BOB McCARTHY

SOUND SYSTEMS: DESIGN AND OPTIMIZATION

MODERN TECHNIQUES AND TOOLS FOR SOUND SYSTEM DESIGN AND ALIGNMENT

Focal Press

Sound Systems: Design and Optimization

Modern techniques and tools for sound system design and alignment

Bob McCarthy

ELSEVIER

AMSTERDAM • BOSTON • HEIDELBERG • LONDON • NEW YORK • OXFORD
PARIS • SAN DIEGO • SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO
Focal Press is an imprint of Elsevier

Focal Press

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

First edition 2007

British Library Cataloguing in Publication Data
McCarthy, Bob
Sound systems design and optimization modern techniques and tools for sound system design and alignment
1. Sound - Recording and reproducing
I Title
621 3'893

Library of Congress Control Number: 2006933362

ISBN-13: 978-0-240-52020-9
ISBN-10: 0-240-52020-3

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

Typeset by Charon Tec Ltd (A Macmillan Company), Chennai, India
www.charontec.com
Printed and bound in Italy

07 08 09 10 11 11 10 9 8 7 6 5 4 3 2 1

Contents

Preface ix
Acknowledgements xvi

SECTION 1: SOUND SYSTEMS **1**

1. Transmission **3**
Transmission Goals 3
Audio Transmission Defined 4
Time and Frequency 4
Wavelength 5
The Waveform 8
Transmission Quantified 10
Decibels 10
Power 16
Frequency Response 17
Polarity 19
Latency 19
Analog Audio Transmission 19
Line Level Devices 20
Line Level Interconnection 32
Speaker Level Devices — Power Amplifiers 35
Speaker Level Interconnection — Speaker Cables 40
Digital Audio Transmission 42
Digital Audio Devices 42
Digital Audio Interconnection 43
Acoustic Transmission 44
Power, Pressure and Surface Area 44
Environmental Effects: Humidity and Temperature 46
Acoustic Transmitters: Loudspeakers 47
Reference 59

2. Summation **60**
Overview 60
Properties of Audio Summation 61
Audio Summation Defined 61
Summation Criteria 61
Summation Quantity 63
Electrical vs. Acoustical Summation 63
Acoustical Source Direction 64
Summation Math 64
Summation Amplitude 65
Summation Phase 66
Response Ripple 70
Summation Zones 71
Comb Filtering: Linear vs. Log 78
Summation Geometry 79
Acoustical Crossovers 87
Acoustic Crossover Defined 87
Crossover Classes 88

Spectral Dividers and Spectral Crossovers 89
Spatial Dividers and Spatial Crossovers 101
Speaker Arrays 112
Introduction 112
Speaker Array Types 112
Coupled Arrays 115
Uncoupled Arrays 125
Speaker/Room Summation 140
Analogous Functions 140
Speaker/Room Summation Types 143
Absorption Effects 147
Environmental Effects 147

3. Reception 150
Introduction 150
Loudness 151
Loudness and dB SPL 151
The Equal Loudness Contours 152
Localization 153
Introduction 153
Sonic Image 153
Vertical Localization 154
Front/Back Localization 157
Horizontal Localization 157
Tonal, Spatial and Echo Perception 164
Introduction 164
Tonal Perception 165
Echo Perception 167
Spatial Perception 168
Perception Zone Detection 169
Stereo Perception 169
Introduction 169
Panoramic Field 170
Stereo Scaling 171
Stereo Side-Effects 175
Amplified Sound Detection 176
Distortion 176
Compression 176
Frequency Response Coloration 176
False Perspective 176
Microphone Reception 178
Introduction 178
Comparing Microphones to Our Ears 178
Measurement Microphones 179
References 180

SECTION 2: DESIGN 181

4. Evaluation 183
Introduction 183
Natural Sound vs. Amplified Sound 184
Contrasting Emission, Transmission and Reception Models 185
Transmission Path Differences 186
Relating the Worlds of Acousticians and Audio Engineers 191
Comparing Our Goals 191
The Middle Ground 198
Moving Forward 201
Reference 205

5. Prediction 206
Introduction 206
Drawings 207
2-D Drawing Types 207
3-D Drawing Types 208
2-D Drawings in a 3-D World 208
Scale 212
Acoustic Modeling Programs 212
Introduction 212
A Very Brief History 213
Speaker Data Files 214
Acoustic Transmission Properties 220
Material Absorption Properties 225
Characterizing Summations 226
Applications 229
Conclusions 230

6. Variation 232
Introduction 232
The Minimum Variance Principles 235
Variation Defined 235
Causes and Standard Progressions of Variance 235
Minimum Variance vs. Maximum SPL 242

The Speaker/Room Link: Aspect Ratio 242
Protractors and Pizza 242
Maximum Acceptable Variance 243
Asymmetric Coverage Considerations 243
The Coverage Bow Method 244
The Forward Aspect Ratio 244
The Proximity Ratio 246
Minimum Level Variance 247
Single Speakers 248
Coupled Speaker Arrays 253
Uncoupled Speaker Arrays 260
Minimum Spectral Variance 265
Relating Aspect Ratio, Beamwidth and Speaker Order 265
Speaker Array Methods 274
Minimum Ripple Variance 303
Subwoofer Arrays 303
End Fire 304
Two-Element In-Line Technique 305
Inverted Stack 305
Coupled and Uncoupled Line Source 306
Coupled and Uncoupled Point Source 308
Conclusion 309
Speaker Order and Beamwidth 309
Maximum Power vs. Minimum Variance 310
Minimum Variance Coverage Shapes 310

7. Specification 314
Introduction 314
The Principles of Specification 314
Specification Defined 314
Goals and Challenges 315
Specific Questions 315
Specific Answers 318
Compromise 319
Channel/System Types 320
Mono 321
Stereo 321
Surround 322
Source Effects 324
System Subdivision 324
Subsystem Types 325
Main Systems 325
Sidefill 327
Infill 327
Downfill 328
Frontfill 328
Delays 328
The Scalable Design 330
Scaling for Power 330
Overlapping for Power 330
Scaling the Shape 331
Array Design Procedures 331
Main System Design 331
Symmetric Coupled Point Source 332
Asymmetric Coupled Point Source 333
Asymmetric-Composite Coupled Point Source 337
Symmetric Uncoupled Line Source 341
Asymmetric Uncoupled Line Source 342
Symmetric Uncoupled Point Source 342
Asymmetric Uncoupled Point Source 342
Symmetric Uncoupled Point Destination 342
Asymmetric Uncoupled Point Destination 343
The Diagonal Plane 343
Balcony Battles 346
Multichannel Sound 350

SECTION 3: OPTIMIZATION 353

8. Examination 355
Examination Defined 355
Physical Measurement Tools 356
Inclinometer 356
Protractor 356
Origami Angle Finder 356
Laser Pointer 357
Thermometer 358
Hygrometer 358
Simple Audio Measurement Tools 358
Volt/ohm meter (VOM) 358
Polarity Tester 359
Listen Box 359
Impedance Tester 359

Oscilloscope 360
Sound Level Meter 360
Real-Time Analyzer (RTA) 360
Complex Audio Measurement Tools 362
The Fourier Transform 363
Analyzer Basics 363
Signal Averaging 370
Single-Channel Spectrum Applications 372
Transfer Function Measurement 374
Other Complex Signal Analyzers 397
Analysis Systems 399
Reference 400

9. Verification 401
Introduction 401
Test Structure 401
Testing Stages 402
Access Points 403
Test Setup 403
Procedures 406
Noise Over Frequency 406
Total Harmonic Distortion + Noise (THD + n) 408
Maximum Input/Output Capability Over Frequency 409
Latency 412
Polarity 415
Frequency Response 415
Phase Response over Frequency 417
Compression 418
Microphone Verification 419
Microphone Match 420
Microphone Response 420
Post-Calibration Verification 421
Additional Considerations 423

10. Calibration 424
Calibration Defined 424
Goals 425
Challenges 425
Strategies 425
Techniques 426
Approaches to Calibration 426
Anarchy 426
Monarchy 427
Capitalism 427
Democracy 427
TANSTAAFL and Triage 428
Access to Information 428
Calibration Subdivision 428
Physical Access 430
Electronic Access 430
Acoustic Access 432
Microphone Position Strategies 445
Procedures 452
Acoustic Evaluation 452
Level Setting 453
Speaker Position Adjustment 455
Equalization 458
Delay Setting 463
Order of Operations 467
Practical Applications 473
Finishing the Process 473
Listening 473
Ongoing Optimization 474
Using Program Material as the Source 474
Audience Presence 475
Temperature and Humidity 476
Stage Leakage 476
Stage Microphone Summation 477
Feedback 478
Multichannel Program Material 478
Show Mic Positions 479
Subwoofers as a Program Channel 480
Final Word 480

Afterword 481
Glossary 482
Bibliography 492
Index 493

Preface

This book is about a journey. On the one hand, the subject is the journey of sound as it travels through a sound system, then through the air, and inevitably to a listener. On the other hand, it is a journey which concerns my own quest to understand the complicated nature of this sound transmission. The body of this text will detail the strictly technical side of things. First, however, I offer you some of the personal side.

I was supposed to build buildings. Unbeknownst to me at the time, this calling was derailed on February 9, 1964 by the appearance of the Beatles on the Ed Sullivan show. Like so many of my generation, this landmark event brought popular music and an electric guitar into my life. I became a great enthusiast of live concerts which I regularly attended throughout my youth at any chance presented. For years, it remained my expectation that I would enter the family construction business. This vision ended on a racetrack in Des Moines, Iowa on June 16, 1974. The experience of hearing the massive sound system at this Grateful Dead concert set my life in a new direction. On that day I made the decision that I was going to work in live concert sound. I wanted to help create this type of experience for others. I would be a mix engineer and my dream was to one day operate the mix console for big shows. I set my sights on preparing for such a career while at Indiana University. This was no simple matter since there was no such thing as a degree in audio. I soon discovered the Independent Learning Program. Under the auspices of that department, I assembled together a mix of relevant courses from different disciplines and graduated with a college level degree in my self-created program of audio engineering.

Figure 0.1 Ticket stub from the June 16, 1974 Grateful Dead concert in Des Moines, Iowa

By 1980, I had a few years of touring experience under my belt and had moved to San Francisco. There I forged relationships with John Meyer, Alexander Yuill-Thornton II (Thorny), and Don Pearson. These would become the key relationships in my professional development. Each of us was destined to stake our reputations on the same piece of equipment: the dual channel FFT analyzer.

I would like to say that I was involved in live concert measurement with a dual channel FFT analyzer from day one, but this is not the case. The process was pioneered by John Meyer on a Saturday night in May of 1984. John took the analyzer, an analog delay line and some gator clips to a Rush concert in Phoenix, Arizona. There he performed the first measurements of a concert sound system using music as the source while the audience was in place. I was not destined to become involved in the project until the following Monday morning.

From that day forward, I have never been involved in a concert or a sound system installation without the use of a dual channel FFT analyzer. Also from that day I have never mixed another show, resetting my vision to the task of helping mix engineers to practice their art. For Don, John, Thorny and many others, the idea of setting up a system without the presence of the FFT analyzer was unthinkable. The bell could not be unrung. From the very beginning, we saw its importance and its practical implications. Our excitement was palpable, with each concert resulting in an exponential growth in knowledge. We all saw it as a break-through at the time and we introduced it to every one who had an open mind to listen. The first product to come from the FFT analysis process was a parametric equalizer. A fortuitous coincidence of timing resulted in my having etched the circuit boards for the equalizer on my back porch over the weekend that John was in Phoenix with Rush. This side project (a bass guitar preamp) for my friend Rob Wenig was already 6 months late, and was destined to be even later. The EQ was immediately pressed into service when John nearly fell over when he saw that it could create the complementary response (in both amplitude and phase) to what he had measured in Phoenix. The CP-10

Figure 0.2 July 14,1984, Grateful Dead, Greek Theater, Berkeley, California. The author with the primordial version of SIM™ (photo Clayton Call)

Figure 0.3 November, 1984 Photo of Luciano Pavarotti, Roger Gans and the author (back row), Drew Serb, Alexander Yuill-Thornton II, and James Locke (front row) (photo Drew Serb)

was born into more controversy than one might imagine. Equalization has always been an emotional "hot button" but the proposition that the equalizer was capable of counteracting the summation properties of the speaker/room interaction was radical enough that we obtained the support of Stanford's Dr. Julius Smith to make sure that the theory would hold up.

The first one outside of our company to really take the concept of in-concert analysis on in the field was Don Pearson, who was then touring as the system engineer for the Grateful Dead. Don and the band immediately saw the benefit and, lacking patience to wait for the development of what would become the Meyer Sound SIM System, obtained their own FFT analyzer and never looked back. Soon thereafter Luciano Pavarotti followed under the guidance of Roger Gans, sound designer in charge of the arena-scale performances given by that artist. We figured it was a matter of months before this became standard operating procedure throughout the industry. We had no idea it would take closer to 20 years! The journey, like that of sound transmission, was far more complex than we ever expected. There were powerful forces lined up against us in various forms: the massive general resistance of the audio community to sound analyzers and the powerful political forces advocating for alternate measurement platforms, to name a few.

In general, the live sound community was massively opposed to what they conceptualized as an analyzer dictating policy to the creative forces involved in the music side of the experience. Most live concert systems of the day lacked complexity beyond piles of speakers with left and right channels. This meant that the process of alignment consisted of little more than equalization. Since all of the system calibration was being carried out at a single location, the mix position, the scientific and artistic positions were weighing in on the exact same question at the same point in space. Endless adversarial debate about what was the "correct" equalization ensued since the tonal balancing of a sound system is, and always has been, an artistic endeavor. It was an absurd construct. Which is better — by ear or by analyzer?

This gave way to a more challenging and interesting direction for us: the quest beyond the mix position. Moving the mic out into the space left us with a terrible dilemma: the new positions revealed conclusively that the one-size-fits-all version of system equalization was utter fantasy. The precision tuning of parametric filters carried out with great care for the mix position had no justification at other locations. The interaction of the miscellaneous parts of the speaker system created a highly variable response throughout the room. The goal for us shifted from finding a perfect equalization to the quest for uniformity over the space.

This would require the subdivision of the sound system into defined and separately adjustable subsystems, each with individual level, equalization and delay capability. The subsystems were then combined into a unified whole. The rock and roll community was resistant to the idea, primarily because it involved turning some of the speakers down in level. The SPL Preservation Society staunchly opposed anything that might detract from the maximum power capability. Uniformity by subdivision was not worth pursuing if it cost power. Without subdivision, the analysis was pretty much stuck at the mix position. If we are not going to change anything, why bother to look further?

There were other genres that were open to the idea. The process required the movement of a microphone around the room and a systematic approach to deconstructing and reconstructing the sound system. We began developing this methodology with the Pavarotti tours. Pavarotti was using approximately ten subsystems. When we moved into the musical theater world with Andrew Bruce, Abe Jacob, Tony Meola, Tom Clark and other such sound designers, our process had to be honed to take on even more complexity. Our emphasis changed from providing a scientifically derived tonal response to instead providing consistency of sound throughout the listening space, leaving the tonal character in the hands of the mix engineer. Our tenure as the "EQ police" was over as our emphasis changed from tonal quality to tonal *equality.* The process was thus transformed into optimization, emphasizing spatial uniformity while encompassing equalization, level setting, delay setting, speaker positioning and a host of verifications on the system. A clear line was drawn between the artistic and the scientific sectors.

In the early days, people assessed the success of a system tuning by punching out the filters of the equalizer. Now, with our more sophisticated process, we could no longer re-enact before and after scenarios. To hear the "before" sound might require repositioning the speakers, finding the polarity reversals, setting new splay angles, resetting level and time delays and finally a series of equalizations for the different subsystems. Finally, the role of the optimization engineer became clear: to ensure that the audience area receives the same sound as the mix position.

In 1987, we introduced the SIM system — the first multichannel **FFT** analysis system designed specifically for sound system optimization (up to 64 channels). It consisted of an analyzer, multiple mics and switchers to access banks of equalizers and delays. All of this was under computer control which also kept a library of data which could be recalled for comparison of up to 16 different positions or scenarios. It thereby became possible to monitor the sound system from multiple locations and see the effects of changes in one part of the system on other areas. It was also possible to make multiple microphone measurements during a performance and to see the effects of the audience presence throughout the space.

This is not to say we were on Easy Street at this point. It was a dizzying amount of information to keep track of. The frequency response was measured in seven separate sections. A single set of data to fully characterize one location at a point in time was an assembly of 63 traces, of which only two could be seen at any one time on the tiny four-inch screen. Comparison of one mic position to another had to be done on a trace-by-trace basis (up to 63 operations). It was like trying to draw a landscape while looking through a periscope.

The multichannel measurement system opened the door toward system subdivision. This approach broke the pop music sound barrier with Japanese sensation Yuming Matsutoya under the guidance of Akio Kawada, Akira Masu and Hiro Tomioka. In the arenas across Japan we proved that the same techniques of level tapering, zoned equalization and a carefully combined system which we had employed for musical theater and Pavarotti were equally applicable to high-power rock music in a touring application.

The introduction of the measurement system as a product was followed by the first training seminar in 1987. It was during this first seminar that for me a seminal moment would occur from an unexpected direction. As I explained the process of placing the mics and subdividing the system for optimization, I was challenged by the very experienced engineer Dave Robb who felt that my mic placement was "arbitrary." In my mind, the selection was anything but arbitrary. However, I could not, at that moment, bring forth any objective criteria with which to refute that assertion. Since that humiliating moment, my quest has been to find a defensible methodology for every decision made in the process of sound system optimization. It is not simply enough to *know* something works, we must *know why* it works. Those optimization methodologies and an accompanying set of methods for sound system design are the foundation of this book. I knew nothing of sound system design when this quest began in 1984. Almost everything I have learned about the design of sound systems comes from the process of their optimization. The process of deconstructing and reconstructing other people's designs gave me the unique ability/perspective to see what aspects of design were universally good, bad or ugly. I am very fortunate to have been exposed to all different types of designs, utilizing many different makes and models of speakers, with all types of program materials and scales. My approach has been to search for the common solutions to these seemingly different situations and to distill them into a repeatable strategy to bring forward with me to the next application.

Beginning with that very first class, with little interruption, I have been optimizing sound systems and teaching anybody who wanted to attend my seminars everything I was learning. Thorny, meanwhile had moved on and founded a company whose principal focus was sound system optimization services using the dual-channel FFT systems. Optimization as a distinct specialty had begun to emerge.

The introduction of SIA-SMAART in 1995 resulted from the collaboration of Thorny and Sam Berkow with important contributions by Jamie Anderson and others in later years. This low-cost alternative brought the dual channel FFT analyzer into the mainstream and made it available to audio professionals at every level. Even so, it took years before our 1984 vision of the FFT analyzer, as standard FOH equipment, would become reality. Unquestionably, that time has finally arrived. The paradigm has reversed to the point where the practice of tuning a system *without* scientific instrumentation would be looked at with as much surprise as was the reverse in the old days.

Since those early days we have steadily marched forward with better tools — better sound systems, better sound design tools and better analyzers. The challenge, however, has never changed. It is unlikely that it will change, since the real challenge falls entirely in the spatial distribution properties of acoustical physics. The speakers we use to fill the room are vastly improved and signal processing capability is beyond anything we dreamed of in those early days. Prediction software is now readily available to illustrate the interaction of speakers, and we have affordable and fast analyzers to provide the on-site data.

And yet we are fighting the very same battle that we have always fought: the creation of a uniform sonic experience for audience members seated everywhere in the venue. It is an utterly insurmountable challenge. It cannot be achieved. There is no perfect system configuration or tuning. The best we can hope for is to approach uniformity. I believe it is far better to be coldly realistic about our prospects. We will have to make decisions that we know will degrade some areas in order to benefit others. We want them to be informed decisions, not arbitrary ones.

This book follows the full transmission path from the console to the listener. That path has gone through remarkable changes along its entire electronic voyage. But once the waveform is transformed into its acoustic form it enters the very same world that Jean Baptiste Fourier found in the eighteenth century and Harry Olson found in the 1940s. Digital, schmigital. Once it leaves the speaker, the waveform is pure analog and at the mercy of the laws of acoustical physics. These unchanging aspects of sound transmission are the focus of 90 per cent of this book.

Let's take a moment to preview the adversary that we face. The primary player is the interaction of speakers with other speakers, and with the room. These interactions are extremely complex on the one hand, and yet can be distilled down to two dominant relationships: relative level and relative phase. The combination of two related sound sources will create a unique spatial distribution of additions and subtractions over the space. The challenge is the fact that each frequency combines differently, creating a unique layout. The frequency range of our sound systems (30 to 18,000 Hz) spans a 600:1 ratio of wavelengths. A single room, from the perspective of spatial distribution over frequency, is like a 600 story skyscraper with a different floor plan at every level. Our job is to find the combination of speakers and room geometry that creates the highest degree of uniformity for those 600 floor plans. The contribution of every speaker element and surface is factored into the spatial distribution. The part that each element plays will be directly in proportion to the amount of energy it brings to the equation at every point in the space. The final result of the combination will be decided by the extent to which there is agreement between the individual phase responses at each location at each frequency. How do we see these floor plans? With an acoustic prediction program we can view the layout of each floor, and compare them and see the differences. This is the viewpoint of a single frequency range analyzed over the entire space. With an acoustic analyzer we get a different view. We see a single spot on each floor from the foundation to the rooftop through a piece of pipe as big around as our finger. This is the viewpoint of a single point in space analyzed over the entire frequency range.

This is a daunting task. But it is comprehensible. This book will provide you with the information required to obtain the x-ray vision it takes to see through the 600 story building from top to bottom, and it can be done without calculus, integral math or differential equations. We let the analyzer and the prediction program do the heavy lifting. Our focus is on how to read x-rays, not on how to build an x-ray machine.

The key to understanding the subject, and a persistent theme of this book, is sound source identity. Every speaker element, no matter how big or small, plays an individual role, *and that solitary identity is never lost.* Solutions are enacted locally on an element-by-element basis. We must learn to recognize the individual parties to every combination, because therein lie the solutions to their complex interaction.

This is not a mystery novel, so there is no need to hide the conclusion until the last pages. The key to spatial uniformity is isolation of the interactions. If two speaker elements are to combine into a consistent shape over frequency they must have distinct zones of coverage. If they are operating at or near the same levels they must have some amount of angular isolation. The separation may be minute, but their on-axis patterns must never cross. If angular isolation is not to be found, if the patterns cross, then one of the elements must yield the floor, by a reduction in level. The interaction of speakers to the room is similar to the interaction of speakers with other speakers. Those surfaces that return energy back toward our speakers will be the greatest concern. The strength of the inward reflections will be inversely proportional to our spatial uniformity.

There is no single design for a single space. There are alternate approaches and each involves tradeoffs in terms of spatial uniformity and other key criteria. There are, however, certain design directions that keep open the possibility of spatial uniformity and others that render such hopes statistically impossible. A major thrust of the text will be devoted to defining the speaker configurations that have the potential for spatial uniformity.

Once designed and installed, the system must be optimized. If the design has kept open the door for spatial uniformity, it will be our task to navigate the system through that door. There is no single optimization solution for a given design in a space, but once again there are only a limited number of approaches that we can hope will bring us spatial uniformity. The key to optimization is the knowledge of the locations of the decisive events in the battle for spatial uniformity. The interactions of speakers and rooms follow a consistent set of progressions of effect over the space. The layering of these effects over each other provides the ultimate challenge, but there is *nothing* random about this family of interactions. It is logical and learnable. Our measurement mic locations will be the places where we shall view the progressions through the hundreds of building layers and make the adjustments that will affect all locations in the room. Since we have only limited time and resources we must know our exact location in the context of the interaction progressions to discern the meaning of the measured data.

We have often seen the work of archeologists where a complete rendering of a dinosaur is created from a small sampling of bone fragments. Their conclusions are based entirely on contextual clues gathered from the knowledge of the standard progressions of animal anatomy. If such progressions were random, there would be nothing short of a 100 per cent fossil record that could provide answers. From a statistical point of view, even with hundreds of mic positions, we will never be able to view more than a few tiny fragments of our speaker system's anatomy in the room. We must make every measurement location count toward the collection of the data we need to see the big picture. This requires advance knowledge of the progression milestones so that we can view a response in the context of what is expected at the given location. As we shall see, there is almost nothing that can be concluded for our application from a single location. Actions that will benefit more than a single point in space absolutely require contextual information about where the given space fits in to the overall spatial distribution.

Defined speakers, in a defined design configuration, with defined strategies for optimization, is what this book is about. This book is not intended to be a duplication of the general audio resource texts. Such books are available in abundance and no effort is made here to encompass the width and breadth of the complete audio picture. My hope is to provide a unique perspective that has not been told before, in a manner that is accessible to the audio professionals interested in a deeper understanding of the behavior of sound systems in the practical world.

There are a few points that I wish to address before we begin. The most notable is the fact that the physical realities of loudspeaker construction, manufacture and installation are almost entirely absent. Instead they are described only in terms of their acoustic performance properties in a space. Several varieties of speakers are described that serve as representative examples of speaker performance. These performance attributes serve as the framework of the discussion. The means by which manufacturers can create physical systems which meet these general criteria are not within the scope of this book. This is also true of electronic devices. Everything is weightless, colorless and odorless here. It is the common sound transmission characteristics of such devices that are the focus, not the unique features of one model or another.

The second item concerns the approach to particular types of program material such as popular music, musical theater, or religious services, and their respective venues such as arenas, concert halls, showrooms or houses of worship. The focus here is the *shape* of the sound coverage, the scale of which can be adjusted to fit the size of the venue at the appropriate sound level for the given program material. It is the venue and the program material taken together that create an application. The laws of physics are no different for any of these applications and the program material and venues are so interchangeable that attempts to characterize them in this way would require endless iterations. After all, the modern-day house of worship is just as likely to feature popular music in an arena setting as it is to have speech and chant in a reverberant cathedral of stone.

The third notable aspect is that there are a substantial number of unique terminologies found here and in some cases, modification of standard terminologies that have been in general use. In most cases the conceptual framework is unique and no current standard expressions were found. The very young field of sound system optimization has yet to develop consistent methods or a lexicon of expressions for the processes shown here. In the case of some of these terms, most notably the word "crossover," there are compelling reasons to modify the existing usage, which will be revealed in the body of the text.

The book is divided into three sections. The first section, "Sound Systems," explores the behavior of sound transmission systems, human hearing reception and speaker interaction. The goal of this section is a comprehensive understanding of the path the signal will take, the hazards it will encounter along the way and how the end product will be perceived upon arrival at its destination. The second section, "Design," applies the properties of the first section to the creation of a sound system design. The goals here are a comprehensive understanding of the tools and techniques required to generate a design that will create a successful transmission/reception model. The final section is "Optimization." This concerns the measurement of the designed and installed system, its verification and calibration in the space.

This has never been a solitary journey. There are many who have contributed to this emerging field and who share the common goals and interests which are the subject of this book. At the outset of this project I solicited various members of the community to share their perspectives in their own words. Their voices can be heard through the course of the text. In the future I hope that you will add your voice to this field of study.

Acknowledgements

The real development of this book spans more than twenty years in the field of sound system optimization. Were it not for the discoveries and support of John and Helen Meyer, I would have never become involved in this field. They have committed substantial monetary and personnel resources which have directly helped the ongoing research and development leading up to this writing. In addition, I would like to acknowledge the contribution of every client who gave me the opportunity to perform my experiments on their sound systems. Each of these experiences yielded an education that could not be duplicated elsewhere. In particular I would like to thank David Andrews Andrew Bruce, John Cardenale, Tom Clark, Mike Cooper, Jonathan Deans, Francois Desjardin, T. C. Furlong, Roger Gans, Scott Gledhill, Andrew Hope, Abe Jacob, Akio Kawada, Tony Meola, Frank Pimiskern, Bill Piatt, Pete Savel, Mike Shannon, Rod Sintow, Bob Snelgrove and Tom Young, all of whom have given me multiple opportunities through the years to refine the methods described here.

I would also like to thank everyone who has attended my seminars, as the real-time feedback in that context provides a constant intellectual challenge and stimulation for me. My fellow instructors in this field, past and present, have contributed much collaborative effort through discussion and the sharing of ideas. Notable among these are Jamie Anderson, Sam Berkow, Jim Cousins, Mauricio Ramirez and Hiro Tomioka.

I am also grateful to my many compatriots in the field who have shared their experiences with me. Some of them contributed their personal perspectives, which are salted throughout this volume.

In the course of writing this book a number of people provided immeasurably important support and direction. It was truly an endurance test for all of us but every step of the way I had someone to give me feedback and reality checks. At their urging I pushed the envelope beyond my previous experience and opened up avenues of related knowledge that help to link our young field to the previously established disciplines. Many of the graphics in this book contain data from sources that I would like to note here. The following figures contain data that was acquired while I worked at Meyer Sound and their permission for publication is gratefully acknowledged: Figures 1.15, 1.25, 2.35-37, 9.6, 9.12-16, 10.8, 10.21 and 10.27. The data presented in Figures 1.2, 1.7, 1.8 and 8.18 were created using the calculations found in files originated by the "Master of Excel", Mauricio Ramirez. The 3-D wraparound graphics (Figure 8.14) were adapted from the animations created by Greg Linhares.

Thanks go to all of the people who aided in the process of bringing this book to physical reality such as my editor Catharine Steers, Margaret Denley, Lisa Jones and Stephanie Barrett at Elsevier. Additional thanks go to Margo Crouppen for her support throughout the entire publishing process.

The following people aided the effort by proofing, submitting photos or other related items, all of which were immensely helpful and for which I am deeply grateful: Jamie Anderson, Justin Baird, Sam Berkow, David Clark, Francois Desjardin, Larry Elliott, Josh Evans, John Huntington, Luke Jenks, Dave Lawler, Greg Linhares, Mauricio Ramirez, Tom Young, and Alexander (Thorny) Yuill-Thornton.

Finally, my wife Merridith took on the enormous burdens that resulted from my extremely near-sighted focus on these issues over the course of a full year. She was my agent, manager, copy editor, proofreader and cheerleader.

Section page images, from left to right: Section 1—Francois Desjardin, Josh Evans, Josh Evans, Author, Mauricio Ramirez; Section 2—Author, Author, Author, Mauricio Ramirez, Author; Section 3—Bob Maske, Bob Hodas, Author, Kazayuki Kado, Miguel Lourtie.

Front cover photos, from left to right: TC Furlong, Author, Josh Evans, Miguel Lourtie.

Section 1:
Sound Systems

1

Transmission

transmission *n.* *transmitting or being transmitted; broadcast program*

transmit *v.t.* **1.** *pass on, hand on, transfer, communicate.* **2.** *allow to pass through, be a medium for, serve to communicate (heat, light, sound, electricity, emotion, signal, news)*

Concise Oxford Dictionary

Transmission Goals

Transmission is the conveyance of a waveform from one place to another. The quality of the transmission is judged by how accurately it tracks the original waveform. We capture the original acoustical or electronic waveform, with the intent of eventually reconstituting it into an acoustic waveform for delivery to our ears. The odds are against us. In fact there is absolutely zero probability of success. The best we can hope for is to *minimize* the distortion of the waveform, i.e. damage control. That is the primary goal of all efforts described in this book. This may sound dispiriting, but it is best to begin with a realistic assessment of the possibilities. Our ultimate goal is one that can be approached, but never reached. There will be large numbers of decisions ahead, and they will hinge primarily on which direction provides the least damage to the waveform. There are precious few avenues that will provide none, and often the decision will be a very fine line.

Our main study of the transmission path will look at three modes of transmission: line level electronic, speaker level electronic and acoustic. If any link in the transmission chain fails, our mission fails. By far the most vulnerable link in the chain is the final acoustical journey from the speaker to the listener. This path is fraught with powerful adversaries in the form of copies of our original signal (namely reflections and arrivals from the other speakers in our system), which will distort our waveform unless they are exact copies and exactly in time. We will begin with a discussion of the properties of transmission that are common to all parts of the signal path.

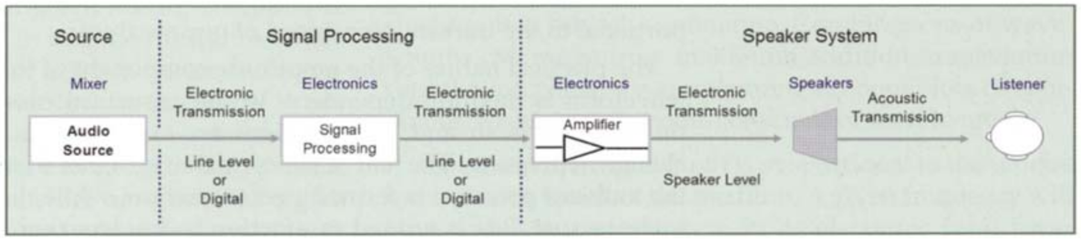


Figure 1.1 Transmission flow from the signal source to the listener

Audio Transmission Defined

An audio signal is constant change: the motion of molecules and electrons transferring energy away from a vibrating source. When the audio signal stops changing it ceases to exist as audio. As audio signals propagate outwards, the molecules and electrons are displaced forward and back but never actually go anywhere. They always return right back to their origin. The extent of the change is the **amplitude,** also referred to as **magnitude.** A single round trip from origin and back is a **cycle.** The round trip takes time. That length of time is the **period** and is given in seconds, or for practical reasons milliseconds (ms). The reciprocal of the period is the **frequency.** This is the number of cycles completed per second and is given in hertz (Hz). The round trip is continuous with no designated beginning or end. The cycle can begin anywhere on the trip and is completed upon our return to the same position. The radial nature of the round trip requires us to find a means of expressing our location around the circle. This parameter is termed the **phase** of the signal. The values are expressed in degrees, ranging from 0 degrees (point of origin) to 360 degrees (a complete round trip). The half-cycle point in the phase journey, 180 degrees, will be of particular interest to us as we move forward.

All transmission requires a **medium,** i.e. the entity through which it passes from point to point, made of molecules or electrons. In our case the primary media are wire (electronic) and air (acoustic), but there are interim media as well such as magnetic and mechanical. The process of transferring the audio energy between media is known as **transduction.** The physical distance required to complete a cycle in a particular medium is the **wavelength** and is expressed in some form of length, typically meters or feet. The size of the wavelength for a given frequency is proportional to the transmission speed of our medium.

The physical nature of the amplitude component of the waveform is medium-dependent. In the acoustical case, the medium is air and the vibrations are expressed as a change in pressure. The half of the cycle that is higher than the ambient pressure is termed **pressurization,** while the low-pressure side is termed **rarefaction.** A loudspeaker's forward motion into the air creates pressurization and its rearward movement away from the air creates rarefaction.

The movement of the speaker cones does not push air across the room. The air is moved forward and then pulled right back to where it was. The transmission passes *through* the medium, an important distinction. Multiple transmissions can pass through the medium simultaneously.

For electronic signals, the electrical pressure change is expressed as **voltage.** Positive and negative pressures are expressed simply as positive and negative voltage. This movement is also termed **alternating current** (AC) since it alternates above and below the ambient voltage known as **direct current** (DC).

It is critical to our efforts to have a thorough understanding of the relationship of frequency, period and wavelength. The relationship of these three parameters plays a major part in our design and optimization strategies.

Time and Frequency

Let's start with a simple tone, called a sine wave, and the relationship of frequency (F) and period (T):

$$T = 1/F \text{ and } F = 1/T$$

where T is the time period of a single cycle in seconds and F is the number of cycles per second (Hz).

To illustrate this point we will use a convenient frequency and delay for clarity: 1000 Hz (or 1 kHz) and 1/1000th of a second (or 1 ms).

If we know the frequency we can solve for time. If we know time we can solve for frequency. Therefore

$$\begin{aligned} F = 1/T \quad & 1000\,\mathrm{Hz} \Leftrightarrow 1/1000\,\mathrm{s} \\ & 1000\,\mathrm{Hz} \Leftrightarrow 0.001\,\mathrm{s} \\ & 1000\,\mathrm{Hz} \Leftrightarrow 1\,\mathrm{ms} \end{aligned}$$

$$\begin{aligned} T = 1/F \quad & 0.001\,\mathrm{s} \Leftrightarrow 1/1000\,\mathrm{Hz} \\ & 1\,\mathrm{ms} \Leftrightarrow 1/1000\,\mathrm{Hz} \end{aligned}$$

For the bulk of this text we will abbreviate the time period to the term "time" to connote the time period of a particular frequency.

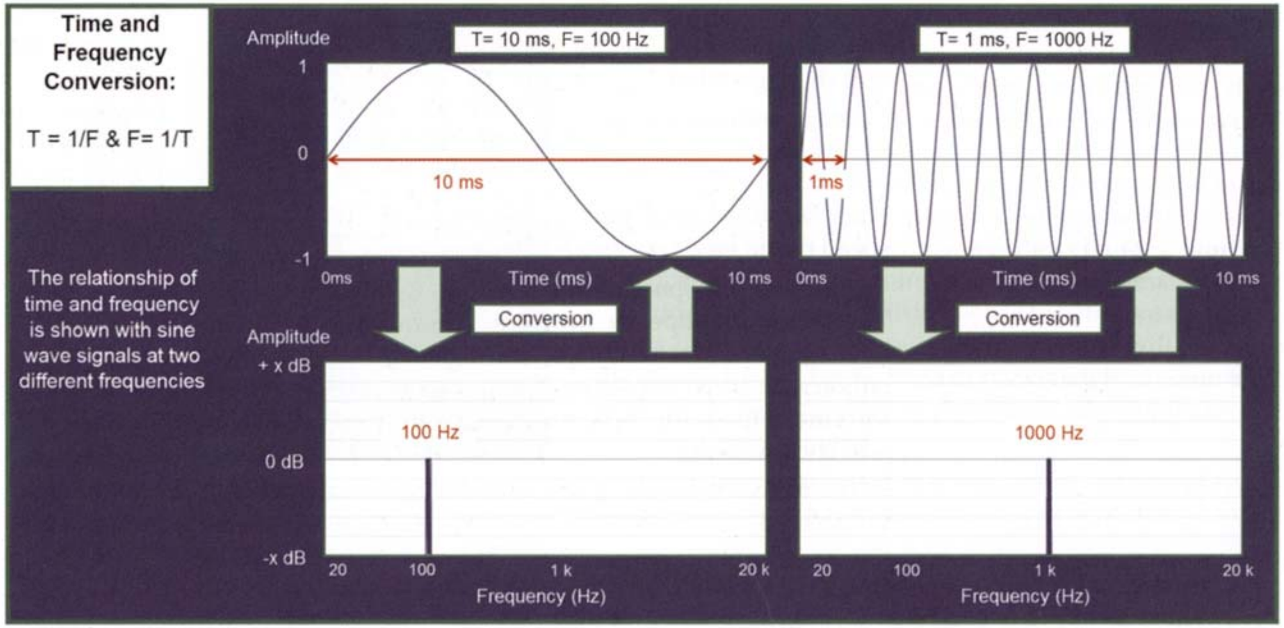


Figure 1,2 Amplitude vs. time converted to amplitude vs. frequency

Frequency is the best-known parameter since it is closely related to the musical term "pitch." Most audio engineers relate first in musical terms since few of us got into this business because of a lifelong fascination with acoustical physics. We must go beyond frequency/pitch, however, since our job is to "tune" the sound system, not tune the musical instruments. In this world we must make an ever-present three-way link between frequency, period and wavelength. The frequency 1 kHz exists only with its reciprocal sister 1 ms. This is not medium-dependent, nor temperature-dependent, nor is it waiting upon a standards committee ruling. This is one of audio's few undisputed absolutes. If the audio is traveling in a wire, those two parameters will be largely sufficient for our discussions. If it is in the air we will need to add the third dimension: wavelength. A 1 kHz signal only exists in the air as a wavelength about as long as the distance from our elbow to our fist. All behavior at 1 kHz will be governed by the physical reality of its time period and its wavelength. The first rule of optimization is to *never* consider an acoustical signal without consideration of all three parameters!

Wavelength

Why it is that we should be concerned about wavelength? After all, there are no acoustical analyzers that show this on their readout. There are no signal processing devices that depend on this for adjustment. There are some applications where we can be blissfully ignorant of wavelength, for example: when we use a single loudspeaker in a reflection-free environment. For all other applications wavelength is not simply relevant: it is decisive. Wavelength is the critical parameter in acoustic summation. The combination of signals at a given frequency is governed by the number of wavelengths that separate them. There is a lot at stake here, as evidenced by the fact that Chapter 2 is dedicated exclusively to this subject: summation. Combinations of wavelengths can range from maximum addition to maximum cancellation. Since we are planning on doing lots of combining, we had best become conscious of wavelength.

The size of the wavelength is proportional to the unique transmission speed of the medium. A given frequency will have a different wavelength in its electronic form (over

500,000X larger) than its acoustic version. If the medium is changed, its transmission speed and all the wavelengths will change with it.

The wavelength formula is

$$L = c/F$$

where L is the wavelength in meters, c is the transmission speed of the medium, and F is the frequency (Hz).

Transmission speed through air is among the slowest. Water is a far superior medium in terms of speed and high-frequency response; however, the hazards of electrocution and drowning make this an unpopular sound reinforcement medium (synchronized swimming aside). We will stick with air.

The formulas for the speed of sound in air are as shown in Table 1.1.

Plain language	Imperial/American measurement	Metric measurement
Speed of sound in air at 0°	1052 feet/second	331.4 meters/second
+ Adjustment for ambient air temperature	+ (1.1 × T) Temperature T in °F	+ (0.607 × T) Temperature T in °C
= Speed of sound at ambient air temperature	= c feet/second	= c meters/second

For example at 22 C:

$$c = (331.4 + 0.607 \times 22) \text{ meters/second}$$
$$c = 344.75 \text{ meters/second}$$

Now that we can solve for the transmission speed we can determine the wavelength for a given frequency/time period:

$$L = c/F$$

Where L is the wavelength in meters, c is the transmission speed of sound, and F is the frequency (Hz).

The audible frequency range given in most books is 20 Hz to 20 kHz. Few loudspeakers are able to reproduce

Wavelength Reference Chart

Frequency (Hz)	Period (ms)	Wavelength (Room temp) (m)	(ft)	Comparable size
20	50.00	17.24	56.56	
25	40.00	13.79	45.07	Intermodal shipping container
32	31.75	10.94	35.77	
40	25.00	8.62	28.17	Band gear truck length
50	20.00	6.90	22.54	1/2 size intermodal container
63	15.87	5.47	17.89	Gas guzzling SUV length
80	12.50	4.31	14.09	Full Size car length
100	10.00	3.45	11.27	Compact car length
125	8.00	2.76	9.01	Too wide for the truck
160	6.25	2.15	7.04	Shaquille O'Neal
200	5.00	1.72	5.63	Average height
250	4.00	1.38	4.51	Shoulder Height
315	3.17	1.09	3.58	
400	2.50	0.86	2.82	
500	2.00	0.69	2.25	Arm's length
630	1.59	0.55	1.79	
800	1.25	0.43	1.41	
1,000	1.00	0.34	1.13	Elbow to fist
1,250	0.80	0.28	0.90	Man's foot
1,600	0.63	0.22	0.70	Woman's foot
2,000	0.50	0.17	0.56	Eight fingers
2,500	0.40	0.14	0.45	
3,150	0.32	0.11	0.36	CD/DVD
4,000	0.25	0.086	0.28	Four fingers
5,000	0.20	0.069	0.23	
6,300	0.16	0.055	0.18	
8,000	0.13	0.043	0.14	Two fingers
10,000	0.10	0.034	0.11	
12,500	0.08	0.028	0.09	
16,000	0.06	0.022	0.07	One finger
20,000	0.05	0.017	0.06	

Figure 1,3 Chart of frequency, period and wavelength (at room temperature) for standard 1/3rd octave frequencies

the 20 Hz or 20 kHz extremes at a power level sufficient to play a significant role. It is more useful to limit the discussion to those frequencies we are likely to encounter in the wild: 31 Hz (the low B note on a five-string bass) up to 18 kHz. The wavelengths within this band fall into a size range of between the width of a finger and a standard intermodal shipping container. The largest wavelengths are about 600 times larger than the smallest.

Temperature Effects

As we saw previously, the speed of sound in air is slightly temperature-dependent. As the ambient temperature rises,

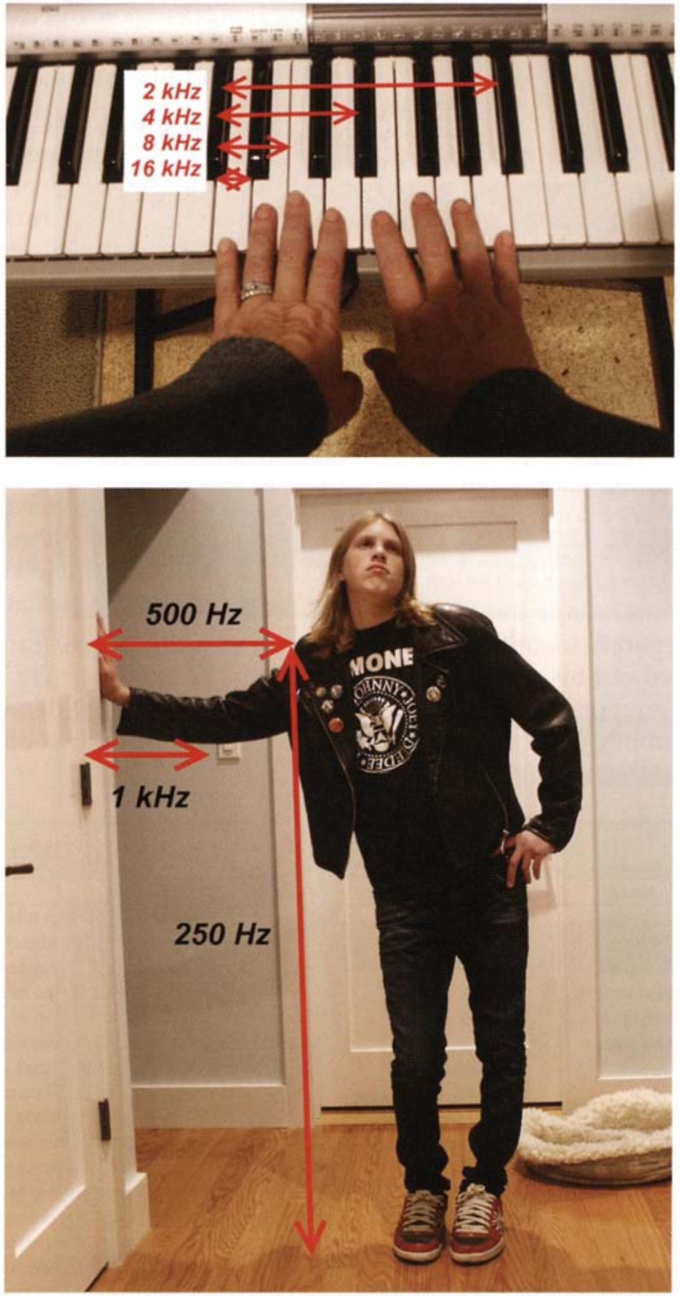


Figure 1.4 A handy reference for short wavelengths

sound speed increases and therefore the wavelengths expand. This behavior may slightly affect the response of our systems over the duration of a performance, since the temperature is subject to change even in the most controlled environments. However, although it is often given substantial attention, this factor is not a major one in the big scheme of things. A poorly designed system is not likely to find itself rescued by weather changes. Nor is it practical to provide ongoing environmental analysis over the widespread areas of an audience to compensate for the drafts in the room. For our discussion, we will consider the speed of sound to be fixed approximately at room temperature. When variable temperature is relevant to the discussion it will be noted.

The relationship between temperature and sound speed can be approximated as follows: A1 per cent change in the

Temperature Effects on Speed of Sound

°C	C Sound speed	Δ	ΔC	°F	C Sound speed	Δ
Metric				English		
Temp Deg (C)	Speed m/sec)	Temp change Deg (C)	Sound speed change (%)	Temp Deg (F)	Speed (ft/sec)	Temp change Deg (F)
0.0	331.4	-22.2	3.9%	32	1087	-40
1.1	332.1	-21.1	3.7%	34	1089	-38
2.2	332.7	-20.0	3.5%	36	1092	-36
3.3	333.4	-18.9	3.3%	38	1094	-34
4.4	334.1	-17.8	3.1%	40	1096	-32
5.6	334.8	-16.7	2.9%	42	1098	-30
6.7	335.4	-15.6	2.7%	44	1100	-28
7.8	336.1	-14.4	2.5%	46	1103	-26
8.9	336.8	-13.3	2.3%	48	1105	-24
10.0	337.5	-12.2	2.2%	50	1107	-22
11.1	338.1	-11.1	2.0%	52	1109	-20
12.2	338.8	-10.0	1.8%	54	1111	-18
13.3	339.5	-8.9	1.6%	56	1114	-16
14.4	340.2	-7.8	1.4%	58	1116	-14
15.6	340.8	-6.7	1.2%	60	1118	-12
16.7	341.5	-5.6	1.0%	62	1120	-10
17.8	342.2	-4.4	0.8%	64	1122	-8
18.9	342.9	-3.3	0.6%	66	1125	-6
20.0	343.5	-2.2	0.4%	68	1127	-4
21.1	344.2	-1.1	0.2%	70	1129	-2
22.2	344.9	0.0	0.0%	72	1131	0
23.3	345.6	1.1	-0.2%	74	1133	2
24.4	346.2	2.2	-0.4%	76	1136	4
25.6	346.9	3.3	-0.6%	78	1138	6
26.7	347.6	4.4	-0.8%	80	1140	8
27.8	348.3	5.6	-1.0%	82	1142	10
28.9	348.9	6.7	-1.2%	84	1144	12
30.0	349.6	7.8	-1.4%	86	1147	14
31.1	350.3	8.9	-1.6%	88	1149	16
32.2	351.0	10.0	-1.8%	90	1151	18
33.3	351.6	11.1	-2.0%	92	1153	20
34.4	352.3	12.2	-2.2%	94	1155	22
35.6	353.0	13.3	-2.3%	96	1158	24
36.7	353.7	14.4	-2.5%	98	1160	26
37.8	354.3	15.6	-2.7%	100	1162	28

Figure 1.5 Chart of speed of sound, period and wavelength at different temperatures

speed of sound occurs with either a 5°C or 10°F change in temperature.

The Waveform

There is no limit to the complexity of the audio signal. Waves at multiple frequencies may be simultaneously combined to make a new and unique signal that is the summation of the contributing signals. This composite signal is the waveform, containing an unlimited mix of audio frequencies with variable amplitude and phase relationships. The complex shape of the waveform depends upon the components that make it up and varies constantly as they do. A key parameter is how the frequency of each contributing signal affects the combined waveform. If two signals of different frequencies are added, the waveform will carry the shape of both waveforms independently. The higher frequency will be added on to the shape of the lower-frequency waveform. The phase of the individual frequencies will affect the overall shape but the different frequencies maintain their separate identities. These frequencies can later be separated out by a filter (as in your ear) and heard as separate sounds. When two signals of the same frequency are combined, a new and unique signal is created and cannot be filtered apart. In this case, the phase relationship will have a decisive effect upon the nature of the combined waveform.

Audio waveforms exist in many forms. The components listed above are present regardless of the form. All forms track the original audio waveform in the same basic way. Analog waveform types include electronic, magnetic, mechanical, optical, and acoustical. Digital audio signals are typically electronic, magnetic or optical, but the mechanics of the digital data transfer are not critical here. It could be punch cards as long as we can move them fast enough to read the data. Each medium tracks the waveform in different forms of energy, suitable for the particulars of that transmission mode, complete with its own vulnerabilities and limitations. Digital audio is most easily understood when viewed as a mathematical rendering of the waveform. For these discussions, this is no different from analog, which in any of its resident energy forms can be quantified mathematically.

The audio signal can be visualized in three different forms as shown in Fig. 1.9. A single cycle is broken into 4 quadrants of 90 degrees each. This motion form illustrates the movement of the signal from a point of rest to

Waveform Terminology Reference

Name	Symbol	Units	Description
Cycle			One round trip of an audio signal's journey
Period	T	(Seconds)	Time required to complete one cycle
Frequency	F	(Hz)	The number of cycles completed in one second
Wavelength	λ		The distance traveled to complete one cycle.
Phase	Φ	(Degrees)	The radial expression of the percentage of cycle completion
Amplitude	V	Volts	The level of the audio signal as an absolute electrical value
	Vrms	Volts	The root/mean/squared voltage. This represents the DC voltage equivalent for AC signals above and below the zero crossing point.
	Vpeak	Volts	The highest voltage above (or below) the zero crossing point.
	Vp-p	Volts	The voltage span between the highest positive and negative peaks
	dB	Decibels	The level ratio between two signals
	dBV	Decibels	The level ratio relative to 1.0 volts RMS
	dBu	Decibels	The level ratio relative to 0.775 volts RMS
	dB SPL	Decibels	The acoustical level ratio relative to the threshold of hearing
Crest Factor		Decibels	The ratio between the peak and rms values of the signal
Waveform			The audio signal expressed as amplitude over time

Figure **1.6** A reference chart of some of the common terms used to describe and quantify an audio waveform

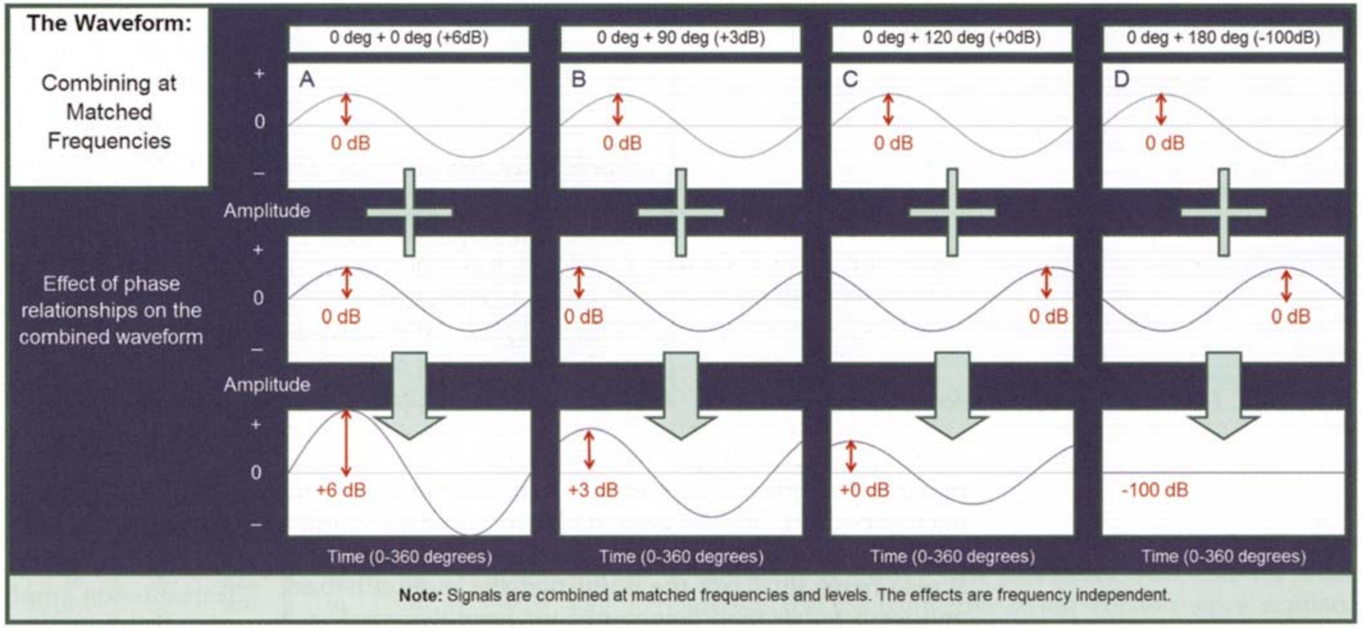


Figure 1.7 Combination of waveforms of the same frequency at the same level with different phase relationships (A) 0 degrees relative phase combines +6 dB amplitude, (B) 90 degrees relative phase combines to +3 dB amplitude, (C) 120 degrees relative phase combines to +0 dB, (D) 180 degrees relative phase cancels

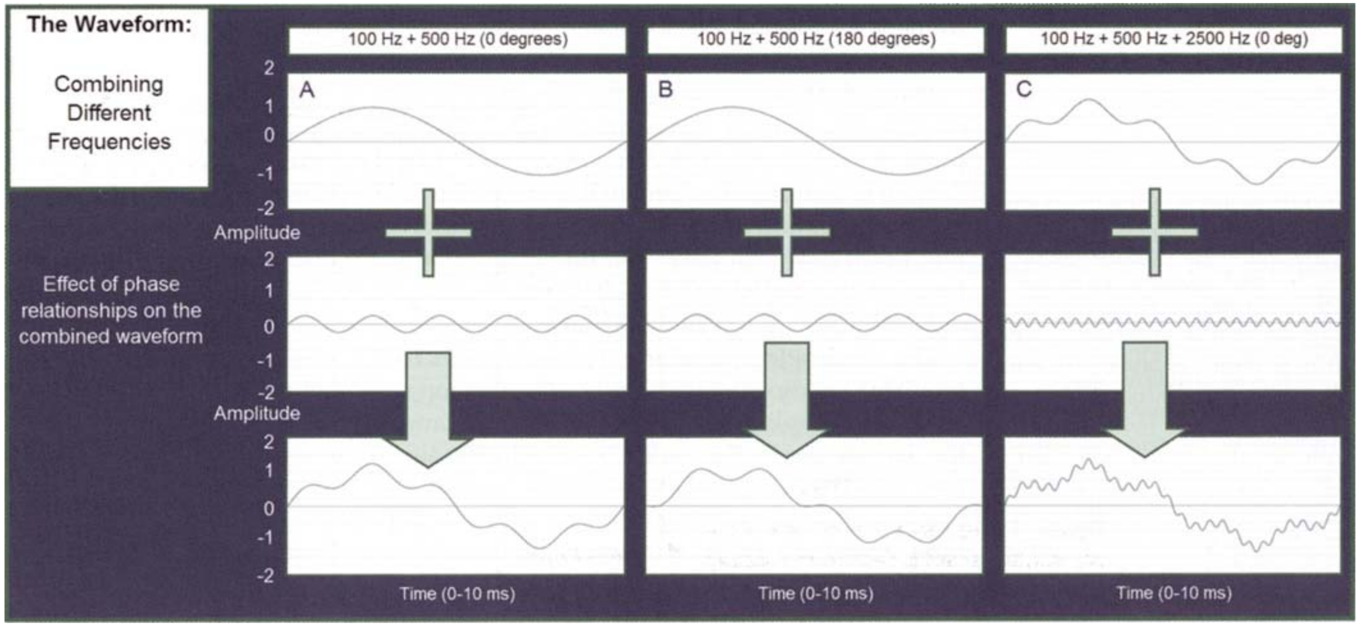


Figure 1.8 Combination of waveforms of different frequencies with different levels and phase relationships. (A) Second frequency is 5x higher and 12dB down in level from the first. Phase relationship is 0 degrees. Note that both frequencies can be seen in the combined waveform. (B) Same as (A) but with relative phase relationship at 180 degrees. Note that there is no cancellation in the combined waveform. The orientation of the high-frequency trace has moved but the low-frequency orientation is unchanged. (C) Combined waveform of (A) with third frequency added. The third frequency is 25x the lowest frequency and 18 dB down in level. The phase relationship is matched for all frequencies. Note that all three frequencies can be distinguished in the waveform

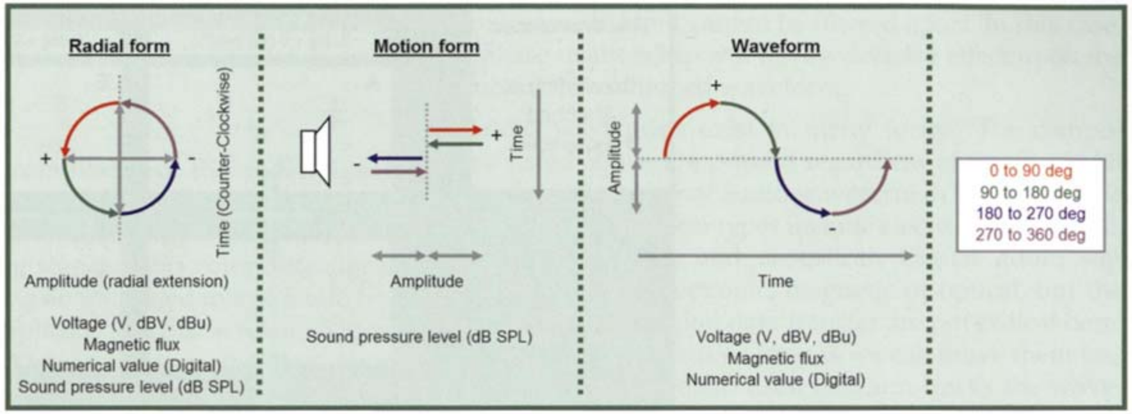


Figure 1.9 Three representations of the audio wave form

maximum amplitude in both directions and finally returning to the origin. This is representative of the particle motion in air when energized by a sound source such as a speaker. It also helps illustrate the point that the motion is back and forth rather than going outward from the speaker. A speaker should not be confused with a blower. The maximum displacement is found at the 90 and 270 degree points in the cycle. As the amplitude increases the displacement from the equilibrium point becomes larger. As the frequency rises, the time elapsed to complete the cycle decreases.

The radial form represents the signal as spinning in a circle. The waveform origin point corresponds to the starting point phase value, which could be any point on the circle. A cycle is completed when we have returned to the phase value of the point of origin. This representation shows the role that phase will play. The difference in relative positions on this radial chart of any two sound sources will determine how the systems will react when combined.

The sinusoidal waveform representation is the most familiar to audio engineers and can be seen on any oscilloscope. The amplitude value is tracked over time and traces the waveform in the order in which the signal passes through. This is representative of the motion over time of transducers and changing electrical values of voltage over time. Analog to digital converters capture this waveform and create a mathematical valuation of the amplitude vs. time waveform.

*Perspectives: I have tried to bring logic, reasoning, and physics to my audio problem applications. What I have found is that, when the cause of an event **is** attributed to "magic," what this really means is that we do not have all the data necessary to understand the problem, or we do not have an understanding of the forces **involved in** producing the observed phenomena.*

Dave Revel

Transmission Quantified

Decibels

Transmission amplitudes, also known as **levels,** are most commonly expressed in decibels (dB), a unit that describes a ratio between two measures. The decibel is a logarithmic scaling system used to describe ratios with a very large range of values. Using the decibel has the added benefit of closely matching our perception of sound levels, which is generally logarithmic. There are various dB scales that apply to transmission. Because decibels are based on ratios they are always a relative scale. The question is: relative to what? In some cases we want to compare a level relative to a fixed standard. Because audio is in constant change it is also useful to have a purely relative scale that compares two unknown signals. An example of the latter type is the comparison of the output level of a device relative to its input by means of a ratio. This ratio is known as the **gain** of the device. The ratio between the input and output can be quantified even though a signal such as music is constantly changing. If the same voltage appears at the input and output, the ratio of input to output is 1, also known as unity **gain,** or 0 dB. If the voltage at the output is greater than the input, the gain value is greater than 1, and expressed in dB is positive. If the output is less than the input the gain is less than 1 and in dB is a negative number, in other words, it's a loss. The actual value

Decibel Ratio Reference							
20 x log (Level$_1$/Level$_2$)				10 x log (Power$_1$/Power$_2$)			
Pressure (SPL)							
Voltage (V)				Power (P)			
Current (I)							
Value (dB)	Ratio (L$_1$/L$_2$)	Value (dB)	Ratio (L$_1$/L$_2$)	Value (dB)	Ratio (P$_1$/P$_2$)	Value (dB)	Ratio (P$_1$/P$_2$)
0.0	1.00	0.0	1.00	0.0	1.00	0.0	1.00
1.0	1.12	-1.0	0.89	0.5	1.12	-0.5	0.89
2.0	1.26	-2.0	0.79	1.0	1.26	-1.0	0.79
3.0	1.41	-3.0	0.71	1.5	1.41	-1.5	0.71
4.0	1.59	-4.0	0.63	2.0	1.59	-2.0	0.63
5.0	1.78	-5.0	0.56	2.5	1.78	-2.5	0.56
6.0	2.00	-6.0	0.50	3.0	2.00	-3.0	0.50
7.0	2.24	-7.0	0.45	3.5	2.24	-3.5	0.45
8.0	2.51	-8.0	0.40	4.0	2.51	-4.0	0.40
9.0	2.82	-9.0	0.35	4.5	2.82	-4.5	0.35
10	3.16	-10	0.32	5.0	3.16	-5.0	0.32
12	4.00	-12	0.25	6.0	4.00	-6.0	0.25
14	5.00	-14	0.20	7.0	5.00	-7.0	0.20
15	5.63	-15	0.18	7.5	5.66	-7.5	0.18
18	8.00	-18	0.13	9.0	8.00	-9.0	0.13
20	10	-20	0.10	10	10	-10	0.10
26	20	-26	0.05	13	20	-13	0.05
32	40	-32	0.03	16	40	-16	0.03
38	80	-38	0.01	19	80	-19	0.01
40	100	40	0.01	20	100	20	0.01
60	1,000	60	0.00	30	1,000	30	0.00
80	10,000	80	0.00	40	10,000	40	0.00
100	100,000	100	100,000	50	100,000	50	0.00

Figure 1.10 The ratio of output to input can be converted to dB with this easy reference chart. To compare a given level to a standard one, then Level, is the given level and Level 2 is the standard. To derive the gain of a device, then Level 1 is the output level and Level 2 is the input level. Likewise the power gain can be found in the same manner by substituting the power parameters

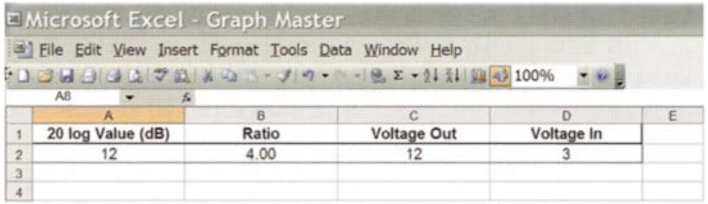
at the input or output is unimportant. It is the change in level between them that is reflected by the dB gain value.

There are two types of log formulas applicable in audio:

$$\text{Relative level (dB)} = 20 \times \log_{10} \frac{\text{Level}_1}{\text{Level}_2}$$

$$\text{Relative power (dB)} = 10 \times \log_{10} \frac{\text{Power}_1}{\text{Power}_2}$$

Power-related equations use the 10 log version, while pressure (SPL) and voltage related equations use the 20 log version. It is important that the proper formula be



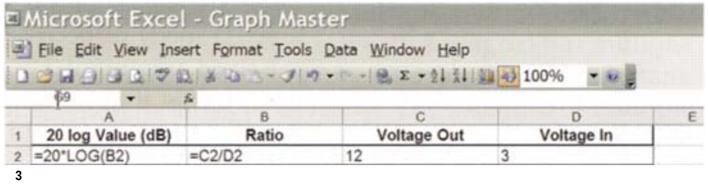


Figure 1,11 Microsoft Excel log formula reference

used since a doubling of voltage is a change of 6 dB while a doubling of power is a change of 3 dB. For the most part we will be using the 20 log version since acoustic pressure (dB SPL) and voltage are the primary drivers of our decision-making. Figure 1.10 provides a reference chart to relate the ratios of values to their decibel equivalents.

Here is a handy tip for Microsoft Excel users. Figure 1.11 shows the formula format for letting Excel do the log calculations for us.

The Electronic Decibel: dBV and dBu

Electronic transmission utilizes the decibel scale to characterize the voltage levels. The decibel scale is preferred by operators over the linear scale for its relative ease of expression. Expressed linearly we would find ourselves referring to the signal in microvolts, millivolts and volts with various sets of number values and ranges. Such scaling makes it difficult to track a variable signal such as music. If we wanted to double the signal level we would have to first know the voltage of the original signal and then compute its doubling. With dynamically changing signals such as music, the level at any moment is in flux, making such calculations impractical. The decibel scale provides a relative change value independent of the absolute value. Hence the desire to double the level can be achieved by a change of 6 dB, regardless of the original value. We can also relate the decibel to a fixed standard which we designate as "OdB." Levels are indicated by their relative value above (+dB) or below (-dB) this standard. This would be simplest, of course, if there were a single standard, but tradition in our industry is to pick several. **dBV** and **dBu** are the most common currently. These are referenced to different values of 1.0 volt and 0.775 volt (1mW across a 600 ohms load) respectively. The difference between these is a fixed amount of 2.21 dB. Note: For ease of use we will use dBV as the standard in this text. Those who prefer the dBu standard should apply +2.21 dB to the given dBV values.

$$\text{Level (dBV)} = 20 \times \log_{10} \frac{\text{Level}_1}{1\,\text{V}}$$

$$\text{Level (dBu)} = 20 \times \log_{10} \frac{\text{Level}_1}{0.775\,\text{V}}$$

The voltage-related dB scales serve the important purpose of guidance toward the optimal operating range of the electronic devices. The upper and lower limits of an electronic device are absolute, not relative values. The noise floor has a steady average level and the clip point is a fixed value. These are expressed in dBV. The absolute level of our signal will need to pass between these two limits in order to prevent excess noise or distortion. The area enclosed by these limits is the linear operating **area** of the electronic device. Our designs will need to ensure that the operating levels of electronic devices are appropriately scaled for the signal levels passing through.

Once we have captured the audio waveform in its electronic form, it will be passed through the system as a voltage level with negligible current and therefore minimal power dissipation. Low impedance output sections, coupled with high impedance inputs, give us the luxury of not considering the power levels until we have reached the amplifier output terminals. Power amplifiers can then be seen as voltage-driven input devices with power stage outputs to drive the speakers. The amplifier gives a huge current boost and additional voltage capability as well. Figure 1.12 provides a reference chart showing the standard operating voltage levels for all stages of the system signal flow. The goal is to transmit the signal through the system in the linear operating voltage range of all of the devices, without falling into the noise floor at the bottom.

Audio System Voltage Level Reference

Audio Signal Level			Mic Level	Line Level	Speaker Level		
Voltage (Volts)	dBV	dBu	250 Ω (Watts)	1k Ω (Watts)	16 Ω (Watts)	8 Ω (Watts)	4 Ω (Watts)
89	39	41.2			512	1024	2048
63	36	38.2			256	512	1024
45	33	35.2			128	256	512
32	30	32.2			64	128	256
22	27	29.2			32	64	128
16	24	26.2		250 m	16	32	64
11	21	23.2		125 m	8	16	32
8.0	18	20.2		63 m	4	8	16
5.6	15	17.2		32 m	2	4	8
4.0	12	14.2		16 m	1	2	4
2.8	9	11.2		8 m	500 m	1	2
2.0	6	8.2		4 m	250 m	500 m	1
1.4	3	5.2		2 m	125 m	250 m	500 m
1.0	0	2.2		1 m	63 m	125 m	250 m
707 m	-3	-0.8			31 m	63 m	125 m
500 m	-6	-3.8			16 m	31 m	63 m
356 m	-9	-6.8			8 m	16 m	31 m
250 m	-12	-9.8			4 m	8 m	16 m
178 m	-15	-12.8			2 m	4 m	8 m
125 m	-18	-15.8			1 m	2 m	4 m
89 m	-21	-18.8				1 m	2 m
63 m	-24	-21.8					1 m
45 m	-27	-24.8					
32 m	-30	-27.8	250 m				
22 m	-33	-30.8	125 m				
16 m	-36	-33.8	63 m				
11 m	-39	-36.8	32 m				
8.0 m	-42	-39.8	16 m				
5.6 m	-45	-42.8	8 m				
4.0 m	-48	-45.8	4 m				
2.8 m	-51	-48.8	2 m				
2.0 m	-54	-51.8	1 m				
1.4 m	-57	-54.8					
1.0 m	-60	-57.8					
707 µ	-63	-60.8					
500 µ	-66	-63.8					
356 µ	-69	-66.8					
250 µ	-72	-69.8					
178 µ	-75	-72.8					
125 µ	-78	-75.8					
89 µ	-81	-78.8					
63 µ	-84	-81.8					
45 µ	-87	-84.8					
32 µ	-90	-87.8					
22 µ	-93	-90.8					
16 µ	-96	-93.8					
11 µ	-99	-96.8					
8.0 µ	-102	-99.8					
5.6 µ	-105	-102.8					
4.0 µ	-108	-105.8					
2.8 µ	-111	-108.8					
2.0 µ	-114	-111.8					
1.4 µ	-117	-114.8					
1 µ	-120	-117.8					

Figure 1.12 A reference chart for the typical operational voltage and wattage levels

There is still another set of letter appendices that can be added on to the dB voltage formulas. These designate whether the voltage measured is a short-term peak or the average value. AC signals are more complex to characterize than DC signals. DC signals are a given number of volts above or below the reference common. AC signals, by nature, go both up *and* down. If an average were taken over time we would conclude that the positive and negative travels of the AC signal average out to 0 volts. Placing our fingers across the AC mains will quickly alert us to the fact that averaging out to 0 volts over time does not mean there is zero energy there. Kids, don't try this at home.

The AC waveform rises to a maximum, returns to zero, falls to a minimum and then returns again. The voltage between the peak and the zero point, either positive or negative, is the **peak voltage** (V_{pk}). The voltage between the positive and negative peaks is the **peak-to-peak voltage** (Vp-p) The equivalent AC voltage to that found in a DC circuit is expressed as the **root-mean-square** (RMS) voltage (VRMS). The peak-to-peak value is naturally double that of the peak value. The RMS value is 70.7 per cent of the peak value.

All of these factors translate over to the voltage-related dB formulas and are found as dBV_{pk}, dBV_{p-p} and dBVrms respectively. The 70.7 per cent difference between peak and RMS is equivalent to 3 dB.

Crest Factor

The 3 dB difference between the peak and RMS values only holds as long as the input signal is a continuous single frequency, i.e. a simple sine wave. If the signal has multiple frequencies, the peak to RMS ratio is no longer constant. It is highly volatile, or dynamic. The presence of multiple frequencies creates momentary confluences of signals that can sum together for a fleeting moment into a peak that is higher than any of the individual parts. This is known as a **transient** peak. Most audio signals are transient by nature since we can't dance to sine waves. A strong transient, like a pulse, is one that has a very high peak value and a minimal RMS value. Transient peaks are the opposite extreme of peak-to-RMS ratio from the sine wave. The term to describe the variable peak-to-RMS ratio found in

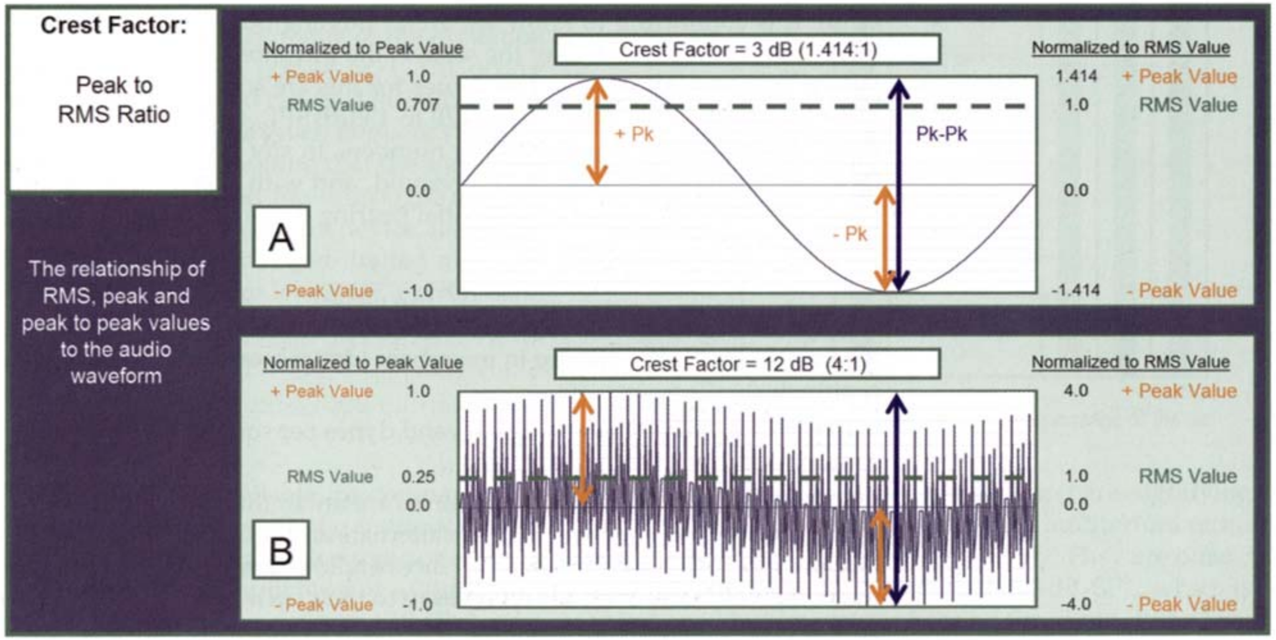


Figure 1.13 Crest factor, RMS, peak and peak-to-peak values. (A) A sine wave has the lowest crest factor of 3 dB. (B) An example complex waveform with transients with a 12 dB crest factor

Perspectives: System optimization starts with good system design. From there it is a process that should be guided by structured thinking and critical listening.

Sam Berkow, founder, SIA Acoustics llc & SIA Software Company, Inc.

different program materials is the **crest factor.** The lowest possible crest factor is 3 dB (sine wave). There is no limit to the maximum. The typical crest factor for musical signals is 12 dB. Since our system will be transmitting transients and continuous signals, we must ensure that the system has sufficient dynamic range to allow the transient peaks to stay within the linear operating range. The presence of additional dynamic range above the ability to pass a simple sine wave is known as **headroom,** with 12 dB being a common target range.

The Acoustic Decibel: dB SPL

The favored expression for acoustic transmission is **dB SPL** (sound pressure level). This is the quantity of measure for the pressure changes above and below the ambient air pressure. The standard is the threshold of the average person's hearing (O dB SPL). The linear unit of expression is one of pressure, dynes/square centimeter, with O dB SPL being a value of 0.0002 dynes/cm^2 (1 microbar). This lower limit approaches the noise level of the air medium, i.e. the level where the molecular motion creates its own random noise. It is comforting to know we aren't missing out on anything. At the other end of the scale is the threshold of pain in our hearing system. The values for this are somewhat inconsistent but range from 120 to 130 dB SPL, with more modern texts citing the higher numbers. In any case this number represents the pain threshold, and with that comes the obvious hazards of potential hearing damage.

$$\text{Level (dB SPL)} = 20 \times \log_{10} \frac{P}{0.0002}$$

where P is the RMS pressure in microbars (dynes/square centimeter).

Excuse me: did you say microbars and dynes per square centimeter?

These air pressure measurement terms are unfamiliar to most audio engineers, along with the alternate and equally obscure term of 20 micropascals. For most audio engineers the comprehension of dB SPL is relative to their own perspective: an experiential correlation to what we hear with what we have read over the years on SPL meters. The actual verification of 0 dB SPL is left to the Bureau of Standards and the people at the laboratories. Few of us have met a dyne, a microbar, or a micropascal at the venue or ever will. Even fewer are in a position to argue with someone's SPL meter as to whether it is calibrated correctly — unless we have an SPL meter and a calibrator of our own to make the case. Then we can argue over whose calibrator is accurate and eventually someone must yield or we will have to take a trip to the Bureau of Standards. This is one place where, in our practical world, we will have to take a leap of faith and trust the manufacturers of our measurement microphones. For our part we will be mindful of even small discrepancies between different measurement instruments, microphones, etc., but we will not be in a good position to question the absolute dB SPL values to the same extent.

The 130 dB difference between the threshold of audibility and the onset of pain can be seen as the dynamic range of our aural system. Our full range will rarely be utilized, since there are no desirable listening spaces with such a low noise floor. In addition, the ear system is generating substantial harmonic distortion before the pain threshold which degrades the sonic experience (for some) before actual pain is sensed. In practical terms we will need to find a linear operating range, as we did for the electronic transmission. This range runs from the ambient noise floor to the point where our hearing becomes so distorted that the experience is unpleasant. Our sound system will need to have a noise floor below the room and sufficient continuous power and headroom to reach the required maximum level.

dB SPL Subunits

dB SPL has average and peak level in a manner similar to the voltage units. The SPL values differ, however, in that there may be a time constant involved in the calculation.

- **dB SPL peak:** The highest level reached over a measured period is the peak (dB SPL_{pk}).
- **dB SPL** continuous **(fast):** This is the average SPL over a time integration of 250 ms. The integration time is used

14

in order to mimic our hearing system's perception of SPL. Our hearing system does not perceive SPL on an instantaneous level but rather over a period of approximately 100ms. The fast integration is sufficiently long enough to give an SPL reading that corresponds to that perception.

- **dB SPL** continuous (slow): This is the average SPL over a time integration of 1 second. The slower time constant mimics the perception of exposure to extended durations of sound.

dB SPL can be limited to a specific band of frequencies. If a bandwidth is not specified, the full range of 20-20kHz is assumed. It is important to understand that a reading of 120dB SPL on a sound level meter does not mean that a speaker is generating 120dB SPL at all frequencies. The 120dB value is the integration of all frequencies (unless otherwise specified) and no conclusion can be made regarding the behavior over the range of speaker response. The only case where the dB SPL can be computed regarding a particular frequency is when only that frequency is sent into the system. dB SPL can also be determined for a limited range of frequencies, a practice known as banded SPL measurements. The frequency range of the excitation signal is limited; commonly in octave or 1/3rd octave bands and the SPL over that band can be determined. The maximum SPL for a device over a given frequency band is attained in this way. It is worth noting that the same data cannot be attained by simply band-limiting on the analysis side. If a full range signal is applied to a device, its energy will be spread over the full band. Band-limited measurements will show lower maximum levels for a given band if the device is simultaneously charged with reproducing frequencies outside of the measured band.

Perspectives Regardless of the endeavor, excellence requires a solid foundation on which to build. A properly optimized sound system allows the guest engineer to concentrate on mixing, not fixing.

Paul Tucci

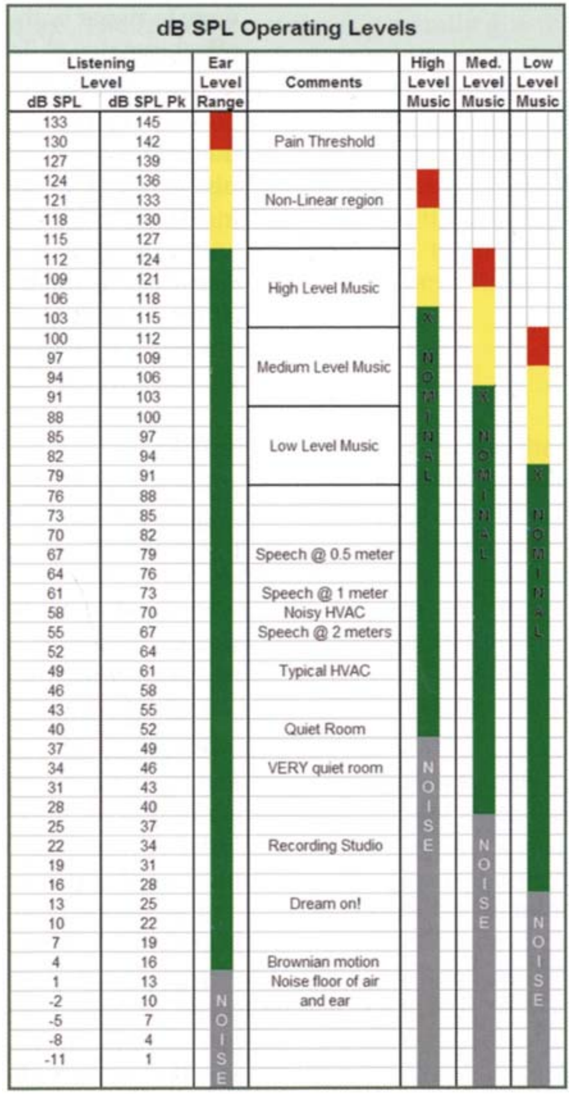


Figure 1 14 Typical operational level over the dynamic range of the ear

The Unitless Decibel

The unitless decibel scale is available for comparison of like values. Everything expressed in the unitless scale is *purely relative.* This can be applied to electronic or acoustic transmission. A device with an input level of —20dBV and an output level of —lOdBV has a gain of +10dB. Notice that no letter is appended to the dB term here, signifying a ratio of like quantities. Two seats in an auditorium receive 94 and 91dB SPL readings respectively. They are offset in level by 3dB. This is not expressed as 3dB SPL, which is a sound level just above our hearing threshold. If the levels were to rise at both seats to 98 and 95 respectively, the level offset remains at 3 dB.

The unitless dB scale will be by far the most common decibel expression in this book. The principal concern of this text is relative level, rather than absolute levels. Simply put: *absolute* levels are primarily an *operational* issue, inside the scope of mix engineering, whereas *relative* levels are primarily a design and *optimization* issue, under our control. The quality of our work will be based on how closely the final received signal resembles the original. Since the transmitted signal will be constantly changing, we can only view our progress in relative terms.

