ACTA UNIVERSITATIS UPSALIENSIS
Uppsala Dissertations from the Faculty of Science and Technology
134

Simon Widmark

Causal MMSE Filters for Personal Audio

A Polynomial Matrix Approach

UPPSALA UNIVERSITET

To Jenny who made it possible,
to Elias and Ida who made it so much more difficult

Sammanfattning

Zoner för personligt ljud (eng. Personal Audio eller Personal Sound Zones) innebär att ett ljudsystem levererar ljud till vissa platser i ett rum och samtidigt undertrycker detsamma på andra platser. Ett sådant ljudsystem har flera möjliga och attraktiva tillämpningar. Ett exempel är individuella lyssningszoner, där flera olika personer i samma rum kan lyssna på olika musik utan att störa varandra. Andra är lokal ljudförstärkning där ljudet från t.ex. en TV förstärks i en position eller riktning i vilken en person med nedsatt hörsel befinner sig, samt dämpning av ljud i vissa bilsäten relativt de andra.

Den grundläggande fysiken bakom detta har varit känd, både i teori och praktik, sedan den första halvan av artonhundratalet. Enkelt sagt använder man en kombination av destruktiv interferens, alltså att ljudvågor möts i motfas, och släcker ut varandra, och konstruktiv interferens, alltså att vågorna istället möts i fas och förstärker varandra.

Trots god kännedom om den underliggande fysiken och trots att närliggande tillämpningar har utforskats under en längre tid dök den första vetenskapliga genomgången av zoner för personligt ljud upp så sent som 1994. Sedan dess har forskningsfältet varit aktivt och forskning och marknad står nu i begrepp att mötas i form faktiska produkter.

Det finns dock fortfarande några aspekter av detta fält som ännu inte undersökts uttömmande. Vi har för tillfället en ganska stor mängd algoritmer som använder skiftande angreppssätt för att generera zoner för personligt ljud. Trots att alla dessa algoritmer är avsedda att implementeras i kausala system,[1] är försvinnande få av dem framtagna med detta mål klart formulerat från början. Det innebär att en 'fix' måste göras efter att filtret konstruerats. Tyvärr blir ofta ljudkvaliteten lidande av detta tillvägagångssätt.

Denna avhandling fokuserar på framtagandet av digitala filter för formandet av zoner för personligt ljud, under antagandet att dessa filter senare ska tillämpas i kausala system. En mängd metoder för att generera filter som fokuserar på olika aspekter av grundproblemet härleds och diskuteras, både utifrån sina egna egenskaper, utifrån dessa egenskaper i relation till de andra här föreslagna metoderna och i relation till några metoder ur den befintliga litteraturen.

Denna avhandling har zoner för personligt ljud i bilar som motiverande tillämpning. Bilkupéer (och generellt fordonskupéer) är platser där flera personer ofta vistas samtidigt och där de befinner sig inom relativt små och väl definerade områden. Det är också i många fall platser där dessa personer spenderar

[1] Kausalitet är här en signalbehandlings-matematisk term som i princip innebär att vi inte kan veta vad som händer i framtiden.

mycket tid. Alla dessa egenskaper sammatagna gör bil- (fordons-) kupéer till attraktiva och väl lämpade miljöer att implementera zoner för personligt ljud i.

Fordonskupéer är dock inte den enda tänkbart intressanta miljön för användandet av zoner för personligt ljud. För att hålla de häri undersökta metoderna så pass generella att de kan behålla sin relevans även utanför denna specifika tillämpning har omsorg lagts på att hålla filterstrukturer och andra antaganden så pass flexibla att inga sådana tillämpningsbegränsningar byggts in i metoderna av misstag. Flera av de undersökta metoderna utvärderas även i en vardagsrumslik miljö i form av ett ljudlaboratorium på Uppsala Universitet.

Sju specifika metoder för design av kausala filter som genererar zoner för personligt ljud föreslås häri.

Den första är en metod där vikten av att generera tilltalande ljud i zonen där ljud ska höras (hädanefter den ljusa zonen) relativt dämpandet av ljud i den tysta zonen (mörka zonen) samt filtrets förstärkning i kvadratmening manuellt justeras. Denna metod är rättfram att använda och tillåter väldigt fin kontroll över resultatet men det manuella intrimmandet av filterparametrarna kan vara tidsödande, framförallt i situationer där man från början vet att ett visst resultat eftersträvas i någon kvalitetsmening.

Tre bivillkorade varianter av den viktade metoden föreslås därför också. Den ena härleds med ett bivillkor på att filtrets förstärkning i kvadratmening inte får överstiga ett visst värde. Den andra istället med ett bivillkor på att det instrålande ljudets effekt i den mörka zonen inte ska överstiga en viss nivå. I den tredje formuleringen begränsas avvikelsen mellan ett önskat ljudfält[2] och det faktiska ljudfältet. I samtliga tre bivillkorade metoder ges möjligheten att manuellt vikta de två återstående storheterna mot varandra. Detta tillvägagångssätt ger stor flexibilitet i att optimala filter genereras som garanterar att den bivillkorade storheten inte överskrids men i övrigt tillåter att de två återstående storheterna fritt prioriteras inbördes.

Den femte föreslagna metoden genererar ett filter under bivillkoret att ljudeffektskillnaden mellan den ljusa och den mörka zonen (akustisk kontrast) måste överstiga ett visst angivet värde. Detta är användbart när en viss separation mellan den mörka och den ljusa zonen krävs.

Utav dessa metoder utökas även den sista, med bivillkorad ljudeffektskillnad (eller kontrast), till att vara robust mot slumpartade avvikelser från den matematiska modell på vilken filtret baseras. Detta innebär att det producerade filtret ska åstadkomma ett godtagbart resultat även i de fall då den akustiska miljön skiljer sig från den modell på vilken filtret baserats. Detta är viktigt till exempel i massproduktion då det inte går att garantera att alla exemplar är identiskt lika och det inte finns tid att mäta upp den akustiska miljön i varje enskilt exemplar. Dessa avvikelser är formulerade på ett generellt sätt och tillåter hänsyn till en stor mängd olika felkällor av olika art.

[2]ung. ljudets sammantagna beteende i ett begränsat område

Denna robusta metod utökas dessutom med ett bivillkor på tidsenvelopet av för-ringningarna.[3]

Avhandlingen består av nio kapitel som i korthet är disponerade på följande sätt:

Kapitel 1
I det första kapitlet introduceras och motiveras det undersökta problemet: design av digitala filter för genererandet av zoner för personligt ljud under bivillkoret att dessa ska implementeras i kausala system. En kort introduktion ges även till de matematiska metoder och modeller som senare kommer att användas, och en översikt över den tillgängliga forskningslitteraturen på området presenteras.

Kapitel 2
Här presenteras den viktade metoden och undersöks utifrån flera olika aspekter i en bil.

Kapitel 3
Detta kapitel tar ett kort avsteg ifrån det långsiktiga målet, att generera filter för personliga ljudzoner, för att istället fokusera på den matematik som krävs för att härleda den undersökta typen av filter under kvadratiska bivillkor. Denna typ av bivillkor utgör stommen i de bivillkorade metoderna som vi senare kommer undersöka men tidigare beskrivningar av hur detta kan göras saknas i litteraturen. I detta kapitel presenteras och undersöks även det filtereffektbegränsade filtret. Detta filter kan direkt användas för att generera filter för personliga ljudzoner men undersökningarnas fokus ligger här på bivillkoren och deras egenskaper.

Kapitel 4
I detta kapitel presenteras metoden för att generera filter under bivillkor på att de ska producera en viss kontrast. Metoden kan sägas ta avstamp i matematiken som presenterades i Kapitel 3 men denna måste även modifieras för att kunna formulera just detta krav.

Kapitel 5
Här härleder vi återigen det filter som föreslås i kapitel 4 men i den matematiska formalism som är den mest populära konkurrent till den som används i denna avhandling, för att generera filter under bivillkor på kausalitet. Målet med detta är att jämföra de två metoderna och vi gör gällande att den konkurrerande metoden stöter på stora beräkningstekniska problem då antalet högtalare

[3]För-ringningar (eng. pre-ringing) är ett fenomen som ofta uppstår vid ljudfältskorrigering och som generellt har en negativ effekt på den upplevda ljudkvaliteten, bl.a. i form av en utsmetning, eller förlust av 'distinkthet' hos det producerade ljudet

blir stort eller när långa filter krävs. Vi visar även, för ett exempelsystem, att den filterlängd som krävs för att inte förlora någon prestanda överstiger vad denna metod kan prestera, utan specialiserade implementationer, på en relativt kraftfull dator.

Kapitel 6
I detta kapitel presenteras de två återstående bivillkorade filterformuleringarna (som baseras på matematiken som presenterades i kapitel 3). Dessa metoder jämförs med varandra, den viktade metoden från kapitel 2 samt den kontrast-bivillkorade metoden från kapitel 4. Detta för att belysa relativa fördelar och nackdelar hos de olika metoderna och för att ge stöd då någon metod ska väljas och implementeras.

Kapitel 7
I den första delen av detta kapitel utökas den kontrastbivillkorade metoden från kapitel 4 till att vara robust mot en stor mängd möjliga modellfel. I den andra delen utökas denna formulering till att även begränsa tidsenvelopet av för-ringningarna.

Kapitel 8
Avhandlingens näst sista kapitel ägnas åt en fallstudie där den robusta metoden används för att generera ett filter i en bil. Detta filter uppvisar god robusthet mot de störningar det är designat för, de presenterade experimenten visar även på vikten av denna typ av robusthet i normala tillämpningar.

Kapitel 9
Detta avslutande kapitel sammanfattar avhandlingens innehåll och pekar ut riktningen för några intressanta, potentiella, framtida forskningsprojekt med bäring på de undersökningar som gjorts häri.

List of papers

This thesis is in part based on the following papers and patents, which are referred to in the text by their Roman numerals and capital letters respectively.

I **Simon Berthilsson**, Annea Barkefors, Lars-Johan Brännmark, Mikael Sternad
Acoustical Zone Reproduction for Car Interiors Using a MIMO MSE Framework, *AES 48:th Conference Munich, Germany* September (2012)

II **Simon Widmark**
Causal IIR Audio Precompensator Filters Subject to Quadratic Constraints, *IEEE Transactions on Audio Speech and Language Processing, Vol. 26, Issue 10, October 2018 pp. 1897–1912.*

III **Simon Widmark**
Causal MSE-Optimal Filters for Personal Audio Subject to Constrained Contrast, *IEEE Transactions on Audio, Speech and Language Processing*, submitted June (2018)

A **Simon Widmark**
Audio Precompensation Filter Optimized With Respect to Bright and Dark Zones, *Patent Application, PCT/SE2016/051228*, December (2016)

Reprints were made with permission from the publishers.

List of contributions not included in this thesis

IV Annea Barkefors, **Simon Berthilsson**, Mikael Sternad
Extending the Area Silenced by Active Noise Control Using Multiple Loudspeakers, *IEEE/ACM ICASSP, Kyoto, Japan* March (2012)

V **Simon Berthilsson**, Annea Barkefors, Mikael Sternad
MIMO Design of Active Noise Controllers for Car Interiors: Extending the Silenced Region at Higher Frequencies, *American Control Conference, Montreal, Canada* June (2012)

VI Annea Barkefors, **Simon Berthilsson**, Mikael Sternad
An Investigation of a Theoretical Tool for Predicting Performance of an Active Noise Control System, *International Conference on Sound and Vibration Control, Vilnius, Lithuania*, July (2012)

VII **Simon Berthilsson**, Mikael Sternad
Acoustic Zones for Aircraft Seat Rows, *International Conference on Sound and Vibration Control, Bangkok, Thailand* July (2013)

Contents

Acknowledgements

All that is gold does not glitter,
Not all those who wander are lost;
The old that is strong does not wither,
Deep roots are not reached by the frost.
J.R.R. Tolkien,
The Fellowship of the Ring

This thesis is the result of a long and arduous journey. Although this thesis reflects the work that I myself have done over the past few years, no man is an island and I owe a lot of people a lot of gratitude.

First and foremost, to my principal supervisor, Professor Mikael Sternad. Thank you for taking me on in the first place and for your generosity both with time and with credit. Thank you also for the entirely immense amount of effort you have spent on reading, understanding, correcting, disputing and ultimately improving every little thing I have produced during the last eight years. Your endless enthusiasm and stubborn belief in my abilities have been inspiring and an invaluable help in ultimately pushing through.

I also wish to express my sincere gratitude to my assistant supervisors, Professor Anders Ahlén and Dr. Lars-Johan Brännmark. Your critical eye, curiosity, insight and alternative point of views have all been immensely helpful. A specific thank you to Anders, for keeping an open door policy and for providing direction when the need has arisen. Thank you also to Lars-Johan, for patiently answering an endless array of stupid questions with an equally endless array of insightful and well thought through answers.

To the staff of the Signals and Systems group, the UNO group and the CoRe group at Uppsala University: The inspiring and friendly atmosphere you all participate in providing has been a great motivator for getting out of bed in the mornings when the research has been going less than great. I would like to extend a special thank you to my fellow PhD students, Rikke Apelfröjd, Adrian Bahne, Annea Barkefors, Sinchan Biswas, Joachim Björsell, Markus Eriksson, Mathias Grudén and Magnus Jobs for outstanding efforts towards salvaging my remaining sanity also after office hours. A special thank you also to Ylva Johansson, for making things other than research run smoothly around the office.

To the great people at Dirac Research: Thank you for believing in me and for providing great discussions and a practical end-user perspective of the applications for all types of audio in general and my research in particular. A special thank you to the company as a whole for offering me the opportunity to be an industrial PhD student. It has been quite instructive to work directly, as well as indirectly, with the industrial side of my research and it has definitely helped broaden my mind.

I want to thank all my friends, for providing great support and space to ventilate, and for tolerating neglect during the final months of thesis writing. I want to thank my extended family, particularly my in-laws and their families for expressing genuine interest in me and my work. I am deeply grateful for the support you provide our family with, emotional as well as practical.

I want to thank my father, Petra, for encouraging me to not accept things at face value and to keep an inquisitive mind and I want to thank my mother, Ruth, for giving me the ability to do so with a positive mindset. I want to thank my siblings, Jakob and Paulina, for training my patience already at a young age. It has been most helpful. I want to thank all of you for always, always being there for me and my family.

Finally, although no words can ever suffice, I want to thank my children, Elias and Ida, for making my heart sing with joy on a daily basis. I want to thank my wife, Jenny, for making this possible at all and for making these past years, in spite of all the hard work, some of the happiest of my life.

Simon Widmark

Acronyms

ACC Acoustic Contrast Control.
ARMA Auto-Regressive Moving-Average.

BACC-RV Broadband Acoustic Contrast Control with Response Variation minimization.

DS Delay-and-Sum beamforming.

EDM Energy Difference Maximization.

FIR Finite Impulse Response.

GSVD Generalized Singular Value Decomposition.

IIR Infinite Impulse Response.

JPVM Joint Pressure and Velocity Matching.

LQG Linear Quadratic Gaussian.
LS Least Square.

MFD Matrix Fraction Description.
MIMO Multiple-Input Multiple-Output.
MISO Multiple-Input Single-Output.
MSE Mean Square Error.

PC Planarity Control.
PM Pressure Matching.
PVC Particle Velocity Constraint.

SFS Sound Field Synthesis.
SIMO Single-Input Multiple-Output.
SNR Signal to Noise Ratio.

Nomenclature

$A,\ a$	Scalar value.
$\boldsymbol{a}$	Constant vector.
$\mathbf{A}$	Constant matrix.
A	Scalar polynomial function.
$\boldsymbol{A}$	Polynomial matrix.
$\boldsymbol{\mathcal{A}}$	Rational matrix.
$\mathscr{A}$	Scalar rational function.
adj $(\dots)$	The adjungate of a matrix.
$\{\dots\}_+$	The causal part of a rational/polynomial matrix/function.
$\{\dots\}^*$	Complex conjugate.
det $(\dots)$	The determinant of a matrix.
$\mathscr{F}\{\dots\}\|_\omega$	Discrete Fourier transform evaluated at angular frequency ω.
$r\|c$	Matrix dimension, r rows and c columns.
dim $(\dots)$	The dimension of a matrix.
$Y(x)\|_z$	The function $Y(x)$ evaluated at $x = z$.
$E\{\dots\}$	Expected value with respect to a stochastic signal.
$\bar{E}\{\dots\}$	Expected value with respect to a probabilistic error model.
$\{\dots\}^H$	Hermitian transpose.
$\{\dots\}^{-1}$	Inverse.
j	Imaginary number, $j = \sqrt{-1}$.
λ	Lagrange multiplier, $\lambda \in \mathbb{R} \geq 0$.
ω	Normalized angular frequency.
q^a	Discrete time advance operator, $q^a y(t) = y(t+a)$.
$\{\dots\}_*$	Conjugate, $\boldsymbol{A}_*(q) = (\boldsymbol{A}(q^{-1}))_* = \boldsymbol{A}^H(q)$.
$\{\dots\}^T$	Transpose.
t	Discrete time index.
tr $[\dots]$	The trace (sum of diagonal elements) of a matrix.

