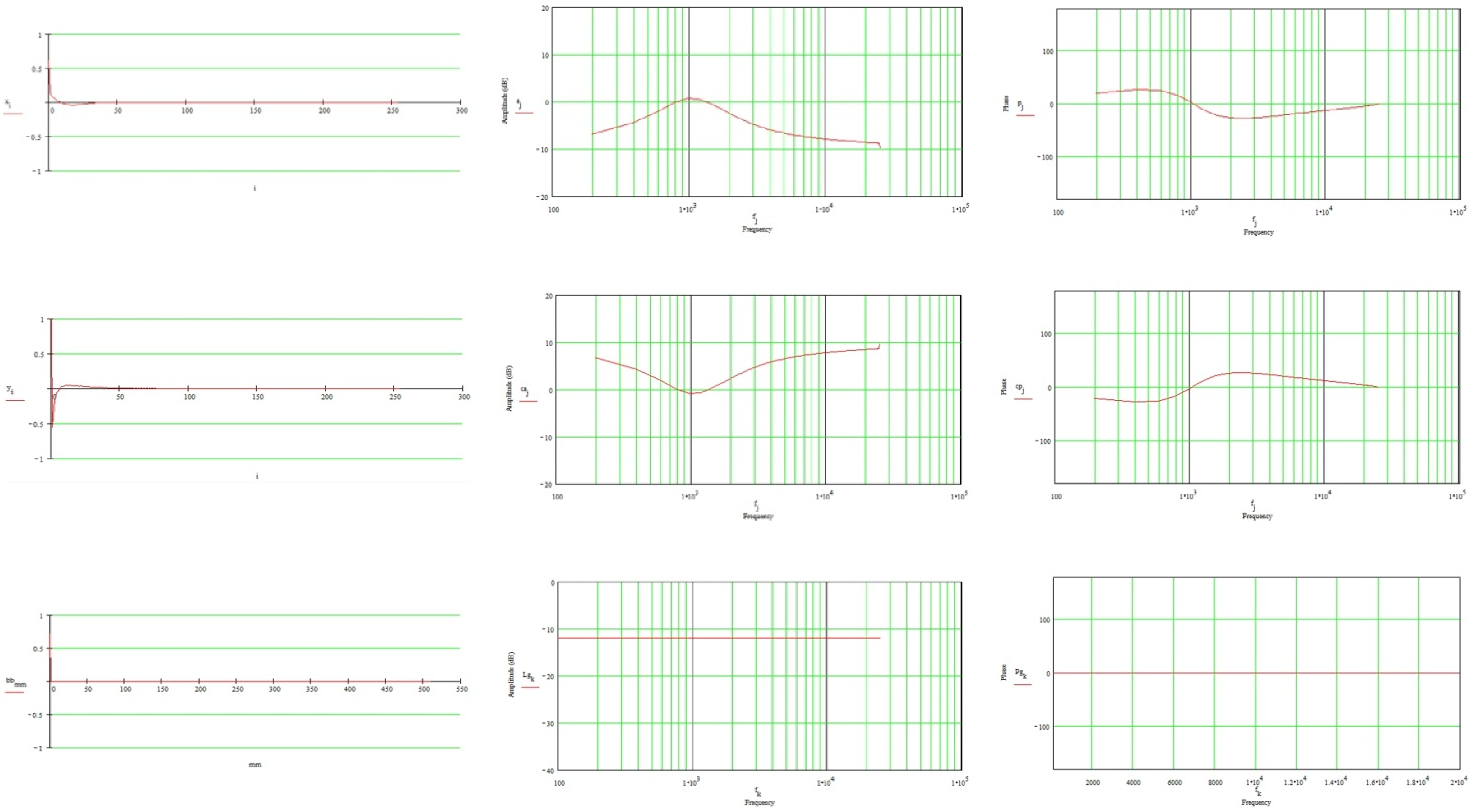


Figure F2: A mathematically simulated equaliser correction with a 256-tap FIR filter. The top section shows a calculated impulse response of an IIR boost band. The middle section shows the filter derived from an inverted frequency response of the boost band. The lower section shows the result of the filtering. On the left: Impulse response. Middle: Corresponding frequency response. Right: Phase response. Please bear in mind the linear frequency scale on the lower right phase response.