Power

Electrical energy is a combination of two measurable quantities: voltage and current. The electrical power in a DC circuit is expressed as:

$$P = EI$$

where P is the power in watts, E is the voltage in volts, and *I* is the current in amperes.

Voltage corresponds to the electrical pressure while **current** corresponds to the rate of electron flow. A simplified analogy comes from a garden hose. With the garden hose running open, the pressure is low and the flow is a wide cylinder of water. Placing our thumb on the hose will greatly increase the pressure, and reduce the width of the flow by a proportional amount. The former case is low voltage and high current, while the latter is the reverse.

The power is the product of these two factors. 100 watts of electrical power could be the result of 100 volts at 1 ampere, 10 volts at 10 amperes, 1 volt at 100 amperes, etc. The power could be used to run a single heater at 100 watts, or 100 heaters at 1 watt each. The amount of heat generated, and the electrical bill, will be the same either way. Power is the measure of energy, which can be distributed in different ways but remains unchanged in its overall quantity. Both voltage and current must be present for power to be transmitted. Voltage with no current is potential for power, but none is transferred until current flows.

A third factor plays a decisive role in how the voltage and current are proportioned: **resistance.** Electrical resistance limits the amount of current flowing through a circuit and thereby affects the power transferred. Provided that the voltage remains constant, the power dissipation will be reduced as the resistance stems the flow of current. This power reduction can be compensated by an increase in voltage proportional to the reduced current. Returning to the garden hose analogy, it is the thumb that acts as a variable resistance to the circuit, a physical feeling that is quite apparent as we attempt to keep our thumb in place. This resistance reapportions the hose power toward less current flow and higher pressure. This has very important effects upon how the energy can be put to use. If we plan on taking a drink out of the end of the hose it would be wise to carefully consider the position of our thumb.

The electrical power in a DC circuit is expressed as:

$$P = IE$$

$$P = I^2R$$

$$P = E^2/R$$

where P is the power in watts, *E* is the voltage in volts, *I* is the current in amperes, and *R* is the resistance in ohms.

These DC formulas are applicable to the purely resistive components of the current-limiting forces in the electrical circuit. In the case of our audio waveform, which is by definition an AC signal, the measure of resistive force differs over frequency. This complex term for resistance with a frequency component is **impedance.** The impedance of a circuit at a given frequency is the combination of the DC resistance, and the **reactance.** The reactance is the value for variable resistance over frequency and comes in two forms: capacitive and inductive. The impedance for a given circuit is a combination of the three resistive values: DC resistance, **capacitive reactance** and **inductive reactance.** These factors alter the frequency response in all AC circuits; the question is only a matter of degree of effect. For our scope here we will not go into the internal circuit components but rather limit the discussion of impedance and reactance to the interconnection of audio devices. All active audio devices present an input and output impedance and these must be configured properly for optimal transmission. The interconnecting cables also present variable impedance over frequency and this will be discussed.

Frequency Response

If a device transmits differently at one frequency than another, it has a **frequency response.** A device with no differences over frequency, also known as a "flat" frequency response, is actually the absence of a frequency response. In our practical world this is impossible, since all audio devices, even oxygen-free hypoallergenic speaker cable, change their response over frequency. The question is the extent of detectible change within the frequency and dynamic range of our hearing system. Frequency response is a host of measurable values but we will focus our discussion in this section on two representations: amplitude vs. frequency and phase vs. frequency. No audio device can reach to an infinitely low frequency (we would have to go all the way back to the Big Bang to measure the lowest frequency) and none can reach to an infinitely high frequency. Fortunately this is not required. The optimal range is some amount beyond that of the human hearing system. (The exact extent is subject to ongoing debate.) It is generally accepted that high-frequency extension beyond the human hearing limits is preferable to those systems that limit their response to exactly 20 Hz to 20 kHz. This is generally attributed to the reduced phase shift of the in band material, which leaves the upper harmonic series intact. Anyone familiar with the first generations of digital audio devices will remember the unnatural quality of the band-limited response of those systems. The debate over 96 kHz, 192 kHz and higher sampling rates for digital audio will be reserved for those with ears of gold.

Amplitude vs. Frequency

Amplitude vs. frequency (for brevity we will call this the **amplitude response)** is a measure of the level deviation over frequency. A device is specified as having an operational frequency range, and a degree of variation within that range. The frequency range is generally given as the —3 dB points in electronic devices, while —6 and —10 dB figures are most often used for speakers. The quality of the amplitude response is determined by its degree of variance over the transmission range, with minimum variance corresponding to the maximum quality. The speakers shown in Fig. 1.15 have an amplitude response that is ±4 dB over their operating ranges. The operating ranges (between —6 dB points) differ in the low frequency (40 and 70 Hz) and high frequency (18 kHz and 20 kHz).

Phase vs. Frequency

Phase vs. frequency (for brevity we will call this the phase response) is a measure of the time deviation over frequency. A device is specified as having a degree of variation within the operational range governed by the amplitude response. The quality of the phase response is determined by its degree of variance over the transmission range, with minimum variance again corresponding to the maximum quality. The phase responses of the two speakers we compared previously in Fig. 1.15 are shown as an example.

The phase response over frequency is, to a large extent, a derivative of amplitude response over frequency. Leaving the rare exception aside for the moment we can say that a flat phase response requires a flat amplitude response. Deviations in the amplitude response over frequency (peak and dip filters, high- and low-pass filters, for example) will cause predictable deviations in phase. Changes in amplitude that are independent of frequency are also independent of phase; i.e. an overall level change does *not* affect phase.

The exception cited above refers to filter circuits that delay a selective range of frequencies, thereby creating phase deviations unrelated to changes in amplitude. This creates interesting possibilities for phase compensation in acoustical systems with physical displacements between drivers. The final twist is the possibility of filters that change amplitude without changing phase. The quest for such circuits is like a search for the Holy Grail in our audio industry. We will revisit this quest later in Chapter 8. Phase will be covered from multiple perspectives as we move onward. For now we will simply introduce the concept.

Phase response always merits a second-place finish in importance to amplitude response for the following reason: if the amplitude value is zero there is no level, and the phase response is rendered academic. In all other cases, however, the response of phase over frequency will need to be known.

There has been plenty of debate in the past as to whether we can discern the phase response over frequency of a signal. The notion has been advanced that we cannot hear phase directly and therefore a speaker with extensive phase shift over frequency was equivalent to one that exhibited flat phase. This line of reasoning is absurd and has few defenders left. Simply put, a device with flat phase response sends all frequencies out in the temporal sequence as they went in. A device with non-linear phase selectively delays some frequencies more than others. These discussions are typically focused on the performance of loudspeakers, which must overcome tremendous challenges in order to maintain a reasonably flat phase response for even half of their frequency range. Consider the following question: all other factors being equal, would a loudspeaker with flat phase over a six-octave range sound better than one that is different for every octave? The answer should be self-evident unless we subscribe to the belief that loudspeakers *create* music instead of *recreating* it. If this is the case, then I invite you to consider how you would feel about your console, cabling or amplifiers contributing the kind of wholesale phase shift that is justified as a musical contribution by the speaker.

A central premise of this book is that the loudspeaker is not given any exception for musicality. Its job is as dry as the wire that feeds it an input signal: track the waveform. There will be no discussion here as to which forms of coloration in the phase or amplitude response sound "better" than another.

Let's apply this principal to a musical event: a piano key is struck and the transient pressure peak contains a huge range of frequency components, arranged in the distinct order that our ear recognizes as a piano note. To selectively delay some of the portions of that transient

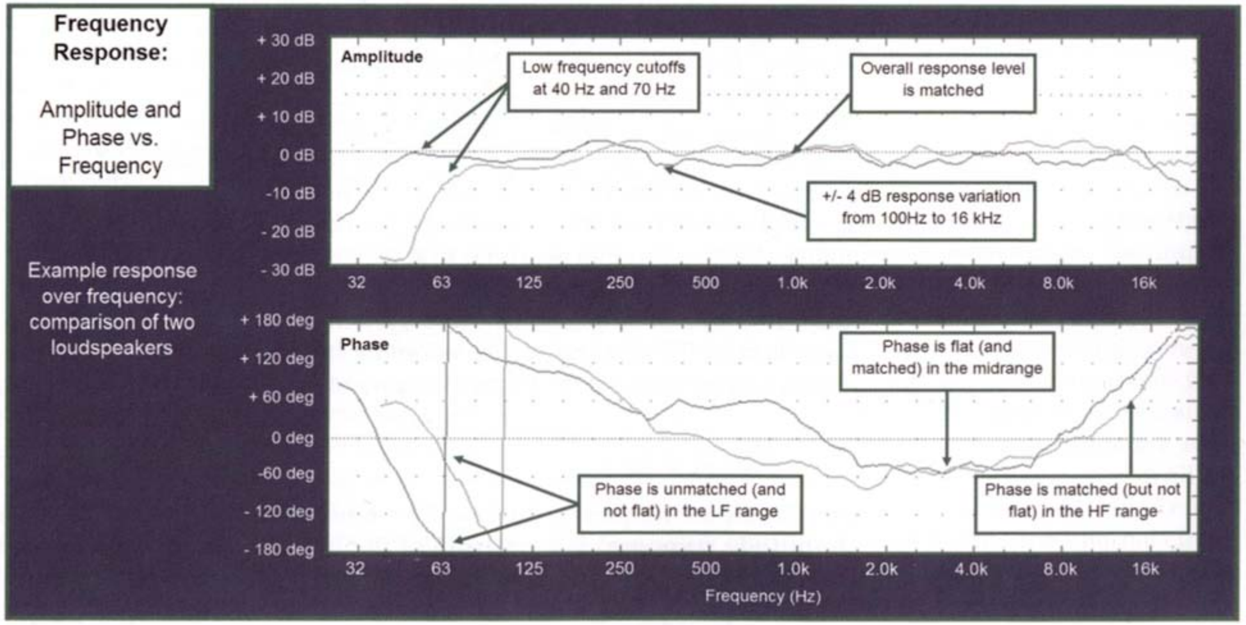


Figure 1.15 Relative amplitude and relative phase responses of two loudspeakers. The two systems are matched in both level and phase over most of the frequency range, but become unmatched in both parameters in the LF range

18

Perspectives: Every voodoo charm and ghost in the machine that I dismissed early in my career about why anything audio was the way it was, has little by little been replaced by hard science. I have no reason to believe that the voodoo and ghosts that persists can't be replaced as well by experience and truth.

Martin Carillo

peak rearranges the sequences into a waveform that is very definitely *not* the original and is less recognizable as a piano note. As more phase shift is added, the transient becomes increasingly stretched over time. The sense of impact from a hammer striking a string will be lost.

While linear phase over frequency is important, it is minor compared to the most critical phase parameter: its role in summation. This subject will be the detailed in Chapter 2.

Polarity

The polarity of a signal springs from its orientation to the origin point of the waveform. All waveforms begin at the "ambient" state in the medium and proceed back and forth from there. The same waveform shape can be created in opposite directions, one going back and forth while the other goes forth and back. There is plenty of debate as to whether we can discern the absolute polarity of a signal. A piano key is struck and the pressure peak arrives first as a positive pressure followed by a negative. If this is reproduced by an otherwise perfect speaker but reversed in polarity would we hear the difference? The debate continues.

In our case the critical parameter regarding polarity is ensuring that no reversals occur in parts of the transmission chain that will be combined either electrically or acoustically. Combining out polarity signals will result in cancellation, a negative effect of which there is little debate.

Perspectives: Keeping your gain structure optimized will ensure the whole system will clip simultaneously, thus ensuring the best signal-to-noise ratio.

Miguel Lourtie

Latency

Transmission takes time. Signal is sent through the signal path from the source to the listener. Every stage in the path takes some amount of time to pass the signal along. This type of delay, known as latency, occurs equally at all frequencies, and is measured in (hopefully) ms. The most obvious form of latency is the "flight time" of sound propagating through the air. It does not take much distance to rack up many milliseconds of delay in the acoustic path. The electronic path is also fraught with latency issues, and this will be of increasing importance in the future. In purely analog electronic transmission the latency is so small as to be practically negligible. For digital systems the latency can never be ignored. Even digital system latencies as low as 2 ms can lead to disastrous results if the signal is joined with others at comparable levels that have analog paths (0 ms) or alternate network paths (unknown number of ms). For networked digital audio systems latency can be a wide-open variable. In such systems it is possible to have a single input sent to multiple outputs each with different latency delays, even though they are all set to "0 ms" on the user interface. The modern audio engineer must keep a vigilant eye on their digital systems to ensure that no surprises arise. This is especially true now that we have digital mixers designed for live audio and which *can* distort (alter) the arrival times of multiples of signals when summed into a mix bus.

Analog Audio Transmission

We have discussed the frequency, period, wavelength and the audio waveform above. It is now time to focus on the transmission of the waveform through the electronic and acoustic mediums. We will begin with the far less challenging field of electronic transmission.

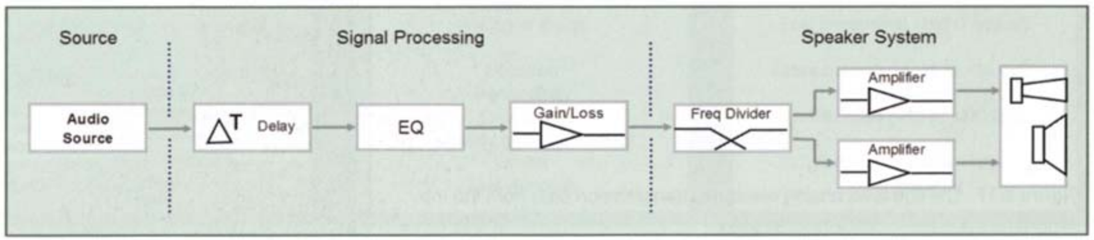


Figure 1.16 The analog electronic transmission path from the mix console to the loudspeaker

Electronic audio signals are variations in voltage, current or electromagnetic energy. These variations will finally be transformed into mechanical energy at the speaker. The principal task of the electronic transmission path is to deliver the original signal from the console to the mechanical/acoustic domain. This does not mean that the signal should enter the acoustic domain as an exact copy of the original. In most cases it will be preferable to modify the electrical signal in anticipation of the effects that will occur in the acoustic domain. Our goal is a faithful copy at the final destination: the listening position. To achieve this we will need to precompensate for the changes caused by interactions in the space. The electronic signal may be processed to compensate for acoustical interaction, split apart to send the sound to multiway speakers and then acoustically recombined in the space.

The analog audio in our transmission path runs at two standard operational levels: line and speaker. Each of these categories contains both active and passive devices. Active line level devices include the console, signal processing such as delays, equalization, level controls and crossovers and the inputs of power amplifiers. Passive devices include the cables, patchbays and terminal panels that connect the active devices together. Active devices are categorized by their maximum voltage and current capability into three types: mic, line and speaker level. Mic and line both operate with high impedance-balanced inputs (receivers) and low impedance-balanced outputs (sources). Input impedances of 5-100 kohms and output drives of 32-200 ohms are typical. Mic level devices overload at a lower voltage than line level, which should be capable of approximately 10 volts (+20dBV) at the inputs and outputs. Since our scope is focused on the transmission side of the sound system we will be dealing almost exclusively with line level signals. Power amplifiers have a high-impedance line level input and an extremely low-impedance speaker level output. Speaker level can range to over 100 V and is potentially hazardous to both people and test equipment alike.

Line Level Devices

Each active device has its own dedicated functionality, but they all share common aspects as well. All of these devices have input and output voltage limits, residual noise floor, distortion and frequency-response effects such as amplitude and phase variations. Looking deeper, we find that each device has latency, low-frequency limits approaching DC and high-frequency limits approaching light. In analog devices these factors can be managed such that their effects are practically negligible — but this cannot be assumed. The actual values for all of the above factors can be measured and compared to the manufacturer's specification and to the requirements of our project.

The typical electronic device will have three stages: input, processing and output as shown in Fig. 1.18. The nature of the processing stage depends upon the function of the device. It could be an equalizer, delay, frequency divider or any audio device. It could be analog or digital in the

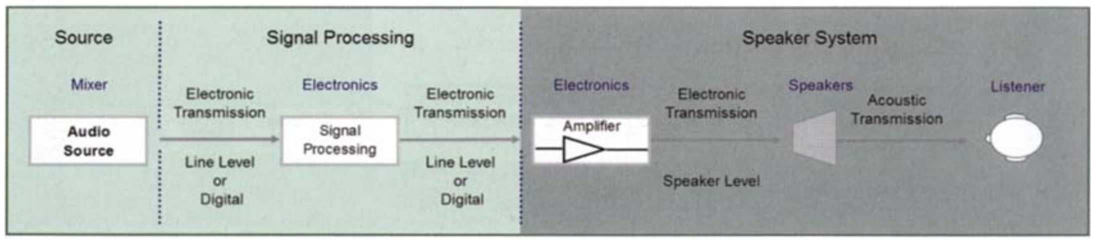


Figure 1.17 The line level analog electronic transmission path from the mix console to the input of the power amplifiers

processing section but the input and output stages are analog. Professional line level electronic devices utilize a fairly standard input and output configuration: the balanced line. Balanced lines provide a substantial degree of immunity from noise induced on the line (electromagnetic interference) and grounding problems (hum) by virtue of the advantages of the differential input. The differential input is discussed later in this chapter. The standard configuration is a voltage source system predicated on low-impedance outputs driving high-impedance inputs. This relationship allows the interconnection to be relatively immune to distance and number of devices.

Key specifications for line level active electronic devices:

- Active balanced high-impedance input: 10 kohms
- Active balanced low-impedance output: 150 ohms
- Frequency response range: 8 Hz to 22 kHz
- Amplitude response: ±0.5 dB, 20 Hz to 20 kHz
- Phase response: <45 degrees from 20 Hz to 20 kHz
- Maximum input capability: > +16 dBV, 20 Hz to 20 kHz
- Maximum output capability: > +16 dBV, 20 Hz to 20 kHz
- Hum and noise: < -90 dBV, 20 Hz to 20 kHz
- Dynamic range: >100 dB, 20 Hz to 20 kHz
- THD: <0.1 per cent, 20 Hz to 20 kHz
- Latency (analog): <0.01 ms (10 nanoseconds)
- Latency (digital): <10 ms
- Polarity: non-inverting
- Indicators: input and output overload.

Now let's look at some of the particular types of devices in common use in our systems. The list that follows is by no means comprehensive. The descriptions for this limited list are for the most generic and established features and applications of these devices. There are simply too many devices to describe. The description of and advocacy for (or against) the specific features and benefits of any particular make and model will be left to the manufacturers. Therefore, consider the insertion of text such as "in most cases," "typically but not always," "usually," "in every device I ever used in the last twenty years," or "except for

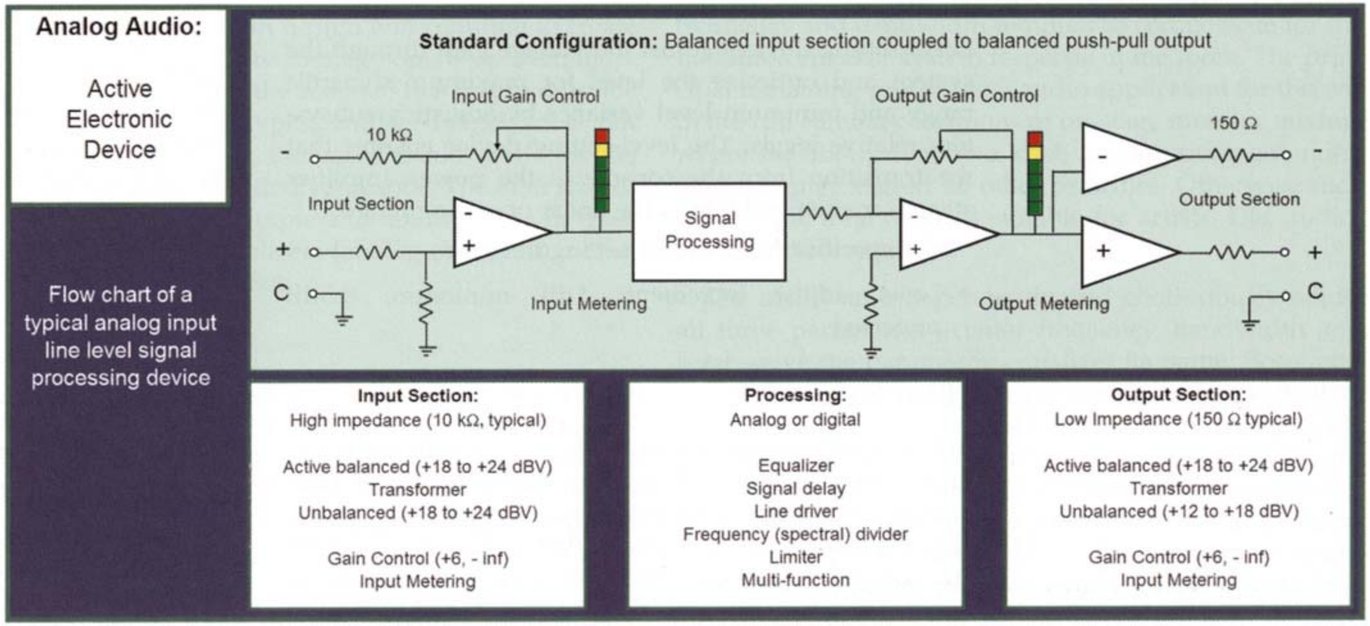


Figure 1.18 Typical analog electronic device flow chart

the model X" as applicable to any of the descriptions that would otherwise seem to exclude a particular product.

Audio Sources

The device transmitting our original transmitted signal must be capable of delivering the signal in working condition. The most common delivery device is the mix console outputs. The console outputs should meet the above criteria in order for us to have an input signal worth transmitting.

Signal Processing

The signal-processing devices will be used during the alignment process. The processing may be in a single chassis, or separated, but will need the capability to perform the basic functions of alignment: level setting, delay setting, and equalization. Since these will be covered separately we will consider the functions of the signal processor as the combination of individual units.

Level-setting Devices

Level-setting devices adjust the voltage gain through the system and optimize the level for maximum dynamic range and minimum level variance by adjusting subsystem relative levels. The level-setting device ensures that the transition from the console to the power amplifier stage is maintained within the linear operating range.

Key specification for level-setting devices:

- Level adjust increment: 1 dB minimum, 0.5 dB preferred.

Delay Lines

Just as level controls adjust the relative levels, the delay lines control relative phase. They have the job of time management. They aid the summation process by phase aligning acoustical crossovers for maximum coupling and minimum variance. Most delay lines have a resolution of 0.02 ms or less, which is sufficient. The minimum latency value is preferred. Most delay line user interfaces give false readings as to the amount of delay they are actually adding to the signal. The indicator on the device gives the impossible default setting of 0ms, which is accurate only as an indication of the amount of delay *added* to the latency base value. An accurate unit of expression for this would be 0ms(R); i.e. relative. If every device in the system is the same model and they are not routing signals through each other on a network, the relative value will be sufficient for our purposes. If we are mixing delay line models or networking them together, that luxury is gone. Any latency variable in our transmission path will need to be measured on site and taken into account.

Key specifications for delay lines:

- Delay resolution increment: 0.1 ms minimum, 0.02 ms or less preferred
- Maximum latency: <5 ms preferred.

Equalizers

Filter Types

Equalizers are a user-adjustable bank of filters. Filters, in the most basic sense, are circuits that modify the response in a particular frequency range, and leave other areas unchanged. The term equalization filters generally refers to two primary types of filters: shelving and parametric. Shelving filters affect the upper and lower extremes and leave the middle unaffected. Parametric filters do the opposite, affecting the middle and leaving the extremes unaffected. Used together, these two filter types can create virtually any shape we require for equalization.

Parametric-type equalization filter characteristics:

- Center frequency: the highest, or lowest point in the response, specified in Hz
- Magnitude: the level, in dB of the center frequency above, or below, the unity level outside of the filter's area of interaction.
- Bandwidth (or Q): The width of the affected area above and below the center frequency. This has some complexities that we will get to momentarily. In simple terms, a "wide" bandwidth filter affects a broader range of

frequencies on either side of the center than a "narrow" bandwidth filter with all other parameters being equal.

Shelving-type equalization filter characteristics:

- Corner frequency: the frequency range where the filter action begins. For example a shelving filter with a corner frequency of 8kHz will affect the range above this, while leaving the range below unaffected. It is more complicated than this in practice due to the various options available in filter circuits.
- Magnitude: the level, in dB of the shelved area, above or below the unity level outside of the filter's area of interaction.
- Slope (in dB/octave or filter order): this controls the transition rate from the affected and unaffected areas. A low slope rate, like the wide band filter described above, would affect a larger range above (or below) the corner frequency than a high slope.

There are a great variety of different subtypes of these basic filters. For both filter types the principal differences between subtypes are in the nature of the transitional slope between the specified frequency and the unaffected area(s). Advances in circuit design will continue to create new versions and digital technology opens up even more possibilities. It is beyond the scope of this text to describe each of the current subtypes, and fortunately it is not required. As it turns out, the equalization needs for the optimized design are decidedly unexotic. The filter shapes we will need are quite simple and are available in most standard parametric equalizers (analog or digital) manufactured since the mid-1980s.

Filter Functions

Since these devices are called "equalizers" it is important to specify what it is they will be making equal. The frequency response, of course. But how did it get unequal? There are three principal mechanisms that "unequalize" a speaker system in the room: an unequal frequency response in the speaker system's native response (a manufacturing or installation issue), air loss over frequency and acoustic summation. For the moment we will assume that a speaker system with a flat free field response has been installed in a room. The equalization scope consists of compensation for frequency-response changes due to air loss over distance, summation response with the room reflections and summation response with other speakers carrying similar signals.

Any of these factors can (and will) create peaks and dips in the response at frequencies of their choosing and at highly variable bandwidth and magnitudes. Therefore we must have the maximum flexibility in our filter set. The most well known equalizer is the "graphic" equalizer. The graphic is a bank of parametric filters with two of the three parameters (center frequency and bandwidth) locked at fixed settings. The filters are spread evenly across the log frequency axis in octave or 1/3rd octave intervals. The front panel controls are typically a series of sliders, which allows the user to see their slider settings as a response readout (hence the name "graphic"). The questionable accuracy of the readout is a small matter compared to the principal limitation of the graphic: fixed filter parameters in a variable filter application. This lack of flexibility severely limits our ability to accurately place our filters at the center frequency and bandwidth required to compensate for the measured speaker system response in the room. The principal remaining professional audio application for these is in the full combat conditions of on-stage monitor mixing, where the ability to grab a knob and suppress emerging feedback may exceed all other priorities. Otherwise such devices are tone controls suitable for artists, DJs, audiophiles and automobiles.

The ability to independently and continuously adjust all three parameters — center frequency, bandwidth and level — give the parametric equalizer its name. Boost and cut maximums of 15 dB have proven more than sufficient. Bandwidths ranging from 0.1 to 2 octaves will provide sufficient resolution for even precise studio applications. There is no limit to the number of filters that can be placed in the signal path; however, the point of diminishing returns is reached rapidly. If we find ourselves using large quantities of filters (more than six) in a single subsystem feed of a live sound system tuning application, it is worth considering whether solutions other than equalization have been fully explored. Recording studios, where the size of the audience is only as wide as a head, can benefit from higher numbers of filters. The equalization process will be discussed in detail in Chapter 10.

Complementary Phase

The phase response of the filters is most often the first derivative of the amplitude response, a relationship known as "minimum phase". This relationship of phase to amplitude is mirrored in the responses for which the equalizer is the proper antidote. This process of creating an inverse response in both amplitude and phase is known as complementary phase equalization.

It is inadvisable to use notch filters for system optimization. Notch filters create a cancellation and thereby remove narrow bands completely from the system response. A notch filter is a distinct form of filter topology, not simply a narrow parametric band. The application of notch filters is not equalization; it is elimination. A system that has eliminated frequency bands can never meet our goal of minimum variance.

Some equalizers have filter topologies that create different bandwidth responses depending on their peak and dip settings; a wide peak becomes a narrow dip as the gain setting changes from boost to cut. Is the bandwidth marking valid for a boost or a dip (or either)? Such filters can be a nuisance in practice since their bandwidth markings are meaningless. However, as long as we directly monitor the response with measurement tools, such filters should be able to create the complementary shapes required.

Bandwidth and Q

There are two common descriptive terms for the width of filters: bandwidth (actually percentage bandwidth) and Q or "quality factor." Both refer to the frequency range between the −3 dB points compared to the center frequency level. Neither of these descriptions provides a truly accurate representation of the filter as it is implemented. Why? What is the bandwidth of a filter that has only a 2.5 dB boost? There is no −3 dB point. The answer requires a brief peek under the hood of an equalizer. The signal path in an equalizer follows two paths: direct from input to the output bus and alternately through the filter section. The level control on our filter determines how much of the filtered signal we are adding (positively or negatively) to the direct signal. This causes a positive summation (boost) or a negative summation (cut) to be added to the full range direct signal. Filter bandwidth specifications are derived from the internal filter shape (a band pass filter) before it has summed with the unfiltered signals. The bandwidth reading on the front panel typically reflects that of the filter before summation, since its actual bandwidth in practice will change as level is modified. Some manufacturers use the measured bandwidth at maximum boost (or cut) as the front panel marking. This is the setting that most closely resembles the internal filter slope.

The principal difference between the two standards is one of intuitive understanding. Manufacturers seem to favor Q, which plugs directly into filter design equations. Most audio system operators have a much easier time visualizing one sixth of an octave than they do a Q of 9.

All that will matter in the end is that the shape created by a series of filters is the right one for the job. As we will see much later in Chapter 10, there is no actual need to ever look at center frequency level or bandwidth on the front panel of an equalizer. Equalization will be set by visually observing the measured result of the equalizer response

Bandwidth Versus Q Reference

BW (Octaves)	Q (rounded)
2	0.7
1.4	1
1	1.4
0.7	2
0.5	3
0.35 (1/3rd)	4
0.25 (1/4th)	6
.167 (1/6th)	9
0.125 (1/8th)	12
0.08 (1/12th)	18

Figure 1.19 Bandwidth vs. Q conversion reference (after Rane Corporation, www.rane.com/library.html)

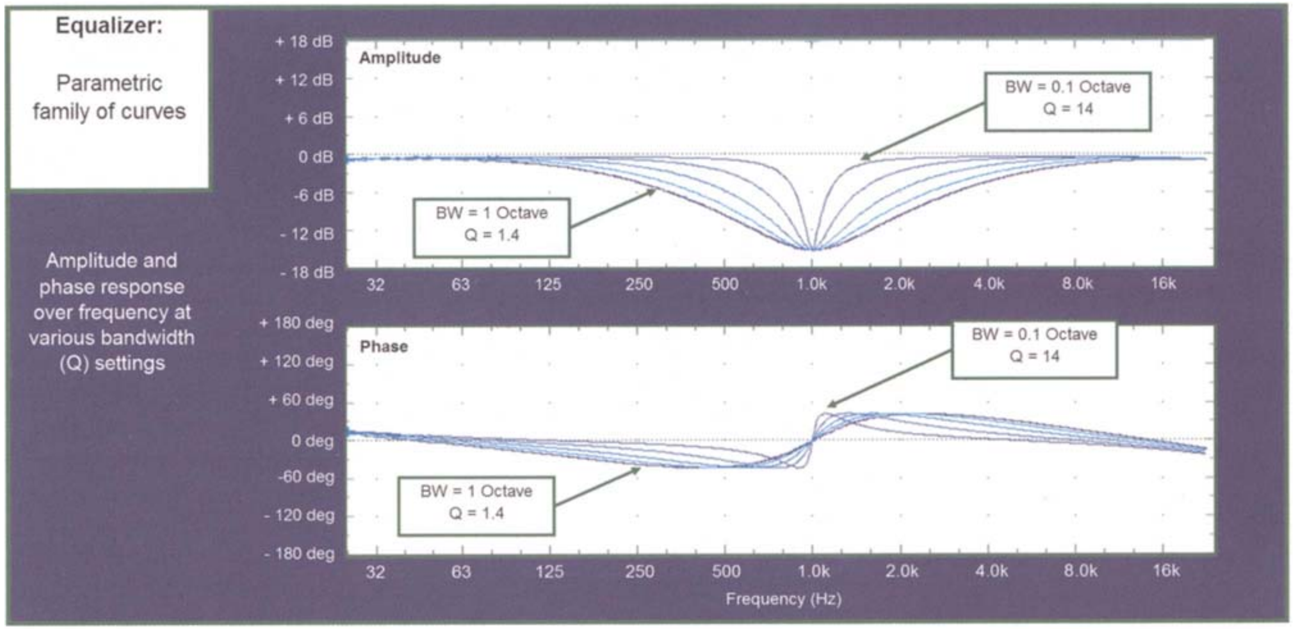


Figure 1.20 Standard parametric equalizer curve family

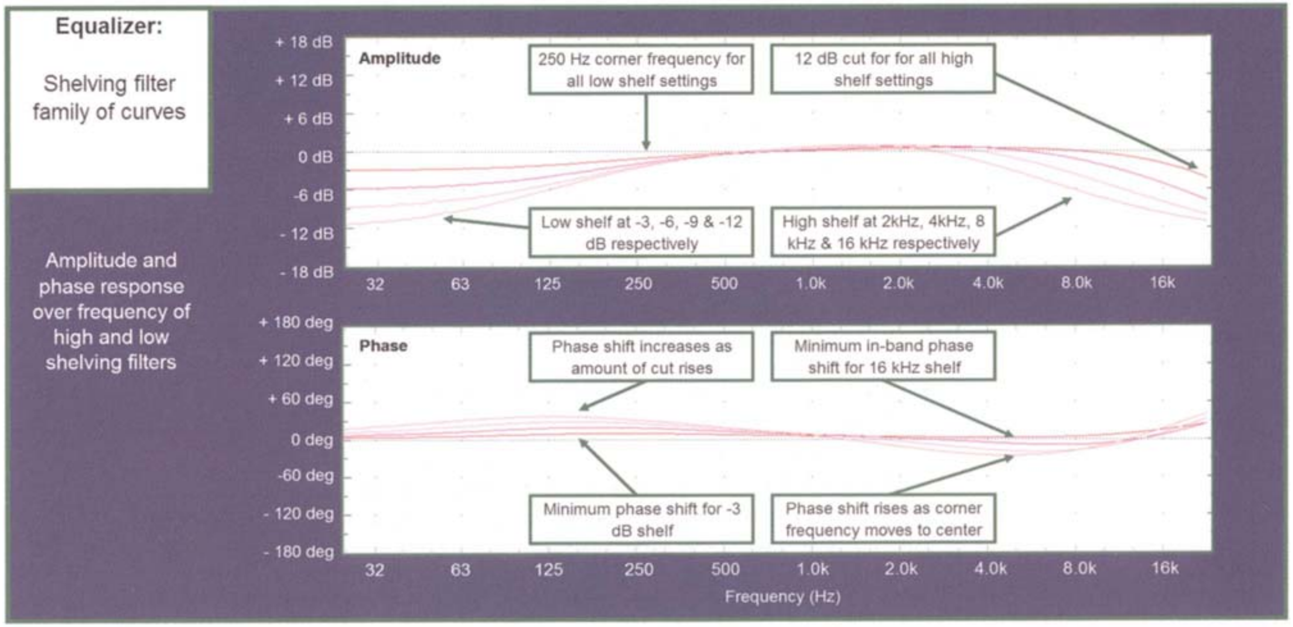
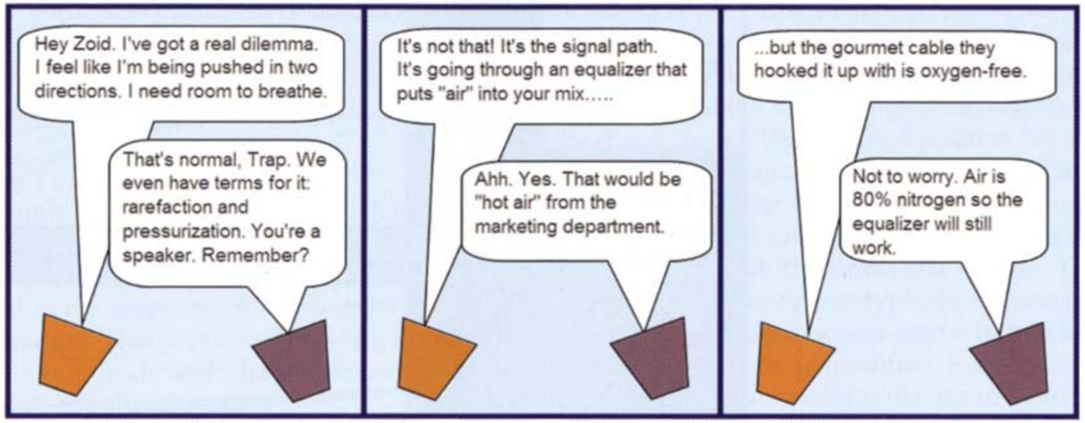


Figure 1.21 Shelving filter family of curves

Trap 'n Zoid by 6o6



Perspectives: 1 have had to optimize a couple of systems using 1/3 octave equalizers and I don't recommend it. I don't think that there is any argument about what can be done with properly implemented octave-based equalizers, but much more can be done with parametric equalizers that have symmetrical cut/boost response curves. The limitless possibilities for frequency and filter skirt width make fitting equalizers to measured abnormalities possible. With fixed filters the problems never seem to fall under the filters.

Alexander Yuill-Thornton II (Thorny)

and viewing it in context with the acoustic response it is attempting to equalize.

An additional form of filter can be used to create broad flat changes in frequency response. Known as shelving filters, they differ from the parametric by their lack of bandwidth control. The corner frequency sets the range of action and the magnitude control sets the level to which the shelf will flatten out. This type of filter provides a gentle shaping of the system response with minimal phase shift.

Key specifications for equalizers:

- Continuously adjustable center frequency, magnitude and bandwidth (Q)
- Minimum of five filters (as described above)
- Bandwidth range from 0.1 (or less) to 1 octave (or more)
- 15 dB boost or cut
- High and low shelving filters with adjustable level and optional slope adjust.

Frequency Dividers

The job of the frequency divider (also termed the spectral divider in this text) is to separate the audio spectrum so that it can be optimally recombined in the acoustic space. The separation is to accommodate the physics which make it impossible (currently) to have any one device that can reproduce the full audio range with sufficient quality and power.

Note: this device is commonly known as an electronic crossover, which is a misnomer. The electronic device *divides* the signal, which is then recombined acoustically. It is in the acoustical medium that the "crossing over" occurs, hence the term "acoustical crossover." This is not just a case of academic semantics. Acoustical crossovers are present in our room at any location where signals of matched origin meet at equal level. This includes much more than just the frequency range where high and low drivers meet, and also includes the interaction of multiple full range speakers and even room reflections. A central component of the optimized design is the management of acoustic crossovers. Frequency dividers are only one of the components that combine to create acoustic crossovers.

Frequency dividers have user-settable corner frequency, slope and filter topology. Much is made in the audio community of the benefits the Bessel, Butterworth, Linkwitz-Riley or some other filter topologies. Each topology differs somewhat around the corner frequency but then takes on the same basic slope as the full effects of the filter order become dominant. There is no simple answer to the topology question since the acoustic properties of the devices being combined will play their own role in the summation in the acoustical crossover range. That summation will be determined by the particulars of the mechanical devices as well as the settings on the frequency divider. A far more critical parameter than topology type is that of filter order, and the ability to provide different orders to the high-pass and low-pass channels. Placing different slope orders into the equation allows us to create an asymmetrical frequency divider, an appropriate option for the inherent asymmetry of the transducers being combined. Any frequency divider that can generate up to fourth order (24 dB per octave) should more than suffice for our slope requirements.

An additional parameter that can be found in some frequency dividers is phase alignment circuitry. This can come in the form of standard signal delay or as a specialized form of phase filter known as an all-pass filter. The standard delay can be used to compensate for the mechanical offset between high and low drivers so that the most favorable construction design can be utilized. The all-pass filter is a tunable delay that can be set to a particular range of frequencies. The bandwidth and center frequency are user-selectable. All-pass filters can be found in dedicated speaker controllers and active speakers, where conditions are sufficiently controlled such that the parameters can be optimized. The all-pass filter is of extremely limited practical field use since the field conditions make clear discernment of the phase response extremely challenging. It will be a happy day in the future when we reach a point where we have speaker systems in the field that are so well optimized that the only thing left to do is to fine tune all-pass delays.

Key specifications for frequency dividers:

- Continuously selectable corner frequencies
- Independent high-pass and low-pass parameters
- First-, second- and third-order minimum. Fourth order optional
- Phase alignment capability (signal delay and optional all pass).

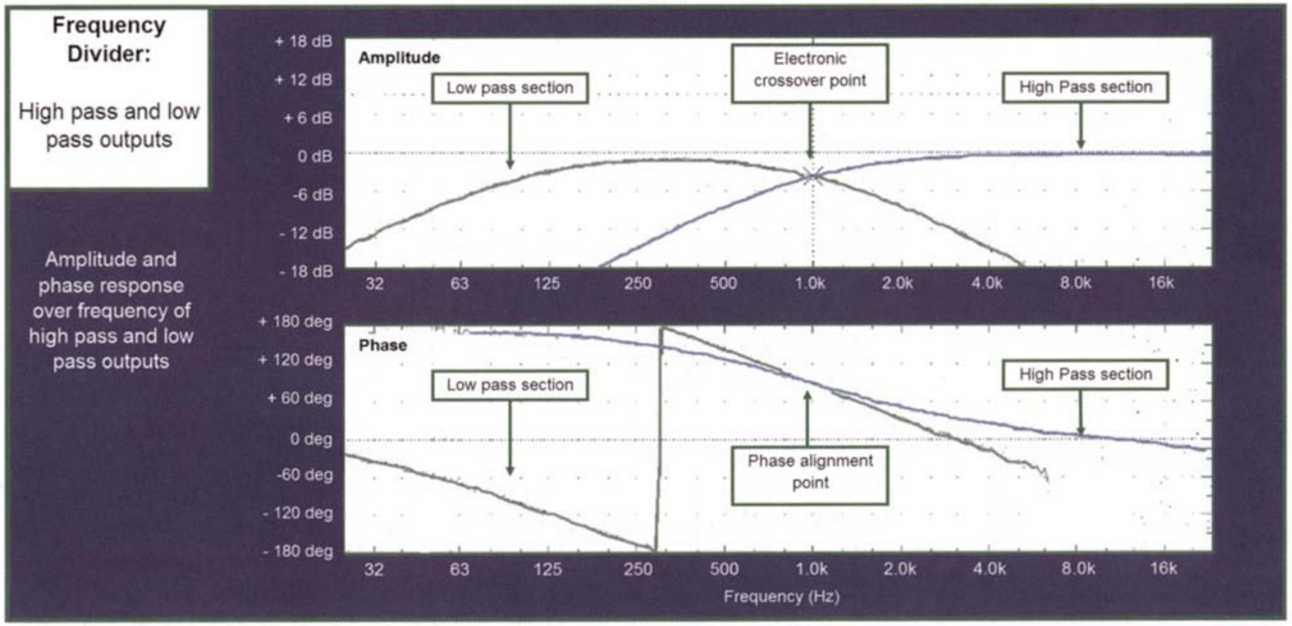


Figure 1.22 Frequency divider curve family

Limiters

Limiters are voltage regulation devices that reduce the dynamic range of the signal passing through. They can be applied at any point in the signal chain including individual input channels, output channels or the post-frequency-divider signal driving a power amplifier. The scope of our interest is those limiters that manage the input signal to the power amplifiers in order to provide protection for the loudspeakers and the amplifier. Limiters are not a requirement in the transmission path. If the system is always operated safely within its linear range there is no need for limiting. This could happen in our lifetime, as could world peace. But unfortunately we have to assume the worst-case scenario and that is that the system will be subjected to the maximum level of abuse fathomable, plus 6 dB. Overload conditions are a strain on both the amplifiers and the speakers, and have many undesirable sonic characteristics (to most of us). Limiters are devices with a threshold controlled variable voltage gain. The behavior of the device is characterized by its two operating ranges: linear and non-linear and by the timing parameters associated with the transition between the states: attack and release. The ranges are separated by the voltage threshold and the associated time constants that govern the transition between them. If the input signal exceeds the threshold for a sufficient period of time, the limiter gain becomes non-linear. The voltage gain of the limiter decreases because the output becomes clamped at the threshold level, despite rising levels at the input. If the input level recedes to below the threshold for a sufficient duration, the "all clear" sounds and the device returns to linear gain.

The limiter landscape is complicated these days by the many devices that now contain them. Limiters may be found inside the mix console where the application is principally for dynamic control of the mix, not for system protection. This is not our concern here. Additional locations include the signal processing chain both before and after active frequency dividers. A limiter applied before frequency division is difficult to link to the physics of a particular driver, unless the limiter has frequency-sensitive threshold parameters. The most common approach to system protection is after the frequency divider, where the limiters are charged with a particular driver model operating in a known and restricted frequency range. Such limiters can be found as stand-alone devices, as part of the active frequency divider (analog or digital) or even inside the power amplifier itself.

Where does the limiter get its threshold and time constant settings? Most modern systems utilize factory settings based on the manufacturer-recommended power and excursion levels. There are two principal causes of loudspeaker mortality: heat and mechanical trauma. The heat factor is managed by RMS limiters which can monitor the long-term thermal conditions based upon the power dissipated. The mechanical trauma results from over-excursion of the drivers resulting in collision with the magnet structure or from the fracturing of parts of the driver assembly. Trauma protection must be much faster than heat protection; therefore the time constants are much faster. These types of limiters are termed peak limiters. Ideally, limiters will be calibrated to the particular physics of the drivers. If the limiters are not optimized for the drivers they will degrade the dynamic response or possibly fail in their primary mission of protection. The best designed systems incorporate the optimal balance of peak and RMS limiting. This is most safely done with the use of amplifiers with voltage limits that fall within the driver's mechanical range and peak limiters that are fast enough to prevent the amplifier from exceeding those limits or clipping. The results are peaks that are fully and safely realized and RMS levels that can endure the long-term thermal load.

A seemingly safe approach would be to use lower-power amplifiers or set the limiting thresholds well below the full capability so as to err on the side of caution. This is not necessarily the best course. Overly protective limiters can actually endanger the system because the lack of dynamic range causes operators searching for maximum impact to push the system into a continual state of compression. This creates a worst-case long-term heat scenario as well as the mechanical challenges of tracking square waves if the amplifiers or drive electronics are allowed to clip. The best chances of long-term survival and satisfaction are a combination of responsible operation and limiters optimized for maximum dynamic range.

Limiters can be split into two basic categories: predictive or negative feedback loop. Predictive limiters are inserted in the signal path before the power amplifier(s). They have no direct correlation to the voltage appearing at the amplifier outputs. Therefore, their relationship to the signal they are managing is an open variable that must be calibrated for the particulars of the system. For such a scheme to be successful the factors discussed above must become known and enacted in a meaningful way into the limiter. Such practices are common in the industry. I do not intend to be an alarmist. Thousands of speakers survive these conditions, night after night. The intention here is to heighten awareness of the factors to be considered in the settings. Satisfactory results can be achieved as long as the limiters can be made to maintain an appropriate tracking relationship to the output of the power amplifiers. Consult the manufacturers of speakers, limiters and amplifiers for their recommended settings and practices.

Required known parameters for predictive limiters:

- voltage limits (maximum power capability) of the amplifier
- voltage gain of the amplifier (this includes user-controlled level settings)
- peak voltage maximum capability of the loudspeaker
- excursion limits of the loudspeaker over its frequency range of operation
- long-term RMS power capability of the loudspeaker.

Negative feedback systems employ a loop that returns the voltage from the amplifier terminals. This voltage is then used for comparison to the threshold. In this way the voltage gain and clipping characteristics of the amplifier are incorporated into the limiting process. This is common practice in dedicated speaker controllers and can afford a comparable or greater degree of protection with lesser management requirements for the power amplifiers;

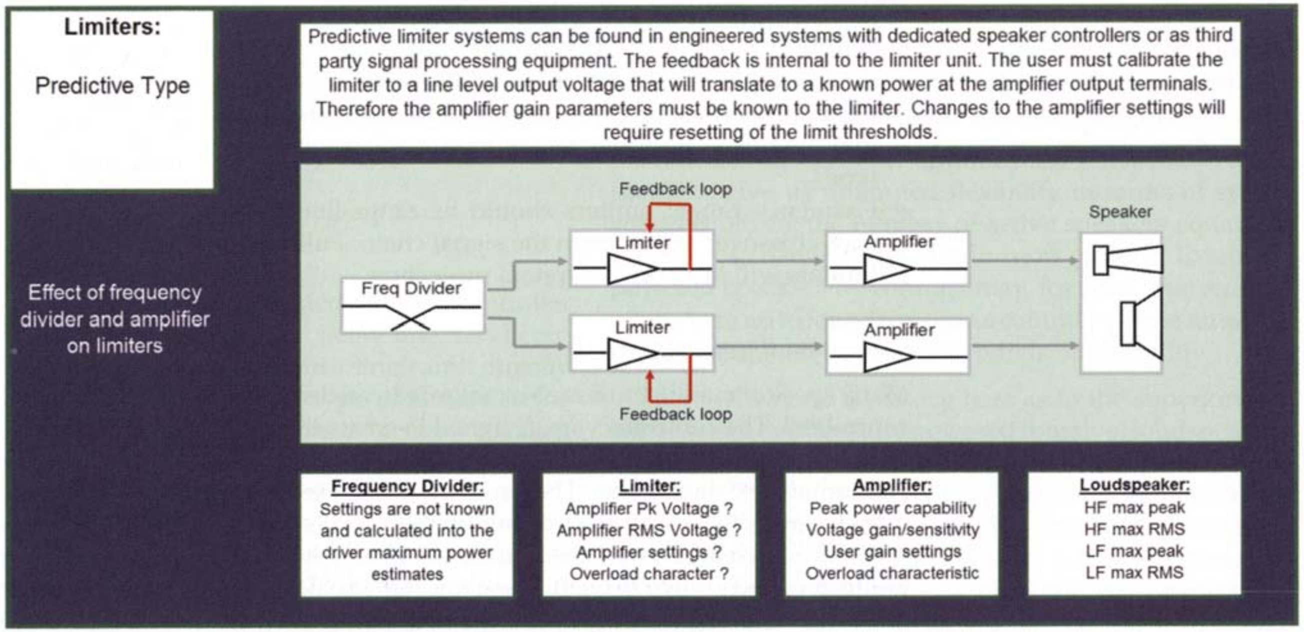


Figure 1.23 Predictive limiter scheme. The limiter is calibrated independently of the amplifier. For this scheme to be effective the limiter must be calibrated to the known voltage gain and peak power characteristics of the amplifier. Changes in the amplifier input level control will decalibrate the limiter

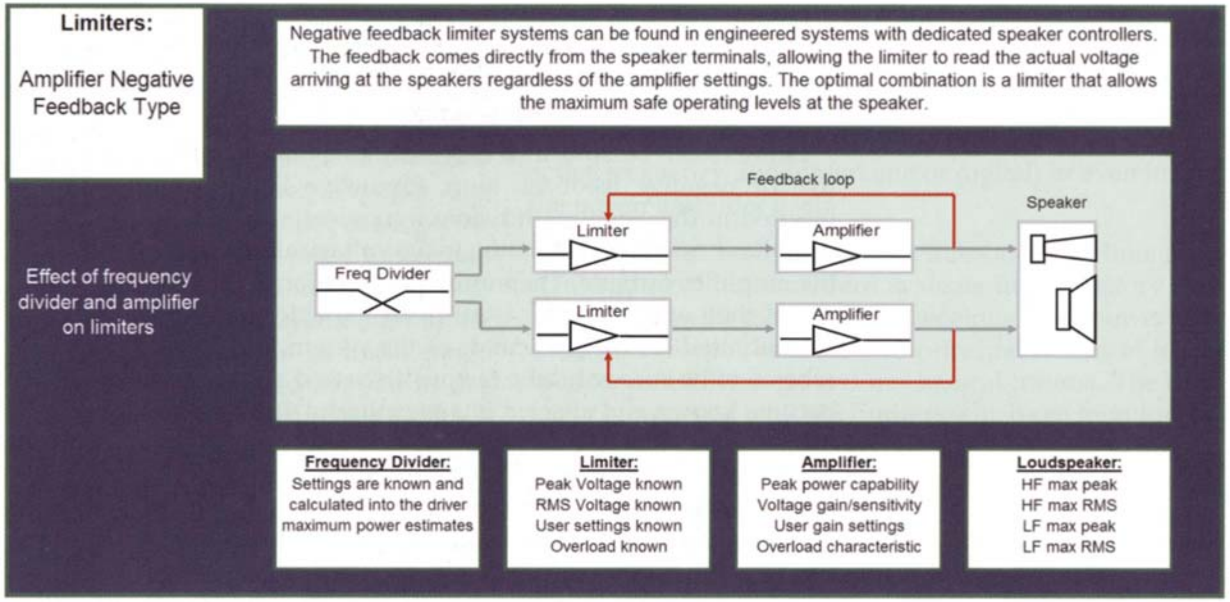


Figure 1.24 Negative feedback scheme. The limiter is calibrated to the output of the amplifier. The limiter remains calibrated even with changes in the amplifier level control

e.g. the amplifier input level controls may be adjusted without readjusting the limit threshold.

Key specifications for limiters:

- Limiter types: need both peak and RMS
- Calibration: must be calibrated to mechanical and thermal overload condition respectively for each speaker type
- Frequency range: limiters should be range-limited to post-crossover positions in the signal chain. Full-range limiters will not provide system protection.

Dedicated Speaker Controllers

Many speaker manufacturers also make dedicated speaker controllers. The controllers are designed to create the electronic processing parameters necessary to obtain optimal performance of the speaker. The parameters are researched in the manufacturer's controlled environment and a speaker "system" is created from the combination of optimized electronics and known drivers. The dedicated speaker controller contains the same series of devices already described above: level setting, frequency divider, equalization, phase alignment and limiters, all preconfigured and under one roof. The individual parameters are manufacturer- and model-dependent so there is little more that needs be said here. This does not mean that equalizers, delays and level-setting devices can be dispensed with. We will still need these tools to integrate the speaker system with other speakers and to compensate for their interaction in the room. We will, however, have reduced the need for additional frequency dividers and limiters and should have a much lighter burden in terms of equalization.

Two recent trends have lessened the popularity of these dedicated systems in the current era. The first is the trend toward third-party digital signal processors (DSPs). These units are able to provide all of the functionality of the dedicated controllers, usually with the exception of negative feedback limiters. Manufacturers supply users with factory settings that are then programmed into the DSP. These tools have the advantage of relatively low cost and flexibility but the disadvantages include a lack of standardized response since users of a particular system may

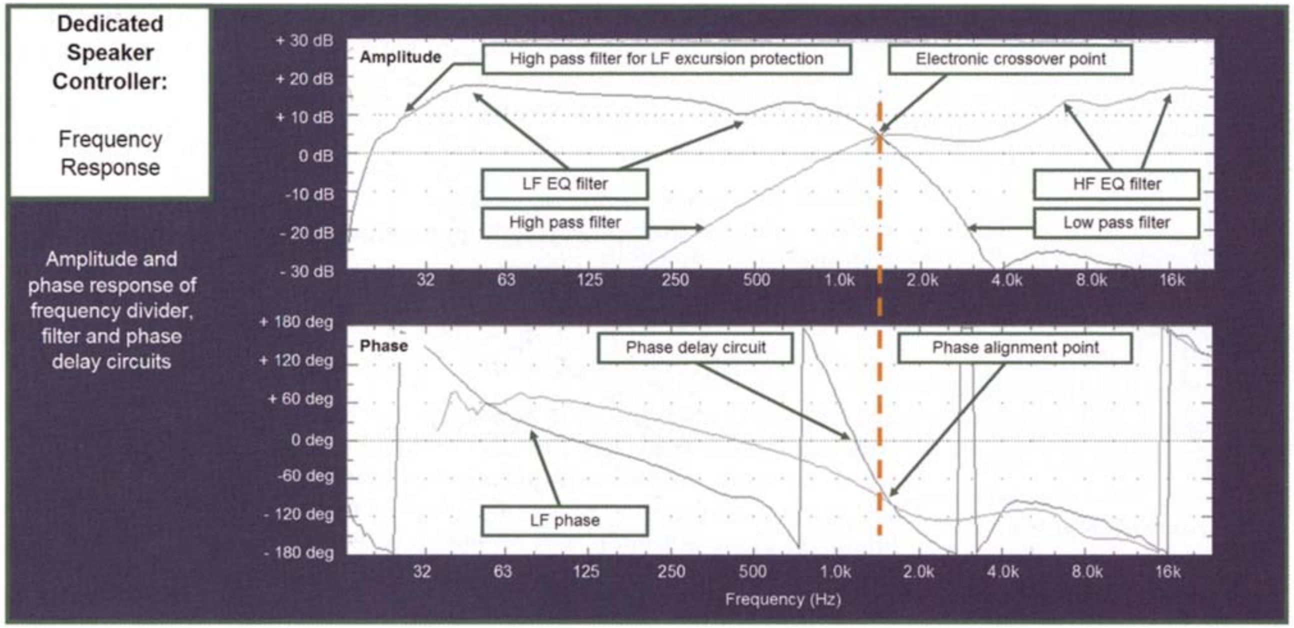


Figure 1.25 Frequency and phase response of an example two-way dedicated speaker controller

choose to program their own custom settings instead of those recommended by the manufacturer. There is also considerable latitude for user error since the programming, copying and pasting of user settings is often poorly executed even by the best of us. The second trend is toward completely integrated systems, inclusive of the amplifier in the speaker cabinet.

Active Speakers

The ultimate version of the dedicated speaker controller is this: a frequency divider, limiter, delay line, level control, equalizer and power amplifier in a single unit directly coupled to the speaker itself. This type of system is the self-powered speaker, also termed the "active" speaker. Active speakers operate with a closed set of variables: known drivers, enclosure, known physical displacement, maximum excursion and dissipation. As a result they are designed to be fully optimized for the most linear amplitude and phase response over the full frequency range and to be fully protected over their full dynamic range. Active speakers, like power amplifiers, have an open polarity variable, since they are electronically balanced at the input and moving air at the output. Since most active speakers came on the market after industry polarity standardization, it would be rare to find a non-standard system.

From our design and optimization perspective active speakers give us unlimited flexibility in terms of system subdivision; i.e. the number of active speakers equals the number of channels and subdivision options. Externally powered speakers by contrast may, for economic reasons, often share up to four drivers on a common power amplifier channel and thereby reduce subdivision flexibility.

There will be no advocacy here as to the superiority of either active or externally powered (passive) loudspeakers. We will leave this to the manufacturers. The principal differences will be noted as we progress, since this choice will affect our optimization and design strategies. In this text we will consider the speaker system to be a complete system with frequency divider, limiters and power amplifiers inclusive. In the case of the active speaker this is a physical

Perspectives: Mixing on a well-designed / well-tuned system is the difference between mixing "technically" and "musically." Too many times with the poorly tuned system it is a struggle all night long trying to make things work sonically and seems many times to be a battle that can't be won. The more frequencies we filter out of the system, the deeper we get in problems with the clarity and cohesiveness of the system. Correct tuning will let us mix in the "musical" format and therefore we then can think about how the kick drum and bass guitar are working together, how a piano pad and guitar rhythm patterns gel and so on. When we lift a guitar for its solo, we are then thinking, "What instruments can I bring up to accent that guitar solo?" That's what I call mixing musically and this requires a well-designed/well-tuned system.

Buford Jones

reality, while in the case of the externally powered system, the components are separated. The techniques required to verify the proper configuration of these two different system types will be covered in Chapter 9.

Key specifications for active speakers:

- Balanced line level input
- Able to reach full power within nominal operation limits
- 110 dB dynamic range.

Line Level Interconnection

This subject is often the source of confusion due in large part to practices that were developed in the distant past when our audio world was filled with transformer inputs and outputs connected to vacuum tubes. In that case the line level interconnection was an electrical power transfer (as is still the case with an amplifier and speaker). The line level power transfer strategy relied upon matched impedance between input and output and was based upon the old telephone transmission standard of 600 ohms. This approach was critical for the very real power transfer needs between the primitive carbon button transducers on either end of the telephone line. The line level devices of modern professional sound systems are active and do not operate under these restrictions and need transfer only negligible amounts of power. The signal transfer is effectively reduced to voltage only; hence the term voltage source describes the transmission system between line level devices. A voltage source is capable of signal transfer that is virtually without loss as long as the driving impedance is very low compared to the receiver. Hence, the modern system seeks an impedance mismatch in direct contrast to the practices of the past.

$$\text{Line loss (dB)} = 20 \times \log_{10} \frac{\text{Input} + \text{Cable} + \text{Output impedance}}{\text{Input impedance}}$$

The amount of loss over the line depends upon the combined impedance of the output relative to the combined impedance of the cable and the input. We will now apply this formula to the typical system shown previously in Fig. 1.18 (neglecting the cable impedance for the moment).

$$\text{Line loss (dB)} = 20 \times \log_{10} \frac{10,000 + 0 + 150}{10,000}$$

The combined 10 kohms input impedance and 150ohms output impedance are divided by the 1Okohms input impedance (10,150 ohms/10,000 ohms) to create a ratio of 203:200, a loss of 0.13 dB. Each additional input fed by this output would reduce the input impedance seen by the output device. Each parallel addition increases the loss by an additional 0.13 dB increment. Therefore we need not be concerned about splitting line level output signals to multiple sources until the number of inputs becomes quite large. The remaining factor is the cable impedance. Cable impedance is a product of the combined effects of cable DC resistance, and AC reactance, which comes in the form of cable capacitance and inductance. The reactive properties are frequency-dependent. Cable capacitance acts as a low-pass filter while inductance acts as a high-pass filter. The DC resistance acts as a full-range attenuator. The band-pass combination of the cable reactance can be a cause of concern if we are running long lines. Everything you ever wanted to know on this subject (and more) can be found in *Audio System: Design and Installation* by Phil Giddings. According to that text, as long as our runs are less than 305 m (1000 ft) there is no reason for concern. An additional consideration regarding the line level interconnection loss: it is the easiest loss to compensate. If this were not the case we would never have been able to communicate via telephones. To overcome this loss requires a voltage gain boost which is readily available to us in the form of active level controls. The principal side-effect would be a rise in the noise floor. This is not to say that such a loss is recommended — only that the interconnection loss in line level systems is far less critical than in speaker lines, where a 6 dB drop would mean a loss of 75 per cent of your power. The difference is that line level systems are transferring negligible power. They are transferring voltage, which is a simpler matter to restore.

The input stage is balanced in order to reject noise that is injected onto the cable driving the input. The term balanced line refers to the wiring configuration that uses two signal conductors and one common (which may or may not be connected to chassis or ground). The two conductors always contain identical copies of the input signal but are reversed in polarity from each other. The two signals are fed into a "differential" input stage which amplifies only signals that are unmatched — which is exactly the case with our input signal. Any noise that has entered the cable (electromagnetic interference and radio-frequency interference being the most typical) will be found equally in both conductors and will be cancelled by the differential input. This result is termed "common-mode rejection" for its ability to suppress induced signals.

The input section may have a level control, which will determine the drive level into the processing module. Once the processing is completed we proceed to the balanced push-pull output stage. The name comes from its function to prepare a balanced differential output signal: two identical signals, one of which is reverse polarity. As in the input stage, a level control may accompany this stage. Additionally input and/or output meters may be provided to monitor levels. One might assume that if the input and output level controls are set to their nominal setting that the signal is unity gain throughout the internal gain structure of the unit. This is not necessarily the case at all and will need to be determined during the verification process outlined in Chapter 9. The overall voltage gain can be anything, and the internal gains can also be unexpected. The most pressing case is that of pseudo-unity gain, which is still a fairly common practice in manufacturers of digital audio processors. In this case the voltage gain of the overall device is unity but the unity is created by as much as 20 dB of gain at the input and a tracking amount of loss at the output. The result of this is that the device driving the input can overload the unit by 20 dB unless it is forced to operate well below its standard levels. Why would manufacturers do this? The reason is to force users to drive the analog-to-digital (A/D) converters hard so that the noise floor limitations of the device are not detected. In most cases this can be remedied if detected in advance. Solutions include front panel and software settings and, worst case, the movement of internal jumpers inside the chassis.

Active Balanced

There is a variety of interconnection wiring schemes used in the industry. We will touch on this only briefly to discuss the most common schemes found for our usual signal path. The most common wiring scheme is, of course, not necessarily the best. This scheme ties the balanced inputs and outputs directly together along with the common (shield) as happens with a direct connection through a 3-pin XLR cable. The drawback of this scheme is the ground loop path that is created by the shield connection. An alternate scheme is able to provide shielded balanced operation without introducing a ground loop. The shield is connected on the source side only. This removes the ground loop. These schemes are shown in Fig. 1.26.

Transformer Balanced

A balanced transformer can be substituted for the active input with advantages in isolation between the systems. The degradation effects of transformers are well documented but there are some applications where the isolation advantages outweigh the degrading effects. This configuration is shown in Fig. 1.27.

Unbalanced

There are occasions when the equipment provided does not have balanced inputs or outputs. In such instances we will want to preserve the performance of balanced lines as much as possible. Wiring schemes have been developed that most closely approximate balanced line performance. These are shown in Fig. 1.28. When an unbalanced output drives a balanced input the differential inputs are fed by the signal and common respectively. This allows for the common mode rejection to suppress any interference that is injected into the line. The success of this scheme will depend upon the isolation between the grounds of the two devices.

Important note: unbalanced interconnections are one of the most common sources of polarity reversals. Verify that the connection is made across the non-inverting source and receivers terminals.

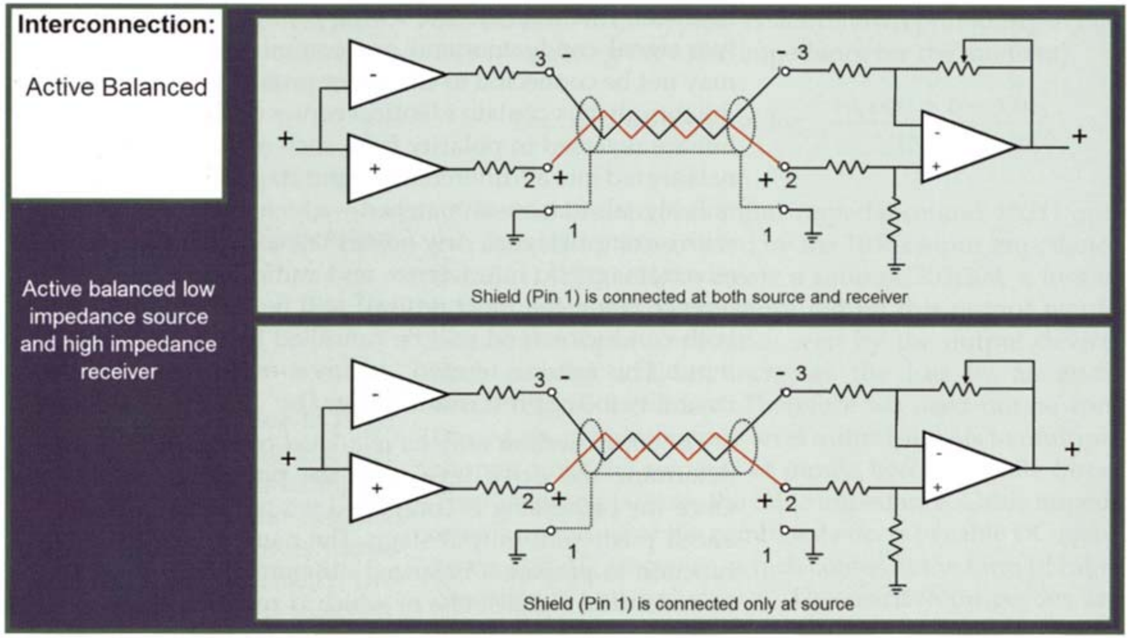


Figure 1.26 Top: typical active balanced line interconnection (possible ground loop). Bottom: improved active balanced line interconnection (no ground loop because the shield is not connected to the input ground) (after Giddings, 1990, pp. 219-220)

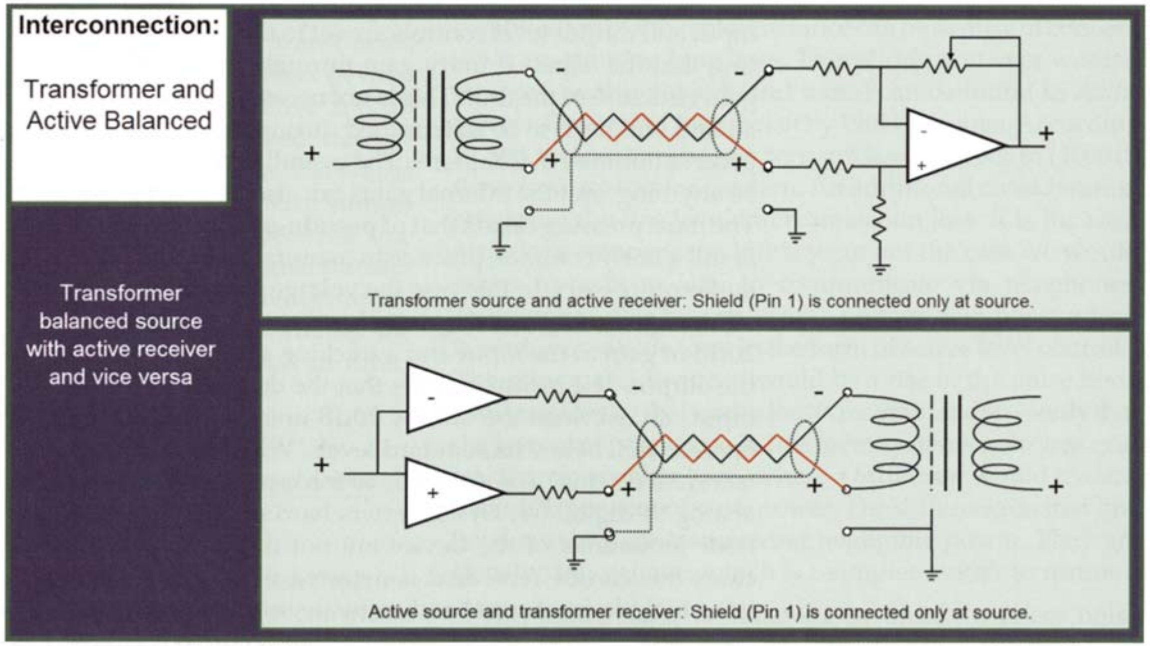


Figure 1.27 Top: balanced transformer output interconnection to active balanced input. Bottom: active balanced line interconnection to balanced transformer input (after Giddings, 1990, pp. 221-223)

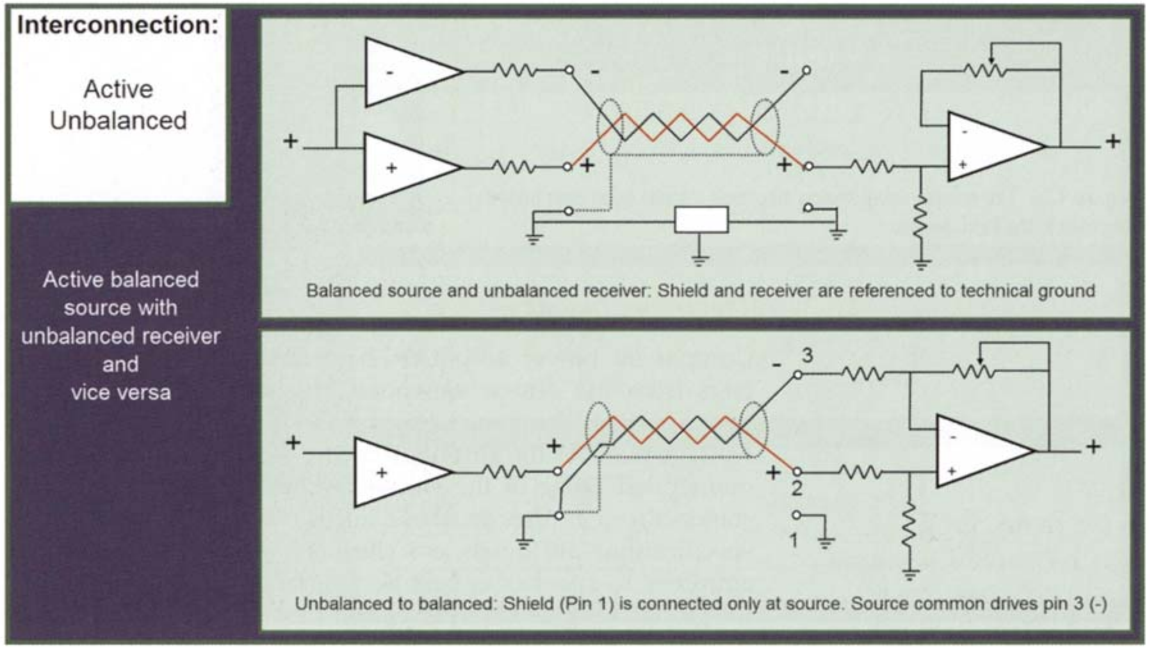


Figure 1.28 Top: active balanced output interconnection to active unbalanced input. Note that the commons must be linked by a technical ground. Bottom: active unbalanced output interconnection to active balanced input (after Giddings, 1990, pp. 226-229)

In cases where an active balanced output drives an unbalanced input the system will not have improved performance over an unbalanced-to-unbalanced configuration. The reverse polarity side of the push-pull output will not be used and the shield connection will be made at only one end. The ground connection will need to be made through a common technical ground rather than through the shield where a loop would be introduced.

Speaker Level Devices — Power Amplifiers

Speaker level transmission moves us into the realm of real power. The speaker level voltages run higher than line by a factor of 10:1 or more, reaching 100 volts RMS. But that ratio pales in comparison to the difference in current levels which can easily reach 250:1. This translates to real power. The minuscule power dissipation of a line level output gives way to the 1000 watts transmitted from a speaker level output. We are no longer able to operate with the one-dimensional voltage frame of mind. Speaker level transmission requires us to work with the more complex properties of electrical power.

In our professional audio transmission application, the amplifier is the source of current and voltage, and the speaker and cabling are the resistance. The motion of the speaker coil tracks the voltage of the audio waveform. The amplifier supplies current as dictated by the output voltage and the voice coil impedance and sets the speaker in motion. Speakers are not unlike a typical teenager: they are resistant to change and require a great deal of motivation to be moved from their natural state of rest. Current provides the motivation, while voltage guides the direction. A positive voltage will cause the speaker to move in one direction while a negative signal reverses the motion. The extent of the speaker movement (its excursion) is proportional to the voltage component of the waveform. The supplied current must be high enough to move the low impedance speaker coil in the magnetic structure, thereby providing the required mechanical force to keep the speaker on track with the waveform.

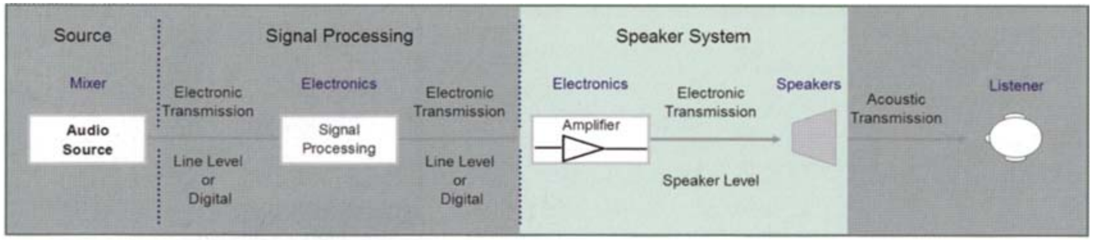


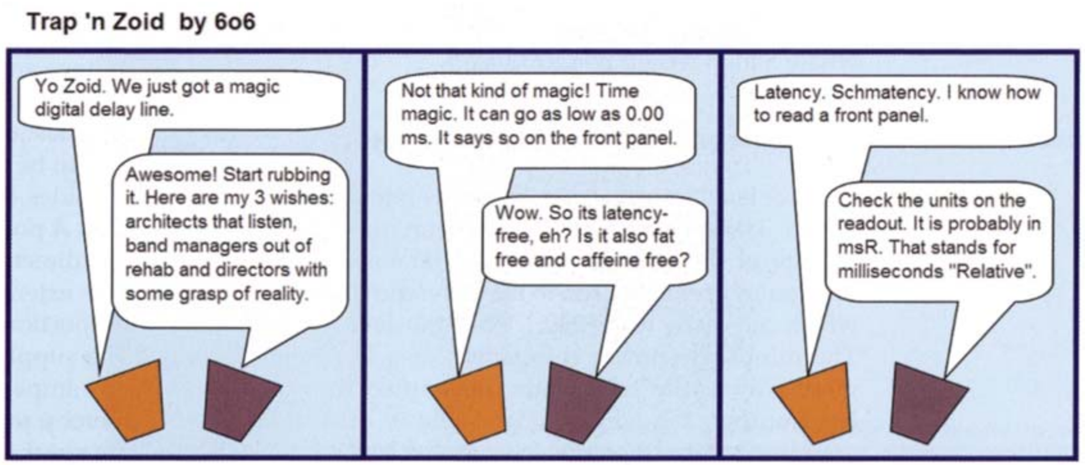
Figure 1.29 The speaker level analog electronic transmission path from the amplifier to the loudspeaker

Power and Impedance

Component power amplifiers have several key parameters from the design viewpoint: the maximum power (wattage) and the minimum impedance. The maximum wattage matches the amplifier-operating range to the recommended range of the speaker. Amplifiers have fairly standardized ratings in these terms, while the speaker specifications are much less clear cut. The matching of amplifier to speaker is best accomplished by following the guidelines given by the speaker manufacturer.

The minimum impedance governs how many speakers can be loaded onto the output terminals before fire accompanies the signal. Most amplifiers claim to be able to operate down to 2 ohms. This allows up to four 8 ohms speakers to be connected in parallel. The temptation is high to load an amplifier down to 2 ohms because we are able to drive the maximum number of speakers for the price of a single amplifier. This is rarely done, for two reasons. First, the sound quality is seriously degraded due to the reduced damping factor control of the load that occurs when amplifiers must deliver such massive amounts of current. Secondly, the amplifiers tend to live very short lives.

The minimum standard operating impedance for component amplifiers is 4ohms. This will take two 8ohms drivers or four 16 ohms high-frequency drivers. Since an amplifier costs the same whether you use it at 8 ohms or 4 ohms the second speaker is essentially driven for free. In the case of HF drivers you get four amplifier drives for the price of one. This is serious money, but it leads to serious compromise by limiting system subdivision. This is the "econohmic" factor of component amplifiers. Our designs will need to be defended against "value engineering" that forces us to operate multiple drivers at matched levels.



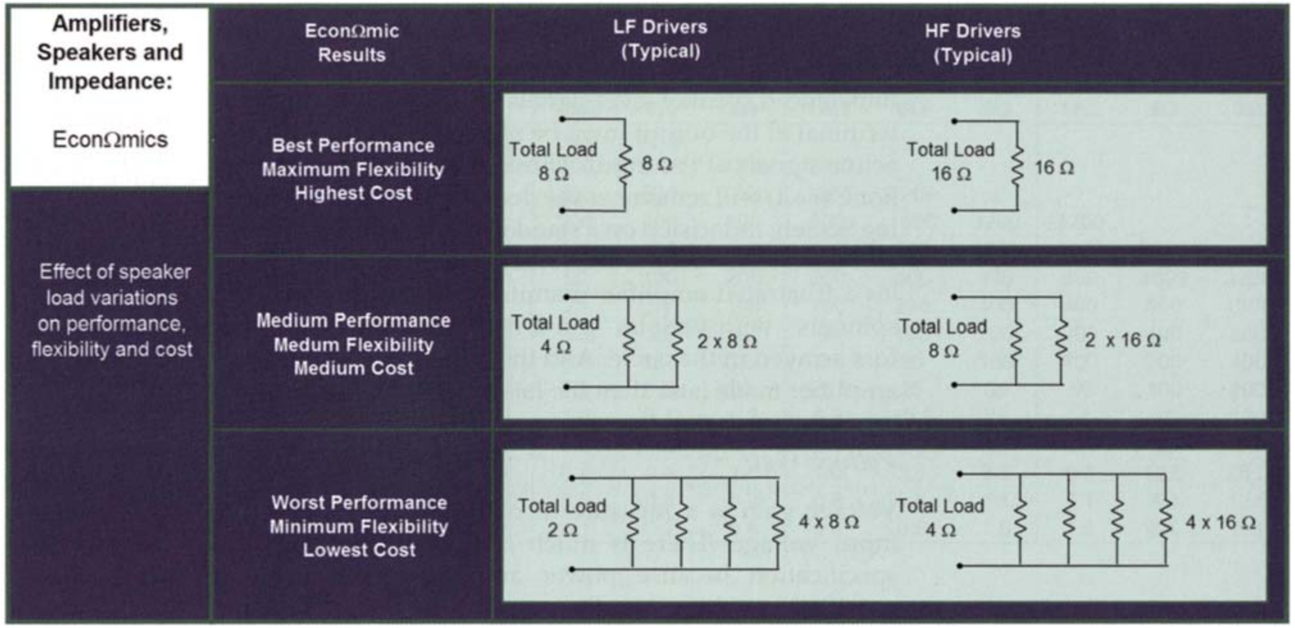


Figure 1.30 The considerations of impedance, performance and cost related to speaker loading of the power amplifier

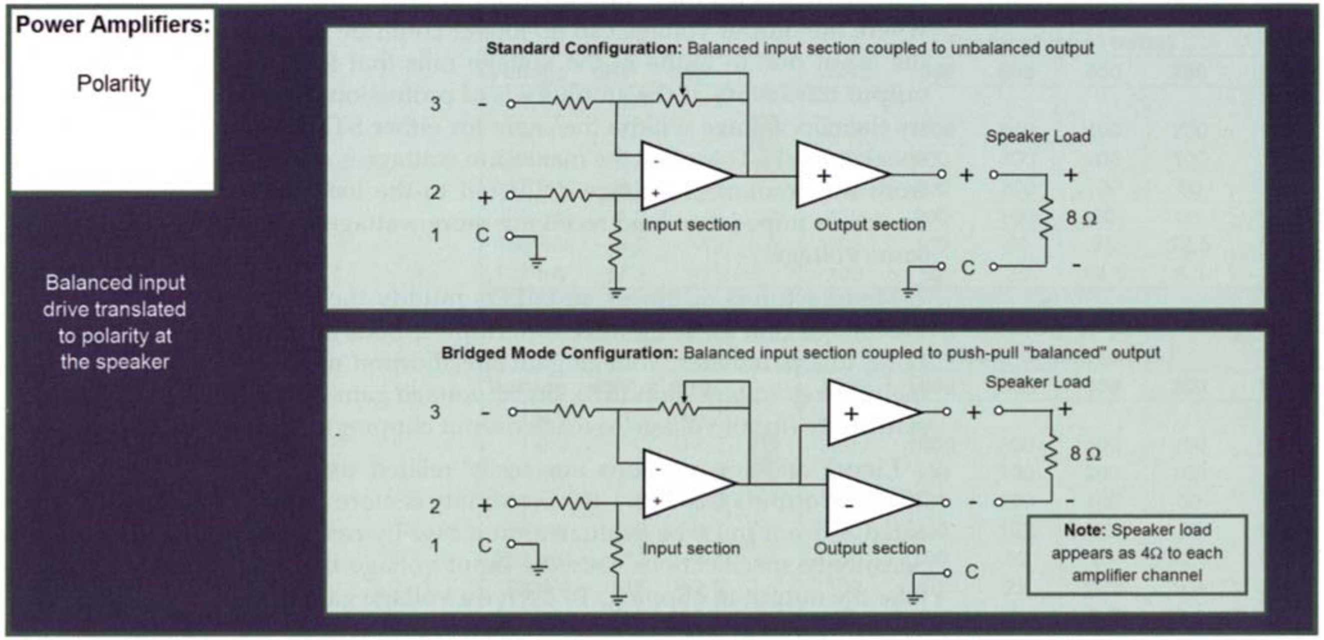


Figure 1.31 Flow block of amplifier polarity for standard and bridged mode configurations

Polarity

Amplifiers take balanced line level signals in and put unbalanced speaker level signals out. Therefore, the "hot" terminal at the output must be matched to one of the two active signals at the input. Those of us with gray hair (or none at all) will remember the decades of Audio Engineering Society indecision on a standard for the question of pin 2 or pin 3 "hot." At the 1985 AES convention in Los Angeles a frustrated amplifier manufacturer distributed "AES spinners" with various "pin 2 hot" and "pin 3 hot" sectors arrayed in the circle. And the winner is .. . pin 2. Any amplifier made later than the late 1980s will be pin 2 hot.