1. Introduction

The field of personal audio, personal sound, or personal sound zones concerns the generation of sound only at select locations within a single volume of space. We shall in the present thesis investigate the construction of causal filters (pre-compensators) that act on the inputs to a set of loudspeakers with the purpose of generating personal sound zones. An archetypal example where this problem arises is when one passenger in a car wishes to enjoy a piece of music while the other passengers do not. Comfort and, particularly in the case of the driver, safety concerns sometimes render the headphone solution insufficient, motivating the generation of personal sound zones. The volume of space in which we want to reproduce sound is referred to as the *bright* zone in the literature. The volume of space in which we want the sound source to be muted is conversely referred to as the *dark* zone. Carrying the analogy further, we refer to the ratio of acoustic power in the bright zone to the acoustic power in the dark zone as *acoustic contrast*.

Reproduction of sound in one region of space with simultaneous preservation of silence in a different region is possible via the combination of destructive and constructive interference of sound waves. This phenomenon has been known for more than 180 years [1, p. 1], but while the closely related problem of cross-talk cancellation has been investigated since the 1960's [2], personal audio has been an active research field only since the early 1990's [3,4].

The reason why the field of personal audio is so young, relatively speaking, is not clearly established. A contributing factor is, however, almost certainly the complexity involved in generating filters for systems of several loudspeakers and several design positions with several, more often than not conflicting, design objectives. The recent surge in interest in the problem is on the other hand easy to understand by the diversity of the proposed applications of such filters. Proposed, and evaluated practical applications include generation of personal zones in a car [5], array beamforming for external warning signals of electric vehicles [6], and directionally selective TV sound amplification for the hard of hearing [7]. Other works investigate filter generation for personal computers in public spaces [8] and directive sound for mobile phones [9]. With the varying use-cases comes also great diversity in loudspeaker geometry, encompassing, e.g., dense line arrays, circular arrays, and irregular loudspeaker placements as is typically found in cars or living rooms. A few different, common loudspeaker geometries are illustrated in Figure 1.1.

The personal audio problem is by its nature a Multiple-Input Multiple-Output (MIMO) problem with several outputs (microphones, or control points

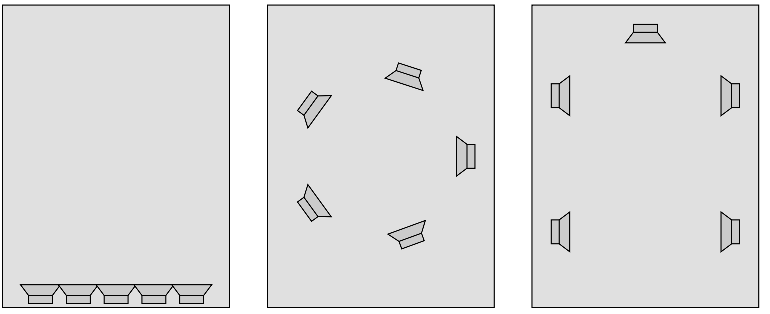

Figure 1.1. A few example loudspeaker geometries.

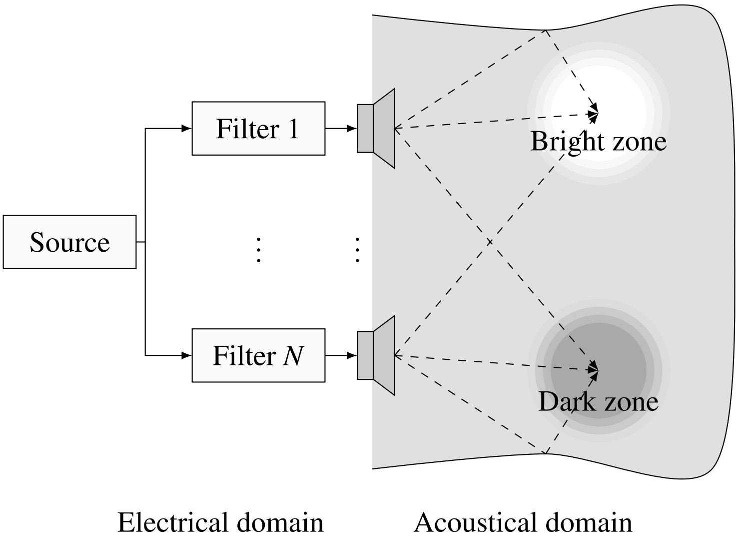

Figure 1.2. A generic electro-acoustical system generating personal sound zones in a reverberant environment.

distributed in space) and several inputs (loudspeakers).[1] A conceptual drawing of the electro-acoustical personal audio generating system is found in Figure 1.2.

Utilizing destructive interference is associated with the risk of phase errors resulting in interference amplification rather than reduction. These errors generally grow with increasing frequency. For this reason, it is widely accepted that the personal audio problem at higher frequencies must be treated with methods that do not rely on room dependent phase cancellation to generate

[1] While it is theoretically possible to design a Single-Input Multiple-Output (SIMO) system for personal audio, with a single loudspeaker reflecting off a surface, such implementations become extremely limited. Design of Multiple-Input Single-Output (MISO) systems are normally not considered within the personal audio literature, as they preclude the simultaneous definition of both a bright and a dark zone.

contrast. The problem thus consists of two sub-problems: the low frequency problem, where active methods, compensating the room dynamics can be utilized, and the high frequency problem, where passive methods that do not attempt to control any room reflections are employed. The focus of the present thesis lies on the first, active part of the problem. This does not mean that higher frequencies can be ignored all together: If pleasing sound is to be generated in the bright zone, co-optimization of the low and high frequency part is beneficial. It does however mean that we herein will assume that at least one input channel passively produces good separation in the high frequency region.

While the contrast generated in a personal audio system is very important, it is rarely the only objective of importance. For instance, the listener in the bright zone will normally want as 'good' sound as possible, and there will be constraints on how much power the loudspeakers can output. We will, in the following, refer to the quality of the sound in the bright zone by the term Sound Field Synthesis (SFS) error. This error quantifies the deviation between the desired bright zone properties and the actual attained properties of the bright zone in the compensated system. The amount of power the loudspeakers output is moderated by the filter amplification, or array gain.

These central metrics are normally in opposition of each other. While there may be situations where, e.g., phase manipulation allows us to improve the sound in the bright zone without altering the attained contrast or the filter gain, such cases are rare. The design trade-offs inherent to the personal audio problem are summarized in Figure 1.3.

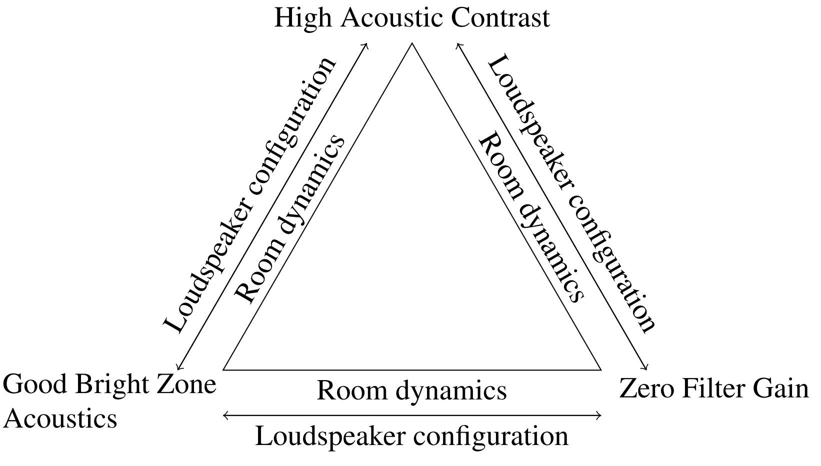

Figure 1.3. Conceptual overview over the main design trade-offs encountered in the general personal audio problem.

Any given filter occupies a single point in the triangle, Figure 1.3. This means that a filter that attains a high level of acoustic contrast, relative to what can be generated in the specific system, generally produces poor bright zone acoustics and does so with significant filter gain. A filter *design method* may however be able to generate a host of different filters, spanning a large area of any such triangle.

As indicated in Figure 1.3, the degree of opposition between the three different objectives can be modified by the room dynamics (including the locations and the spatial extent of the bright and dark zones), and also by the loudspeaker placement and layout. A suitable physical layout will better align the three goals, and we could represent this by shrinking one or several sides of the triangle. For example, a good natural acoustic separation between the bright and dark zone reduces the opposition between attaining a high contrast and 'good' sound in the bright zone.

The not so easily defined metric 'good bright zone acoustics' is to be taken to mean 'as good sound reproduction as possible in the given system by any conceivable audio quality measure'. Similarly, 'high contrast' is to be interpreted as 'the greatest possible contrast attainable in the given physical system'.

We will in the present thesis focus on the generation of personal audio filters derived subject to a constraint on causality. This type of filter design scheme extends further toward the lower left corner of the conceptual filter design triangle than the methods that are not subjected to this constraint.

We will furthermore take a car-centric approach in design, analysis, and in what simplifications and assumptions we allow ourselves to make, but we will at the same time keep the methods generic enough that they may be applied also in other environments. An important aspect of the generation of filters for such complex problems as the personal audio problem, is the effort and amount of background knowledge required for a filter designer to attain desired results. This perspective also plays an important role in the choices and motivations that underpin the filter design methods proposed and investigated herein.

A reader who would like to get a preview of the type of results that can be accomplished with the methods to be developed in this thesis can take a quick look at Figure 8.14, from the case study of Chapter 8. The bright zone here covers the two front seats and the dark zone the two main rear seats in a car with the audio system outlined in Figure 8.1.

1.1 Design Choices and Motivations

The motivating goal of this thesis is to develop a comprehensive framework for the theoretical and practical design of pre-compensation filters for personal audio. The main focus of the discussion concerns filters for automotive cabins which serves as the main motivating application. The resulting filter design schemes will however be quite general and we will take care so as to not inadvertently preclude other possible applications.

Practical filter design cannot exist in a void but is always dependent on the practical realities of the implementation. In this section, we shall review some of the physical aspects of the systems in focus in the present thesis. With this in mind we will then discuss and motivate the over-arching design choices and assumptions that underpin the discussions to follow.

1.1.1 The Acoustic Environment

The physical reality of the car cabin differs markedly from that of a 'regular' room. These differences have specific impacts on the acoustic environment. The car cabin is far more crowded and physically smaller than a normal room is, which means that the average distance between a loudspeaker and a control point is smaller than is the case in a room. Further, as also the distance between a control point and the nearest reflective surface is smaller, the reverberant field will be relatively stronger and harder to separate from the direct field than is the case in the larger room. Since the physical volume of the car is smaller than the room and the walls and interior are normally padded, the reverberation time of the car interior can be expected to be shorter than that of the room, particularly in higher frequencies. All these phenomena are illustrated by the impulse response between one loudspeaker and one microphone estimated in a car cabin and in the living-room-like acoustical environment of the Audio lab (Appendix A). See Figure 1.4.

Assumptions that are based on the dominance of the direct sound field, e.g., for motivating the use of anechoic models that are common in the personal audio literature, are therefore not generally valid in the automotive environment.

We also note that the span of reverberation times, and therefore also model orders, of the relevant acoustical environments can be quite large.

1.1.2 Causality and Pre-Ringing

A process is said to be *causal* if it can be described using past and/or present signals (events) only. Formalizing this mathematically, the output of a causal system at time t may only depend on input and/or output signals from times up to, and including t.

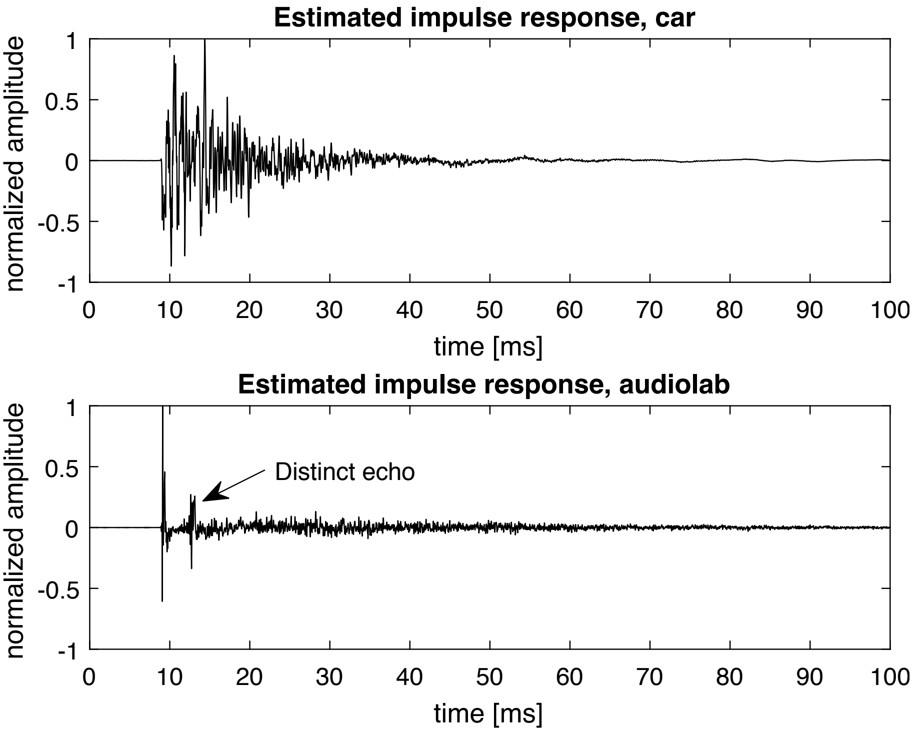

Figure 1.4. Estimated impulse responses in a car and in the audio lab described in Appendix A.

In practical filter design, this matters since we can normally not know future input signals to a system, and utilizing non-causal filters would force us to work with, often crude, predictions of the relevant signals.

One common way of designing pre-compensation filters for personal audio is to simply ignore the required causality of the generated filters, compute filters per-frequency and, if necessary, apply a delay to them, so that they become completely causal. This approach, however, comes at a cost. For instance, the required delay of the causal filter may be large and this may influence the user experience when using the filters.

Another approach is to simply truncate the non-causal part of the filters but this generally has a severely detrimental impact on the properties of the filters.

Some middle ground can often be found in practice, where a combination of truncation and delays may produce filters with improved behaviour as compared to either of the two above approaches. The delay may however still not be freely specifiable and the resulting filters are no longer optimal with respect to the criterion by which they were designed.

An issue with different causes and effects, that is interconnected with that of causality is *pre-ringing*. Pre-ringing is here defined as non-zero outputs of the compensated system that occur before the start of the desired system

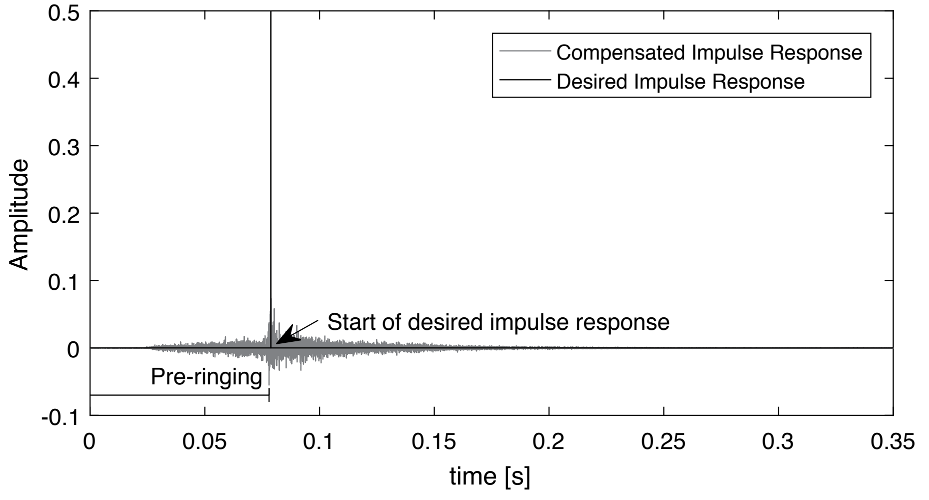

Figure 1.5. Illustration of pre-ringing in a compensated impulse response.

behaviour, see Figure 1.5. The cause of pre-ringing is mismatched zero-compensation of the filters [10]. This mismatch may originate from poor knowledge of the system-to-be-compensated. Pre-ringing may also be generated in perfectly modelled systems when the zeros of the transfer functions from one loudspeaker to two or more positions in space do not match perfectly. Pre-ringing errors have a detrimental effect on perceived sound quality. The human auditory system may however mask these effects (pre-masking) to some extent, at least if they occur shortly before a peak of larger magnitude in the impulse responses of the compensated system [11].

By designing causal filters,[2] we guarantee that the filters generate no output before a time of our choosing. We can thus limit the temporal extent of any generated pre-ringing errors in the compensated system. This approach can be utilized to completely remove pre-ringing of the compensated system (generating minimum phase filters), or, by allowing a *modelling delay*, precisely specify the duration of any possible pre-ringing.

Omission of taking causality into consideration in the filter optimization stage often leads to significant problems with pre-ringing or echoes in the compensated system. Some methods do exist that generate non-causal filters (in some sense), see e.g., [12] while retaining control over the generated pre-ringing in the compensated system. Such methods are, however, rare and have not yet seen application to the personal audio problem.