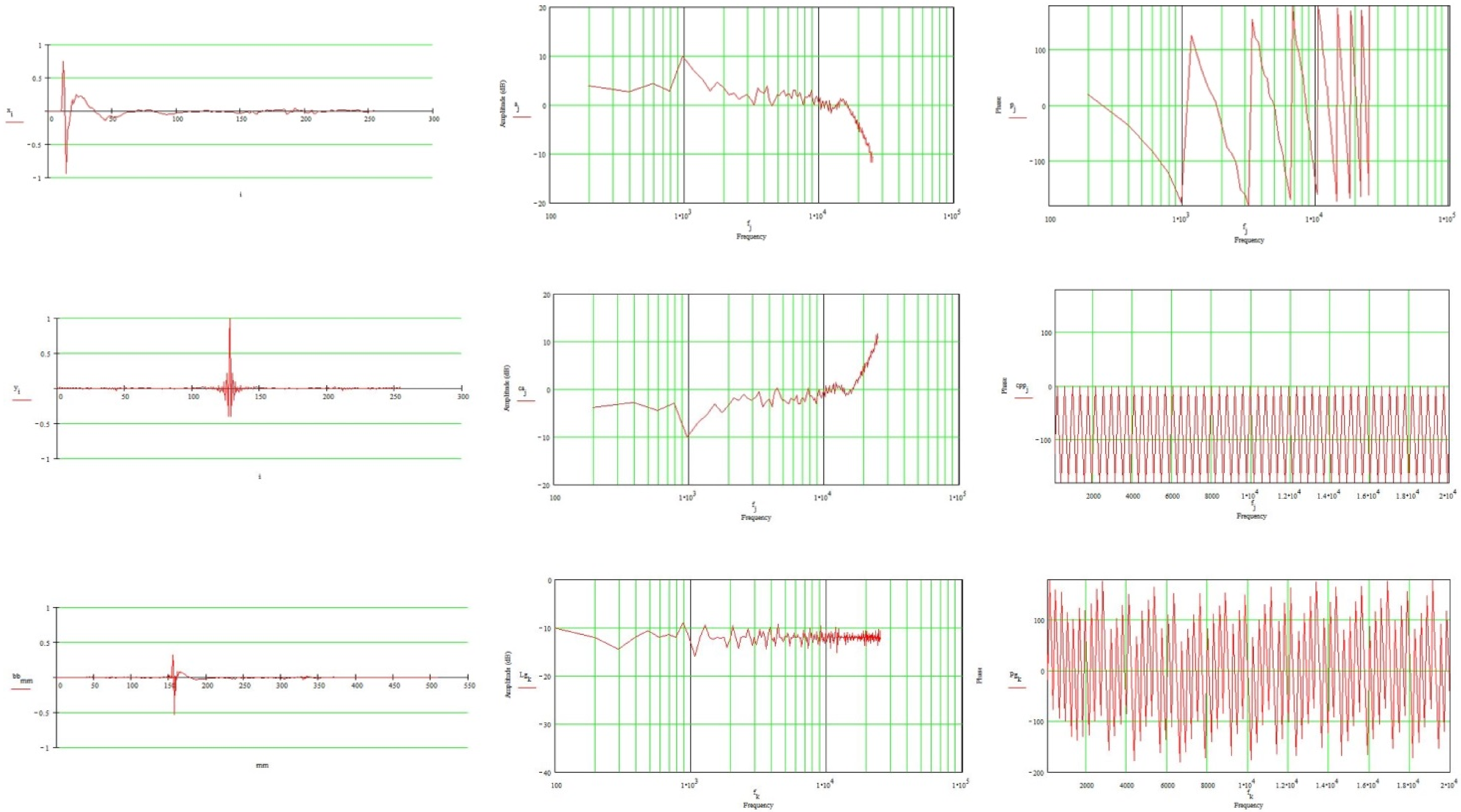


Figure F3: A mathematically simulated loudspeaker correction with a 256-tap linear phase FIR filter. The top section shows the original loudspeaker measurement. The middle section shows the filter derived from an inverted amplitude response of the measurement combined with an artificial linear phase response. The lower section shows the result of the filtering. On the left: Impulse response. Middle: Corresponding frequency response. Right: Phase response. Please bear in mind the linear frequency scale on the middle and lower right phase responses.

Appendix G: Construction details of the improved coaxial loudspeaker

The improved loudspeaker construction uses a 4" coaxial driver from SB-Acoustics. According to the datasheet (see Figure G1) the free-air resonance of the woofer is at 58Hz with a Vas of 4.8 litres. Mounted in a box with 2 litres volume the resonance will rise to 107Hz (detailed calculation at the end of this appendix). The tweeters' resonance frequency can be seen at 1.3kHz. The structural breakdown of the membrane of the woofer starts at about 5kHz, and partial oscillations in form of larger bumps in the frequency diagram become visible. The crossover frequency sounded best at 2.5kHz. According to Figure G1 this is a frequency where both woofer and tweeter work without problems.