Voltage Gain

Voltage gain is a measure of output voltage related to input voltage. There is much confusion regarding this specification because power amplifier input drive is specified in volts, while the output is in watts. The input and output voltages rise together in a fixed linear gain relationship until clipping is reached. That fixed gain ratio is termed the amplifier voltage gain. Clipping is the point where the output voltage can no longer continue to track the input due to limits in the voltage rails that feed the output transistors. If the amplifier is of professional quality the clip voltage will be the same for either 8 ohms or 4 ohms speaker loads. Therefore the maximum wattage is derived from the amount of voltage delivered to the load, with the lower impedance load receiving more wattage for the same voltage.

Manufacturers of power amplifiers muddy the waters on voltage gain by using three differing methods of specifying this parameter. Voltage gain specification methods include log voltage gain (dB), linear voltage gain (X) and sensitivity (input voltage to reach output clipping).

Linear and log versions are easily related using the 20x log formula (see Fig. 1.10). Sensitivity is more complicated, since it must be evaluated on a case-by-case basis. Sensitivity specifications state the input voltage that will take the output to clipping. To derive a voltage gain from this, we will need to know the output clip voltage. If the amplifier clips at 10 volts and the sensitivity is 1 volt then the amp has 20 dB (10 X) voltage gain.

Here is where it gets confusing. Some manufacturers use a standard sensitivity for all amplifiers, regardless of maximum power rating. A 100 watt and a 400 watt model will both reach full power when driven at the input level of, for example, 0.775 volts. This is done by giving the amplifiers 6 dB difference in voltage gain. Other manufacturers offer standard voltage gains. In those cases the 400 watt amp will require 6 dB more drive level to reach full power. Still others can't make up their mind and change from model to model.

Which is better? A standard drive level that brings all amps to clipping, or a standard voltage gain that has all amps tracking at the same voltage level?

The most important factor is that the gain be known. If the amplifiers are sensitivity-based we can deduce their voltage gain. If they are voltage-gain-based the job is done for us. Either way they can be made to work sufficiently by adjustment of the amplifier level controls.

The answer to which is better lies in two places: the frequency dividers and the limiters.

When we have frequency dividers ahead of the amplifier the signals will need to be recombined acoustically. If amplifiers have matched voltage gain, the job of setting the acoustical crossover is made much easier. For example, if we are using a dedicated speaker controller, the device is designed with the assumption of matched amplifier gains. If the gains are not matched then the acoustical crossover settings will not be transferred at the intended relative levels, causing a shift of crossover frequency (see Figs 2.29 and 2.37). This is also true if we are using factory-recommended settings programmed into a DSP.

Why would the voltage gains be unmatched? This would occur in a sensitivity - based amp with high and lower drivers being driven by amplifiers with different maximum power. As in the example above, a 400 watt low-driver amp paired with a 100 watt HF driver amp would have 6 dB of difference in gain at the crossover if the amps were sensitivity-based. A 12 dB per octave crossover point could shift by half an octave upward!

Amplifier Voltage Gain Reference

Input Drive Level				Voltage Gain=20 dB Power at load (watts)			Voltage Gain=26 dB Power at load (watts)			Voltage Gain=32 dB Power at load (watts)		
Voltage	dBV	dBu		16Ω	8Ω	4Ω	16Ω	8Ω	4Ω	16Ω	8Ω	4Ω
16 V	24.0	26.2		1600	3200							
11 V	21.0	23.2		800	1600	3200	3200					
8.0 V	18.0	20.2		400	800	1600	1600	3200				
5.6 V	15.0	17.2		200	400	800	800	1600	3200	3200		
4.0 V	12.0	14.2		100	200	400	400	800	1600	1600	3200	
2.8 V	9.0	11.2		50	100	200	200	400	800	800	1600	3200
2.0 V	6.0	8.2		25	50	100	100	200	400	400	800	1600
1.4 V	3.0	5.2		12.5	25	50	50	100	200	200	400	800
1.0 V	0.0	2.2	X	6.3	12.5	25	25	50	100	100	200	400
707 mV	-3.0	-0.8		3.1	6.3	13	12.5	25	50	50	100	200
500 mV	-6.0	-3.8		1.6	3.1	6.3	6.3	12.5	25	25	50	100
356 mV	-9.0	-6.8		0.8	1.6	3.1	3.1	6.3	13	12.5	25	50
250 mV	-12.0	-9.8		0.4	0.8	1.6	1.6	3.1	6.3	6.3	12.5	25.0
178 mV	-15.0	-12.8		0.2	0.4	0.8	0.8	1.6	3.1	3.1	6.3	12.5
125 mV	-18.0	-15.8		0.1	0.2	0.4	0.4	0.8	1.6	1.6	3.1	6.3

Figure 1.32 Amplifier voltage gain reference chart

Amplifier Sensitivity Reference

Input Drive Level				.775 volt sensitivity by power rating Power at load (watts)					
Voltage	dBV	dBu		3200	1600	800	400	200	100
1.1 V	+0.8	+3							
775 mV	-2.2	0		3200	1600	800	400	200	100
550 mV	-5.2	-3		1600	800	400	200	100	50
337 mV	-8.2	-6		800	400	200	100	50	25
275 mV	-11.2	-9		400	200	100	50	25	12.5
168 mV	-14.2	-12		200	100	50	25	12.5	6.3
137 mV	-17.2	-15		100	50	25	12.5	6.3	3.1

Input Drive Level				1.0 volt sensitivity by power rating Power at load (watts)					
Voltage	dBV	dBu		3200	1600	800	400	200	100
1.4 V	+3	+5.2							
1.0 V	0	+3.2		3200	1600	800	400	200	100
707 mV	-3	-0.8		1600	800	400	200	100	50
500 mV	-6	-3.8		800	400	200	100	50	25
356 mV	-9	-6.8		400	200	100	50	25	12.5
250 mV	-12	-9.8		200	100	50	25	12.5	6.3
178 mV	-15	-12.8		100	50	25	12.5	6.3	3.1
125 mV	-18	-15.8		50	25	12.5	6.3	3.1	1.6

Figure 1,33 Amplifier sensitivity reference chart

The second factor is related to the use of limiters. Limiter action is based upon voltage level. Estimates of the power (wattage) at the speaker are based upon the voltage that is seen by the limiter circuit. If the voltage gain of the amplifier is unknown the limiter is uncalibrated. An uncalibrated limiter will either decrease the dynamic range unnecessarily or not provide the needed protection.

Level Controls

An additional case of non-standardization in the amplifier market is the various level control markings. As if it is not enough to have us confused about voltage gains let's add level controls with meaningless markings. It is not unusual to hear that the amplifiers have been set to "three clicks down." Huh? Marking schemes include: dB relative to maximum voltage gain (linear or log), blank with tick marks, the numbers 0 to 10 and there is at least one manufacturer that displays dB voltage gain. If a system is made up entirely of one model of power amplifier it is possible to use the markings for some relative calibration. If it is in "clicks" the calibration only holds if all are at the same click or if the clicks actually change the level in uniform dB increments. But since the different speakers have different power needs there is very little chance that a single amp will suffice. Once we introduce different models we have opened a Pandora's box of unrelated scales. Then we get to the system upgrade that has eight different amplifier models from three different manufacturers. Help!

It is also worth noting that the amplifier level control does not reduce the maximum power capability of the amp as one might believe from listening to too much audio folklore or watching the film *This is Spinal Tap.* Turning the level control down merely resets the drive level required to reach full power. Resetting the start position of the accelerator does not add horsepower to our engine. It only changes the position of our foot relative to our speed. So it is with amplifier level controls. As long as the amp level control is not turned down so low that the drive electronics will not be able to clip the amp we are fine. There is usually over 20 dB of excess headroom in commercial power amplifiers.

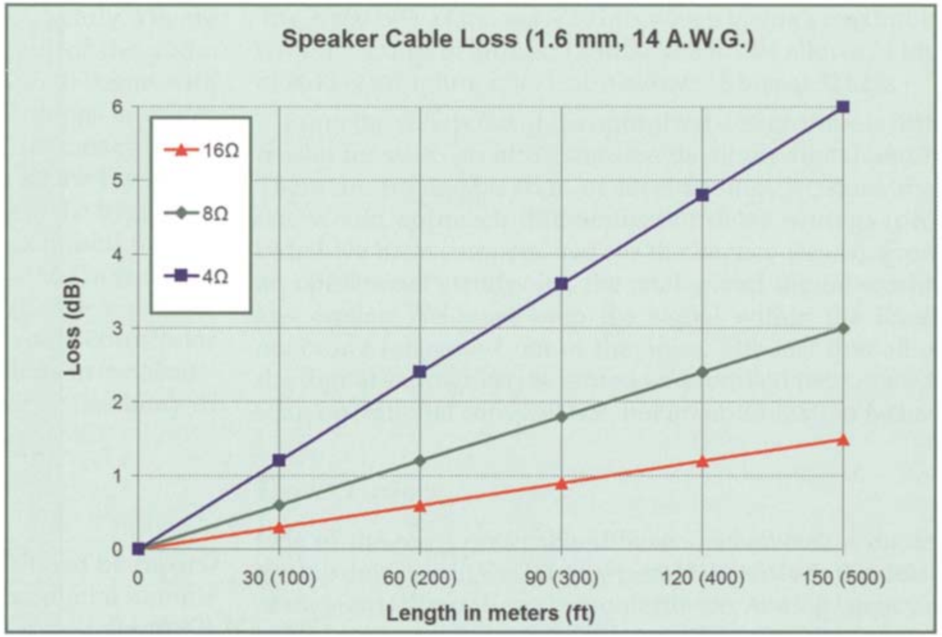
The reaction of most engineers has been to fear the lack of calibration in amplifiers and impose an edict of "all amps set to fully open." This is a regrettable yet understandable consequence of the lack of standards. It does, however, reduce our optimization options regarding relative level between system components. And it increases audible speaker system noise.

Speaker Level Interconnection — Speaker Cables

Relatively speaking, the connection between the amplifier and speaker is another case of a low-impedance output driving a high-impedance load. While 8 ohms would not seem to qualify as a high-impedance load, it is when compared to the extremely low 0.1 ohms output impedance of the amplifier. This familiar impedance scaling allows the amplifier to drive the speaker reliably. The low overall impedance allows lots of current to flow, and therefore lots of power.

Speaker cable runs are a different story than line level cables. They are prone to substantial amounts of resistive loss because the impedance of the cable can become significant in proportion to the speaker load. The loss rate depends primarily upon three factors: cable length, conductor diameter and load impedance. The loss increases with the cable length. Decreasing load impedance and cable diameter also increases the loss.

The relationship between amplifiers and speakers is driven by impedance economics. Amplifiers are capable of driving loads as low as 2 ft, which could be as many as four 8 ohms speakers wired in the standard parallel fashion. This can be driven by a single cable split to each speaker at the far end. On the other hand we can run each speaker from a dedicated amplifier (an 8 ft load) and cable. Every single performance factor goes in favor of the latter approach. The 2 ft load scenario has higher distortion, inferior control of the speaker load, lower reliability, higher cable loss and less power delivered to the load. On the other hand it is much cheaper: a case of econohmics in action. The most common load is the compromise value of a 4 ft load. A pair of 8 ft speakers or four 16 ft HF drivers are the most common scenarios. This seemingly benign situation can be a serious limitation to system flexibility. There are many



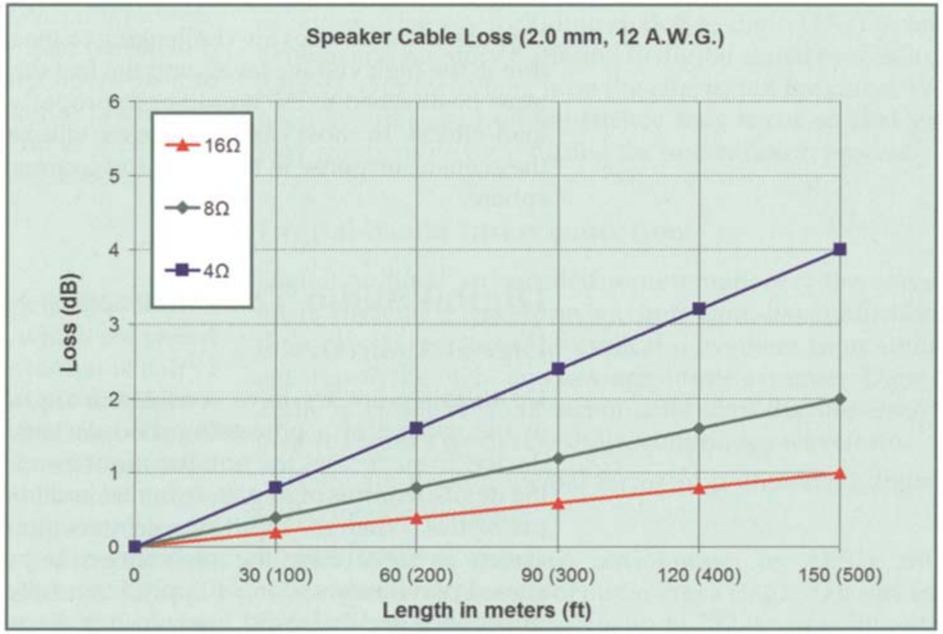


Figure 1.34 Charts of speaker level cable transmission loss (after Giddings, 1990, pp. 333-335)

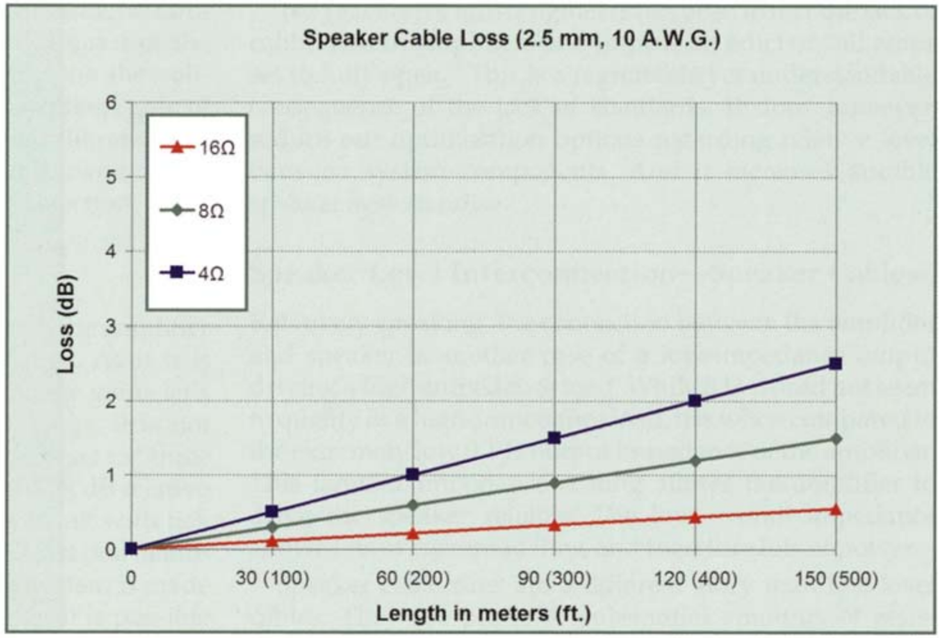


Figure 1.34 *(Continued)*

cases where being limited to one set of processing settings for four HF drivers is a serious performance degradation.

Speaker cable losses are challenging to measure directly, due to the high voltage levels and the fact that the speaker must be attached to the amplifier to properly monitor the load effects. In most cases the losses will be detected in the acoustic response in the room and compensated in that sphere.

Digital Audio Transmission

Digital Audio Devices

Up to this point we have considered digital audio only in the context of a processing module inside an analog device. In such cases the unit functions as a line level analog device in terms of its gain structure and interconnection properties. When the input signal enters our transmission system in digital form the interconnection properties discussed previously no longer apply. Forget about line level, impedance and balanced lines. In our discussion of line level we saw that current considerations were of minimal importance. Now even voltage is out. It has been converted into numbers. Welcome to the information age.

Digital audio transmission operates under a new set of rules. In fact, the term transmission as we have been using it is not fully applicable to digital audio. Analog transmission requires a medium, and digital audio is medium-independent. Whether the digital signal moves through optical fiber, wire or wireless Internet will not affect the frequency response, as with all other mediums. Digital audio transmission is a data transfer. If the data is transferred faithfully we will have a matched copy at both ends of the transference. This is never the case in analog transmission through a medium, which always has a loss.

This is not to say that there are no concerns regarding data transmission. There are a great number of opportunities for data damage. If it becomes corrupted it is unlikely to manifest itself as something that we can recognize as an "audio" problem. The more likely outcomes of digital transmission errors are no audio, audio from the planet Zircon or, worse yet, maximum level noise.

The world of digital audio is evolving rapidly. We are witnessing a steady erosion of the portion of the audio transmission path left in the analog domain. It began with delay lines. Yes, there actually were such things as *analog* delay lines. Next came digital equalizers, frequency dividers and onward. Is there any doubt that in the future we will have an analog-to-digital conversion at the beginning of the transmission path and remain digital until the last possible link in the chain to our analog ears? These developments, however significant, do not change our basic mission: monitor the transmission and enact corrections as required. Since most of the damage is done in the acoustical world, the digital age offers us superior flexibility of corrective tools, but not a new world order.

Bandwidth

The digital audio bandwidth is not determined by capacitance and inductance, but rather by the digital sample frequency. The most common digital transmission sample frequencies are 44.1kHz, 48kHz and 96kHz with 192kHz becoming more common. The frequency range upper limit is no more than half of this sample rate. We might wonder why we would need anything more than bandwidth above 20kHz. The reason is that the sharp filters used to limit the upper range in digital systems cause phase shift below their corner frequency. As the corner frequency rises, the amount of in-band phase shift decreases.

Dynamic Range

The dynamic range is determined by the bit resolution. The analog signal is encoded voltage values, which are stored as bits. Each represents a voltage threshold that is half of the previous. The top of the dynamic range is the highest number and is termed full-scale digital. It is important to note the relationship of full scale digital to analog voltage is not fixed and can range from +20 dBV down to OdBV The number of bits, i.e. the resolution, determines how far down we go before the encoding process no longer differentiates the signal. Each additional bit gives us an additional 6 dB of signal to work with at the bottom of the range. The AES/EBU standard is 20 bits which yields a maximum dynamic range of around 120 dB. AES3-2003 allows 24 bits at 96kHz sampling rate. It also allows 16 bits at 32kHz.

From the viewpoint of the optimized design there is little reason for us to go into extensive detail on digital audio. There are no equalization or level setting decisions that one would approach differently, nor delay settings (provided we have compensated for the latency issues). From an operational standpoint, the analog and digital worlds are similar. We must keep the signal within the linear operating range and out of the noise. The fact that all of the digital settings can be stored and recalled (and erased) is an operational convenience, not an optimization factor.

Device Latency

One of the most noticeable differences between a digital analog device is in the latency period. Latency is the delay of the signal through an electronic device. Analog latency is measured in nanoseconds (10^{-9} seconds), which is so small that it is considered negligible compared to digital audio which is several milliseconds at best. There are a variety of causes of latency including analog-to-digital (A/D) conversion and memory buffering in digital signal processing devices. Our focus here is on the effects, not the cause. We need awareness of where latency may occur so that we may compensate for it during the optimization process.

Digital Audio Interconnection

Digital audio is an encoded representation of the waveform. The analog waveform is a continuous linear function: there are no breaks in the cyclical movement from ambient state to the high- and low-amplitude extremes. Digital audio is a finite series of incremental steps that are assembled to create a rendering of the continuous waveform.

There are three principal forms of professional digital audio interconnection:

1. AES/EBU (a standard also known as AES3): this encodes two channels of audio into a single data stream (or connection). It supports up to 24 bit sampling with

preferred sampling frequencies of 44.1 kHz, 48 kHz and 96 kHz. The standard connection is made with a shielded pair cable and XLR connectors. The signal is balanced and requires a cable with a 110 ohms characteristic impedance. Optical transmission is accomplished by the F05 connector. S/PDIF is an unbalanced consumer variant.

2. Network: there are various network protocols, the most common being Ethernet/Cobranet. Networked interconnections follow the standard wiring practices of the computer industry. "Thank you for calling the technical support hot line. Your call is important to us ... "
3. Proprietary: manufacturer-specific interconnection that does not conform to any standard (or theirs alone) and only connects to their own equipment. See manufacturer's specifications for details.

Each of the transmission interconnection paths has limited cable length. The principal factor limiting the length is the type of cable used. Unlike cables used for analog signals, cables used for digital signals are built to have controlled impedances, and it is essential that a cable has the proper characteristic impedance that matches the system being used. By driving these cables properly at the transmitter end, and terminating them correctly at the receiver, they will move the digital data streams around safely. All of this driving and receiving is built into the equipment we are using, so all we have to do is supply the correct cable type. The routine practice of splitting signals in the analog world is not advisable for digital interconnections. Very short cable runs may continue to work in such cases, but this is a risky practice.

The data stream is a series of binary states: "l's" and "0's" that are represented by voltage levels. There is no tolerance for half states. The receiving end of the digital transmission line must differentiate the l's from the 0's, a task that is made more challenging as distortion of the signal increases. Using the wrong type of cable results in distortion that limits the distance the signal can travel before becoming unreliable. As the distortion increases, the decoder can no longer reliably discern the data states and errors are generated that can result in dropouts and audible distortion.

Network Latency

Transmission of analog electronic signals travels in a continuous stream at two-thirds the speed of light. Digital audio is transmitted in packets of data. The speed of the transfer is network-dependent. The interconnection between devices and over digital audio networks is an area full of opportunities for differences in latency delay. This is an area that will require careful monitoring during the verification of the system (see Chapter 9).

Access to Information

The analog audio world contains multiple interconnections. This is a mixed blessing. On the negative side are the myriad opportunities for interconnection errors, ground loops and component failures. On the other side is the fact that we can monitor the signal at various points along the way, re-route it if necessary and, most important in our case, we can measure it. Digital audio systems in general, and networked systems in particular, tend to leave us unable to monitor the signal en route. The reason is simple: to bring a signal out of the system requires a digital-to-analog converter, a connector and so on. This drives the cost up. Since our interest here is to design systems that can be fully optimized it is vital that the signal-processing parameters, i.e. equalization, delay and level setting, be visible to our analyzer. Any system that leaves us fenced out of the signal path is less able to be fully optimized, and is therefore less desirable. The user interface renderings of EQ, level and delay should *never* be trusted on blind faith. They must be measured. The criteria for access to measurement are detailed in Chapters 8 and 9.

Acoustic Transmission

Power, Pressure and Surface Area

Acoustic power operates under the same conditions as the electrical power discussed previously, with analogous aspects. Power in this case is acoustic power and is measured in watts. Sound pressure is analogous to voltage,

Perspectives: As the need for multisource, multiple isolated zone line arrays and surround sound increases, the need for true system design and system tuning will also augment. Often misused, most of the time mistrusted and always undervalued, the system, tuner, system engineer, PA tech., call it what you will, is bound to become one of the most important parts of the audio team. Without a system tuned with precision, there cannot be a perfect mix in a live show,

Francois Bergeron

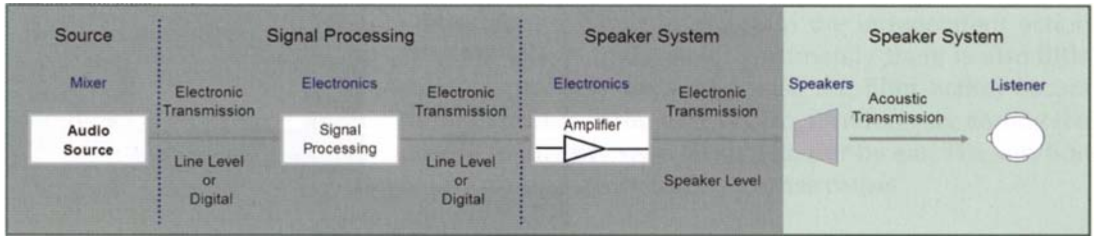


Figure 1,35 Acoustic transmission from the speaker to the listener

and is measured in dB SPL (sound pressure level). Surface area is like current. The acoustic watt could be the product of high pressure in a small area or low pressure in a large area. To illustrate this, let's consider a favorite source of acoustic power, fireworks.

An explosive charge pressurizes the air in direct contact with burning powder and propagates spherically outward from the point of origin. The amount of acoustic energy (power) created by the explosion remains constant as the sound wave moves outward from the source. The outward movement stretches the surface area of the wave. Since the quantity of energy is fixed, it is not possible to maintain the same pressure level while expanding the surface area. As the surface area increases, the pressure decreases proportionally, leaving us with the original quantity of acoustic energy. Does this mean that it does not matter how close we are to the explosion? Of course not. If we get too close we will damage our ears, or worse. But the damage is not due to the acoustic power. It is a consequence of excessive sound *pressure.*

Our ears are sound pressure sensors. There is no way for us to sense acoustic power. We would have to be spread over the entire surface area to experience it. We cannot detect the presence of sound anywhere else but our ears. We only hear reflected sounds because their paths lead them to our ears.

The quantity of acoustic power a device generates is of interest to the loudspeaker designers but few others spend time considering it. A loudspeaker, like the fireworks charge, has a fixed amount of acoustic power at its source. From the perspective of audio engineers, the measure that matters for sound system performance is **sound pressure.** How much sound pressure level can we deliver to our ears? However, if we are to concern ourselves with creating the desired sound pressure at locations other than the exclusive confines of the mix position, we need to be very conscious of the surface area aspect of sound propagation. As mentioned earlier, sound propagates spherically from the acoustic source. If the audience members are all located equidistant from the source, they will experience the same sound pressure. If not, it will be necessary to create an asymmetrical acoustical power source that can maintain constant sound pressure level into the locations that are more distant. This is done by steering the sound so that the surface area is decreased, thus allowing for higher pressure at a given distance. Our choice is simple. Either we construct our listening environments such that all members of the audience are equidistant or we learn how to steer the sound.

We are finished with the concept of acoustic power. From here forward we will focus on sound pressure level spread over the surface area. It is common, however, for audio engineers to use the term "power" and SPL interchangeably. To the extent that this may occur in this text, please bear in mind that we are referring to pressure.

As sound propagates away from the source, the SPL decreases at a rate 6 dB for each distance doubling. This rate holds for a source in a free field (reflection-free) acoustic environment. This loss rate is known as the **inverse square law.** In the presence of reflections the SPL will drop off at a lower rate. The effects of reflections on dB SPL loss

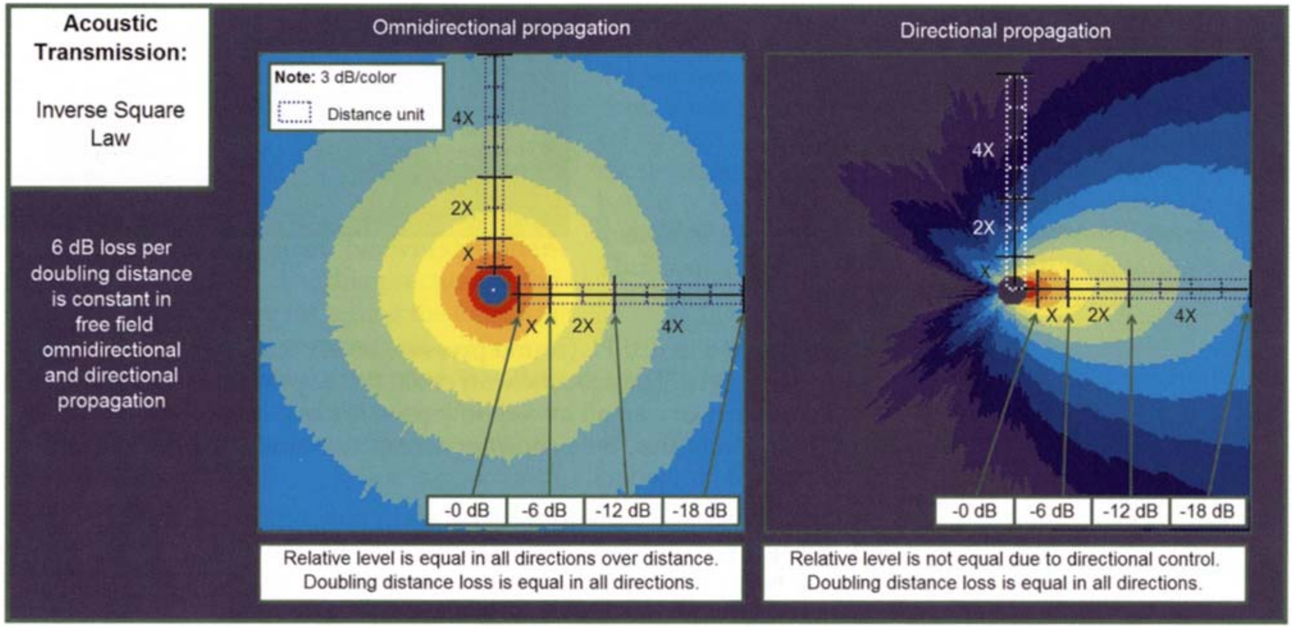


Figure 1,38 Spherical sound propagation: (left) omnidirectional source; (right) directional source

are not consistent over frequency or over location in the listening area. These complications make it impractical to go about computing SPL loss rates for every frequency at each position. For our purposes then, it is assumed that the loss will occur at the free field rate.

Environmental Effects: Humidity and Temperature

There are additional factors above and beyond the inverse square law that affect transmission loss. Air is a non-linear transmission medium, i.e. the high frequencies are attenuated at a greater rate than lows. The highest frequencies are the most strongly affected with the losses gradually decreasing as frequency falls. The losses accumulate over distance; therefore concerns in this regard become quite significant in long throw applications. In the near field, a speaker's very high frequency (VHF) range extension is at its maximum. As we move further away the VHF area will become low-pass-filtered. As more distance is traveled, the corner frequency of the filter action moves downward.

Most audio engineers are familiar with the environmental effects on their sound systems. The sun goes down and the system seems to undergo radical changes in the high-frequency range. Battle stations!

There are three factors whose combined effects create the values of transmission loss through air as a medium. They are distance, humidity and temperature. These mix together in some unexpected combinations, much like weather in general. For the most part the rate of HF loss is greatest as the humidity falls. This holds true as long as the temperature is moderate and humidity levels are between 30 and 70 per cent. This would be sufficient for indoor venues. At cold temperatures the dry air has the best transmission capability while the reverse is true in hot weather. The loss values over temperature and humidity can be found in Fig. 1.38.

Perspectives: A long time ago, during the ancestral times when rigging was a very unusual and expensive option, my friends and I were in charge of the sound of a large outdoor festival on the banks of Lac Lemon in Switzerland. This setup was very directional in the vertical plane and fine tuning of the overall tilt angle made by the stacks with wooden cleats was necessary. We never measured this PA with relevant machines (that did merely exist in 1984) but a simple prediction using a good computer program today will show that there is no coherent energy outside of the disc that has the thickness of the stack; i.e. 2.6 meters. We had to cover about 100 m long so intuitively it was not convenient to place the PA too high above the audience so we had a chance to reach the back seats. The cabinets on top started at 2.5m from the floor above. During the day we had to deal with all the bands that wanted to do their sound check and we experienced quite a good sound, very promising for the evening show. The weather was cool, not too hot under the sun and fresh in the shadow But when the night came it was almost cold and the vicinity of the lake made the air change of temperature very quick from 30°C to 15°C.

The audience had been attending the festival since late afternoon and things came to a peak for the last two bands of the night: thirty thousand enthusiastic fans ready to absorb the music as an enormous sponge. After a long change over on stage the show began at night and astoundingly, we had lost some 12 dB SPL at the desk and some places even more over the audience. We checked every possible thing but we

dBSPL loss over Distance Reference

dBSPL loss (relative to 1m)	Distance: Meters (m)	Distance: Feet (ft)
SPL 0	1	3.3
-3.0	1.4	4.6
-6.0	2.0	6.5
-9.0	2.8	9.3
-12.0	4.0	13
-15.0	5.7	19
-18.0	8.0	26
-21.0	11	37
-24.0	16	52
-27.0	23	74
-30.0	32	105
-33.0	45	148
-36.0	64	209
-39.0	90	296
-42.0	128	418

Figure 1.37 Sound propagation loss over distance

Weather conditions can have discernible effects upon the directional transmission of sound as well. Temperature gradients can separate the atmosphere into layers, thereby providing a refractive component to the sound transmission as it encounters the change in temperature. This response is well known in the field of underwater acoustics, where the hunting of submarines is made more difficult by the refractive effects of layers of water at different temperatures. In airborne sound this is often experienced as the arrival of a distant sound at a much higher than expected level, such as a train whistle on a foggy night. In the world of outdoor concert sound this can cause unexpected vertical redirection of the speaker system transmission.

It is very difficult to ascertain the independent action of these factors in the field. Fortunately there is also little need to precisely do so. Because the filter action occurs almost entirely in the VHF region, its effects are among the easiest to identify with an analyzer or by ear. The solution is the same regardless of its precise cause.

Acoustic Transmitters: Loudspeakers

We have reached the end of the electronic transmission chain: the loudspeaker. The role of the speaker is delivery of the original transmission to the listeners. In the ideal world we might be able to do the job with a single full-range speaker capable of creating the surface area shape that spreads even SPL over the listening area. There are a limited number of shapes that a single speaker can create. Therefore, the likelihood of a perfect fit is poor. Most applications have complex shapes that will require multiple speakers to evenly spread the SPL to conform to the particular shape. But before we can begin to attempt this we will need to investigate the transmission characteristics of a single loudspeaker and create a framework to evaluate its behavior.

The first item to establish is that we will not discuss loudspeakers at all. We will only discuss loudspeaker *systems.* This book is not targeted to the research scientists of professional loudspeaker manufacturers. Nor is it intended to aid garage scientists who wish to find the ultimate mix of components and turn lead into gold. The age of audio alchemy is long gone. We use engineered systems, i.e. speakers in tuned enclosures with documented characteristics, repeatable construction and professional quality standards.

Let's begin with a generic specification list of the expected characteristics of loudspeakers.

Professional loudspeakers system generic specifications:

1. Amplifier(s) shall be capable of driving the system to full SPL without input stage overload.
2. Systems shall have known operational limits. Amplifiers and limiters shall be calibrated such that

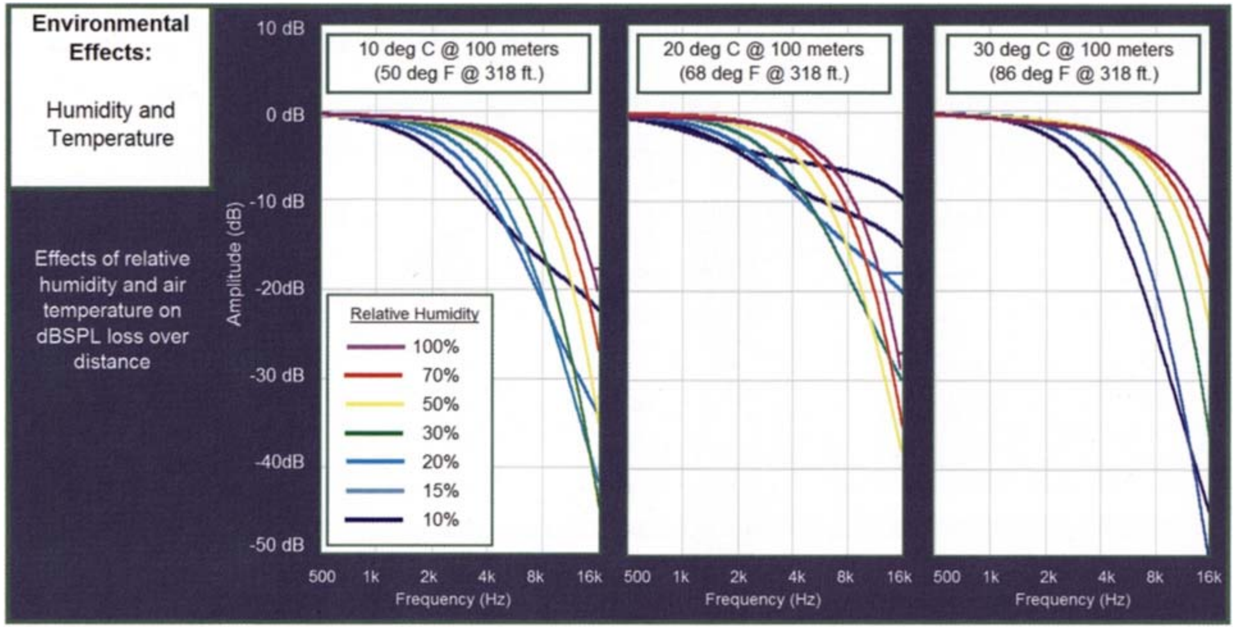


Figure 1.38 Transmission loss due to air absorption over temperature and humidity at: (left) 10°C, (centre) 20°C, (right) 30°C. Conversions are approximately 50, 68 and 86°F at a distance of 318 feet (courtesy of Meyer Sound Laboratories Inc.)

just could not understand what was happening!

An hour later we were still experiencing the same trouble and we received a phone call from a farmer living 10km away in the mountain complaining that we were too noisy and he could not sleep! The sound never reached the audience, but went up in the air so well that it disturbed somebody very far away. How could this be? We never changed anything in the settings and position of the PA from the afternoon to the evening.

The extreme vertical directional characteristics of our PA turned out to be a weakness. Thermal change between the audience layer (hot = faster propagation) and close above (cold = slower propagation) had done a planar acoustical diopter that bent

the speakers may be operated within the limits without fear of damage. System shall be self-protecting.

3. System shall be capable of reaching maximum SPL with graceful overload characteristics.
4. System shall be low noise, with dynamic range > 100 dB.
5. Frequency range of 70 Hz to 18 kHz to be covered as a two-way system (minimum), four-way system (maximum) in a single enclosure. Low-frequency range may optionally extend below 70 Hz.
6. Frequency range of 30 Hz to approximately 125 Hz to be covered as a single subwoofer system in a separate enclosure.
7. Frequency range of approximately 60 Hz to 160 Hz to be covered as a single mid-bass system in a separate enclosure.
8. THD <1 per cent over the frequency range at SPL levels within 12 dB of maximum SPL.
9. Free field frequency response ±3 dB over the frequency range of the device.
10. Coverage pattern shall either maintain beamwidth as frequency rises or narrow as frequency rises. Beamwidth shall not narrow in the low-frequency range or mid-range and then widen in the high-frequency range.
11. Acoustical crossover between drivers in a single enclosure shall be phase-aligned with fixed level relationship.
12. Acoustical crossover between drivers in separate enclosures shall have variable phase and level adjustments to compensate for variable relative quantities and placement.

These specifications provide a common basis for our consideration of loudspeaker systems. After these have become familiar we will move toward differentiation of the types of loudspeakers and their respective applications.

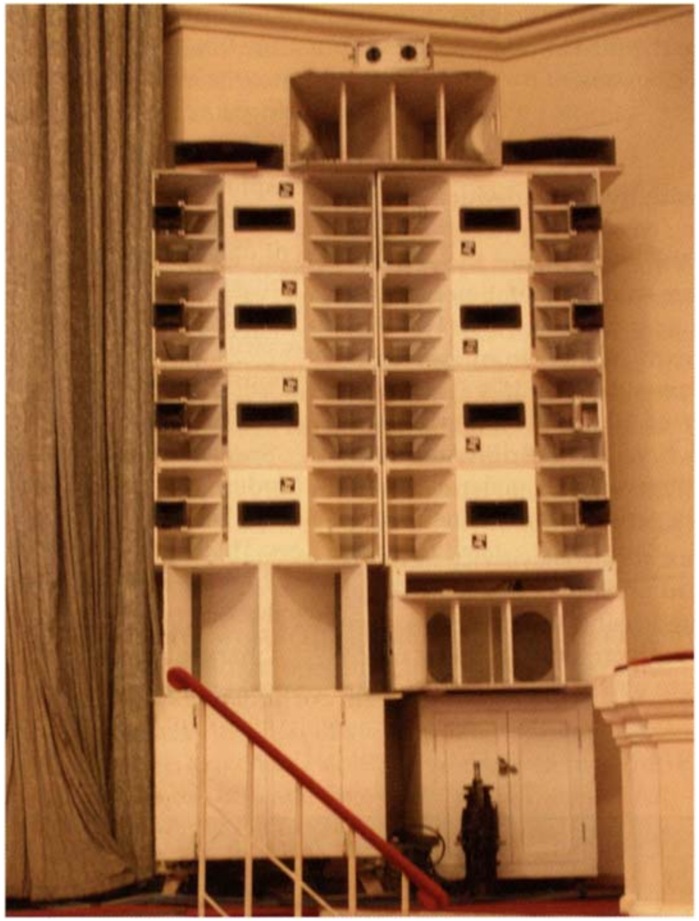
Transmission Transition: Electronic to Acoustic

Before we can expect to hear sound coming from our loudspeaker system we must deliver the signal from the

the propagation upwards so that the aim axis for the audience was under the refraction angle: All the energy was reflected on this diopter, just like the image you can get from underwater through the water surface.

In order to compensate for this phenomenon we should have flown the PA and aimed it downward so the incidence angle is larger than the refraction angle. This is what we did the next years and later on. Since that day I always warn users that are doing outdoor shows to he sure they can fly the PA; then they will not have to fight against nature since they already have a lot to do with musicians.

Marc de Fouquieres



This loudspeaker escaped from its garage. This is *not* considered an engineered system (photo courtesy of Dave Lawler)

electronic signal chain. Our goal is to ensure that we can get the maximum level out of the acoustical system while still operating the electronics within their limits. This requires a conversion between the systems. We will exchange electrical volts for acoustic pressure: dBV for dB SPL. The mix console is at the head of the electronic chain. The mix engineer's most obvious concern is this: how much SPL can we get out of this console? That depends on how far you drop it.

How do we bridge the gap between the dBV and dB SPL worlds so that we can operate the console in its linear range and ensure that maximum pressure is obtained from the system? The complicating factor here is the presence of transducers in the transmission chain between the electronic and acoustical worlds. Transducers convert energy from one domain to another. Microphones and speakers are our most well-known transducers. They convert acoustical energy to/from mechanical, magnetic and electrical energy, in opposite orders. The domain of mechanical and magnetic energy conversion will be left to the designers and manufacturers of microphones and loudspeakers. We will abridge the matter to an electronic/acoustic transduction, also known as electro-acoustic transmission. How much acoustical pressure corresponds to a particular voltage? How much voltage correlates to a particular SPL value?

Sensitivity

One method of expressing this is termed the **sensitivity.** This parameter bridges the gap by including both expressions. For microphones the standard sensitivity is given in mV/pascal. What is a pascal? Ninety four dB SPL, of course. If a standard SPL level is found at the microphone diaphragm (94 dB SPL) then the voltage at the terminals determines it sensitivity. This is accomplished by a microphone calibration device, known as a pistonphone, which generates a tone and couples to the microphone capsule. For example, the Danish Pro Audio 4007 microphone has a sensitivity of 2.5 mV/Pa. This specification is based only upon the open circuit voltage at the microphone terminals. Because microphones generate both current and voltage, a second sensitivity rating can also be given that factors in the mic's output impedance. This is the "power level" and is specified for a relationship to a standard of O dB = 1 mW/pascal. Typical microphone power levels run in the −60 dBV to −40 dBV range. Two mics with the same open circuit voltage and different output impedances will have unmatched power level specifications. This complicates our prospects of matching microphones. Fortunately this is not a concern for us. Since our primary application for microphones is acoustic measurement we will not be loading down the microphones by splitting them to monitor mix consoles and/or recording trucks. Therefore we will be fortunate to be able to use the far simpler open circuit voltage as our sensitivity specification.

1 watt/1 meter

A full-range loudspeaker sensitivity can be deduced by the inversion of the microphone sensitivity concept. A fixed amount of power is sent to the speaker and the acoustic output is measured. The common representation of this is one watt input drive at one meter distance (1 W/1 m). The SPL generated can then be extrapolated from there provided we know how much power the amplifier is sending. A speaker with a sensitivity of 100 dB (1 W/1 m) would create 110 dB when driven at 10 W and 120 dB with 100 W. This does not mean that it will necessarily generate 130 dB when driven at 1000 W. It may generate smoke instead.

A secondary nuisance factor to the sensitivity figure is that it is rated for the standard one meter distance. Other distances must be extrapolated using the inverse square law to determine the SPL at a given location. How much level can this speaker generate at the mix position? Take the sensitivity value and extrapolate that to the maximum rated wattage value, then apply the inverse square law distance loss to the mix position. This provides some insight as to why this figure has limited appeal.

There's more. The 1 W / 1 m rating fails to factor in the amplifier voltage gain, since it is based on the output power only. This means that speakers driven with different amplifier settings still have matched 1 W / 1 m values. Things get more complicated for actively biamplified (or triamplified, etc.) speaker systems since 1 W will not necessarily appear at the output of both amplifiers simultaneously. Some manufacturers provide separate sensitivity specifications for each driver.

Speaker sensitivity is a holdover from a bygone era. Modern professional audio systems select the power amplifier based on the maximum capabilities of the drivers. The levels are set by how the systems combine acoustically at the acoustic crossover, not by comparing sensitivity values on paper.

Chances are high that we will be listening to an array of speakers out there at the mix position. Modern arrays are

Speaker Sensitivity Reference

Amplifier Power Rating			Calcuted SPL at 1m 97 dB 1w /m			Calcuted SPL at 1m 100 dB 1w /m			Calcuted SPL at 1m 103 dB 1w /m			Calcuted SPL at 1m 106 dB 1w /m			Calcuted SPL at 1m 109 dB 1w /m			Calcuted SPL at 1m 112 dB 1w /m		
16Ω	8Ω	4Ω	16Ω	8Ω	4Ω	16Ω	8Ω	4Ω	16Ω	8Ω	4Ω	16Ω	8Ω	4Ω	16Ω	8Ω	4Ω	16Ω	8Ω	4Ω
512 W	1024 W	2048 W	124	127	130	127	130	133	130	133	136	133	136	139	136	139	142	139	142	145
256 W	512 W	1024 W	121	124	127	124	127	130	127	130	133	130	133	136	133	136	139	136	139	142
128 W	256 W	512 W	118	121	124	121	124	127	124	127	130	127	130	133	130	133	136	133	136	139
64 W	128 W	256 W	115	118	121	118	121	124	121	124	127	124	127	130	127	130	133	130	133	136
32 W	64 W	128 W	112	115	118	115	118	121	118	121	124	121	124	127	124	127	130	127	130	133
16 W	32 W	64 W	109	112	115	112	115	118	115	118	121	118	121	124	121	124	127	124	127	130
8 W	16 W	32 W	106	109	112	109	112	115	112	115	118	115	118	121	118	121	124	121	124	127
4 W	8 W	16 W	103	106	109	106	109	112	109	112	115	112	115	118	115	118	121	118	121	124
2 W	4 W	8 W	100	103	106	103	106	109	106	109	112	109	112	115	112	115	118	115	118	121
1 W	2 W	4 W	97	100	103	100	103	106	103	106	109	106	109	112	109	112	115	112	115	118
0.5 W	1 W	2 W	94	97	100	97	100	103	100	103	106	103	106	109	106	109	112	109	112	115
0.25 W	0.5 W	1 W	91	94	97	94	97	100	97	100	103	100	103	106	103	106	109	106	109	112

Figure 1.40 Speaker sensitivity reference

complex mixtures of active multiway speakers with different drive levels, different amplifiers and different amounts of acoustic addition over frequency. Again we ask the question: "How many dB SPL can I get out of the console?"

Fortunately there is a better way.

dB SPL/volt

If the amplifier and speaker are viewed as an integrated system, the sensitivity can be viewed in a context relevant to our modern systems with line level console and signal-processing drive levels. The first question is: "Can the amplifier/speaker system be driven to full power by the drive electronics with a reasonable amount of headroom?" The second question is: "How loud will the system get at a given location in the room when driven at a nominal level?" The answer can be found with a modernized sensitivity value which denotes the dB SPL value generated by the speaker when driven at line level: dB SPL/volt.

How does this work? The electronic transmission components have standard operating level centering around 1 volt (0 dBV) and ranging to a maximum of 18 to 24 dB above that. How many dB SPL will the amplifier / speaker system generate when driven at 1 volt? Drive the system at 0 dBV and measure the acoustic level with an SPL meter. Anywhere. With any size or complexity of array. The acoustic gain of the speaker coupling, the equalization, the delay settings, and the amplifier drive levels, even the room, are all included in the data. Add more speakers and it will get louder. The dB SPL / volt value will reflect this increase in system capability. Add 20 dB to the dB SPL / volt figure and you have the absolute maximum SPL the system can create before the drive electronics clip. This is a measurement we can apply directly to our system.

Before we move on let's take a brief moment to consider the case of self-powered speakers. The 1 W / 1 m sensitivity rating is rendered truly academic in a system that has a line level input. The dB SPL / volt figure is able to illustrate what the speaker can do with a nominal drive level.

How much SPL can we get out of this system at the mix position?

The 1 W / 1 m method:

1. Determine the maximum output voltage of the console.
2. Determine gain structure and maximum output voltage of all electronic processing.
3. Determine amplifier sensitivity with current gain settings.
4. Determine the speaker sensitivity.
5. Determine the distance to the console and calculate inverse square loss.
6. Determine axial orientation loss at the mix position.
7. **Do** steps **2-6 for** every driver **in** every cabinet **that** contributes sound **at the mix** position.
8. Factor in the summation of every driver based on their relative level and phase at the mix position.
9. Take a wild guess how much the room reflection summations are adding.
10. Look at your computer, rub your chin thoughtfully and announce the answer as an "approximation."

The dB SPL / volt method:

1. Determine the maximum output voltage of the console.
2. Align the system.
3. Place a mic with known sensitivity at the mix position. A given SPL produces a known voltage.
4. Compare the console output voltage to that at the mic. The dB SPL / volt number is now known for the complete system in the room.
5. Prorate the number to the maximum console output voltage. This is the maximum SPL of the system.

Maximum Power Capability

We have seen how we can drive the speakers to their maximum levels. But what are their maximum levels? Power is the scalar factor for speakers. High SPL comes from big (expensive) speakers. Because the scalar factor translates so directly to cost, the power capability decision will be one of the most critical in our design.

The specifications for modern professional speakers go beyond the 1 watt / 1 meter sensitivity rating. The maximum levels, for both short duration and long term, are specified.

The transient nature of our signal makes both of these important. The specifications are fairly straightforward.

Distance and Orientation

These specifications are normally given as an axis at a distance of 1 meter from the speaker. The exceptions to this are loudspeakers that are so large that the 1 m distance is too close to characterize the response. The maximum SPL that we can expect to achieve is extrapolated from the maximum SPL data by employing the inverse square law.

dB SPL Peak

This is the absolute maximum pressure that the speaker can create. This figure is derived by driving the speaker with an instantaneous burst of pink noise or music. This number does not mean that the speaker can reach this point without distortion, or that it can be sustained for an extended duration. The speaker can reach this level. Period. This specification is relevant to the reproduction of material with high peak content such as drums.

dB SPL Continuous

This is the sustainable pressure level over an extended period of time. The time is at least long enough for the system limiters to have engaged. This number should be 6-12 dB lower than the peak. If it is not then one of three things is likely occurring: the peak limiters are too aggressive, the amplifier is too small and is clipping off the peaks, or the RMS limiters are too loose and the speaker will likely live a short life.

Weighting functions are used to tailor the response to mimic the equal loudness contours in the human hearing mechanism. Our hearing is non-linear, giving us different frequency responses over level. At the low levels our hearing favors the range centered around the vocal spectrum, which helps us to hear when people are whispering about us. At high levels our hearing flattens out and then reduces the high- and low-frequency extremes. **"A" weighting** is a filtering added to SPL measurements that mimics the quiet response. "C" **weighting** mimics the very loud response. "A" weighting has its place in noise floor measurements — not in maximum SPL. "C" weighting has some applicability to high-level measurements. If no weighting is specified the specification is assumed to be unweighted (linear).

dB SPL will be a key selection criteria for our design process. Different types of program material will require different dB SPL levels to satisfy our expectations. The maximum dB SPL capability of a single speaker unit will be used to classify speakers. Different levels of speakers can be mixed together as long they are traveling proportionally different distances. There are no "heavy metal" speakers. A low-power speaker at a short distance can keep up with a high-power speaker at a long distance.

The relationship between dB SPL, distance and program material is shown in Fig. 1.41. For a given speaker its capability to satisfy the power needs falls over distance. Power capability is the most intuitive of all of the speaker parameters. Everyone has their own experiential feeling as to the power needs of a particular genre, and of how well a particular speaker fares in that regard. The figures here should be considered as a reference, but not as definitive. What one of us calls extreme level, others wonder when the show is going to start. We will return to this later when we look at how speakers can sum together to increase SPL.

Half-Space Loading

Many speaker specifications give their SPL numbers with the speaker measured on the ground. This couples the low-frequency range and raises the SPL number. When comparing models we need to make sure all speakers were measured under similar conditions.

Frequency Range

The frequency range of our transmission was defined earlier as 31 Hz to 18 kHz. Professional loudspeaker systems are incapable of spanning this entire range with a single

Maximum SPL Over Distance by Program Material

Max dBSPL (Peak)	Distance (Meters)														
	1.0	1.4	2.0	2.8	4.0	5.7	8.0	11	16	23	32	45	64	90	128
154	154	151	148	145	142	139	136	133	130	127	124	121	118	115	112
151	151	148	145	142	139	136	133	130	127	124	121	118	115	112	109
148	148	145	142	139	136	133	130	127	124	121	118	115	112	109	106
145	145	142	139	136	133	130	127	124	121	118	115	112	109	106	103
142	142	139	136	133	130	127	124	121	118	115	112	109	106	103	100
139	139	136	133	130	127	124	121	118	115	112	109	106	103	100	97
136	136	133	130	127	124	121	118	115	112	109	106	103	100	97	94
133	133	130	127	124	121	118	115	112	109	106	103	100	97	94	91
130	130	127	124	121	118	115	112	109	106	103	100	97	94	91	88
127	127	124	121	118	115	112	109	106	103	100	97	94	91	88	85
124	124	121	118	115	112	109	106	103	100	97	94	91	88	85	82
121	121	118	115	112	109	106	103	100	97	94	91	88	85	82	79
118	118	115	112	109	106	103	100	97	94	91	88	85	82	79	76
115	115	112	109	106	103	100	97	94	91	88	85	82	79	76	73
112	112	109	106	103	100	97	94	91	88	85	82	79	76	73	70
	3.3	4.6	6.5	9.3	13	19	26	37	52	74	105	148	209	296	419
	Distance (Feet)														

130 Aural Overload | 124 High level | 112 Medium level | 100 Low level

Figure 1.41 Typical maximum SPL (peak) over distance by program material

transducer. This is a simple case of "you can't get there from here." If the tweeter is made big enough to reproduce the lows it is too heavy to reproduce the highs. If the woofer is made light enough to reproduce the highs it will fall apart while trying to push out the lows. We bring in specialists to do the work of the different sections of the frequency range. The division generally goes at least three ways: into subwoofer, lows and highs. Together we have a full-range system. Note the obvious: the subwoofer is part of the system. It is not a piece of circus equipment that is brought in to provide sonic pyrotechnics. While this may be the case in some applications, this is not the role of a subwoofer in an optimized design. If the subwoofer is intended to provide special effects it will stand alone as a separate system, and will not be part of the minimum variance transmission strategy. If our intent with the subwoofer is to extend the low-frequency range and SPL capability of the system, then it must be integrated into the system. Our discussion will treat subwoofers and mid-bass systems as range extension and SPL boosters within the main system.

Full-Range Speakers

The term "full range" connotes a speaker that covers the entire range of the human voice. Most full-range speakers have a low-frequency cutoff range around 60-70 Hz. Larger units with 15" drivers will reach lower frequencies, while those with 10" LF drivers or less will roll off closer to 100 Hz. The high-frequency range of such devices usually extends up to 18 kHz. Smaller format speakers with very-low-mass HF drivers will have range extension above the high-power systems, which have heavier diaphragms to accommodate their power requirements. The low-frequency range of these systems will not be required to do the work alone in the bottom end. They may overlap the subwoofers or possibly be crossed over above their LF cutoff and be relieved of low-frequency transmission.

Mid-Bass

Mid-bass systems can be used to provide additional SPL capability in the lower mid-range (60-160 Hz). This frequency range has extraordinarily high SPL requirements

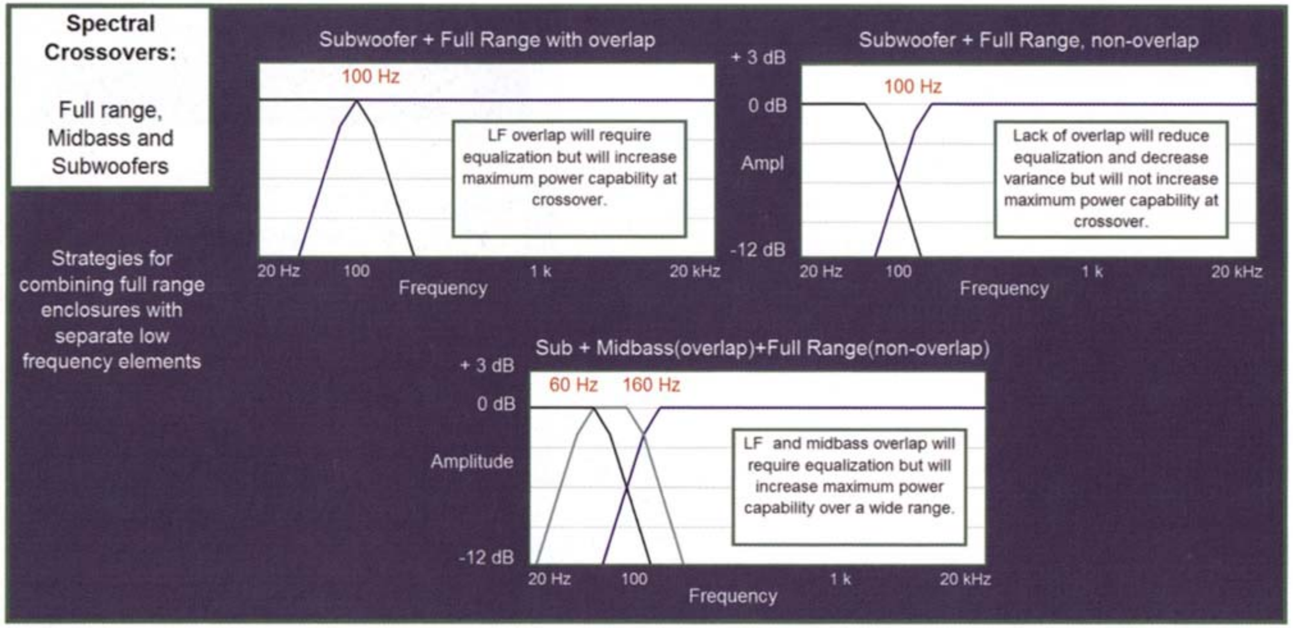


Figure 1.42 Overlapping frequency ranges "full range" + sub, full + mid-bass

in the popular music genre. Dedicated mid-bass cabinets may also have improved low-frequency directional control over the full-range systems in the shared range. If the systems are allowed to overlap their frequency ranges, the coverage pattern will be altered by the acoustic summation. Summation will be discussed in detail in Chapter 2. For now we will note that allowing some degree of overlap is a viable option in this case. Mid-bass systems can also be allowed to have substantial overlap with the subwoofers.

Subwoofers

Subwoofers generally run from 30 Hz to 125 Hz. Subwoofers may overlap with the full-range systems or operate in the low-frequency range alone. If mid-bass systems are used the subwoofers may optionally overlap them.

Speaker Directivity

In the case of the fireworks explosion discussed previously, the radiation from the source was omnidirectional, i.e. equal pressure was sent in all directions. The upper half of the acoustic power of the explosion was wasted since no audience members were located in the air above the sound source. If the fireworks company were able to invent an explosive device that only radiated below it, they would need only half the acoustic power and the audience would be none the wiser. Black powder is cheap compared to the technology it would take to control the sound, so don't look for much research and development on this front.

This is not the case in sound systems. The potential benefits of controlling speaker directionality make it worthwhile to expend the time and money. The most compelling reason for controlling speaker directionality is the degrading effects of echoes. The sound of fireworks echoing off surfaces further enhances our experience. For our speakers a little bit of echo goes a long way. Excessive reflections will degrade the intelligibility and modify the tonal content of the music. Prevention begins with controlling the directionality so that minimal energy is sent into spaces where we don't have audience members.

There are two principal mechanisms that create directional control in a loudspeaker system: the interaction of the speaker with a boundary and the interaction of a speaker with another speaker. Boundary interaction includes horns, walls, wave guides, manifolds and an assortment of names invented by marketing departments. The shape of the radiating element, such as a cone driver, will also affect the free-field directionality of a single loudspeaker. This is beyond our scope here and the circular cone driver is assumed. Speaker interaction with other speakers can occur both inside of a single enclosure and in separate units.

These two mechanisms share much more than one might expect. The reflected energy from a boundary is essentially a secondary sound source. It will combine with the direct sound in much the same way as will the direct sound of two speakers. This is consistent with the principals of acoustical summation which will be described in detail in Chapter 2. Directional control is the result of positive phase addition in one direction and cancellation in the other. Cancellation gets a bum rap in our audio society but without it we have very few options for pattern control.

The facility with which directionality is controlled is frequency-dependent or, more precisely, wavelength-dependent. In the case of boundary steering, the length of the boundary must be sufficient compared to the radiated wavelength, to achieve control. As wavelength increases, the boundary must increase proportionally to maintain the same directionality. One quarter wavelength is commonly considered the minimum boundary size to achieve a strong effect. This is easily achieved at high frequencies, but would require a horn that is two meters deep to control 30 Hz. For lower frequencies the steering is achieved more practically by using the summation properties of multiple speakers.

Defining the Coverage Pattern

The directional aspects of loudspeakers are the result of filter effects that operate spatially in the vertical and horizontal planes. The radial shape that this creates around the speaker is the **coverage pattern.** The coverage pattern is a shape, not a number. A subset of the coverage pattern is the area where the filter effects are less than 6 dB of attenuation. This is the coverage angle. The coverage angle is expressed as a number in degrees.

Using the on-axis level of the speaker as a reference, the coverage angle edges are found by moving off-axis in an arc until the response has dropped 6 dB. This is done separately for both the vertical and horizontal planes and for different frequency ranges.

Representations of both coverage pattern and angle have the common feature of normalization to the on axis response. For the given frequency range the specified values are *relative* to the on axis level, no matter what that level is in *absolute* terms. For example, if the high-frequency horn generates 100 dB SPL on-axis at 1 kHz one meter away, the off-axis points are found when the level falls to 94 dB SPL at that same distance. If the frequency is changed to 30 Hz, the polar pattern can still be determined even though the on-axis response might be 60 dB down from the 1 kHz on-axis response. Coverage renderings give relative level over angle information only. They give us no indication of the power capability of the speaker nor of its on-axis frequency response.

Just as we had a variety of methods for expressing voltage in the previous section, so it is for coverage pattern. Coverage patterns are relevant because off-axis sound does not simply cease after the —6 dB point has been exceeded.

Common representations of speaker coverage include:

1. Coverage angle: the angle between the equidistant —6 dB points at a given frequency, or range of frequencies. Known as the "radial" or "protractor" method. Specified separately for vertical and horizontal planes.
2. Polar pattern: a radial plot of relative level over angle. The data is plotted on a series of concentric rings which represent relative level loss. The outermost ring is O dB loss and the inner rings represent a given number of dB down. The most common formats are 6 dB and 10 dB per ring. A continuous radial function is created with the values of loss over angle. A pair of typical polar

plots is shown in Fig. 1.44. Specified for vertical and horizontal planes over frequency.

3. Equal level contours (isobaric contours): a radial plot of SPL loss over distance from the source. Take a polar plot, turn it inside out, plot it on a linear axis and you have the equal level contour map. The equal level contour map places O dB at the 1 meter distance from the speaker (the standard measurement point for speaker data). The plot traces the radial shape that maintains the same level. Successive rings show the level drop over distance over angle. Specified for vertical and horizontal planes over frequency (also shown in Fig. 1.44).
4. Directivity index (DI): this parameter describes the directional capability of the system over the entire sphere of radiation. The index value is a ratio of the amount of energy forward of the speaker to the energy that would be present if the speaker were omnidirectional. Recall our previous discussion of the fireworks. We have a fixed amount of energy at our source. The DI tells us how much of that energy we are focusing in the forward direction. The value is given in dB, with increasing numbers indicative of greater directional control. The term "front to back ratio" is also used to describe this relationship. The DI values are given as a single value for the specified frequency range.
5. Directivity factor (Q): this is a linear version of the directivity index. The DI value is the 10 log equivalent of the Q factor value. These two values (DI and Q) can

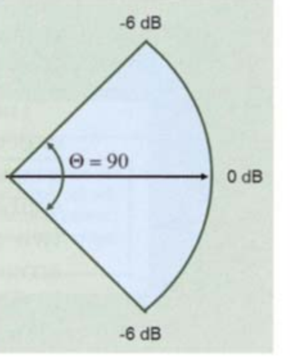


Figure 1.43 Coverage angle as found by the radial (protractor) method

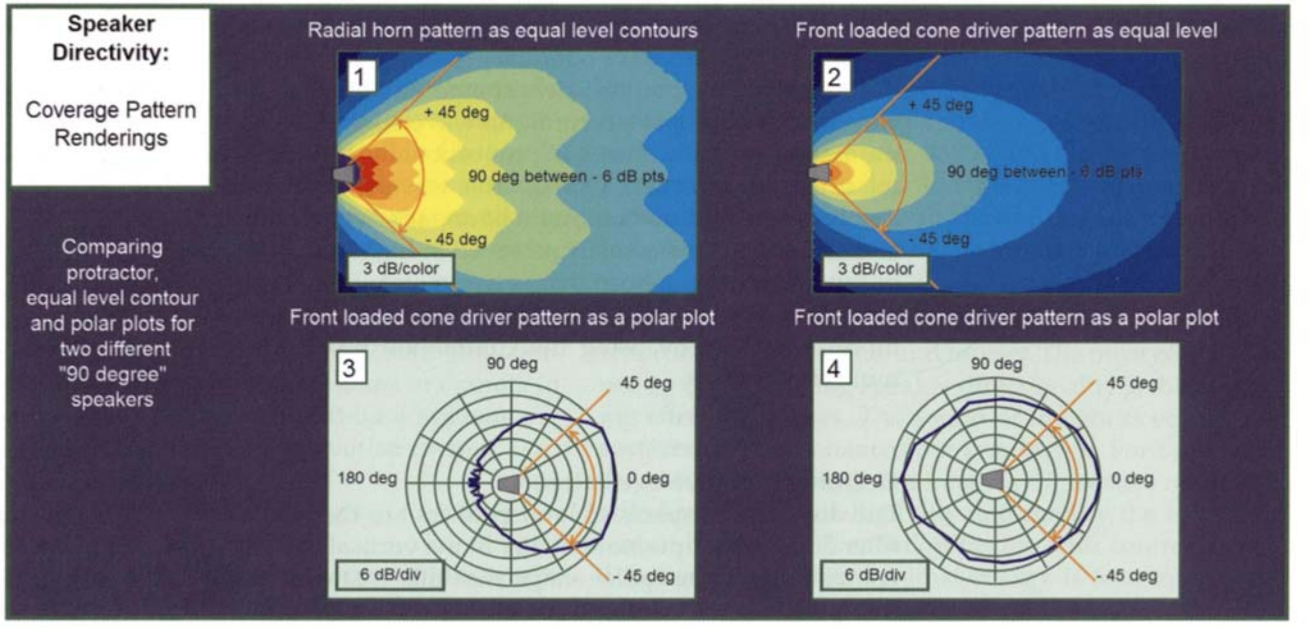


Figure 1 44 Equal level contour method of HF (1) and l
Polar plots of the same HF driver (3) and LF drivers (4)

Figure 1.45 Beamwidth vs. frequency of a small-format full-range loudspeaker. Coverage angle is 90 degrees (nominal)

be plotted on the same graph with different vertical axis numberings.

6. Beamwidth vs. frequency: the term beamwidth is interchangeable with coverage angle (the sound beam). Beamwidth vs. frequency plots create a composite of the full frequency range of the speaker from a series of individual coverage angles. Typical beamwidth vs. frequency plots are 1/3rd octave or 1 octave frequency resolution. Beamwidth plots allow us to view the coverage angle trends over the full range of the speaker in a single chart. Since the beamwidth is made up of coverage angle values, the off-axis response is not included. See Fig. 1.45.

90 Degrees of Separation

Each of these coverage representations tells us about our speakers. Do we need to factor all of them in to making our decisions? This myriad of coverage data can be overwhelming to those of us that just want to figure out what is the best speaker and where to point it, within our short lifespan. Which of these representations is most relevant to our task? We will cover this in more detail later in Chapter 6, but for now let's present them all with a simple task and see how they fare. The test will be to determine the speaker coverage angle required for the most uniform level over four different shapes. Let the games begin.

Figure 1.46 shows a section view with four vertical coverage lines labeled A through D. The same speaker location is used for all shapes. From the speaker's perspective, the angle between the first and last seat is unchanged in all cases. It is the distance *relative* to the lower part of the shape that changes in each case. The first question is: what is the required coverage angle? The answer is found in panel 2. It is 90 degrees in all cases, since we will cover from the top rear to the front near. The differing distance to the floor does not change the angular relationship from the speaker to either the top or bottom extremes. It does, however, introduce a gross asymmetry between the distances to the respective coverage edges. The asymmetry is highest for shape A, and lowest for shape D. This is the principal vulnerability of the radial or protractor method.

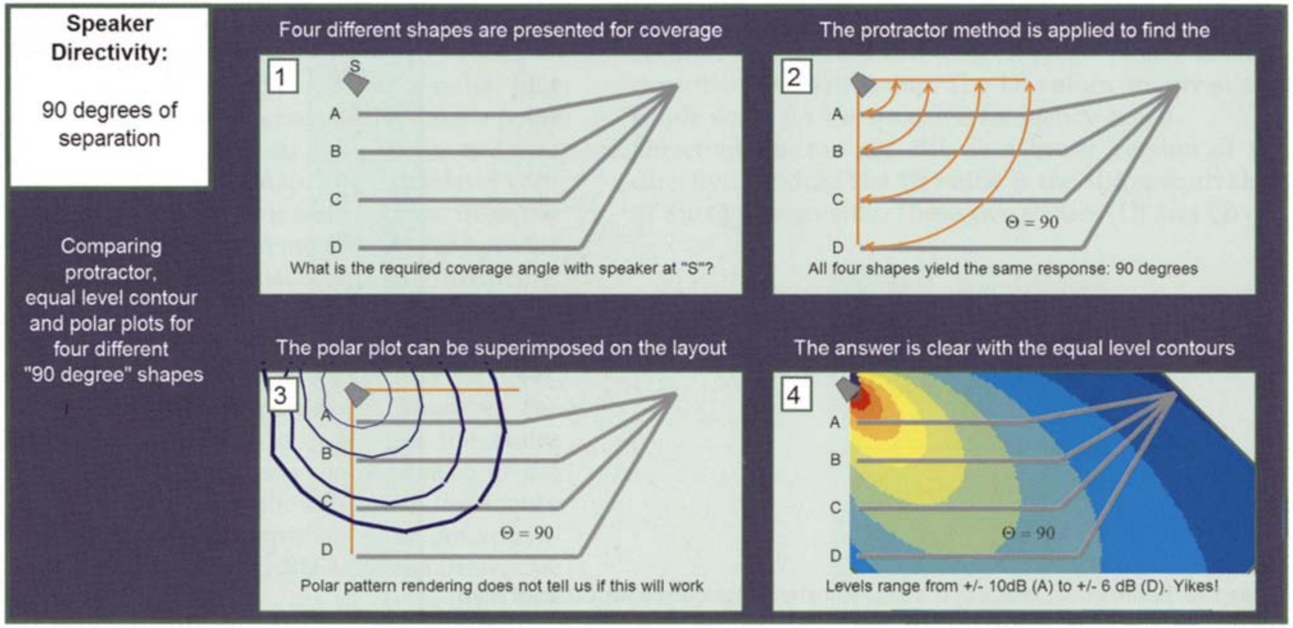


Figure 1.46 Four shapes are shown for coverage from a single speaker location. (1) Listening location targets A - D. The speaker is closest to shape A, and farthest from shape D. (2) Protractor method indicates the application requires 90 degrees of coverage in all cases since the angle between the first and last seat is constant. (3) The polar plot shape using 90 degree coverage plots provides no clear linkage to the listening area shapes. (4) Equal level contours method shows the level uniformity to be poor in all cases, and worst for shape A

Panel 3 employs the polar plot to attempt to discern the best angle. Unfortunately the polar plot gives us no further insight beyond seeing the off-axis shape of the speaker. The frontal lobe is only slightly different from the radial arc so the asymmetry issue persists. The vulnerability of the polar plot is that it follows the radial coverage angle method, rather than leads it. We view the polar plot only after we have determined the angle. Next we apply the equal level contour method. Panel 4 shows the 90 degree contours with the 45 degree down tilt that would center the speaker over the complete coverage angle. In all cases (A - D) the entire seating area is within the coverage angle of the speaker. In all cases the people in the rear of the listening area are more than 10 dB down from the front. This is obviously a very unsatisfactory result. The good news is we have proven that this won't work, so now we can embark on finding out what *will* work. The equal contour method gives us the clues we need to find the angles that will best serve the four different shapes.

There is no figure for the directivity index (DI) or directivity factor (Q) methods. What would those methods be? With a single number that comprises vertical and horizontal directivity how can we possibly do anything but guess? DI and Q are once again followers of the combined horizontal and vertical coverage angles, and we have been through that already.

The answer lies in four different speaker coverage patterns applied to the four different audience shapes. In all cases the speaker is oriented asymmetrically — at the farthest seat, in order to compensate for the asymmetry of the space. The equal level contours are the best indicator of how well we compensate for asymmetry, since they indicate level directly. Since many coverage applications are asymmetrical, this is a critical parameter. Figure 1.47 utilizes the equal level contour method to find the best fit for each shape. Each color gradient represents 3 dB of level change. In each case a different coverage *angle* is found, as determined by the coverage *pattern.* The coverage pattern shape that most closely follows the line of coverage is the determining factor. (1) The asymmetry between the on axis and off axis distance is at its maximum. The coverage angle is at its minimum (20 degrees) and the proportion of the

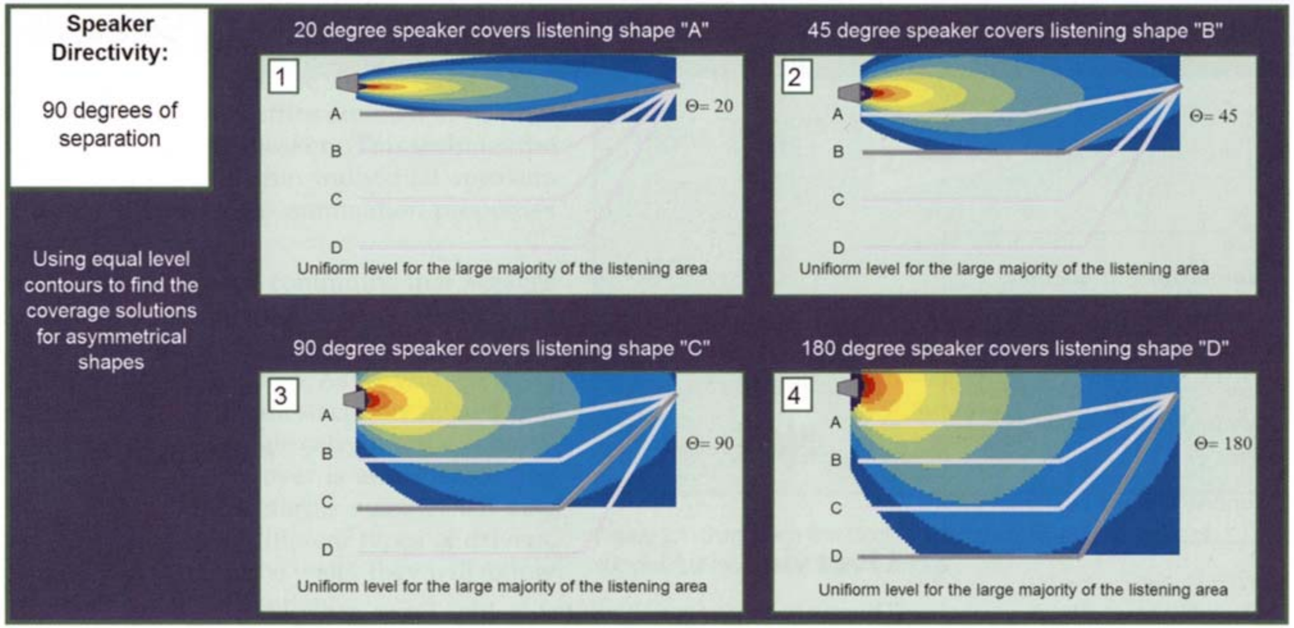


Figure 1.47 Equal level contours method using four different coverage angles. The shape of the contours reveals the best fit for each room

line that can be uniformly covered is at a minimum. (2) The increased distance from the floor decreases the asymmetry. Less asymmetry yields to wider coverage (45 degrees) and more of the shape is covered. (3) Coverage angle opens to 90 degrees. Note that while the 90 degree angle is equal to that of the radial coverage angle method, the orientation of the speaker is different. (4) Coverage reaches maximum symmetry. Coverage widens to 180 degrees and all the seats are covered. The equal level contours lead us toward selection of both coverage angle and orientation, even when faced with an asymmetrical coverage situation. Asymmetry will be a consistent issue in our designs, e.g. every vertical coverage application for starters. Our analysis of coverage angle will need to be able to deal with asymmetry.

Naturally we are concerned about having the upper half of the pattern bouncing off of the walls. At first glance this approach would be a cause of great concern in this regard, perhaps even a lapse of sanity on my part. Take a breath. Bear in mind that we are discussing the behavior of a single speaker. This is the building block, not the finished product. The optimal solutions for creating asymmetric coverage patterns will come from arrays of speakers. The behavior of speaker arrays is the summation of its individual speaker parts. Before we learn to run we will first need to walk.

Reference

Giddings, P. (1990), Audio System: Design and Installation, Sams.

summation *n. addition, finding of total or sum; summing up*

sum *n. & v.* **1.** *total amount resulting from addition of items*

Concise Oxford Dictionary

Summation

Overview

There are very few absolutes in life. There are so many things that we first see in absolute terms and later learn of more complexity. Like right and wrong. My right might be your wrong. Such terms require qualification and context. Well, here is one that we can be sure of as an absolute: one plus one equals two.

Not in audio.

Audio waveforms can be summed electrically or acoustically. We would expect a simple case of addition when combining waveforms with matched amplitude over frequency content. That outcome is possible, but not guaranteed. There is a hidden parameter in the equation that exerts a decisive force on the outcome: phase. The summed signal level may be more than, equal to, or less than its individual components. Perhaps even zero. The multi-dimensional aspect of audio waveform combination gives rise to its proper name: complex summation. And complex it is, but fortunately not so much that we will have to cover the walls with mathematical equations. The properties of audio summation are one of the most important subjects for the sound system designer and optimization engineer. This mechanism governs the outcome at every place that audio is summed: every electrical summing junction, every acoustical interaction between speaker elements and every interaction with a reflection in the room.

As designers we must know the parties that meet at each summation junction, identify the location and manage their interaction. Uncontrolled summation causes widespread level and frequency-response variation through the space. The spatial properties of summation cause every listening position to sound different and the mechanism appears very mysterious or random. Consequently there is much folklore that has grown around it. Once understood, the mechanism will be shown not to be random but rather a repeat offender, whose effect on the sound system is completely predictable. Once understood, the techniques of summation management will become clear, and we are on the road toward the desired minimum variation for the optimized design.

In the previous chapter the transmission signal became airborne. Learning to fly, however, does not end with takeoff. Now that we are in the air it might be a good time to learn how to control the steering mechanisms. This will be our only hope of making a soft landing. That steering mechanism is summation.

This discussion will progress through several stages. We will begin with a study of the mechanism of audio summation in the abstract in both electronic and acoustical systems. The roles of relative level and relative phase are introduced and their effects on the summed frequency response are shown. These properties are then applied to the interaction of multiple loudspeakers. This includes the role of acoustical crossovers within individual speakers and arrays. We conclude with the summation properties of room reflections.

This section draws a line of continuity that weaves together the three most interactive and volatile aspects of our amplified sound system: the spectral divider (crossover), the speaker array and the room. As we will see, these three seemingly distinct entities are so closely related that they can all be lumped into a single category: the acoustical crossover. The acoustic crossover is anywhere in the system where copies of the waveforms meet. When they meet, whether they come from different types of drivers, different speaker cabinets or off the walls, they will follow the same rules: the rules of summation.

Properties of Audio Summation

Audio Summation Defined

Summation occurs when two or more audio signals are combined together and create a new waveform. The summation could be only a momentary event, in which case there is little we can do to manage it. As long as certain criteria are met, the summation will be stable over time and the outcome of the combination is predictable and manageable.

Summation Criteria

Stable summation occurs only when the signals maintain a consistent level and phase relationship. This is not to say they must be matched in level and phase. They may be drastically mismatched. But whatever the relationship, it must be constant. The necessary conditions for stable summation are matched sources and overlapped duration at the summing junction.

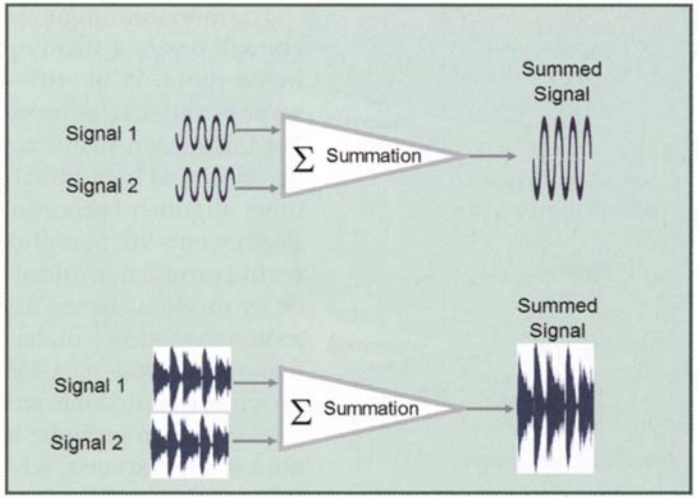


Figure 2.1 Summation flow block. The input signals may be simple or complex, but must have common origins

Source Matching

Recall that in Chapter 1 we discussed the distinction between adding signals at the same frequency (summation), and adding signal at different frequencies (mixing). Stable summation is evaluated on a frequency-specific basis. Stable summation at a single frequency requires that the input signals have a fixed differential at that frequency. To expand this over the full frequency range requires that the input signals have a stable differential at all frequencies. The signals must, then, have related waveforms, i.e. they must come from the same original source waveform. In genetic terms, they must be the children of the same parental waveform. This occurs in our audio system in two forms: electrical copies and acoustical copies. These copies are found all over our audio world with mixing consoles (electrical), speaker arrays and reflections (acoustical) being obvious examples. If we add perfect clones of the source signal, the summation will behave like simple mathematical addition. If the cloning is unsuccessful the copies combine with the original in complex form. The outcome can be predictable and stable, but not necessarily addition.

The monaural signal arriving from two displaced speakers will create a steady-state summation at a given point in the room. While this summation may have a different response at each frequency the response is stable over time. For this reason such a response is measurable and has the possibility of being treatable with delay, equalization and other alignment procedures. By contrast the summation of Beethoven's 9th Symphony with Black Sabbath's *Iron Man* would produce an unstable summation, since it would only be by random chance that the two pieces would contain a moment of source matching. A hybrid case between these extremes is stereo. A stereo feed from two speakers will provide a semi-stable summation. The portions of the mix that appear in both the left and right channels would create a stable response, while those which appear in only one side would not. Because the two signals are not fully correlated, the summed response will change over time at a rate proportional to the degree of difference between the two channels. Therefore, the summation of such a system is not treatable by equalization since the combined frequency response is in constant flux. This unstable summation property of stereo music is easily demonstrated by a simple listening test: perform an electrical summation of the left and right signals at the console. The resulting unstable electrical summation is then reproduced in the sound system and can be confused for acoustical summation in the space.