[2]We will, throughout the thesis, refer to filters that are derived subject to constraints on causality, as 'casual filters'. This is perhaps somewhat inexact as also other filters may be converted to be causal via, e.g., a delay but this convention will greatly reduce the complexity of some sentences used herein. All implemented and investigated filters will be causal, in the strict sense of the word.

1.1.3 Design Choices

Given the above, we will now formalize a list of requirements on the filters and the design methods:

- Causality: In view of the discussion above, there are clear advantages to taking the causality of the resulting filters into account already in the optimization stage.
- Designer perspective: If a design method is to stand any chance of making it into production in the automotive industry (or, indeed any field), it will not be sufficient only to produce a pleasing result. The design process itself must be fairly straight-forward and to the point, so that these pleasing results can be attained by the on-site staff with minimal training.
- Robustness: The performance of any model based design is always dependent on the validity of the utilized model. The filters under investigation herein aim to utilize both destructive and constructive interference to achieve the specified goals. Such operations are particularly sensitive. Hence, a certain robustness to model errors must be considered. In addition, the implementation in a series production setting raises the need for a design that is valid for a large amount of similar, but different actual systems. This need originates from the inevitable variability due to manufacturing tolerances of a large set of components and assembly variations.
- Efficiency of implementation: Ideally, the design method should not depend on assumptions of, e.g., the particular reverberation times of any particular system. For example, if Finite Impulse Response (FIR) filters are assumed, then long decay times are particularly problematic as they lead to very long filters. In addition, the design of FIR filters for systems with long decay times often lead to computational difficulties in that constraints on computation times or memory usage are not met. It is, on the other hand, quite straight forward to obtain a set of FIR filters from a given set of Infinite Impulse Response (IIR) filters, if this is desired from an implementation perspective.

These requirements will herein be addressed using multipoint Mean Square Error (MSE) minimisation in a rational matrix framework. This framework has previously been successfully utilized to produce causal filters for robust sound field synthesis, see e.g., [13] and loudspeaker pre-compensation, e.g., [14]. The framework is based on the manipulation of impulse responses, rather than poles and zeros, which allows for improved intuitive understanding of the resulting filters and the process. The framework has also been applied to treat similar robustness issues as those discussed herein, see e.g., [15]. Finally, the framework has proven to provide attractive computational properties and produces IIR filters, which may alleviate some implementation issues related to long filter lengths. The computational times and memory requirements are also often lower than is the case with other alternative choices.

Two potential alternative design strategies for generating causal pre-compensators were identified and rejected:

- The Toeplitz convolution FIR matrix design methodology was not chosen, since systems that require many filter taps, as would be needed to capture the full dynamics of the systems at hand, generate problems with matrices with large dimensions. These systems of linear equations present a considerable computational burden with long associated computational times, see e.g. Chapter 5. Further, FIR filters that capture both the fine details of the higher frequencies and the longer ringing of the lower frequencies normally need to be of high degrees. This poses a problem when they are to be implemented in a system with limited memory and computational resources such as an automotive audio head unit. The rational matrix framework generates IIR filters, which require fewer filter taps to model long impulse responses, and avoids the large matrices needed in the FIR Toeplitz matrix convolution filter design methodology.
- The reason for choosing the rational matrix framework over the state-space approach lies in the design step. Using the rational matrix framework, we model the system and filters using plain impulse responses. This allows for good insight into the inner workings of the algorithm in the design stage and avoids translations and factorizations to move between the state-space form and the impulse response domain.

1.2 Contributions and Thesis Overview

A brief chapter-by-chapter overview over the thesis is given with the main contributions highlighted. The main non-chapter-specific contribution is the detailed overview of the problem of generating causal filters for personal audio with design insights and other considerations.

Chapter 2: Weighted causal MMSE design for Personal Audio filters

In this chapter, we propose and investigate a first example of a causal personal audio filter design strategy based on minimization of a weighted objective function. The specific objective function investigated is the weighted sum of an SFS error term, a filter power term, and a dark zone power term. Contrast is here generated by the joint minimization of the acoustic power in the dark zone and the generation of a user-specified sound field in the bright zone. This formulation can be seen as an extension of the rational matrix framework designed for SFS, as explored in [16, 17], to the personal audio problem. It can also be seen as the rational matrix formulation of the causal weighted Pressure Matching (PM) dark zone power minimization investigated in [18]. The proposed algorithm is evaluated by design of a set of filters that are in turn examined via simulations and measurements in a mass produced car.

This chapter is based, in part, on Paper I.

Chapter 3: Causal IIR Wiener Pre-Compensator Design Subject to Quadratic Constraints

There are several situations in general filter design where a solution has to adhere to certain hard constraints.[3] These can be, e.g., on the maximally allowed filter gain or on the maximal tolerated SFS error. The personal audio problem also provides a few situations in which the ability to derive filters that satisfy hard constraints may facilitate or speed up the design process. One such example is the possibility to use the masking effect of the background noise in an automotive environment to mask the leakage of sound from the bright zone into the dark zone. If we know the approximate ambient acoustic power in the dark zone we can then constrain the expected acoustic power generated by the personal audio system in the dark zone below this level. It is plausible that the experienced disturbance caused by leakage from the bright zone is then reduced significantly.

No method of deriving rational matrix based filters subject to this type of constraints is however previously known in the literature. Incorporation of

[3] A constraint is considered 'hard' if the constraint function *must* conform to the constraint condition. A corresponding 'soft' constraint is akin to a weighted function, were larger deviations are more discouraged, but may still be accepted.

quadratic constraints into the rational matrix optimization framework is therefore the focus of this chapter. The methodology is presented and motivated and a set of example filters are derived.

This chapter is based on Paper II.

Chapter 4: Constraints With Respect to Contrast

While the absolute sound power level in the dark zone is important, it can be argued that the acoustic contrast, or the acoustic power ratio between the bright and the dark zone is even more so. The case for explicitly using contrast as compared to the approach taken in Chapter 2, is that it alleviates the inadvertent trade-off in formulations where the dark zone power and SFS error terms are traded against each other. In such cases, the SFS error and the separation between the bright and the dark zone are not independent variables. Greater contrast can in these formulations not be attained directly but becomes a by-product of balancing the different objectives of the objective function.

A method that minimizes a quadratic criterion (or objective) function of an SFS error term and a filter power gain term subject to constraints on the generated contrast and on the causality of the generated filter is proposed. This formulation can be seen as the causal design corresponding to the non-causal design proposed by Cai et. al in [19]. The constrained design method is particularly beneficial when truly personal audio is required, i.e., the situation when different source material is to be enjoyed in different zones. In such a case, the required contrast must exceed a prescribed level, which is easily implementable by the contrast constraint. In other applications, such as generation of a relative reduction of the sound level in the dark zone, the contrast constraint can be viewed as a tuning parameter that directly controls the attained contrast. The theory of Chapter 3 is also elaborated to accommodate the additional algorithmic challenges that arise in the specific constraint formulation.

This chapter is based on Paper III.

Chapter 5: The FIR Toeplitz Solution

In this thesis, we investigate causal filter generation based on the rational matrix framework. The generally preferred approach for causal filter generation in the literature is however Toeplitz matrix convolution based methods. In this chapter, we derive a causal, contrast constrained FIR pre-compensator on Toeplitz convolution form. This pre-compensator corresponds to an FIR version of the pre-compensator proposed in Chapter 4. The Toeplitz convolution based pre-compensator is investigated in terms of memory requirements and computational time, and compared to the rational matrix IIR pre-compensator of Chapter 4.

Chapter 6: Constrained Methods for Personal Audio

Having established a method of deriving rational matrix based filters subject to quadratic constraints in Chapter 3, we now put it to use by deriving two constrained causal filter generation algorithms for personal audio. The first is the motivating design for the problem of Chapter 3, i.e., filters constrained with respect to the expected acoustic power generated in the dark zone. The second is constrained with respect to the maximal allowed SFS error in the bright zone. These formulations can be said to be the causal designs corresponding to the non causal designs proposed in [20] and in [21] (albeit with a variable constraint level). The proposed methods are compared to each other and to the method proposed in Chapter 2 from a mathematical vantage point. They are also compared to the contrast constrained filter derived in Chapter 4 via both simulations based on measured impulse responses and validation measurements.

Chapter 7: Robustness to Modelling Errors and Treatment of Pre-ringing

In this chapter, the contrast constrained method proposed in Chapter 4, and investigated in Chapter 4 and Chapter 6 is expanded upon further. First, we derive the optimal, causal, contrast constrained controller that is robust with respect to a probabilistic error model. This robust formulation is further extended with an implicit (but fully user defined) constraint on the allowed amount of pre-ringing in the impulse responses of the compensated system.

Chapter 8: A Case Study: Robust Zone Design in a Car

In this chapter, we undertake a case study in which the robust contrast constrained controller is put to work in a series production car. The different design steps and considerations are explicitly stated so as to guide a reader through the process of designing a satisfactory filter for personal audio.

We also make a first, tentative, investigation into the effects of a change in passenger constellation, with respect to contrast and bright zone behaviour.

Chapter 9: Concluding Remarks

In this, final, chapter of the thesis, we summarize the main findings and discuss briefly some potential directions of future investigations, with bearing on the work presented herein, that may be of interest to the research field at large.

1.2.1 Summary

The main technical contributions of the thesis can be summarized as follows:

- A general method for including multiple quadratic constraints in the rational matrix LQG pre-compensator framework
- An optimal linear causal controller constrained w.r.t. multiple constraints on the expected average generated SFS error.
- An optimal linear causal controller constrained w.r.t. multiple constraints on the expected average generated acoustic power in the dark zone
- An optimal linear causal controller constrained w.r.t. multiple constraints on the expected average generated contrast
- A *robust* optimal linear causal controller constrained w.r.t. the expected average generated contrast
- A robust optimal linear causal controller constrained w.r.t. the expected average generated contrast *and* the amplitude envelope of the generated system pre-ringing
- A Toeplitz based FIR design, optimizing sound field reconstruction with weighted filter power gains, subject to a contrast constraint
- A general investigation into the applicability and benefits of treating personal audio problems using filters that are optimized with causality in mind.

1.3 The Rational Matrix Framework for Causal Audio Filter Design

While rational (and the special case polynomial) equation methods have been applied to active control problems since the 1950's [22], their application to audio signal processing is less mainstream and a brief introduction therefore follows.

The rational matrix approach has been applied to audio pre-compensation problems for nearly two decades [14], with applications in loudspeaker compensation [14], room compensation [13], sound field synthesis [23], and active noise control [24].

In the rational matrix equations for audio pre-compensation, linear time-invariant discrete time filters and models are represented by rational functions in the time shift (delay) operator q^d, where $q^{-d}y(t) = y(t-d)$ and $q^d y(t) = y(t+d)$. These operator functions yield the IIR filtered response when applied to an input signal.

Example 1.3.1.
Consider a linear time-invariant discrete time system, with normalized sampling rate, input signal $u(t)$ and output signal $y(t)$, described by the difference equation

$$y(t) = 0.5y(t-1) + u(t) + 0.1u(t-1). \tag{1.1}$$

Using the backward shift operator q^{-1}, this difference equation can equivalently be represented

$$(1-0.5q^{-1})\,y(t) = (1+0.1q^{-1})\,u(t), \tag{1.2}$$

or by using the rational transfer operator

$$y(t) = \mathscr{F}(q^{-1})u(t), \tag{1.3}$$

where

$$\mathscr{F}(q^{-1}) = \frac{1+0.1q^{-1}}{1-0.5q^{-1}}. \tag{1.4}$$

The corresponding frequency domain transfer function is obtained by substituting the complex variable z^{-1} or $e^{-j\omega}$ for the operator q^{-1}, resulting in the rational transfer function

$$\mathscr{F}\left(e^{-j\omega}\right) = \frac{1+0.1e^{-j\omega}}{1-0.5e^{-j\omega}}. \tag{1.5}$$

For matrices of transfer operators, each matrix element contains a rational expression. Such matrices are denoted rational matrices.

Polynomial and Rational matrices that describe causal processes can be expressed in the backward shift operator, q^{-1}, exclusively. Such matrices are sometimes referred to as causal (rational) matrices. Conversely, matrices that describe exclusively non-causal processes, i.e. rational or polynomial matrices that can be expressed only using the forward shift operator, q, will be referred to as non-causal matrices.

By changing operator of the hermitian transpose of a rational matrix, e.g., from q^{-1} to q, we attain the *conjugated* rational matrix. Since we here investigate causal systems and filters, conjugated matrices will in general be non-causal.

We here adopt the notation of a subscript star, e.g., $\boldsymbol{\mathcal{M}}_*(q)$ to denote non-causal matrices that are obtained by conjugation of a causal matrix $\boldsymbol{\mathcal{M}}(q^{-1})$.

The rational matrix framework provides a convenient way of parametrizing electro-acoustical systems in terms of input-output relationships, time-domain transfer operators or frequency domain transfer functions.

Just as FIR filters are special cases of general IIR filters, polynomial matrices are rational matrices with with elements with constant denominators only.

A rational function, such as (1.4) or (1.5), can be specified in terms of its numerator and denominator polynomials. In analogy, rational matrices can be specified in terms of two polynomial matrices representing a 'numerator' and a 'denominator'. Such a description is called a Matrix Fraction Description (MFD) [25, Ch. 6]. Since matrices in general do not commute, two types of MFDs can be defined: the right MFD $\boldsymbol{\mathcal{F}}(q^{-1}) = \boldsymbol{B}_R(q^{-1})\boldsymbol{A}_R^{-1}(q^{-1})$ and the left MFD $\boldsymbol{\mathcal{F}}(q^{-1}) = \boldsymbol{A}_L^{-1}(q^{-1})\boldsymbol{B}_L(q^{-1})$. The right MFD is in general more convenient to use in control problems, and will be used in this thesis. An example is given below.

Example 1.3.2.
A simple example of a (right) MFD is the decomposition of the rational matrix

$$\begin{aligned}&\begin{bmatrix} \frac{b_{11_1}+b_{11_2}q^{-1}}{1+a_1q^{-1}} & \frac{b_{12_1}+b_{12_2}q^{-1}}{1+a_2q^{-1}} \\ \frac{b_{21_1}+b_{21_2}q^{-1}}{1+a_1q^{-1}} & \frac{b_{22_1}+b_{22_2}q^{-1}}{1+a_2q^{-1}} \end{bmatrix} \\ &= \begin{bmatrix} b_{11_1}+b_{11_2}q^{-1} & b_{12_1}+b_{12_2}q^{-1} \\ b_{21_1}+b_{21_2}q^{-1} & b_{22_1}+b_{22_2}q^{-1} \end{bmatrix} \begin{bmatrix} 1+a_1q^{-1} & 0 \\ 0 & 1+a_2q^{-1} \end{bmatrix}^{-1}.\end{aligned} \tag{1.6}$$

This simple example is based on the denominators of the elements in the first column being common and likewise of the second column. This is, however, not a necessary assumption as arbitrary element denominators can be constructed using multiplication in both numerator and denominator of non-common denominator factors.

A *stable* system on rational matrix form, $\boldsymbol{\mathcal{F}}(q^{-1})$, can be expressed as a right (or left) MFD $\boldsymbol{F}(q^{-1})\boldsymbol{A}_R^{-1}(q^{-1})$, with all roots of $\det(\boldsymbol{A}_R(z^{-1})) = 0$ contained in $|z| < 1$.

If $\det(\boldsymbol{A}_R(z^{-1})) = 0$ has all roots in $|z| < 1$, then all roots of $\det(\boldsymbol{A}_{R*}(z)) = 0$ are located in $|z| > 1$.

Rational matrices and scalar functions are in the present thesis represented by bold, capital, script font symbols $\boldsymbol{\mathcal{A}}$, and capital, script font symbols $\mathscr{A}$, respectively. Similarly, polynomial matrices and scalar functions are denoted by bold, capital, italic symbols $\boldsymbol{A}$, and capital italic symbols A, respectively. In contrast, constant matrices are denoted by bold capital symbols, $\mathbf{A}$, and constant vectors by bold lower case symbols $\boldsymbol{a}$. Scalar constants and variables are denoted by regular capital or lower case letters A, a. The discrete time index is denoted t. The time shift operator argument of rational or polynomial matrices, q^{-1}, will sometimes be omitted, where there is no risk of confusion.

The (generic) degree of a polynomial matrix $\boldsymbol{M}(q^{-1})$, $\deg(\boldsymbol{M})$ is defined as the highest possible degree of any of its polynomial elements, if $\boldsymbol{P}(q^{-1}) = \boldsymbol{M}(q^{-1})\boldsymbol{N}(q^{-1})$, then

$$\deg(\boldsymbol{P}) \leq \deg(\boldsymbol{M}) + \deg(\boldsymbol{N}). \tag{1.7}$$

The inverse of a polynomial matrix is a rational matrix that can be expressed

$$\boldsymbol{M}^{-1}(q^{-1}) = \frac{1}{\det(\boldsymbol{M}(q^{-1}))} \operatorname{adj}(\boldsymbol{M}(q^{-1})), \tag{1.8}$$

where the right factor is a polynomial matrix.

If $\boldsymbol{M}(q^{-1})$ is causal (if it contains only polynomials in the backward shift operator q^{-1}), then its inverse, $\boldsymbol{M}^{-1}(q^{-1})$ will also be causal. If the polynomial $\det(\boldsymbol{M}(q^{-1}))$ has all zeros within $|z| < 1$, then all denominator polynomials of the rational matrix $\boldsymbol{M}^{-1}(q^{-1})$ will only have roots within $|z| < 1$. The matrix $\boldsymbol{M}^{-1}(q^{-1})$ will thus be stable.