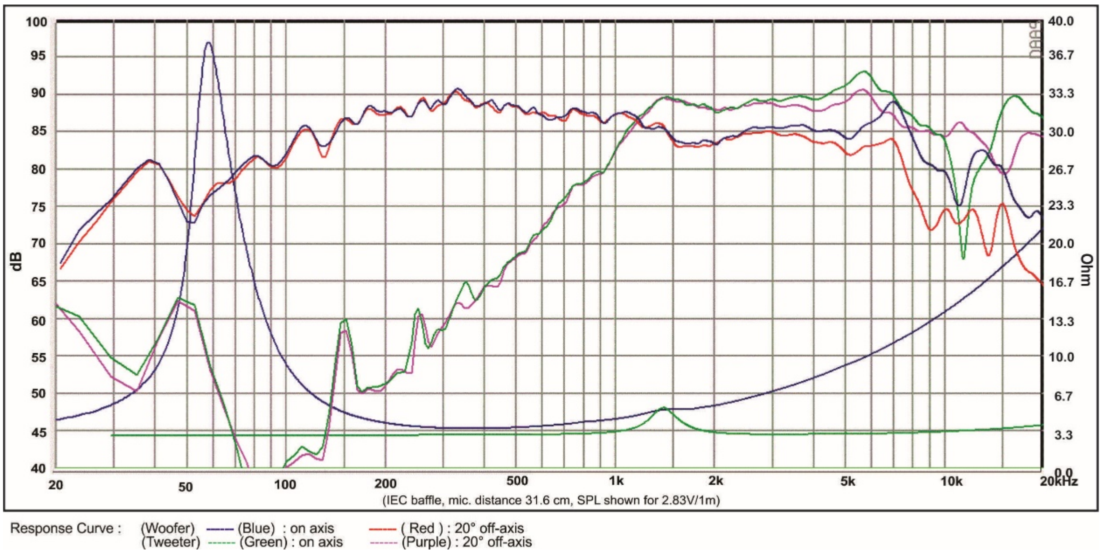


Figure G1: Amplitude diagrams for the SB12PFC25-4-COAX (SB-Acoustics, 2020).

The newly designed port was folded twice in the box. This construction could be seen in Figures G2 and G3. The total length of the port is 35cm. This led to a measured resonance frequency of 46Hz (detailed calculation and measurement at the end of this appendix).

Before mounting the driver, the lower, left, right and back walls of the box were covered with 2cm Basotect foam. This foam absorbs reflections inside the box and therefore avoids standing waves. It further broadens the port-resonance as the Q of the Helmholtz resonator is decreased. The finished box can be seen in Figure G4. Figure G5 shows the equalisation of the woofer and Figure G6 the impact of time alignment on the phase response. Finally, Figure G7 presents the overall frequency response with and without equaliser. Especially from 9000Hz to 14000Hz the linearization of the phase by the equaliser can be seen.

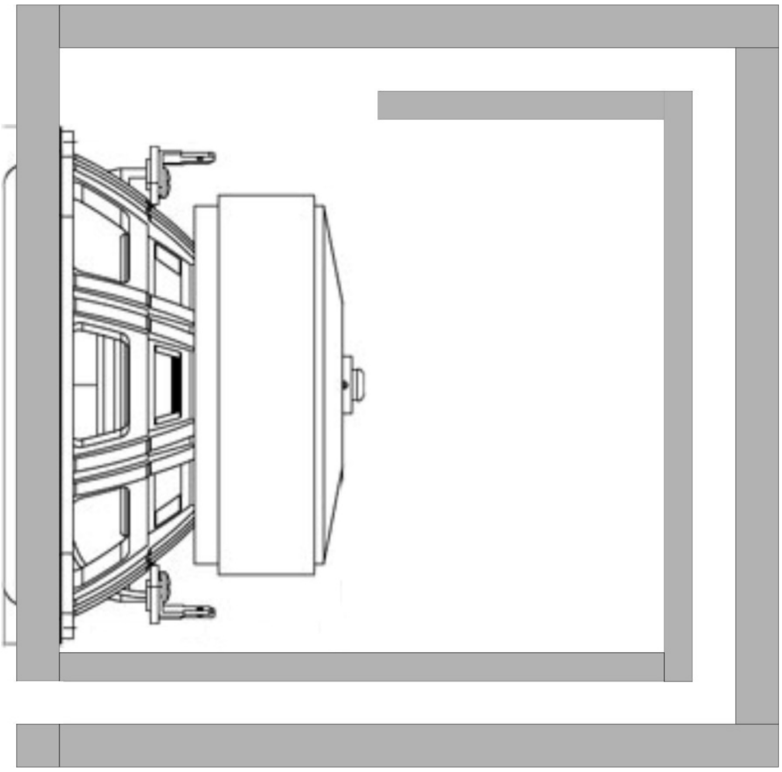


Figure G2: Part of the construction plan for the new loudspeaker. (Embedded driver drawing from SB-Acoustics (2020)).