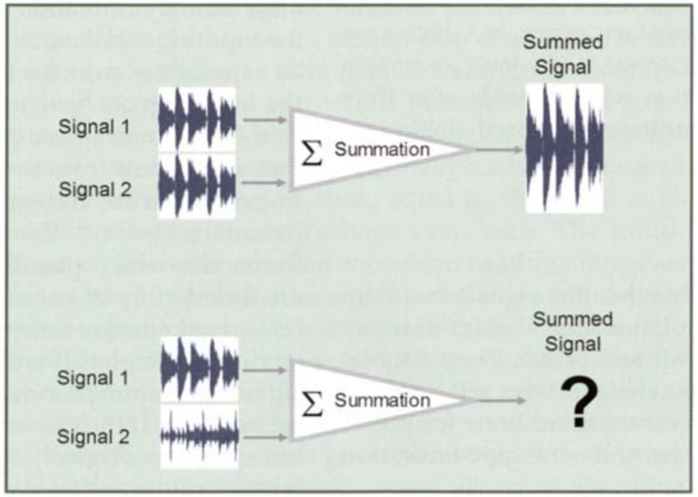


Figure 2.2 Effects of source matching on summation. Top sources have originated from a related waveform for stable summation. Bottom unrelated sources will have a random summation relationship

Duration

The duration of the summation depends upon the length of time the two frequencies share the same location. Let's consider a single frequency from a random signal source such as music or noise. The level at that frequency changes over time. If two copies of this signal are summed together and synchronized, the duration of the summation is infinite. They both rise and fall together, always maintaining a matched relationship. If two signals are offset in time, the duration of the summation will be limited to the time when both signals are present. The first signal is alone at the meeting point until the second signal arrives. If the signal duration is long enough, the two will meet and the summation will produce a stable value. When the source ceases, the early signal will depart the meeting point ahead of the later arrival. If the signal duration is shorter than the time arrival difference, the meeting will not occur. Therefore, with late arriving signals such as echoes, the signal duration must be long compared to the time offset for a stable summation to occur. The implications of this would seem to be that we will never be capable of having a steady summation state in anything but a synchronized system or a continuous unchanging source, such as a sine wave. In practice this is not the case. Music and speech typically contain durations of sufficient length for us to obtain a stable response. The signal must sustain for sufficient duration in order for the ear to discern pitch. As the duration extends beyond a single wavelength our pitch perception improves, and therefore most music will far exceed that duration. Because the perception is wavelength-dependent, the amount of time elapsed before pitch is discerned will change over frequency. For example, 25 ms is long enough to provide around 100 wavelengths at 4 kHz but only a single wavelength at 40 Hz. Pitch sensing of the lower frequency will take more time than the highs. Most music and speech comprise more than single-wavelength durations and spend more than enough time

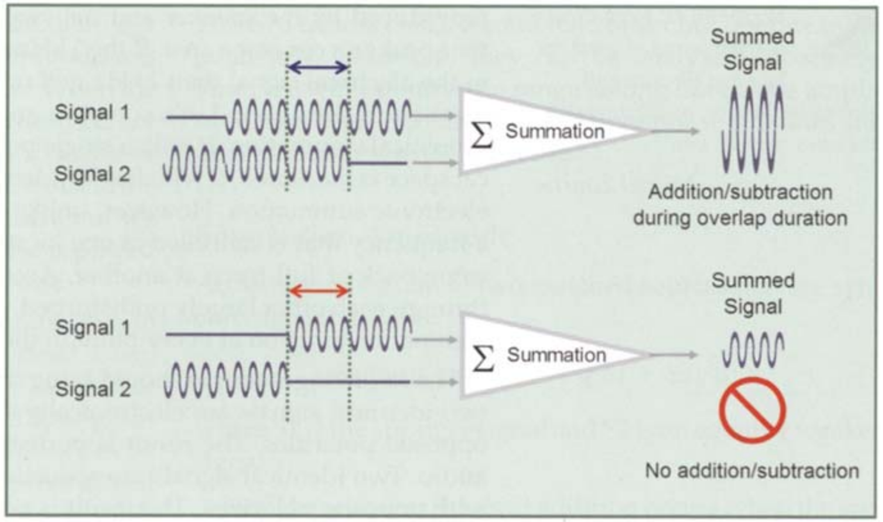


Figure 2.3 The signal must be of sufficient duration to overlap at the summing point for stable ad
channels, the summation will show no stable addition or subtraction

at a given frequency for us to be able to find a steady summation value in spite of the lack of synchronization.

In this text, unless stated otherwise, we will assume that the signal is of sufficient duration to provide us with a stable summation. There are limits to this, and when they are exceeded they will be noted.

Summation Quantity

There is no limit to the number of signals that can be joined together as long as the signals meet the stable summation criteria outlined previously, e.g. room reflections are the summation of a virtually unlimited quantity of individual signals. When the reflection time offset exceeds the signal duration, the summation is no longer stable.

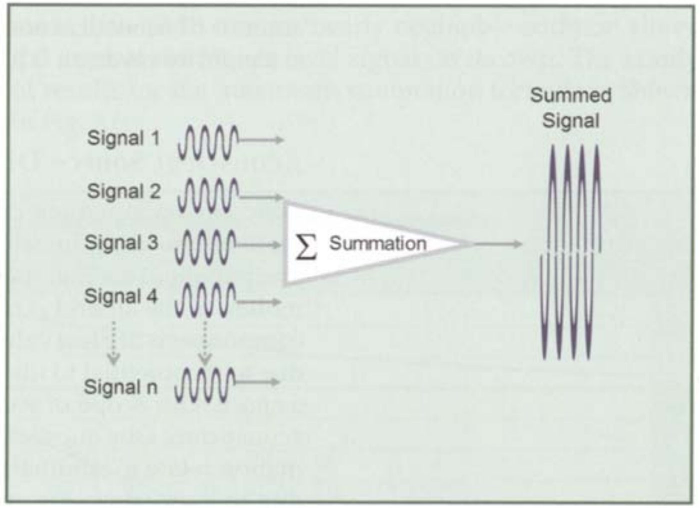


Figure 2.4 The maximum quantity of summation inputs is unlimited. The potential addition a

Electrical vs. Acoustical Summation

Most of the properties of summation apply identically to both electrical and acoustical systems. The key difference is in the spatial placement of the summation. Electrical summation has no geometric dimension, whereas acoustical summation cannot be considered without it. An electrical summation occurs inside the circuit and once joined becomes a new signal, complete with all of the coloration that will soon be explored here. When this signal becomes an acoustical signal, the electrical summation signature is

Perspectives: Regardless of what brand of speakers you use, physics will always be physics, no matter what the sales rep tells you.

Miguel Lourtie

reproduced by the speaker and delivered to all points in the speaker's coverage area. If the 1 kHz range is cancelled in the electrical signal then 1 kHz will not be heard at any location in the space. Let's compare and contrast this to acoustical summation. If only a single point in the acoustical space is measured, it will display identical properties to electronic summation. However, unlike electrical signals, a frequency that is cancelled at one location can, and will, come back at full force at another. Acoustic signals pass through each other largely undisturbed, forming a unique summation junction at every point in the space.

The following example should bring the point to clarity: two identical signals are electronically summed but with opposite polarities. The result is perfect cancellation. No audio. Two identical signals are acoustically summed but with opposite polarities. The result is perfect cancellation at a single location. Other locations will have less than perfect cancellation, and some will have perfect addition. The amount of acoustical energy does not change when there is a polarity reversal. It just moves to other places.

Acoustical Source Direction

Acoustical summations contain signals arriving from wave motions traveling in different directions. The directional component of each sound wave is a product of the particle motion in the air and is termed the intensity. The intensity component is of great value and importance to acousticians, due to its potential to identify specific reflection sources in a space. Our scope of work, however, is distinct from the acoustician. Our question is: how does the intensity information relate to summation? Does the summation change due to the particle directional relationship of the signals? For a given point in space: no. Just like the electrical signal above, the summation is governed by relative amplitude and relative phase at the junction point. The source direction will, however, have a huge effect upon the *distribution* of the summation over the whole of the space. The directional relationship between sound sources will be a prime factor in the *rate of change* of the summation over the space. The rate of summation change translates into the degree of response variation over the space. Fortunately we will not need an intensity probe microphone setup to see this, since they are expensive and impractical for our application. We can see the source direction of the major players in summation with our eyes. They are the speakers and the walls.

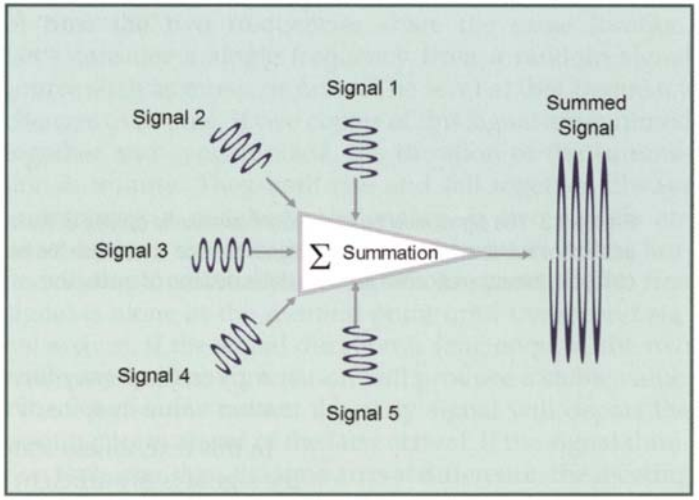


Figure 2.5 Summation of acoustic sources from different directions. The summatio

Summation Math

The relationship between two summed signals in its most simplified form can be expressed by the formula:

$$1 + 1 = 1\ (\pm 1)^*$$

*Depends upon relative phase

This formula illustrates the critical dependency upon relative phase between two signals. The summed signals can add to as much as a value of two (+6dB) and as low as a value of zero (—oo dB). The nature of the summation appears balanced when viewed numerically but when viewed logarithmically (in dB) we can see that the loss effects are much greater than the addition effects. Summation is a form of acoustical gambling where relative amplitude sets the stakes and relative phase decides the winner. When the relative amplitudes are equal the stakes are highest. We are risking everything. We can double our money (+6dB), or lose it all (—100 dB). As the level difference increases, the betting becomes lightweight. We can't win much, but we can't lose much either. The part of relative phase is played by the cards, which decide whether we win or lose.

The odds in Las Vegas are in favor of the house due to the laws of economics. The odds here are dead even, due to the laws of physics (energy can not be created or destroyed). But the game is very asymmetrical. Our gains are small, but spread over a wide area. Our losses can be very large, but confined to smaller areas. A professional gambler learns their trade by studying all of the possible outcomes of the game. So shall we. We cannot changes the rules, but we can learn to place our bets on the winners.

The previous formula can be reworked as follows:

Summation = (Relative amplitude limit factor)
X (Relative phase multiplication factor)

The two factors create a combined effect but operate independently. Therefore, they can be analyzed separately. First, we will examine the range setting of relative amplitude, and then examine how relative phase provides the final placement inside the limits.

Summation Amplitude

Acoustic summation of two matched sources that are synchronized in time is:

$$\text{Summation} = 20 \times \log(S1 + S2)/S1$$

where SI is the stronger signal and S2 is an equal or weaker signal.

It can be seen that the most addition occurs when the signals are equal in level and that the summed signal decreases as they move apart from one another. Level differences of more than 12 dB create a nearly negligible addition above the level of the higher level signal on its own. The family of results for the maximum summation formula is shown in Fig. 2.6.

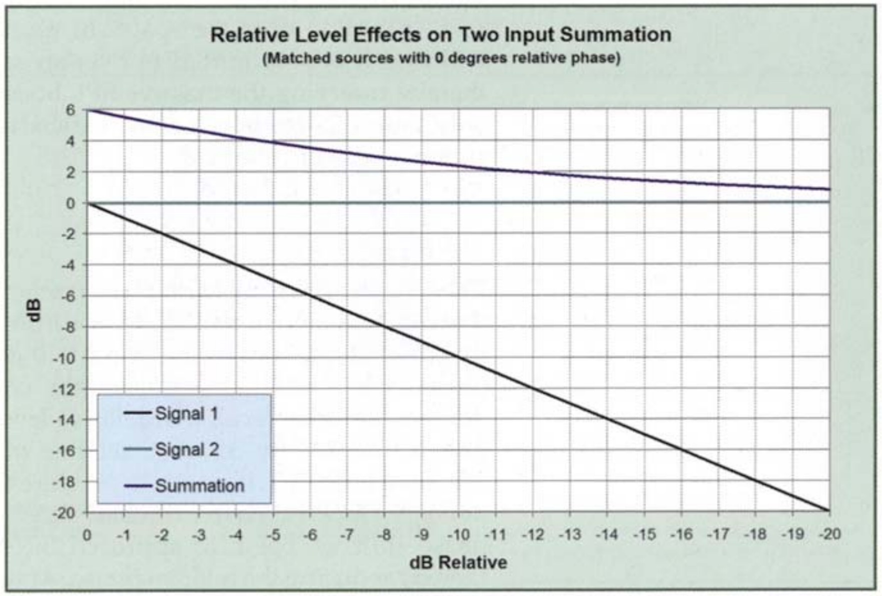


Figure 2.6 The summation level depends upon the relative level between the two input signals. Matched signals have the highest potential. The values of this chart represent the maximum limits of two-element summation, i.e. this is what will occur only when the phase responses of the two signals are perfectly matched

Figure 2.7 Summation level rises as multiple inputs are combined. Each doubling of inputs has the potential of an additional 6 dB at the output

Multiple Input Summation Level Reference												
# of Inputs	**2**	**3**	**4**	**5**	**6**	**8**	**10**	**12**	**16**	**20**	**24**	**32**
Level relative to single input	+6 dB	+10 dB	+12 dB	+14 dB	+17 dB	+18 dB	+20 dB	+22 dB	+24 dB	+26 dB	+28 dB	+30 dB
	Note: This is the **maximum** addition and will only occur when all inputs at matched level and phase											

The limits of addition don't end at two sources — they end at our budget. The addition of more sources opens the possibility for further addition. Each quantity doubling has the potential to add 6 dB of additional level to the response. The maximum addition is easily obtained in the case of a summed electronic signal, and is equivalent to supplying 6 dB of voltage gain. The rationale for multiple unit summation is the desire for increased SPL, so this is where action is. Before we start counting up our cabinets and concluding we will soon be reaching 170 dB, there are some limiting factors to consider. In order to get the 6 dB addition the systems must sum just like an electrical signal, i.e., they must be exact copies in level and phase. They must have 100 per cent acoustical overlap. The spatial nature of acoustical summation dictates that this cannot happen at all but for a minute percentage of locations. As will be discussed later, maximum addition comes with a cost — variation over the space. In most applications we will find it best to minimize overlap as frequency rises, thereby reserving the massive SPL boosts to the low end only. Figure 2.7 contains a reference chart for multi-element summation.

Summation Phase

Relative phase is a measure of the fraction of a wavelength that separates two signals, expressed in degrees. Zero wavelengths of separation correspond to 0 degrees of relative phase, while one-half wavelengths correspond to 180 degrees. We may recall that relative level can range from O dB to infinity. By contrast, relative phase is a circular function, limited in its range to no more than ±180 (0-360) degrees. Once we have exceeded one-half wavelength of phase shift we begin to approach the next wavelength, thereby reducing the relative phase. At the point that a full wavelength of phase shift has occurred, the relative phase value will have returned to zero. This is not to say that 0 degrees and 360 degrees are the same. They are different insofar as different solutions will be indicated to manage their effects, but the effect on summation for a given frequency will be the same. More on this later.

The Phase Cycle

Relative phase can first be visualized as a circle, which we will term the relative phase **cycle** (for brevity "phase cycle" will be used). The phase cycle features 0 degrees at the top and proceeds in either direction until the 180 degree point is reached at the bottom. The addition / subtraction effects of relative phase are based on the radial position of the relative phase value, i.e. our position on the phase cycle.

Relative phase summation properties:

- Maximum addition occurs at 0 degrees
- Addition occurs at values of less than ±120 degrees
- No addition or subtraction occurs at 120 degrees
- Subtraction occurs at values of more than ±120 degrees
- Maximum subtraction occurs at 180 degrees.

The phase cycle's effect on amplitude is not symmetrical as one might have expected. The break-even point is not at 90 degrees and the rates of gain and loss are not symmetrical. The addition side comprises two-thirds of the circle and its action is gradual. The subtraction side comprises only one-third of the circle but its action is extremely steep. The precise amount of addition and subtraction cannot be known from the phase cycle alone. As the relative level approaches unity the action of the phase wheel intensifies. However, in all cases the effect is asymmetrical. The area of addition will always be wider than the area of subtraction. This is balanced by the fact that the level of the subtraction will always be greater than the gains. The asymmetry inherent in the phase wheel will be a key factor in our summation management strategies.

Let's explore the action of the phase cycle in its most extreme form, i.e. when two signals are summed at equal levels. In this case the maximums of addition and subtraction will result. Figures 2.8 and 2.9 show the amplitude values that result from unity gain summation of two signals. The asymmetry of the addition is revealed by viewing the difference in the amount of phase shift required to make a 6dB change. On the addition side the change is spread over 120 degrees, whereas a comparable loss occurs in only 30 degrees.

The matched amplitude condition that permits maximum addition also permits the maximum subtraction when the phase responses are opposite. Therefore the relative amplitude produces both an upper and a lower limit, as shown in Fig. 2.10. Both limits reach their maximums with a unity gain summation and decrease as the levels are offset.

The addition and subtraction effects are a continuous function, i.e. it is not an all or nothing proposition as is often believed. This is worth addressing because of the pervasive misunderstanding regarding the terms "in phase" and "out of phase." These terms are generally referring to the polarity of a signal which has simple duality: in or out, normal or reversed, inverted or non-inverted. For our purposes, "in phase" can be seen as a summation where the relative phase is less than 120 degrees and therefore additive. The subtractive side of the phase wheel would therefore be the "out of phase" side of the equation.

360 degrees and Beyond

We cannot tell the date from looking at a clock face. The fact that the date on my wristwatch is wrong does not mean that I will miss today's 8:00 train.

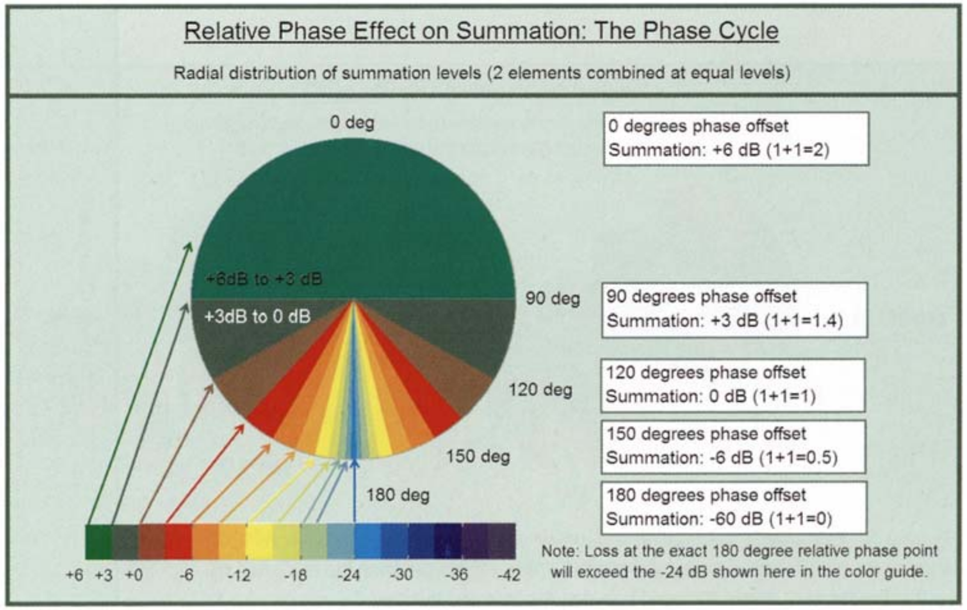


Figure 2.8 The effect of relative phase on summation is shown in the radial form of the phase cycle. Each color shape represents a 3dB level change

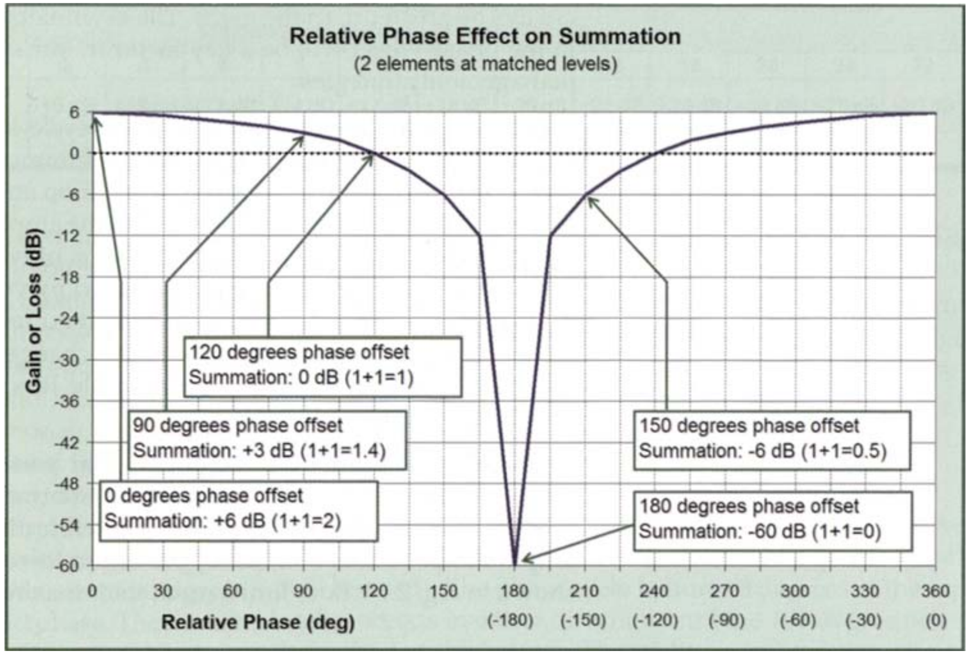


Figure 2.9 The effect of relative phase on summation is shown in the horizontal form of level versus phase offset

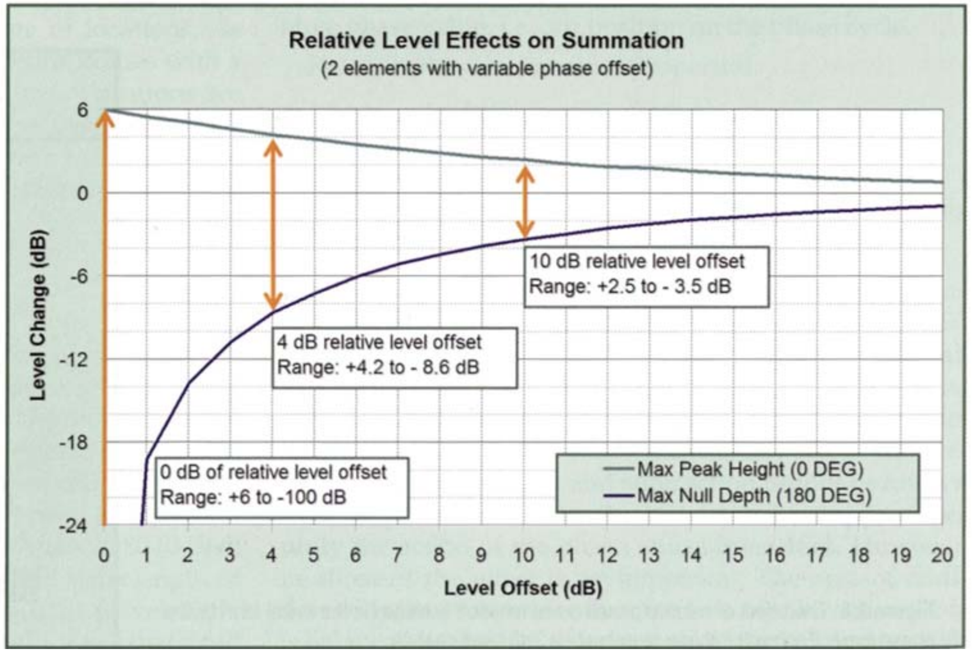


Figure 2.10 Extremes of summation (0 degrees phase offset) and cancellation (180 degrees) for signals with various level offsets. Phase offsets between the extremes will result in the range of levels enclosed between maxima and minima

Is there a difference between 0 degrees and 360 degrees? Is one wavelength apart the same as zero or two or two hundred? Yes, there is a difference, but the summation equations still hold until the duration of the signal has been exceeded. In terms of phase addition, 0 degrees and 360 degrees (one cycle behind) will create the same addition. Likewise the subtraction that occurs from 180 degrees of relative phase summation will also occur with 540 degrees (180 + 360). We are now ready to take the phase cycle to the next level, and observe the action of multiple rotations. First we must relabel the wheel to run from 0 to 360 degrees. Then we can begin to visualize each revolution as a wavelength. As one wavelength is passed a second wheel appears behind the first, creating a spiral of phase cycles. The position on the cycle remains the key to the summation formula for that frequency, as long as its depth in the spiral does not exceed the signal duration.

If the signal is a continuous sine wave there is no way to tell how many wavelengths apart two sources are. Looking at the summation phase of a single frequency will only reveal its position on the phase wheel. Discerning the number cycles that the phase wheel has turned can only be deduced by its context with other frequencies. This context is given by the phase slope, which will be demonstrated as we put the next piece of the summation puzzle in place: time.

Phase delay characteristics applied to the phase cycle:

- A fixed amount of time delay creates a different amount of phase shift for each frequency (the phase cycle turns at a different rate for each frequency)
- A fixed amount of phase shift creates a different amount of delay at each frequency (a given position on the phase cycle creates a different amount of delay at each frequency)
- For a given delay the phase slope increases with frequency (the phase cycle turns faster as frequency rises)
- For a given frequency the phase slope increases with delay (the phase cycle turns more as delay increases)
- The number of "wraparounds" indicates the number of wavelengths of delay (the number of rotations of the phase cycle).

The complexity of the above assertion may be aided by a mechanical analogy with a bicycle: the phase bicycle. When the pedal of a bicycle is turned, the large sprocket

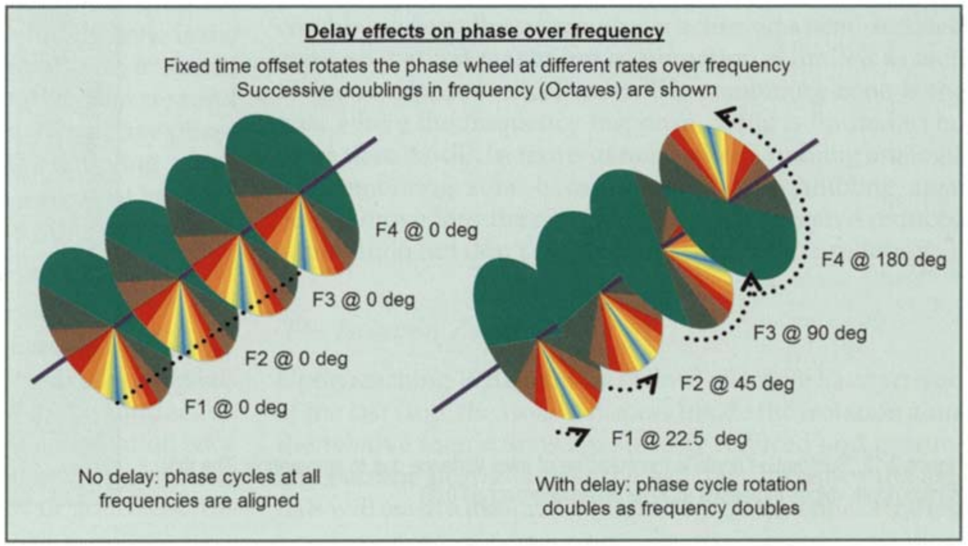


Figure 2.11 The phase cycle as it is affected by delay. The four cycles represent four frequencies at one octave spacing. Left no delay offset. The cycles are all aligned at the same point in their rotation. Right a delay equivalent to ½ wavelength at the highest frequency (F4) has been introduced. Each frequency has been rotated a different amount

attached to the pedal shaft turns the small sprockets via a connecting chain. The speed at which the driving wheel turns depends upon the gear ratio between two sprockets. At one setting the gear ratio is 1:1 while a second setting is 2:1. The same number of turns on the pedal would create twice the number of turns at the drive wheel at the latter setting. For our analogy, time is the turning pedal shaft, while the rear wheel is the phase cycle. The above difference in gear ratio is analogous to an octave change in frequency. Now we expand the analogy to take in the full range of our audio systems. A single pedal shaft is attached to 20,000 different gears attached to 20,000 phase cycles. Each gear ratio is unique and range from 1:1 to 20,000:1. At the beginning, all of the wheels are lined up with 0 degrees on top. Time offset is zero and we have perfect summation at all frequencies. Once the pedal of time begins to move forward, all of the wheels are set in motion, all at different rates. The first wheel to pass the subtraction line will be #20,000 and all others will move at their own rate toward the same point and beyond. The result is a different frequency response for every change in time offset.

Response Ripple

The variance over frequency for a given response can be boiled down to a single number known as the **ripple.** The ripple is the total range in level from the highest to the lowest point in the response. A system that ranges from a +4 to — 8 dB is described as having 12 dB of ripple. Our major concern is to minimize the variation rather than singling out the peaks or the dips. By this reasoning the above response becomes ±6 dB of ripple. We will use ripple as a qualitative measure of summation effects, with minimum ripple being the ultimate goal.

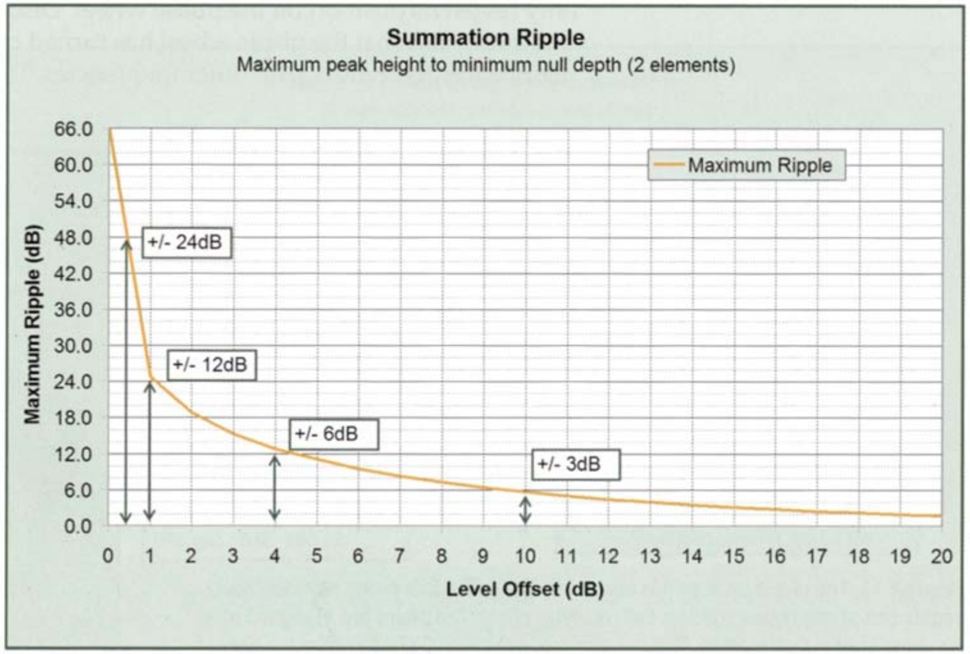


Figure **2.12** Summation ripple is the measure of level variance due to summation. The risk of high ripple values increases as level offset approaches O dB

Summation Zones

When two sources are combined together there is a continuous range of possible mixes of the two decisive factors: relative level and phase. On one end of the continuum are matched level and phase, while at the other extreme they are miles apart. It is impractical to discuss every possible combination; therefore we will focus on trends and group them into five categories based on their type of interaction. These categories will serve as the mileposts along the continuous road of summation.

The Coupling Zone

The coupling zone is found where the relative phase of the summed signals is confined to the addition side of the phase cycle. The sources must remain within ± 1/3rd wavelength (±120°). The effects in the coupling zone are only additive, since the phase offset is never large enough to cause a loss. The amount of addition will range from 0 to 6 dB depending on the degree of phase and level offset. The ripple likewise ranges from 0 to 6dB (<±3 dB). The coupling zone is the most sought-after response when speakers are summed. This is most easily achieved at low frequencies due to the large wavelengths. It is also guaranteed to occur at the exact centerline between two matched speakers (unless one is polarity-reversed). The coupling zone is the most effective means of creating power addition in sound systems. The question becomes: over how large an area and how high a frequency can we go before the relative phase surpasses 120° and pushes us out of the coupling zone? Pursuit of coupling is a risky business, and must be managed carefully to avoid the pitfalls of the combing zone.

The Cancellation Zone

The cancellation zone is the evil twin brother of the coupling zone. This is defined as where two or more signals are combined and the phase is confined to the subtraction side of the phase wheel. The effects in the cancellation zone are only subtractive, since the phase offset is never small enough to cause addition. The amount of subtraction is potentially huge but could be less if there is some degree of level offset. The ripple likewise ranges from 0 to 100 dB (±50 dB). The cancellation zone is the least ideal response when speakers are summed, unless we are intending to steer the sound away from a particular area.

The Combing Zone

When the phase offset reaches the point where subtraction begins (>±120 degrees) we have reached the most volatile area, the combing zone. In the combing zone there is less than 4 dB of isolation between the two signals, but an unspecified amount of phase offset. In contrast to the happy story of the coupling zone, there is a big price to be paid for the addition that we will get at some frequencies: deep and narrow dips at others. The addition may be as large as 6 dB but the losses can be total. Combing zone ripple ranges from ±50 dB to ±6 dB. The combing zone is to be avoided as much as possible, since it is the highest form of variance over frequency.

The Combining Zone

The combining zone ranges from 4 dB to 1O dB of level offset, and again an unspecified amount of phase offset. In the combining zone the systems have achieved a semi-isolated state, so that the maximum combination is limited as well as the maximum cancellation. The combining zone is the area where the frequency response ripple is limited to no more than ±6 dB. In terms of our earlier gambling analogy, the combining zone is our lower-stakes gambling area. As we move into the combining zone we receive reduced summation but don't pay the price of deep cancellation.

The Isolation Zone

Upon reaching 10 dB or more of level offset we have arrived at the last stop: the isolation zone. Inside the isolation zone the relative interactions are steadily reduced and eventually become negligible. As the isolation increases the signals will reach a minimally interactive state, rendering their

Summation Zone Reference Chart	Level Offset (dB)	Maximum (dB) 0 Degrees Phase Offset	Minimum (dB) 180 Degrees Phase Offset	Ripple (dB)	
	0.01	6.0	-60.0	+/- 33	
	0.1	6.0	-38.4	+/- 22	
	0.25	5.9	-30.5	+/- 18	Combing Zone
	0.5	5.8	-25.0	+/- 15	Level offset range: 0 to 4 dB
	0.75	5.7	-21.7	+/- 13	Phase offset range: unlimited
	1	5.5	-19.2	+/- 12	Summation range: +6 dB to -60 dB
	2	5.1	-13.7	+/- 9.4	Ripple range: > +/- 6 dB
	3	4.6	-10.7	+/- 7.7	
	4	4.2	-8.7	+/- 6.5	
	5	3.9	-7.2	+/- 5.5	Combining Zone
	6	3.5	-6.0	+/- 4.8	Level offset range: 4 to 10 dB
	7	3.2	-5.2	+/- 4.2	Phase offset range: unlimited
	8	2.9	-4.4	+/- 3.7	Summation range: +4 dB to -8 dB
	9	2.6	-3.8	+/- 3.2	Ripple range: +/- 6 dB to +/- 3 dB
	10	2.4	-3.3	+/- 2.9	
	11	2.1	-2.9	+/- 2.5	
	12	1.9	-2.5	+/- 2.2	
	13	1.8	-2.2	+/- 2.0	Isolation Zone
	14	1.6	-1.9	+/- 1.8	Level offset range: > than 10 dB
	15	1.4	-1.7	+/- 1.5	Phase offset range unlimited
	16	1.3	-1.5	+/- 1.4	Summation range: +2 dB to -3 dB
	17	1.1	-1.4	+/- 1.3	Ripple range: < +/- 3 dB
	18	1.0	-1.2	+/- 1.1	
	19	0.9	-1.1	+/- 1.0	
	20	0.8	-0.9	+/- 0.9	
		Coupling Zone Level offset range unlimited Phase offset range: 0 to 120 degrees Summation range: +6 dB to 0 dB Ripple range: < than +/- 3 dB	Cancellation Zone Level offset range unlimited Phase offset range: 120 to 180 deg Summation range: 0 dB to -60 dB Ripple range: < +/- 30 dB		

Figure 2.13 This summation zone reference shows the ranges of addition, subtraction and ripple for each of type of interaction

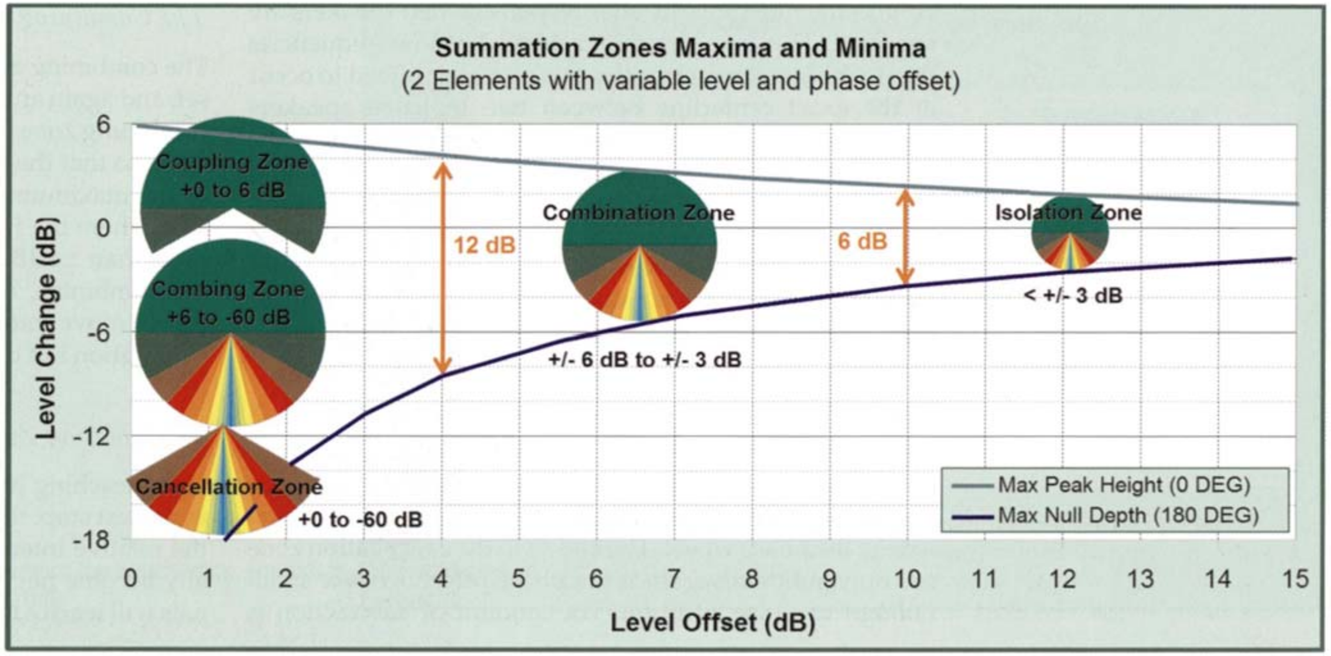


Figure 2.14 The summation zones graphed by level change. The coupling and cancellation zones use only part of the phase cycle, while all others use the complete cycle

relative phase academic. The ripple in the isolation zone does not exceed a total of 6 dB, which corresponds to the accepted minimum standard of coverage uniformity.

Summation Zones over Frequency

Now we can combine the roles of relative level and phase together and observe the finished product: summation effects over the full frequency range. To complete this we will need to add the final ingredient: time. Up to this point we have discussed the summation zones without regard to a specific frequency or time offset. To discuss full-range summation the time offset must be specified. A given time offset creates a different phase offset at each frequency which will affect its summation zone placement. When two full-range signals are added together, the summation values are derived separately for each frequency, based on their amplitude and phase differences. Over the course of a full-range response the interaction may fall into as many as four of the five summation categories. When the time offset is 0 ms the summation is entirely positive (coupling). As the time offset increases the frequency range where addition and subtraction both occur moves progressively downward (combing). As the level offset rises, the interaction lessens (combining) and finally becomes a minimal effect (isolation). This zonal progression will be a central theme in this book since it is the standard sequence of events for audio summation interaction.

This series of frequency responses (Figs 2.15 and 2.16) will serve as an abridged dictionary of the summation family of responses. The frequency range effects are continuously variable as are the level effects. However, the traces represent the major milestones. These traces will be reduced into icons for use in the remainder of this book. The icons will be used to designate positions in the hall or around arrays where the conditions that cause them exist. We will need to become fluent enough with these icons to be able to identify them in the wild on our analyzer.

(a)

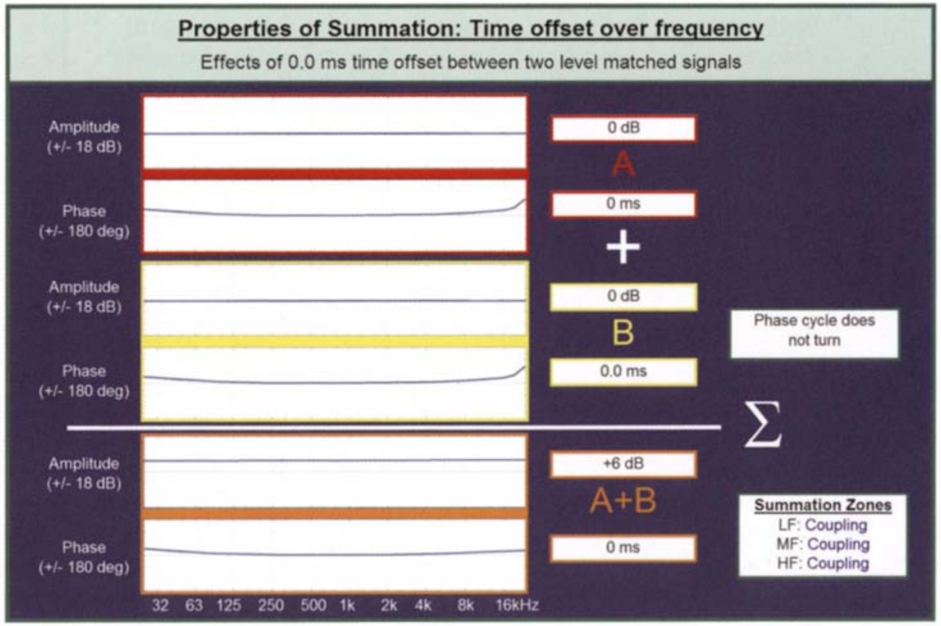
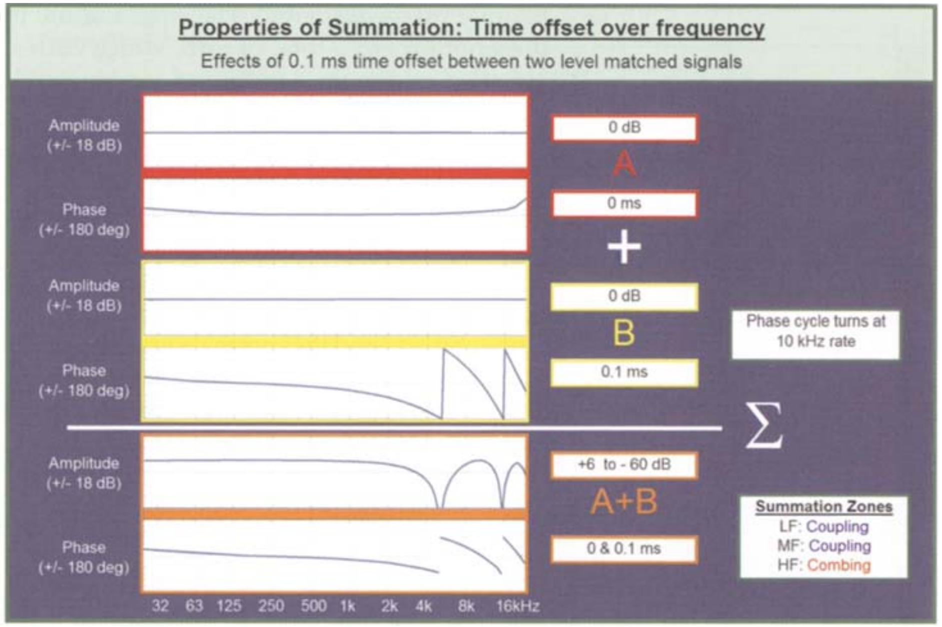


Figure 2.15 Summation over frequency with O dB level offset, with various time offsets. (b) 0.1 ms, (c) 1 ms, (d) 10ms. Note the range above 1 kHz on the 10ms screen is blank due to the analyzer resolution limitations. The actual response ripple continues all the way to 20 kHz

(b)



(c)

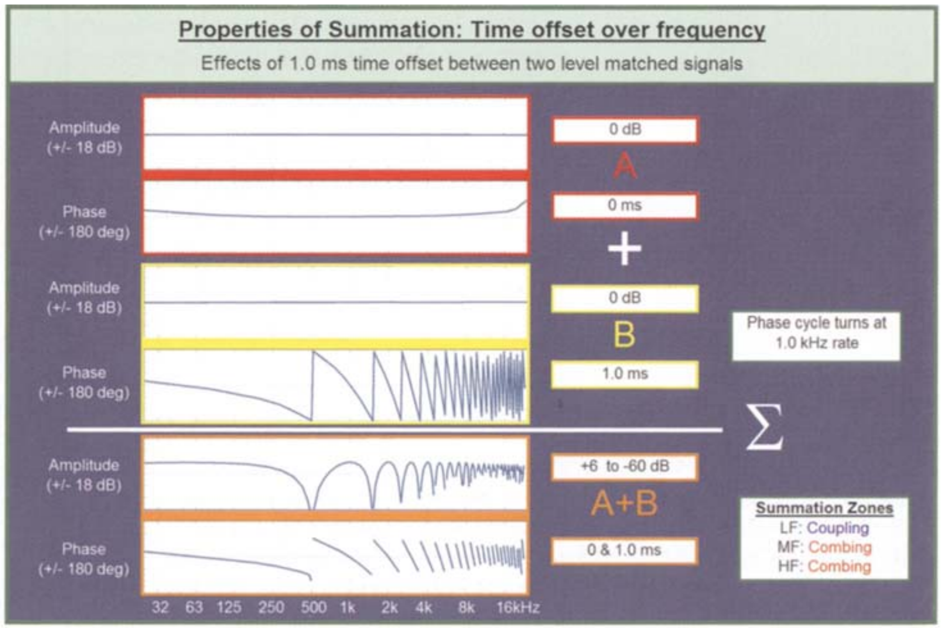


Figure 2.15 (*Continued*)

(d)

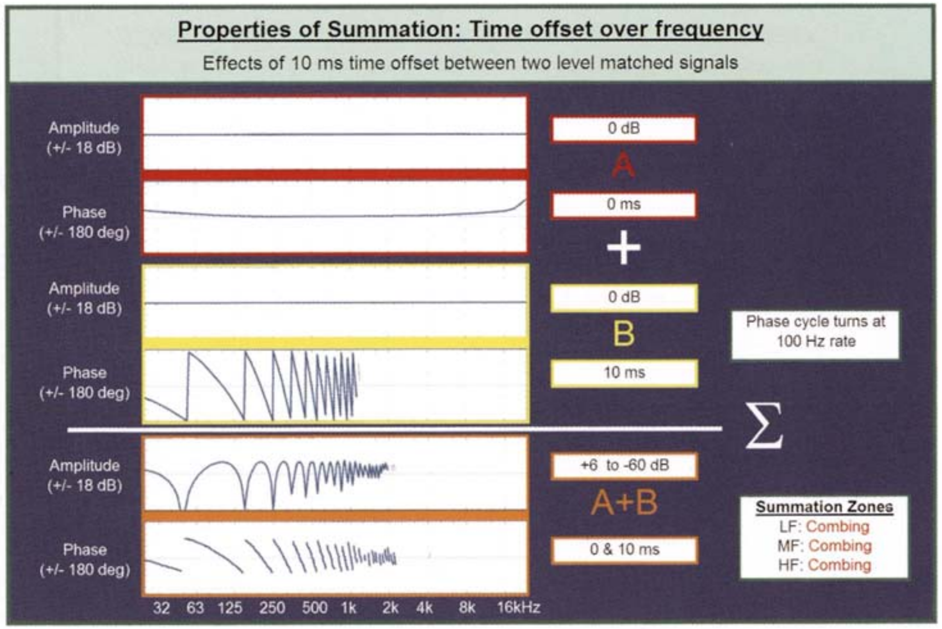


Figure 2.15 *(Continued)*

(a)

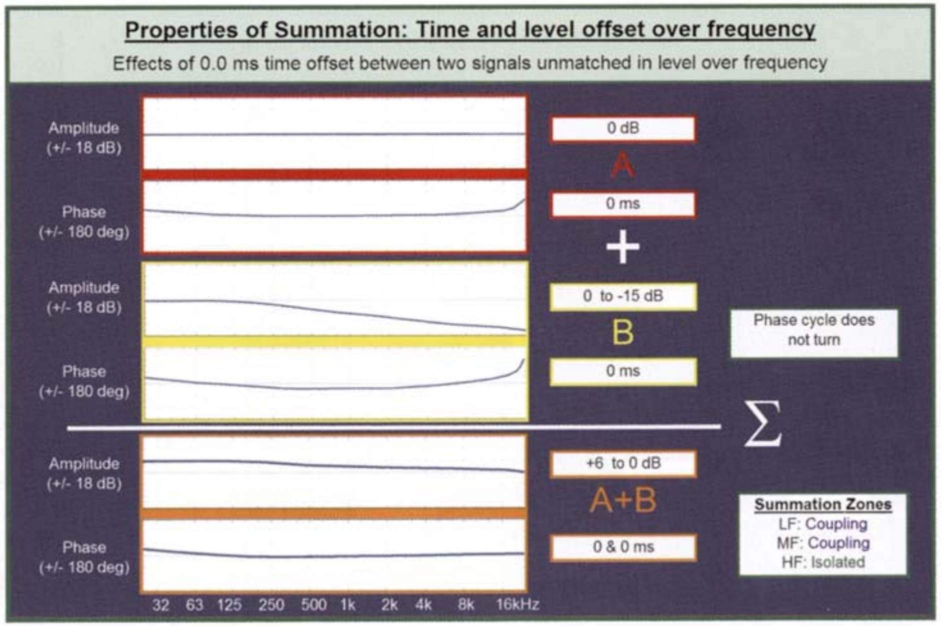
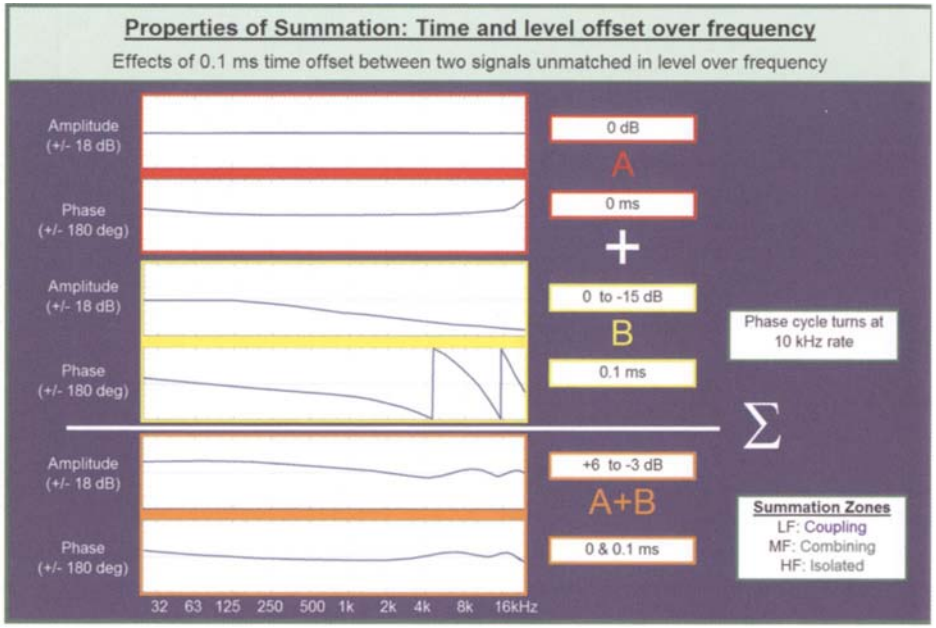


Figure 2.16 Summation over frequency with a variable level offset (due to filtering), with various time offsets. As time offset increases the ripple variance spreads downward in frequency. The range of the ripple is reduced by the filter induced level offset (a) O ms, (b) 0.1 ms, (c) 1 ms and (d) 10 ms. Note The range above 1 kHz on the 10ms screen is blank due to the analyzer resolution limitations. The actual response ripple continues all the way to 20 kHz

(b)



(c)

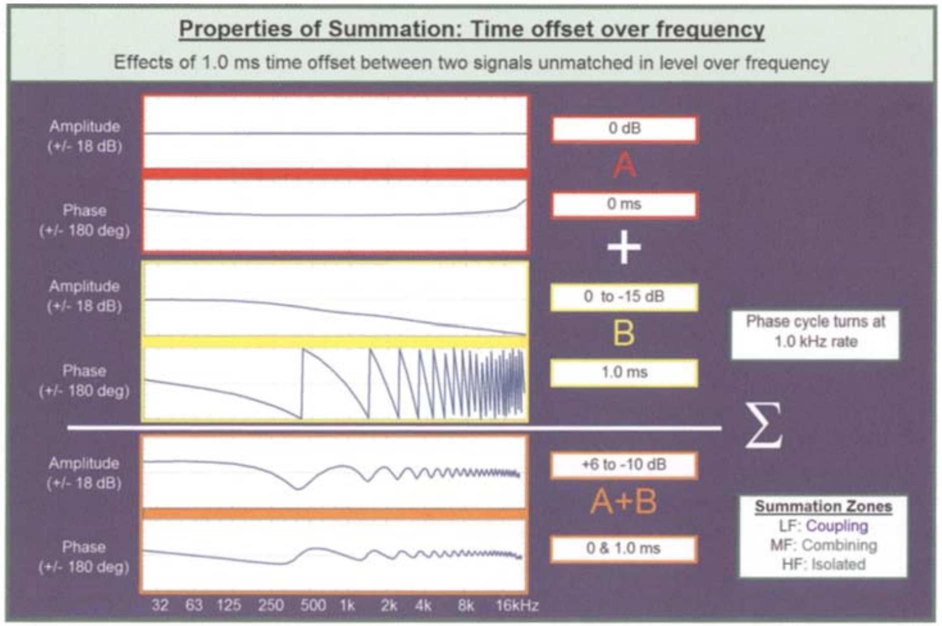


Figure 2.16 *(Continued)*

(d)

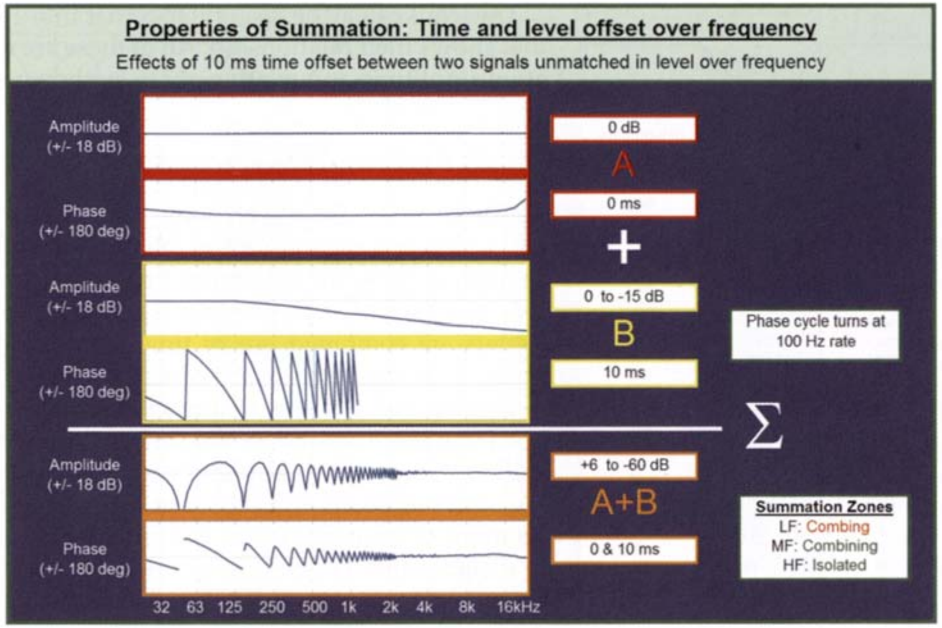


Figure 2.16 *(Continued)*

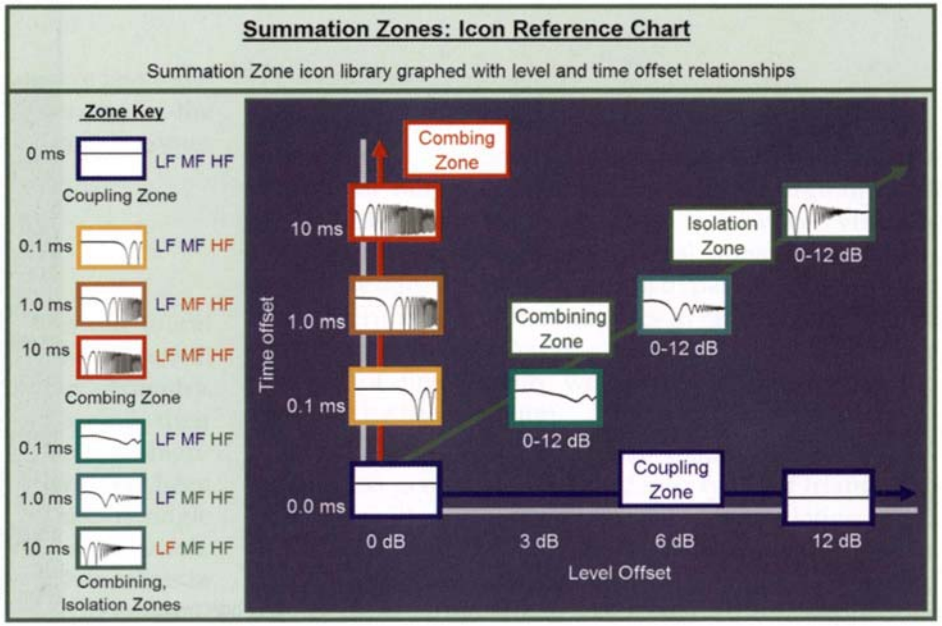


Figure 2.17 Summation icons displayed as time offset vs. level offset

The icons can all be brought together into a single graph that shows their relationship. All of these are related to relative amplitude and relative phase as shown in Fig. 2.17.

Comb Filtering: Linear vs. Log

Recall the phase bicycle analogy. We saw that when a signal is delayed the amount of phase shift will vary over frequency. Now we will apply this to the combination of a delayed signal with one that is not delayed. When two signals are combined out of time the amount of *relative* phase shift will vary over frequency. The variance is due to the fact that phase is frequency-dependent, but time is not. For a given amount of time offset, the amount of phase shift will increase as frequency rises. The result is a repeating series of peaks and dips that are spaced apart by the rotations of the phase cycle. For brevity here, we will refer to a signal that is delayed by one cycle (360 degrees) as being one "wavelength" delayed. A 100 Hz signal that is delayed 10 ms and a 10 kHz signal that is delayed 0.1 ms are both "one wavelength delayed." A 10 ms delay will cause 10 kHz to be 100 wavelengths delayed. This is significant because this is how we hear the frequency response effects of summation with our log-based hearing. The number of wavelengths difference between two signals is the decisive factor in the percentage bandwidth of the filter shape created. A delay of a single wavelength will create a peak that spans one octave between its surrounding nulls. A two-wavelength delay will narrow the peak to a half octave and so on.

The conventional terminology for the time offset summation is "comb filtering." The name connotes the visual similarity to the teeth spacing of a comb. This similarity is based upon a linear frequency axis display, where the spacing comes out perfectly even over frequency. Since we do not hear linear, this visualization has limited applicability. With a log display the effects of summation appear as a series of peaks and dips that get progressively narrower as frequency rises. This is where the tonal perception characteristics of the ear discussed in the next chapter will come into play. The wide peaks provide the most discernible tonal coloration,

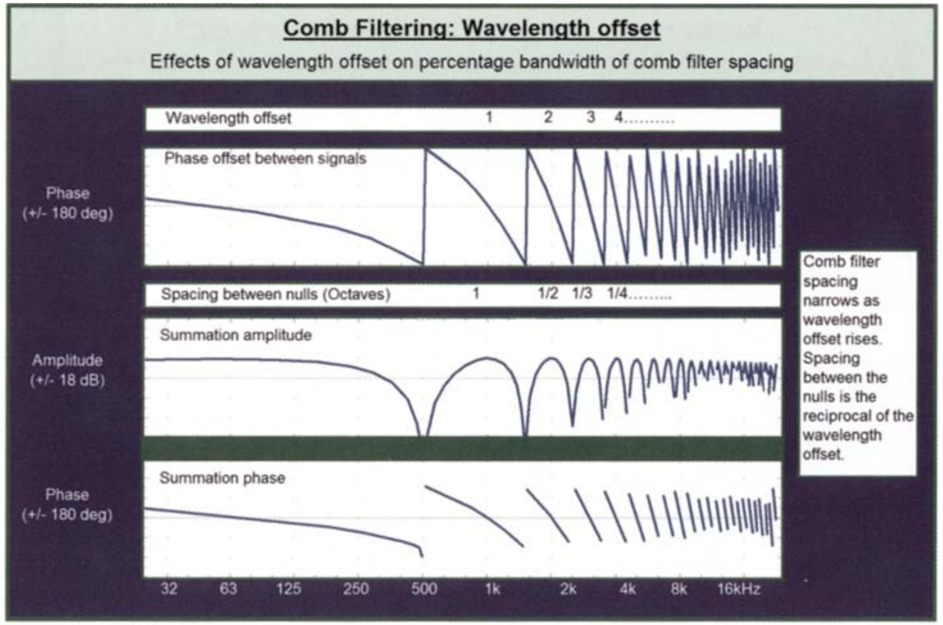
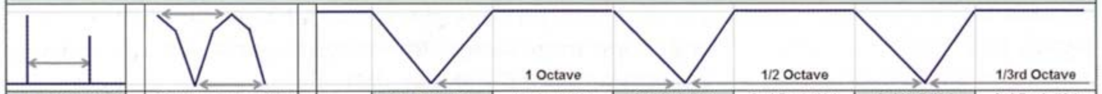


Figure 2.18 The relationship of comb filtering and wavelength offset over frequency

Comb Filter Identification Reference



Time Offset (ms)	Comb Frequency (Hz) Spacing (Hz)	1st Null (Hz) .5 x Comb Freq	1st Peak (Hz) 1 x Comb Freq	2nd Null (Hz) 1.5 x Comb Freq	2nd Peak (Hz) 2 x Comb Freq	3rd Null (Hz) 2.5 x Comb Freq	3rd Peak (Hz) 3 x Comb Freq
0.1	10,000	5,000	10,000	15,000	20,000	25,000	30,000
0.2	5,000	2,500	5,000	7,500	10,000	12,500	15,000
0.3	3,333	1,667	3,333	5,000	6,667	8,333	10,000
0.4	2,500	1,250	2,500	3,750	5,000	6,250	7,500
0.5	2,000	1,000	2,000	3,000	4,000	5,000	6,000
0.6	1,667	833	1,667	2,500	3,333	4,167	5,000
0.7	1,429	714	1,429	2,143	2,857	3,571	4,286
0.8	1,250	625	1,250	1,875	2,500	3,125	3,750
0.9	1,111	556	1,111	1,667	2,222	2,778	3,333
1.0	1,000	500	1,000	1,500	2,000	2,500	3,000
2.0	500	250	500	750	1,000	1,250	1,500
3.0	333	167	333	500	667	833	1,000
4.0	250	125	250	375	500	625	750
5.0	200	100	200	300	400	500	600
6.0	167	83	167	250	333	417	500
7.0	143	71	143	214	286	357	429
8.0	125	63	125	188	250	313	375
9.0	111	56	111	167	222	278	333
10.0	100	50	100	150	200	250	300
20.0	50	25	50	75	100	125	150
30.0	33	17	33	50	67	83	100
40.0	25	13	25	38	50	63	75
50.0	20	10	20	30	40	50	60
60.0	17	8	17	25	33	42	50
70.0	14	7	14	21	29	36	43
80.0	13	6	13	19	25	31	38
90.0	11	6	11	17	22	28	33
100.0	10	5	10	15	20	25	30

Figure 2.19 A reference chart showing the relationship of time offset and the frequency of peaks and nulls in comb filtering

Perspectives: Cirque du Soleil has been an amazing testing ground for me. Since 1990, I have been adding complexity for each show I designed. This wouldn't have been possible without a lot of measurements and tuning to arrive at this type of complex distributed system. The more sources I have to play with, the happier I am as a creative person but this can rapidly become a time-alignment nightmare. Of course, reconfiguring time delays on the fly satisfies the creative and the system engineer in me but it exhausts my programming side.

Francois Bergeron

while the narrowly spaced peaks and dips have less tonal distinction. When the spacing becomes too narrow, i.e. the time difference becomes too large, the perception becomes a discrete echo.

Summation Geometry

Acoustic summation contains spatial aspects not found inside a piece of wire. To prepare ourselves for this we will need to acquaint ourselves with summation geometry. Whenever two displaced sources propagate sound, their summation properties for a given frequency at any singular point are determined by the same factors we have been facing all along: relative level and relative phase. A single point is negligibly different from electronic summation. But what about all the other points? Are the other points in the room randomly related to each other? Hardly.

Triangulation

The answer lies in spatial geometry and array theory. The more scholarly textbooks discuss array theory with omnidirectional radiating sources. There is no arguing with the validity of this approach. There is, however, virtually no practical use for omnidirectional transmitters in our system design, so we are going to bypass that theory, and its staggering piles of mathematics in favor of a simplified rendering: triangles. As long as the speakers have some measure of directionality, we can model their summation behavior by triangulation.

The triangle is the elementary shape for two-speaker summation geometry. There are four types of triangles to consider as shown in Fig. 2.20. The differentiating factor is the angle seen by Speaker A to the remaining points. The different triangle types will be shown to correspond closely to the summation zones detailed in the previous section, providing a link from the electrical to the acoustical domains. For this linkage to solidify, we must narrow down the three factors affecting the response at the summation point of the triangulation (C).

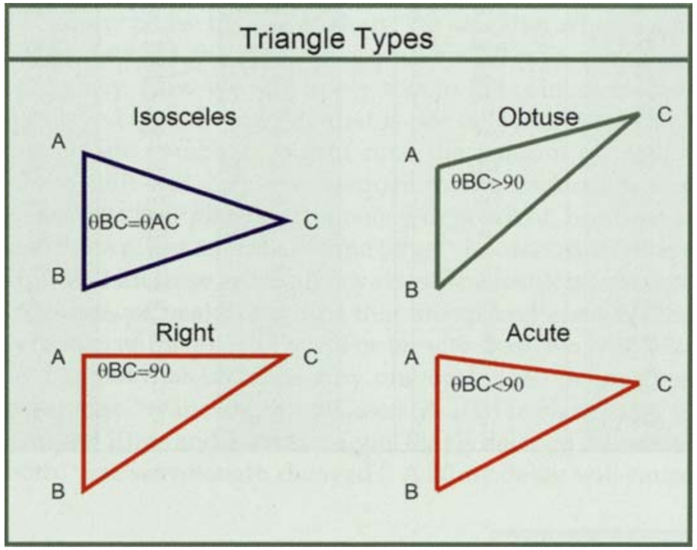


Figure 2.20 Summation triangle types. Sound sources are at locations A and B. The listening position is C

Factors affecting the response at the summation point (C):

- the level offset due to distance offset
- the level offset due to axial response offset
- the time offset.

To begin we will explore triangulation's most basic equation, the Pythagorean theorem: the sum of the squares of the smaller sides of a right triangle is equal to the square of the remaining side (the hypotenuse).

Let's apply this to a two-point summation in space. Identical speakers A and B radiate sound. Our listening reference point is C. We have a triangle. The displacement between the speakers is side AB. The paths to the listening position are AC and BC respectively. The summation response at C is the result of the level offset and the time offset between paths AC and BC. There is, however, a key difference between these two factors. The level offset is the *ratio* between the level of the two speakers. The time offset is the difference between arrivals from the speakers. The first factor is governed by multiplication and division, the second by addition and subtraction. The dilemma that this imposes for us is illustrated in Fig. 2.21 where we can

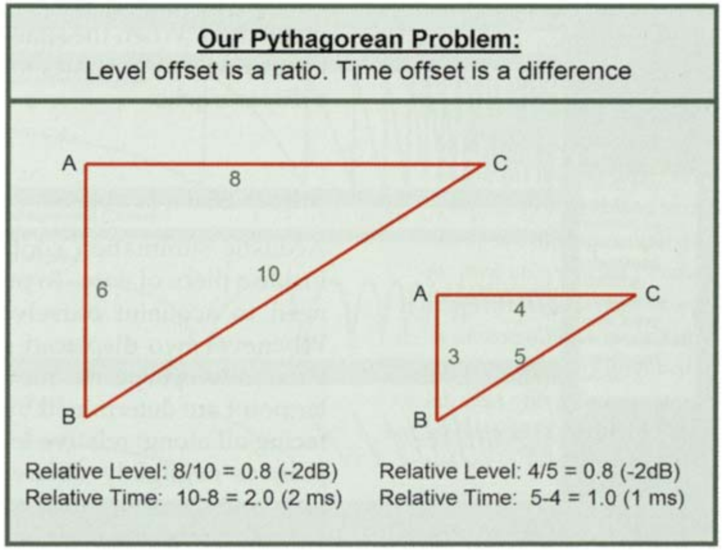


Figure 2.21 The level offset and time offset do not scale together. Level offset scales as a ratio between lengths (multiplication). Time offset is the difference between lengths (subtraction)

HO

see that the relationship between speakers does not scale equally in level and time. In this case we have modified two factors: the distance to the summation point and the displacement between the sources. Both factors were doubled, the level offset ratio remained constant, but the time offset difference doubled.

Now let's look at each of the four triangle types and see how they are affected by those two parameters: variable distance to the summation point and variable displacement between the sources. These two variables will affect the relative distance, relative axial loss and time offset.

We begin with the isosceles triangle as shown in Fig. 2.22(a). The isosceles triangle gives us zero offset in distance, axial loss and time. The coupling zone is found at the summation point of the isosceles triangle regardless of distance or displacement.

Next we can look at the right triangle (Fig. 2.22(b)). The right triangle, as we saw before, presents changing time offsets and level offsets over both distance and displacement. The icons representing the combing zone can be found along the various milestones of the right triangle. These variations are affected by the axial offset, which can reduce the combing when the offsets are highest. However, since most systems exhibit poor directional control in the low frequencies, the degree of reduction will be limited. The range of action for the right triangle is directly in front of speaker A. The sound from B is going to meet it at every point on its journey forward. Their time and level relationships will never stabilize. As distance increases the interaction becomes more stable since the axial loss becomes negligible but some degree of time offset persists.

The next entry is the obtuse angle found in Fig. 2.22(d). The time and level offsets will never match up as above, but a vital difference emerges in the axial offset category. Because the summation point has moved outward from the on-axis point of speaker A, this unit has the ability to gain dominance in level via differences in axial loss. The isolation capability increases as the angle becomes more obtuse. The level offsets created by the axial offset push us into the combination zone and eventually beyond to the isolation zone. As the inevitable time offsets are accrued the isolation reduces their degrading effects. The obtuse angle is the principal driving mechanism of the most popular array type: the point source.

The final angle is the acute (Fig. 2.22(c)), and we have an acute problem here. This is the inverse of the obtuse angle and affords us the least amount of isolation possible, and yet rampant opportunity for variable time offset. The combing zone is found whenever displaced speakers face inward toward a mutual destination.

Now we will plot out the triangles onto a layout of two speakers as shown in Fig. 2.23. We will place the summation icons at the appropriate places on the layout and see how the various zones come together. For our example we will look at the triangles in the context of an equal level contour rendering of a two-speaker point source array in space. This series shows the response at the three frequency ranges from which the summation icons are based (100 Hz, 1 kHz and 10 kHz).

Summation example key features:

- All frequencies show maximum addition at the center of the isosceles triangle (coupling zone).
- 1 kHz and 10 kHz show deep cancellations around the right triangle point (combing).
- 10 kHz is the first to show decreased combing as we move off-center into the obtuse triangle (isolation zone). 1 kHz follows this trend as the obtuse angle increases.
- The area inside the isosceles triangle, the acute angle area, has the highest degree of variation over position and over frequency (combing zone).
- The isolation increases as the obtuse angle increases.
- The 100 Hz pattern does not show a cancellation at the right angle position. This is due to the fact that the displacement between the sources is less than a wavelength at 100 Hz (see next section).

The spatial effects of summation can be dizzying if we do not know what to expect. We need a road map with

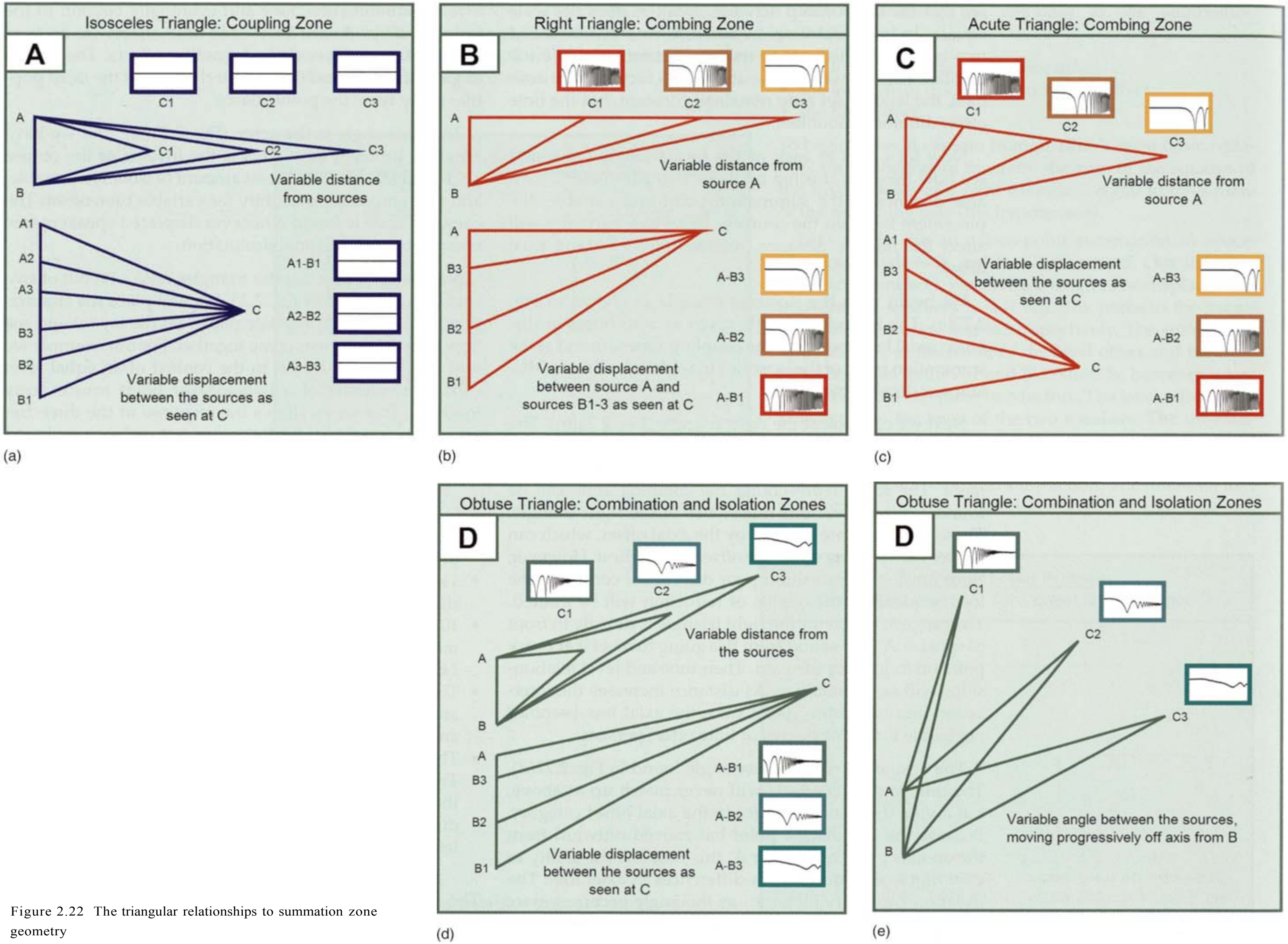


Figure 2.22 The triangular relationships to summation zone geometry

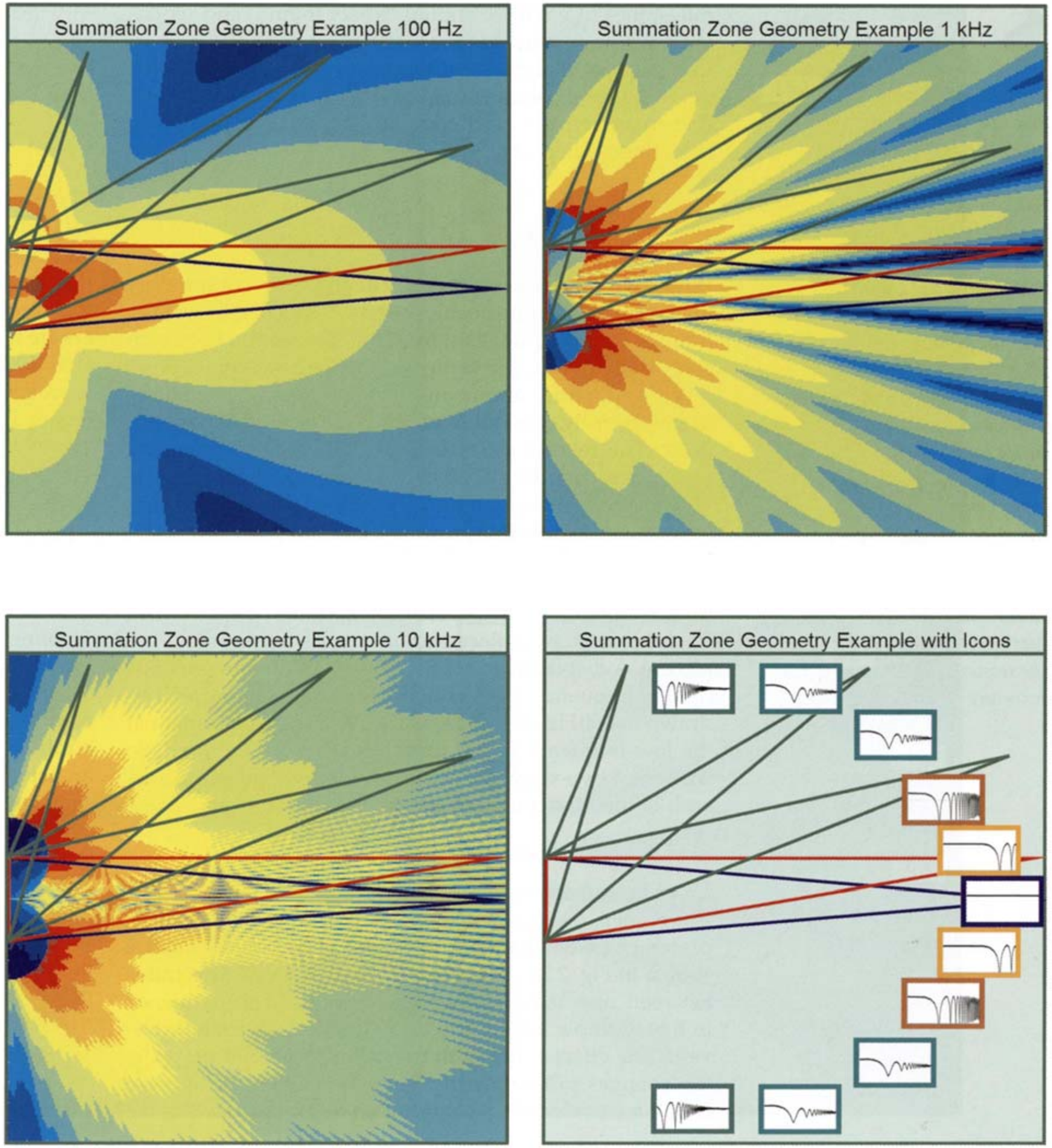


Figure 2.23 Summation geometry examples for two-element summation. Each element has a coverage pattern of 100 degrees. Spacing is 2m, with a 50 degree splay angle.
12m x 12m area

Perspectives: A lack of understanding of the interaction between components contributes to the idea that more is better. I was asked to tune a system with three horizontally arrayed speakers that were much too wide for this space. The system had many problems with interaction between components and bounce off the side walls. There was significant pressure from the client to use all of the rented resources. As a solution, I waited until the dinner break, and while everyone was inside eating, I unplugged two of the three speakers. I received many compliments for how well the event sounded, and especially comments that it was significantly improved after dinner as compared to our afternoon tuning session. Not a single person could identify the reason for this improvement, however!

Dave Revel

milestones of what we should expect to find, and where. We have now seen that the summation triangles give us that map for placement of the summation zone icons on the equal level contours map. As we move further we will find that even the most complex array summations will have these mileposts.

Wavelength Displacement

When two sources are physically displaced, it is a certainty that only the listening positions at the center of the isosceles triangle will be equidistant from the sources and have synchronized arrivals. As we move off of the array center point it is only a matter of time until we fall into a cancellation. More precisely, it is not a matter of *time* — it is a matter of *phase*. The relative delay leads to differences in relative phase over frequency. When the relative phase exceeds 120 degrees we are falling. If the displacement and frequency are known, it is possible to predict precisely where the summation will cancel and if and where it rises again. The decisive factor is the ratio of source displacement to wavelength. The wavelength and the displacement track together. A displacement of one wavelength draws a picture. A displacement of two draws a different one. The ratio is frequency-independent. The same picture will be drawn at 100Hz as 10kHz, as long as the displacement at the low frequency is 100 times greater than the high. So we have a two-variable system, and we will now illustrate each of their independent effects.

Fixed Frequency, Variable Displacement

If we keep the frequency constant and change the displacement, we will modify the ratio. The effects of this are shown in Fig. 2.24. The spacing is shown in various ratios between one-half and four wavelengths. The frequency in this example happens to be 100 Hz, but that is not relevant. The effect scales with frequency, provided the spacing changes proportionally. Note that when the spacing is half a wavelength, a single beam is concentrated in the forward (and rearward, although not shown) axis. This ability to narrow coverage with beam concentration will have an important effect on array design and will be covered later in Chapter 6.

Fixed Displacement, Variable Frequency

This scenario is of great concern to us, since in the practical world speakers exist at fixed distances from each other and cover wide frequency ranges. The speakers shown in the previous example (Fig. 2.24) will be reused to illustrate this second factor of wavelength displacement. The speakers are spaced 2 m apart. At the three frequencies of interest (100,1000 and 10,000 Hz) the displacement is 0.6, 6 and 60 wavelengths respectively. The number of nulls with a 90 degree polar quadrant will equal the number of displaced wavelengths between the sources. This can be seen as the incomplete null 100 Hz, the 6 nulls and 1 kHz and 60 nulls at 10 kHz. The cause of the null spacing is found in Fig. 2.25(a) where a series of concentric circles propagates from each speaker. The spacing of these phase contour lines corresponds to a single wavelength. Along the horizontal center line the circles cross and the familiar coupling is found since there is zero offset (marked as 0lamda). As we move off-center (up or down) the circles begin to diverge. The next point where the circles cross again represents a complete phase cycle (marked as 1lambda), and will result in addition. The one wavelength displacement (1 ms) creates an octave-wide peak as shown in the frequency response (our 1 ms icon). Further off-axis we turn additional phase cycles until the maximum of six cycles is reached (6 ms) at 90 degrees. The positions where the phase contour lines are farthest apart correspond to a null, since the phase cycles are 0.5 wavelengths apart. This point is shown as a red line on the 100 Hz screen. In all screens the frequency response at the 1 ms offset point is shown, with its coupling at 100, octave-wide combing at 1 kHz and 0.1 octave-wide combing at 10 kHz.

Speakers must be placed a real physical distance apart. Therefore, their displacement causes a different amount of phase shift over frequency.