Polynomial or rational matrices in both the delay *and* the advance operator, q^{-1} *and* q are referred to as *double-sided*. They generally have two degrees, one in the q^{-1} dimension and one in the q dimension. If the two associated degrees are equal, then we may refer to this as the *double-sided degree*.

1.4 The Personal Audio Design Problem

Several design methods are examined throughout this thesis. Most are designed explicitly with personal audio in mind. It will for this reason be useful to establish a baseline system model and clarify and motivate the relevant assumptions so that this does not need to be covered in detail in every chapter. The system model will be the mathematical foundation for the formulations and design strategies proposed herein.

1.4.1 The Electro-Acoustical System

Let us first assume that we are working with an electro-acoustical system comprising N loudspeakers. We shall also assume that the system is linear and time-invariant so that it can be perfectly described using linear, time invariant, models of dynamic systems. Of course, no conceivable electro-acoustical system is linear and time-invariant in reality. No loudspeaker is for instance perfectly linear and something so trivial as temperature changes violates the time-invariance assumption. We will for the time being, however, trust that these deviations from the assumed linearity and time-invariance are small, with negligible effects and revisit them in Chapter 7 for a more systematic treatment.

The transfer functions between all electrical signal inputs and all electrical outputs as measured at M_B positions in the room are described by the causal and stable rational matrix model $\boldsymbol{\mathcal{H}}_B(q^{-1})$ of dimensions $M_B|N$. The M_B positions define the physical extent of the bright zone. These transfer functions are normally estimated via impulse response measurements using measurement microphones, but purely theoretical models may also be utilized.

The dark zone is likewise defined by M_D positions in the room and a model of the electro-acoustical system as sampled in these positions is contained in the $M_D|N$ causal and stable rational matrix $\boldsymbol{\mathcal{H}}_D(q^{-1})$. The $M_B + M_D$ positions defining the bright and dark zones are in the following referred to as *control points*.

In the present system parametrization, element (m,n) in, e.g., the bright zone system model matrix $\boldsymbol{\mathcal{H}}_B$ contains the rational transfer operator that describes the input-output relationship between loudspeaker n and design point m. A conceptual image of a generic electro-acoustical personal audio system, with $M_B = M_D = 16$ and $N = 5$, is illustrated in Figure 1.6.

Note that the transfer functions between each loudspeaker and each control point, in the figure illustrated by microphone symbols, only need to be estimated once in a time-invariant system. We may therefore remove any microphones once all transfer functions are estimated, without any consequence for the generated zones.

If we employ more than one control point in either zone, the corresponding rational system model describes not only the sound field in the control points

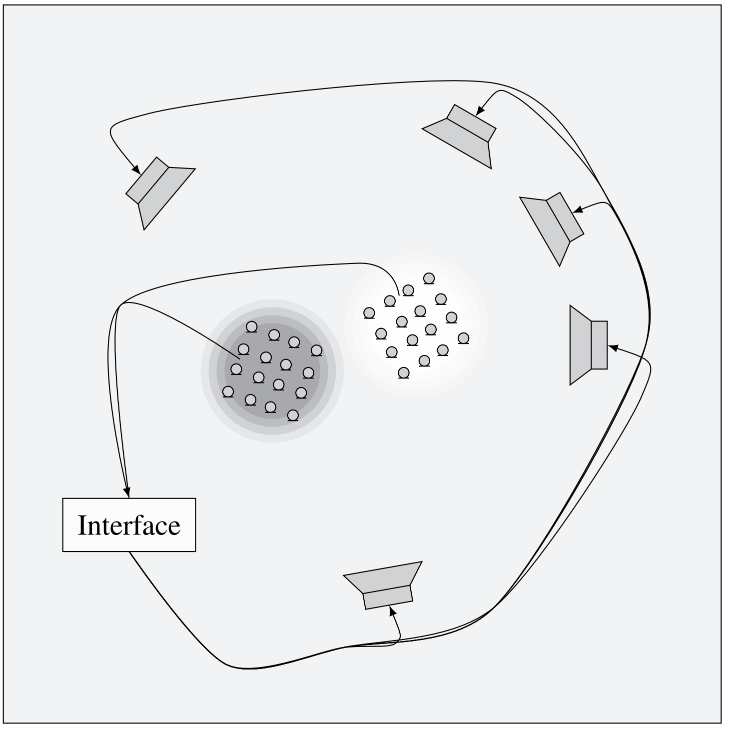

Figure 1.6. A generic electro-acoustical system generating personal sound zones in a reverberant environment. Here illustrated in the system model estimation phase, after which any measurement microphones may be removed.

but also the sound field in between them up to the spatial Nyquist frequency. The spatial Nyquist frequency is defined, for a uniform spatial sampling interval of d_m in the far-field, by

$$f_{nyq} = \frac{1}{2}\frac{c}{d_m}, \tag{1.9}$$

where c is the speed of sound. In practice, however, we may partly sample in the near-field of the loudspeakers and we will therefore use the more conservative estimate

$$f_{kir} = \frac{1}{3}\frac{c}{d_m}, \tag{1.10}$$

also known as Kirkeby's rule of thumb [26] to account for this.

The rational matrix system descriptions for the bright and dark zones can be expressed on polynomial matrix form, using right MFDs, as

$$\begin{aligned} \boldsymbol{\mathcal{H}}_B(q^{-1}) &= \boldsymbol{B}_B(q^{-1})\boldsymbol{A}^{-1}(q^{-1}) \\ \boldsymbol{\mathcal{H}}_D(q^{-1}) &= \boldsymbol{B}_D(q^{-1})\boldsymbol{A}^{-1}(q^{-1}), \end{aligned} \tag{1.11}$$

where both the $M_B|N$, $M_D|N$ numerator matrices $\boldsymbol{B}_B(q^{-1})$ and $\boldsymbol{B}_D(q^{-1})$ and the diagonal $N|N$ denominator matrix $\boldsymbol{A}(q^{-1})$ are polynomial, rather than rational, matrices.

It may seem overly optimistic to expect that both the bright zone system model and the dark zone system model can be expressed using the same denominator matrix $\boldsymbol{A}^{-1}(q^{-1})$. However, using a well motivated parametrization, this is perfectly reasonable. Resonant frequencies in a room are frequencies whose wavelength perfectly match the wall-to-wall distance or floor-to-ceiling distance in that room so that a whole number of half wavelengths fit in the room. These resonant frequencies (or room modes) do not change with position, as they are linked to the geometry of the room itself. A property of these resonant frequencies is, further, that they generally have long decay times and therefore would require a large number of FIR filter taps to model. Using the feedback property of the denominator of an IIR filter on the other hand, slowly decaying resonances can be succinctly modelled with relatively few filter taps. Admittedly, the effect of a room mode is not equally strong in every point in the room. This variability in coupling between the room modes and each point in the room can, however, be modelled by zeros in the IIR numerator that match the corresponding poles, thus reducing the gain of that frequency in that position. The common denominator matrix in the right MFD description for all control points in a room is thus quite reasonable both from a mathematical and physical point of view. We shall therefore adopt the common denominator model (1.11) in the remainder of the thesis. A far more elaborated argument than the above can be found in [16, Chapter 2].

We now have a physically motivated description of the input-output *relationship* defining our electro-acoustical system. Introducing a vector of N arbitrary input signals, $\boldsymbol{u}'(t)$, we may describe the outputs at the control points,

$\boldsymbol{\sigma}'(t)$, as a function of the input signal by

$$\begin{aligned}\boldsymbol{\sigma}_B'(t) &= \boldsymbol{\mathcal{H}}_B(q^{-1})\boldsymbol{u}'(t)\\ \boldsymbol{\sigma}_D'(t) &= \boldsymbol{\mathcal{H}}_D(q^{-1})\boldsymbol{u}'(t).\end{aligned} \tag{1.12}$$

Only modelling the sound will not, however, allow us to generate personal sound zones and so we also introduce a stable causal linear time-invariant pre-compensator, represented by the rational matrix $\boldsymbol{\mathcal{R}}(q^{-1})$ in the electro-acoustical path.

From vector $\boldsymbol{r}(t)$ of L sound source signals as inputs, this pre-compensator produces the control input vector $\boldsymbol{u}'(t)$ by

$$\boldsymbol{u}'(t) = \boldsymbol{\mathcal{R}}(q^{-1})\boldsymbol{r}(t). \tag{1.13}$$

The output from the compensated system can now be described by

$$\begin{aligned}\boldsymbol{\sigma}_B'(t) &= \boldsymbol{\mathcal{H}}_B(q^{-1})\,\boldsymbol{\mathcal{R}}(q^{-1})\boldsymbol{r}(t) = \boldsymbol{B}_B(q^{-1})\boldsymbol{A}^{-1}(q^{-1})\,\boldsymbol{\mathcal{R}}(q^{-1})\boldsymbol{r}(t)\\ \boldsymbol{\sigma}_D'(t) &= \boldsymbol{\mathcal{H}}_D(q^{-1})\,\boldsymbol{\mathcal{R}}(q^{-1})\boldsymbol{r}(t) = \boldsymbol{B}_D(q^{-1})\boldsymbol{A}^{-1}(q^{-1})\,\boldsymbol{\mathcal{R}}(q^{-1})\boldsymbol{r}(t).\end{aligned} \tag{1.14}$$

As will be discussed in the literature review, the perceived distraction in the dark zone is highly signal dependent and a solution would benefit from utilizing signal dependent filters. This approach is however infeasible in practice and we will instead take a probabilistic approach to the filter design. The arbitrary input signal vector $\boldsymbol{r}(t)$ will in the following therefore be modelled by a white noise signal of zero mean, with covariance matrix $E\left\{\boldsymbol{r}(t)\boldsymbol{r}^T(t)\right\} = \mathbf{P}$, under the assumption that we know nothing else about its nature.

If we were to know *something*, for instance the average spectral colouration of the input signal, this could also be taken into account. Statistical modelling of a coloured input signal can be done using an Auto-Regressive Moving-Average (ARMA) model, as $\boldsymbol{r}(t) = \boldsymbol{E}(q^{-1})\boldsymbol{F}^{-1}(q^{-1})\boldsymbol{e}(t)$ where $\boldsymbol{E}(q^{-1})$ and $\boldsymbol{F}(q^{-1})$ are $L|L$ polynomial matrices and $\boldsymbol{e}(t)$ is a vector of L stationary white zero mean noises sampled at time index t. A general solution which permits coloured driving noises is presented in Chapter 3, but we will otherwise generally assume that $\boldsymbol{E}(q^{-1}) = \boldsymbol{F}(q^{-1}) = \mathbf{I}$ throughout this thesis.

As the quality of the sound in the bright zone is of importance, we require a means of specifying an ideal, *desired* or *target* sound field. In the following, an $M|L$ polynomial matrix $\boldsymbol{D}(q^{-1})$ will represent the desired sound field. In the target matrix, we specify the desired (finite) impulse response, in the bright zone, of the electro-acoustical system. In order to avoid forcing the filter to try to predict the input signal (which, in the white noise input signal situation is, by definition, impossible), it is advisable to include a delay in the target matrix impulse response corresponding to the propagation delay from the nearest loudspeaker to the bright zone. In addition to the sound propagation delay, we may sometimes also allow an additional delay in the target matrix. This additional delay is referred to as a *modelling delay* and a longer such delay will normally reduce the per-frequency SFS error. Unfortunately, increasing the modelling delay also normally means that the temporal duration of any pre-ringing in the compensated system is increased. It is therefore important to find a balance in what modelling delay is allowed. The modelling delay will henceforth be denoted d_0 and the target matrix free of modelling delay will be denoted $\boldsymbol{D}_0(q^{-1})$, so $\boldsymbol{D}(q^{-1}) = q^{-d_0}\boldsymbol{D}_0(q^{-1})$.

Using the difference between the desired system response, as parametrized by $\boldsymbol{D}(q^{-1})$ and a compensated system, $\boldsymbol{\mathcal{H}}(q^{-1})\,\boldsymbol{\mathcal{R}}(q^{-1})$, we may express the Sound Field Synthesis error at time t by

$$\boldsymbol{\varepsilon}'(t) = \boldsymbol{\mathcal{H}}(q^{-1})\,\boldsymbol{\mathcal{R}}(q^{-1})\boldsymbol{r}(t) - \boldsymbol{D}(q^{-1})\boldsymbol{r}(t) = (\,\boldsymbol{\mathcal{H}}(q^{-1})\,\boldsymbol{\mathcal{R}}(q^{-1}) - \boldsymbol{D}(q^{-1}))\,\boldsymbol{r}(t). \tag{1.15}$$

In most problem formulations in the coming chapters, the SFS error $\boldsymbol{\varepsilon}'(t)$ will be considered in the bright zone only, in which case $\boldsymbol{\mathcal{H}}(q^{-1}) = \boldsymbol{\mathcal{H}}_B(q^{-1})$. In some problems, the SFS error in both the bright and the dark zone will be considered, in which case $\boldsymbol{\mathcal{H}}(q^{-1})$ represents the rational matrix $\begin{bmatrix} \boldsymbol{\mathcal{H}}_B(q^{-1}) \\ \boldsymbol{\mathcal{H}}_D(q^{-1}) \end{bmatrix}$.

In many optimization problems, where we aim to minimize the SFS error, it is helpful to be able to assign more importance to reducing the error at certain frequencies or at certain control points. For this reason, we introduce the polynomial weighting matrix $\boldsymbol{V}(q^{-1})$, and obtain the weighted SFS error description by

$$\boldsymbol{\varepsilon}(t) = \boldsymbol{V}(q^{-1})\,(\,\boldsymbol{\mathcal{H}}(q^{-1})\,\boldsymbol{\mathcal{R}}(q^{-1}) - \boldsymbol{D}(q^{-1}))\,\boldsymbol{r}(t). \tag{1.16}$$

It is also often useful to have precise control over the resulting filter amplification factor and in what bands each filter is active. We therefore, similarly to the SFS error vector $\boldsymbol{\varepsilon}(t)$ above, introduce the polynomial weighting matrix $\boldsymbol{W}(q^{-1})$ on the control signal vector $\boldsymbol{u}'(t)$ and obtain a weighted filter output signal by

$$\boldsymbol{u}(t) = \boldsymbol{W}(q^{-1})\,\boldsymbol{\mathcal{R}}(q^{-1})\boldsymbol{r}(t). \tag{1.17}$$

The above system description is summarized in Figure 1.7.

In the practical experiments performed herein, we will often refer to the results in terms of input-output gain of acoustic power at a set of points in space. It should be noted that the quantities referred to are generally not absolute sound powers, but rather (squared) numerical representations of sums or means of the square of the electric microphone outputs. These outputs are proportional to the actual sound pressure at each point, but a numerical scaling is needed if it is to represent the correct physical quantity of pressure. Further, acoustic power is also proportional to the area over which it is measured and to the angle of incidence of the impinging sound field. We here assume that the investigated area is proportional to the number of control points. Reported acoustic powers herein thus also differ from the actual sound powers by

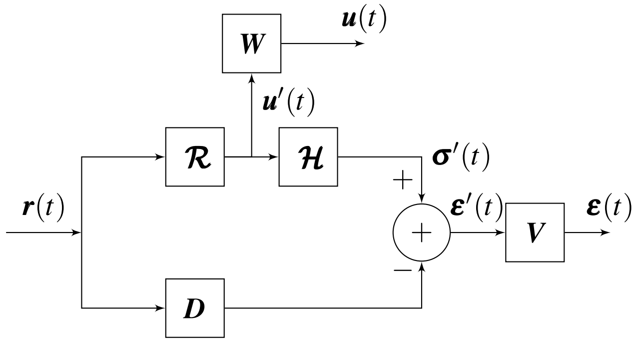

Figure 1.7. A pre-compensated electro-acoustical system with desired response $\boldsymbol{D}(q^{-1})$, weighted control signal $\boldsymbol{u}(t)$ and weighted error vector $\boldsymbol{\varepsilon}(t)$.

some scale factor. Reported acoustic power *gains* should however be quite representative.

1.4.2 Definition of Measures of Performance

To be able to qualitatively evaluate the filters-to-be designed, it is important to establish some measures of performance. We shall in the following adopt the notational convention $\mathscr{F}\{\boldsymbol{M}(q^{-1})\}|_\omega$ to denote the discrete Fourier transform of the polynomial matrix $\boldsymbol{M}(q^{-1})$ evaluated at the normalized angular frequency ω.[4]

The first and foremost performance measure in the context of personal audio is *acoustic contrast*. The acoustic contrast, at an arbitrary angular frequency ω, as generated by the pre-compensator designed for input signal l (column l of $\boldsymbol{\mathcal{R}}(q^{-1})$) is

$$C_l = \frac{M_D \mathbf{R}_l^H \mathbf{H}_B^H \mathbf{H}_B \mathbf{R}_l}{M_B \mathbf{R}_l^H \mathbf{H}_D^H \mathbf{H}_D \mathbf{R}_l}. \tag{1.18}$$

Above, the constant matrices $\mathbf{H}_B$ and $\mathbf{H}_D$ represent the rational matrices $\boldsymbol{\mathcal{H}}_B(q^{-1})$ and $\boldsymbol{\mathcal{H}}_D(q^{-1})$ respectively, evaluated at frequency ω. Similarly, the constant matrix $\mathbf{R}_l$ represents column l of the rational pre-compensator matrix evaluated at frequency ω

The contrast at frequency ω is thus the ratio of the acoustic power in the bright zone, at angular frequency ω to the acoustic power in the dark zone, at angular frequency ω.