Figure G3: The folded port in the half-finished enclosure.

Figure G4: The finished loudspeaker (frontal view).

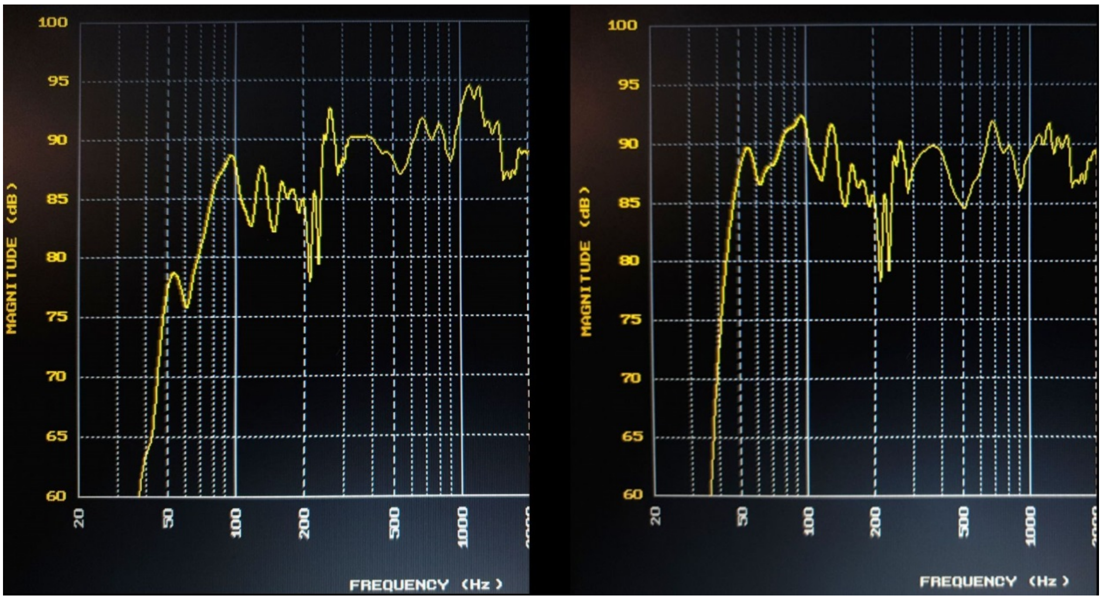


Figure G5: Woofer frequency response of the new loudspeaker. Left diagram: without equaliser, right diagram: with equaliser. The notches around at 220Hz were not corrected as they are too narrow to be a resonance error.

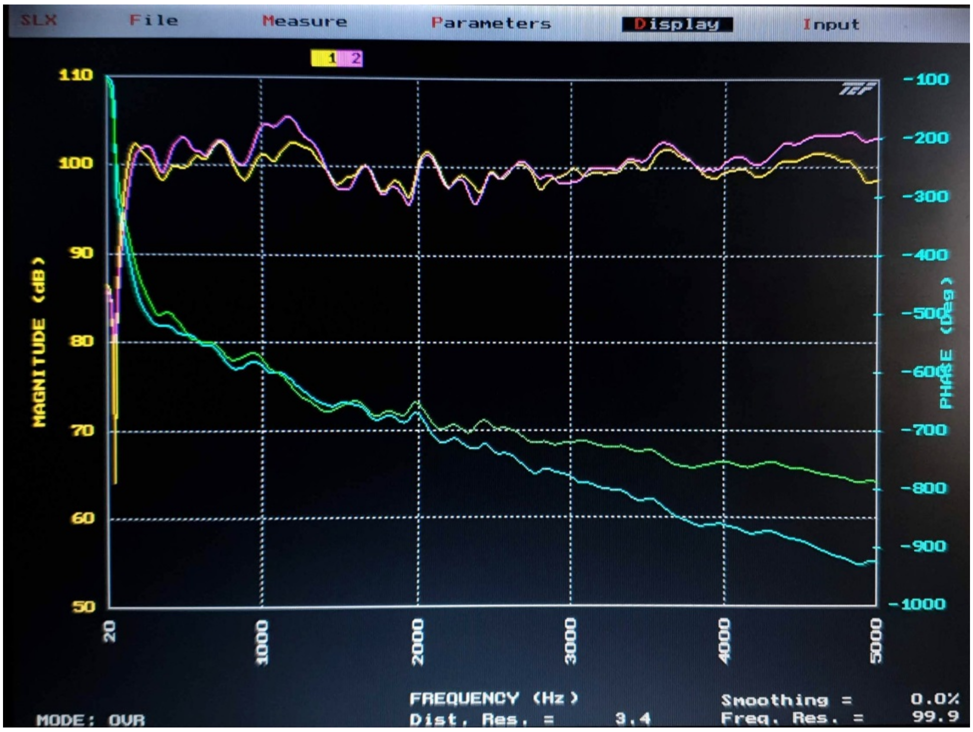


Figure G6: Frequency response of the new loudspeaker with and without time alignment and equaliser (up to 5000Hz). Without time alignment and equaliser: pink: level & green: phase. With time alignment and equaliser: yellow: level & blue: phase. The time alignment delays the tweeter (crossover frequency 2,5kHz) and so the phase (blue) gets closer to linear. Around 1200Hz the decrease in level from the equaliser also linearises the phase.