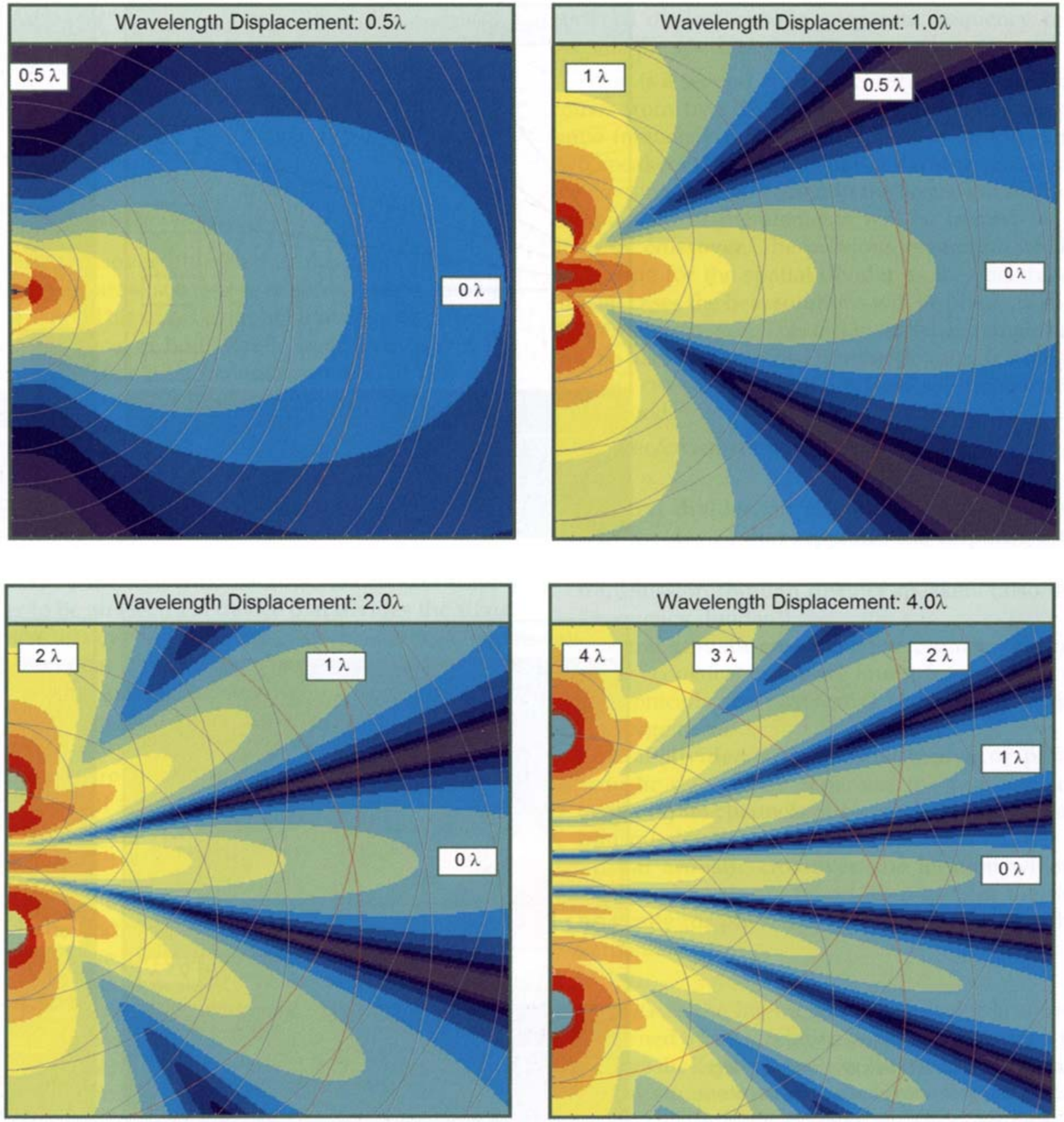


Figure 2.24 Wavelength displacement effects fixed frequency and variable displacement. The number of nulls/quadrant is equal to the wavelength displacement

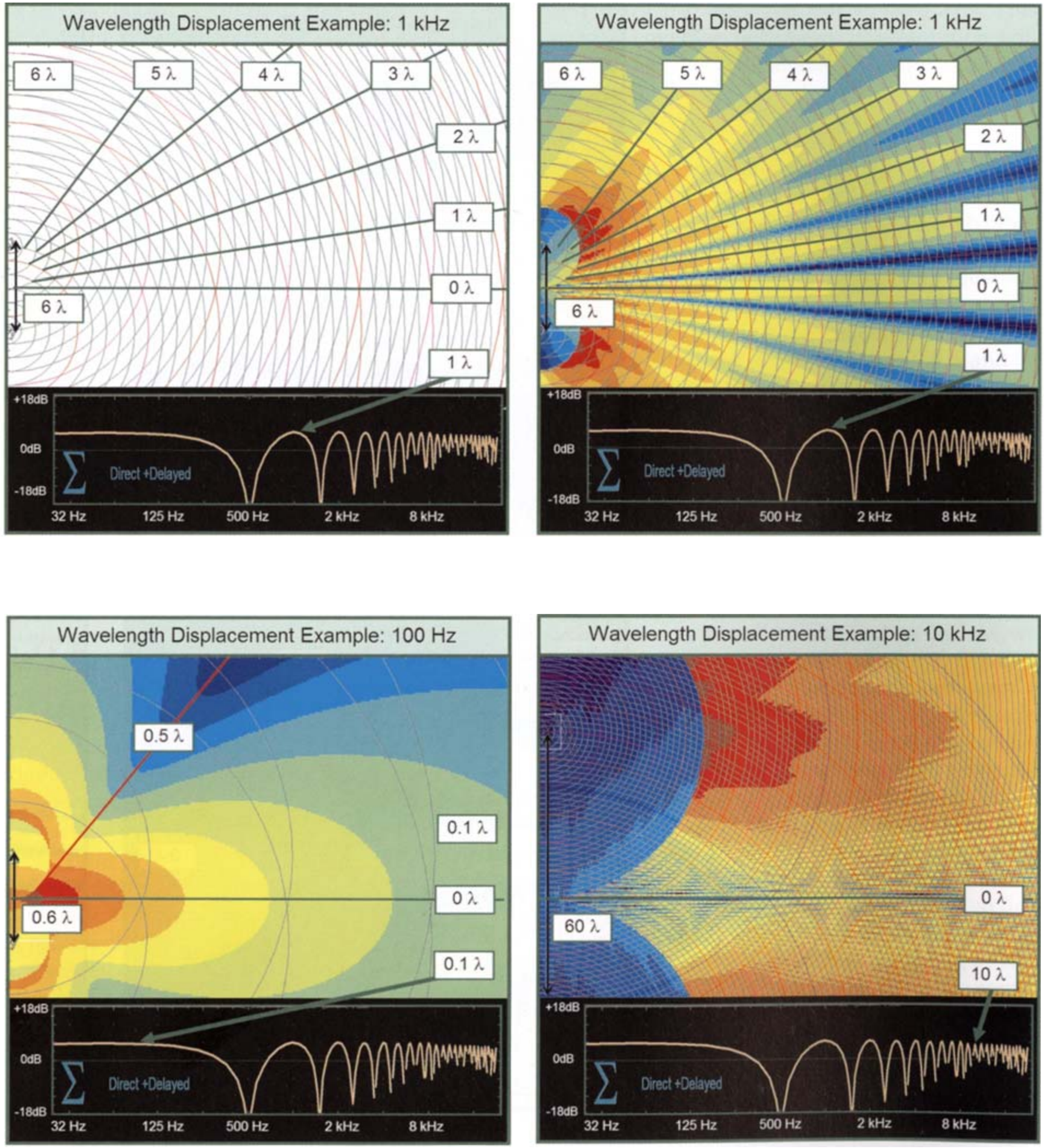


Figure 2.25 Wavelength displacement effects fixed displacement (2 m) and variable frequency (100 Hz, 1 kHz, 10 kHz). Concentric phase contour circles represent one wavelength each. The frequency response for a position with 1 ms time offset is shown at the bottom

86

Acoustical Crossovers

Acoustic crossover performance is governed by the rules of summation outlined previously. The summation zones will now take on practical context.

Acoustic Crossover Defined

An **acoustic crossover** is defined as the point where two signals from separate sources of common origin combine at equal energy. This may be a particular frequency range, a position in space, or both. An acoustic crossover is the most important type of summation junction. It is the one where the levels of two or more sources are equal, therefore the stakes are the highest. It is this position that will benefit the most from being "in phase" and be damaged the most from being "out of phase." As a result we will make every effort to align the phase responses at the acoustical crossover.

The users of the conventional audio lexicon consider a crossover to be an electrical device that divides the signal into low-frequency and high-frequency channels to be sent to separate drivers. That device is more properly termed a **spectral divider** (or frequency divider) since the crossing over will occur in the acoustical medium. Why does this matter? If we consider a spectral divider to be the "crossover" then we are assuming that the electrical response alone is sufficient for an expected outcome at the actual acoustical crossover. Such a leap of faith is likely to set us up for a fall.

The outcome at the acoustical crossover is dependent upon the individual electrical responses, acoustical responses, relative levels at the amplifier and the physical displacement of the speakers. The acoustical crossover is where the summation is made up of equal parts between the two contributors, regardless of what they had to go through to get there. Once that is known the investigation can begin as to how to best phase-align the crossover for maximum performance. That said we will now augment the conventional nomenclature of the acoustical crossover between devices covering separate frequency ranges to become the **spectral acoustic crossover.**

There is a second type of acoustical crossover, one that comes from two distinct speaker elements covering the same frequency range. The separation of the signal into two speakers is termed a **spatial divider,** and once again the crossing over will occur in the acoustical medium. The combination of the elements will be termed the **spatial acoustic crossover.** The previous paragraph holds every bit as true for the spatial divider as the spectral divider. In both cases the best solutions will be phase alignment in the crossover region. This fact means that a single language can be used to describe the analogous features of the acoustic crossovers and the related solutions more easily understood.

Divider / crossover terminology:

- **Spectral divider:** an electronic (active or passive) or acoustic device that separates the response into separate high- and low- (or multiple) frequency channels for transmission through distinct speakers (also termed a "frequency divider").
- **Spectral acoustic crossover:** the frequency range where equal acoustical levels are found from each of the two electronically (or physically) separated elements that converge there.
- **Spatial divider:** an electronic (active or passive) or acoustic device that separates the response into two (or multiple) channels of common frequency range for transmission through distinct speakers.
- **Spatial acoustic crossover:** the location where equal acoustical levels are found from each of the two electronically (or physically) separated elements that converge there.

In both cases the interaction of the multiple sources will be governed by their relative level and phase responses. In both cases the keys to success will be in careful management of the five summation zones. In those areas where coverage is desired we must limit the overlap in coverage to areas where the additive side of the phase cycle is maximized

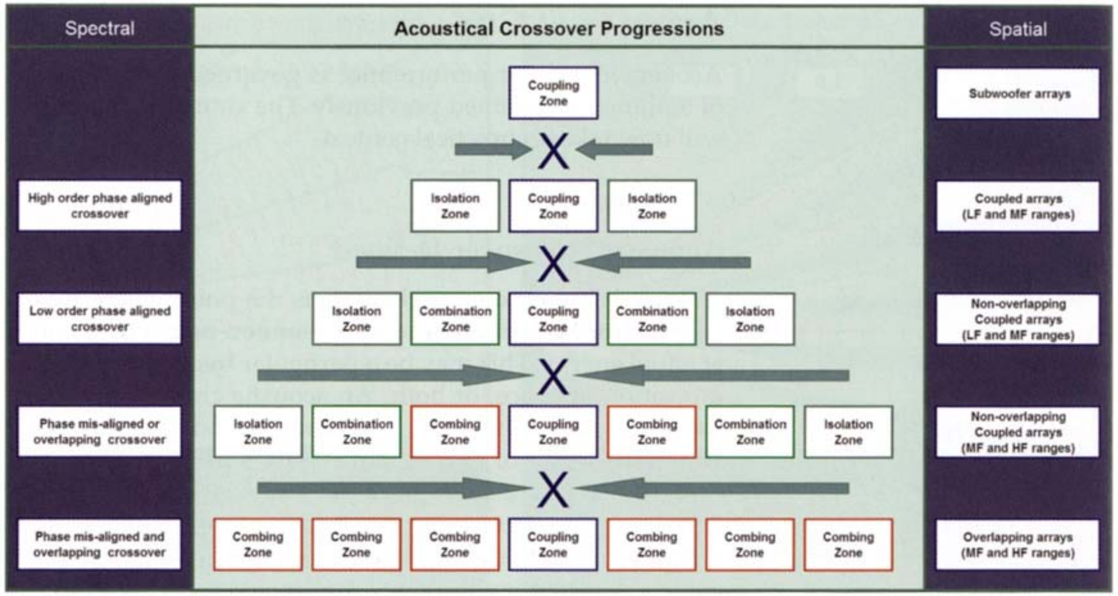


Figure 2.26 The acoustical crossover progression scheme. There ar
point. The ideal cases are those without combing zone interaction (upper
tiers) and the worst-case scenarios (lower tier) contain widespread combing. Typic

Perspectives: Optimization tips from a monitor mix engineer

1. *Taking sufficient time at the beginning for optimization of the system is the surest and fastest way to success.*
2. *For checking the system, listening to the pink noise has become much easier for me now than listening to unfamiliar CD.*
3. *Stage monitors can get the maximum feedback margin by taking the same method as the optimization for FOH speakers.*
4. *We used to be given no time for optimizing the stage monitors because of the time budget. People now understand that the optimization of stage monitors*

and the subtractive side minimized. For these areas we will design our crossover to make maximum use of the coupling zone. For areas where coverage is not desired we will use the isolation or cancellation zones. As we move away from the crossover areas we will attempt to escape through the combining zone with the minimum effects of the combing zone. The journey through crossover is shown in Fig. 2.26. The best-case scenario is shown at the top of the pyramid, where the coupling zone is the only product of the summation. This is typical in the spatial crossover of subwoofer arrays where the displacement is small compared to the low-frequency wavelengths. In the next level the transition goes directly through crossover without any significant loss as it moves from coupling to isolation. Such a scenario is practically attainable with a spectral divider, or the low- and mid-frequency ranges is a closely coupled spatial divider. The lower part of the pyramid represents a typical journey for a full-range speaker through a spatial crossover, with the lowest level showing the hazards of high overlap and large displacements. The challenge is to find ways to spend the minimum amount of our transition in the combing zone. The middle ground of the pyramid is typical of spectral dividers and some spatial dividers.

Welcome to the summation race.

A similar scenario plays out in reverse in cases where acoustic addition is not desired. In such cases the cancellation zone may be used to steer sound away from a given area, or the isolation zone may be placed at the crossover center in order to reduce the sound there. In either case we will strive to transition out of those zones with a minimal percentage of combing zone effects.

Crossover Classes

Acoustic crossovers can be characterized as having certain common qualities. Their response through the crossover can be classified as either unity, overlapped or gapped. Each crossover class will have specific applications in our system design.

helps rehearsal to go much smoother.

5. *Ever since 1 started doing this job, I have become more careful about things such as the position and response of microphones we use on stage.*
6. *I have realized that precisely aiming speakers, setting levels and arraying them are far more meaningful than EQ and delay setting.*
7. *In the case of sound engineers who do not care for our way of optimization (EQ in particular), it is worthwhile, even just to help them verify the speaker system.*

Hiro Tomioka
ATL Inc., Japan

Crossover classes:

- **Unity:** the level through the crossover matches the surrounding response level. The unity class crossover is expressed by the equation: (—6 dB) + (— 6dB) = O dB.
- **Overlapped:** the level through crossover is higher than the surrounding response level. The overlap class crossover is expressed by the equation: (X dB + X dB) > O dB.
- **Gapped:** the level through crossover is lower than the surrounding response level. The gap class crossover is expressed by the equation: (X dB + X dB) < O dB.

Acoustical crossovers share common features. These features will control how well our crossover designs will suit our needs.

Crossover features:

- Crossover location (frequency or location)
- Degree of overlap (unity, overlap or gap)
- Crossover slope (rate of transition towards isolation)
- Degree of symmetry (level and slope)
- Degree of audibility (how perceptible the transition is to our hearing).

Spectral Dividers and Spectral Crossovers

Spectral dividers will normally fall into just two of the three crossover classes: unity or overlap. A key factor for acoustical crossover success is the confinement of the overlap to the range where phase addition occurs. A spectral divider will use four of the summation zones: coupling, combing, combining and isolation. This discussion will begin with the unity crossover class, wherein all of the terms will be introduced. The overlap crossover discussion will follow, giving comparison and contrast as needed.

The coupling zone is the region at the very center of the crossover, where the phase responses match. To put crossover coupling in perspective we will introduce another simple formula.

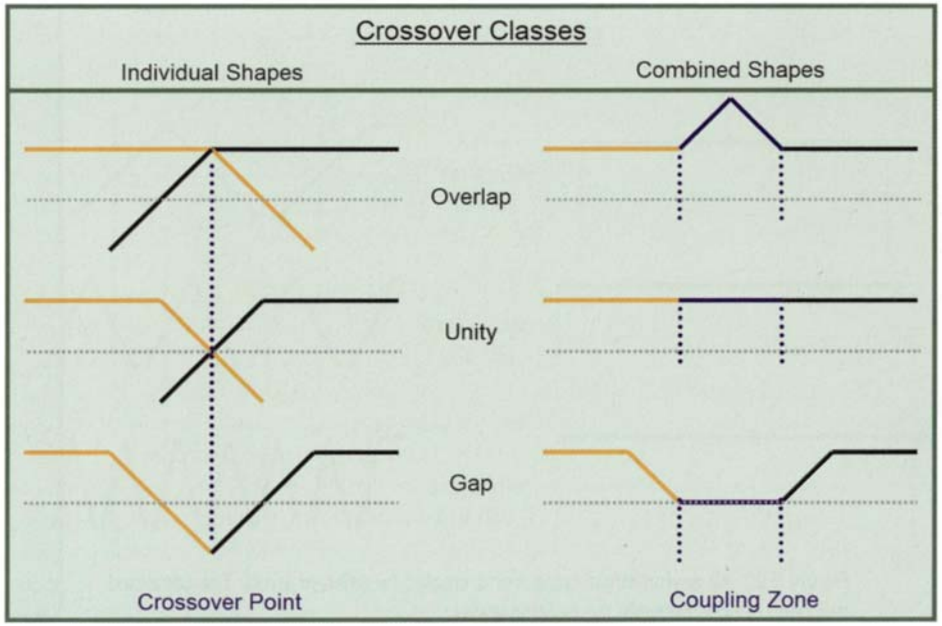


Figure 2.27 Crossover classes. The degree of overlap determines the class. The unity class acoustic crossover has equal level in the combined and isolated areas

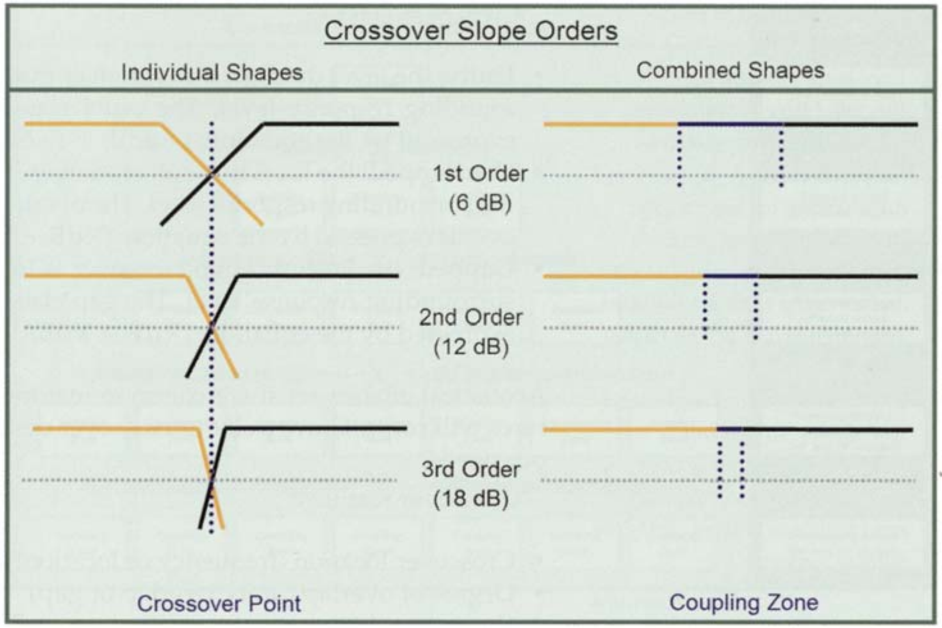


Figure 2.28 Crossover slope order. The slope determines the size of the combined area in proportion to the isolated areas. As slope increases the coupling zone shrinks and the individual elements must be closely spaced. High-order elements must be closely spaced to maintain unity class crossover performance

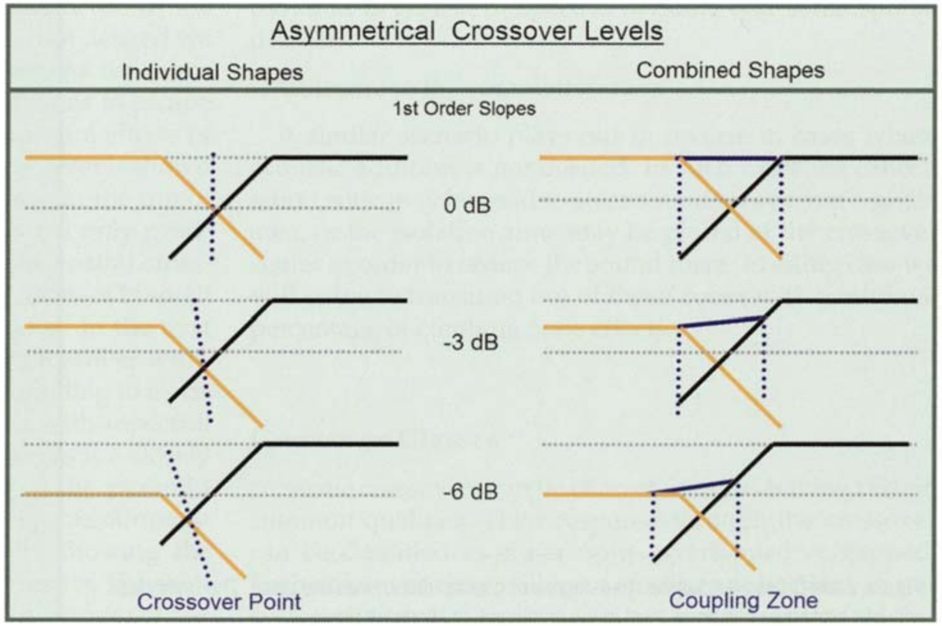


Figure 2.29 An asymmetrical crossover is created by different levels. The combined response bridges between the isolated areas

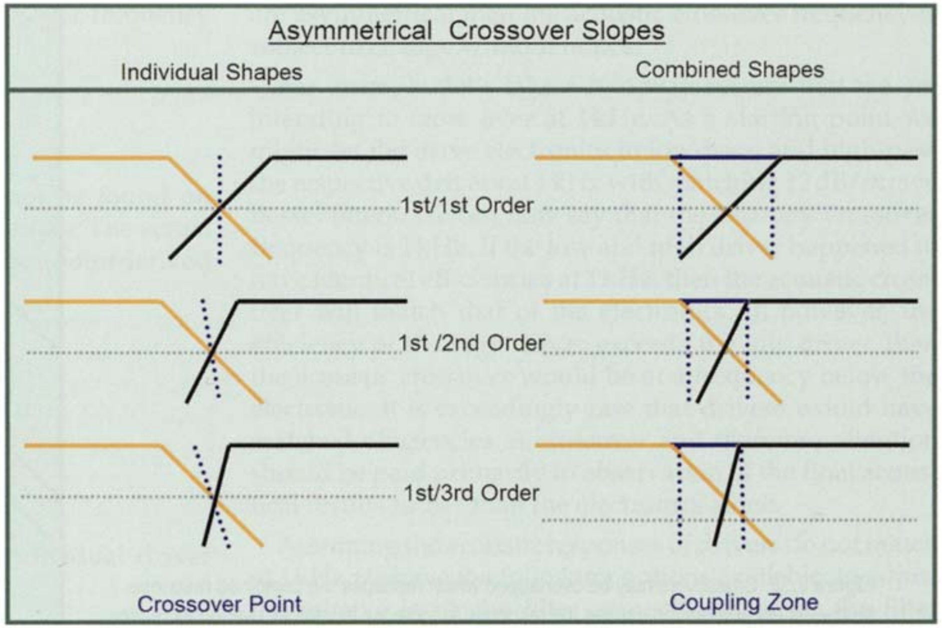


Figure 2.30 An asymmetrical crossover is created by unmatched slopes. The size of the combined area and the spacing between the elements is affected

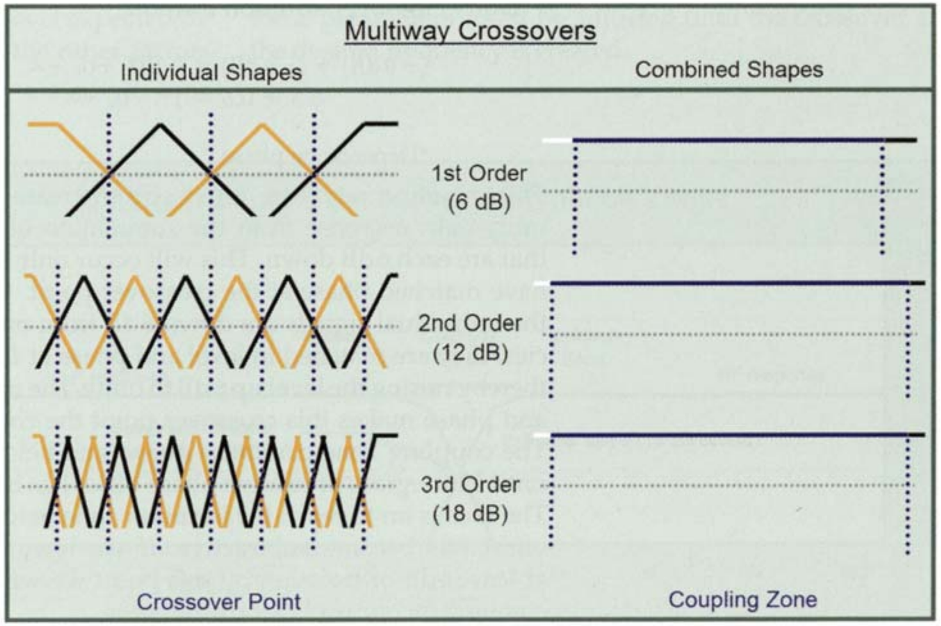


Figure 2.31 The overall range may be covered by an infinite number of multiway crossovers. A comparable response can be made from small quantities of widely spaced low order elements or from larger quantities of closely spaced high order elements

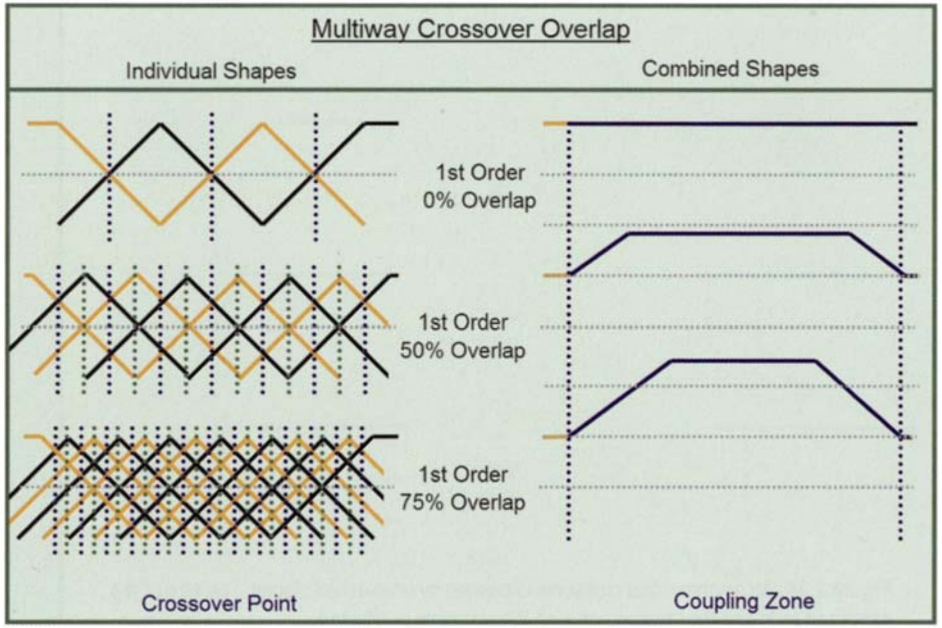


Figure 2.32 Crossovers may be overlapped which reshapes the combined response. Multi-element summation causes the response to increase in level as overlap increases

Unity crossover addition formula:

$$(-6\ \text{dB}) + (-6\ \text{dB}) = 0\ \text{dB}^* \ +0, -\infty \text{ or}$$
$$0.5 + 0.5 = 1^* \ +0, -\infty$$

*Depends on phase

This equation refers to our desire to create a combined unity gain response from the summation of two signals that are each 6 dB down. This will occur only if the signals have matched phase at the crossover point. Even though the individual signals are down 6 dB from other frequencies, they are matched in level and phase at the crossover, thereby raising the level up 6 dB to unity. The matched level and phase makes this crossover point the coupling zone. The coupling zone continues above and below crossover until 120 degrees of relative phase offset has been reached. The race is on to reach the isolation zone before the phase offset has become subtractive. If we have not reached at least 4 dB of isolation by this point we will experience combing on our way toward isolation.

The shape of the crossover is controlled by the slopes of the filters that comprise it. The size of the overlap zone is dependent upon three factors: filter corner frequency, topology and slope. These factors together will govern the degree of interaction between the two elements and to what extent we are able to isolate before the phase responses move into cancellation. This is not simply a matter of "steeper is better." Phase delay increases as filter slope increases. There is a tedious give and take to the choice of center frequency and slope to optimize a filter response in a spectral divider. Add to this the individual characteristics of the driver elements and their physical placement and the complexity of the challenge becomes clear. Because the two speaker elements must always be displaced in space there are resulting challenges to the polar response due to the relative proximity of one driver to the other in space. For example, a high and low driver in a vertical orientation can be aligned and in phase at the physical center point. However, above and below that center point the distance between the drivers is no longer equal. This will result in a different phase relationship at the crossover frequency over the course of the vertical coverage.

Crossover Frequency

The acoustical crossover frequency cannot be found on the front panel of an electronic spectral divider. The actual crossover frequency is an acoustic summation point derived from the confluence of five factors.

Factors affecting crossover frequency:

- Relative filter slope
- Relative filter corner frequency
- Relative filter topology
- Relative drive levels
- Relative speaker position
- Relative speaker efficiency and/or individual driver parameters.

If the first five factors are made symmetrical they will cancel each other out, leaving only the driver efficiency as unknown on the equation. If the high driver is more efficient the crossover will be below the level expected by the electronics and vice versa. If any of the other factors are asymmetrical then the acoustic crossover frequency is subject to change without notice.

For example, let's take a two-way system that we are intending to cross over at 1 kHz. As a starting point we might set the drive electronics to low-pass, and high-pass the respective drivers at 1 kHz with matching 12 dB/octave Bessel filters. We can now say that the *electronic* crossover frequency is 1 kHz. If the low and high driver happened to have identical efficiencies at 1 kHz, then the acoustic crossover will match that of the electronics. If, however, the efficiency of the high driver exceeds the low driver, then the acoustic crossover would be at a frequency below the electronic. It is exceedingly rare that drivers would have matched efficiencies at crossover and therefore attention should be paid primarily to observation of the final acoustical results rather than the electronics alone.

Assuming the acoustic responses of drivers do not match at 1 kHz we have the following options available: to adjust the relative level, the filter corner frequencies, the filter topologies, or the filter slopes. Some asymmetry must be introduced in the electrical system to compensate for the asymmetry in the acoustical system. A combination of these parameters could be adjusted until the crossover at the desired frequency is created.

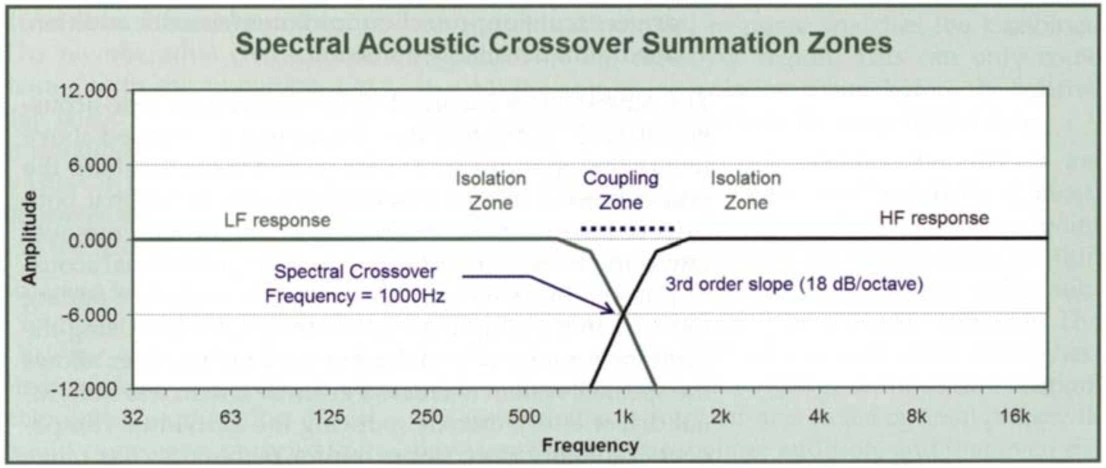


Figure 2.33 Spectral acoustic crossover summation zones

Filter Order

The "crossover range" is defined as the area around the crossover center frequency where the two signals remain highly interactive. The crossover range may contain a variable mix of the coupling, combing and combining zones, but is limited at either end by emergence into the isolation zone. The slope of the filters defines the size of the crossover range, with steeper filters having a range-shortening effect. The goal is to isolate before cancellation, a race of amplitude vs. phase. The strategy would seem very simple, i.e. use steep filters which isolate quickly, and this simple solution has gained great popularity in the industry. However, this practice has an aspect of "a tiger chasing its tail" in that the phase delay is greatly increased by the steeper filters. Therefore, as the filters decrease the size of the crossover range, they increase the relative phase shift therein. Steep filters also require the corner frequencies to become closer together to prevent gapping. The result is that while the size of the crossover range is reduced, the degree of volatility inside that range is increased. Like most everything in life, we have a tradeoff.

Crossover filter slope tradeoffs:

- Low order: large range, low phase shift, minimum HF driver excursion protection, maximum acoustic addition, minimally detectable transition.
- High Order: small range, high phase shift, maximum HF driver excursion protection, minimum acoustic addition, maximum detectable transition.

The other factors involved in the tradeoff include acoustic addition. The properties of summation outlined above reveal the opportunity for the two devices to share the power burden in the crossover area. Assuming that both drivers have suitable power capability in the crossover range, there is a potential for up to 6 dB of additional acoustic power. This power addition comes from drivers adding together in the coupling and combining zones. Using the combined power of two drivers working together allows the speaker system to create a given SPL with less individual driver effort, thereby reducing the individual component excursion. Steep slopes minimize the crossover range in the hopes of minimizing the negative phase interaction (combing zone). This practice also minimizes the effects of additive phase interaction, i.e. the positive effects of two drivers coupling around the crossover center (coupling zone). This compromise hinges on the following question: are the positive effects of power addition able to outweigh the negative effects of potential phase cancellation? The answer is not simple but the following trend is noteworthy: the lower the crossover frequency, the more favorable the conditions are for a wide crossover range. This is due to the large wavelength size relative to the physical driver offset and to the tangible benefit of utilizing every available ounce of low-frequency power addition possible. As frequency increases, the wavelengths are small compared to the physical displacement, and isolation from phase cancellation becomes more crucial. In addition, the HF drivers tend to be more fragile mechanically, which bodes well for steeper filters to minimize excursion.

Crossovers slopes may optionally be asymmetrical. The slopes of the high and low drivers may be different orders. Asymmetrical crossovers usually contain steeper filters for the high driver to improve excursion protection, and broader filters for low driver to extend power assistance into the crossover area.

Filter Topology

There are various filter topologies such as Bessel, Butterworth and Linkwitz-Riley. As discussed in the first chapter, these differ primarily in how the area near the corner frequency begins its descent. The merits of one type over the other can be debated until the cows come home. In the end, what matters is the result of the confluence of the electrical and acoustical factors, as seen in the final combination. I restate my position that the creation of engineered systems with phase-aligned crossovers is the responsibility of the manufacturer of the system. These should be designed in anechoic conditions and thoroughly researched for maximum component reliability, stability and maximum power addition. Attempting to design acoustical crossovers of this type in the decidedly

Perspectives: The key to successful system "tuning" is: understanding how changing one parameter can affect a seeming unrelated issue. Examples of this are the role that crossover and delay settings can play in affecting the directional response of a loudspeaker or array of loudspeakers.

Sam Berkow

challenging conditions of the field is something I try never to do. If we don't trust the spectral crossover parameters given us by the manufacturers, why would we trust their speakers?

Crossover Audibility

Our efforts to create an accurate transmission include suppression of clues that we are listening to a speaker rather than the original artist. Violins do not have crossover frequencies. An ideal multiway speaker system will be able to play all notes of the scale without the listener noticing a transition between drivers. Efforts to mask the presence of a crossover in our speaker will help greatly in this regard. There are several mechanisms that can expose the transition between drivers. The first is driver displacement. The probability of distinct localization increases as drivers move apart. This is not an issue for most modern sound systems since the boxes contain both high and low drivers in close proximity. This can be an issue for subwoofers, which are often a large distance apart from the mains. The second factor is overlap, which can expose the crossover transition by having too much or too little overlap. Those with extremely steep filters provide a sudden transition that can occur in a single note of the musical scale. This becomes most audible if the transition moves between elements which exhibit large differences in pattern control as is found when we transition from a front-loaded cone driver to a narrow horn. The reverberation character of the sound can suddenly change with the transition. On the other hand, if the overlap is too high, the filters will not have achieved sufficient isolation to prevent combing around the crossover center. This can occur either above or below crossover (or both) and leaves the crossover exposed by the presence of dropouts in the response.

Since we are lacking an accepted term for this phenomenon it is hereby dubbed "crossover audibility" and is defined as: the ability of a listener to hear the transition from a loudspeaker component working in one frequency range (or location) into another driver working in a different range (or location). It should come as no surprise that our goal would be to make the crossover as inaudible as possible. Likewise it should come as no surprise that minimal displacement, minimal coverage angle transition, minimal combing and more gradual filter slopes have the highest prospects of slipping under our sonar. The clues that our hearing mechanism picks up are abrupt changes in the sonic character between notes in a scale. Displacement causes an abrupt change in localization. Combing causes an abrupt change in level as one note disappears and the next returns. Coverage angle transition causes a change in the reverberant field, leading the listener to feel as if one frequency is far away in a reverberant space while the next is nearby in a dry space.

Crossover Overlap

The elements will overlap in the crossover region. If the elements combine at their -6 dB points, the center will sum to 0 dB (a unity class crossover), provided the phase response is matched. At frequencies above and below the crossover point, the amplitude and phase responses will both begin the process of moving apart. The result is successively lesser amounts of addition, which is exactly what is required to maintain a constant summed level. Our design intent is that the remaining addition will fill in for the individual response rolloffs. Once isolation has been achieved, the extent of the addition tapers such that the individual nominal response matches the combined responses in the crossover region. This can only come about if the amplitude isolation occurs before the relative phase response has moved into the cancellation side.

With non-coaxial two-way systems the drivers are physically displaced in at least (and hopefully at most) one direction: vertical, or horizontal. The crossover point can only be perfectly optimized for a single point within the coverage of the system. The choice of this point, subject to discussion, is not one to be selected arbitrarily. The best choice for such point is the one that yields the highest degree of usable crossover coverage without cancellation. It is given that any point off-axis to the optimal point will have less than a perfect phase addition, and that even the best alignment will have cancellation points somewhere, but that does not prevent us from having objective criteria for choosing this point.

Specification:

- The spectral crossover optimization point should be chosen to provide the maximum angle of coverage without cancellation in the crossover range.
- The angle should remain as constant as possible over distance.

The most common mistake in this matter is the tendency to optimize the system for on-axis to the HF horn. The roots for this practice are understandable since the HF horn provides the principal source of angular control for the system. The vulnerability of this approach can be found by returning to the triangulation concepts discussed earlier. Alignment to the horn axis introduces the right triangle summation response and therefore cannot hold its synchronicity over distance. The phase optimization can only work at one distance. At distances closer than this point the HF driver leads and at points behind, the LF driver leads. Such approach also results in an asymmetrical response at angles off the center line of the horn. This is due to the different rates of change over angle. In one direction you move away from both drivers, in the other you move first toward the LF and then away from both.

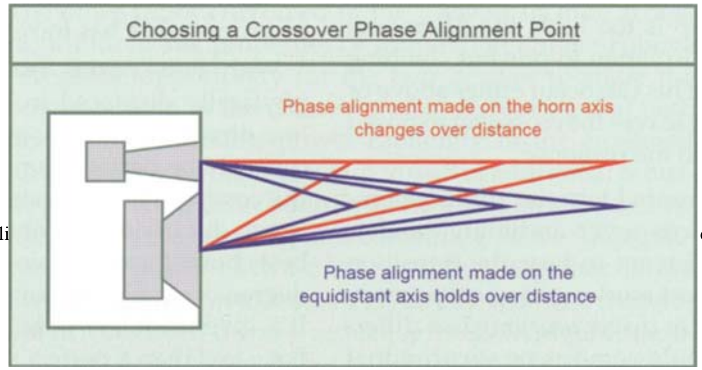


Figure 2.34 Choosing a phase ali... er distance. Note that the spectral crossover (HF range to LF range) is also a spatial crossover (HF location to LF location)

A superior method comes from using the equidistant center between the drivers. Any difference in depth will be compensated by delay. We now have the familiar isosceles triangle of the coupling zone. This approach creates a situation where the relative angle change and relative distance change occur at the same rate. With this method, an optimization done in the relative near field will hold up over long distances. It is also worth noting that we have already begun to apply the concepts of the spatial divider (in this case between the HF and LF drivers) in order to optimize the combination effects of the spectral divider. The linkage between the two domains has begun.

An alternate approach regarding spectral crossovers is that the center frequency should be joined with the two elements being at their —3 dB points. If the responses are phase-aligned, the summation response will show a 3dB rise, creating an overlap class crossover. This can be returned to a unity class crossover by setting the relative phase responses to 90 degrees apart. However, 90 degrees apart as a starting point is dangerously close to the edge of the phase abyss that begins at 120 degrees. Therefore, the isolation between the drivers must occur very rapidly to avoid the slide into cancellation. This is challenging since bringing the drivers together at their —3 dB points causes the overlap area to be more interactive. Such an approach requires the use of steep filters in order to isolate quickly. This limits the usable crossover range substantially and reduces the power sharing between drivers. The zero degree overlap method has a 3 dB combined power advantage and 90 degree phase addition advantage over the —3 dB, 90 degree method in the race against cancellation.

An overlap class spectral crossover is used primarily in low-frequency applications, where the wavelengths are sufficiently long to remain in the coupling zone. The typical application is the overlap of subwoofers and full-range enclosures. These systems can share the range from 60 Hz to 120 Hz, and there are good reasons to take advantage of the redundancy. First, this frequency range has some of the highest power demands in the entire system. We can use all of the headroom we can get. Second, the physical displacement of the drivers tends to be close enough that the phase

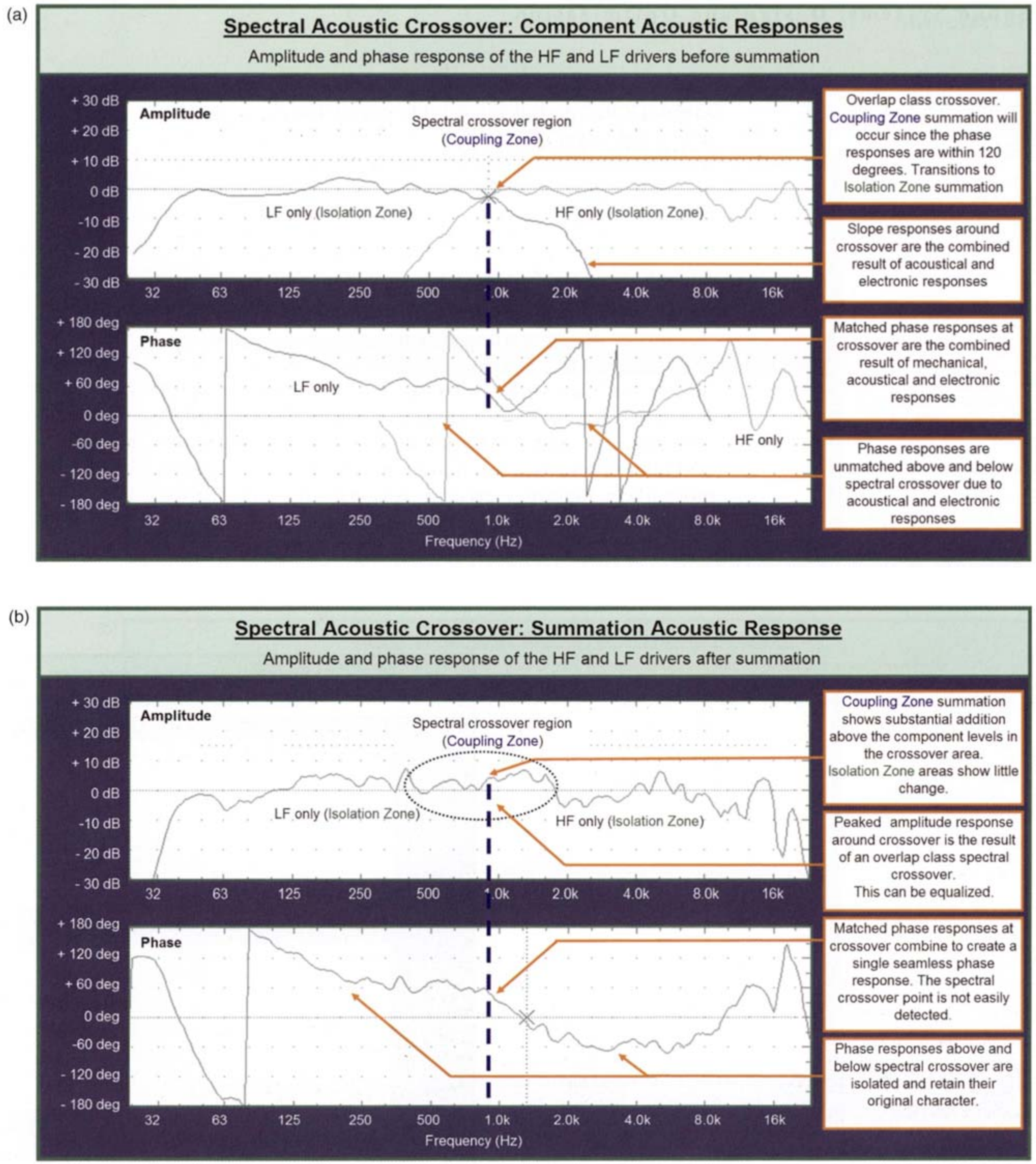


Figure 2.35 Spectral crossover alignment example (a) Individual responses of two-way loudspeaker. Note the convergence of the phase traces around the spectral acoustic crossover point of 900 Hz. (b) The combined response of the same speakers shows a coupling zone transition to isolation without combing. The crossover progression shown here is found in the third tier of the pyramid shown in Fig. 2.26

responses remain in the coupling zone over a substantial range. Third, crossover audibility can be decreased in cases where the main speakers are flying and the subwoofers are grounded. Finally, the directional control of the system may be improved by the additional beam concentration obtained through the presence of additional sources.

Crossover Asymmetry

There are two common forms of asymmetry in acoustical crossovers: slope and level. If the two components contain different rolloff rates, the transition out of the coupling zone toward the isolation zone will be asymmetrical. Asymmetrical slope rates can be used effectively, but their action must be anticipated. The mixing of even and odd filter orders (e.g. second order and third order) would most likely require a polarity reversal and mismatched corner frequencies to achieve a unity crossover result.

Asymmetrical level setting can also cause unexpected results if the effects are not compensated. It is given that the relative level of the component systems is matched at the crossover frequency If the level of one of the components is changed, the crossover frequency shifts. If the HF level is raised the crossover moves downward and vice versa. There are a number of important ramifications of this action, which illustrate that such a decision should not be taken lightly. First let's clarify an important aspect regarding acoustic power and spectral crossovers.

A mix engineer needs a certain amount of power capability in the system to obtain the desired sound. That quantity is not found by a single number. Even if we were able to give an engineer a system that delivers 130 dB ,SPL, this is not necessarily enough level to mix even soft lullaby music. Why? Because the proposed 130 dB SPL system might be 64 subwoofers and a single bookshelf speaker. It can deliver 130 dB SPL all the way up to 100 Hz. After that, good luck.

What is required is a system able to deliver the required level at all frequencies. That does not mean that 130 dB

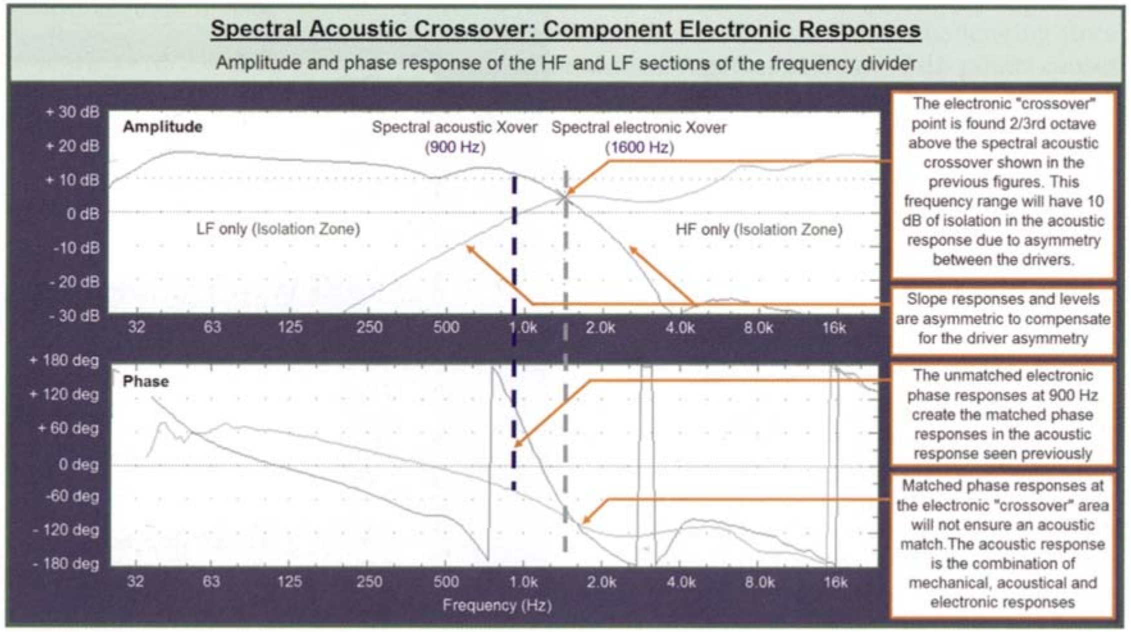


Figure 2.36 Asymmetrical spectral crossover slopes. The slope of the low-pass filter is steeper (third order) than the high-pass filter (second order). Manufacturers and end users employ a huge variety of asymmetrical filter sets for such applications

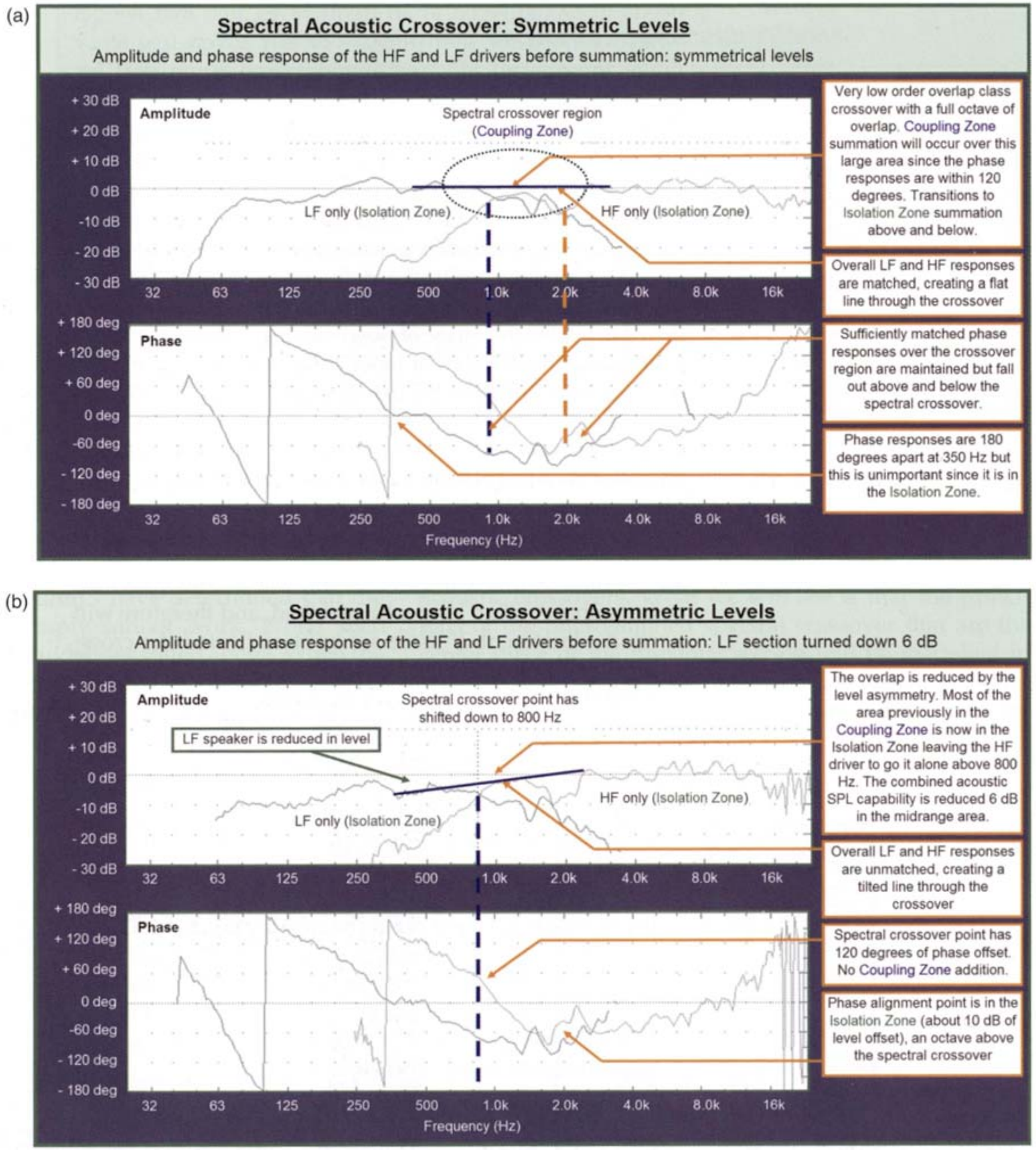


Figure **2.37** Asymmetrical crossover level setting, (a) The HF and LF channels overlap in the 1200 Hz range. The phase responses are also aligned in this range, ensuring a coupling zone summation with minimal combing zone, (b) The level of the LF amplifier was reduced 6 dB. The spectral acoustic crossover is now around 800 Hz. The phase responses are nearly 150 degrees apart at this frequency, ensuring combing zone summation around the spectral acoustic crossover needs to be delivered at all frequencies, just that we get what we need for that music. For any given frequency range there is an SPL requirement, and some part (or parts) of the system must deliver it.

Here is how this relates to the spectral crossover. The mix engineer has a power need and is not concerned about which driver will deliver it. If we set our two-way crossover to 1kHz we will deliver the power needed in the 1kHz range via the combined response of both drivers. The areas above and below will take delivery from their isolated components alone. If we were then to move the crossover down to 500Hz the range of power sharing moves down an octave and the division of labor proceeds from there. This means that at 1kHz the HF driver must now go it alone and it will need to supply an additional 6dB of power to make up for the loss of LF driver assistance. The mixer's needs have not changed. The acoustic power is still needed. The question is simply who will deliver it.

Effects of a relative level on the spectral crossover range:

- The relative phase response, if matched at the previous frequency, will no longer be matched, and therefore will need adjustment. If these are left unadjusted the crossover will not add to full power and the efficiency and reliability of the system will go down.
- The polar response of the HF and LF components changes over frequency. Therefore the rate of polar change through crossover may be affected. For example, HF horns tend to become wider as the cutoff frequency is lowered, while LF drivers narrow as the cutoff rises.
- The choice of cutoff frequency has a very large impact on the excursion required of the HF driver to produce the same acoustic power at the cutoff frequency. This is a complex subject beyond the scope of this text, but basically amounts to this: the excursion rises exponentially as the frequency drops. For example, moving the frequency down an octave (0.5f) requires 4x the excursion to produce the same acoustic power.
- The lowering of a crossover may cause the device to operate below the resonant range of the horn, making for lower efficiency and requiring still more excursion to achieve the demanded SPL.
- For systems with steep crossovers, an offset in level creates a "step function" in the frequency response, i.e. the crossover area shows a sudden, steep, cliff-like rise in response. As the crossover slope rises, the level change will have a lesser frequency-shifting effect and instead will create a "fault zone" at crossover. Needless to say, this is an undesirable listening condition.

The most common justification for crossover level changes is the perception that the user can reduce the system equalization needs. The concept is that systems couple much more extensively at the low frequencies than they do at the highs. Therefore a spectral balancing is performed which reduces the need for large cuts in the low mid-range areas by turning down the LF drivers. The fallacy is that few systems couple evenly from the low frequencies up to the HF driver crossover frequency. The result is excess HF driver wear and tear in order to keep from wearing out the parametric filters. I have seen many a fractured HF driver from this practice but never yet seen a filter damaged because it had worked too hard removing a low mid-peak.

The rate at which the center frequency shifts due to level changes is dependent upon the slope of the crossover. Low slopes will change at the highest rate, high slopes at the slowest. An infinite slope will never change the center frequency. This should not be misunderstood as an advantage for the steep slopes. Steep crossovers with large level offsets are extremely audible as the combined response has an abrupt, step-like rise in level at the crossover. To picture this, visualize an infinitely steep crossover as two blocks placed side by side. As you slide one block up, the center between them does not change, but rather becomes a vertical cliff at crossover. As crossover slope increases greater care must be taken to level match at the center.

Final note: the recommendations of manufacturers regarding spectral dividers should be treated with respect. A reputable manufacturer invests considerable research into obtaining optimized parameters. They have knowledge of the system power capabilities which cannot be obtained in the field (except the hard way). Makers are able

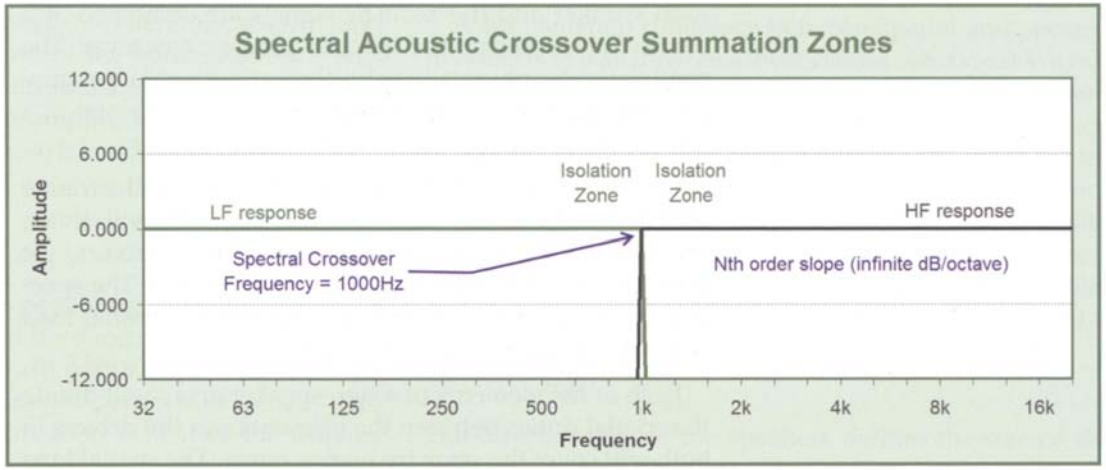


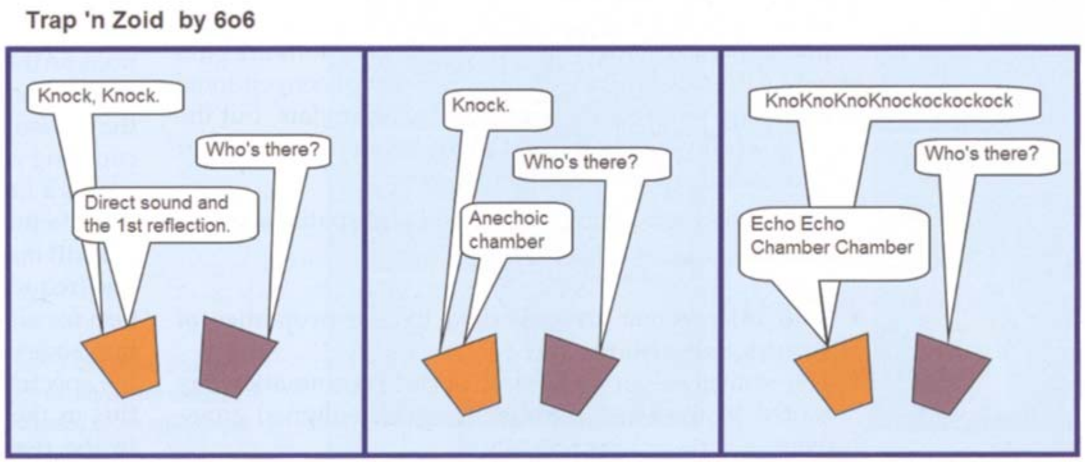
Figure 2.38 Spectral acoustic crossover with infinitely steep (theor

to measure in controlled acoustic conditions, etc. The complexity of the task is daunting. Although I have been a practitioner in system alignment for over twenty years, I do not make it a practice to deviate from manufacturers' settings unless and until I have determined that the delivered parameters simply cannot work. When such cases occur, I follow the guidelines stated above. When the manufacturers' recommended parameters work, I happily move on to other matters.

Spatial Dividers and Spatial Crossovers

Spectral vs. Spatial

This section initiates a paradigm shift in the approach to acoustic crossovers. What we will see is that the principals of the phase-aligned spectral crossover that are the center piece of the previous section will be reapplied in a new context: the spatial crossover. The spectral divider physically splits the spectrum into separate signals for



each speaker and the acoustic signals are combined at a particular frequency: the spectral acoustic crossover. The spatial divider physically splits the entire signal into separate speakers and the acoustic signals are combined at a particular location: the spatial acoustic crossover.

Let's take a pair of two-way speakers as an illustrative example. Each of the single two-way speakers will divide the spectral duties between the HF and LF drivers, yet both drivers will cover the same listening area. The spectral load (the high and low) is split, but the spatial load (the coverage area) is shared equally.

Each of the elements of a two-speaker array will divide the spatial duties between the elements, yet the drivers in both will cover the same frequency range. The spatial load is split, but the spectral load is shared equally.

Such an array contains both species of acoustic crossover, and the approach that we will use to optimize the response in this most sensitive area will be the same: the phase-aligned crossover; i.e. wherever two signal are found to be equal in level, they must be made to be equal in phase.

What we will see in this section is that spatial division, the process of separating our listening area into zones, is so directly analogous to the separation of high, low and mid-range drivers that a single set of principals will apply to a four-way crossover in a single enclosure (spectral) or a four-element array (spatial). The final piece of the puzzle will fall into place as we find that the walls of the room are the ultimate spatial dividers and that their reflections are governed by the same principals. The revision of conventional terminology will require some time to assimilate, but the effort is worthwhile in the end as the mysteries of speaker and room interaction yield their secrets.

Common ground between spectral and spatial acoustic crossovers:

- Both interactions are governed by the properties of acoustic summation.
- The strategies for achieving optimal summation are rooted in the same concept: the phase-aligned crossover.

Differences between spectral and spatial acoustic crossovers:

- The spatial crossover can run through the full audio range and therefore is much more sensitive in terms of our ability to isolate without cancellation.
- Whereas a relative phase change in the spectral divider affects only the frequency range of that crossover, the relative phase in a spatial divider affects all frequencies.

Analogous functions:

- The position in the room where two speakers share equal level is analogous to the spectral divider crossover frequency. This area will once again be the coupling zone.
- The directional control capability of the speaker is analogous to the spectral divider slope. A highly directional speaker is like a steep crossover.
- A change in relative level between speakers becomes a change in the position of the spatial crossover point. This is analogous to the crossover frequency shift in the spectral divider that results from changing the level of one of the drivers.
- The power addition capability that comes from horizontal or vertical overlap in the spatial crossover is analogous to the usable overlap in the crossover frequency range of the spectral divider.

Spatial acoustic crossovers are much more complex than their spectral counterparts, yet they are really just variations on the same theme. Place two identical speakers in any orientation and find the mid-point between them. That's the crossover for all frequencies. That's the center of the coupling zone for all frequencies. That was easy. The next part is a little harder. Now we have to find our way out of there to the isolation zone, the position where one speaker is 1O dB more dominant in level. This is easily found at any one frequency. If the speaker has the same directional pattern for all frequencies then the break point into the isolation zone would be as simple as single-line rolloff found in the spectral divider. Unfortunately, the chances of finding this in the field are less than that of spotting a unicorn. In the real world, we will find that the low frequencies

Perspectives: For me, a complex sound system is an assemblage of the edges of a number of simpler systems. Successful management of the transitions between two speakers in an array and the transitions between arrays help to create seamless coverage.

*Alexander Yuill-Thornton **II** (Thorny)*

overlap much more than the highs. Therefore, we will have different points in the room for the borders of the combing, combining and isolation zones over frequency. These borders depend on both the amplitude and phase relationships between the speakers so it is a complex task.

Crossover Overlap

The overlap considerations pertaining to the spatial divider are closely related to that of the spectral divider. If the elements combine at their —6 dB points the center will sum to 0 dB, provided the phase response is matched. This matches the crossover point level to that of the isolated areas on-axis to the individual elements; a unity class crossover. Just as in the spectral divider we will expect the amplitude and phase responses to move apart off-centre and aspire to isolate before cancellation occurs. Once isolation has occurred the extent of the addition tapers off such that the individual nominal response matches that of the combined responses in the crossover region.

A fundamental difference between spatial and spectral dividers is that all three crossover classes can coexist in a two element spatial summation (a two-way spatial crossover). The transitions occur over distance and over frequency. In the near field we will have a gapped crossover, which finds unity and eventually will overlap. Over frequency we find that the transition rate changes as the directionality changes. Low frequencies may immediately overlap at locations where the high frequencies are still gapped. Examples of these effects are found in Figs 2.39 and 2.40 respectively.

Speaker Order

The directionality of the speakers defines the size of the crossover range, with narrow speakers having a range-shortening effect. The goal of spatial acoustic crossover design is the same as the spectral acoustic crossover discussed previously: to transition out of the crossover region before the relative phase responses reach the subtractive side of the circle. Once again we have a race of amplitude isolation vs. phase cancellation but this time with a much

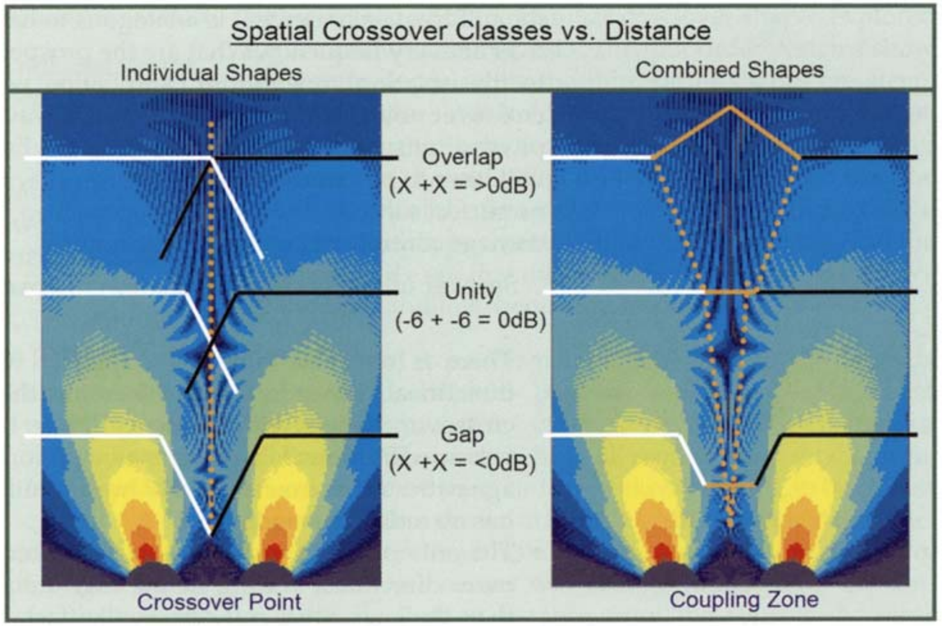


Figure 2.39 Spatial crossover classes change over distance. A pair of displaced sources will go through all three crossover classes. The transition sequence expands over distance if the speakers are moved apart, the element coverage angle is decreased or the splay angle increased

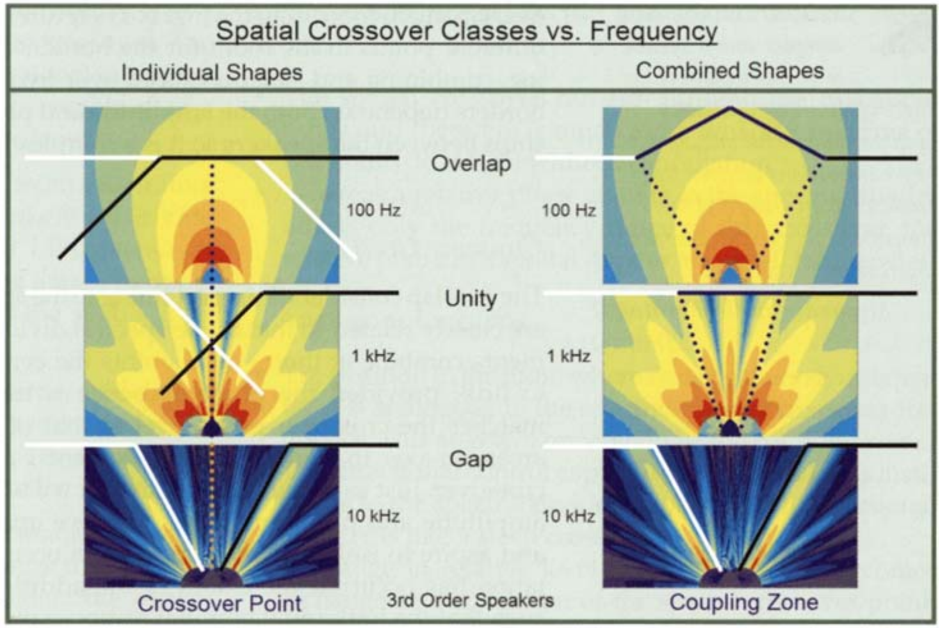


Figure 2.40 Spatial crossover class changes with frequency. As frequency falls the crossover will overlap due to lack of pattern control

lower margin for error. Since the spatial divider runs over the full frequency range it is analogous to having a crossover at every frequency. What are the prospects of getting to the isolation zone before cancellation with a 16 kHz crossover point? Not good. The reason is wavelength. The physical offset between the arrivals at the listening point will have to be less than half of our finger width! There is no perfect solution, but it can be approached. This is about damage control, and making compromises.

Speaker order concerns for spatial crossovers:

- There is no such thing as a speaker with uniform directionality over frequency. Therefore the slope of the crossover changes with frequency. (Note: coverage pattern specifications for speakers give a "nominal" coverage pattern. The exact nature of how nominal is defined has no industry standard.)
- The only prospect for success is a speaker that has the more directional control in the high-frequency range than the lows. Since our slope will affect our degree of overlap it will be a defining factor in the choice of relative placement of the speakers.
- Since the wavelengths at the high end are the smallest, they will have the least margin for phase offset. Therefore, controlling the high-frequency overlap will be crucial. Mid-range and low-frequency overlap can be accommodated to a larger extent since the wavelengths are long enough to minimize cancellation.
- The high frequencies will need to have the shortest possible differences in path length if they are overlapping. If the displacement is large compared to the wavelength the systems must be isolated.

While there are no speakers with uniform directional control over frequency, this does not mean that there is no difference between highly directional speakers and wide coverage units. They all go wide in the low end, but differ in the mids and highs. These can be viewed as having different ratios of LF, MF and HF directionality.

Spatial crossover example: an example first-order speaker (see Fig. 2.41) has a 270 degree pattern at 250 Hz

Perspectives: An anecdote about using a line array system of a certain manufacturer. The line array was flown at the height of 6 meters on both sides of the stage in a concert hall. I was introduced to a mixing engineer who exclusively worked for the concert hall and was told that the tuning of sound system was done by him. The mixing engineer was very confident of the outcome. I checked the sound by playing a CD as the music source. There was much reverberation as well as some peaks that 1 could not ignore. So I suggested we take measurements and tune the system using SIM machine. I brought in SIM3 and started from readjusting the crossover among individual speaker units within each element of speaker system.

At first, I set the phase responses and made the amplitude value even. Then I gave angles to the rigged element aiming at the appropriate direction to adjust energy balance as a whole. Also, I performed the alignment of subsystem, making full use of parametric equalizer and delay machine.

As the result of my attempt, the sound image that had been scattered all over the place concentrated toward the center of a stage. Timbres and orientation of each musical instrument became distinctive. Even the reverberation decreased. All those who were present on that occasion praised, "This is MASU MAGIC!"

I know that I owe this magic to SIM that took measurements of phases and carried out optimization.

Akira Masu

and settles down to a 90 degree pattern over the range of 1 kHz to 16 kHz. This speaker could be characterized as having 3:1:1 ratio of directionality over frequency. A unity class crossover will be created if speakers are splayed apart by 90 degrees. At frequencies below 1 kHz the overlap would increase and there would be a wider combination zone and smaller isolation zones.

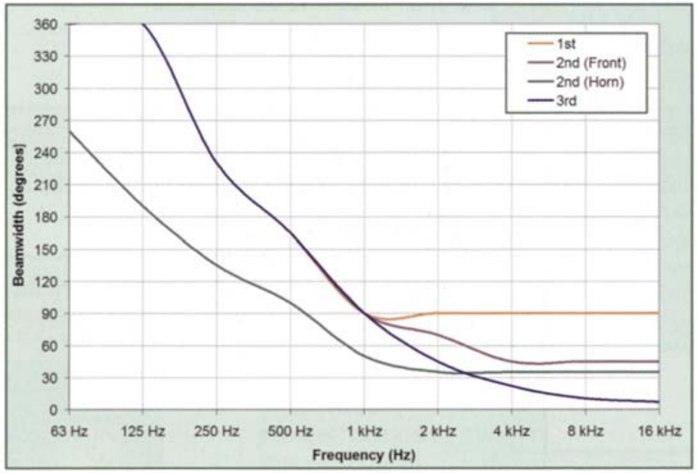


Figure 2.41 Speaker order as a function of beamwidth. High order speakers have the widest ratio between low-frequency and high-frequency coverage angle

An example third-order speaker has a similar low- and mid-frequency response but narrows down to 7 degrees at 16 kHz. This speaker could be characterized as having a 40:12:1 ratio of directionality over frequency. A unity class crossover now requires a 7 degree splay. The amount of overlap at 1 kHz exceeds 90 per cent of the patterns!

In both cases the splay angle of the speakers must be set to no more than the narrowest angle: that of the high end to prevent a gapped crossover. In the overlapped ranges the speakers will couple, comb or combine, depending upon the relative level and phase but they will not isolate. The first-order speaker would have much less overlap over frequency and therefore could withstand a higher amount of physical displacement. The third-order speaker has so much overlap that physical displacement must be absolutely minimized.

There is no hard line between the speaker orders like there are for filters. The change in coverage is gradual. Nonetheless the distinction is useful, as it will allow us to classify speakers without needing to separate them into hundreds of levels. The classification is derived from viewing the directional control of a speaker as a form of spatial filtering. The nominal response is the on-axis response. The classifying factor is the degree of filtering over a quadrant of coverage (a 90 degree slice) over frequency. A "zero"-order speaker would be one that has the same level at the center and edges of the quadrant; a flat line (360 degrees of coverage). A first-order system has dropped 6 dB at the ±45 degree edges, while second- and third-order speakers fall 12 and 18 dB respectively (see Fig. 2.42). The rolloff rates correspond roughly to 90, 45 and 22 degrees respectively. These are approximations but serve to prevent us from falsely visualizing that a 7 degree speaker is a thin slice of pizza at all frequencies.

The shape created by the speaker coverage pattern can be visualized with the familiar crossover shapes explored earlier as shown in Fig. 2.43. The first-order systems show the gentle slopes while the third-order shapes are steep. The shape does not hold over the full frequency range. The relationship between the speaker-order shapes over frequency is shown in Fig. 2.44. Here we can see that the gentle slopes of the first-order system closely approximate the coverage shape over the full range. As we move to the third order we see that the steep shape is unrepresentative of the mid- and low-frequency responses.

The spatial crossover point at a given distance depends upon the individual patterns and the displacement between them. A unity class crossover can be created for all three speaker orders, but will require different displacements, with the third-order speakers needing to be closest together. This is shown in Fig. 2.45. Alternatively if the spacing were to remain constant the distance at which the unity class crossover occurs will change (third being the longest). Note that as distance from the sources increases

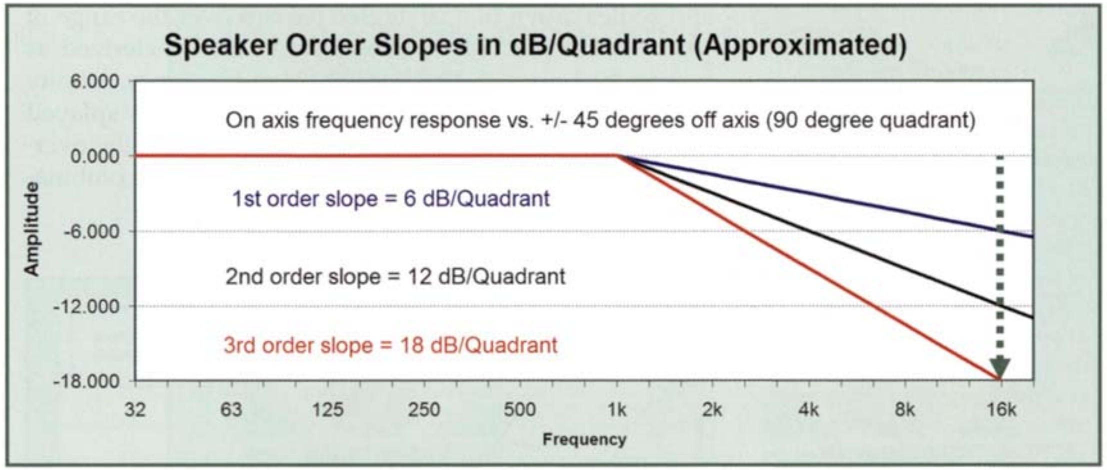


Figure 2.42 Speaker order slopes over frequency normalized to the on-axis response. The rate of high-frequency rolloff over a quadrant (±45 degrees from on axis) is shown

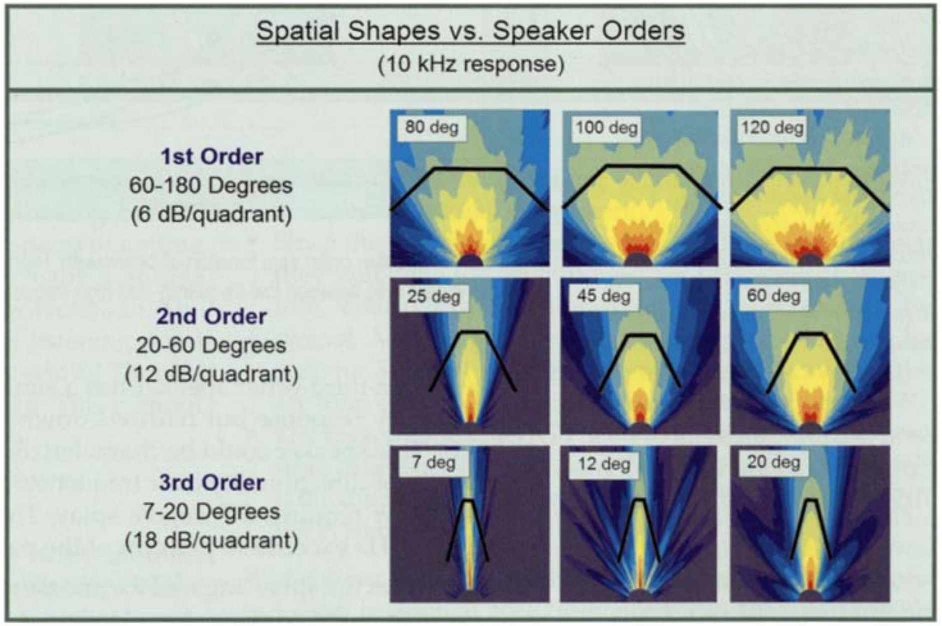


Figure 2.43 Spatial shapes change as a function of coverage angle. The speaker order categorized the shapes into three basic shapes

the proportion of the combing and combination zones becomes a larger proportion of the coverage area.

over shapes will be modified accordingly (Fig. 2.46). The

speakers can be splayed at any angle up to their unity class angle. The unity crossover distance will come from the order, Speakers

angle and displacement must compress proportionately.

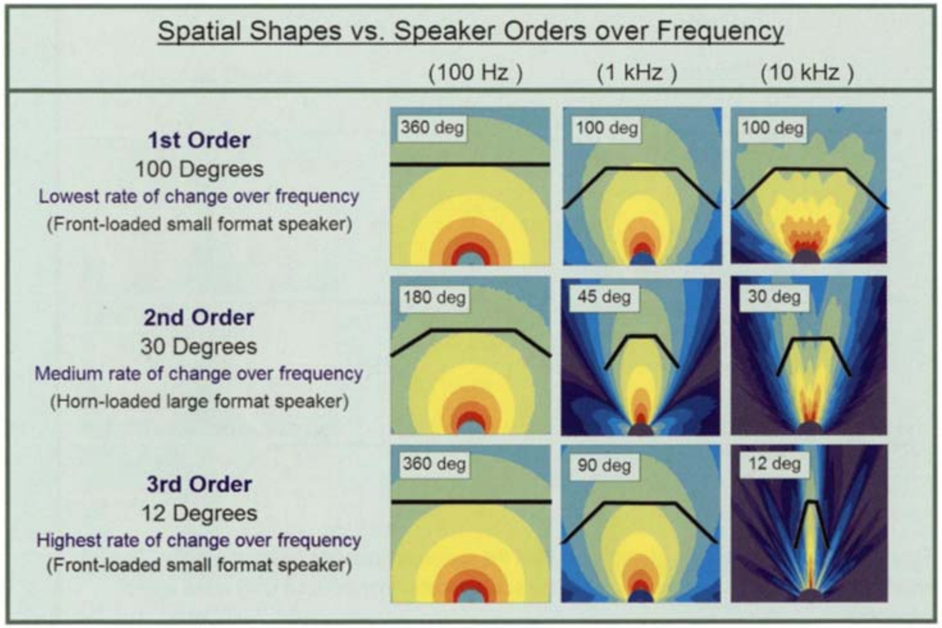


Figure 2.44 Speakers are unable to hold their coverage angle over their full frequency range, which makes the shape applicable to a limited range. As speaker order increases, the variation from its high frequency shapes increases

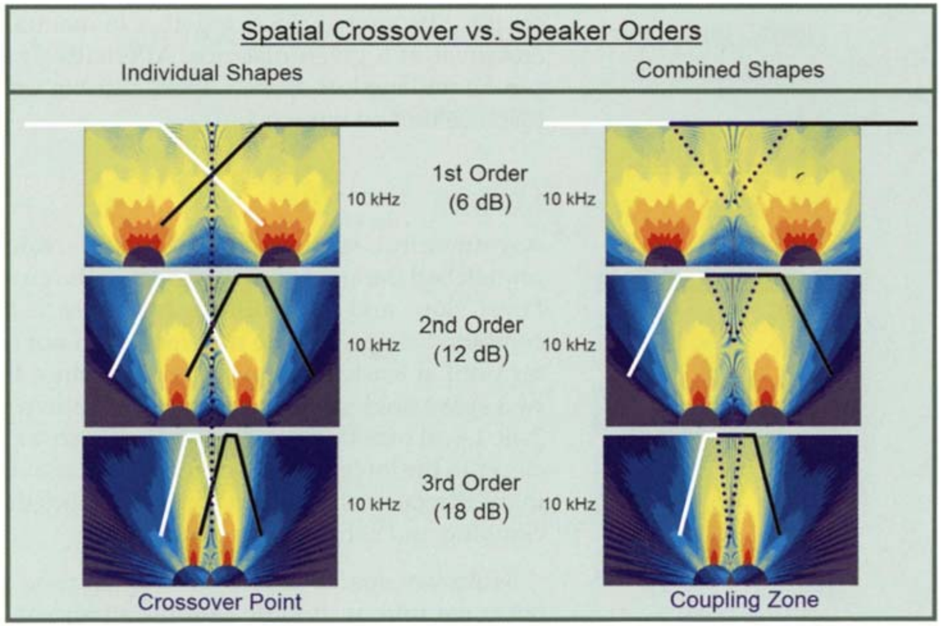


Figure 2.45 The speaker order affects the spacing between the units in order to achiev point

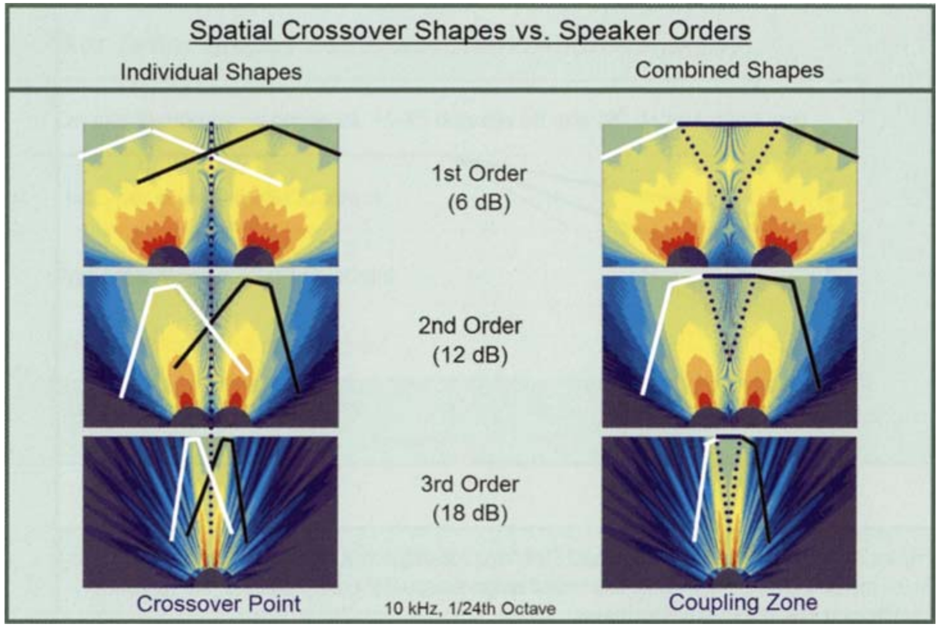


Figure 2.46 The introduction of angle between the speakers creates a combined shape. The spacing and angle must be proportional to the speaker order to maintain a unity class spatial crossover

For a given speaker order, as the angle opens up the speakers must be moved closer together to maintain unity class crossover at a given distance. Alternatively the speakers can be widened or spaced apart and the unity crossover distance moved outward.