The second measure that will be employed herein is the *filter power gain*. This measure describes the filter power amplification, as compared to a white,

[4]In the rational matrix formalism, this is equivalent to substituting $z^{-1} = e^{-j\omega}$ for the argument q^{-1}.

unit gain, zero delay filter (i.e., the equivalent of not performing any filtering). The filter power gain, of the l:th column of the pre-compensator is expressed, at angular frequency ω by

$$G_F = \mathbf{R}_l^H \mathbf{R}_l. \tag{1.19}$$

Third, a quantifiable measure of perceived sound quality in the bright zone would be helpful. Such a measure is, however, extremely complicated due to the various psychoacoustic effects associated with the human auditory system as well as the variability of personal preference between listeners. Attempts exist, such as, e.g., the planarity measure, which describes how well a generated sound field can be described by a plane wave propagating across a volume of space. These measures, however, generally only capture parts of the perceived sound quality.

In the investigations to follow, we will therefore generally refrain from using quantified sound quality measures and instead investigate the resulting sound fields by their spectral and temporal adherence to the desired sound field.

1.5 An Overview of the Field With Relations to the Present Thesis

Although the personal audio research field is relatively young, it has been quite active and it is useful, to at least have a summary overview of the various ways in which the problem has been treated since the first papers by Druyvesteyn et. al [3, 4].

1.5.1 Acoustic Contrast Control algorithms

The personal audio problem is revisited by Choi and Kim in [27] where the Acoustic Contrast Control (ACC) method is proposed. The Acoustic Contrast Control (ACC) algorithm produces a filter that maximizes the acoustic contrast (1.18) between two zones, and a more suitable name may be the acoustic contrast *maximization* algorithm. This is one of the most influential works in the personal audio literature and has been investigated and expanded upon in a large number of publications, see, e.g., [8, 28–36].

This work is followed up by Shin et al. in [37] and elaborated in [38] by the introduction of the Energy Difference Maximization (EDM) algorithm. Instead of maximizing the *ratio* between the acoustic energy in the bright zone and the dark zone, the EDM algorithm maximizes the *difference* between the acoustic energy in the bright zone and the dark zone. This produces a problem formulation with improved numerical properties as compared to the ACC algorithm, since the inversion of a likely poorly conditioned matrix is avoided. The EDM formulation also allows for some control over the radiation efficiency into the bright zone, while the maximum attainable contrast is generally slightly poorer than that of the ACC formulation [38].

Another important method to combat poorly conditioned inverse matrices (that is an integral problem also to, e.g., many Pressure Matching (PM) based algorithms) is regularization [31, 32]. Too much regularization, however, comes with adverse effects on the filter performance [32]. Additionally, the need for regularization is not frequency independent and it may be wise to consider different regularization for different frequencies, e.g., as investigated in [32, 39].

Relating back to Figure 1.3, we can summarize by noting that the ACC algorithm produces a filter that occupies the upper tip of what is accessible to linear filters of the triangle while the EDM algorithm yields a filter that accesses a point somewhat below (and to the left of) the upper tip of the triangle. Applying regularization, we move both points accessed by the ACC and EDM algorithms down towards the lower right corner of the triangle, reducing both the attained contrast and the filter gain.

While acquiring good acoustic contrast is important in personal audio problems, there are, as mentioned in the introduction, many situations where this is not the only important measure. This has led to a large body of work aiming to extend the ACC algorithm to incorporate a plethora of other qualitative aspects in addition to contrast maximization, spanning a larger part of our conceptual triangle of Figure 1.3.

A method of limiting the variability of the power over the bright zone, and thus expanding the range of the design method towards better bright zone acoustics, is proposed and investigated in [34]. The method is found to reduce the pressure difference between control points as compared to ACC while producing contrast on par with the ACC method.

It is shown in [9, 31, 40, 41] that the ACC method can be formulated as a problem of maximizing the bright zone acoustic power while keeping the dark zone power at a constant level. This is termed the direct formulation. Alternatively, an indirect formulation is also presented [31, 41] where the dark zone power is minimized subject to fixed bright zone power. The indirect formulation is better conditioned due to how the introduced array power[5] constraint acts as a regularizing term. The indirect method is also shown to produce an identical result to the EDM method but it is argued that the indirect method provides better insight into the inner workings of the filter design process. Although the solution to find the optimizing filters are equivalent to that of the ACC method, the final filters may not be equal since the methods proposed in [9, 31, 40, 41] are formulated subject to constrained acoustic potential energy in either the bright or the dark zone, which is not the case for the original ACC method.

1.5.2 Pressure Matching and Other Multi-Point and MSE Techniques

While the ACC method is perhaps the most widely explored method for personal audio problems, other more general methods, often drawing from the sound field synthesis literature, have also been investigated.

Of these, the most extensively investigated methods are versions of the PM method, where a filter is tuned to minimise the synthesis error between the sound field as generated by the compensated electro-acoustical system and a desired sound field. Often, a Least Square (LS) approach is used to achieve the minimum SFS error, see e.g., [7, 42]. The PM method has been extended with constraints on, e.g., filter power [43], acoustic power in the dark zone [20], and bright zone SFS error [21, 32].

A comparative investigation between a version of the PM method with weighted SFS error minimization and a version where the SFS error is constrained to be zero, is undertaken by simulation in [21], and by using anechoic measurements in [44]. The constrained method appears to produce better performance in terms of SFS error while producing roughly the same contrast as the weighted method does in [21]. On the other hand, it is reported in [44] that

[5] which relates to the filter power gain

the weighted version, with a certain weight, produces a better result in terms of directivity than the constrained version. This implies that the directivity may benefit from making trade-offs in both the bright and the dark zone as opposed to only compromising with the quality of either. There is also a possibility that the constrained method is less robust with respect to modelling errors.

An approach that combines PM with the EDM method via user specified weights is proposed by Møller et al. in [45]. This approach allows the filter designer to trade generated contrast against some freely specified bright zone properties. Cai et al. similarly combine a PM sound field synthesis term with a constraint on minimum allowed contrast by the EDM formulation in [19]. This means that a contrast can be specified by the user and a filter is produced by the proposed method that generates the smallest possible SFS error given that the specified contrast is attained.

A weighted particle velocity term is added to the PM method in [46] with the intention to describe the sound field on the border of the dark zone. The performance of filters computed with the proposed Joint Pressure and Velocity Matching (JPVM) method is compared to filters designed using the regular PM method on the border of the control zone. This approach is also investigated and modified in [47]. The paper [47] indicates that the proposed JPVM methods are more robust to low observability ill-conditioning than the PM method in anechoic conditions. The effects of particle velocity are also investigated in [48]. There, a constrained PM method is compared to a method that is likewise constrained but also constrained with respect to the maximally allowed particle velocity at the control points in the dark zone (termed the Particle Velocity Constraint (PVC) method). The idea is to define a dark zone around the bright zone and the loudspeakers and limit leakage from the bright zone by limiting the escaping particle velocity. The two methods are computed using different norms and compared in terms of acoustic contrast and pressure matching error levels. It is found that there are benefits to the sound field synthesis of reducing the external radiation in a reverberant room. It is also found that the PVC method is superior to the PM method in terms of localization performance and acoustic contrast.

Another popular cost function is the Planarity Control (PC) method, introduced in [49]. This method minimizes the dark zone power at the control points, while also reproducing a plane wave within a range of acceptable propagation directions in the bright zone. The method is evaluated via simulations and compared to the ACC method and the EDM-PM hybrid proposed in [45]. It is demonstrated that allowing the method to choose the best plane wave propagation direction may indeed be beneficial in certain circumstances. The PC method is compared to the PM method in [50] and to the indirectly formulated ACC method in [35].

Beside the PM approach, there are also other examples of filter generating algorithms for personal audio that use mathematical frameworks from the general sound field synthesis literature. One such is the null space based approach to sound field control, investigated in [51]. Personal audio filter generation by Generalized Singular Value Decomposition (GSVD) is proposed and evaluated in [52]. The analytical properties of the GSVD as a tool for analysing personal audio problems are also highlighted. This is expanded upon and compared to an optimal LS method in [53].

In [54], the authors draw instead from the field of wireless communication and arrive at an iterative algorithm for maximizing the signal-to-interferer (and noise) ratio subject to constrained loudspeaker power emission.

An interesting aspect of the PM method is that, while it is obvious that the properties of the bright zone sound field are affected by the choice of target sound field, it is less obvious that this also affects the maximum attainable contrast. Assume that we derive a filter that maximizes the contrast in a given electro-acoustical system. We may then use the resulting compensated system as the target sound field for a PM optimization. Such a design should produce the contrast maximizing filter again, save from the differences in what matrices are inverted. This does not necessarily mean that PM methods can be used to attain the same level of contrast in practice as, e.g., the ACC algorithm, since we have no structured way of designing such a contrast maximizing target. It does however mean that the basic PM algorithm spans a larger area in our conceptual trade-off triangle (Figure 1.3) than the directly contrast maximizing algorithms do and that any comparison of generated contrast involving a PM based algorithm depends on the particular choice of target sound field.

With all the recently proposed extensions to both the ACC method and to the PM based method, the dividing line between them has become blurred. Particularly methods that utilize properties of both approaches [19, 45] make the distinction obsolete. It may still be useful however, on a conceptual level, to keep the different origins and motivations of the methods in mind.

1.5.3 Wave Equation Based Methods

There is a rich tradition in sound field synthesis of using algorithms that are based on mathematical knowledge of the wave equations. Prominent examples are Wave Field Synthesis [55] and Higher Order Ambisonics [56, 57]. These methods often need to make restrictive assumptions about the source properties or distribution, or about the shape of the zones to be controlled. Nevertheless, such techniques have been used to produce some interesting and promising results, also for personal audio applications. Note that the distinction between multipoint methods and wave equation based methods is perhaps flawed and an argument can be made for some methods belonging to both categories. We have here made the distinction based loosely on design motivation, where the more explicit wave equation theory based designs are described in the following.

Generation of two individual sound zones by elimination of lower order modes in a cylindrical expansion of the sound field under a free-field assumption is investigated in [58]. Generation of personal sound systems using ambisonics is investigated by Poletti in [43]. Mode Matching as an approach to personal audio filter generation is investigated in [59]. It is found in [60] that the PM method produces approximately equal results as the Mode matching method for frequencies below the spatial Nyquist frequency.

Generation of bright and dark zones using the spatial Fourier transform was investigated both analytically and experimentally in [61,62], where the proposed method was compared to the Delay-and-Sum beamforming (DS) method, the ACC method, and the LS method under anechoic conditions. It is reported that the proposed, analytical methods are more robust with respect to modelling errors than the inversion based, measurement driven methods to which they are compared.

With a similar motivation as in [46, 47], a sparse sampling framework is utilized to reduce the number of microphones needed to correctly quantify the sound field and used as a basis in a personal audio problem formulation in [63,64]. The method is roughly based on describing the sound field in each zone by a free field Green's function together with a 'corrective soundfield'. It is assumed that the corrective soundfield can be constructed from a relatively small number of user definable basis Helmholtz solutions. It is further assumed that no scatterers are present inside the circular 'reproduction region' which is surrounded by the loudspeakers and includes both the (disc shaped) bright and dark zone. The latter assumption is hardly valid for the general automotive case and must be relaxed if such applications are to be explored.

The personal audio problem is also investigated from a privacy perspective, in which a goal is to reduce the intelligibility in the dark zone, of a material leaking from the bright zone, in [65]. A design method is proposed in which the sound quality in the bright zone is traded against reduced intelligibility in the dark zone. Contrast in conjunction with a masking noise, with spectrum matched to the leakage from the bright zone, is employed to reduce the intelligibility of the leaked bright zone material in the dark zone. The method is derived with an emphasis on reducing artefacts that occur when the idealized framework is implemented in the real world and the results are promising.

1.5.4 Causality

A causal version of the ACC method, by the indirect, constrained, formulation is proposed in [41]. The causal ACC method [41] is expanded upon in [66] where a bright zone inter-frequency variability term is introduced. This method is termed Broadband Acoustic Contrast Control with Response Variation minimization (BACC-RV) and a robust version is proposed in [67]. A modified approach with a reduced number of tuning parameters is further proposed in [68]. The method is improved by minimizing the variability over a moving average of frequencies in [69] as opposed to the variability of two neighbouring frequency bins as is the approach proposed in [68].

A causal version of the PM formulation, where SFS error reduction in the bright zone is traded against reduction of dark zone sound power is proposed and investigated in an automotive setting in Paper I, and in free-field simulations in [18].

A similar formulation is also explored in [70], with an added penalty term that is used for temporal impulse response shaping which can be used to reduce pre- or post-ringing in the compensated impulse responses.

Of the investigated causal methods, only Paper I and [18,70] allows explicit specification of a desired bright zone target sound field with phase properties.

1.5.5 The Higher Frequencies

As noted already by Druyvesteyn and Garas in [4], the fact that the sound field becomes more erratic at higher frequencies, coupled with the increased sensitivity to changes in the environment, renders active control of sound inappropriate at higher frequencies. The two main ways in which the community has dealt with the higher frequency part of the personal audio problem are by *proximity* and *directivity*.

Directivity

Directivity is attained either by utilizing loudspeakers with strong directive properties or, more commonly, by beamforming for arrays of loudspeakers. The latter approach has a long and rich history, both in audio, telecommunication and RADAR applications (see, e.g., [71], treating the dual receiver side problem). These methods differ from the more specialized personal audio algorithms in that they do not consider the dark zone, but try to reduce the energy radiated anywhere but into the desired direction or region. Beamforming has the advantage that it can be made less dependent on the room acoustics (and therefore more robust with respect to changing dynamics) than the methods that optimize for certain regions of space. It may therefore constitute a possible approach for controlling the higher frequency part of the personal audio problem. Note that the methods presented in this section may be used also in the lower frequencies, some also are in the cited publications. The beamforming problem does normally not take room reflections into account, however, and does not attempt to generate bright or dark zones. For these reasons, beamforming approaches are here sorted under the treatment of the higher frequencies.

Most methods that were originally designed for personal audio have also been applied to the beamforming problem. Instead of a dark zone, we then define a set of directions in the beam radiation pattern into which little power should be transmitted.

For instance, a version of the ACC method with weighted input signal power is devised and compared to the LS approach with respect to beamforming in [29]. The EDM approach is likewise compared to the PM approach in [72]. In [33], the ACC and PM methods are compared to each other and to the DS beamforming method in a reverberant room and evaluated via listening tests. It is found that the PM method attains a better planarity score[6] but produces much less contrast and under-performs in terms of perceived distraction from sound leakage.

A version of the ACC method that is constrained to produce a specified bright zone power while keeping the filter power gain below a pre-specified level is compared to a PM method derived subject to filter power gain constraints in [5]. It is found in [5, 73] that while the ACC method may attain higher theoretical contrast levels (assuming a certain bright zone target sound field), the LS-PM problem is better conditioned and more robust to measurement errors.

Curved sound beams have been explored using accelerating acoustical beams in [74], and formalized by the Tangent Line method as introduced in [75, 76] where it is also compared to the ACC method. Such precise beamforming may bring about increased control over the contribution to the ambient sound field in the automotive environment of scattered sound beams but the robustness and flexibility of the method is not very well known as of yet.

A method referred to as the Sound Power Minimization algorithm is presented in [6] and compared to the ACC algorithm, the EDM algorithm, the LS algorithm and the DS algorithm from a beam forming point of view. It is reported in simulations that all algorithms perform similarly with respect to directivity at higher frequencies.

Overall, as with the personal audio problem, the PM methods generally generate less contrast in the extreme than the contrast maximizing methods do. The PM methods do however provide better control over the bright zone properties, or in this case, the properties of the sound emitted in the beam. At wavelengths shorter than the inter-ear distance of a listener, however, the importance of phase from a subjective sound quality perspective is reduced and algorithms with more emphasis on directivity, or contrast generation, may be preferable.

Loudspeaker configuration

The problem of array design is two-fold. The issue of designing a filter that produces good directivity with an arbitrary array geometry is discussed above, but not all arrays are created equally. The other aspect of array design is choosing the loudspeaker placement and configuration that provides the most attainable directivity with the smallest number of loudspeakers or the lowest power

[6]The planarity score is a measure of how closely the sound field in the bright zone resembles a plane wave propagating through the zone from an arbitrary direction, see e.g., [49].

consumption. This may also extend beyond beamforming problems into the lower frequency part of the problem but choices regarding the placement of the low and mid frequency range loudspeakers in an automotive environment are often severely restricted.