Figure G7: Frequency response of the new loudspeaker with and without equaliser (full range). Without equaliser: pink: level & red: phase. With equaliser: yellow: level & white: phase. It could be seen that between 9000 and 14000 Hz the equaliser linearises the phase while correcting the amplitude errors.

Detailed calculations

Basics:

The resonance frequency for a spring-mass system is:

$$f_r = \frac{1}{2 \cdot \pi \cdot \sqrt{m \cdot C}}$$

where:

f_r: Resonance frequency

C: Compliance (inverse of spring constant)

m: Mass

The compliance for a closed air volume is given by (Clemen, 1998, p.16):

$$C = \frac{V}{S^2 \cdot c^2 \cdot \rho}$$

where:

V: Volume

S: Surface of the membrane or port

c: Speed of sound (343m/s (20°C))

ρ: Volumic mass of the air (1.2kg/m³ (20°C))

The loudspeaker cabinet has a net volume of approx. 2 litres = 0.002m³:

2.8l (empty box) – 0.5l (folded port) – 0.3l (driver) = 2l

Woofer resonance:

According to the datasheet (SB-Acoustics, 2020) the free-air resonance of the woofer is 58Hz, the woofer has an equivalent volume of 4.8 litres, a moving mass of 4.5 g and a piston surface of 45 cm². This leads to a compliance of 0.00167 m/N which is confirmed by the datasheet which provides a compliance of 0.00166 m/N.

Mounting the woofer in a box with 2 litres volume leads to a new volume of:

$$V_t = \frac{1}{\frac{1}{Vas} + \frac{1}{Vb}} = \frac{1}{\frac{1}{4.8l} + \frac{1}{2l}} = 1.4l$$

where:

Vt: New 'total Volume' for the woofer in a box

Vas: Equivalent volume of the woofer

Vb: Volume of the box

The new compliance for this volume will be:

$$C_t = \frac{V_t}{S^2 \cdot c^2 \cdot \rho} = \frac{0.0014m^3}{0.0045m^2 \cdot \left(\frac{344m}{s}\right)^2 \cdot 1.2\frac{kg}{m^3}} = 0.00049\frac{m}{N}$$

The new resonance frequency is therefore:

$$f_r = \frac{1}{2 \cdot \pi \cdot \sqrt{m \cdot C_t}} = \frac{1}{2 \cdot \pi \cdot \sqrt{0.0045kg \cdot 0.00049\frac{m}{N}}} = 107Hz$$

Port-resonance:

The port in a vented box is a Helmholtz resonator (Stark, 2003, p.117). The resonance frequency of a Helmholtz resonator is also given by a mechanical spring-mass system and the formula for the compliance will be used here, as well. The compliance will be calculated with the volume of the box and the surface of the port. The mass of the air in the port is:

$$m = S \cdot l \cdot \rho$$

where:

m: Mass of the air in the port

S: Surface of the port

l: Length of the port

ρ: Volumic mass of the air (1.2kg/m³ (20°C))

This can be solved for f_r:

$$f_r = \frac{1}{2 \cdot \pi \cdot \sqrt{\frac{V \cdot S \cdot l \cdot \rho}{S^2 \cdot c^2 \cdot \rho}}} = \frac{1}{2 \cdot \pi \cdot \sqrt{\frac{V \cdot l}{S \cdot c^2}}} = \frac{1}{\frac{2 \cdot \pi}{c} \cdot \sqrt{\frac{V \cdot l}{S}}} = \frac{1}{0.0183 \cdot \sqrt{\frac{V \cdot l}{S}}} = 55 \cdot \sqrt{\frac{S}{V \cdot l}}$$

The port has a surface of 9mm · 90mm = 0.00081m²

The length of the port is 0.35m.