Crossover Asymmetry

Asymmetrical spatial crossovers are created whenever unmatched parameters are present in the crossover range. Level, slope and angle differences create an asymmetrical combined shape. (Angle differences do not become a factor until at least three units are used, since then there are two splay angles.) These factors are shown in Figs 2.47-2.49. Level offset causes the spatial crossover point to shift closer to the lower level speaker. The crossover point shifts in the direction of the quiet speaker as does the inter-active combing and combination zones.

Multiway spatial crossovers can be used to subdivide coverage into an infinite number of slices. As with two units, the spacing and speaker order will determine the depth of the unity class crossover (Fig. 2.50). As more devices are used, we will need to employ higher-order speakers and closer spacing for the job. This brings us to a critical point: the fact that the low- and mid-frequencies will have large amounts of overlap in such dense multiway systems. The unity class spatial crossover cannot be maintained over distance and over frequency except for a very small number of cases. We are building up to how to do this, but the answers will not come for several chapters yet.

The nature of the multiway interaction is shown in Figs 2.50 and 2.51. The familiar crossover class zones are shown as they progress over distance. At wide spacings the transitions are spread apart. As the spacing decreases the zones compress. When a third element is added two new crossovers are added: the second and third element, and that of the first and third element. The additional elements extend the unity crossover zones outward (horizontally) but have the inverse effect in the overlap zone. There are actually three overlap zones, which stack together in a pyramid-type

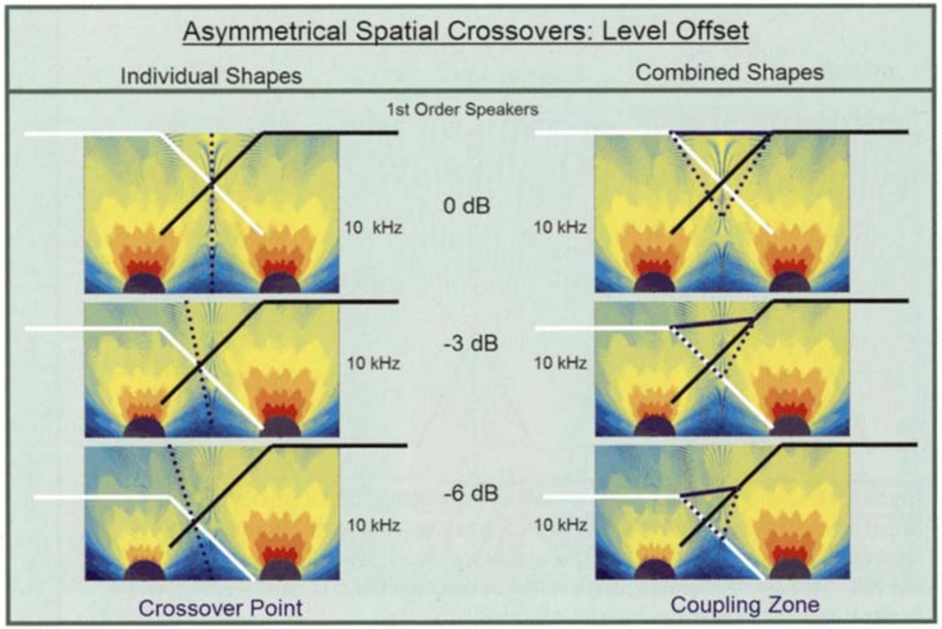


Figure 2.47 Level offset creates an asymmetrical crossover. The spatial crossover position shifts in the direction of the lower-level speaker

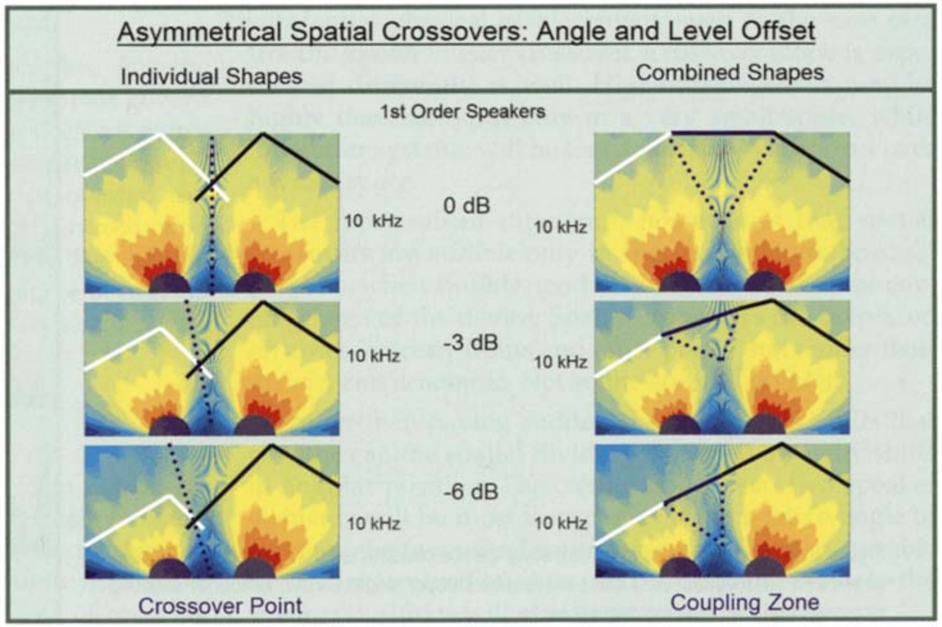


Figure 2.48 Level offset affects the combined shape of an angled pair. As offset increases the shape bulges forward of the louder speaker

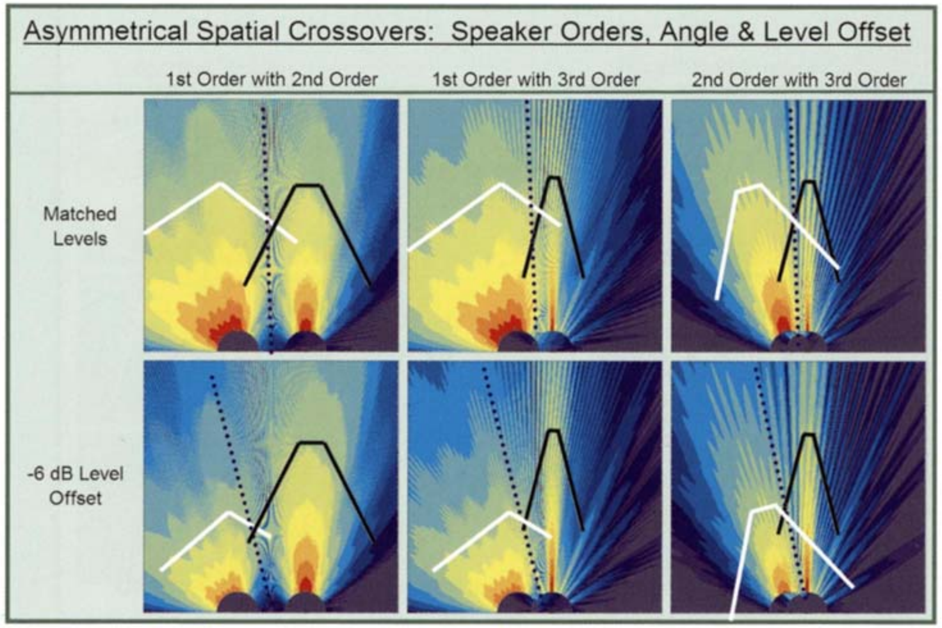


Figure 2,49 Mixing of speaker orders creates an asymmetrical spatial crossover. This is further affected by changes in level and angle. A great variety of spatial shapes can be constructed by the mixing of asymmetrical speaker orders, angles and levels. The three factors of speaker order, level and angle can all be used together to create the optimal shape for those applications where asymmetry is required

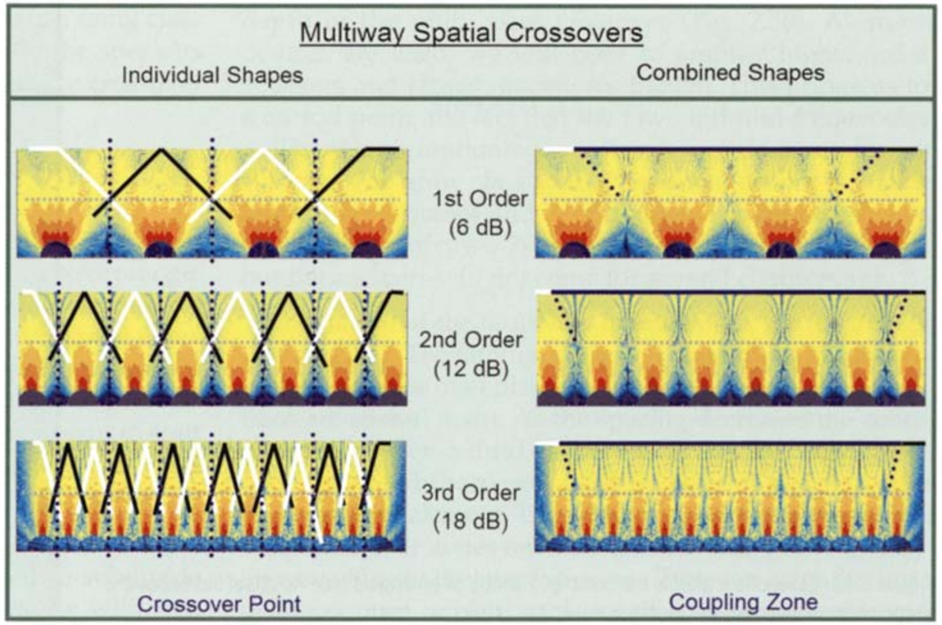


Figure 2.50 A unity class spatial crossover can be constructed from an infinite number of sections. To maintain unity class will require steeper slopes as the number of sections increases and their spacing decreases

110

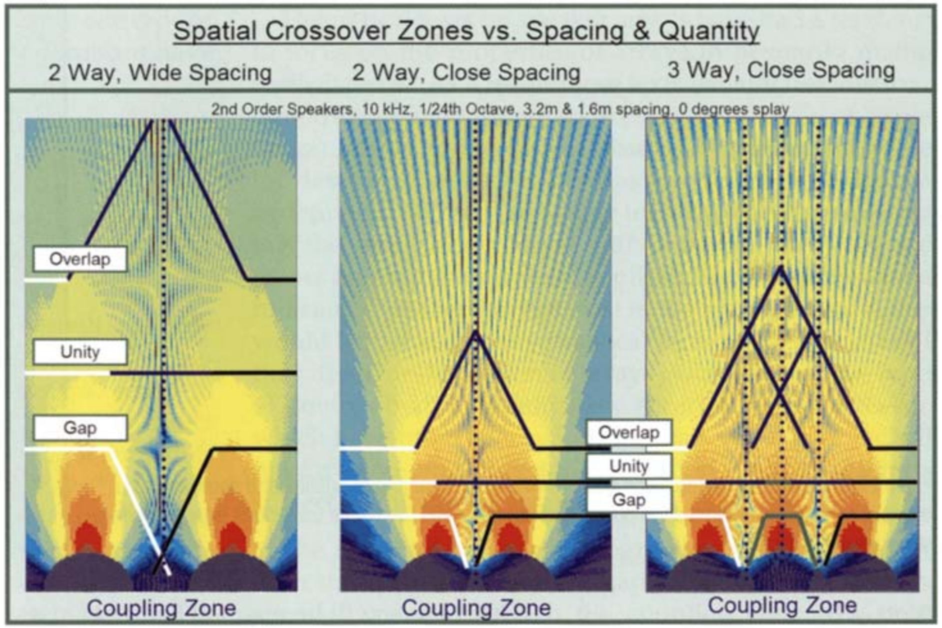


Figure 2.51 The spacing and quantity of elements will affect the crossover class transition points. As spacing decreases (center) the transitions compress to a shorter range. As a third speaker is added (right) the overlap area increases and the combined response bulges forward

shape. As frequency goes down the coverage pattern widens and all of the zones compress downward.

The nature of multiway arrays is complex. For the moment we are introducing the mechanisms that govern them. In such systems there is no more decisive factor than crossover overlap. A lack of overlap causes coverage pattern extension. A large quantity of overlap can cause pattern narrowing or scattering. An introduction to the multiway spatial crossover is found in Fig. 2.52. In all cases a pyramid shape emerges. The degree of overlap will decide the steepness of the pyramid steps.

Crossover Audibility

Crossover audibility takes on a different form in the spatial divider. The three factors outlined previously for spectral dividers all come into play again. Spatial dividers are much more difficult to hide since it is virtually impossible to transition through one without some area of combing in the high frequencies. They are also very likely to have substantial physical displacement, such as the case of a frontfill / center cluster crossover. Crossover slope is experienced differently as well. High-order systems may be highly detectable but only in a very small space, while low-order systems will be less detectable, yet spread over a wider space.

The most salient difference, however, is that spatial crossovers are audible only in specific locations. Spectral dividers, when audible, tend to be so over most of the coverage area of the device. Spatial crossovers can be placed on aisles, balcony fronts and other places that render their deficiencies academic. Not so the spectral divider.

Rather than having sudden spectral quality shifts that clue the ear, the spatial divider gives itself away by shifts in angular position. The crossover between two speaker elements will be most transparent if the relative angle of origin of the two sound sources is small. Our ears are able to pick up localization clues, and large angles between the speaker positions will give them away.

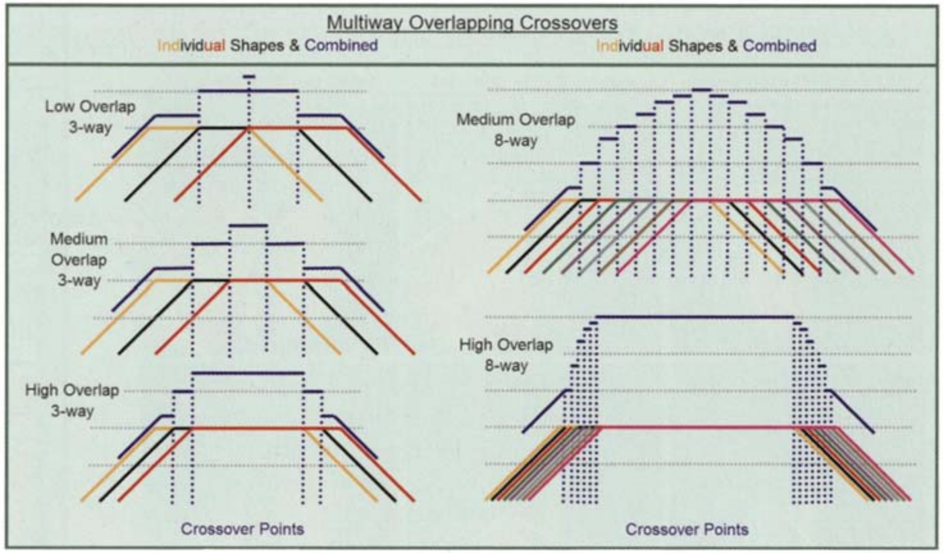


Figure 2.52 Multiway systems can be constructed with varying quantities and degrees of overlap. As overlap increases the power addition rises and the shape transitions from a gentle slope toward a beam

Perspectives: The most difficult aspect of system tuning is resisting the temptation to draw nice straight lines on the computer screen when using any kind of analysis tool, he it RTA or FFT based. Many times, I have seen and heard the results of blindly drawing pretty curves on the computer screen and pronouncing the system "tuned." Most times this sounds awful. Without an understanding of all the interactions that take place when loudspeakers are arrayed in an environment, electronic and acoustic, the display on the screen is little more than a glorified Etch-a-Sketch!

Dave Revel

Other clues will be mismatched responses in the crossover area. The most obvious is when one speaker has HF range extension well above the other. This is a factor that must be carefully managed when crossing over small format speakers such as frontfill and under balcony speakers with main systems.

Another clue comes in the form of the relative distance. Speakers that are close in proximity will have superior signal to noise over those at distant locations. This is due to the increased proportion of reflected energy in the response of the far speaker. When near speakers are crossed over with far speakers great care must be taken to manage their relative levels so that the near speaker is not allowed to stand out above the distant speaker.

Speaker Arrays

Introduction

We are now ready to apply our study of summation and the acoustic crossover to the practical construction of speaker arrays. Add two speakers and the sum will depend on their relative amplitude and phase. Add ten speakers and they will behave exactly as the summation of the summations. The spatial distribution of the arrays will depend upon a limited set of key parameters: the individual elements, their displacement, their relative angles, relative distances and relative levels. The points of confluence of these parameters are at the spatial crossovers. If we can successfully merge the systems at these meeting points, the rest of the coverage area will become predictable and manageable. All of these factors can be independently controlled in our design and, to a lesser extent, in our optimization process.

Our study here will be a methodical treatment of the role that each of these parameters plays in the creation of the different types of arrays. Once exposed, these independent mechanisms become manageable, and the seemingly infinite complexity of a multiple element speaker system becomes a composite of known and understood parts.

Speaker Array Types

There are a number of different types of speaker arrays. If we spend much time reading the trade news we could come up with the conclusion that there were hundreds of different kinds, based on the trademarked names. Alternatively we could conclude these days that there is only one type of array: the line array and that all other configurations have gone the way of the steam engine.

The actual number of array types is three: speakers in parallel, speakers angled outward, or speakers angled inward. The speakers may be placed together or apart but the same properties apply albeit on different scales. The angular orientation is our first level of subdivision, and the most important.

For the purposes of our discussion here, we will secondly divide the speaker array types into two versions each: coupled and uncoupled. The term "coupled arrays" refers to those where the speakers are physically placed in direct proximity. Standard touring arrays, or blocks of subwoofers are examples of this type. "Uncoupled arrays" are those where the elements are displaced with some unspecified distance between them. These serve as guidelines regarding the behavior of the arrays when closely or more distantly spaced. However, bear in mind that to 16 kHz the HF drivers that are a hand's length apart have the same relative spacing as 100 Hz does to two subwoofers on opposite sides of the stage.

All speaker arrays will begin the crossover class progression (the gap crossover) and possibly move on to the unity and overlap class behavior. Speakers spread apart will take longer to meet but provided the splay angle is less than the coverage angle of the elements, they will eventually meet. Does this mean that two speakers spaced 3 meters apart are the same as two spaced 300 cm apart? No and yes. To 100 Hz the behavior is quite different between them. But the behavior at 100 Hz for the wide spaced units is the same as the behavior at 1 kHz in the closely spaced units. This is because the source displacement of the units relative to the wavelength is preserved. The behavior of arrays will always have scaling factors associated with frequency.

Note: unless specified otherwise, this section assumes the arrays to comprise identical speakers, driven with identical signals. The arrays are described on a single plane but the behaviors would be the same in either the vertical or horizontal orientations. The individual speakers are termed the elements of the array.

Theoretical texts on speaker arrays have had a tendency to focus on the properties of arrays in primarily mathematical terms. As a result these texts usually use idealized omnidirectional speaker sources as their basis of explanation. These are, of course, absolutely the correct model for describing array behavior as a series of mathematical equations. The premise of this text is to focus on describing the optimized design with minimal math requirements and maximum field applicability. Omnidirectional full-range speakers do not exist in the real world, and they would be useless in any practical form of speaker array if they did. Our discussion of arrays is limited to those types of devices that we might have occasion to use: speakers which have more energy in front than in the rear. We will not focus on omnidirectional radiation until we reach the subject of subwoofer arrays.

The standard crossover progression for arrays begins with the isolation zone (the gap crossover area near the speaker), passes through the coupling zone (the unity class crossover area where the patterns first meet) and finally to the combing and combination zones (the overlap crossover area where one element has spilled into the coverage area of the other). Coupled arrays are in such close proximity that the gap area is very short, essentially only in the near field of the system. The unity class area follows. The range of the unity class crossover will be completely dependent upon the array type and on the choices of elements. Only one array type will hold on to unity crossover performance over a virtually unlimited range. Others will hold it for a limited range while some will blow though unity and into the overlap zone in the blink of an ear. This section charts out the behaviors of the different array types in terms of how each one constructs its crossovers. We will explore the variables that can be manipulated to shape arrays by controlling their spatial crossovers. These are the steering mechanisms for our arrays. Each of the arrays has predictable behaviors, and most, but not all, will have practical applications.

The six different array types are shown in Fig. 2.53. There are two drawing series that will dominate our treatment through most of the remainder of this chapter. The first

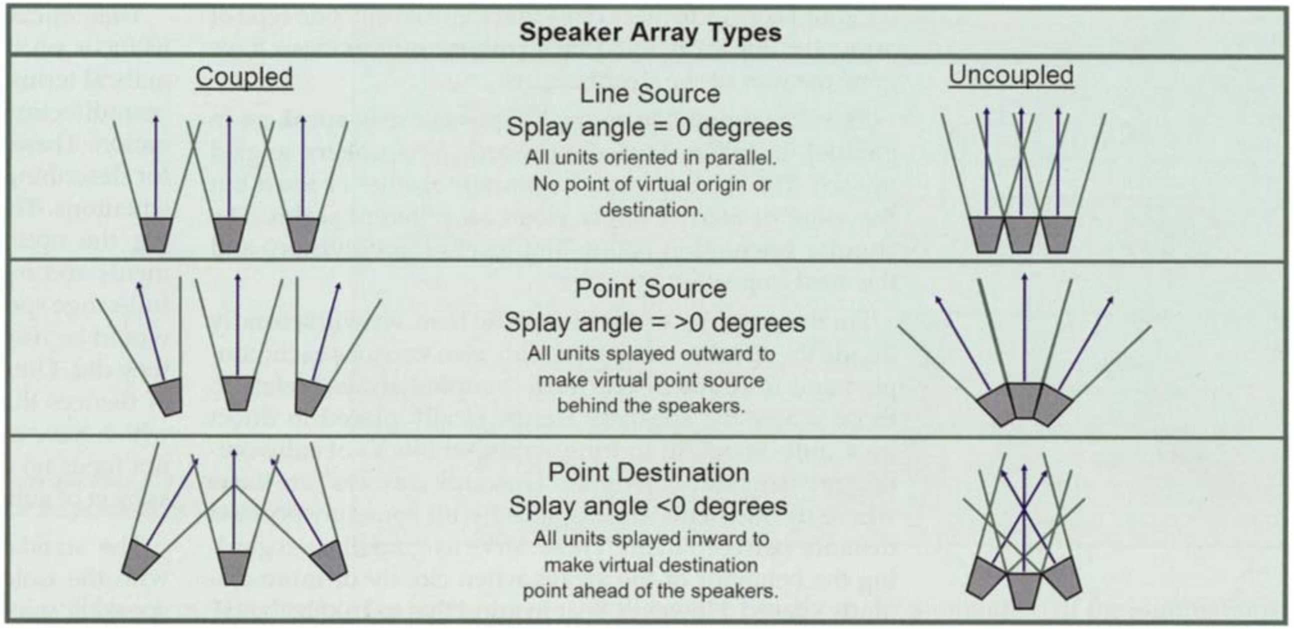


Figure 2.53 Speaker array types

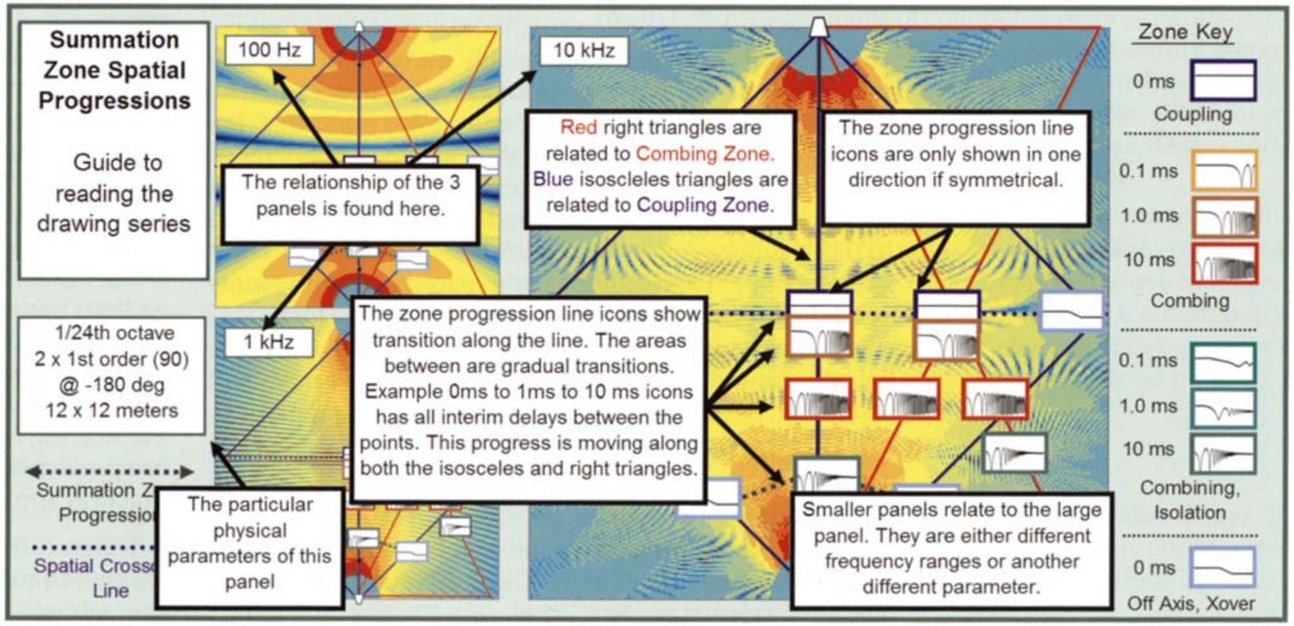


Figure 2.54 Guide for summation zone drawing s

details the spatial crossover class transitions for each of the array types. The second details the progression of the summation zones over the space. Thankfully, these panels will do most of the talking for the summation behavior of the array types. A guide to reading the summation zone panels is provided in Fig. 2.54.

Note that the coupled point destination array will not be covered in any detail here. The reasons follow.

For the coupled point destination array:

- This array is functionally equivalent to the coupled point source in terms of behavior over angle, frequency, etc.
- It is impractical to array beyond 90 degrees because the elements are aimed through other array elements.
- The geometry of most manufactured systems is unfavorable due to the horn driver placement at the cabinet rear. A point destination array will have a higher displacement than its comparably angled point source counterpart.
- With the exception of the rare and special occasions where physical logistics require the use of this array, they are not preferred over their point source counterparts.

Coupled Arrays

We will begin with coupled arrays. The speakers are in close enough proximity that the gap coverage area is restricted to the very near field, an area where we are not expecting to find sober audience members. As a result, the progression will move quickly into the highly interactive unity and overlap zones. Coupled arrays are principally concerned with the management of overlap, and as a result they are the arrays that will bring power addition. The challenge will be to obtain the maximum power addition without excessively costly combing interaction.

Coupled Line Source

There is no simpler array to describe than the line source. There are no angles to be concerned with, only the number of elements and their spacing. The behavior of coupled line sources is also the simplest to describe, as well. Coupled line source arrays contain an unlimited number of elements at the most limited numbers of angles: one. All units are facing in the same direction with identical orientation. The crossover is described as a line originating at the geometric center between the elements and proceeding ahead and behind to infinity.

Crossover Class Progression

The spatial crossover class progression for the coupled line source is shown in Fig. 2.55. Here we can see the principal mechanisms that will drive this array's behavior. There is a perfect consistency in the effects of the three major parameters. Speaker order, displacement and frequency all have the same effect: as they rise the transition points move outward from the speakers. Another consistent feature of the coupled line source is that overlap class behavior comprises virtually the entirety of the system's response. This is both its most attractive and most ominous feature. The overlap gives it the maximum power addition, but at the cost of minimum uniformity.

Summation Zones

The first in this series is Fig. 2.56, where we see the results of a pair of first-order speakers arrayed as a coupled line source. The response is shown in the 100 Hz, 1 kHz and 10 kHz format which is used throughout this chapter. This series brings together the summation icons presented earlier and gives them context with the spatial crossover locations. As for this particular scenario the results are clear. The only position in the entire room that will enjoy a ripple-free frequency response is the exact center line; the spatial crossover. All other positions will vary. This is the result of overlap behavior at all locations. The summation zone progressions can be seen moving symmetrically to the left and right. The overlap is the dominant feature in the 10 kHz range causing deep combing. The combing is not reduced by distance, and finally relents when we have gone off-axis. The proximity of the sources prevents the combing from reaching down to low frequencies.

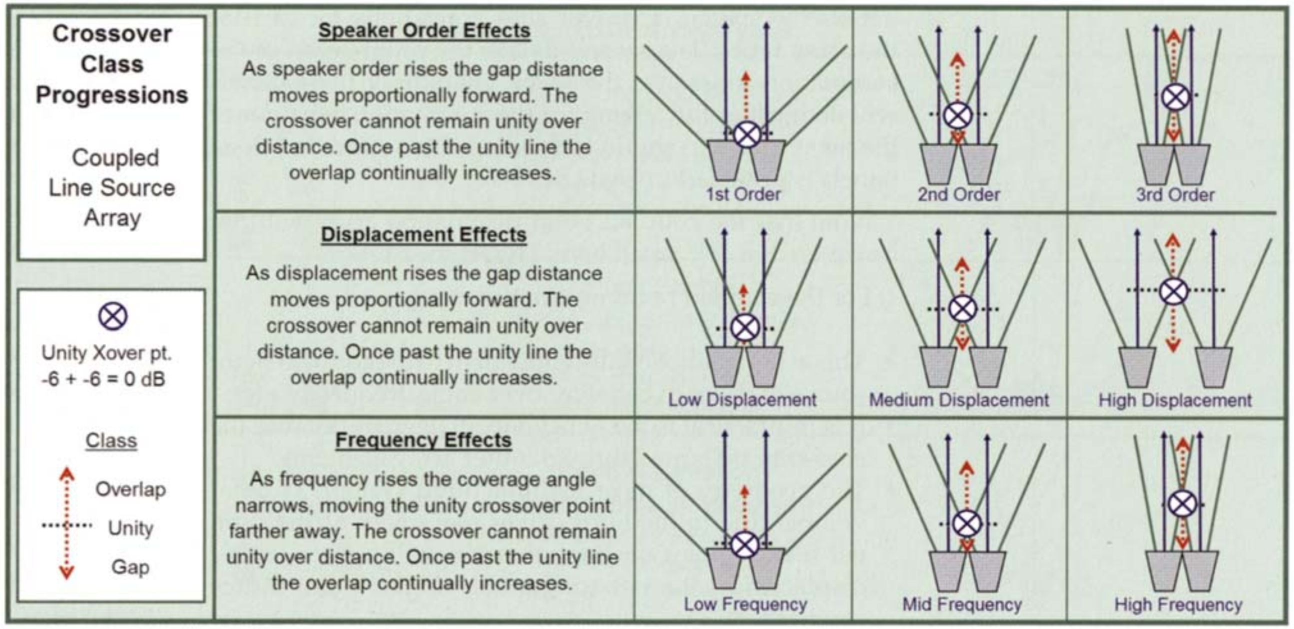


Figure 2.55 Crossover class progression fact

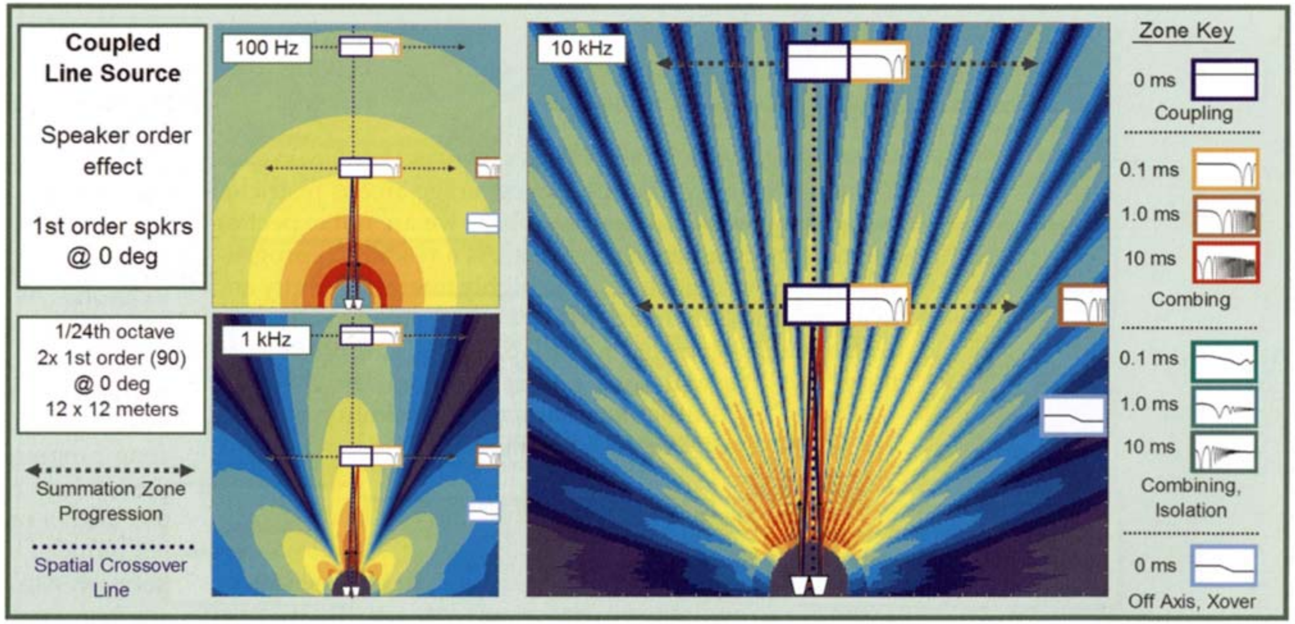


Figure 2.56 Summation zone progression factors for the couple line source array, first-order speakers

The area enclosed by the blue isosceles triangle is the coupling zone. The close proximity of the elements confines this and the on-axis right triangle area to a minute percentage of the coverage. The obtuse angle outside of these confines does not yield isolation zone performance because there is no angular separation. The 90 degree patterns overlap out of time in the off-centre areas. The source displacement can be deduced from the response. The 10kHz response reveals 10 nulls (10 wavelengths displaced, or 34 cm).

The next panel (Fig. 2.57) shows the effects of a change of speaker order. In this case the horn-loaded second-order speaker has increased directional control at all frequencies. The area of destructive interference is not reduced (as a percentage of coverage). The high-frequency combing covers the entire coverage area. Note that the 1 kHz response is narrower than the 10 kHz response. The HF responses are essentially passing through each other whereas the mid-range wavelengths are long enough to concentrate into a summed beam.

A pair of third-order elements is seen in Fig. 2.58. In this case the individual high-frequency coverage is so narrow that the gap coverage zone can be seen in the near field of the speakers. This quickly gives way to overlap coverage and the summation concentrates the high-frequency response into a single beam. The high-frequency response has no combing zone in the far field since sufficient isolation has been achieved to move us into the combination zone. On the other hand we have virtually no coverage, since the beam is concentrated to cover such a minute angle. Note that the mid- and low-frequency responses are only nominally changed for all three speaker orders in this array.

An additional third-order element is added to the array in Fig. 2.59. This has the effect of extending the gap crossover range into two sections. The first section awaits the convergence of adjacent elements and the second gap section ends when all three elements have converged. This series is plainly visible in the detail panel. The addition of the third element adds power coupling to the array and narrows the response at all frequencies further. It does nothing

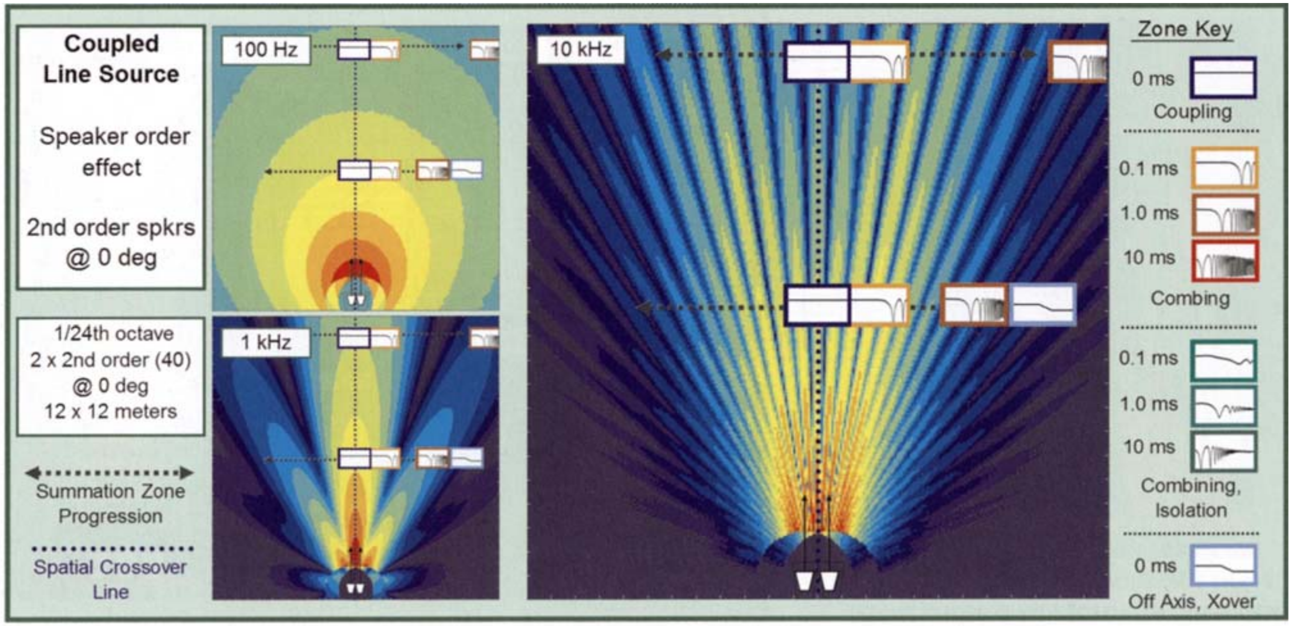


Figure 2.57 Summation zone progression factors for the couple line source array, second-order speakers

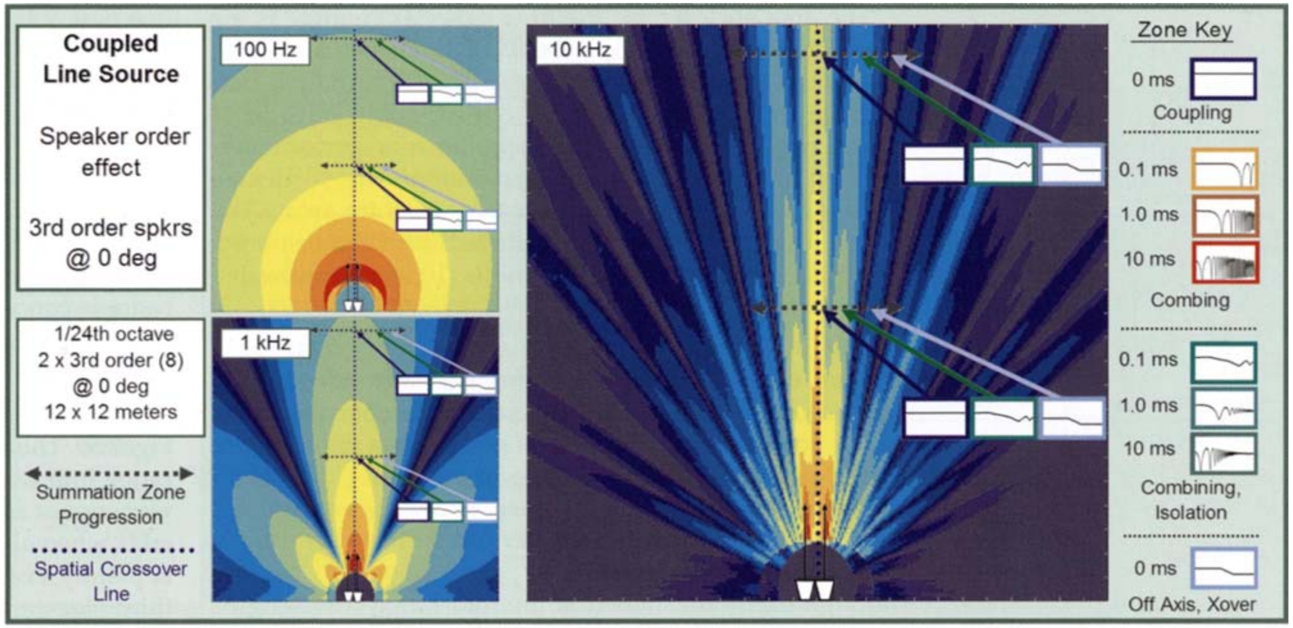


Figure 2,58 Summation zone progression factors for the couple line source array, third-order speakers

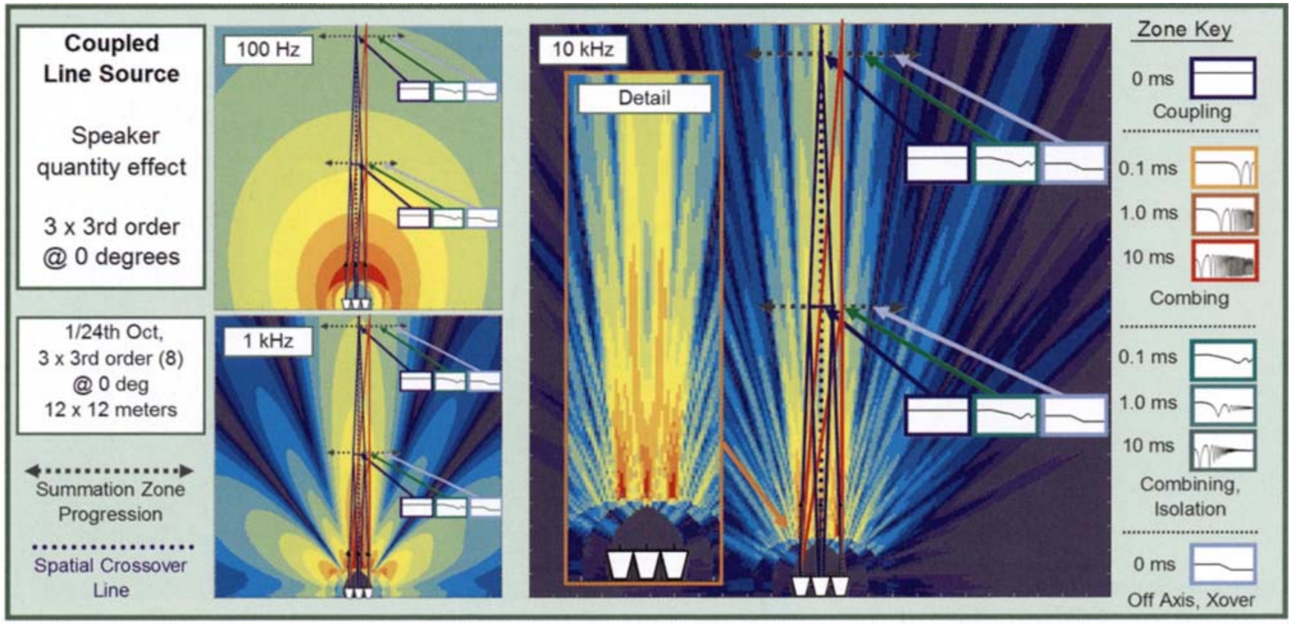


Figure 2.59 Summation zone progression factors for the coupled line source array, change of quantity

to address the fact that there is high-frequency coverage for a tiny area while much larger areas receive mids and lows.

The complexity of the triangulation geometry increases with the addition of the third element. There are now multiple triangles stacked together which increases the overlap over distance, causing the coverage angle to narrow.

We digress for a moment in order to address a fundamental property of the line source (both the coupled and uncoupled versions). The parallel nature of line sources creates a pyramid-shaped series of summations. The number of steps in the pyramid is equal to one less than the number of array elements. The pyramid effect comes from the summation of previously summed systems. A three-element system contains three crossovers, two between adjacent system and one that sums all three. This can be viewed as a pair of overlapping two-way crossovers and a single overlapping three-way crossover. The forces that create the pyramid are shown in Fig. 2.60. The phase contours of the three elements converge initially into two zones of addition (and one of cancellation). As we move further away the three phase responses converge to form a single beam, the pyramid peak.

The parallel pyramid does not stop with three elements but rather continues moving upward for every additional element. An eight-way line source pyramid is shown in Fig. 2.61. The foundation begins with the isolated elements (the gapped crossover) and then continues with 7 two-way overlapping crossovers until it eventually converges into a single eight-way overlapped summation. The distance at which the pyramid steps transition will depend upon how far we must travel to reach the overlap point. As elements become more directional or displacement increases the distance extends. Since directionality is variable over frequency, the slope of the pyramid will vary over frequency. As a result the beamwidth of the array will be highly variable over frequency. The parallel pyramid operates over the full range of frequencies. If the highs are

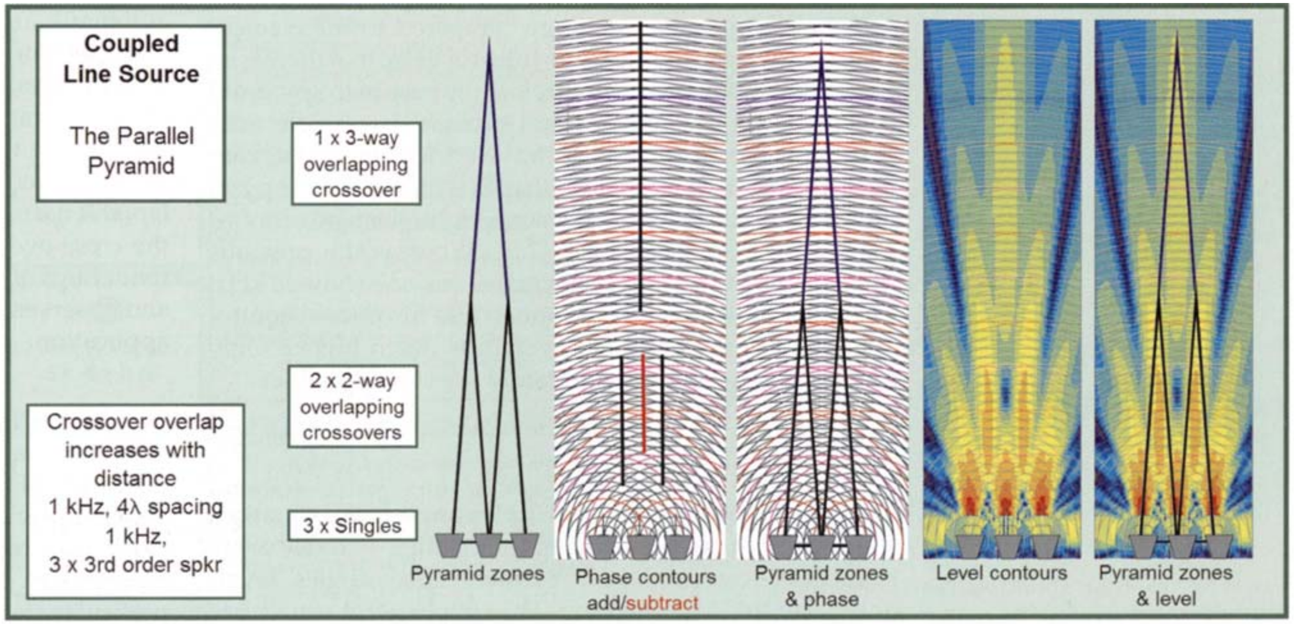


Figure 2.60 Parallel pyramid for three elements

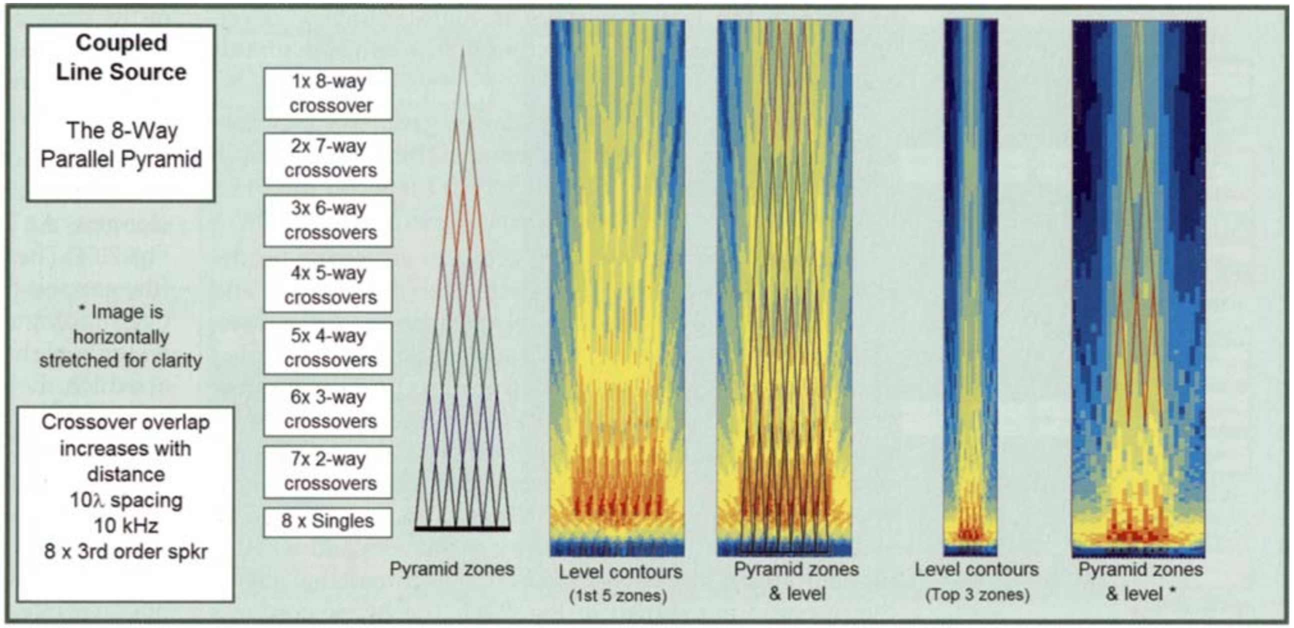


Figure 2.61 Parallel pyramid for eight elements

narrow compared to the lows for a single element, then the relationship will be preserved with additional elements.

If the wavelengths are large compared to the element spacing, the pyramid base is so full of overlap that there is no gap class crossover zone. Therefore the base of the pyramid collapses and the steps do not become discernible until sufficient distance has been traveled for the phase contours to take on the coherent shapes which drive the pyramid. Figure 2.62 shows the same eight-element array as previously discussed in Fig. 2.61. Whereas the previous figure showed the 10 kHz response, this one shows 1 kHz. The third-order speaker has much less directional control at 1 kHz and therefore the overlap is much higher. Only the top three levels of the pyramid are clearly visible.

Coupled Point Source

The coupled point source adds splay angle to the equation. This opens many possibilities for creating variable array shapes. We can mix speaker orders, splay angles, levels and delay to do the shaping. The couple point source has one feature that no other array type can duplicate: it can maintain a unity class crossover over distance. This is not automatic, and it is variable over frequency, but no others can duplicate this for even a single frequency This is done by creating an array whose splay angle is equal to the element coverage angle (the unity splay angle). If the coverage angle remains constant over frequency, so will the unity class crossover. If it widens we will move into overlap, if it narrows we fall into the gap. The factors affecting the cross-over progression are described in the unsealed renderings of Fig. 2.63. The next section takes these factors and observes the summation zone progression in a scaled application.

Crossover Class Progression

The crossover class progression for the coupled point source is shown in Fig. 2.63. Here we can see the principal mechanisms that will drive this array's behavior. The unity class crossover can be maintained over distance for each of the three orders. The unity splay angle must be selected for the particular element. This angle becomes smaller as speaker

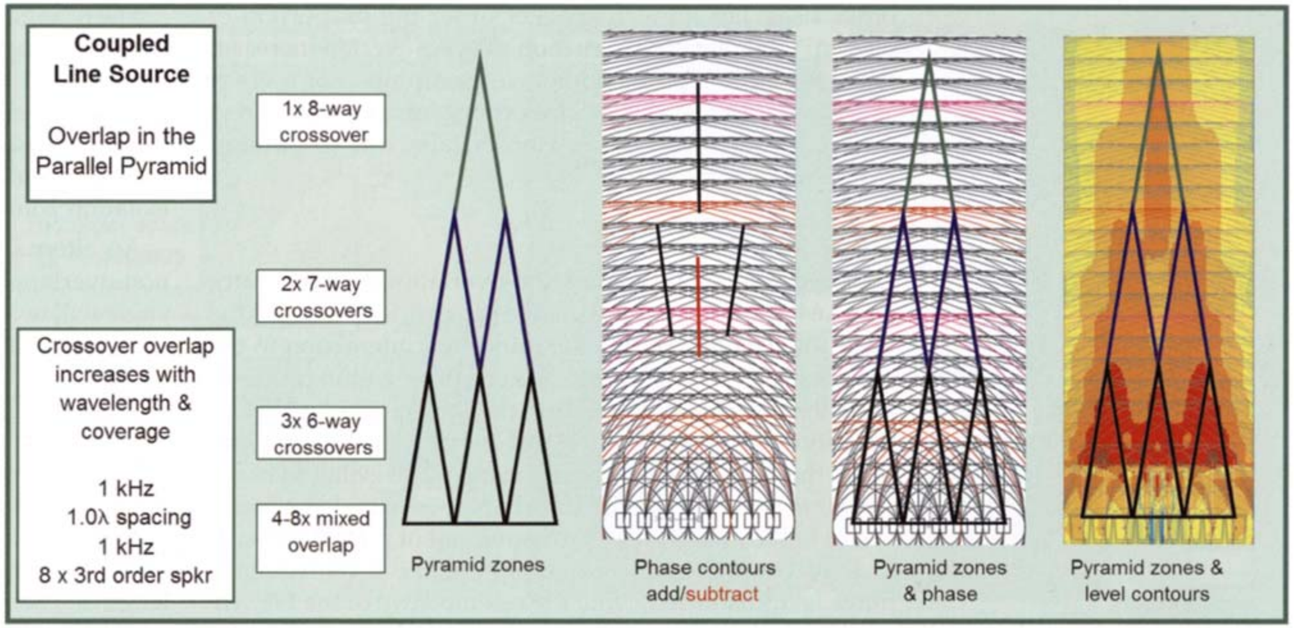


Figure 2.62 Parallel pyramid for eight elements with overlap effects

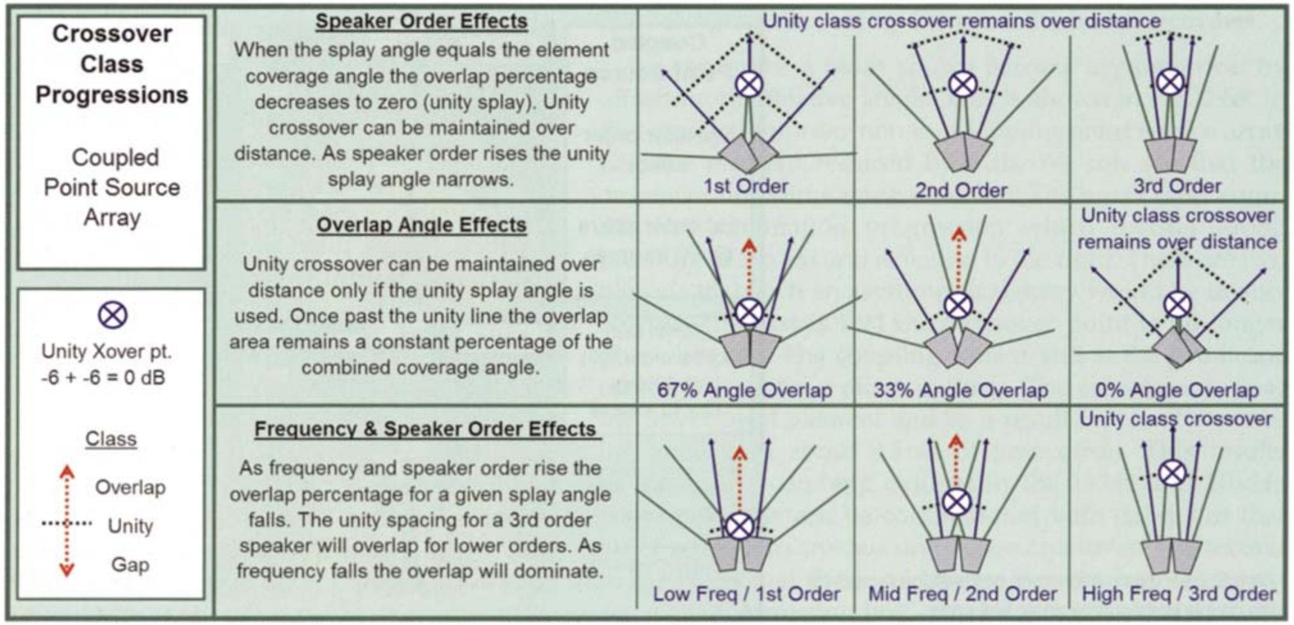


Figure 2.63 Crossover class progression factors for the couple point source array

order rises. For a given speaker order the proportion of overlap will affect the transition rate. As overlap increases the crossover class transitions will compress. For a given splay angle the crossover class transitions will compress as frequency and as speaker order falls, due to pattern-widening in both cases.

Summation Zones

Let's begin with a first-order non-overlapping point source (Fig 2.64). In this case a 90 degree speaker is splayed at the unity angle. The result is the minimum interaction in the crossover area. The summation zone progression proceeds from the crossover line and reaches the combination and isolation zones as it moves off-centre. The progression holds the same qualities over distance. The isolation of the elements is clearly seen in the HF response and to a lesser extent in the mid-range. The displacement is small enough to limit the mid-range combing and the low-frequency range is undisturbed. Note the resemblance of the HF, MF and LF shapes. The uniformity throughout the coverage area for the full range of frequencies is very high.

The triangulation aspects work out quite differently than with the coupled line source, even though the array axis and element locations are identical. The splay angle separation moves the on-axis energy of both elements away from the isosceles and right triangles (coupling and combing) and into the obtuse triangle area (combination and isolation zone).

An alternative form using a second-order speaker in a non-overlapping mode is shown in Fig. 2.65. The splay angle will need to be revised to accommodate the narrower element angle. In this case the element and splay angle are 40 degrees. The summation zone progression in the center is similar to the first-order sequence, but is more compressed. The isolation zone arrives more quickly due to the decreased element coverage angle. The displacement approximates 1 wavelength in the mid-range and results in a single null pattern per quadrant. The displacement at the low-frequency range is approximately 0.1 wavelengths. Therefore no evidence of dual radiation is seen. Also noteworthy is the fact that here the mid-range response is wider than the HF response.

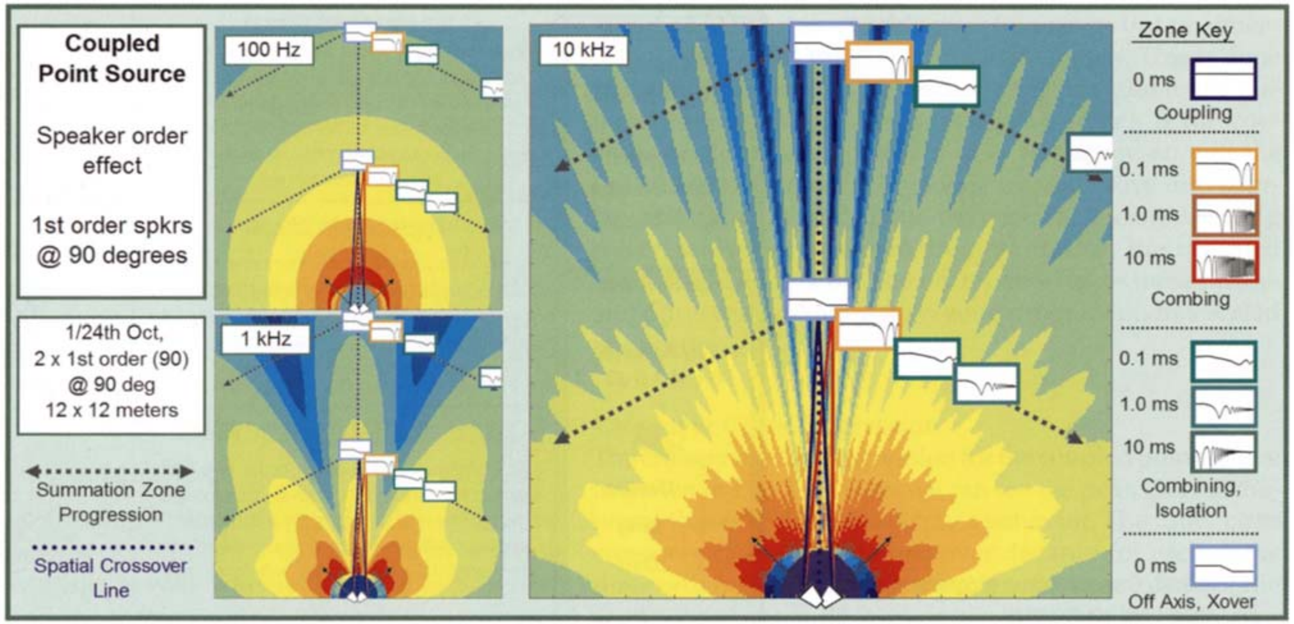


Figure 2.64 Summation zone progression factors for the couple point source array, first order

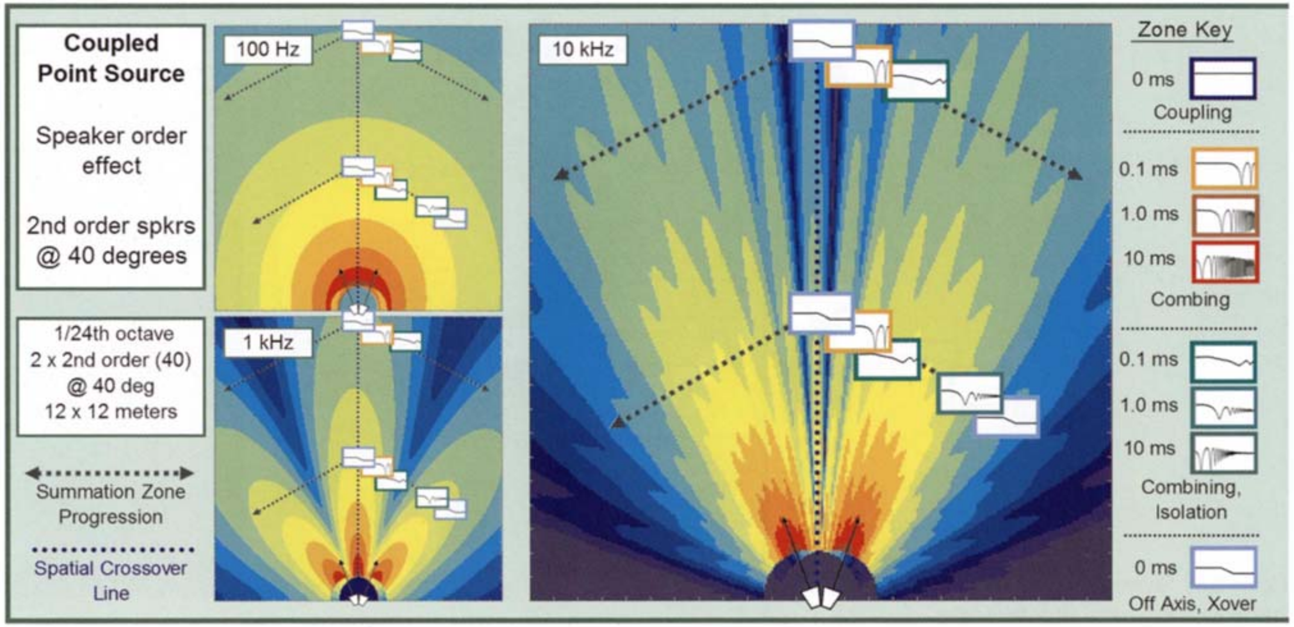


Figure 2.65 Summation zone progression factors for

This approach can also be used for a third-order speaker as shown in Fig. 2.66. The splay angle in this case is only 8 degrees. The gap crossover area is clearly visible here in the near field HF response. By the top of the panel the crossover has reached the unity point which will hold out for an extended range. Note that the third-order system has the highest ratio of low-frequency to high-frequency coverage angle. The unity class point source array has begun to reverse the trend in that the high frequency widened while the mid and lows narrowed. As we will see in Chapter 6, this will be the guiding principal in third-order speaker applications.

We will now look at the effects of various proportions of coverage overlap (Fig. 2.67). In this case we view the HF response with three different percentages of overlap. The 0 per cent version was already viewed in Fig. 2.64 and is here for reference. The other panels show the response with 50 per cent overlap (45 degree splay with a 90 degree element) and 75 per cent overlap (22 degrees of splay). Two trends are apparent. The first is an increased proportion of the HF combing zone over the pattern. The 75 per cent overlap more closely resembles the coupled line source than the non-overlapping point source. The second trend is pattern narrowing. As the overlap increases the pattern narrows.

We can make a point source become asymmetrical by offsetting the relative levels. This is shown in Fig. 2.68. In this case a first-order non-overlapping point source array has one element reduced by 6 dB. We can see that the crossover line shifts toward the left. The result is an asymmetrical summation progression which reveals strong combing to the left and isolation to the right. There are two reasons that such an asymmetrical array would be unsuccessful. The first is that the crossover point is no longer phase-aligned. The coupling zone is still at the geometric center (where time offset is 0 ms). The crossover is near the lower level element and as a result it is in the combing zone with about 0.5 ms of time offset. This results in the strong combing evident in the 1 kHz and 10 kHz responses. This can be compensated with delay but that is not enough to create a unity class crossover. The second challenge is that the level change has not been compensated by a corresponding angular change. At the center

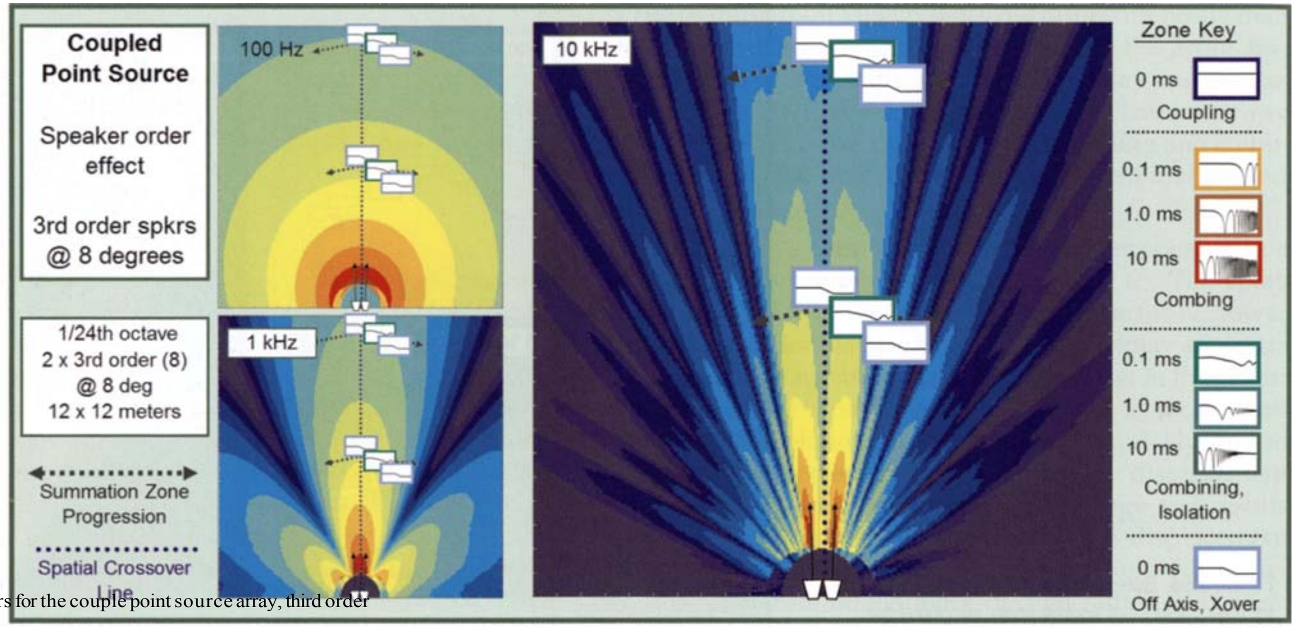


Figure 2.66 Summation zone progression factors for the couple point source array, third order

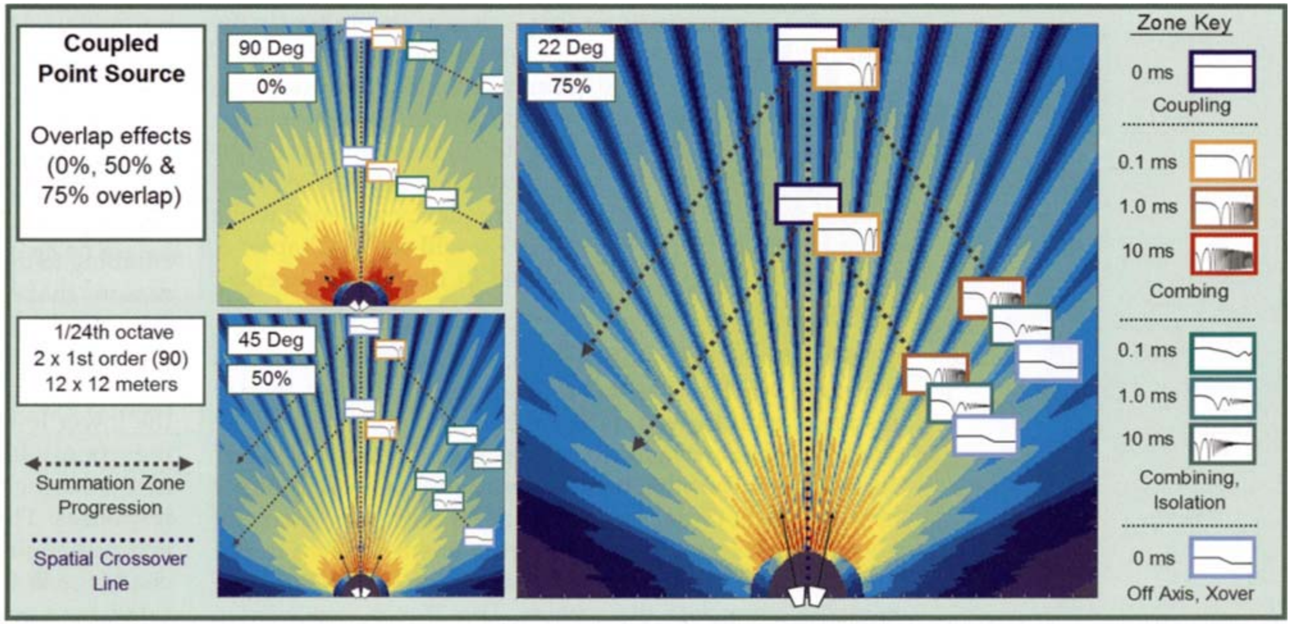


Figure 2.67 Summation zone progression factors for the coupled point source array, overlap effects

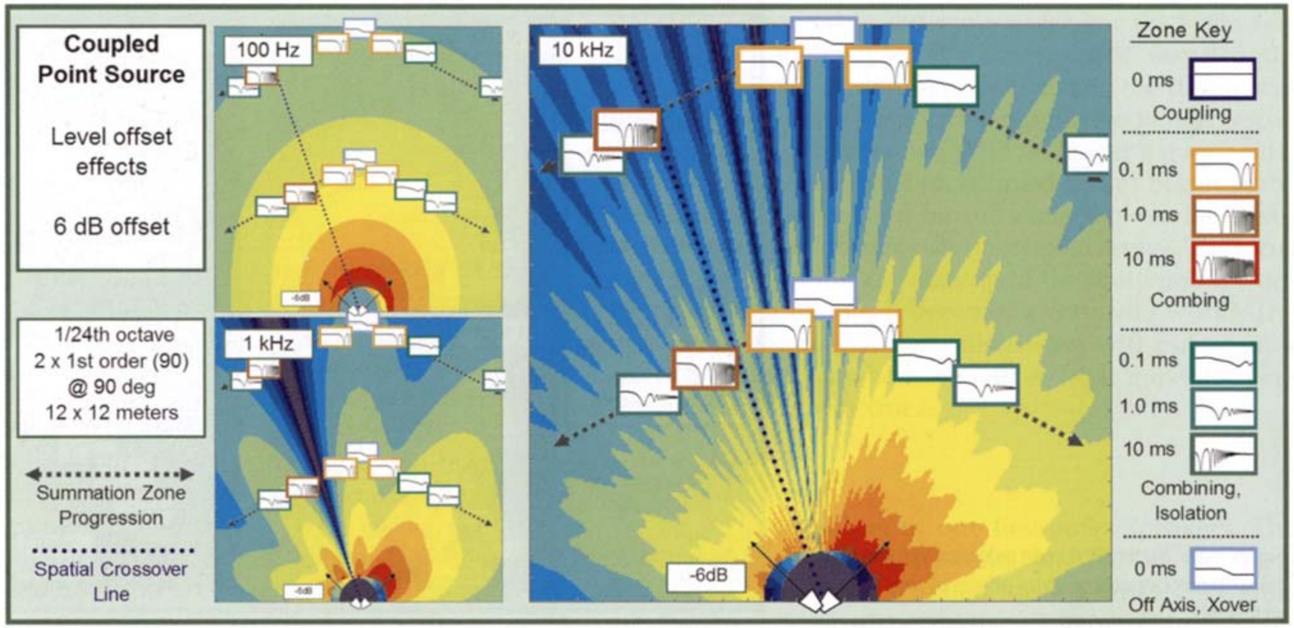


Figure 2.68 Summation zone progression factors for the couple point source array, level offset effects

line (formerly the crossover) the responses can not create a unity crossover. The levels at the original center are —6 dB and —12 dB respectively. The splay angle will need to be closed inward to compensate for the level offset.

Note that the maximum splay angle (where element equals splay) can only be used in a symmetrical crossover.

The phase-aligned asymmetrical point source requires three parameters to be managed. As level offset rises, the splay angle must decrease proportionally in order to prevent a gap in coverage. As level offset rises the delay offset at the newly created crossover point rises, and must be compensated. This process is shown in Fig. 2.69 and can be contrasted to the previous figure (2.68). The level has been offset 6 dB. The lower level element is turned inward until the splay angle is reduced by half (as was the level). The lower level element is then delayed and a phase-aligned asymmetrical array with a unity crossover is created. Note that the summation zone progression is now free of the combing zone interaction found previously and that the progression on either side of the crossover line begins symmetrically.

Three versions of the phase-aligned crossover are shown in Fig. 2.70. Each contains the same two elements but differs in level, splay angle and delay. The symmetrical version features matched levels, no delay and the maximum angle. The most asymmetrical version is the scheme just described in Fig. 2.68 (6dB offset, 0.46ms delay and 50 per cent overlap). The middle ground is found in the large panel where 3 db of level offset is compensated by 25 per cent overlap and 0.15ms of delay. In all cases the crossover area centers at the coupling zone and the summation zone progression moves predictably off of the center line towards isolation.

Uncoupled Arrays

We now move on to the uncoupled arrays. The speakers are no longer in close proximity and the gap coverage area can be extended far into the field. As a result, the progression will move slowly into the highly interactive unity and overlap zones. The large displacement, however, ensures that the interaction will become highly volatile once the

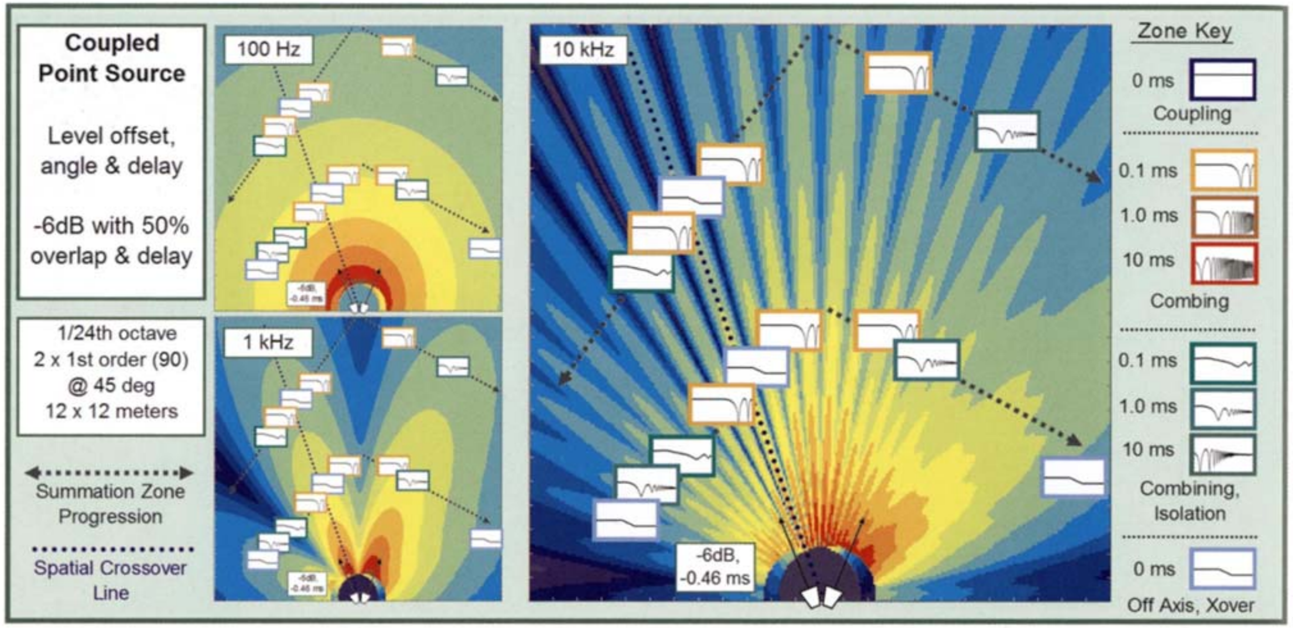


Figure **2.69** Summation zone progression factors for the coupled point source array, effects of level angle and delay

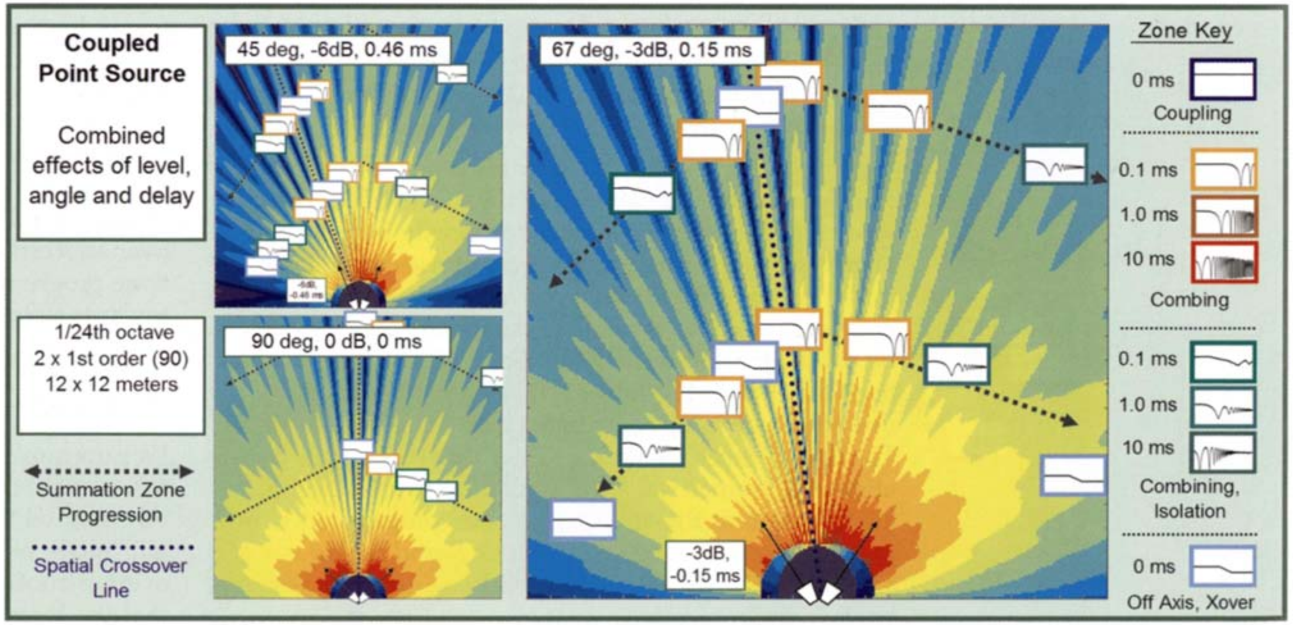


Figure **2.70** Summation zone progression factors for the coupled point source array, effects of level angle and delay overlap has begun. Uncoupled arrays are limited in the usable range of coverage due to these large displacements. The challenge will be to limit the range of coverage without excessively costly combing interaction.

Uncoupled Line Source Arrays

Crossover Class Progression

The uncoupled line source differs from the coupled line source only in scale. All of the crossover class and zone summation progressions will proceed in exactly the same order. The parallel pyramid will continue to stack up summations, albeit over larger distances. The reason for making the differentiation lies in the practical applications of our trade. We have huge issues of scale to deal with when we are tasked to provide coverage for the front row as well as the back of the arena. The coupled arrays will carry the long throw but the uncoupled arrays will rule the near field. We need to understand the behavior at the bottom of the pyramid just as clearly as the top. A glance at the unsealed renderings of Fig. 2.71 will show that the crossover class transitions proceed in the same order, for the same reasons as with the coupled line sources. The wavelengths we are transmitting do not rescale and so the differences in displacement will have very tangible effects. The rescaling makes the gap crossover class no longer a minor sideshow. Whereas our focus on coupled arrays began at the unity class crossover and moved on to the overlap areas, our viewpoint reverses for the uncoupled arrays. The focus instead becomes the gap area up to the unity point. Once we pass the unity line, it becomes a very wild world. The upcoming summation zone discussion will put us in a scaled perspective, where we will see that the old saying "size matters" is very true.