One way of reducing the number of loudspeakers needed is to select, from a set of candidate loudspeakers, the subset of loudspeakers that is best able to reproduce a desired sound field. This approach is explored using the Least absolute shrinkage and selection (Lasso) operator in [77,78]. The computational complexity of the Lasso approach proposed in [77] is reduced for wideband filters in [79]. Loudspeaker selection using a sequential forward-backward search [80, p. 220] is evaluated based on a cost function that aims to model listener distraction due to inter-zone sound leakage (which is source material dependent) in [81], and for several different criteria in [82]. It is found in [83] that in a comparison between loudspeaker selection methods using the Lasso operator, a singular value decomposition method [84], a constrained matching pursuit (CMP) method [85], and a method that tries to optimize the controllability of the soundfield, independently of the desired bright zone properties [86], that the Lasso approach performs the best without any constraints on array effort, but that the CMP method is preferable if the array effort is constrained.

Another approach is to design the array geometry based on basic understanding of the physics and mathematics involved in the beamforming problem. One example can be found in [87] where an optimal array for crosstalk cancellation is devised. The proposed array may not be practically implementable for the personal audio problem in most automotive environments but valuable insight into the considerations in array design may nevertheless be gained.

A simulation based investigation of the performance of different array geometries for arrays located on the ceiling of a car can be found in [88,89]. Three different array locations are proposed for systems with few, distinct reflections in [90]. These source placement strategies are reported to provide a tangible improvement in contrast as compared to random source array placement.

Using two loudspeakers in a coupled back-to-back configuration and only (actively) driving one of them produces increased directivity as compared to a single loudspeaker but requires less power input than using two uncoupled loudspeakers [40]. The second, not actively driven loudspeaker acts as an acoustical resistor, producing supercardioid beampattern sources, see e.g., [91]. The reduced power input also reduces the sensitivity with respect to transfer function estimation errors of the coupled loudspeaker configuration as compared to the uncoupled configuration [92]. An array of supercardioid 'phase shift' sources used to reduce back-radiation for personal audio applications is investigated in [93, 94].

Parametric Arrays

Parametric arrays utilize ultrasonic transducers to attain very focused sound beams with relatively small spatial requirements. Parametric arrays were originally investigated with underwater applications in mind, but has since also been investigated for in-air audio reproduction, which is what we focus on here. An audible sound wave is modulated on top on an ultrasonic carrier wave. The combined waveform is demodulated via non-linear interaction with the medium (air) and the audible frequency part of the signal can then be heard by listeners within the narrow beam. The use of parametric arrays was first discussed (for underwater applications) by Westervelt in [95] and has since attracted considerable research interest, see e.g., [96–99]. Due to the good directivity performance, the method constitutes an attractive solution in both higher and lower frequencies (although the arrays have traditionally struggled to reproduce the very lowest frequencies, see e.g., [98]). The attractive performance to size ratio of these arrays has also appealed to commercial interests and a wide array of patents have been filed, see e.g., [100–105].

There are also, however, drawbacks of using modulated ultrasonic carrier frequencies for localized sound reproduction. The major such drawback, that has thus far been prohibitive in high fidelity sound applications, is the non-linear nature of the physical mechanism for de-modulation of the modulated ultrasonic sound beam into audible frequencies. This causes significant harmonic distortion of the audible sound and a large part of the research literature (aimed at in-air audio reproduction) revolves around addressing this and related issues.

In addition to the audio quality issues, safety concerns have been raised regarding exposure to high ultrasonic sound pressure levels. The research field has not yet reached conclusion as to the health effects of exposure to broadband ultrasound. Consideration should therefore be given when designing such systems that the required sound pressure levels may be kept low. An overview over the state of the ultrasound safety research and regulation as of 2012 can be found in [97, Section 7].

Proximity

Relative proximity between the bright zone and a set of loudspeakers as compared to the dark zone is the other major way in which the higher frequency part of problem is treated. The transmitted power from an omnidirectional sound source to a point in space decays quadratically with distance. If the relative distance between the sound source and the bright zone is several times shorter than the distance between the source and the dark zone, substantial acoustical separation can then be expected.

The proximity to the ears of a listener is leveraged in conjunction with active methods for contrast generation in a system with headrest mounted loudspeakers in [106–108].

Another approach is found in [109] where fast sound power decay with respect to distance using evanescent wave generation is investigated with some interesting results, using a line array in [109], and a circular array in [110]. Any means capable of delivering acoustic power into a small (or preferably arbitrary) volume of space with sharp decay outside of said space is naturally interesting from a personal audio approach. Unfortunately, array geometry constraints and lack in control of the spatial power delivery mechanism lends this particular approach limited applicability in most automotive systems.

1.5.6 Robustness

Robustness is an important aspect of personal audio. In all filter designs, there will be a certain amount of modelling error due to faulty system parametrization, inaccurate model assumptions, sound field variability [111], or other causes. In methods based on destructive interference, these errors become particularly problematic.

One specific type of modelling error that is very hard to overcome is scattering of the incident sound field on the listeners' bodies, which causes the sound power level in adjacent dark zones to increase [112]. This effect becomes more severe with increasing frequency [112, 113].

Greater generated contrast and in particular the higher required array gains makes the ACC algorithm particularly sensitive to modelling errors [113]. Sensitivity of the ACC algorithm with respect to various types of modelling errors is also investigated theoretically in [114] and numerically in [115].

Most literature in the personal audio field which aims to produce robust filters do so using regularization [31, 32, 116]. The regularizations act by promoting filters with reduced gain that therefore amplify model errors to a lesser extent. A version of the ACC algorithm that is robust to certain types of transfer function variability, based on probabilistic modelling was however proposed in [31, 41, 117]. In [117], both a probabilistic error solution and a worst case solution are utilized and investigated both from a robustness point of view and as a means for design point interpolation, with encouraging results.

While the vast majority of all sound field synthesis literature revolves around the assumption of linear and time-invariant (LTI) systems, all practical loudspeakers will produce non-linear artefacts if driven hard enough. It is indicated in [118] that regularization can be used to reduce the level of generated non-linear distortion, but at the expense of reduced contrast and/or increased SFS error.

Robustness is very important also in array design where loudspeaker elements are located close to each other and small deviations between model and reality may have a large effect on the directivity, unless care is taken in the design step. Robustness with respect to loudspeaker manufacturing tolerances is investigated in [119]. Robustness of the listener experience with respect to head position is investigated in [120] with motivations from the closely related crosstalk cancellation problem. As the listener experience is more robust to head movements towards and away from the array [120], the main focus lies on robustness to head movement parallel to the array.

Many publications and methods assume a point source loudspeaker model, this is challenged in [121, 122], where filters for personal audio applications are produced, using an ACC and an LS-PM algorithm respectively. It is shown in [121] that the physical extent of loudspeaker baffles have both detrimental and beneficial effects with respect to array directivity but in the end improve the attained directivity at higher frequencies. It is reported in [122] that the best performance in each setting is produced by the filters that are based on the relevant set of transfer functions. It is also, however, reported that the best *perceived* performance in informal listening tests are often attained by the filters based on the point source model. This could indicate that over fitting is a problem in personal audio filter design and that using idealized point source loudspeaker models may to some extent alleviate this problem.

Another common approach in the literature is to assume that the direct wave is dominant and that free field propagation is a good approximation to the actual reverberant case. It is however shown in [123] that taking the reverberant sound field into account in a statistical sense acts as a regularization, making the generated solution more robust at the expense of attained contrast and/or SFS error.

In addition to robustness issues arising from limited knowledge of the electro-acoustical system dynamics, also mathematical robustness of the algorithms must be considered. It is for instance well documented that the original ACC algorithm struggles with such issues which motivated the derivation of the EDM algorithm [38].

SFS problems in general attempt to replace one system behaviour with another, this is expressed mathematically as an inverse. This inverse (be it explicitly expressed in the filter generation algorithm or implicitly so) is at the root of most mathematical stability issues pertaining to the SFS research field in general, and the personal audio research field in particular. This problem can be alleviated via problem reformulations, promoting inverses of different aspects of the original system, which may affect the involved condition numbers (c.f., e.g., the direct and the indirect methods [31]). This problem is also commonly addressed using regularization, which may degrade the best case performance for improved mathematical properties, often leading to improved actual performance in a real system. Even better is to include a filter power penalty which acts as a physically motivated regularization. This approach simultaneously promotes filters that do not utilize loudspeakers in frequency regions for which they are not designed and improves the mathematical properties of the solution see, e.g., [31].

1.5.7 Metrics of Inter-Zone Interference

Due to its general nature, contrast is, by a wide margin, the most common measure in the way of objectively evaluating an algorithm-system combination in terms of dark zone performance. The raw ratio of bright zone sound power to dark zone sound power is however not always a particularly good measure of how distracting, or otherwise annoying, the sound leaking in from the bright zone is. The actual disturbance, distraction or annoyance perceived in the dark zone is a very complicated topic and the answer varies with many variables including the measured quantity, the source material, if there is any visual stimuli, and the individual listener.

An attempt to quantify the interferer sound power level required for an interferer to be perceived as 'acceptable' in a situation where the listener is focusing on some other source material, or is reading, is described in [124]. The source material on which the listener is focused is further divided into two scenarios, information gathering or entertainment. In the first scenario, the test subjects are tasked with extracting information from an audio source. In the second they are asked to relax and enjoy their source material. It is found that road noise in a moving vehicle helps mask the interferer, increasing the threshold of acceptability in all three scenarios. It is also found that the threshold of acceptability varies strongly with the interferer material. It is finally reported that the relative levels of the interferer to the material on which the test subjects were focusing, that was acceptable for 95% of all participants ranges between approximately -31 dB and -39 dB in the entertainment scenario, depending on listener experience. In the information scenario, the range is between -12 dB and -42 dB depending on how the task was interpreted by the test subjects.

The beneficial effect of road noise masking is also corroborated by Baykaner et. al in [125] although it is reported that this effect is moderate. It is also reported that high pass filtering the interferer is detrimental to the level of acceptability while low pass filtering the interferer is beneficial. The type of source material is also here reported to be the most important variable. Interestingly, the acceptable interferer-to-desired-sound ratio is here reported to be far more optimistic than what is reported in [124], ranging between -5 dB and -30 dB for a range of different interferer and listener target programs.

The CASP model proposed in [126] is a mathematical model of the human auditory system. This model is used in [127] and in [128] as a basis for extraction of features with relevance to the level of acceptable interference in the dark zone. A principal component analysis is performed on the extracted features and used to predict the perceived distraction generated by the investigated interferer materials. When the model of [127] is validated against audio material on which the model design was not based, however, it is apparent that some refinement is needed. Such a refinement was developed in [129], and validated with encouraging results in [130, 131]. A fast computing implementation is further proposed in [132].

A list of descriptors for interference in situations where two different audio source materials are present in the same space is proposed in [133]. It is found that the descriptors ‘distraction’ and ‘balance and blend’ are the most useful parameters when describing this type of audio-on-audio interference.

The compromise between bright zone sound quality and contrast is further investigated in [134] in which listeners are asked to rate the distraction and bright zone sound quality produced by a set of filters. These filters produce bright zones with varying planarity score and contrast. It is found that while planarity correlates positively with perceived bright zone sound quality and contrast correlates negatively with perceived distraction, neither correlation is of unit magnitude.[7] This implies that also other factors matter. One such factor was, again, identified to be the source materials and their psycho-acoustical interaction.

1.5.8 Summary of the State of the Art

The major algorithms in personal audio are based on either of two objective functions, the ACC/EDM or the PM based methods. Today, there are several hybrid schemes that incorporate both of these and/or other objective function terms with similar but more relaxed objectives. A non-exhaustive list of the algorithms included in the overview above, and in what publication they are investigated is presented in Table 1.1. The table contains two columns, where the first provides references in which the algorithm in question is ‘significantly modified’, meaning that a non-trivial modification of the algorithm is presented. The second column contains more analytic references, in which each algorithm is investigated and/or compared to another algorithm.

Three main metrics are central to the generation of personal audio filters: the generated contrast, the bright zone acoustical properties, and the required filter gain. These have subsequently been addressed in the literature and hybrid methods now exist that allow explicit control over all these physical parameters [19, 45].

One aspect of filter generation, causality, that has bearing on both the quality of the sound in the bright zone and on the implementability of the generated filters has been left largely untreated. A few causal methods are proposed [18, 41, 66–70] but they offer limited control over either the desired bright zone sound field or over the contrast generating abilities of the filter.

It is widely accepted that robustness is an important issue in personal audio filter generation. There are several studies that dissect different methods from a robustness viewpoint and others that try to quantify the impact of cer-

[7] Due to various sources of estimation uncertainty, the actual numbers reported are not expected to be reliably representative.

tain types of model errors on the generated contrast. Thus far, the control effort constrained ACC algorithm and the basic LS based PM algorithm (see, e.g., [135]) have also been extended with systematic, error model based robust formulations.

It is also consensus that a broad-band personal audio implementation needs to be divided into at least two sub-systems, one dealing with low frequencies by active control and the other passively treating higher frequencies. Considerable effort is therefore also spent on loudspeaker geometry optimization, both for the low frequency and the high frequency sub-problems.

Efforts have also been directed towards determining a good model for describing and predicting the quality of a generated personal audio system in terms of inter-zone leakage and bright zone sound properties. It is clear that the contrast and planarity measures capture a substantial part of the perceived quality, but also that these measures alone do not sufficiently model all aspects of a well performing filter.

Table 1.1. *A break-down over what algorithms are investigated in what publication, sorted by year of publication.*

Method	Significantly modified in	Investigated in
Acoustic Contrast Control (ACC)	[27] [29] [41] [31] [66] [34] [67] [19] [76]	[108] [9] [8] [112] [28] [136] [31] [42] [33] [121] [113] [116] [121] [34] [35] [67] [19] [134] [137] [6] [118] [138] [89] [62]
Energy Difference Maximization (EDM)	[37] [38] [18]	[41] [72] [31] [49] [6]
Pressure Matching (PM)	[48] [46] [21] [44] [32] [70]	[139] [42] [72] [33] [122] [113] [116] [121] [34] [48] [35] [50] [123] [134] [6] [39] [46] [21] [137] [44] [138] [89] [62] [53]
Planarity Control (PC)	[49]	[34] [35] [50] [134] [118]
Delay-and-Sum beamforming (DS)		[33] [6] [137] [62] [21] [44]
Tangent Line	[75]	[76]
Sparse particle velocity	[63]	[64]
Evanescent waves	[109] [110]	
Null space based Sound Field Synthesis (SFS)	[51]	
Generalized Singular Value Decomposition (GSVD)	[52]	[53]
Filtered-X LMS		[4] [53]

1.6 Contributions of the Present Thesis in Relation to the Existing Literature

The present thesis is geared towards generation of *causal* pre-compensation filters for personal audio. This particular branch of the field has, as we saw in Section 1.5.4, received limited attention in the literature.

During the work that is reported herein, six novel filter designs for causal personal audio filters were proposed:

1. The causal weighted SFS formulation for personal audio, or alternatively, the causal version of the weighted PM approach.
2. The causal filter for personal audio with constrained maximal power contribution to the dark zone.
3. The causal filter for personal audio with constrained maximal SFS error.
4. The causal filter for personal audio with constrained minimal generated sound power difference between the bright and dark zone.
5. The robust causal filter for personal audio with constrained minimal generated sound power difference between the bright and dark zone.
6. The robust causal filter for personal audio with constrained minimal generated sound power difference between the bright and dark zone and constrained pre-ringing envelope magnitude.

These designs expand the library of available causal design methods for personal audio. All designs (save for the one described in item 6 above) are evaluated via simulations based on measured impulse responses of echoic environments.

In addition, an experimental investigation of the effect of small passenger induced variations to the system on the resulting filter performance is presented in Chapter 8. Even though this is an important type of uncertainty, at least for automotive applications, this aspect of robustness seems to be underexplored.

Finally, a method for deriving causal, stable IIR pre-compensators on rational matrix form subject to quadratic constraints is proposed. This method is utilized in the design of two of the filters listed above but may also have a wider applicability also outside of audio pre-compensation.

1.7 Summary

In this introductory chapter, we have presented the fundamental concepts of the personal audio problem. The present thesis has been outlined and the over-arching design choices have been presented and motivated. We have presented a pre-cursory overview over the mathematical specifics that define the rational matrix framework and the mathematical model that is the foundation for the discussions and derivations of the personal audio filters that are in focus in the following chapters. An overview over the research field was also given. We have seen that the problem of designing casual filters for personal audio is not very thoroughly treated in the previous literature, in spite of some convincing positive practical properties. We have also noted that robustness is particularly important in personal audio filter design.

References

[1] Jöns Jacob Berzelius. *Årsberättlese om framstegen i fysik och kemi*. P.A. Norstedt & Söner, 1837.

[2] Benjamin B Bauer. Stereophonic earphones and binaural loudspeakers. *Journal of the Audio Engineering Society*, 9(2):148–151, 1961.

[3] W. F. Druyvesteyn, Ronald M. Aarts, P. Gelat, and A. Ruxton. Personal sound. In *Proceedings of the Institute of Acoustics*, volume 16, 1994.

[4] W. F. Druyvesteyn and John Garas. Personal sound. *Journal of the Audio Engineering Society*, 45(9):685–701, September 1997.

[5] Jordan Cheer, Stephen J. Elliot, and Marcos F. Simón Gálvez. Design and implementation of a car cabin personal audio system. *Journal of the Audio Engineering Society*, 61(6):412–424, June 2013.

[6] Raymond van der Rots and Arthur Berkhoff. Directional loudspeaker arrays for acoustic warning systems with minimised noise pollution. *Applied acoustics*, 89:345–354, 2015.