This leads to a theoretical resonance frequency of:

$$f_r = 55 \cdot \sqrt{\frac{0.00081}{0.002 \cdot 0.35}} = 59Hz$$

In practice the length of the port appears longer as there is air at both ends that is carried along with the air in the port (mouth correction) and therefore adds to the mass of the spring-mass system. For small ports (like this) the increased length could be many centimetres on both ends. The so-called 'mouth correction' of a small rectangle port like this is not well studied and therefore the apparent length of the port must be determined empirically. Further, the volume of the box appears larger as there is damping material inside that acts like a change in the volumic mass of the air. Increasing the port length and the volume of the box will lower the resonance frequency. For approximating the port-resonance a test cabinet was built with a port that was 25cm long and only folded once. Measuring the resonance frequency repeatedly the port was lengthened by 2cm in multiple turns. The desired resonance frequency was 45Hz. At 35cm (46Hz) the port reached its maximum length as there should still be some free air travel at the back opening of the port. This length was kept for the final version. The measurement in Figure G6 shows a measurement 2mm in front of the woofer's membrane (yellow) and a measurement at the opening of the port (pink). The membrane excursion will be at a minimum at the port-resonance frequency (Panzer, 1994, p.41) – this should be visible as a dip in the amplitude of the woofer. This minimum can be seen at about 46Hz. Normally, the signal from the opening of the port (pink) shows a band-pass characteristic with its maximum at the resonance frequency. This cannot be seen here. The port clearly emits energy below the woofer's frequency range but also towards higher frequencies, which is untypical. Note: The two measurements cannot be added to each other as the distance to the membrane was different.

Figure G6: Frequency response of the woofer (yellow) and the port (pink) in the bass range.

The performance of the new port was measured and compared to the problematic port in the first version of the coaxial speaker (see discussion in subsection 4.3.2.3 on page 76). It was expected that the new port with a larger surface area should have fewer air compression problems. For this measurement the resonance frequency of 46Hz was applied to the loudspeaker. The output signal of the power amplifier and the sound pressure near the port was recorded with a 2-channel storage oscilloscope (see Figure G7).

The microphone was positioned 90° to the side of the port opening and 2.5cm away from the port opening to avoid capturing wind noise from the air movement in front of the port (see Figures G8 and G9).

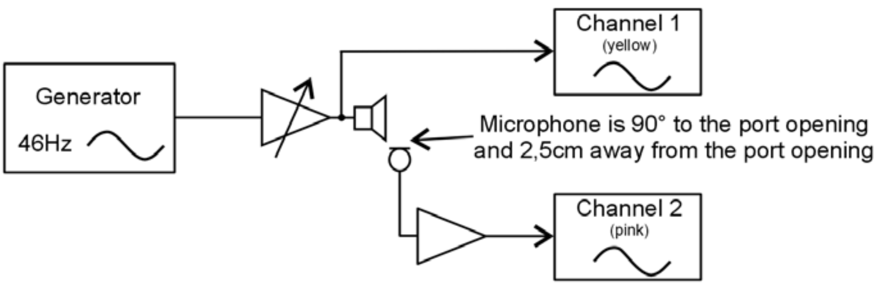


Figure G7: Measurement setup for the port measurements

Figure G8: Measurement positions for measuring the old loudspeaker with the round ports.

Figure G9: Measurement positions for measuring the new loudspeaker with the rectangular port.

The output voltage from the amplifier was increased in steps from 0.2V to 0.5V, 1.0V, 1.5V, 2.0V and 2.5V. For better comparison the measurements from both loudspeakers are combined in one figure for each voltage step. The old one is on the left and the new one is on the right of each figure (see Figures G10 to G16).

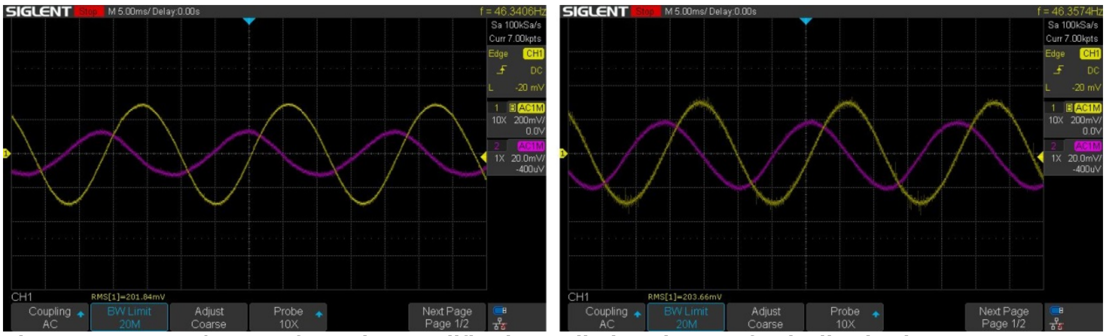


Figure G10: 46Hz/200mV from the amplifier is applied to the woofer (yellow). The port response of the old ports (left) and of the new port (right) is shown in pink.