Summation Zones

We begin, as usual, with our first-order speakers. The uncoupled line source is shown in Fig. 2.72. The spacing is 3 m, which will be our standard for this series. Five units are displaced along a line and the response observed over distance and width. The most obvious features are that the

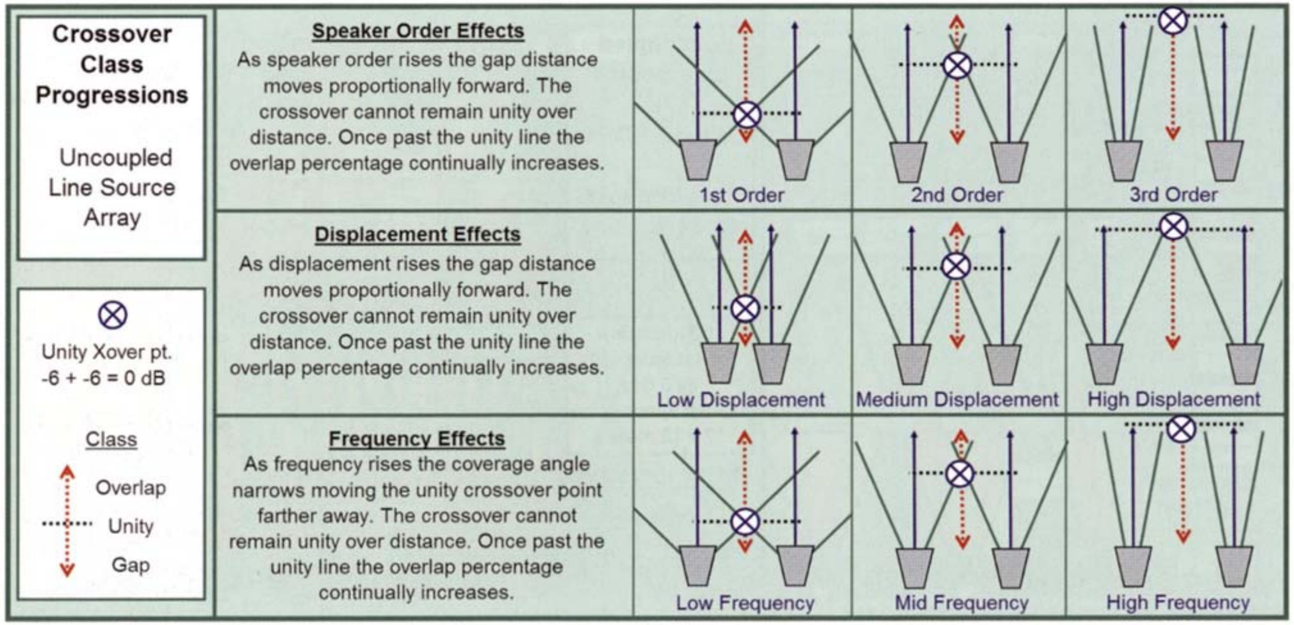


Figure 2.71 Crossover class progression factors

response has repeating horizontal themes but is drastically different from front to back. This contrasts sharply with any of the coupled arrays, which created a basic shape and held to it over distance. The presence of gap class crossover activity is clearly seen in the HF response. The first summation zone progression is shown moving along the unity class crossover line. As we move further away the overlap of additional elements brings in deep full-range combing zone interaction. There is no escape. The further we go the deeper it gets. A look at the LF response will reveal the familiar parallel pyramid. The 3 m spacing is wide enough to see multiple levels on their way up to the peak. The near area from the unity point to the next step in the pyramid is the place where the most uniform response can be found over a horizontal line. Note the difference between the height of the parallel pyramid in the HF and MF responses. The MF is lower due to the wider coverage of the elements. The difference in height is indicative of the level of uniformity in the early coverage regions. The uniformity over frequency rises as the pyramid heights approach equality.

If we keep the same spacing and change to a second-order speaker we will extend the gap area forward and move the location of the unity line. This is shown in Fig. 2.73. This is true, of course, only for those frequencies where we have increased control. In this case we have a front-loaded second-order system that has similar MF pattern control to its first-order counterpart. The result is range extension in the HF range only. The MF and LF ranges are overlapped far ahead of the HF. Again the most uniform response over the horizontal plane occurs between the unity crossover line and the first double overlap crossover (level 2 of the parallel pyramid). Note that the pyramid extension only occurred in the HF range, which increases the difference in the MF and HF heights. This is due to this particular element being a front-loaded LF/MF driver and therefore having the same MF coverage as the first-order elements.

We will select yet another speaker while maintaining our standard spacing. This time (Fig. 2.74) we will use a horn-loaded second-order system that provides increased MF directional control. The result is an extension of the crossover classes in the HF and MF regions and very closely

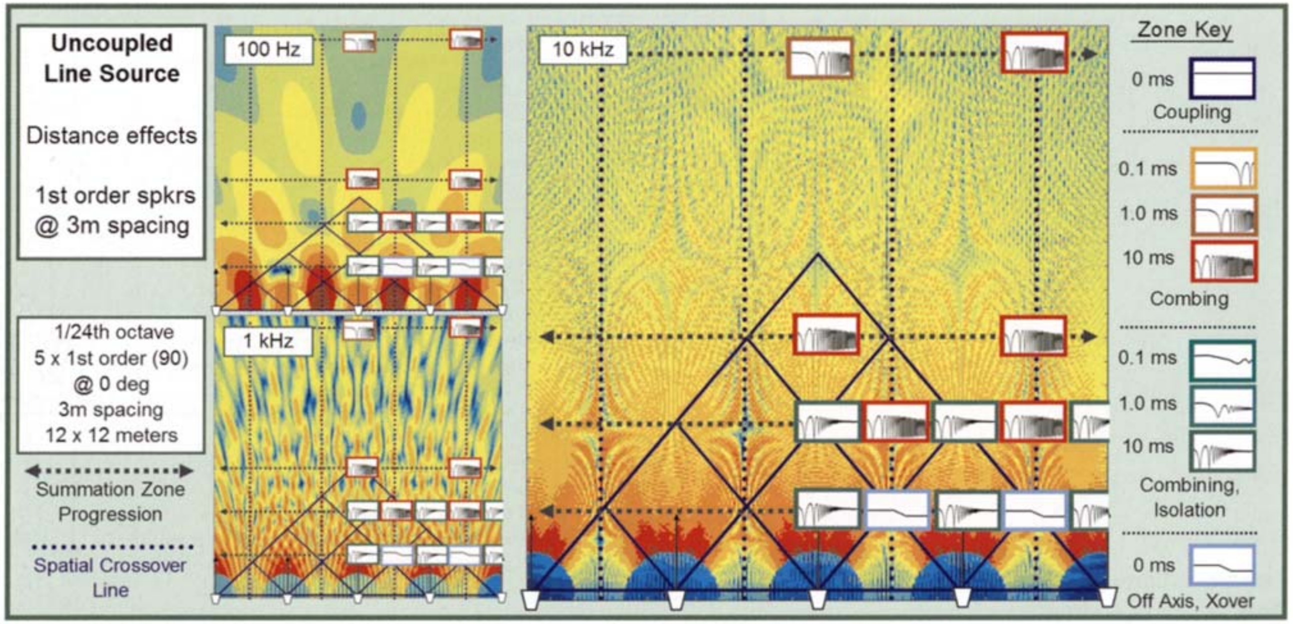


Figure **2.72** Summation zone progression factors for the uncoupled line source array, first-order speaker

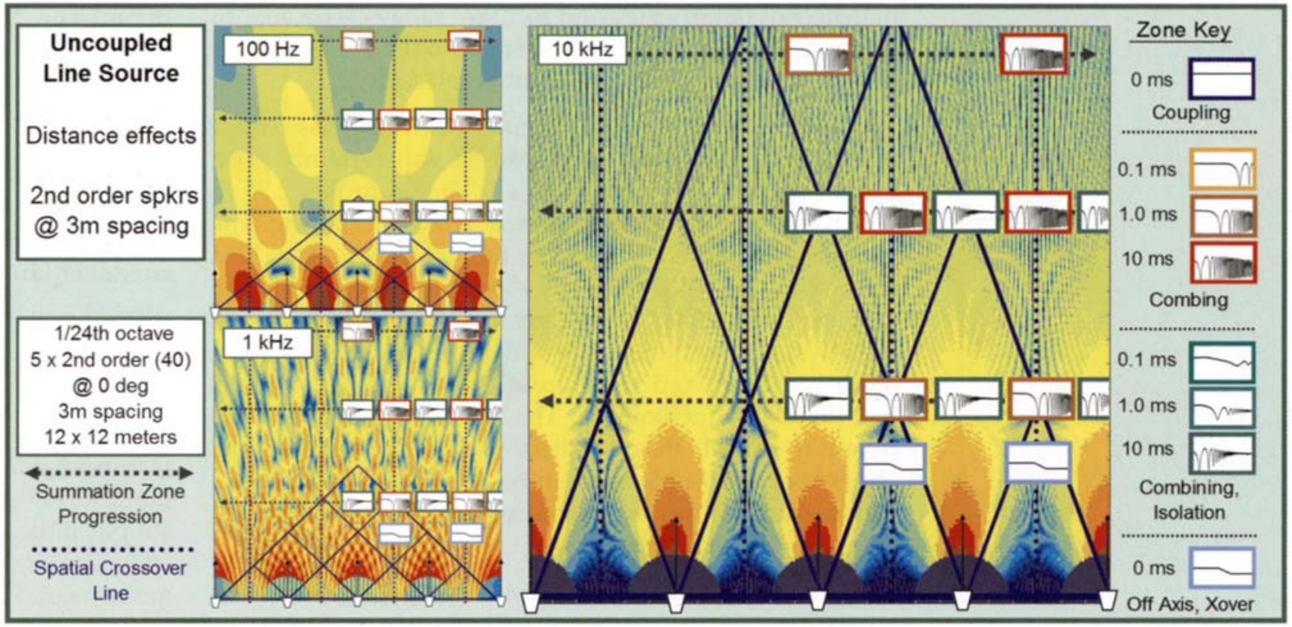


Figure 2.73 Summation zone progression factors for the uncoupled line source array, second-order speaker

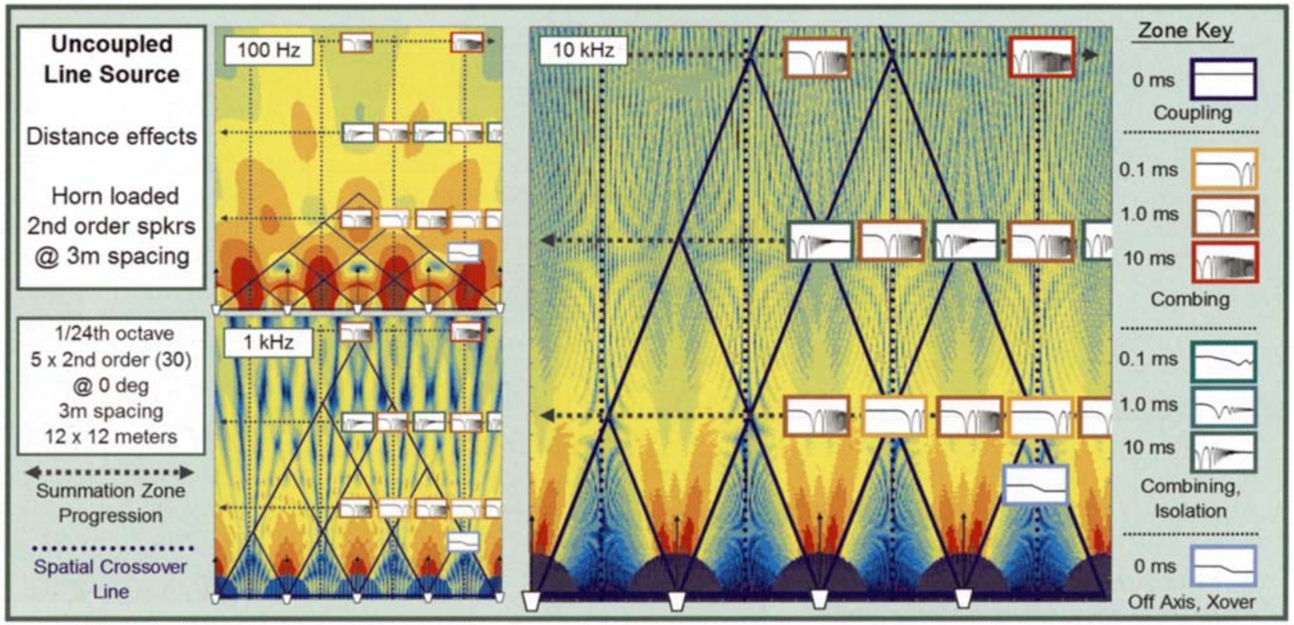


Figure 2.74 Summation zone progression factors for the uncoupled line source array, second-order horn-loaded speaker

matched pyramid extension in these ranges. The area of uniform coverage has been extended over the first-order system with comparable element displacement.

If a second-order system can move the unity class crossover line back, a third-order system will certainly push it back further. The truth of this is shown in Fig. 2.75 in which a third-order system is used at the 3 m spacing. The gap zone in the HF range has moved beyond the 12 m screen limit. It will eventually meet and have the longest range yet before it is "double crossed." Note, however, that the MF and LF responses are still quite similar to the first-order response, and therefore are overlapping within a few meters of the sources. The differential between the pyramid heights has reached its most extreme state. Unless our third-order system has MF directional control commensurate with its HF control we are doomed to have early overlap as frequency falls.

It is time to change the quantity and spacing of the elements (Fig. 2.76). In this case we will return to our first-order speaker, reduce the quantity to three and double the spacing to 6 m. The result is a scalar doubling of the position of the crossover class division lines. They move outward in proportion to the displacement change. Note the LF parallel pyramid has finally seen some change. The increased spacing expands the height of the pyramid steps and the decreased quantity reduces the total number of steps to two (from the previous four). Once again the zone of maximum uniformity lies between the unity and double crossover lines.

Uncoupled Point Source Arrays

We now add splay angle to the uncoupled equation. The uncoupled point source does not share the foremost feature of its coupled counterpart in that it cannot maintain a unity class crossover over distance. The reason is simple: if there is a 10 m gap between two elements whose splay angles meet their element coverage angles, they will never meet. 100 meters away the patterns will still be 10m apart. We can maintain gap coverage over distance, but if we want unity somewhere we will need angular overlap

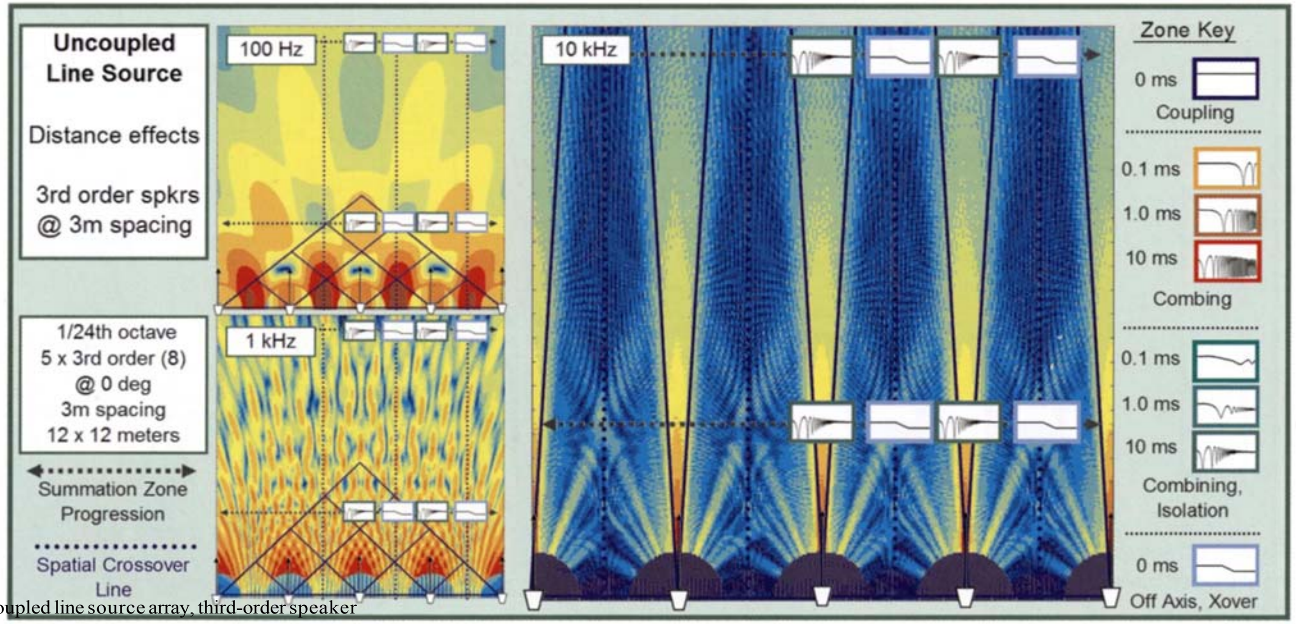


Figure 2.75 Summation zone progression factors for the uncoupled line source array, third-order speaker

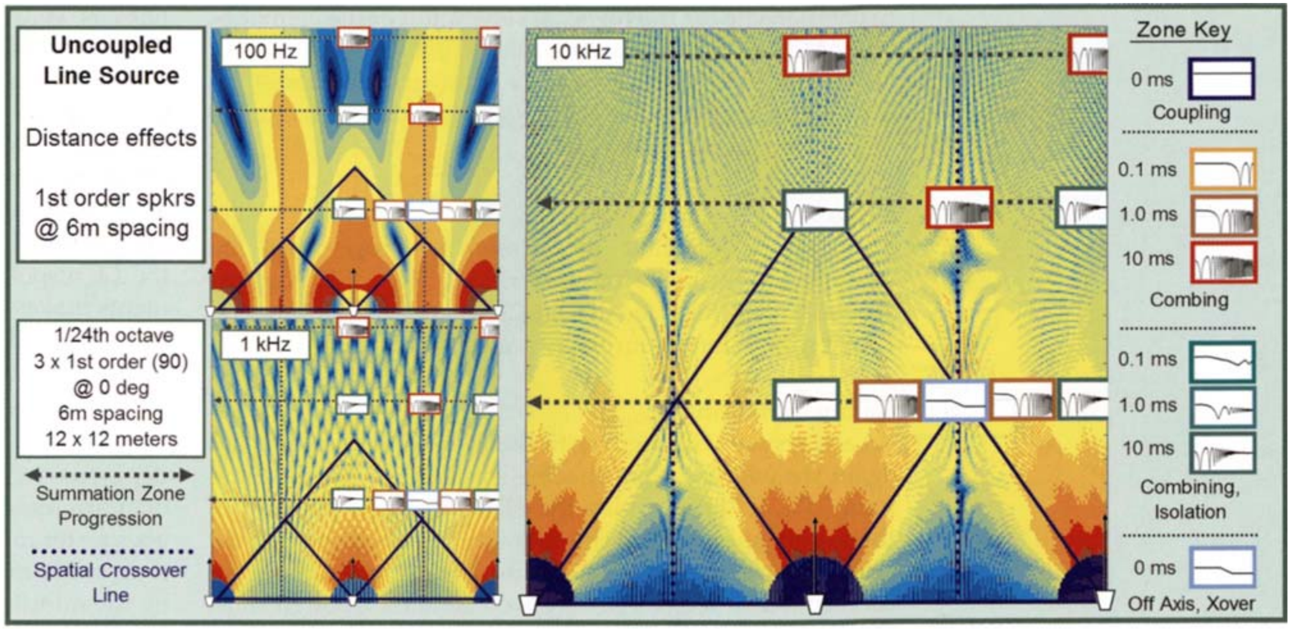


Figure 2.76 Summation zone progression factors for the uncoupled line source array, effects of spacing

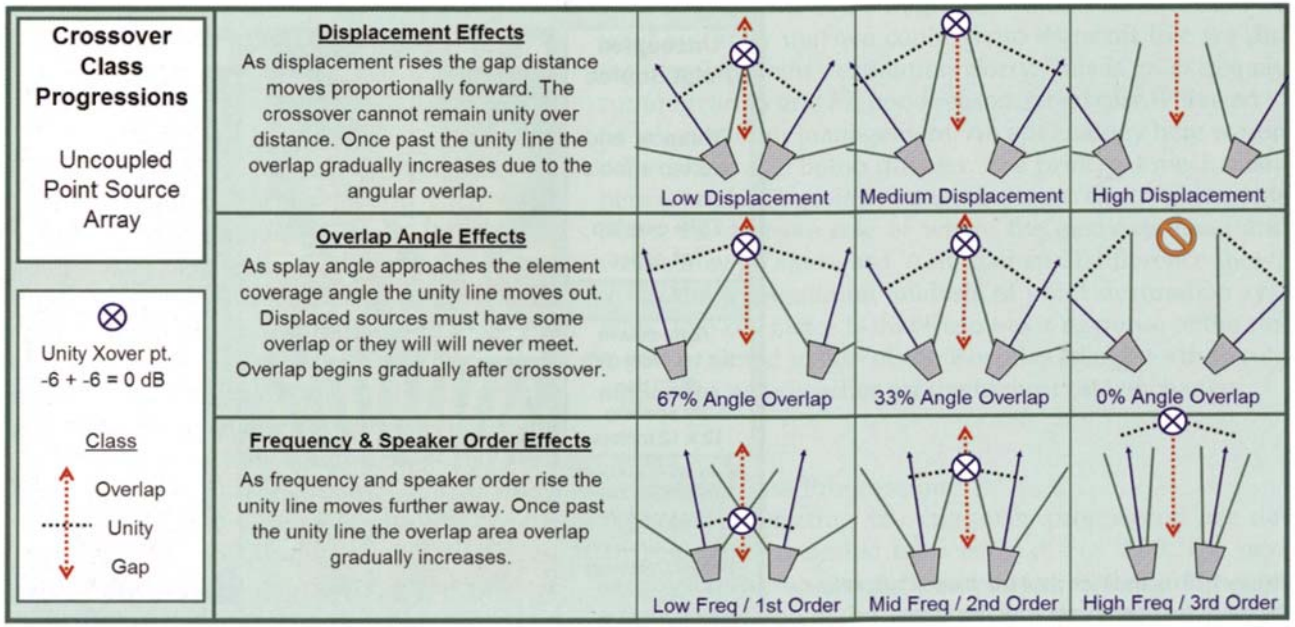


Figure 2.77 Crossover class progression factors for the uncoupled point source array

to compensate for the physical separation of the elements. It should be noted, however, that of the uncoupled array types, the point source has the most gradual transition into the overlap zone. In this regard it is closely related to its coupled counterpart.

Crossover Class Progression

The factors affecting the crossover progression are described in the unsealed renderings of Fig. 2.77. The next section takes these factors and observes the summation zone progression in our usual scaled application.

Summation Zones

As always we will begin with the first-order speakers and again we will use a 3 m spacing standard. In the first scenario (Fig. 2.78) we splay the elements at 50 per cent of their unity splay angle. This allows us to observe all of the crossover class transitions. In the HF response panel we can see that the elements are well isolated through most of their individual patterns. The isolation decreases as we move outward and the addition along the two crossover lines is evident. The summation progression gradually degrades over distance with combing zone interaction becoming more apparent. This degradation is far less than was found in the uncoupled line source systems regardless of spacing. There is evidence of mid-frequency isolation here as well. The displacement and angular splay yield areas of reduced ripple in the sides and, to a lesser extent, the center area. Note the lingering presence of the parallel pyramid in the LF response. The omnidirectional nature of these elements makes their relative angles a minimal factor in their summation. Therefore the response, though stretched outward, still shows the tell-tale signs of triangulation.

A change of splay angle compresses the range of action for the array as shown in Fig. 2.79. The result is earlier departure of the gap crossover and earlier arrival of the overlap crossover. The summation zone progression undergoes faster rates of change and more combing. The central area is the most affected since it becomes triple-covered by the middle of the panel. Another notable effect here is the correspondence of the LF response shape to that of the MF and HF shapes over distance. The three ranges have

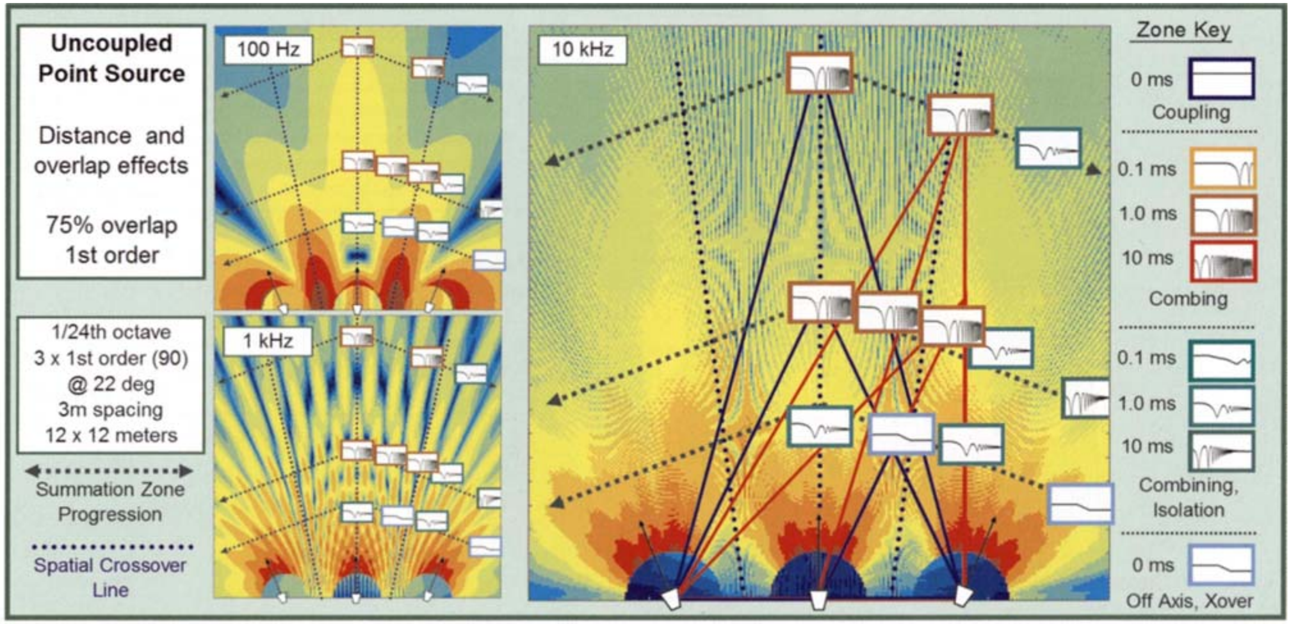


Figure 2.78 Summation zone progression factors for the uncoupled point source array, first order

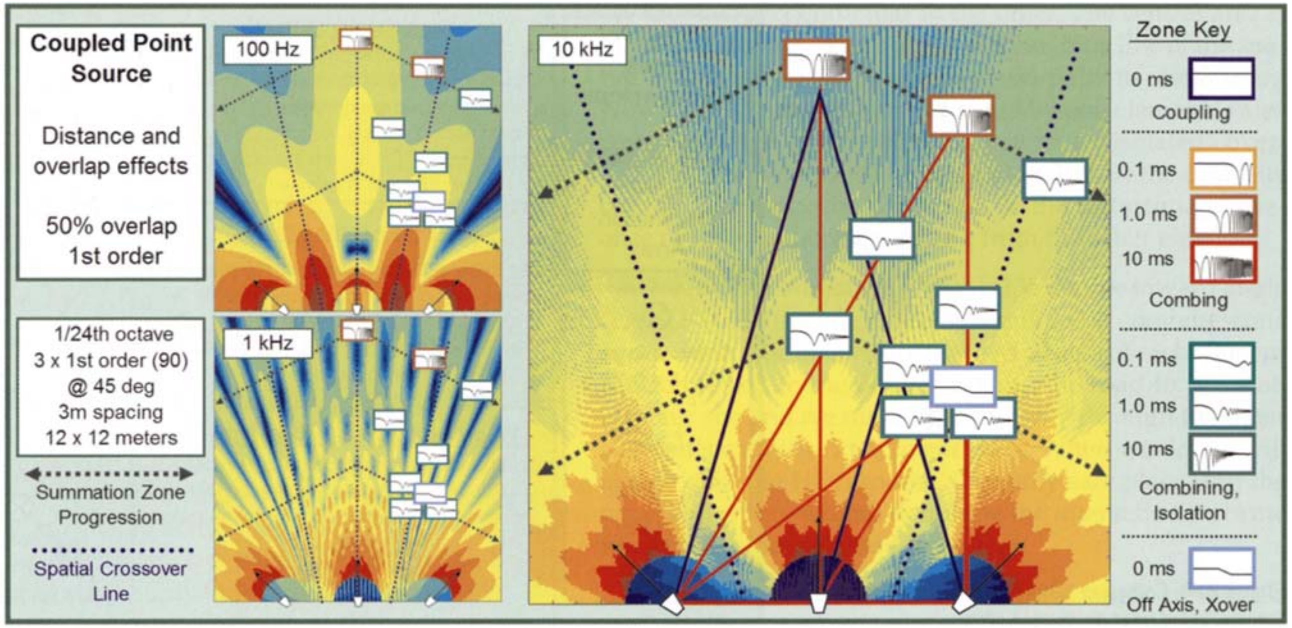


Figure 2.79 Summation zone progression factors for the uncoupled point source array, overlap effects

similar overall contours at the unity class crossover line (the first line across). Beyond that, the low frequency narrows at a much faster rate than the other ranges (the pyramid effect).

Uncoupled Point Destination Arrays

When speakers are pointed inward toward each other we are taking overlap to an entirely new level. In all other cases the overlap was something that happened to the sides and edges of our response patterns, rather than right on the nose. Not so with the point destination where we fire the on-axis beams of the speakers directly into each other. Since the speakers are uncoupled they will enjoy some period of peace before crashing through the opposing beam. Once on the other side they emerge again and head back towards isolation. Well sort of, at least. Unfortunately, over on the other side is the low-frequency presence of the other element to contend with. The uncoupled point destination is by far the most complex array type. Its response is the most spatially variable and can not hold a unity class crossover for more than a single point in the space. With all this going for it we might figure that these would be used so rarely that we could write them off like we did the coupled point destination. Sorry. This is an extremely common array and for good reason. Great care will need to be taken in their management. An apt analogy here would be working as a bomb diffuser. The principal mechanism here is angle. The relative angle between the two elements plays the decisive role of where the elements meet and where they go afterward. A fundamental difference shows up in the triangulation analysis of point destination systems. The key factor is that the on-axis response of the elements is aimed inside of the isosceles triangle — the acute angle area noted earlier for the highest rate of change.

Crossover Class Progression

The factors affecting the crossover progression are described in the unsealed renderings of Fig. 2.80. The next section takes these factors and observes the summation zone progression in our usual scaled application.

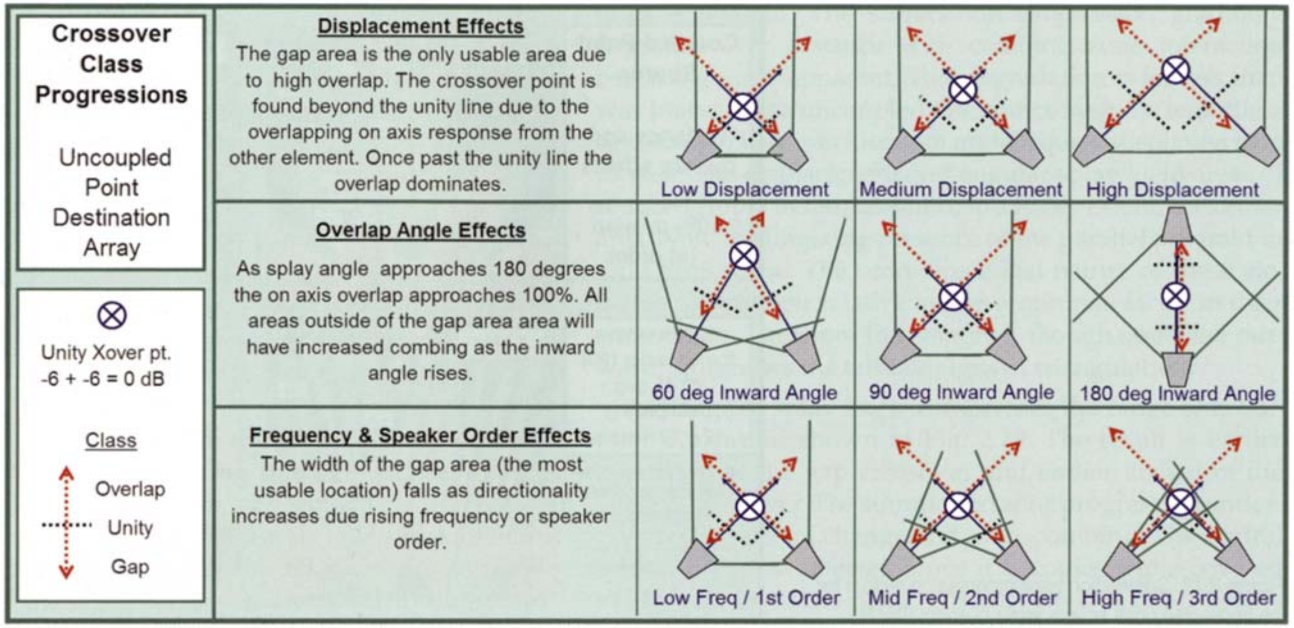


Figure 2.80 Crossover class progression factors for the uncoupled point destination array

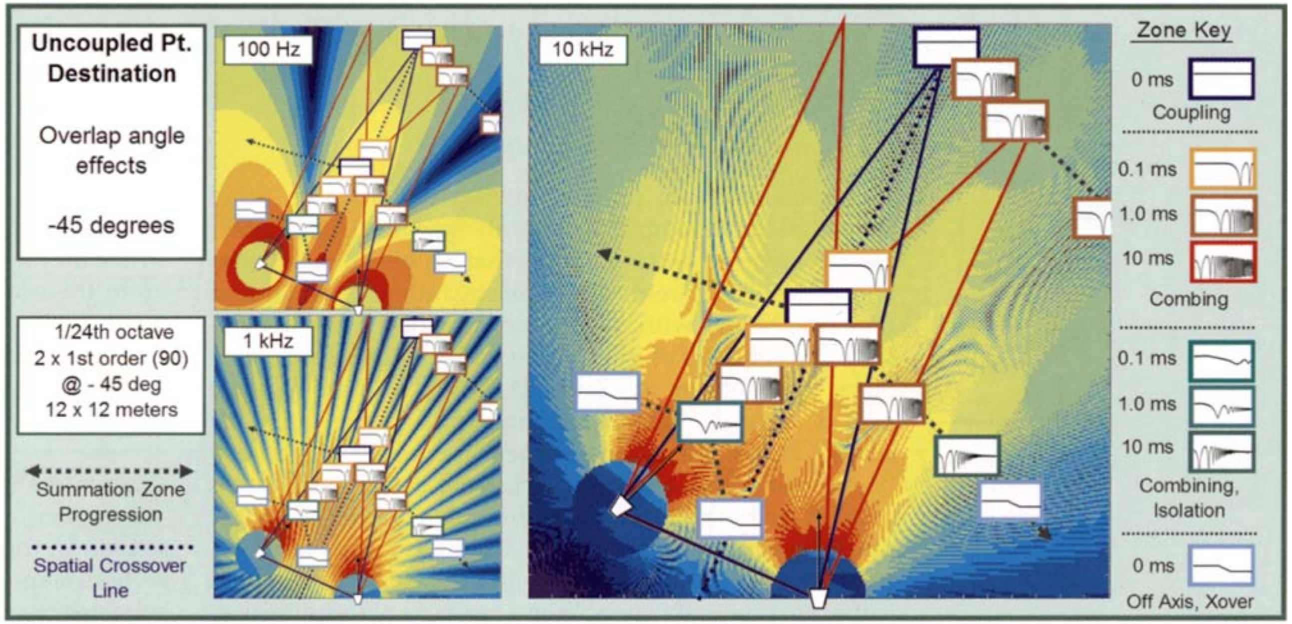


Figure 2.81 Summation zone progression factors for the uncoupled point destination array, 45 degree angle effects

Summation Zones

Here we go again with our first-order speaker. This time we will face the elements at a 45 degree inward angle (Fig. 2.81). The unity crossover point is the on-axis meeting point (the destination). The summation zones progress in multiple directions. There is a line that moves over distance (front to back along the on-axis line of each speaker), and others that move across the array from side to side. The crossover line bisects the two elements and is the line of equal relative level and time. This is the only place where combing zone interaction is not dominant except in the areas near to the elements. The rapid rate of change and large ripple depth is due to the shared on-axis response and the high rate of time offset change. Recall the triangulation model discussed early on in this chapter. In that model the obtuse triangle moved us in the direction of isolation. By facing the elements inward we are in the acute angle region and have given up the angular means of escape. Only the brute-force-level dominance of the near field provides relief.

We can open up the angle further to 90 degrees between the elements (Fig. 2.82). As we approach one speaker we are moving perpendicular to the other. The central area is all combing zone, except along the thin blue line of the spatial crossover. We have reached a mid-point in our journey. Now all of the angular area covered by one element is also covered by the other. The only way to get isolated coverage is in the near field (by level alone) or in the opposite corners. The opposite corners have the unfortunate presence of low-frequency combing from the other element.

We continue our journey with a 135 degree inward angle (Fig. 2.83). As the angle rises there is more combing zone summation. The majority of the on-axis area of each element is also on-axis to the other element. The road to isolation is along one and we must fight our way through the other element's coverage area to get there. When we do we will isolate only the HF response. The only safe place is in the near field where we can dominate the interaction by brute force of distance related level offset.

Finally we reached the most inward angle of all: 180 degrees. This array has the dubious distinction of being the most rapid movement into the combing zone and the

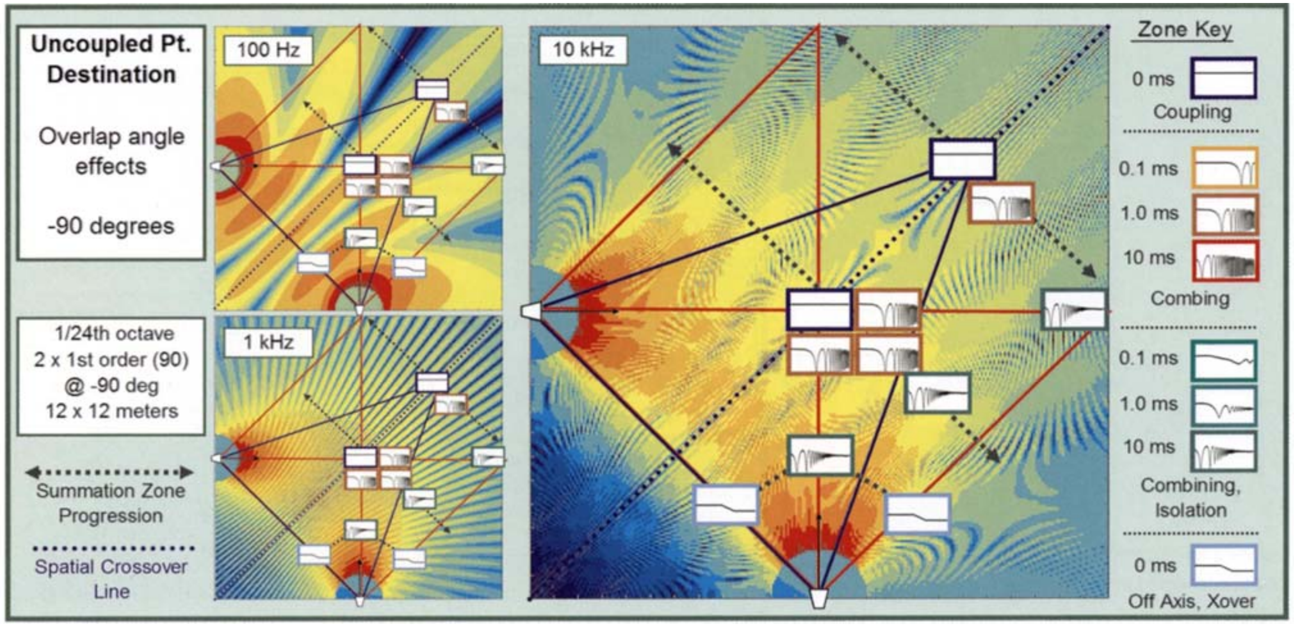


Figure 2.82 Summation zone progression factors for the uncoupled point destination array, 90 degree angle effects

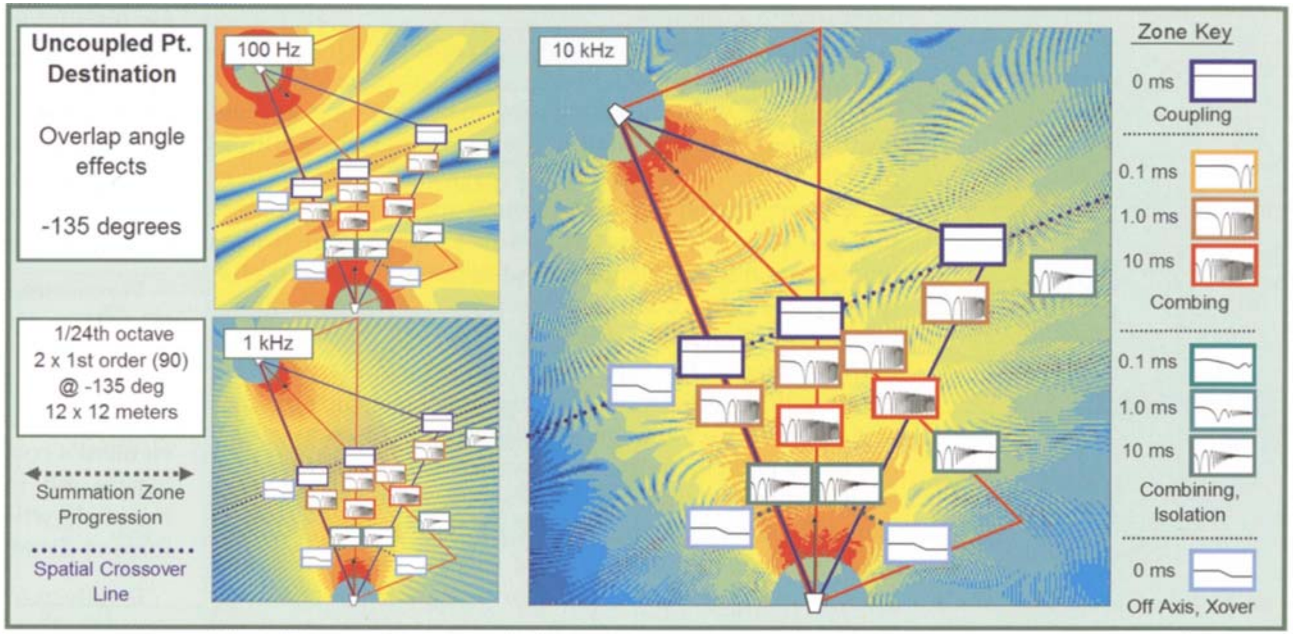


Figure 2.83 Summation zone progression factors for the uncoupled point destination array, 135 degree angle effects

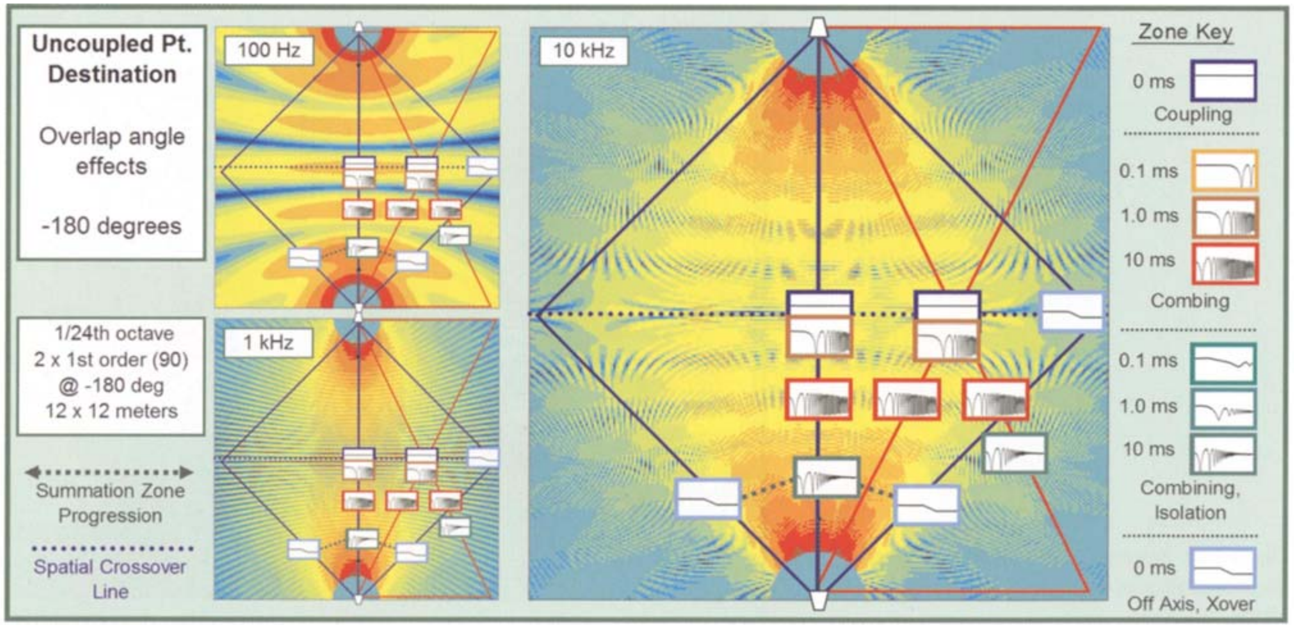


Figure **2.84** Summation zone progression factors for the uncoupled point destination array, 180 degree angle effects least prospect of escape. There is no angular mechanism to move us toward the isolation zone. Only the brute force of distance-related level offset will provide relief. If we are on-axis to one element, we are equally on-axis to the other. It's guaranteed. The rate of progression is the highest because a movement toward one element is de facto a movement away from the other. The rate of time offset change is therefore the highest possible. Now more than ever, the only safe place is in the near field where we can dominate the interaction by the dominance of distance related level offset.

Figure 2.85 shows the three different speaker orders arrayed at the same inward angle. The affected area narrows as the speaker order rises but there is no movement of the basic lines. The crossover is still at the forward destination. The summation zone progression completes its journey to the off-axis point more quickly but the sequence is otherwise unchanged. The low-frequency distribution (not shown) is largely unaffected. The triangulation aspects are unchanged by the speaker order. The difference is simply how much of the space is filled with interaction. The lessening of interaction is accompanied by a proportional lack of coverage, in the HF response. Since the LF and MF responses will continue the highly variable interaction found earlier (see Fig. 2.81) the third order is actually the worst of the three choices. All orders are usable only in the near field, and totally destructive beyond.

A change in level will cause the crossover line to shift toward the lower level element. Level offset in displaced arrays causes both a movement and bending of the crossover line. The bending is due to the different level loss rates due to the different doubling distances in a displaced asymmetrical summation. A level asymmetry of 6 dB is introduced in a 90 degree point destination array in Fig. 2.86. The crossover line has moved toward the upper left speaker and begun to bend inward toward the quieter speaker. The bending action can most easily be visualized by considering the consequences of further level offsets. As the level goes down the quiet speaker gradually concedes equal level around its perimeter, until eventually it is surrounded by the more powerful speaker. This panel illustrates a change that reapportions the combing zone

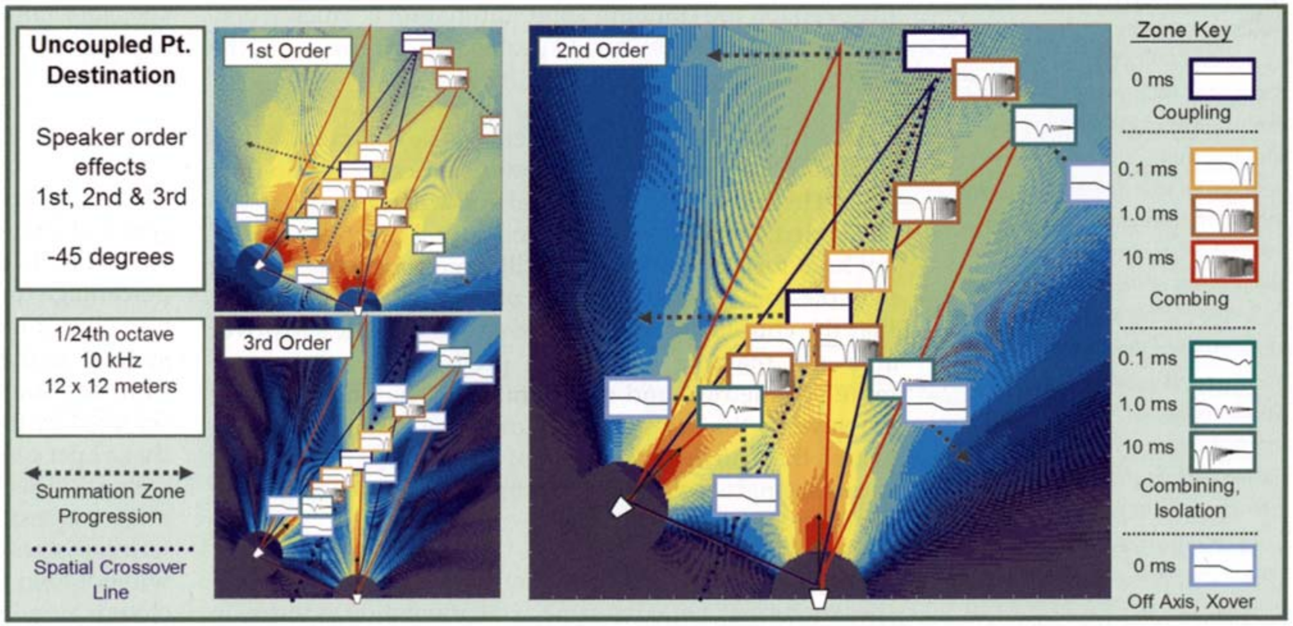


Figure 2.85 Summation zone progression factors for the uncoupled point destination array, speaker order effects

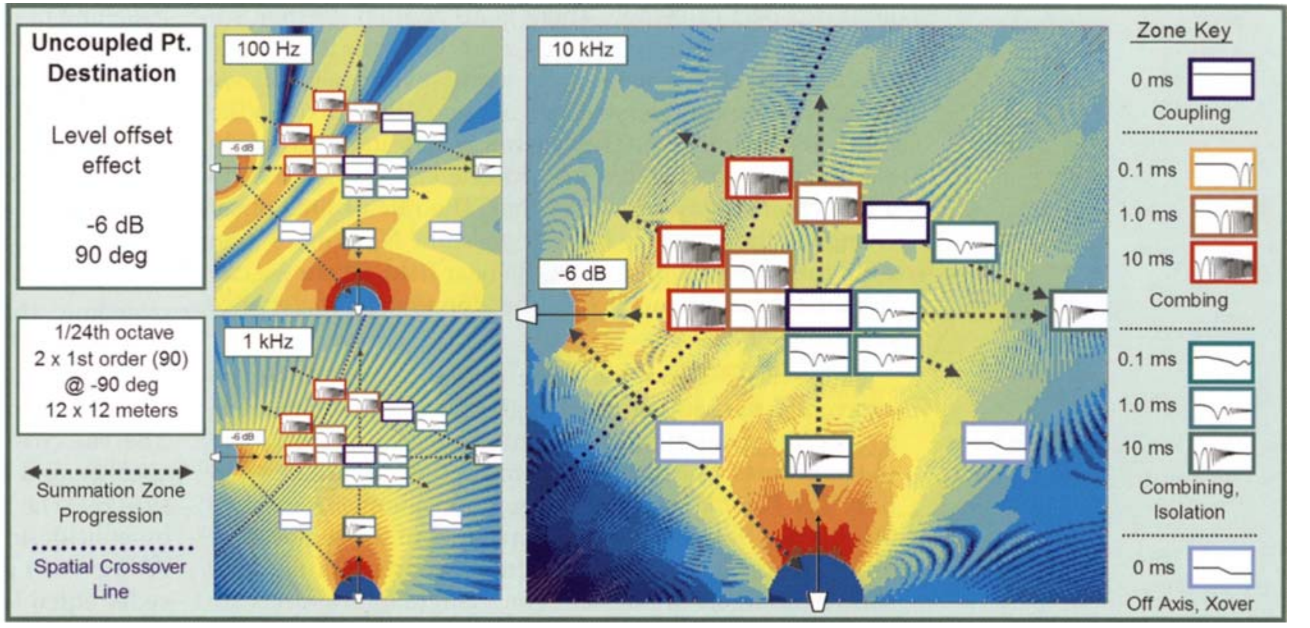


Figure 2.86 Summation zone progression factors for the uncoupled point destination array, level effects

Perspectives: Upon arriving at the gig for 8,000 with the console about 400' back in the room, I was given a brief explanation of the sound design. "The system has been set up as a distributed system instead of a delayed system. The farthest delay clusters only have about 80 ms of delay."

Seeing the big stacks of speakers at the front of the room, the question was posed, "What about the subwoofers?" "The subwoofer information is going to arrive at the farthest clusters about 200 ms after the full-range program."

The reply: "Oh, I forgot about the subwoofers."

RS

summation but does not reduce it. Compare this figure to the level symmetrical version (Fig. 2.82). In this asymmetrical version the combing zone summation is much worse in the area near to the quiet speaker, while reduced in the area of the louder speaker.

The shifted crossover placement can be compensated by insertion of a delay for the quieter element scenario (Fig. 2.87). This allows the equal level and equal time points to be merged; the phase-aligned crossover. The crossover cannot hold for all positions but will have substantial benefit over the non-delayed approach of the previous page. Note that the coupling zone has moved back to the crossover area and that the summation progression has become more symmetrical and more uniform. The phase-aligned crossover is the key to spatial uniformity. The proportion of area that can benefit is highest as the angle between the sources is reduced. In this case the angle is 90 degrees but the benefits are discernible.

Figure 2.88 shows a speaker set up in the typical "delay" speaker format, i.e. at the same axial orientation as the main speaker. Such a setup obviously requires a delay, and an appropriate amount is inserted in this case. Note that the speakers fall out of sync as we move off-axis from the delay speaker. This is due to the familiar right triangle distance differences as we have found earlier.

Environmental Effects

We have previously discussed that temperature affects the speed of sound. The components of an acoustical summation will all have their sound speed changed by the same percentage (provided they all undergo the same temperature change). This will create no effect on the summation as long as the path lengths are identical. If the distance from one source to another differs, the time offset will be changed by the temperature movement. Consider the fact that a 1 per cent change in sound speed causes a 1 ms change over a 100 ms transit time, while only 0.1 ms changes over a 10ms transit time. The result is that two speakers at differing distances which had been phase-aligned by delay will not hold the alignment over temperature. A reference chart is provided in Fig. 2.89.

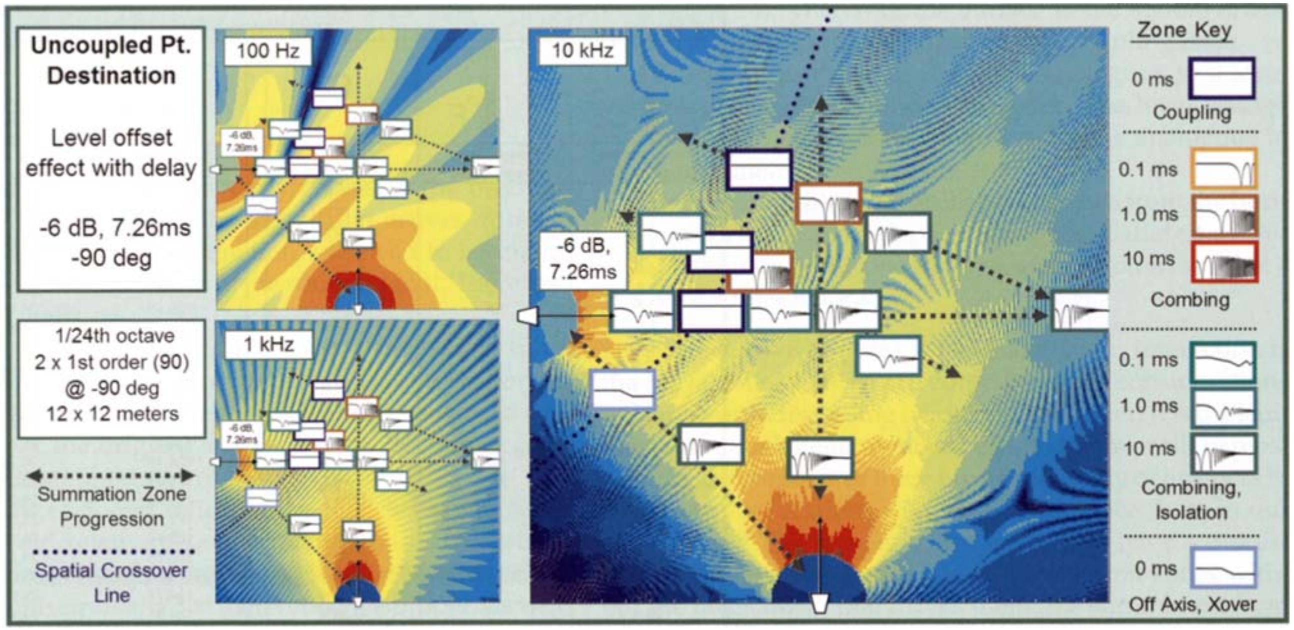


Figure **2.87** Summation zone progression factors for the uncoupled point destination array, level and delay effects

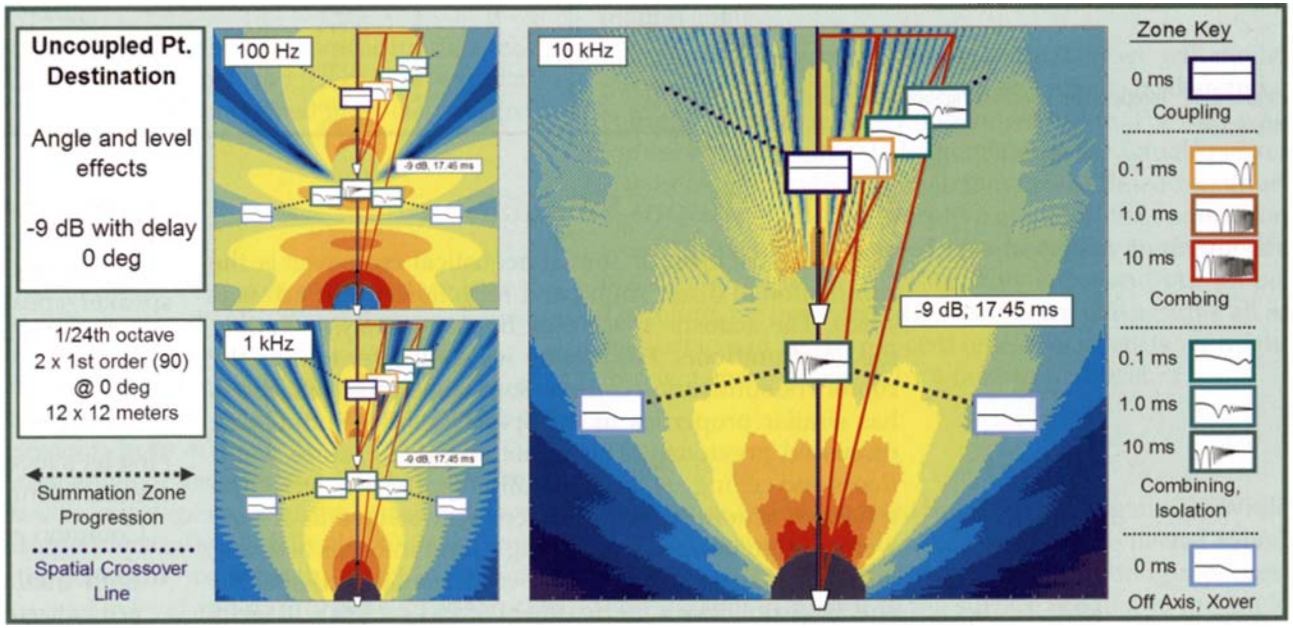


Figure **2.88** Summation zone progression factors for the uncoupled point destination array, angle and level effects

Temperature Effects on Speaker/Speaker Summation

°C Temperature Degrees (C)	°F Temperature Degrees (F)	ΔC Speed of sound change (%)	T Main / T Delay / T Delay Spkr Main Spkr Arrival (ms)	Delay Spkr Arrival (ms)	Synchronised delay time (ms)	ΔT Change from room temp setting (ms)	Comb filter frequency (Hz)
0.0	32	3.91%	129.89	25.98	103.91	-3.91	256
1.1	34	3.72%	129.65	25.93	103.72	-3.72	269
2.2	36	3.52%	129.40	25.88	103.52	-3.52	284
3.3	38	3.32%	129.15	25.83	103.32	-3.32	301
4.4	40	3.13%	128.91	25.78	103.13	-3.13	319
5.6	42	2.93%	128.66	25.73	102.93	-2.93	341
6.7	44	2.74%	128.43	25.69	102.74	-2.74	365
7.8	46	2.54%	128.18	25.64	102.54	-2.54	394
8.9	48	2.35%	127.94	25.59	102.35	-2.35	426
10.0	50	2.15%	127.69	25.54	102.15	-2.15	465
11.1	52	1.96%	127.45	25.49	101.96	-1.96	510
12.2	54	1.76%	127.20	25.44	101.76	-1.76	568
13.3	56	1.56%	126.95	25.39	101.56	-1.56	641
14.4	58	1.37%	126.71	25.34	101.37	-1.37	730
15.6	60	1.17%	126.46	25.29	101.17	-1.17	855
16.7	62	0.98%	126.23	25.25	100.98	-0.98	1020
17.8	64	0.78%	125.98	25.20	100.78	-0.78	1282
18.9	66	0.59%	125.74	25.15	100.59	-0.59	1695
20.0	68	0.39%	125.49	25.10	100.39	-0.39	2564
21.1	70	0.20%	125.25	25.05	100.20	-0.20	5000
22.2	72	0.00%	125.00	25.00	100.00	0.00	
23.3	74	-0.20%	124.75	24.95	99.80	0.20	5000
24.4	76	-0.39%	124.51	24.90	99.61	0.39	2564
25.6	78	-0.59%	124.26	24.85	99.41	0.59	1695
26.7	80	-0.78%	124.03	24.81	99.22	0.78	1282
27.8	82	-0.98%	123.78	24.76	99.02	0.98	1020
28.9	84	-1.17%	123.54	24.71	98.83	1.17	855
30.0	86	-1.37%	123.29	24.66	98.63	1.37	730
31.1	88	-1.56%	123.05	24.61	98.44	1.56	641
32.2	90	-1.76%	122.80	24.56	98.24	1.76	568
33.3	92	-1.96%	122.55	24.51	98.04	1.96	510
34.4	94	-2.15%	122.31	24.46	97.85	2.15	465
35.6	96	-2.35%	122.06	24.41	97.65	2.35	426
36.7	98	-2.54%	121.83	24.37	97.46	2.54	394
37.8	100	-2.74%	121.58	24.32	97.26	2.74	365

Figure 2.89 Temperature effects on speaker/speaker interaction

Speaker/Room Summation

The alternate form of spatial acoustical crossover is the summation of direct sound and reflections in the acoustic space. The acoustic space may be any variety, including the great outdoors. For brevity we shall refer to it as the room. The summation of direct sound and room reflections has similar properties to the spatial divider summation discussed previously. This is not surprising, since walls, floors and ceiling are quite literally "spatial dividers." The reflective action of these surfaces will behave as if they were additional speakers adding sound into our listening area. For each type of surface there are analogous speaker and array qualities. Our treatment of this subject will focus on these common qualities, and contrast them as required. The common ground shared between these two types of crossovers is so substantial that our understanding of speaker/speaker summation tells us most of what we need to know about speaker/room summation.

Analogous Functions

First let's focus on the common ground between speaker/speaker summation and speaker/room summation.

Common ground:

- Both interactions are governed by the properties of acoustic summation.

- The strategies for optimal summation are rooted in the same concepts.

Differences:

- Room reflections are inherently copies of the original signal. They will reflect the signal that arrives at their surface. Therefore no distinction will need to be made in terms of the correlation of the source signal. There are no stereo walls.
- Reflections can continue for extended periods of time as they move from surface to surface around the room. The length of time for these reflections may exceed the duration of the original signal, thereby becoming outside of our summation criteria.
- Most surfaces do not reflect all frequencies at the same relative level, relative phase and same angle of incidence. Surfaces are of such complexity that for our purposes only a rough equivalent of their reflective properties will be asserted. These filter effects will cause the summation properties to be variable over frequency. Once taken on a frequency-by-frequency basis, however, the analogy holds very well.
- Except for the filter effects cited above the room/speaker crossover is "self aligning," i.e. it can not be polarity reversed, it is in time and matched in level at the crossover (the surface).

Analogous organs:

- The distance between a speaker and a surface is analogous to half the source displacement distance between two speaker elements in an array.
- The relative angle between the speaker and the surface is analogous to half the splay angle between two speaker elements in an array.
- The coverage angle of the reflected "virtual" speaker is the same as the coverage angle of the source speaker.
- The surface is a physical spatial divider. Only at the surface does the possibility exist that the direct and reflected sound are equal in level and time (0 per cent absorption is required). Therefore the surface is the spatial crossover. This area will once again be the coupling zone.
- Absorption in the wall surface is analogous to a level reduction of the secondary speaker. The level change causes the relative level to shift, which is analogous to an asymmetrical crossover.
- The power addition capability that comes from the horizontal or vertical overlap in the surface spatial crossover is analogous to the coupling zone addition in speaker/speaker int

Since the crossover point for speaker/room interaction is the surface itself there are some practical considerations that are worth discussing before going further. It is difficult to imagine our listeners embedded into floors, walls or most notably, ceilings. Since there is no probability of listeners in the crossover locations why should we place them in our discussions? The reason is that we cannot have a discussion of the combing, combining and isolation zones without the mechanism that drives them: the summation from the virtual source behind the walls. These virtual speakers "cross over" into our audible world at the walls. Therefore, the discussion must include all of the players in the summation game.

In the case of speaker/room summation we will model the interaction as if it is between the actual speaker and a "virtual" speaker representing the reflection. The only way for the virtual speaker to be matched to the actual speaker is if there was no absorption in the surface. Since every surface has some unique absorptive qualities, the summation would have to be evaluated on a frequency-by-frequency basis to be perfectly accurate. This is beyond our scope. Therefore, for our initial discussions we will assume all of the surfaces to be 100 per cent reflective, such as we might find in a space designed by a famous architect.

Crossover Zones

Reflections can be modeled by the ray tracing model. While this does not explain all of the behavior of reflected sound it serves our purpose for modeling summation properties. The angle of incidence to the surface is equal to the angle of reflection. The ray tracing model sees the reflection of sound off a wall as analogous to light reflecting off a mirror. Place a speaker in any orientation and find its distance to the surface. The virtual speaker is a mirror image that is "placed" at the same distance and opposite orientation behind the surface.

We will also find analogous representation of the crossover classes: the gap crossover, the unity crossover and the overlap crossover. Each different type of speaker orientation will yield different positions for the summation zone and crossover zone divisions, just as was the case for speaker arrays.

Crossover Slope

The directionality of the speakers defines the slope of the crossover range, with narrow speakers having a range-shortening effect. Our goal is to transition out of the coupling zone (the surface) toward the isolation zone before the relative phase responses reach the subtractive side of the circle and land us in the combing zone. The race of amplitude vs. phase is on again. What are the prospects of getting to the isolation zone before cancellation with a 16 kHz crossover point? They are actually better here than with speaker/speaker interaction for a number of reasons. The first is that we are not required to reach the surface with a unity class crossover. If there are no seats at the wall or ceiling, we are not required to reach there at unity gain. If the surface is highly reflective we can aim the speaker such that the last audience members are at the —6 dB point, our maximum acceptable standard of variance. We can create a gap in coverage at the crossover surface whenever possible and thereby reduce the reflected energy. The second is that the wall may have some absorption, especially in the high frequencies, where the wavelengths are most easily sent into the combing zone. There is no perfect solution, but it can be approached.

Speaker room summation considerations:

- There is no such thing as a speaker with uniform coverage over frequency. Therefore the degree of crossover overlap changes with frequency.
- The best prospect for success is a speaker with equal or greater directional control in the high-frequency range than the lows.
- Since the wavelengths at the high end are the smallest we will have the least margin for phase error. Therefore, controlling the high-frequency overlap will be crucial. Mid-range and low-frequency overlap can be more easily accommodated since the wavelengths are long enough to minimize cancellation.
- Overlapped high frequencies will need to have the shortest possible differences in path length. If the displacement is large compared to the wavelength the reflection level will need to be reduced by directional control, positioning or absorption.

Crossover Audibility

The final chapter in crossover audibility plays out in the room. When we find ourselves staring at the source of a reflection at the front of the balcony rail we are looking right at the crossover. Needless to say we do not want our listeners looking through walls at the off-beat drummers playing in the virtual speakers behind them.

Crossover audibility in the room mirrors the spatial divider in that it gives itself away by shifts in angular position. The crossover between the speaker and the room will be most transparent if the relative angle of origin of the two sound sources is small. Our ears are able to pick up localization clues, and large angles between the speaker and the echo will give them away. As was the case in speaker/speaker sum
tive due to our binaural localization.

Excess overlap will give away the position of the echo source. It is impossible to keep the systems synchronized at any position except at the crossover (the surface). As long as the reflection arrival is close behind the direct sound, the listener will experience a tonal change and sense of spaciousness. When the reflections fall too far out of time but remain strong in level, the ear begins to pick up the two sources separately and our perception becomes that of a discrete echo. As was the case in speaker/speaker summation, strong high-frequency content makes it easier to localize as a discrete echo. Position speakers to minimize overlap of the high frequencies into the walls.

Speaker/Room Summation Types

Introduction

Speaker/room behavior is the advanced version of applied summation. Place a speaker system in a room and it will sum with its reflections. How they sum will depend on their relative amplitude and phase. These values depend on the directionality and orientation of the speaker relative to the surface, and its absorption properties. The goal of this section is to review the different types of speaker to surface relationships and learn how the properties of summation affect them. The properties of the different speaker/room summation types will begin to point us toward our principals of design.

There are three different types of relationships between the speaker and the room: speaker and surface at the same angle, surface angled outward, or surface angled inward. Sound familiar?

Perspectives: An underlying principle of system optimization is the definition of the system. In the past there was great focus on separating the response of the sound system and the room. More recently, the understanding is clearer, and the focus of system optimization has been on optimizing the systems interactions with itself and the acoustics of the room. This approach, which is "built-in" to multitime windowed transfer functions, is extremely powerful, as it correlates well with hearing !

Sam Berkow

For the purposes of our discussion here, we will once again divide the summation types into two versions each: coupled and uncoupled. The term "coupled surface" refers to those where the speakers are physically placed in direct proximity to the surface. Ground stacked subwoofers, "half space loading" and soffit mounting are examples of this type. "Uncoupled surfaces" are those where the speaker is placed away from the wall by some unspecified distance. Does a speaker hung three meters in the air respond the same as one sitting just 300 cm off the floor? No and yes. To 100 Hz the behavior is quite different between them. But the behavior at 100 Hz for the flying unit is the same as that of 1 kHz in the ground-based speaker. This is because the spacing of the units relative to the wavelength is preserved. The behavior of speakers with respect to surfaces will always have the same scaling factors associated with distance and wavelength that we found in the speaker/speaker interaction. These terms serve as guidelines regarding the behavior of speakers in rooms when closely or more distantly spaced. However, bear in mind that the back wall of an arena feels as close to your subwoofers as does a sheet of plywood placed one foot in front of your dome tweeter.

Note: Unless specified otherwise, this section assumes the surfaces to be 100% reflective at all frequencies. The summations are described on a single plane, vertical or horizontal.

The six different interaction types are shown in Fig. 2.90. They are the familiar set of possibilities that we have faced previously. The parallel surface is analogous to the line source array, the outward wall to the point source array and the inward wall to the point destination. The following drawing series details the progression of the summation zones over the space. Thankfully, these panels will do most of the talking for the summation behavior of the speaker/room summation types. A guide to reading the summation zone panels was provided earlier in Fig. 2.54.

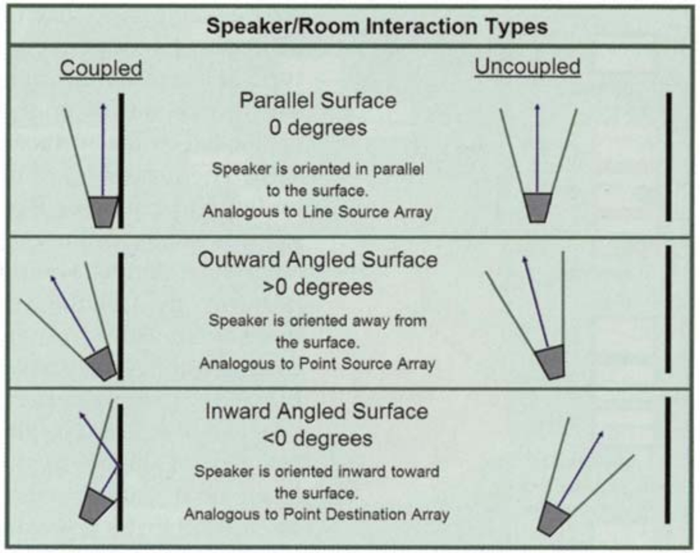


Figure 2.90 Speaker room interaction types

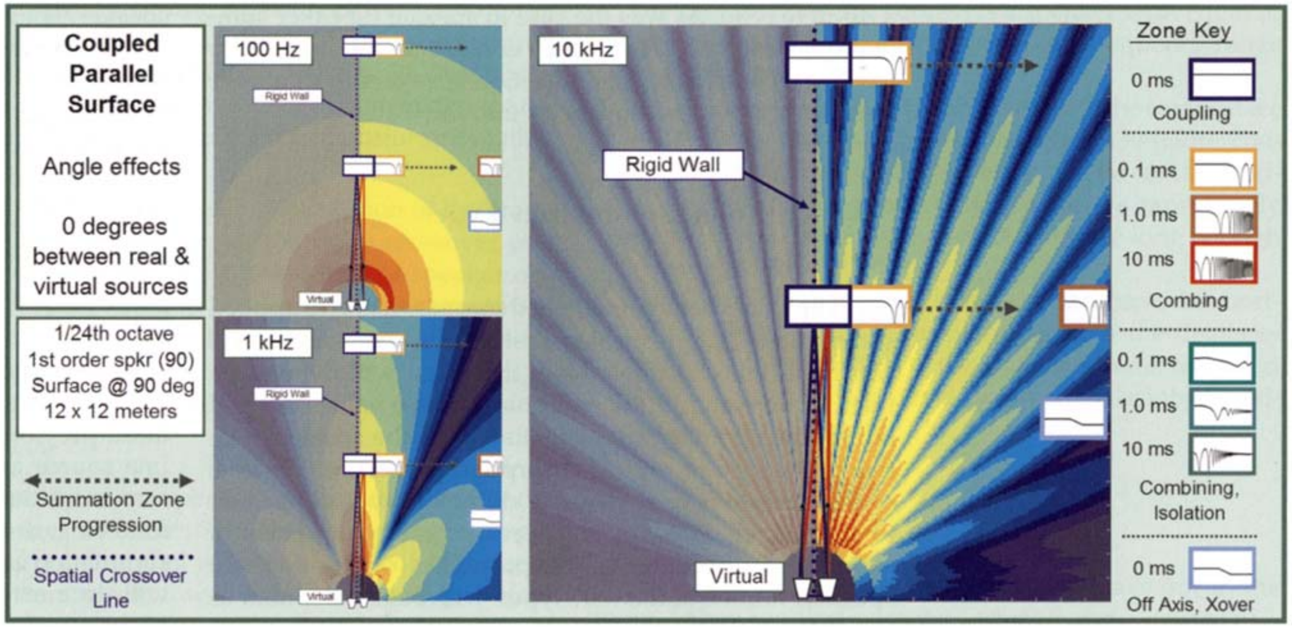


Figure 2.91 Summation zone progression factors for the coupled parallel surface

Coupled Parallel Surface

A speaker placed flat on the floor is the most obvious example of a coupled parallel surface array. This particular type of summation has its own name: half space loading. The term comes from the fact that the spherical radiation is blocked by the surface and forced to radiate hemispherically The properties of half space loading are well known and unsurprising: 6 dB of addition in the low frequencies. What is not as well understood is what happens above the subwoofer range. There is no mystery to us. The surface performs the function of adding an element in a coupled line source array format as discussed earlier. The results over frequency are shown in Fig. 2.91. The panels contain two sets of responses. The colored response is that of the speaker as affected by the surface. The grayed response is one half of the analogous response of an actual coupled line source array. The responses are indeed mirror versions of each other, as we would expect. The analogous nature greatly simplifies the discussion. The crossover class and summation zone progressions move in front and back and side to side in the same manner as discussed for the coupled line source.

Coupled Outward Angle Surface

Next we find ourselves with a speaker coupled to a surface but pointing outward. This is a surface version of the coupled point source and is shown in Fig. 2.92. The splay angle from the surface is equal to the splay angle from center to each of the speakers in a coupled point source array. Once again the analogous features prevail and the progressions will move along as in the speaker/speaker versions. We will stop here for a moment to note that we will not be detailing the response of the coupled inward surface. To face a speaker downward into a coupled surface could only be defended as some form of special effects processing. We will leave that matter to the artists.

Uncoupled Parallel Surface(s)

We quickly move on to the uncoupled surfaces. The first of these is the parallel surface as seen in Fig. 2.93. The

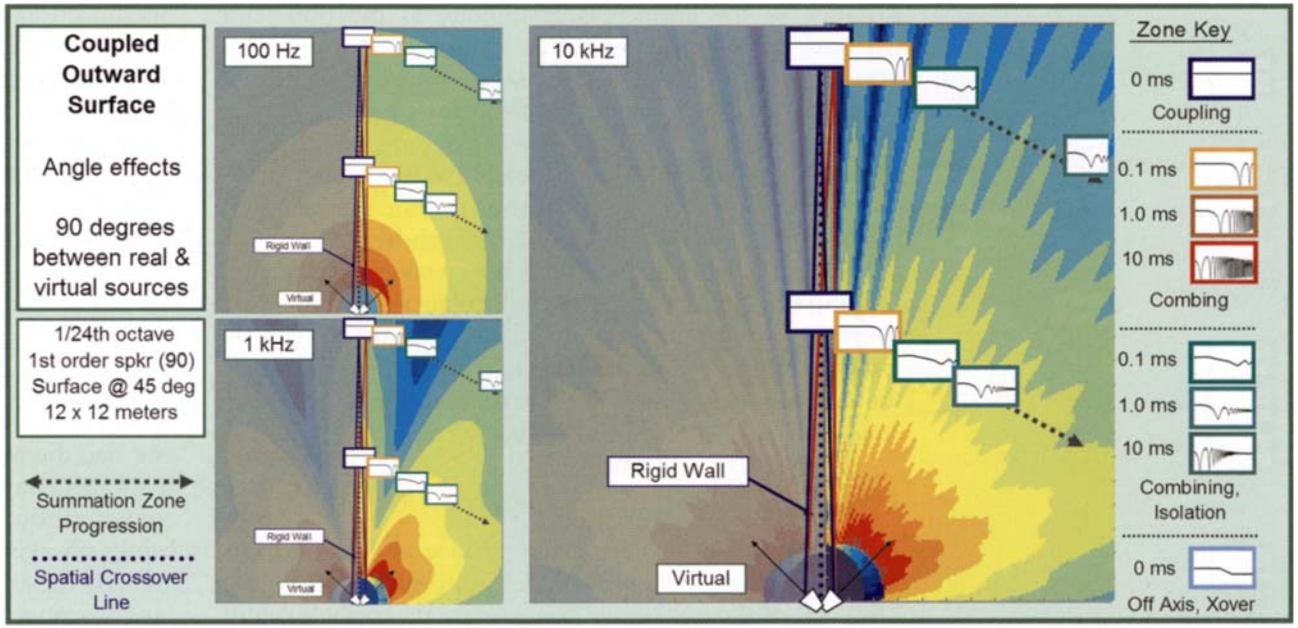


Figure 2.92 Summation zone progression factors for the coupled outward surface

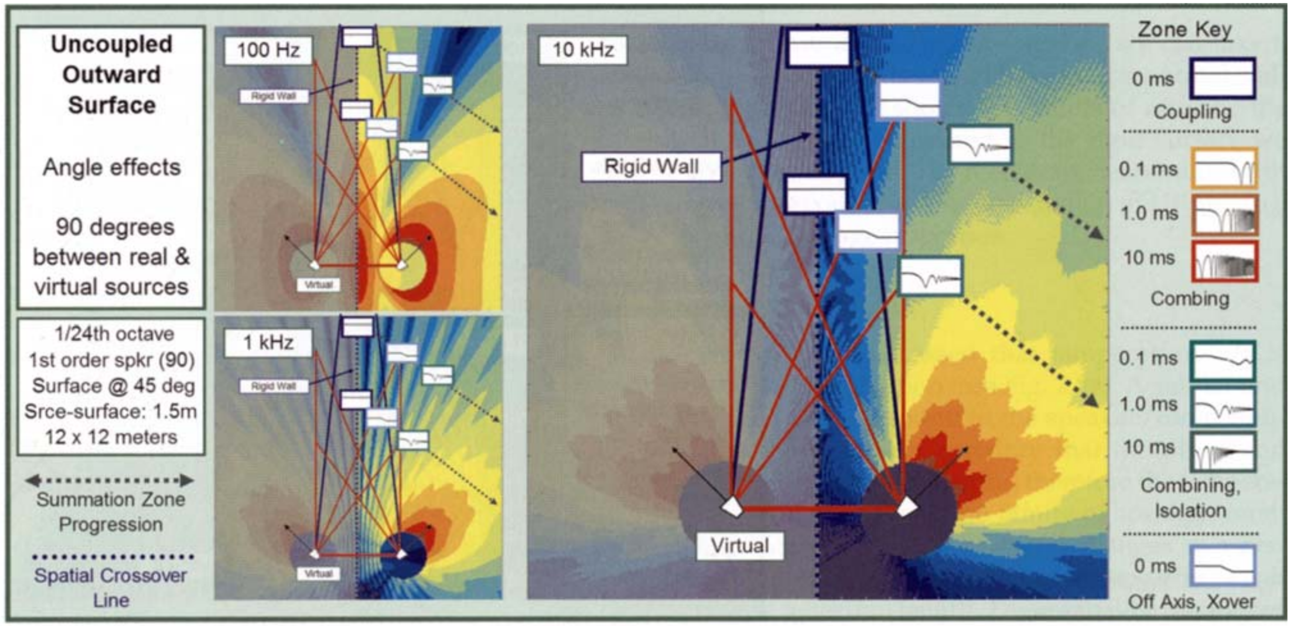


Figure 2.93 Summation zone progression factors for the uncoupled parallel surface

distance to the surface is analogous to the mid-point between speaker elements in an uncoupled line source array, e.g. 3 m to the wall is equal to 6 m from element to element (3 m each to center) in the array model. Just as in the array model, the response will degrade steadily over distance. The distance to the walls has the same range-shortening effects on the speaker as did the displacement between array elements. The example shown here shows a symmetrical surface pair, with the speaker in the center. Moving the speaker off-centre will create an asymmetrical spacing. The near side will have a shorter range than the far side. Note also the presence of the parallel pyramid in the LF response, which will inevitably appear when the conditions are centered and symmetrical.

Uncoupled Outward Angled Surface

A speaker uncoupled from the surface and angled outward is a match to the uncoupled point source array (Fig. 2.94). All of the analogous considerations apply but with one additional note of interest. We previously observed that if the splay angle and element angle were equal then the uncoupled point source array would gap for eternity. This can now be used for practical advantage, giving us a means to skim coverage along seating areas without overlapping into the nearby surface. The outward angle surface configuration has the long-range stability which was found earlier in the point source arrays.

Uncoupled Inward Angled Surface

Previously we found that the uncoupled point destination array created the most rapid summation zone transitions and had the highest amount of combing zone summation. It should come as no surprise that its analogous surface configuration will do the same. The inward wall (Fig. 2.95) brings on-axis energy back into the heart of our coverage area. As the angle of the inward surface increases the need for absorption (or demolition) rises proportionally.