[7] Marcos F. Simón Gálvez, Stephen J. Elliott, and Jordan Cheer. A superdirective array of phase shift sources. *The Journal of the Acoustical Society of America*, 132(2):746–756, 2012.

[8] Ji-Ho Chang, Chan-Hui Lee, Jin-Young Park, and Yang-Hann Kim. A realization of sound focused personal audio system using acoustic contrast control. *The Journal of the Acoustical Society of America*, 125(4):2091–2097, 2009.

[9] Stephen J. Elliott, , Harry Murfet, and Keith R. Holland. Minimally radiating arrays for mobile devices. In *The Sixteenth International Congress on Sound and Vibration (ICSV16)*, 2009.

[10] Lars-Johan Brännmark and Anders Ahlén. Robust loudspeaker equalization based on position-independent excess phase modeling. In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processning (ICASSP)*, pages 385–388, Las Vegas, March–April 2008.

[11] Hugo Fastl and Eberhard Zwicker. *Psychoacoustics*. Springer, Berlin, 3rd edition, 2007.

[12] Lars-Johan Brännmark and Anders Ahlén. Spatially robust audio compensation based on SIMO feedforward control. *IEEE Transactions on Signal Processing*, 57(5):1689–1702, May 2009.

[13] Lars-Johan Brännmark, Adrian Bahne, and Anders Ahlén. Compensation of loudspeaker-room responses in a robust MIMO control framework. *IEEE Transactions on Audio Speech and Language Processing*, 21(6):1201–1216, June 2013.

[14] Mikael Sternad, Mathias Johansson, and Jonas Rutström. Inversion of loudspeaker dynamics by polynomial LQ feedforward control. Presented at the 3rd IFAC ROCOND Symposium, Prague, 2000.

[15] Mikael Sternad and Anders Ahlén. Robust filtering and feedforward control based on probabilistic descriptions of model errors. *Automatica*, 29(3):661–679, May 1993.

[16] Lars-Johan Brännmark. *Robust Sound Field Control for Audio Reproduction: A Polynomial Approach to Discrete-Time Acoustic Modeling and Filter Design*. PhD thesis, Uppsala University, Uppsala, Sweden, 2011.

[17] Adrian Bahne. *Multichannel Audio Signal Processing*. PhD thesis, Uppsala University, Uppsala, Sweden, 2014.

[18] Marcos F. Simón Gálvez, Stephen J. Elliott, and Jordan Cheer. Time domain optimization of filters used in a loudspeaker array for personal audio. *IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP)*, 23(11):1869–1878, 2015.

[19] Yefeng Cai, Ming Wu, and Jun Yang. Sound reproduction in personal audio systems using the least-squares approach with acoustic contrast control constraint. *Journal of the Acoustical Society of America*, 135(2):734–741, February 2014.

[20] Terence Betlehem and Paul D. Teal. A constrained optimization approach for multi-zone surround sound. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 437–440, 2011.

[21] Ferdinando Olivieri, Filippo Maria Fazi, Mincheol Shin, Philip A. Nelson, Simone Fontana, and Lang Yue. A method to improve the quality of an audio signal reproduced by a compact loudspeaker array in a private sound zone. In *22:nd International Congress on Sound and Vibration (ICSV)*, 2015.

[22] Eliahu I. Jury. *Sampled Data Control Systems*. John Wiley and Sons, Inc., 1958.

[23] Adrian Bahne, Lars-Johan Brännmark, and Anders Ahlén. Improved loudspeaker-room equalization for stereo systems regarding channel similarity. In *Proceedings of the IEEE/IET International Conference on Audio, Language and Image Processing (ICALIP)*, pages 254–259, Shanghai, China, Jul. 2012.

[24] Annea Barkefors, Mikael Sternad, and Lars-Johan Brännmark. Design and analysis of linear quadratic gaussian feedforward controllers for active noise control. *IEEE/ACM Transactions on Audio, Speech and Language Processing*, 22(12):1777–1791, December 2014.

[25] Thomas Kailath. *Linear Systems*. Prentice-Hall, Englewood Cliffs, NJ, 1980.

[26] Ole Kirkeby. *Reproduction of Acoustic Fields*. PhD thesis, University of Southampton, UK, 1995.

[27] Joung-Woo Choi and Yang-Hann Kim. Generation of an acoustically bright zone with an illuminated region using multiple sources. *The Journal of the Acoustical Society of America*, 111(4):1695–1700, April 2002.

[28] Jung-Min Lee, Tae-Woong Lee, Jin-Young Park, and Yang-Hann Kim. Generation of a private listening zone; acoustic parasol. In *Proceedings of 20th International Congress on Acoustics*, 2010.

[29] Jung-Woo Choi, Youngtae Kim, Sangchul Ko, and Jung-Ho Kim. Super-directive loudspeaker array for the generation of personal sound zone. In *Audio Engineering Society Convention 125*. Audio Engineering Society, 2008.

[30] Jin-Young Park, Ji-Ho Chang, and Yang-Hann Kim. Generation of independent bright zones for a two-channel private audio system. *Journal of*

the Audio Engineering Society, 58(5):382–393, 2010.

[31] Stephen J Elliott, Jordan Cheer, Jung-Woo Choi, and Youngtae Kim. Robustness and regularization of personal audio systems. *IEEE Transactions on Audio, Speech, and Language Processing*, 20(7):2123–2133, 2012.

[32] Ferdinando Olivieri, Filippo Maria Fazi, Simone Fontana, Dylan Menzies, and Philip Arthur Nelson. Generation of private sound with a circular loudspeaker array and the weighted pressure matching method. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 25(8):1579–1591, 2017.

[33] Marek Olik, Jon Francombe, Philip Coleman, Philip JB Jackson, Martin Olsen, Martin Møller, Russell Mason, and Søren Bech. A comparative performance study of sound zoning methods in a reflective environment. In *Audio Engineering Society Conference: 52nd International Conference: Sound Field Control-Engineering and Perception*. Audio Engineering Society, 2013.

[34] Naomichi Yanagidate, Jordan Cheer, Stephen Elliott, and Takeshi Toi. Car cabin personal audio: Acoustic contrast with limited sound differences. In *Audio Engineering Society Conference: 55th International Conference: Spatial Audio*. Audio Engineering Society, 2014.

[35] Philip Coleman, Philip J. B. Jackson, Marek Olik, and Jan Abildgaard Pedersen. Personal audio with a planar bright zone. *The Journal of the Acoustical Society of America*, 136(4):1725–1735, October 2014.

[36] Mincheol Shin, Filippo M Fazi, Philip A. Nelson, and Fabio C Hirono. Controlled sound field with a dual layer loudspeaker array. *Journal of Sound and Vibration*, 333(16):3794–3817, 2014.

[37] Mincheol Shin, Sung Q. Lee, Hye-jin Kim, Kang Ho Park, Daesung Kim, and Semyung Wang. Maximization of the acoustic energy difference between listening and non-listening regions. In *Proceedings 16 International Congress on Sound and Vibration*, 2009.

[38] Mincheol Shin, Sung Q. Lee, Filippo M. Fazi, Philip A. Nelson, Daesung Kim, Semyung Wang, Kang Ho Park, and Jeongil Seo. Maximization of acoustic energy difference between two spaces. *Journal of the Acoustical Society of America*, 128(1):121–131, July 2010.

[39] Ferdinando Olivieri, Filippo Maria Fazi, Mincheol Shin, and Philip Nelson. Pressure-matching beamforming method for loudspeaker arrays with frequency dependent selection of control points. In *Audio Engineering Society Convention 138*. Audio Engineering Society, 2015.

[40] Stephen J. Elliott, Jordan Cheer, Harry Murfet, and Keith R Holland. Minimally radiating sources for personal audio. *The Journal of the Acoustical Society of America*, 128(4):1721–1728, 2010.

[41] Stephen J. Elliott and Jordan Cheer. Regularization and robustness of personal audio systems. `http://eprints.soton.ac.uk/207989/1/Pub12685.pdf`, 2011. Accessed: 2018-03-29.

[42] Ji-Ho Chang and Finn Jacobsen. Sound field control with a circular double-layer array of loudspeakers. *The Journal of the Acoustical Society of America*, 131(6):4518–4525, 2012.

[43] Mark Poletti. An investigation of 2-d multizone surround sound systems. In *Audio Engineering Society Convention 125*. Audio Engineering Society, 2008.

[44] Ferdinando Olivieri, Filippo Maria Fazi, Philip A. Nelson, and Simone Fontana. Comparison of strategies for accurate reproduction of a target signal with compact arrays of loudspeakers for the generation of zones of private sound and silence. *Journal of the Audio Engineering Society*, 64(11):905–917, 2016.

[45] Martin Bo Møller, Martin Olsen, and Finn Jacobsen. A hybrid method combining syntheis of a sound field and control of acoustic contrast. In *132:nd Audio Engineering Society Convention*, April 2012.

[46] Michael Buerger, Roland Maas, Heinrich W Löllmann, and Walter Kellermann. Multizone sound field synthesis based on the joint optimization of the sound pressure and particle velocity vector on closed contours. In *Applications of Signal Processing to Audio and Acoustics (WASPAA), 2015 IEEE Workshop on*, pages 1–5. IEEE, 2015.

[47] Michael Buerger, Christian Hofmann, and Walter Kellermann. Broadband multizone sound rendering by jointly optimizing the sound pressure and particle velocity. *arXiv preprint arXiv:1709.07269*, 2017.

[48] M. R. Bai, J-C. Wen, H. Hsu, and Y-H. Hsieh. Investigation on the reproduction performance versus acoustic contrast control in sound field synthesis. *J. Acoust. Soc. Am.*, 136(4):1591–1600, October 2014.

[49] Philip Coleman, Philip Jackson, Marek Olik, and Jan Abildgaard Pedersen. Optimizing the planarity of sound zones. In *Audio Engineering Society Conference: 52nd International Conference: Sound Field Control-Engineering and Perception*. Audio Engineering Society, 2013.

[50] Philip Coleman, Philip JB Jackson, Marek Olik, and Jan Abildgaard Pedersen. Stereophonic personal audio reproduction using planarity control optimization. In *The 21:st international congress on sound and vibration (ICSV21)*, 2014.

[51] Yumiko Nakayama and Yosuke Tatekura. Narrow area control for individual sound image generation by combining nbsfc and liner loudspeaker array. In *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, volume 249, pages 3667–3674. Institute of Noise Control Engineering, 2014.

[52] Philippe-Aubert Gauthier, Yann Pasco, and Alain Berry. Generalized singular value decomposition for personalized audio using loudspeaker array. In *Audio Engineering Society Conference: 2016 AES International Conference on Sound Field Control*. Audio Engineering Society, 2016.

[53] Yann Pasco, Philippe-Aubert Gauthier, Alain Berry, and Stéphane Moreau. Interior sound field control using generalized singular value decomposition in the frequency domain. *The Journal of the Acoustical Society of America*, 141(1):334–345, 2017.

[54] Gema Piñero, Carmen Botella, Maria de Diego, Miguel Ferrer, and Alberto González. On the feasibility of personal audio systems over a network of distributed loudspeakers. In *Signal Processing Conference (EUSIPCO), 2017 25th European*, pages 2729–2733. IEEE, 2017.

[55] Augustinus J Berkhout, Diemer de Vries, and Peter Vogel. Acoustic control by wave field synthesis. *The Journal of the Acoustical Society of America*, 93(5):2764–2778, 1993.

[56] Michael A Gerzon. Ambisonics in multichannel broadcasting and video. *Journal of the Audio Engineering Society*, 33(11):859–871, 1985.

[57] Jens Ahrens and Sascha Spors. Analytical driving functions for higher order ambisonics. In *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, pages 373–376. IEEE, 2008.
[58] Mark A Poletti and Filippo Maria Fazi. An approach to generating two zones of silence with application to personal sound systems. *The Journal of the Acoustical Society of America*, 137(2):598–605, 2015.
[59] Wen Zhang, Thushara D Abhayapala, Terence Betlehem, and Filippo Maria Fazi. Analysis and control of multi-zone sound field reproduction using modal-domain approach. *The Journal of the Acoustical Society of America*, 140(3):2134–2144, 2016.
[60] Filippo Maria Fazi, Mincheol Shin, Ferdinando Olivieri, Simone Fontana, and Yue Lang. Comparison of pressure-matching and mode-matching beamforming for methods for circular loudspeaker arrays. In *Audio Engineering Society Convention 137*. Audio Engineering Society, 2014.
[61] Takuma Okamoto. Generation of multiple sound zones by spatial filtering in wavenumber domain using a linear array of loudspeakers. In *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, pages 4733–4737. IEEE, 2014.
[62] Takuma Okamoto and Atsushi Sakaguchi. Experimental validation of spatial fourier transform-based multiple sound zone generation with a linear loudspeaker array. *The Journal of the Acoustical Society of America*, 141(3):1769–1780, 2017.
[63] Wenyu Jin and W Bastiaan Kleijn. Multizone soundfield reproduction in reverberant rooms using compressed sensing techniques. In *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, pages 4728–4732. IEEE, 2014.
[64] Wenyu Jin and W Bastiaan Kleijn. Theory and design of multizone soundfield reproduction using sparse methods. *IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP)*, 23(12):2343–2355, 2015.
[65] Jacob Donley, Christian Ritz, and W Bastiaan Kleijn. Multizone soundfield reproduction with privacy-and quality-based speech masking filters. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 26(6):1041–1055, 2018.
[66] Yefeng Cai, Ming Wu, and Jun Yang. Design of a time-domain acoustic contrast control for broadband input signals in personal audio systems. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, May 2013.
[67] Yefeng Cai, Li Liu, Ming Wu, and Jun Yang. Robust time-domain acoustic contrast control design under uncertainties in the frequency response of the loudspeakers. In *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, volume 249, pages 5775–5780. Institute of Noise Control Engineering, 2014.
[68] Yefeng Cai, Ming Wu, Li Liu, and Jun Yang. Time-domain acoustic contrast control design with response differential constraint in personal audio systems. *The Journal of the Acoustical Society of America*, 135(6):EL252–EL257, 2014.
[69] Daan HM Schellekens, Martin B Møller, and Martin Olsen. Time domain acoustic contrast control implementation of sound zones for low-frequency

input signals. In *Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on*, pages 365–369. IEEE, 2016.

[70] Martin Bo Møller and Martin Olsen. Sound zones: On envelope shaping of fir filters. In *24th International Congress on Sound and Vibration, ICSV24*, pages 1–8, 2017.

[71] Barry D Van Veen and Kevin M Buckley. Beamforming: A versatile approach to spatial filtering. *IEEE ASSP Magazine*, 5(2):4–24, 1988.

[72] Mincheol Shin, Filippo Maria Fazi, Fabio Casagrande Hirono, and Philip Arthur Nelson. Control of a dual-layer loudspeaker array for the generation of private sound. In *41st International Congress and Exposition on Noise Control Engineering 2012 (INTER-NOISE 2012)*, 2012.

[73] Jordan Cheer and Stephen J Elliott. A comparison of control strategies for a car cabin personal audio system. In *Audio Engineering Society Conference: 52nd International Conference: Sound Field Control-Engineering and Perception*. Audio Engineering Society, 2013.

[74] Sipei Zhao, Yuxiang Hu, Jing Lu, Xiaojun Qiu, Jianchun Cheng, and Ian Burnett. Delivering sound energy along an arbitrary convex trajectory. *Scientific Reports*, 4:6628, 2014.

[75] Sipei Zhao, Xiaojun Qiu, and Ian Burnett. Acoustic contrast control in an arc-shaped area using a linear loudspeaker array. *The Journal of the Acoustical Society of America*, 137(2):1036–1039, 2015.

[76] Sipei Zhao, Xiaojun Qiu, Jing Lu, and Jianchun Cheng. Robust time domain acoustic contrast control in an arc-shaped area using a linear loudspeaker array. In *International Congress on Sound and Vibration (ICSV)*. INT INST ACOUSTICS & VIBRATION, 2015.

[77] Nasim Radmanesh and Ian S Burnett. Generation of isolated wideband sound fields using a combined two-stage lasso-ls algorithm. *IEEE Transactions on Audio, Speech, and Language Processing*, 21(2):378–387, 2013.

[78] Nasim Radmanesh and Ian S. Burnett. Effectiveness of horizontal personal sound systems for listeners of variable heights. In *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*, pages 316–320. IEEE, 2013.

[79] Nasim Radmanesh, Ian S Burnett, and Bhaskar D Rao. A lasso-ls optimization with a frequency variable dictionary in a multizone sound system. *IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP)*, 24(3):583–593, 2016.

[80] Pierre A. Devijer and Josef Kittler. *Pattern Recognition: A Statistical Approach*. Prentice Hall, 1982.

[81] Jon Francombe, Philip Coleman, Marek Olik, Khan Baykaner, Philip J. B. Jackson, Russell Mason, Martin Dewhirst, Søren Bech, and Jan Abildgaard Pederson. Perceptually optimized loudspeaker selection for the creation of personal sound zones. In *Audio Engineering Society Conference: 52nd International Conference: Sound Field Control-Engineering and Perception*. Audio Engineering Society, 2013.