Figure G11: 46Hz/500mV from the amplifier is applied to the woofer (yellow). The port response of the old ports (left) and of the new port (right) is shown in pink. With the old port, a slight distortion can be seen.

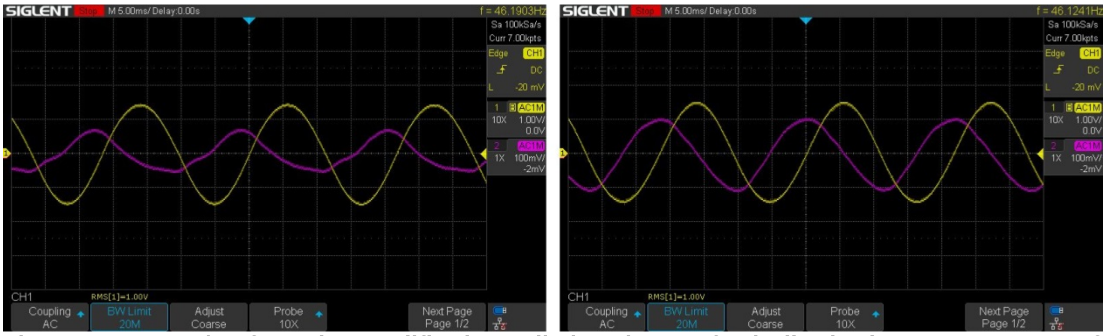


Figure G12: 46Hz/1V from the amplifier is applied to the woofer (yellow). The port response of the old ports (left) and of the new port (right) is shown in pink. With the old port, the distortion is clearly visible; with the new port, a slight distortion can be seen.

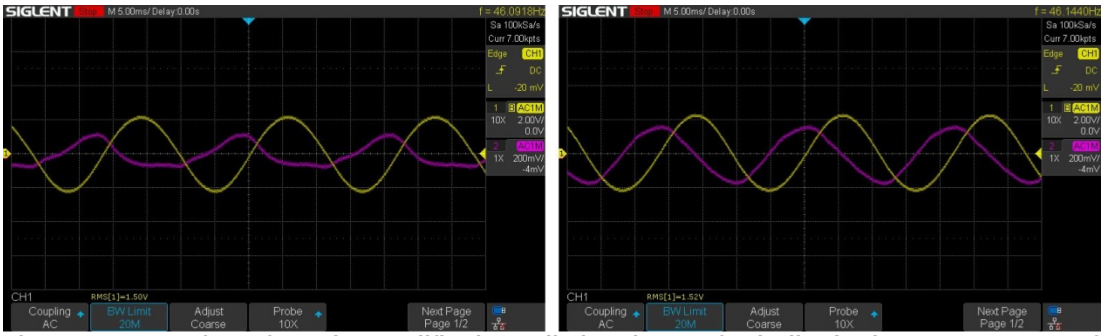


Figure G13: 46Hz/1.5V from the amplifier is applied to the woofer (yellow). The port response of the old ports (left) and of the new port (right) is shown in pink. With the old port, the distortion is strong; with the new port, a slight distortion can be seen.

Figure G14: 46Hz/2V from the amplifier is applied to the woofer (yellow). The port response of the old ports (left) and of the new port (right) is shown in pink. With the old port, the distortion is strong; with the new port, the distortion gets slightly more visible than in Figure G13.

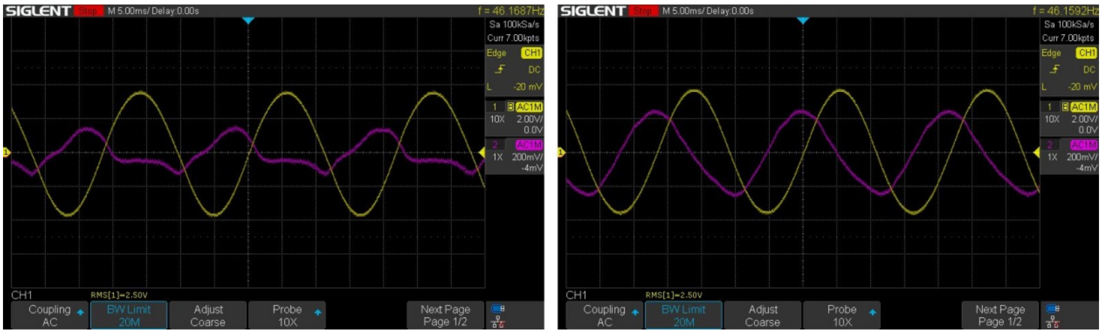


Figure G15: 46Hz/2.5V from the amplifier is applied to the woofer (yellow). The port response of the old ports (left) and of the new port (right) is shown in pink. With the old port, the distortion is strong; with the new port, the distortion is comparable to that in Figure G14. The output of the new port rises proportionally to the input signal while the output from the old port does not.