[82] Philip Coleman, Philip J. B. Jackson, Marek Olik, and Jan Abildgaard Pedersen. Numerical optimization of loudspeaker configuration for sound zone reproduction. In *21:st International Congress on Sound and Vibration*

(ICSV), 2014.

[83] Hanieh Khalilian, Ivan V Bajić, and Rodney G Vaughan. Comparison of loudspeaker placement methods for sound field reproduction. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 24(8):1364–1379, 2016.

[84] Hanieh Khalilian, Ivan V Bajić, and Rodney G Vaughan. Towards optimal loudspeaker placement for sound field reproduction. In *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*, pages 321–325. IEEE, 2013.

[85] Hanieh Khalilian, Ivan V Bajic, and Rodney G Vaughan. Loudspeaker placement for sound field reproduction by constrained matching pursuit. In *Applications of Signal Processing to Audio and Acoustics (WASPAA), 2013 IEEE Workshop on*, pages 1–4. IEEE, 2013.

[86] Futoshi Asano, Yôiti Suzuki, and David C Swanson. Optimization of control source configuration in active control systems using gram-schmidt orthogonalization. *IEEE Transactions on Speech and Audio Processing*, 7(2):213–220, 1999.

[87] Takashi Takeuchi and Philip A. Nelson. Optimal source distribution for binaural synthesis over loudspeakers. *The Journal of the Acoustical Society of America*, 112(6):2786–2797, 2002.

[88] Xiangning Liao, Stephen J. Elliott, Jordan Cheer, and Sifa Zheng. Design array of loudspeakers for personal audio system in a car cabin. In *International Congress on Sound and Vibration (ICSV)*, 2016.

[89] Xiangning Liao, Jordan Cheer, Stephen J. Elliott, and Sifa Zheng. Design of a loudspeaker array for personal audio in a car cabin. *Journal of the Audio Engineering Society*, 65(3):226–238, 2017.

[90] Marek Olik, Philip J. B. Jackson, Philip Coleman, and Jan Abildgaard Pedersen. Optimal source placement for sound zone reproduction with first order reflections. *The Journal of the Acoustical Society of America*, 136(6):3085–3096, 2014.

[91] Thomas J Holmes. The "acoustic resistance box" -a fresh look at an old principle. *Journal of the Audio Engineering Society*, 34(12):981–989, 1986.

[92] Jordan Cheer. Robustness and efficiency of an acoustically coupled two-source superdirective array. In *International Congress on Sound and Vibration (ICSV)*, 2015.

[93] Marcos Simón Gálvez, Stephen J. Elliott, and Jordan Cheer. Loudspeaker arrays for family tv. In *International Congress on Sound and Vibration (ICSV)*. International Institute of Acoustics and Vibration, 2012.

[94] Marcos F. Simón Gálvez and Stephen J. Elliott. The design of a personal audio superdirective array in a room. In *Audio Engineering Society Conference: 52nd International Conference: Sound Field Control-Engineering and Perception*. Audio Engineering Society, 2013.

[95] Peter J Westervelt. Parametric acoustic array. *The Journal of the Acoustical Society of America*, 35(4):535–537, 1963.

[96] F. Joseph Pompei. *Sound from ultrasound: the parametric array as an audible sound source*. PhD thesis, 2002.

[97] Woon-Seng Gan, Jun Yang, and Tomoo Kamakura. A review of parametric

acoustic array in air. *Applied Acoustics*, 73(12):1211–1219, 2012.

[98] Jun Yang, Jing Tian, and Woon Seng Gan. Parametric loudspeaker: From theory to applications. In *Proceedings of the 21st International Congress on Sound and Vibration*, pages 1–18, 2014.

[99] Chuang Shi, Yoshinobu Kajikawa, and Woon-Seng Gan. An overview of directivity control methods of the parametric array loudspeaker. *APSIPA Transactions on Signal and Information Processing*, 3, 2014.

[100] Tsuneo Tanaka, Mikio Iwasa, Youichi Kimura, and Akira Nakamura. Directional loudspeaker system, April 25 1989. US Patent 4,823,908.

[101] James J Croft III, Michael E Spencer, and Joseph O Norris. Modulator processing for a parametric speaker system, June 24 2003. US Patent 6,584,205.

[102] Frank Joseph Pompei. Ultrasonic transducer for parametric array, August 3 2004. US Patent 6,771,785.

[103] James Croft. Parametric loudspeaker system and method for enabling isolated listening to audio material, September 13 2007. US Patent App. 10/575,672.

[104] Joseph O Norris and James J Croft III. System for playback of pre-encoded signals through a parametric loudspeaker system, March 11 2008. US Patent 7,343,017.

[105] Jianchun Dong, Eliot Kim, and Mitchell J Heinrich. Parametric array for a head-mountable device, June 10 2014. US Patent 8,750,541.

[106] Stephen J Elliott and Matthew Jones. An active headrest for personal audio. *The Journal of the Acoustical Society of America*, 119(5):2702–2709, 2006.

[107] Chiho Chung and Stephen J Elliott. Personal audio headrest. In *Audio Engineering Society Convention 120*. Audio Engineering Society, 2006.

[108] Matthew Jones and Stephen J. Elliott. Personal audio with multiple dark zones. *The Journal of the Acoustical Society of America*, 124(6):3497–3506, 2008.

[109] Hiroaki Itou, Ken'ichi Furuya, and Yoichi Haneda. Evanescent wave reproduction using linear array of loudspeakers. In *Applications of Signal Processing to Audio and Acoustics (WASPAA), 2011 IEEE Workshop on*, pages 37–40. IEEE, 2011.

[110] Hiroaki Itou, Ken'ichi Furuya, and Yoichi Haneda. Localized sound reproduction using circular loudspeaker array based on acoustic evanescent wave. In *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pages 221–224. IEEE, 2012.

[111] Martin Olsen and Martin Bo Møller. Sound zones: On the effect of ambient temperature variations in feed-forward systems. In *Audio Engineering Society Convention 142*, May 2017.

[112] Ji-Ho Chang, Jin-Young Park, and Yang-Hann Kim. Scattering effect on the sound focused personal audio system. *The Journal of the Acoustical Society of America*, 125(5):3060–3066, 2009.

[113] Martin Olsen and Martin Bo Møller. Sound zones: Scattering study with head and torso simulator. In *Audio Engineering Society Conference: 52nd International Conference: Sound Field Control-Engineering and Perception*. Audio Engineering Society, 2013.

[114] Jin-Young Park, Min-Ho Song, Ji-Ho Chang, and Yang-Hann Kim. Performance degradation due to transfer function errors in acoustic brightness

and contrast control: sensitivity analysis. In *Proceedings of the 20th International Congress on Acoustics*, 2010.

[115] Jin-Young Park, Jung-Woo Choi, and Yang-Hann Kim. Acoustic contrast sensitivity to transfer function errors in the design of a personal audio system. *The Journal of the Acoustical Society of America*, 134(1):EL112–EL118, 2013.

[116] Philip Coleman, Philip J. Jackson, Marek Olik, Martin Olsen, Martin Møller, and Jan Abildgaard Pedersen. The influence of regularization on anechoic performance and robustness of sound zone methods. In *Proceedings of Meetings on Acoustics ICA2013*, volume 19, page 055055. ASA, 2013.

[117] Qiaoxi Zhu, Philip Coleman, Ming Wu, and Jun Yang. Robust acoustic contrast control with reduced in-situ measurement by acoustic modeling. *Journal of the Audio Engineering Society*, 65(6):460–473, 2017.

[118] Xiaohui Ma, Patrick J Hegarty, Jan Abildgaard Pedersen, Lars G Johansen, and Jakob Juul Larsen. Personal sound zones: The significance of loudspeaker driver nonlinear distortion. In *Audio Engineering Society Conference: 2016 AES International Conference on Sound Field Control*. Audio Engineering Society, 2016.

[119] Ivan Tashev, Jasha Droppo, Michael Seltzer, and Alex Acero. Robust design of wideband loudspeaker arrays. In *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, pages 381–384. IEEE, 2008.

[120] Darren B Ward and Gary W Elko. Optimum loudspeaker spacing for robust crosstalk cancellation. In *Proceedings of the 1998 IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, volume 6, pages 3541–3544. IEEE, 1998.

[121] Jordan Cheer, Stephen J. Elliott, Youngtae Kim, and Jung-Woo Choi. Practical implementation of personal audio in a mobile device. *Journal of the Audio Engineering Society*, 61(5):290–300, 2013.

[122] Ferdinando Olivieri, Mincheol Shin, Filippo M. Fazi, Philip A. Nelson, and Peter Otto. Loudspeaker array processing for multi-zone audio reproduction based on analytical and measured electroacoustical transfer functions. In *Audio Engineering Society Conference: 52nd International Conference: Sound Field Control-Engineering and Perception*. Audio Engineering Society, 2013.

[123] Marcos F. Simón-Gálvez, Stephen J. Elliott, and Jordan Cheer. The effect of reverberation on personal audio devices. *The Journal of the Acoustical Society of America*, 135(5):2654–2663, 2014.

[124] Jon Francombe, Russell Mason, Martin Dewhirst, and Søren Bech. Determining the threshold of acceptability for an interfering audio programme. In *Audio Engineering Society Convention 132*. Audio Engineering Society, 2012.

[125] Khan Baykaner, Christopher Hummersone, Russell Mason, and Søren Bech. The prediction of the acceptability of auditory interference based on audibility. In *Audio Engineering Society Conference: 52nd International Conference: Sound Field Control-Engineering and Perception*. Audio Engineering Society, 2013.

[126] Morten L Jepsen, Stephan D Ewert, and Torsten Dau. A computational model of human auditory signal processing and perception. *The Journal of the*

Acoustical Society of America, 124(1):422–438, 2008.
[127] Jon Francombe, Russell Mason, Martin Dewhirst, and Søren Bech. Modelling listener distraction resulting from audio-on-audio interference. In *Proceedings of Meetings on Acoustics ICA2013*, volume 19, page 035036. ASA, 2013.
[128] Jon Francombe, Russell Mason, Martin Dewhirst, and Søren Bech. The computational prediction of masking thresholds for ecologigally valid interference scenarios. In *Proceedings of Meetings on Acoustics ICA2013*, volume 19. ASA, 2013.
[129] Jon Francombe, Russell Mason, Martin Dewhirst, and Søren Bech. A model of distraction in an audio-on-audio interference situation with music program material. *Journal of the Audio Engineering Society*, 63(1/2):63–77, 2015.
[130] Jussi Rämö, Steven Marsh, Søren Bech, Russell Mason, and Søren Holdt Jensen. Validation of a perceptual distraction model in a complex personal sound zone system. In *Audio Engineering Society Convention 141*. Audio Engineering Society, 2016.
[131] Jussi Rämö, Lasse Christensen, Søren Bech, and Søren Jensen. Validating a perceptual distraction model using a personal two-zone sound system. In *Proceedings of Meetings on Acoustics 173EAA*, volume 30, page 050003. ASA, 2017.
[132] Jussi Rämö, Søren Bech, and Søren Holdt Jensen. Real-time perceptual model for distraction in interfering audio-on-audio scenarios. *IEEE Signal Processing Letters*, 24(10):1448–1452, 2017.
[133] Jon Francombe, Russell Mason, Martin Dewhirst, and Søren Bech. Elicitation of attributes for the evaluation of audio-on-audio interference. *The Journal of the Acoustical Society of America*, 136(5):2630–2641, 2014.
[134] Khan Baykaner, Philip Coleman, Russell Mason, Philip J. B. Jackson, Jon Francombe, Marek Olik, and Søren Bech. The relationship between target quality and interference in sound zone. *Journal of the Audio Engineering Society*, 63(1/2):78–89, 2015.
[135] Laurent El Ghaoui and Hervé Lebret. Robust solutions to least-squares problems with uncertain data. *SIAM Journal on Matrix Analysis and Applications*, 18(4):1035–1064, 1997.
[136] Finn Jacobsen, Martin Olsen, Martin Møller, and Finn Agerkvist. A comparison of two strategies for generating sound zones in a room. In *18th International Congress on Sound and Vibration*. International Institute of Acoustics and Vibration, 2011.
[137] Mingsian R Bai and Yu-Hao Hsieh. Point focusing using loudspeaker arrays from the perspective of optimal beamforming. *The Journal of the Acoustical Society of America*, 137(6):3393–3410, 2015.
[138] Martin Bo Møller and Martin Olsen. Sound zones: On performance prediction of contrast control methods. In *Audio Engineering Society Conference: 2016 AES International Conference on Sound Field Control*. Audio Engineering Society, 2016.
[139] Jung-Woo Choi, Youngtae Kim, Sangchul Ko, and Jung-Ho Kim. Super-directive loudspeaker array for the generation of personal sound zone. In *125:th Audio Engineering Society Convention*, October 2008.
[140] Mikael Sternad and Anders Ahlén. LQ controller design and self-tuning

control. In Kenneth Hunt, editor, *Polynomial Methods in Optimal Control and Filtering*, pages 56–92. Peter Peregrinus, London, UK, 1993.

[141] Vladimír Kučera. *Discrete Linear Control: The Polynomial Equation Approach*. John Wiley & Sons, Chichester, 1979.

[142] Vladimír Kučera. *Analysis and Design of Discrete Linear Control Systems*. Academia, Prague, 1991.

[143] Mikael Sternad and Anders Ahlén. A novel derivation methodology for polynomial-LQ controller design. *IEEE Trans. Autom. Control*, 38(1):116–121, January 1993.

[144] Annea Barkefors, Simon Berthilsson, and Mikael Sternad. Extending the area silenced by active noise control using multiple loudspeakers. In *IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, pages 325–328, 2012.

[145] Mathias Johansson, Lars-Johan Brännmark, Adrian Bahne, and Mikael Sternad. Sound field control using a limited number of loudspeakers. In *AES 36th International Conference on Automotive Audio*, Dearborn, Michigan, 2009. AES.

[146] Simon Berthilsson, Annea Barkefors, Lars-Johan Brännmark, and Mikael Sternad. Acoustical zone reproduction for car interiors using a MIMO MSE framework. In *AES 48th Internat. Conf. Automotive Audio*, Sep 2012.

[147] Lars Lindbom, Mikael Sternad, and Anders Ahlén. Tracking of time-varying mobile radio channels–Prt I: The Wiener LMS algorithm. *IEEE Trans. Commun.*, 49(12):2207–2217, December 2001.

[148] Stephen P. Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge, UK, 2004.

[149] Marcos F. Simón Gálvez. *Design of an Array-Based Aid for the Hearing Impaired*. PhD thesis, University of Southampton, Southampton, UK, 2014.

[150] Lars-Johan Brännmark. Robust audio precompensation with probabilistic modeling of transfer function variability. In *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pages 193–196, New Paltz, NY, Oct. 2009.

[151] Michael Grant and Stephen P. Boyd. CVX: Matlab software for diciplined convex programming.

[152] Lakshmi Krishnan. *Acoustic Room Impulse Response Shaping*. PhD thesis, 2017.

[153] Fuzhen Zhang. *The Schur complement and its applications*, volume 4. Springer Science & Business Media, 2006.

[154] Asger Heidemann Andersen. On multiple sound zones for wideband signals. Master's thesis, School of Information and Communication Technology, Aalborg University, 2014. Accessed: 2018-05-06.

[155] Kenth Öhrn. *Design of Multivariable Cautious Discrete-Time Wiener Filters: A Probabilistic Approach*. PhD thesis, Uppsala University, Sweden, 1996.

[156] Graham C Goodwin and Mario E Salgado. A stochastic embedding approach for quantifying uncertainty in the estimation of restricted complexity models. *International Journal of Adaptive Control and Signal Processing*, 3(4):333–356, 1989.

[157] Kenth Öhrn, Anders Ahlén, and Mikael Sternad. A probabilistic approach to

multivariable robust filtering and open-loop control. *IEEE Trans. Autom. Control*, 40(3):405–418, Mar. 1995.

[158] Rikke Apelfröjd and Mikael Sternad. Design and measurement-based evaluations of coherent jt comp: a study of precoding, user grouping and resource allocation using predicted csi. *EURASIP Journal on Wireless Communications and Networking*, 2014(1):100, 2014.

[159] Angelo Farina. Advancements in impulse response measurements by sine sweeps. In *Audio Engineering Society Convention 122*. Audio Engineering Society, 2007.

[160] Swen Müller and Paulo Massarani. Transfer-function measurement with sweeps. *J. Audio Eng. Soc.*, 49(6):443–471, June 2001.

[161] David H Raab. Forward and backward masking between acoustic clicks. *The Journal of the Acoustical Society of America*, 33(2):137–139, 1961.

[162] Robert Humphrey. Playrec - multi-channel matlab audio. `http://www.playrec.co.uk/`, 2008–2014. Accessed: 2018-04-27.

[163] Claes Tidestav. *The Multivariable Decision Feedback Equalizer: Multiuser Detection and Interference Rejection*. PhD thesis, Uppsala University, Sweden, 1999. Available online at `http://www.signal.uu.se/Publications/pdf/a993.pdf`.

[164] Anders Ahlén and Mikael Sternad. Wiener filter design using polynomial equations. *IEEE Trans. Signal Process.*, 39(11):2387–2399, Nov. 1991.

[165] AP Roberts and MM Newmann. Polynomial approach to wiener filtering. *International Journal of Control*, 47(3):681–696, 1988.