Figure G16: 46Hz/2.8V from the amplifier is applied to the woofer (yellow). The amplifier clips in the positive half-wave. The port response of the new port is shown in pink. While the output of the port still rises proportionally to the increase of the input, additional distortion is seen in the output during the period when the input signal clips.

It could be observed that at 0.2V the output from the port (pink) has a waveform comparable to the input signal (yellow). At 0.5V the first distortions appear with the old port. The signal from the new port still looks unsuspicious. At 1.0V the old port already produces a clearly distorted signal while first small irregularities can be seen with the new port. At 1.5V and above the old port produces strong distortion. With the old port at 2.0V and 2.5V it can be observed that distortion gets stronger and the output of the port is not increasing proportionally with the increasing input signal.
The new port produces small distortion up to 2.5V and its output increases proportionally with the input signal. Finally, in Figure G16 at 2.8V the amplifier shows clipping in the positive half-wave which can also be seen as an additional distortion on the port signal.

In conclusion, it can be stated that the new port is working across the whole power range (up to amplifier clipping) with only little distortion. Compared to the behaviour of the old port this is a large improvement.

Appendix H: Examples of feedback on the Time Alignment Rings

The following examples (Figures H1-H6) contain some recent feedback received on the Time Alignment Rings.

Tag Herr Friesecke,

Frohes Neues!

Ich wollte nur einmal nachhören, ob es noch irgendeine Möglichkeit gibt an Time Alignment Ringe zu gelangen?
Habe die alten NS-10 von [redacted] aus München (vielleicht kennt ihr euch, keine Ahnung) und bin absolut begeistert von den Ringen, hätte gern welche auf Reserve.

Figure H1: A spare pair was ordered as the customer was "… impressed by the rings"

hab heute schon alles erhalten und umgebaut. Danke für die prompte Abwicklung!

Muß sagen, bin wirklich überrascht über das Ergebnis. Kenne die NS10 (M Pro), mit all den "Für und Wider",
ja wirklich schon lange, aber es ist alles "entspannten", im unteren Bereich differenzierter, Räume besser gestaffelt,
obwohl mein Arr.-Raum alles andere als optimal ist (tw. Dachschräge).

Hat mich sehr gefreut :-)

Figure H2: The customer is "… surprised about the result […] the sound is more relaxed with a more differentiated low end" and "depth imaging is better"

2. Herr Friesecke, ihr Time-Alignment Produkt ist der „Wahnsinn“. Genau wie sie sagten: Tiefenstaffelung. Alles klingt wie „entzerrt“, „ach, genau so sollte es sein“. Eines der wenigen Sachen, wo ich mir sicher bin, daß ich nicht zurückgehen möchte. (Ich frage mich, wie 99% der heutigen NS-10 Nutzer, OHNE hören können).

Figure H3: The customer says "…the product is amazing" and "… wonders how users could hear on these speakers without the modification kit"

Sehr geehrter Herr Friesecke,

habe mir vor kurzem ein "neues" Päärchen NS-10 gekauft auf denen sich Time Alignment Rings befinden. Kannte diese bis dato überhaupt nicht, finde das Konzept jedoch interessant und richtig ;)

Nun habe ich jedoch eine Frage, da ich zu diesen Ringen keine Dokumentation finde:

Figure H4: The customer bought a used pair of loudspeakers with the modification kit already attached. He states that "... the concept is interesting and right" and wants more detailed information.

Hallo Andreas,

danke sehr, bin hoch zufrieden. Der Bassbereich hat sich schön angehoben dadurch und die NS 615 klingt nicht mehr so flach. Ebenso ist der Hochtöner im gesamt Sound angenehmer eingebunden und der Sound wärmer dadurch. Kurz um es hat sich das verbessert was ich mir erhofft habe, alles zu positiven und 100 pro paßgenau. Die Weiche ist im übrigen auch die selbe, hab ich so mal

Figure H5: The customer is "...very happy. The low end improved and the localisation is no longer flat". Interestingly he used the kit not for the NS-10 but for the NS-615 which is a larger model with two woofers.

Hallo Andreas,

Die Time Alignment Ringe kamen bei meinem Freund Paul gestern an und er hat sie bereits zusammen mit seinem „Sound Engineering" Freund ausprobiert. Surprise, surprise, die beiden kamen zu folgendem Schluß: Originalton Skype-mail: "On of my close engineer friends came over today and we listened to many kinds of music and we both agree it sounds much better with the rings". LOL!

Figure H6: The customer sent a pair to his friend in America. The friend's statement is in English in the text above