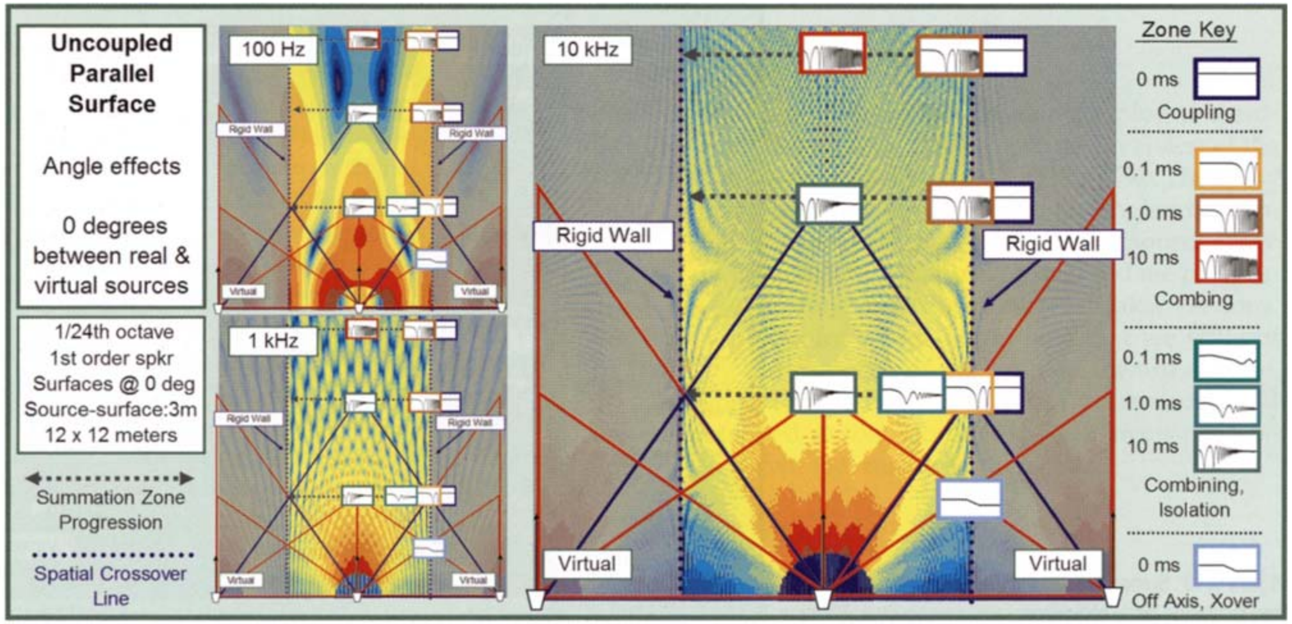


Figure 2.94 Summation zone progression factors for the uncoupled outward surface

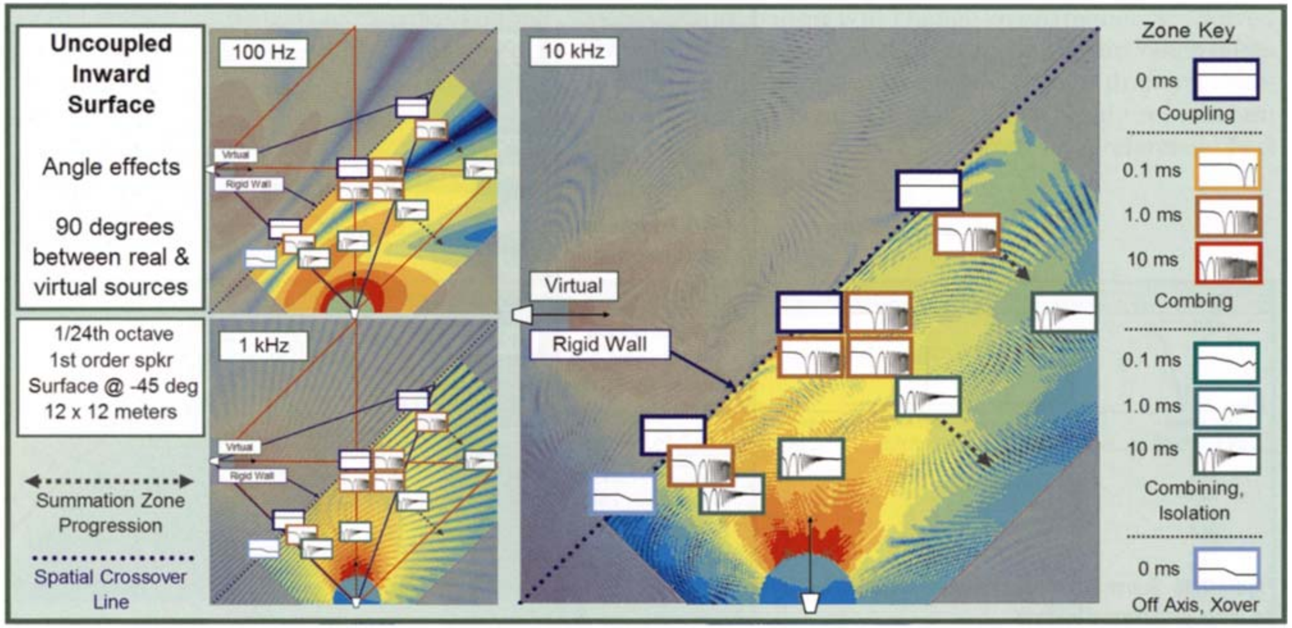


Figure **2.95** Summation zone progression factors for the uncoupled inward surface

The normal surface is the most extreme case of speaker/room summation. The surface line is at 90 degrees to the speaker axis (a "normal angle" in geometric terms) which puts the virtual source at 180 degrees to our speaker axis. This configuration is shown in Fig. 2.96 and is analogous to the 180 degree point destination array. The practical world version of this is the rear wall. The low-frequency coupling of the rear wall is well known to most engineers. If we are looking for a low-frequency boost all we have to do is lean against the back wall. Unfortunately our mix position is likely to be a few meters forward of that, which puts us right in the heart of the combing zone. There is one heartening aspect of our anatomy that helps us to reduce the perceived effects of this summation form: the front/back rejection of our ears. The measured response of combing zone interaction that originates from behind us will be perceived as less audible than a comparable degree of combing from a forward-originated summation. That being said, this type of speaker/room summation needs absorption more than any other.

Absorption Effects

As long as we are discussing absorption, we can take a look at the effects that this would have on the normal wall summation. Figure 2.97 shows the results of acoustic tile placed on the surface compared to the rigid surface we have seen up to this point. The results are heartening, with reductions in the combing zone summation in HF and to a lesser extent mid-range responses.

Environmental Effects

We have previously discussed that temperature affects speaker/speaker summation (see Fig. 2.89). A similar and more profound effect is found in the speaker/room summation model. It is a near certainty that the direct and reflected paths are different lengths; therefore the temperature change will affect the entire family of speaker/room summations. Since the speed of sound changes as a percentage of temperature change, the summation effects will differ for every reflection length. The comb filter frequency

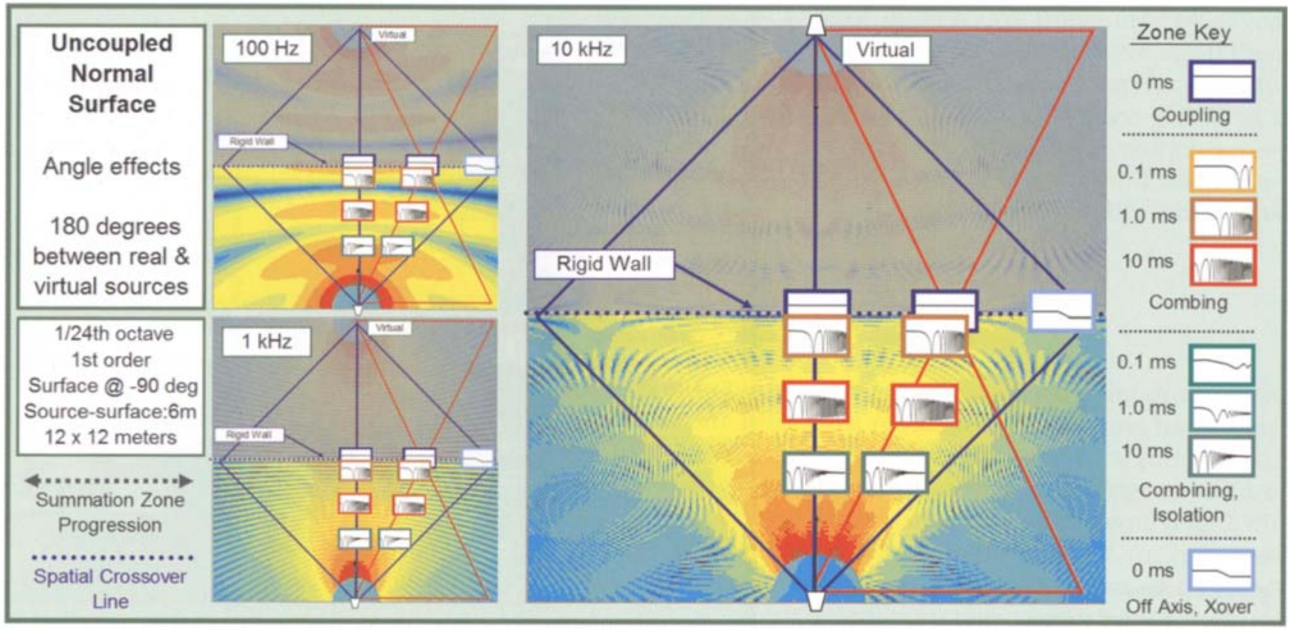


Figure **2.96** Summation zone progression factors for the uncoupled normal surface

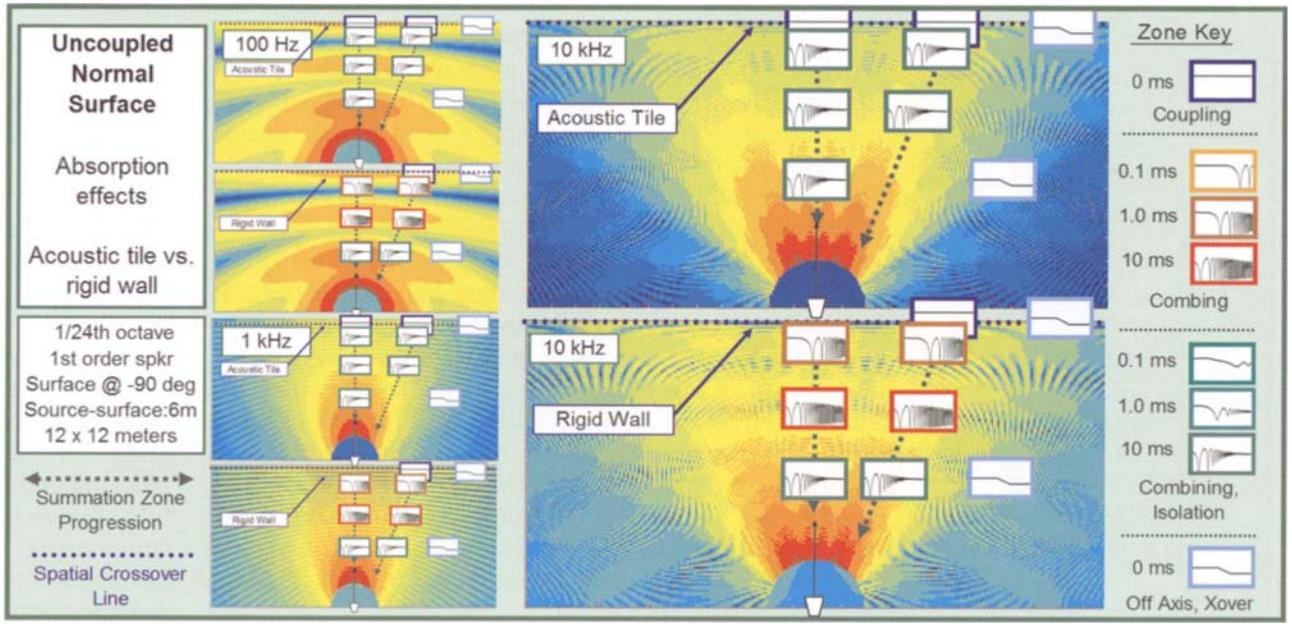


Figure 2.97 Summation zone progression factors for the uncoupled normal surface, absorption effects

is governed by the *absolute* value (ms) of the relative time offset, not the *percentage* of relative time offset. Consider the fact that a 1 per cent change in sound speed causes a 1 ms change over a 100 ms direct sound transit time, while an echo that arrives 100 ms later (a total path of 200 ms) will change by 2 ms. The result is that the entire family of room/speaker summation will change over frequency. Viewed individually the change appears insignificant. When taken as a whole, the composite experience can be dramatic. Naturally the effect is most apparent in a highly reverberant space with large changes in temperature. A reference chart is provided in Fig. 2.98.

Temperature Effects on Speaker/Room Summation

Temperature Degrees (C)	Temperature Degrees (F)	Temperature change (%)	Sound speed change (%)	Direct sound arrival (ms)	Reflection arrival (ms)	Time offset (ms)	Comb filter frequency (Hz)
0.0	32	-7.52%	3.91%	25.98	129.89	-103.91	9.62
1.1	34	-7.14%	3.72%	25.93	129.65	-103.72	9.64
2.2	36	-6.77%	3.52%	25.88	129.40	-103.52	9.66
3.3	38	-6.39%	3.32%	25.83	129.15	-103.32	9.68
4.4	40	-6.01%	3.13%	25.78	128.91	-103.13	9.70
5.6	42	-5.64%	2.93%	25.73	128.66	-102.93	9.72
6.7	44	-5.26%	2.74%	25.69	128.43	-102.74	9.73
7.8	46	-4.89%	2.54%	25.64	128.18	-102.54	9.75
8.9	48	-4.51%	2.35%	25.59	127.94	-102.35	9.77
10.0	50	-4.13%	2.15%	25.54	127.69	-102.15	9.79
11.1	52	-3.76%	1.96%	25.49	127.45	-101.96	9.81
12.2	54	-3.38%	1.76%	25.44	127.20	-101.76	9.83
13.3	56	-3.00%	1.56%	25.39	126.95	-101.56	9.85
14.4	58	-2.63%	1.37%	25.34	126.71	-101.37	9.86
15.6	60	-2.25%	1.17%	25.29	126.46	-101.17	9.88
16.7	62	-1.87%	0.98%	25.25	126.23	-100.98	9.90
17.8	64	-1.50%	0.78%	25.20	125.98	-100.78	9.92
18.9	66	-1.12%	0.59%	25.15	125.74	-100.59	9.94
20.0	68	-0.75%	0.39%	25.10	125.49	-100.39	9.96
21.1	70	-0.37%	0.20%	25.05	125.25	-100.20	9.98
22.2	72	0.00%	0.00%	25.00	125.00	-100.00	10.00
23.3	74	0.38%	-0.20%	24.95	124.75	-99.80	10.02
24.4	76	0.76%	-0.39%	24.90	124.51	-99.61	10.04
25.6	78	1.14%	-0.59%	24.85	124.26	-99.41	10.06
26.7	80	1.51%	-0.78%	24.81	124.03	-99.22	10.08
27.8	82	1.89%	-0.98%	24.76	123.78	-99.02	10.10
28.9	84	2.27%	-1.17%	24.71	123.54	-98.83	10.12
30.0	86	2.64%	-1.37%	24.66	123.29	-98.63	10.14
31.1	88	3.02%	-1.56%	24.61	123.05	-98.44	10.16
32.2	90	3.40%	-1.76%	24.56	122.80	-98.24	10.18
33.3	92	3.77%	-1.96%	24.51	122.55	-98.04	10.20
34.4	94	4.15%	-2.15%	24.46	122.31	-97.85	10.22
35.6	96	4.52%	-2.35%	24.41	122.06	-97.65	10.24
36.7	98	4.90%	-2.54%	24.37	121.83	-97.46	10.26
37.8	100	5.28%	-2.74%	24.32	121.58	-97.26	10.28

Figure 2,98 Temperature effects on speaker/room interaction

Reception

reception *n.* ***1.*** *receiving or being received.* ***2.*** *receiving of ideas or impressions into the mind.* ***3.*** *receiving of broadcast signal; efficiency with which these are received*

Concise Oxford Dictionary

Introduction

We have all heard the proverbial question:

If a tree falls in the forest and no one hears it, does it make a sound?

We can rephrase this in audio engineering terms:

If an acoustic transmission is sent without reception, does the audio signal exist?

If there is a concert without an audience, will we get paid?

The human hearing mechanism is a subject worth great volumes in its own right. Research is ongoing and subject to debate. The study of reception includes both objective and subjective aspects. The objective side includes the mechanisms of reception, both the human hearing system and that of inanimate microphones. The subjective side includes the *experience* of the human hearing system, i.e. perception. It comes as no surprise that the perception side is the center of ongoing debates, disagreements and controversy. We are fortunately able to limit our scope on this subject to those issues that are relevant to the optimized design. While this limits our exposure to some controversial subjects, it does not relieve us entirely. The scope reduction occurs because we are less concerned with hearing perception in an *absolute* sense than we are in a *relative* sense. For example, we are not concerned about the 3 kHz resonance peak in our ear canal. The existence of this peak is a given for all listeners. No design or optimization decisions will factor this non-linearity into account. Instead we focus on factors that could make one listener's experience different from another's. An example of this would be localization, which is affected by the relative arrival times and level between different sources. We will need to factor arrival time and the location of sound sources into our design process in order to create a similar localization experience for different listeners in the room.

Let's begin with a list of exclusions. These are items that we will not need to explore in any depth since they are either equal for all listeners, or are unrelated to our design and optimization decisions.

- The perception of pitch: our transmission process does not change the frequency of the input signal. How the ear decodes pitch will be a common factor to all listeners.
- The non-linear frequency response of the ear: every sound we hear is distorted by the frequency response peaks in our ears.
- The mechanical nature of the ear: our assumption here is that our listeners share common anatomy. Alien

beings are advised to use the interplanetary edition of this text.

- Aural harmonics and beat frequencies: these exist only in our ears and will not be detectable with our acoustic analyzer (not that we would be able to do anything about them anyway).

This leaves us with a limited number of subjects to explore. In regard to these, we will focus on how they affect design optimization.

Loudness

Our perception of loudness comes from a combination of the sound pressure level and its duration at our ears. Our scope is mostly limited to concerns about how sound from loudspeakers is perceived. Therefore we can dispense with the usual listings of average SPL levels for jackhammers, annealing furnaces and Saturn rocket engines. Our concern is with the sounds people are willing to pay us to reproduce through sound systems. These systems must be capable of sufficient loudness to meet the audience (and band management) expectations for the program material. The foremost of these are music and voice. The maximum loudness range for these signals is genre-dependent, i.e. the fact that a human voice in the natural world ranges to 90 dB SPL does not mean that this level will be sufficient for reproducing Mick Jagger's voice in a stadium. Popular music is here to stay and its sound level requirements are fairly well known. We will do well to focus on the realistic SPL needs of our program material.

The secondary focus will be on the level of the noise floor that limits the dynamic range of our reproduction to something far less than the full capability of our hearing system. The noise floor in a sound reinforcement venue is typically 50 dB or more above the bottom of our hearing threshold. The causes range from HVAC systems, moving lights, and people coughing, hacking and spilling their drinks, among others. Suffice to say that time spent discussing the perception of our sound system at 30 dB SPL will be a waste since it will be drowned out by the noise.

We now have established a range of interest: maximum capability vs. noise floor. The actual dynamic range used by the program material will be a matter of artistic concern, squarely in the domain of the mix engineer. Our job will be to provide a system capable of reaching the maximum level with minimal distortion and having its noise floor under that of the venue. If these conditions are met, there is a chance that the listeners will perceive the program material, not the sound system.

Loudness and dB SPL

It is a simple matter to measure dB SPL, but this does not correspond directly to our experience of loudness. Loudness is an expression of human perception. The ear integrates level over a period of roughly 100 ms (Everest, 1994, p. 48) and therefore our experience of loudness depends upon the signal duration. Short-duration bursts at high levels will be perceived as equal to longer bursts at lower levels. Two different values are commonly given for dB SPL peak and continuous. The peak value represents the maximum pressure regardless of duration, while the continuous represents a value sustained over a period of more than 100 ms.

A continuous sine wave has the lowest possible difference between its peak and continuous level: a crest factor of 3 dB. On the other side of the dynamic scale is a pure impulse which is all peak and no continuity. Pink noise, our most common test signal, has a crest factor of 12 dB. Typical music will have a crest factor in this range. Speech has a variable crest factor. Vowels, closest to sine waves, have low crest factors, while consonants are transients. A system that cannot reproduce speech transients will lose distinction between consonants. Articulation loss due to unclear consonants is the foremost means of measuring speech intelligibility. This is expressed as the percentage of articulation loss of consonants (%ALCONS).

The relevance of this to our sound system is that we will need to ensure that the system has both the continuous and peak SPL capability required to create the perception of loudness that is appropriate for the program material.

The role of limiters for speaker protection was discussed in Chapter **1.** Now we will investigate how they affect perception.

Perspectives: When we first used impulse/phase aligned speakers, people asked if we could make them sound more like brand "x" speakers. They didn't understand that because what they were hearing was more intelligible, the apparent volume (dB SPL) was lower. They were looking for an edge. When measured with an SPL meter most people were off by 6-10 dB.

Don (Dr Don) Pearson

A sound system will maintain its maximum crest factor capability until it reaches the overload point. As overload is reached the system will either undergo peak limiting or clipping. Either of these actions reduces the crest factor. Clipping saws off the peaks of the response bringing them closer to the lower level signals. Peak limiting has a similar effect but with fewer harmonic distortion products. Compression reduces the overall level (not just the peaks) and therefore has less effect upon the crest factor. Compression, however, rarely travels alone, and can usually be found along with limiting and/or clipping. The combination of all three is the end of dynamic range for our system as peak and continuous capability merge to become the same. As crest factor is removed from the system, intelligibility, definition and detail go with it.

An additional experience in our loudness perception is that of danger. When the ear senses that the levels are dangerously high our aural muscle, the tensor tympani, contracts. This tightens the eardrum and reduces its sensitivity. The result is that *perceived* loudness goes down until the level is reduced enough to allow the muscle to relax. If allowed to relax, the full dynamic range of the ear is restored. If not, the range is restricted and our perception of crest factor is compressed. The onset threshold for the tensor tympani more closely approximates a slow compressor than a peak limiter. High peaks can be endured without the muscle tightening, but long durations will bring it on.

Mix engineers would be wise to consider the action of this muscle for themselves and the audience. If the mix is allowed to rise and fall dynamically, it is possible to have high impact moments in the music. If the mix maintains a constant high level, the aural compression will be engaged. Subsequent peaks will not be perceived, even if the system has the headroom to deliver them.

Sound systems operated at the overload point can be caught in a cycle of compounding compression. Once the system has been sent into clipping and limiting it is much more likely to engage the ear's compression system since the continuous levels are high due to the crest factor loss. The loss of dynamics can lead to more level being asked of the system to retrieve the lost perception of excitement. The compression increases in both the transmission and reception systems. Continual engagement of the tensor tympani strains the muscle and results in the experience of aural fatigue. A low-power system run into continual overload is more fatiguing to the ear system, and more likely to be perceived as louder over the long term, than a high-power system that is able to make use of the full crest factor without compression. Unfortunately a high-power system run into gross overload is the worst of all, and this is an all too common outcome. This is, however, an operational matter, strictly under the control of the mix engineer. My musings on this matter are most likely to fall on deaf ears.

The Equal Loudness Contours

The ear does not maintain a constant frequency response over level. When the sound levels at the ear are low, the mid-range area becomes more sensitive than the lows and highs. There exists a family of curves that represent the frequency sensitivity of the ear over level known as the **equal loudness contours.** These curves reconcile the perceived loudness over frequency to the dB SPL in the mid-range. The unit used to express this is **phons.** Audiologists are concerned with phons but do they matter to us? Only if they impact our decisions about equalization for the sound system. Do they? Only if we plan on resetting the system equalizers every time the program level changes. The equal loudness contours are *natural* aspects of our perception, not some artificial ingredient introduced by the sound system. If we attempt to counteract their effects we are introducing an unnatural element into the equation. Maintaining a natural tonal balance that is dynamically appropriate is an operational issue for the mix engineer.

This does, however, put us on notice that in order for listeners at different locations to perceive the same frequency response we will need to match both frequency response *and* level. If the level is not matched, the listeners

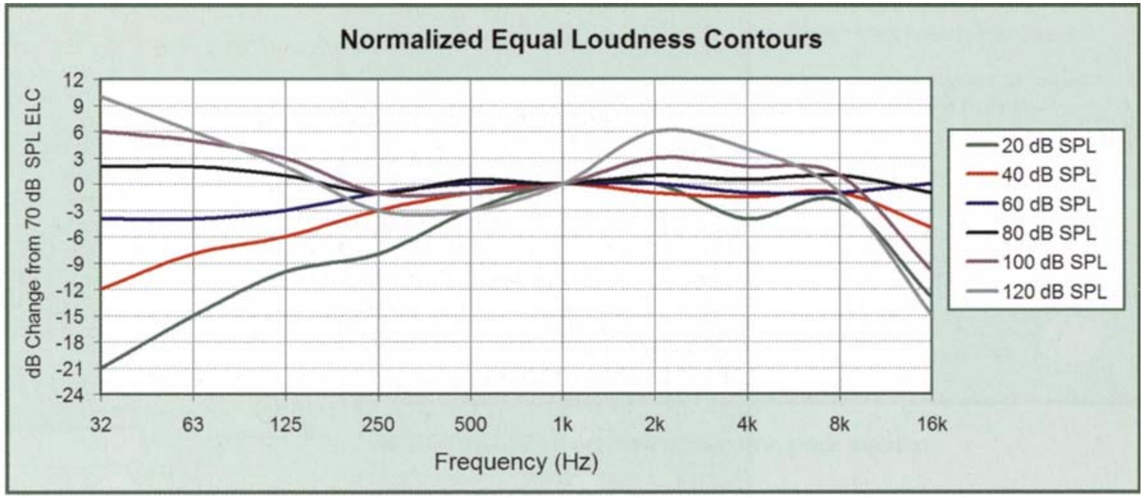


Figure 3.1 Equal loudness contours normalized. The chart shows the perceived frequency response differences over level from the nominal level (after Robinson and Dadson, 1956)

will *perceive* a matched frequency response differently due to their ear's non-linearity over level. Are the perceived dynamic frequency response differences enough to warrant concern? If so we will need one equalization curve for the close seats, and another for the rear. Fortunately not. The changes due to the equal loudness contours are very small when looked at in the context of level distribution over the room. A change of 20 dB in level (from 80 to 100 dB SPL) would change the 30 Hz response relative to 1 kHz by just 3 dB. If we have seats in the auditorium that are 20 dB louder than other seats we have got problems much more serious than the equal loudness contours!

Localization

Introduction

When I hear my name called I turn my head around in the direction that the prompt came from. How did I know? Auditory localization. My personal sonar. Humans and most other animals have a location decoder built into our ear anatomy and brain that guides us in the direction of a sound source.

Auditory localization can be a matter of life and death. Imagine for a moment what would happen if the sound of a nearby tiger appeared opposite to its actual location. We would be lunch. What we have in this example is a sonic image distortion, i.e. the visual, and in this case physical, source of the sound does not match the perceived sonic location. This type of ventriloquism rarely happens in the natural world, and when it does we find the experience disquieting. One example is the sound of a distant jet airplane that appears far behind the visual image. Another effect can be found in spaces with concave surfaces or complex corners where focused concentrations of reflections create a disembodied sound source. In most cases the sound image is right where we see it. This is immensely comforting, allowing us to hear the source without questioning its authenticity. In the world of sound reinforcement we will be asked to perform sonic sleight of hand as standard fare. In many cases our role as sound engineers is to make the sound appear to be originating from actors on stage whose sonic contribution is completely overpowered by our speakers.

Sonic Image

Our ability to steer the sound image to the desired location will depend upon speaker placement and relative levels. The extent to which the perceived sound image differs from

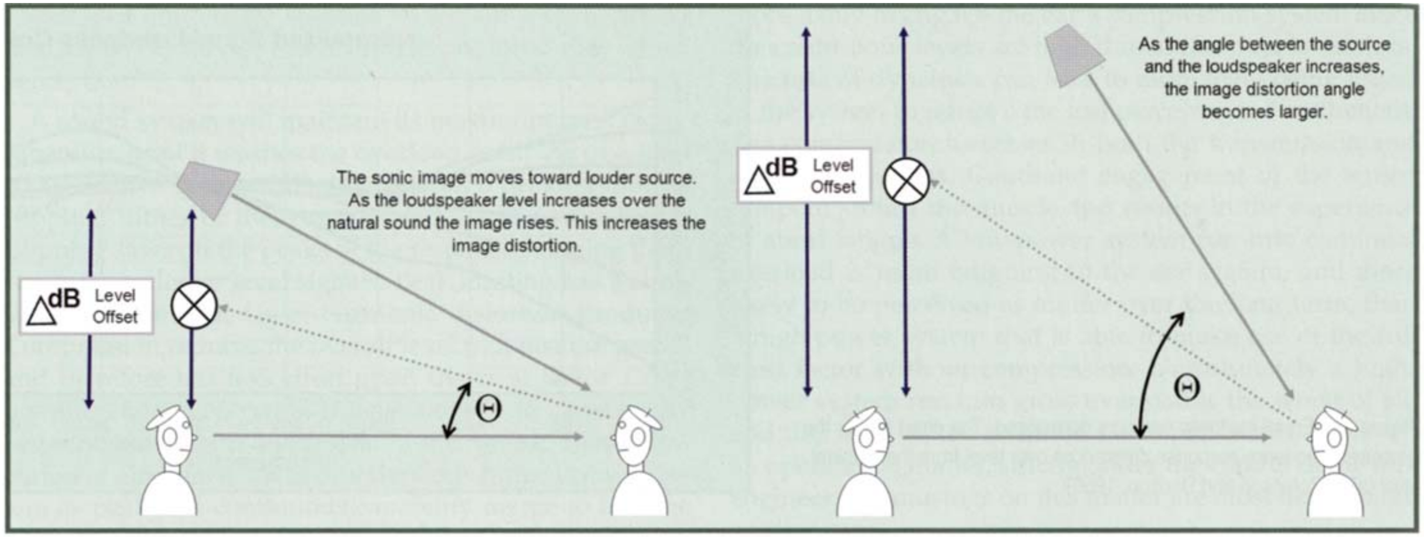


Figure 3.2 Vertical sonic image

the intended sound source is the **sonic image** distortion. This factor is roughly quantified in degrees (the angle between the perceived and intended positions) and less objectively by intensity of effect (whether the image location is perceived as a single point or as spread out over an area). In most cases the intended source corresponds to an actual sound source, such as an actor or musician on stage. The speakers provide sound reinforcement and our intent will be to link the perceived sound location to the visually perceived source. The localization is aided by the arrival of the "natural" sound coming from the stage. If the natural sound arrives first or is significantly louder than the speaker sound, we will have a strong tendency to localize to the natural source. If localization were our solitary concern we would ask the musicians to help us by maximizing their stage volume. The negative implications of this being obvious, we will instead focus on how to minimize sonic image distortion in those cases where the stage levels are not overwhelming.

In some cases there will be no appreciable levels of "natural" sound, e.g. a cinema system. Such circumstances have become common in amplified sound applications as well. Many bands have implemented in-ear monitors which have eliminated stage sound almost entirely. In musical theater the head-mounted microphones have given actors the liberty to sing at a whisper level.

How can we minimize sonic image distortion? Success in this endeavor will require us to understand how the ear's natural localization mechanism functions, so that we can learn how to fool it. The driving forces behind localization are relative level and time, which are the most recurring themes in this book. Our personal sonar systems have independent vertical and horizontal mechanisms, which will be individually discussed.

Vertical Localization

Vincent Van Gogh would have had a difficult time telling whether a sound was coming from above or below him. Why? You may recall that the famously disturbed painter cut off his outer ear. This act would not make him deaf in that ear, but it would remove his vertical localization mechanism, the outer ear or **pinna.** The contours of the pinna create a series of reflections that steer the sound into the inner ear. These reflections cause coloration of the sound as they make their way into the inner ear. The tonal

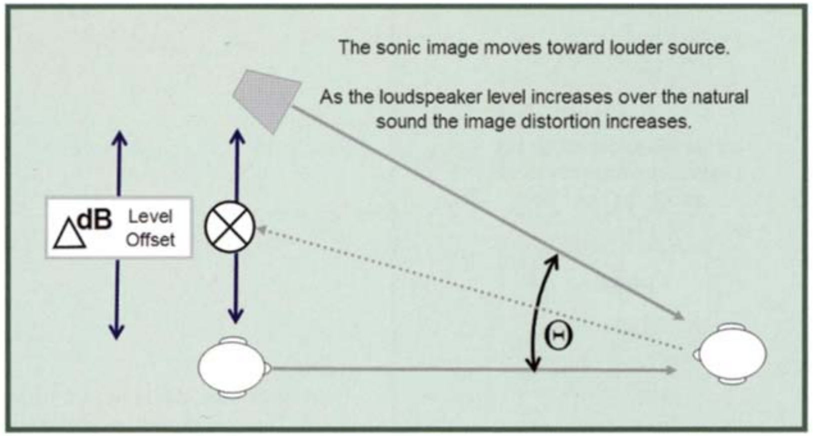


Figure 3.3 Level-offset effects on horizontal image distortion

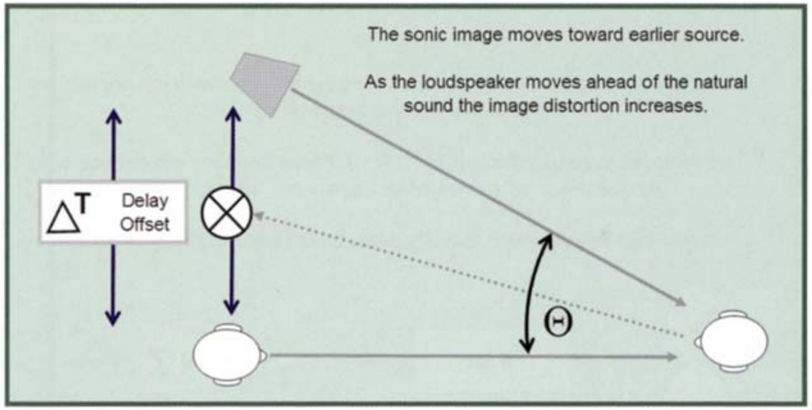


Figure **3.4** Time-offset effects on horizontal sonic image distortion

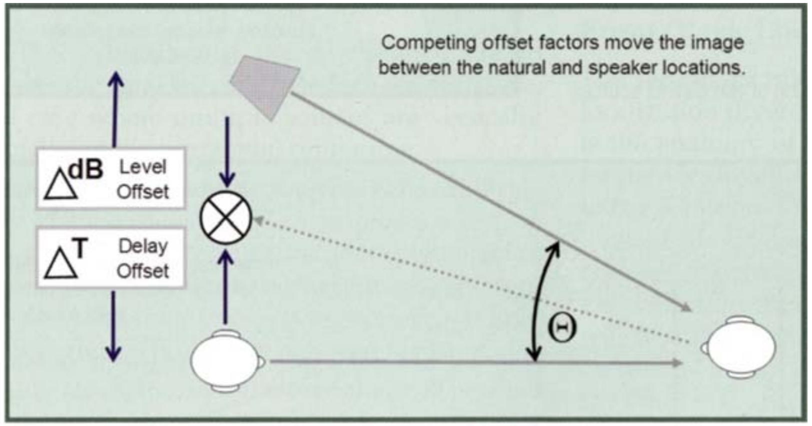


Figure 3.5 Competing time and level offsets can vary the sonic image position

signature of these reflections is encoded in the brain as a vertical map that provides us with our sense of vertical sonic image (Everest, 1994, p. 54). It would seem that the ear/eyes/brain system has learned over time to correlate the visual image and the sound image and that this coloration is mapped into memory as an auditory overlay to our hearing decoder. This response, the head-related transfer function (HRTF), operates independently for each ear. It stands to reason that the restoration of Van Gogh's vertical localization would have required both a prosthetic ear and the time to reset his memory for the new HRTF.

The vertical position of an isolated sound source can be localized with ease and high accuracy. When the ear is stimulated by multiple arrivals of matched signals the process of localization becomes more complex. If the signals are sufficiently close in time, the signals will mix

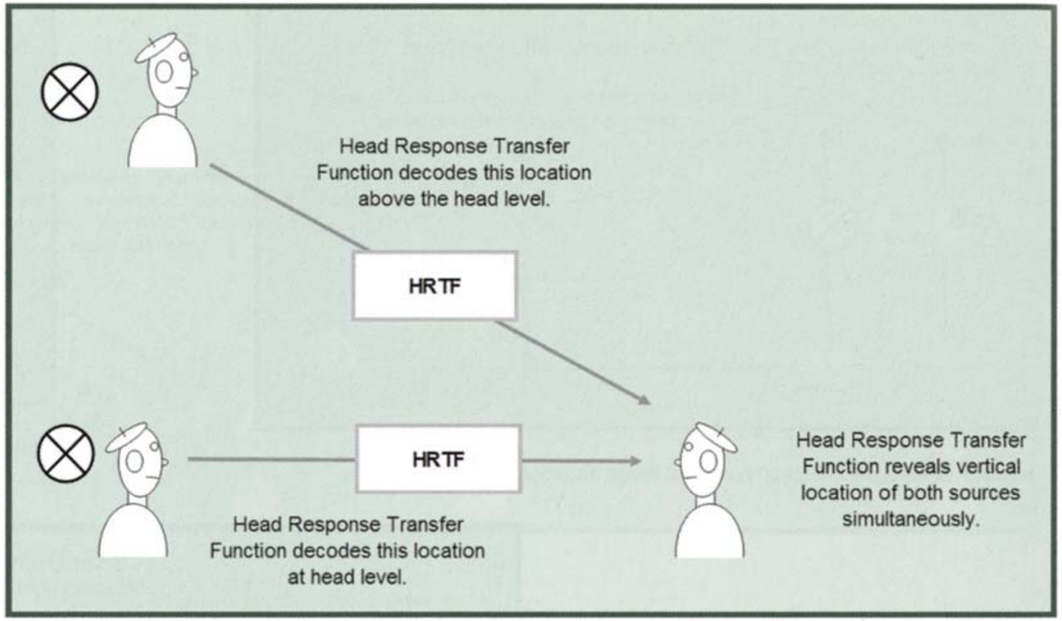


Figure 3.6 Vertical localization of two distinct sources

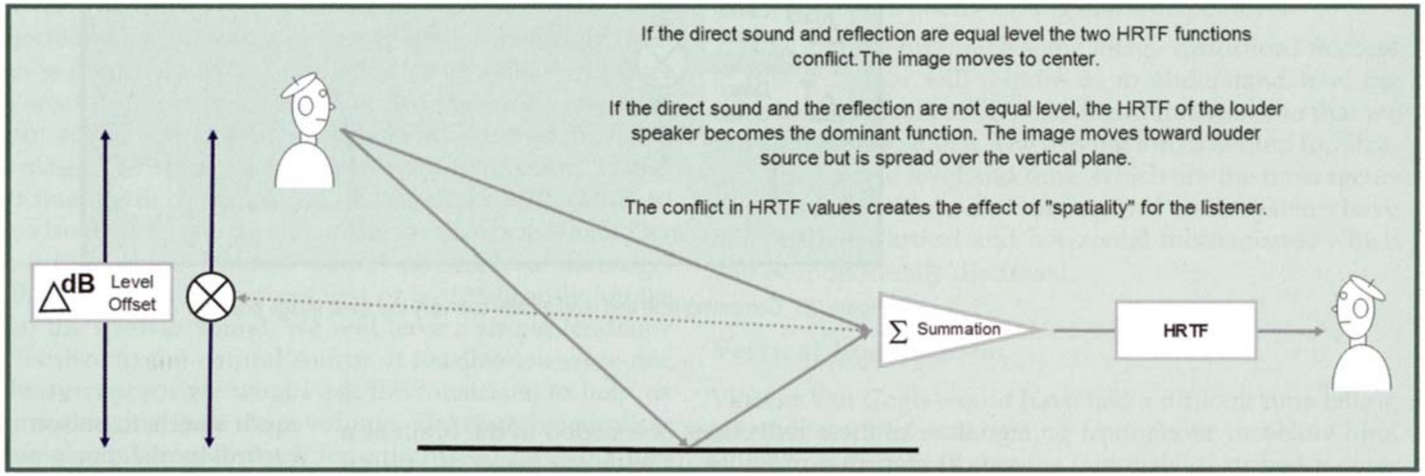


Figure 3.7 Vertical localization in the presence of reflections

in the ear and the HRTF will be modified as a multiple source summation. Each ear independently determines the vertical localization. This contrasts to the horizontal localization which is based upon the *difference* between the ears (Everest, 1994, pp. 51-53). As a result sources that are displaced horizontally and vertically will have conflicting vertical localizations along with the differential horizontal localization. We will discuss this later. For the moment we

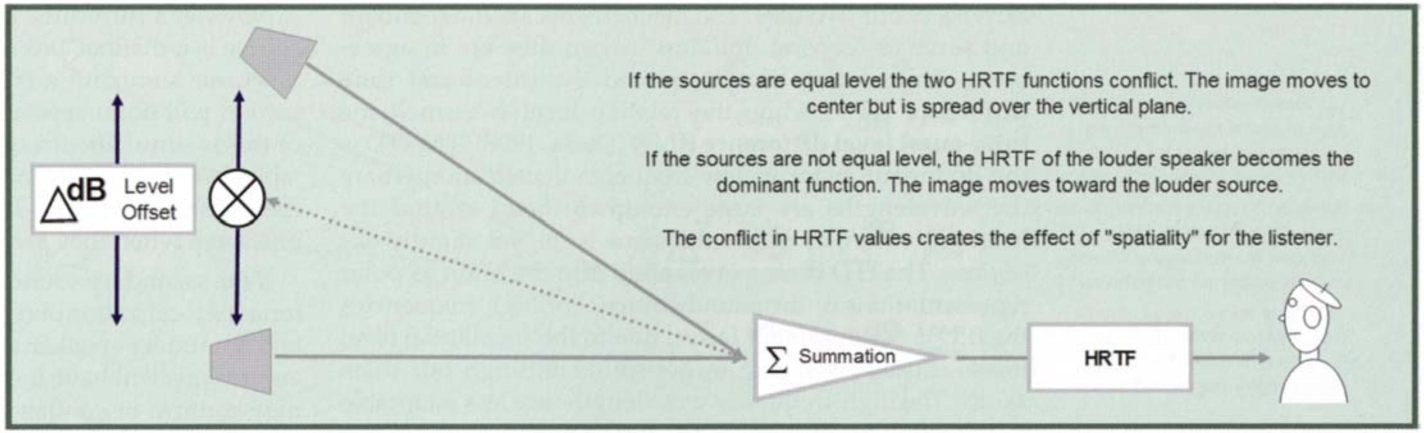


Figure 3.8 Vertical localization in the presence of multiple speakers

Perspectives: Learning how one listens is important. I have found that I can talk myself into hearing almost anything if I am not careful. When optimizing a system I make it a point to not listen to the raw system so it won't disrupt my aural memory. It is only after I have made a first pass on a system that I will listen to something I know very well. I have about three CDs I use to evaluate a system. I have been using them for years, and I get lots of complaints from people about how often they have heard them, but I don't care, and I tell them so! I know the albums so well that after listening to almost any track for a few seconds I know where I am and have a good idea as to where I need to go. They save me and my clients a lot of time. Use nothing less than CD quality, and keep the number of items small.

Alexander Yuill-Thornton II (Thorny)

will focus on the vertical localization where there are no conflicts between the ears. We can consider each ear alone, or the special case where multiple sources are vertically spread but exactly on the horizontal center line.

If two vertically displaced sources arrive at equal levels the conflicting HRTF values reach a compromise central localization between the sources. This is not experienced as a pinpoint sound source as would occur with a single central source. Instead the sonic image is experienced as spatially spread over the vertical plane between the sources.

If the levels are offset, the dominant speaker will become the home of the sonic image. This would be due to its having a recognizable HRTF for the stronger arrival. If the times are offset, there does not appear to be a mechanism for discerning which arrival came first. Early reflections will usually fall into this category. As long as the direct sound is dominant over the reflection the image will favor the direct arrival. Only when the signals fall sufficiently out of time are they heard as distinct sounds, each having separate source locations.

The dominance of level over time contrasts with the horizontal mechanism, which uses a dual-channel system for localization. In the two-channel horizontal system, arrival time and relative level is the dominant player.

Front/Back Localization

The secondary role of the pinna is to provide front/back localization (Everest, 1994, p. 53). The mechanism for this is the anatomy of the pinna which provides some high-frequency directionality and adds another layer of HRTF to the equation. The front/back localization is of minimal interest to our case, although it is worth noting the discrepancies between the behavior of microphones and the ear. Omnidirectional microphones have less front/back rejection than the ear, while cardioid microphones have far more. Neither is a perfect modeling of the ear, although the omni is far closer.

Horizontal Localization

The horizontal mechanism is more sensitive than the vertical. The spaced placement of our ears on opposite sides of our head provides an automatic spatial sensor. Any position on the horizon can be found by triangulation, i.e. the sound is heard in two locations and the source is detected by the difference in their arrivals. This mechanism is known as **binaural localization** and will be of great interest to our application.

The perceived horizontal sonic image location depends upon the both the time and level differences between the arrivals at our two ears. The mechanisms are independent and serve as "second opinions" when they are in agreement. The relative time is termed the **inter-aural time difference** (ITD) while the relative level is termed the **inter-aural level difference** (ILD) (Duda, 1998). The ITD is the dominant factor in low-frequency localization where the wavelengths are large enough to bend around the head and arrive at nearly the same level, yet slightly out of time. The ITD timing is mapped into the brain as polar representations of the sound source. For high frequencies the ILD is the dominant factor, due to the substantial level losses caused by trying to get sound through our thick skulls. The high-frequency wavelengths are less adaptable to refracting around the head, and therefore the level differences are substantial enough to provide the necessary detection. Either of these factors alone are enough to create the localization perception. When listening to a single source, the two parameters will always track together, giving secondary confirmation to the localization data.

If a single sound source is located on the horizontal center line, the ILD and ITD values are both zero. The brain decodes its location exactly where we see it; front and center. If the sound source is moved off the horizontal center it will arrive earlier and at a higher level at the closer ear. In this instance, both the ITD and ILD values confirm the new location.

Now let's introduce a secondary sound source. If the source is a distinct and separate signal (one that does not meet our summation criteria from Chapter 2), the localization will occur separately for each source. An example of this is simultaneous conversations around a conference table. We are able to localize each of the participants by their distinct ITD and ILD values as we would a single one, even when they are speaking at the same time.

If the secondary sound source meets our summation criteria the localization becomes more complex; e.g. reflections and secondary speaker arrivals. In such a case the secondary arrivals will have their own ITD and ILD values which may confirm or conflict with the primary source. We will first consider the earliest reflection arrivals. Provided that we are in the path of the direct sound, there is a predictable relationship between the direct sound and any singular reflected sound. The direct sound will arrive earlier and be higher in level. Our localization system has learned to expect this and does not become confused by the addition of the echo. We continue to perceive the sound source at the location of the direct signal, albeit with a modified tonal quality and sense of "spaciousness," i.e. the sense that sound is coming from a general area rather than from a single point (Everest, 1994, pp. 295-301). The ITD and ILD of the direct sound are the strongest factors since the direct sound is both earliest and loudest. This relationship

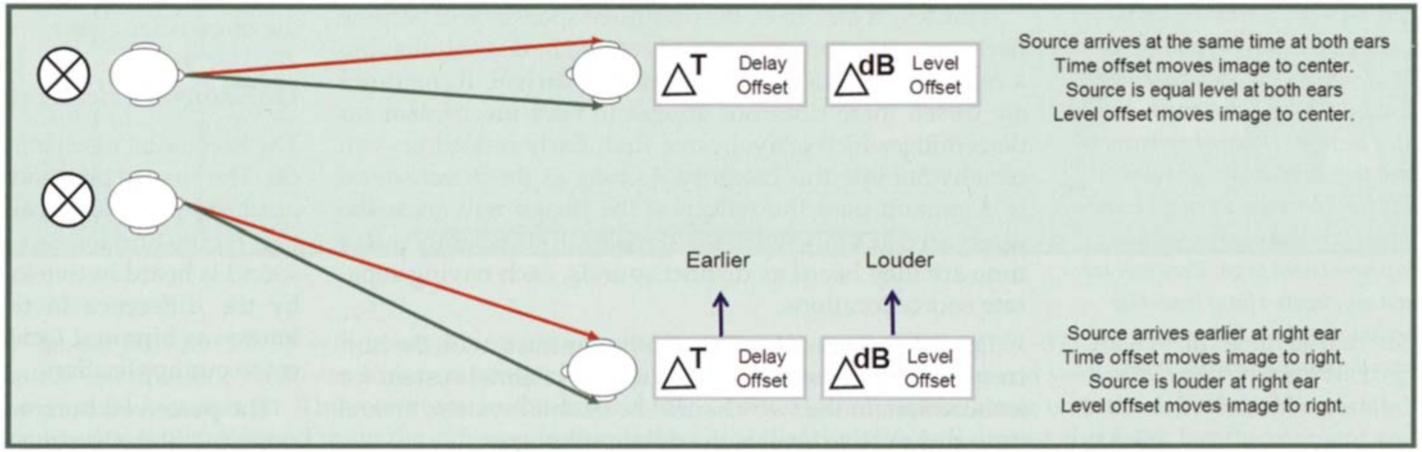


Figure 3.9 Horizontal localization of a single source, on and off center

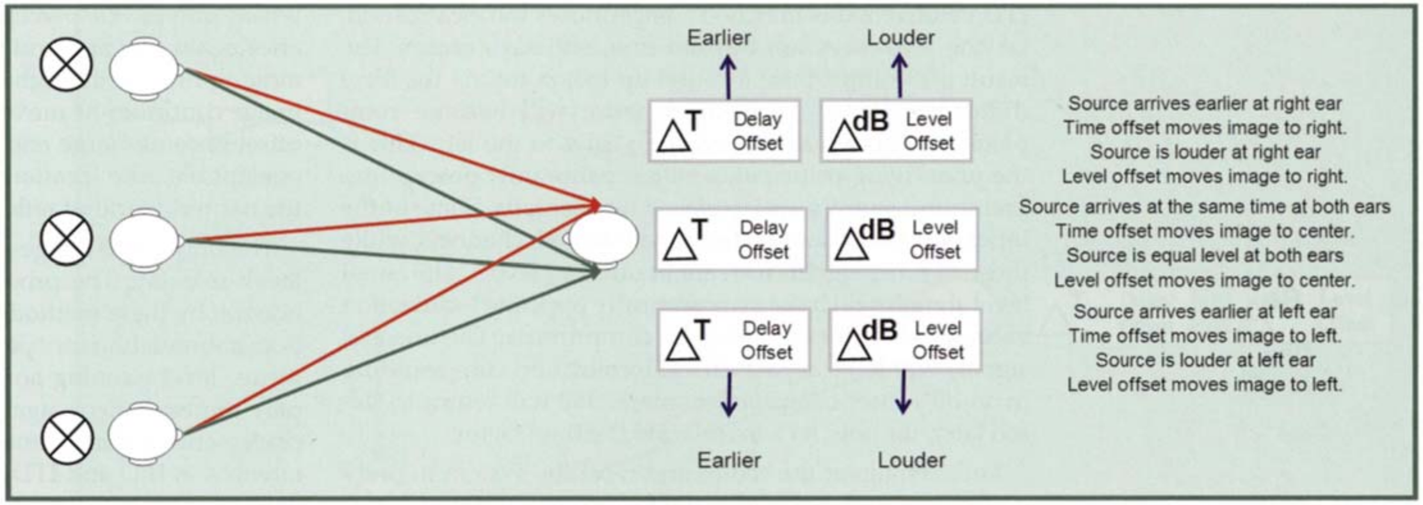


Figure 3.10 Horizontal localization c

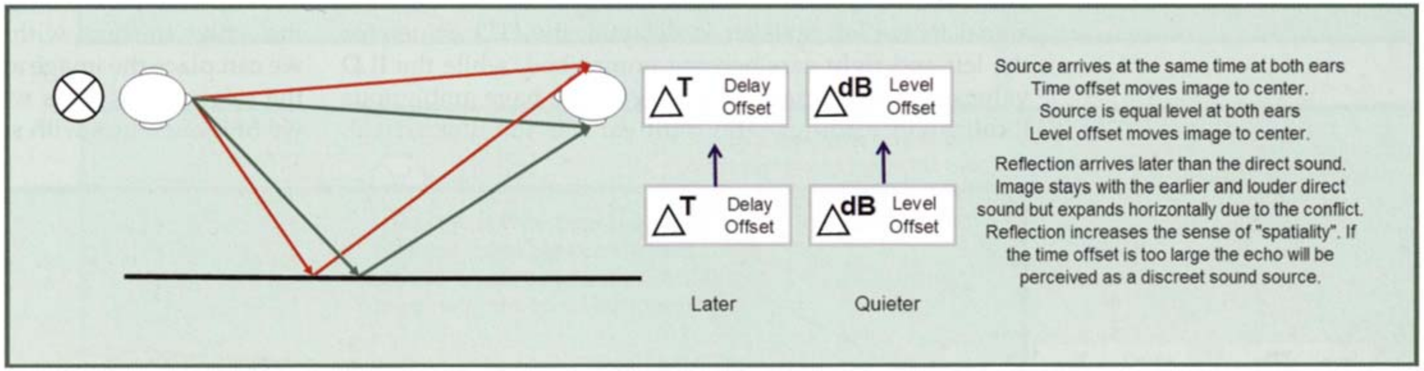


Figure 3.11 Horizontal localization in the presence of a reflection

remains even as the delay time between the direct and reflected sound increases until a point is reached that we perceive the echo as a separately localized source.

When the direct sound is later and quieter than the reflected sound the situation reverses. The ear assumes the early and louder arrival to be the actual location, and the other to be the pretender. This can occur any time there is path obstruction to the direct sound but not to the echo. As sound system designers we should seek to avoid such situations by adding secondary sound sources. Let's consider what happens when that occurs. Let's begin with the most common secondary speaker source: stereo.

Take a seat at the center of a room. The speakers are matched in level, distance, angle and acoustical environment. Where will we perceive the sound image when a matched signal is sent to both speakers? Exactly where there are no speakers: the center. Our sonic ventriloquism has begun! This occurs in spite of the fact that the ITD and ILD values detect speakers off to the sides. Because they are equal and opposite, a compromise value is reached at the center. If the signal is sent to the left speaker at a higher level than the right, the image will shift to the left. The ILD value indicates that the left source is the "direct" sound and the right channel is the "reflection." The fact that the

ITD value remains matched compromises the localization, i.e. the level says left but the time still says center. The result is a compromise location up to a point. As the level difference increases, the dominance will become complete and the localization will go fully to the left. This is the underlying principal of stereo **panoramic perception.** The sound sources are displaced horizontally. Some of the input signals are sent at equal levels to both channels, while those of other signals are sent at unequal levels. The equal level signals will have symmetrically opposite localization values and center the image by compromise. The unequal signals will have asymmetrical localization cues resulting in an off-center compromise image. We will return to stereo later, for now let's investigate the time factor.

Let's remain at the center and reset the system to unity gain between the channels. Instead of offsetting the level between the speakers we will offset the time. When the signal to the left speaker is delayed, the ITD values for the left and right ears become unmatched, while the ILD values remain the same. Once again we have ambiguous localization readings. The right ear has the first arrival, which moves our perception in that direction. However once again, the ambiguity prevents us from immediately moving fully to the right. As the time offset increases, the image continues its move to the right. Eventually the time offset becomes large enough that the left separates into a perceptible echo location on its own, much the same as in the natural world of reflections.

The horizontal image can be moved by offsetting either level or delay. The process of placing images along the horizon by these methods is referred to as "panning". This is an abbreviation for "panoramic" placement and has two forms: level panning and delay panning. It is possible to play the two effects against each other and create image placements that are compromises between substantial differences in ILD and ITD values. This is where our sleight of hand takes full force. It is possible to use level pan to the left and time pan to the right and bring the image back to the center. In short, with control of relative level and time we can place the image at any horizontal location between the two sources. This will come into practical use when we find ourselves with speaker locations that will give us

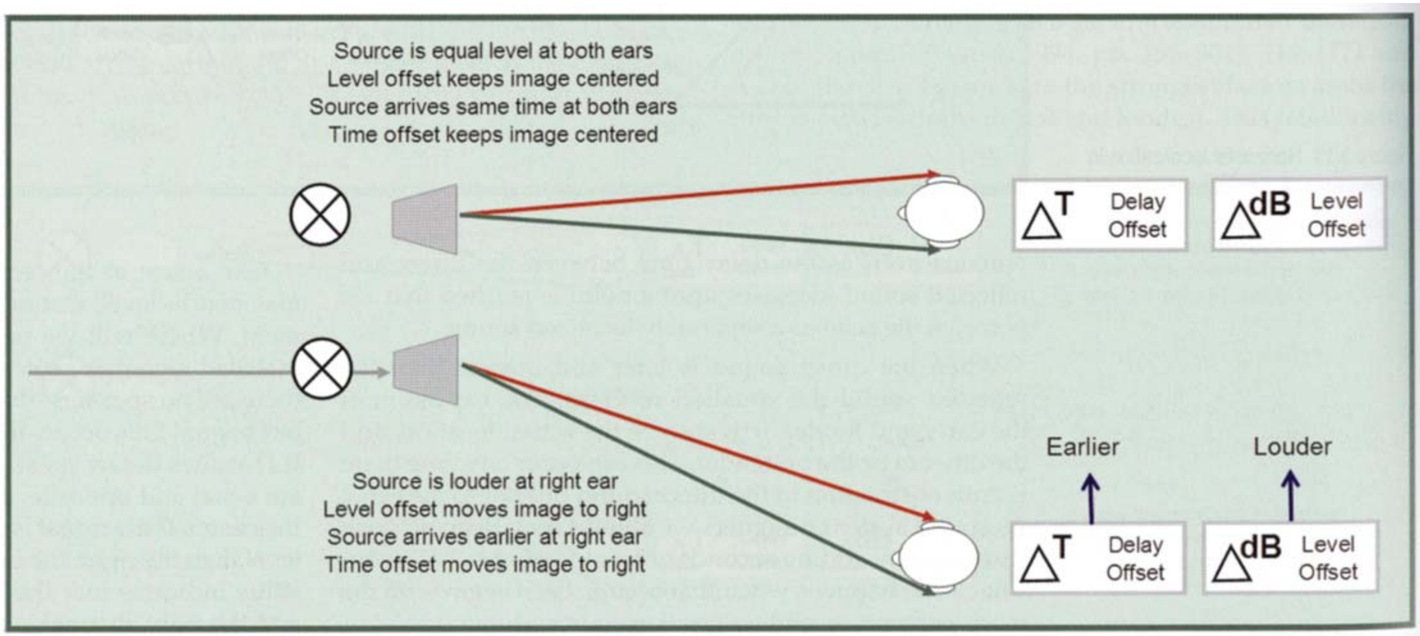


Figure 3.12 Horizontal localization of a single speaker is similar to any natural source

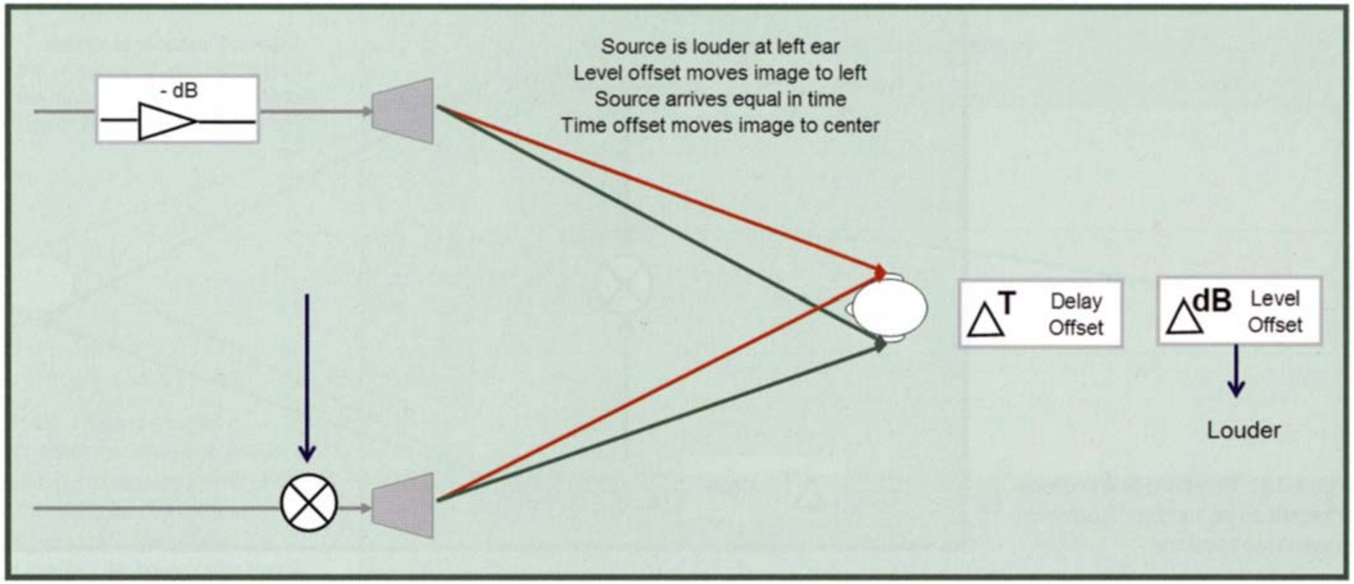


Figure 3.13 Level offset effects on horizontal localization between two speakers

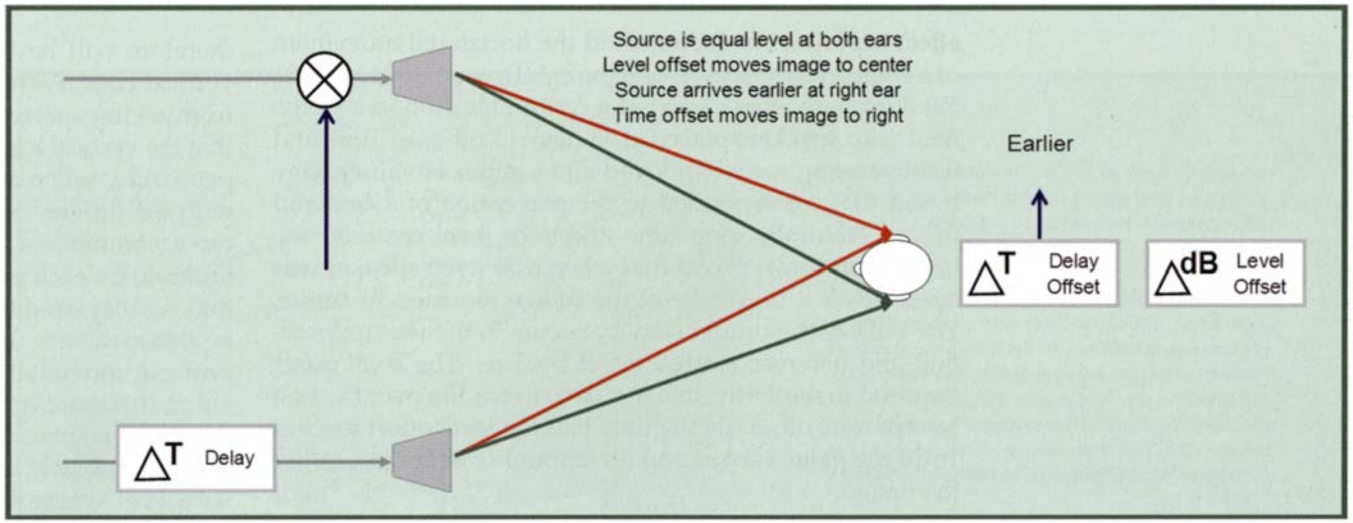


Figure 3.14 Time-offset effects on horizontal localization between two speakers

high levels of sonic image distortion. Speakers at these positions can be level adjusted and delayed so that the image moves away from them on the line toward another speaker at a more favorable angle.

The relationship between these arrivals and our perception of the sound image is known as the **precedence effect.** The research on this is credited primarily to Helmut Haas and many refer to this as the Haas effect. The precedence

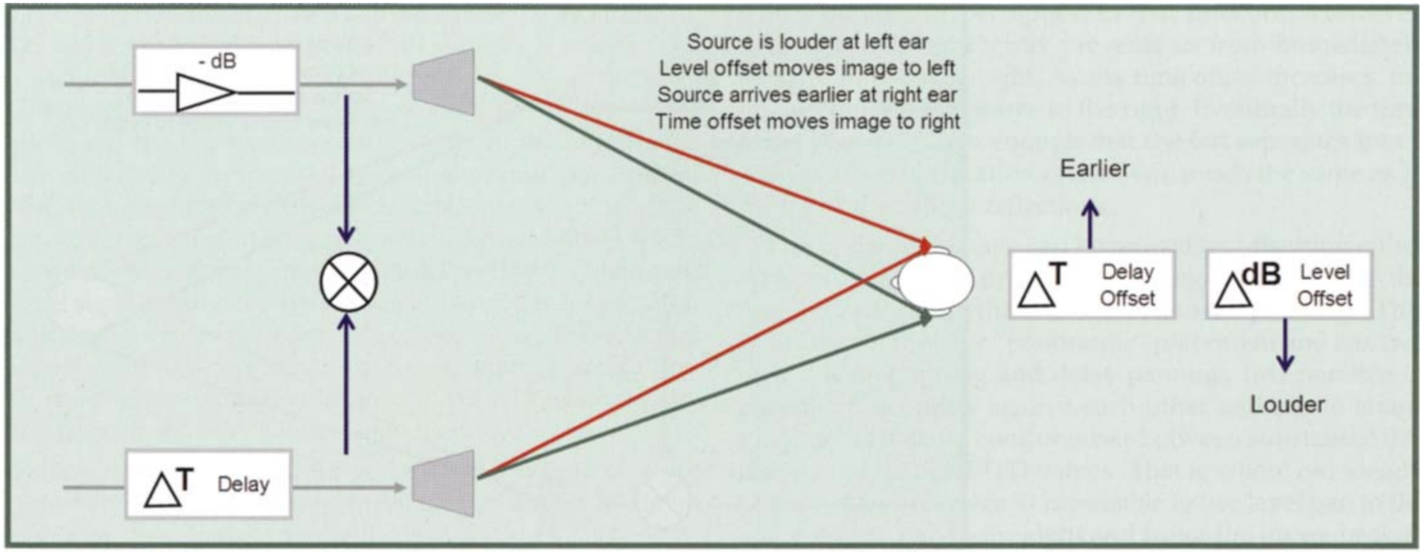


Figure 3.15 Combined time and level offset effects on horizontal localization between two speakers

effect experiments documented the horizontal movement of a sonic by time and level panning (Everest, 1994, pp. 58-59). Listeners were placed at a central location in a stereo field with speakers placed at 45 degrees off-axis. Time and level panning were employed and a series of values were found that corresponded to the perception of a centered image. Naturally zero time and zero level centered the image. But it was found that when a delayed speaker was operated at a higher level the image returned to center. The bulk of the timing activity occurs in the first millisecond and has reached full effect by 5 ms. The level offset required to center the image changes steadily over the first 5 ms of time offset. By the time that the level offset reaches 10 dB the game is over and no amount of delay can center the image.

Image control can be maintained over a limited range of time (5 ms) and level (lO dB) differences between horizontally displaced sources. And that is just the horizontal! What are we to do about the vertical image? Recall that our vertical localization is based on the HRTF and therefore will have little means of sorting out summed vertical cues. Fortunately there is help on the way, albeit from an unexpected source, the horizontal system. Recall that the vertical localization systems for each ear acts independently, whereas the horizontal system is concerned with the difference between the ears. The vertical system can accommodate two vertical values, as long as they are separate for each ear. Therefore, if there are sound sources that are displaced both horizontally and vertically we will be able to discern their locations independently and a compromise horizontal and vertical value is created. Unless we are at the exact horizontal center between two vertically displaced sources we will receive vertical cues from both sources. The only way to perform pure vertical panning is with level, where the level of one source becomes the dominant HRTF signature. If we are hearing the combination of upper left and lower right speakers the precedence effect will compromise the horizontal while the two distinct HRTF values will compromise the vertical. The result will be a sound image where no actual speaker exists in either

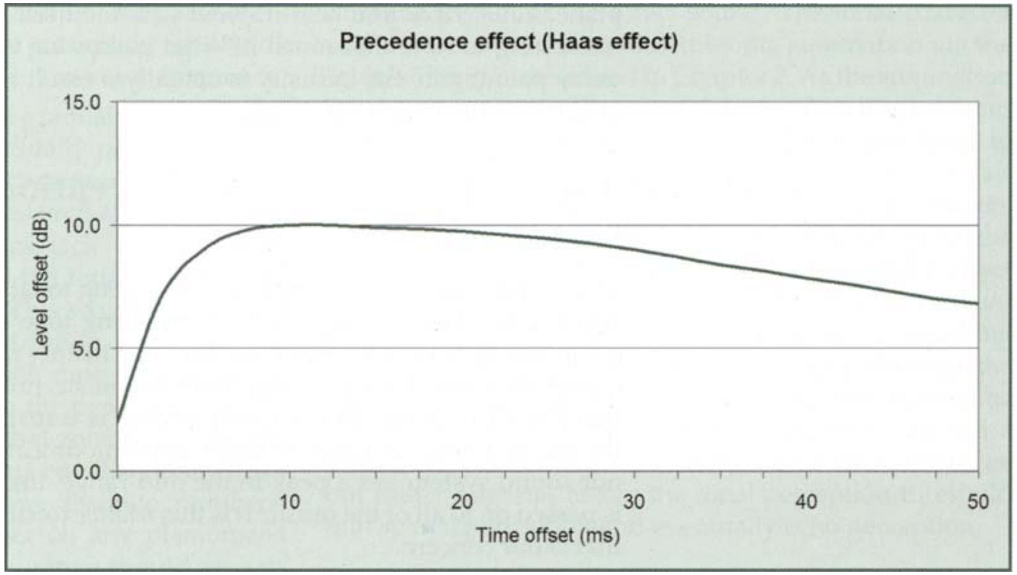


Figure 3.16 The precedence effect. The horizontal sonic image is affected by the time and level offsets between sources. With both offsets at 0 the sonic image appears at the center between the sources. As time-offset increases, the level of the later source must be upwardly compensated in order to maintain a central image. The first 5 ms of time offset are the most critical in the balancing act between time and level. Beyond 5 ms the level offset must be very high to affect the image (after Haas)

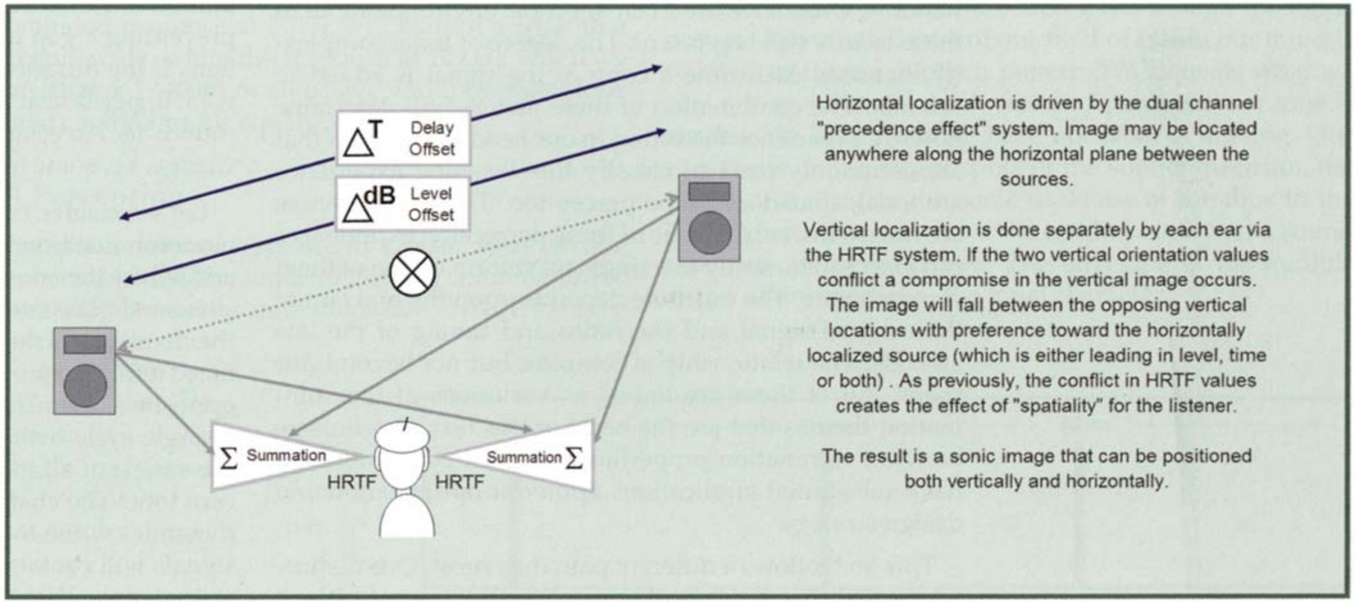


Figure 3.17 Vertical and horizontal localization systems in operation at the same time. Speakers are horizontally separated with asymmetric vertical orientation. The horizontal localization remains binaural. The conflicting vertical HRTF values reach a compromise that follows the dominant horizontal location

plane. Voila! Now you hear it. Now you don't see it. It is interesting to note that much of what passes for vertical delay panning in our industry is actually a result of cues taken from the horizontal plane.

Tonal, Spatial and Echo Perception

Introduction

Music and voice create a constantly changing tonal shape which is tracked by our ears. This changing tone is also compared in our brains against an ongoing history file. We expect the tonal shape to change *with* the music, not independent of it. If the reproduction system is transparent, the music passes through without tonal modification. If our sound system has a peak in the mid-range, that peak is passed on to all of the music. It is this *relative* tonal shape that is our concern.

The tonal character of our listening experience is composed of three parts: the direct sound waveform, copies of the waveform that arrive within the summation duration window, and copies that arrive outside the duration window. Unless we are in an anechoic environment all of these factors will be present. The degree of tonal complexity increases each time a copy of the signal is added to the mix. The combination of these factors will determine how we experience the sound in our heads. The terms that are commonly used to classify this listening experience are tonal, spatial and echo perception. These experiences are not exclusive. All three of these perception experiences can exist concurrently in a single waveform or one of them can dominate. The outcome depends upon the makeup of the original signal and the ratios and timing of the late arrivals. The relationship is complex, but not beyond our scope. All of these are linked as variations of the summation themes that are the heart of this text. The linkage between summation properties and tonal perception will have substantial implications upon our optimization and design strategy.

This text follows a different path than most. One distinction is that we are focused beyond the vocal range. Much of the published findings in these matters are centered on speech perception and intelligibility. We are hard pressed to find statistics for NASTI (noise and snare transmission index) or %ALCONGAS (articulation loss of conga drums) but our perceptions of these and other musical instruments are tremendously important to us. Nothing found in this text relates strictly or decidedly towards the vocal range. We are equal opportunity waveform employers. Another distinction is that we are looking to frame the perception issues in terms that we will be able to translate into action. We are not passive observers. We are sonic battlefield participants. The matter boils down to this: how can we best match what we hear with what we can see in the space, and can read on our analyzer? How can we translate this into the actions of optimized design? A third distinction is that, while other texts emphasize the contrast between tonal, spatial and echo perception, we will be discussing the continuum between them, and their common cause.

Our perceptions lead us to believe that discrete echoes are distinct events, separated by silent space, hence the term "discrete." This is much rarer than we might think, since in most cases the duration of the low-frequency information is sufficient for summation to occur, and thereby preventing a gap in the sound. Duration is a key element here. If the duration is infinite, such as a continuous tone, we will never hear an echo no matter how much glass surrounds us. An echo can only be heard if there is a dynamic change, i.e. some form of transient.

Let's consider two extremes of program material. The monotonous drone music known as Gregorian chant was created in the most echoic acoustical spaces ever made: stone and glass cathedrals. The nearly infinite duration of the chant gave it the maximum tonal perception and immunized it from disturbance by echo perception. Amen. At the opposite extreme is the pure impulse, rising and falling in a single cycle over a wide frequency band. This signal is the easiest of all to discern echoes, and the hardest to discern tone. The chant is maximum duration and minimum dynamics while the impulse is the opposite. Our musical signals will contain rich mixtures of transient and steady-state signals. We will be constantly transitioning between the three perception experiences (tonal, spatial and echo), even in the same seat of the same building. Our goal here is to find a method of discerning these perception thresholds without having to control the music. The key is frequency.

The road to echo perception is a gradual one. It begins at the highest frequencies and gradually peels apart successively larger portions of the spectrum. It is widely published that the dividing mechanism is time offset. The published temporal dividing lines lack perfect agreement. The first perception zone is the tonal fusion zone, which runs for the first 20-30 ms. This would be the area that would most benefit from equalization. From there we enter the area of spaciousness which runs up to between 50 or 60 ms and finally emerges to the perception of discrete echoes (above 60 ms). This final zone would be the area that would least benefit from equalization. These, however, should not be viewed as absolute numbers. Attempts to place a single number on any phenomena that ranges over our full frequency range should give us pause. After all, the time period difference over frequency for our hearing range of 20 Hz to 20 kHz is 1000 to 1! As an example let's consider two combined signals with 30 ms of delay between them, the limit of the tonal fusion zone. The combined response at 30 Hz will feature a peak that is an octave wide, while the response at 12 kHz will have combing that is 1/400th of an octave wide. Do these look like equal candidates for equalization?

Tonal Perception

The perception of tonal response in a system has several layers of complexity. The tonal quality is a combination of the direct sound and the summations that arrive *within* the duration period of the direct sound. The tonal character of the direct sound is modified by the summation via the "comb filtering" described in Chapter 2. As the summation ripple increases, the distortion of the tonal quality becomes increasingly perceptible. The tone will be modified to some extent no matter how late the secondary arrivals are, provided they meet the summation duration criteria. The time offset determines the frequency range where the tonal disturbance is most perceptible. As the time offset increases, the affected frequency range falls. As the tonally apparent wide filters move downward, the combing causes increasingly narrow filters to sweep through the upper regions. Eventually we reach a point where the filtering is so narrow that the ear is unable to discern a tonal quality. The disturbance does not go away, however, but rather migrates across the tonal perception threshold into spatial perception and eventually echo perception.

Tonal Envelope

The room is dark except for a strand of lights. The lights form a recognizable shape, a tree. Even though the lights occupy only a small part of our field of vision our mind is focused on discerning their pattern. We focus on what we see, not on what we don't see. Our perception of tone is similar. We hear what we hear, not what is missing. Our ears focus on the bright spots of the sound spectrum, the peaks, and relegate the sonic darkness of the dips to the background. The pattern of the bright points in the sound spectrum is the **envelope.** The envelope is the audible shape of the spectrum, the tonal character.

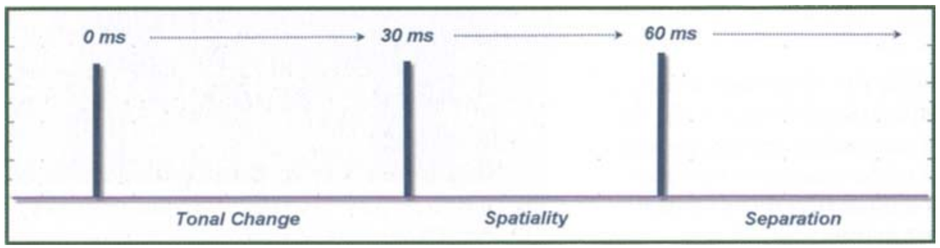


Figure 3.18 Tonal, spatial and echo perception thresholds as traditionally rendered over time offset

In a free field environment the envelope will have the spectral shape of the speaker, and should be free of cancellations. A theoretically ideal speaker envelope would be flat and have no gaps. It is free of tonal character. As summations with room reflections or other speakers are added, the response changes. Peaks and dips are added and the envelope changes. The peaked areas stand out above the median response, and the cancellations move below. Our ears track along the brightest lights of the spectral areas that stand out. A peak in the vicinity of the median response will stand out just as we are able to see a hill on a flat horizon. A median response area in the neighborhood of cancellations will also stand out in a similar fashion. A peak that is bordered by deep cancellations is the most extreme case due to the contrast in levels. In all cases a new tonal character is born. As more summations are added, the response complexity increases. The envelope is reborn into the shape of the leading spectral areas.

The bandwidth of the peaks plays a critical part in the level of tonal perception. The ear has limited tonal resolution. The finite levels of the ear's frequency resolution are extremely high. Very small changes in an oscillator pitch are detectible even to an untrained ear. But discerning pitch and discerning tonal character are different tasks. The frequency resolution to which tonal character is audible is known as **critical** bandwidth (Everest, 1994, pp. 46-47). The implication is that peaks and dips that are narrower than critical bandwidth are not perceived as tonal variations. This does not mean that narrow peaks and dips have no audible effect, but rather that they are perceived as spatial or temporal effects. Recall from Chapter 2 that the bandwidth of these comb filters is related to the wavelength offset between two summed signals. At a given frequency, the percentage bandwidth (in octaves) of the filtering is inversely proportional to the time offset. When the time offset is large enough to narrow the filter below

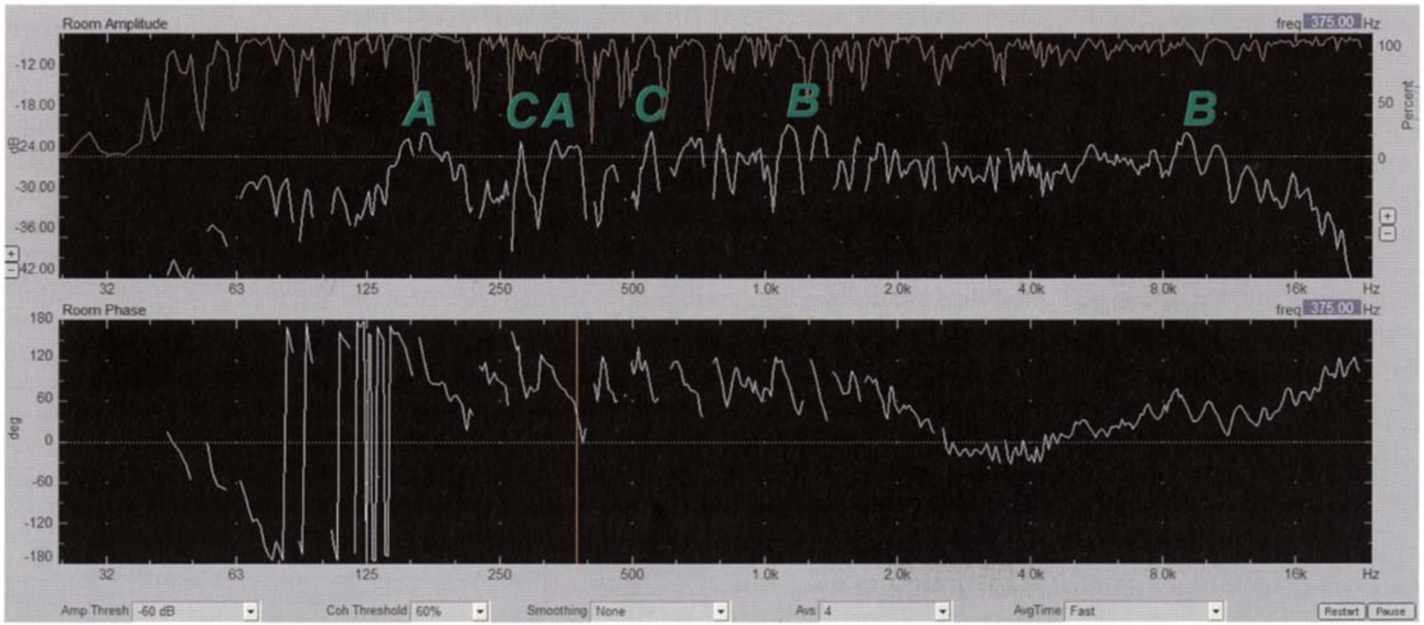


Figure 3.19 Tonal envelope example (A) wide peaks with dips on their sides; (B) wide peaks composed of a series of smaller peaks above the median line; (C) large narrow peaks which are less audible

Perspectives: More than once I have equalized a system with measurements that looked just fine only to find that it sounded awful when I listened to it. At this point it is real important to trust one's ears. It is so easy to get caught up in the technique and forget what the task really is. All it takes is a measurement microphone in the wrong place or using the wrong strategy when equalizing the system. This is when I find out — if it does not sound right, it is not right. I will go hack and start from the beginning and do it again!

Alexander Yuill-Thornton II (Thorny)

the tonal threshold its effects become apparent in the time domain as spatial spreading and, eventually, discrete echo perception.

Field experience confirms that the perceived tonal envelope is consistent with the response trends as we see them on a high-resolution complex analyzer. Wide bandwidth peaks are the most identifiable, even if they are only a few dB above the median. Very narrow peaks of a comparable level are less of a tonal concern. Critical bandwidth is not a single value and is a subject of ongoing debate. The most common value found is around 1/6th octave in the mid- and high-frequency ranges. In the lows the ear's frequency resolution becomes more linear as does the critical bandwidth. The practical implications of this are seen in how we might attempt to compensate for the tonal irregularities with filters. It makes sense to allocate them toward solving the most audible disturbances, i.e. the widest peaks. Prioritization goes down from there, with peaks and dips that are too narrow to be perceived tonally being best left to other solutions.

This is not to say that the presence of dips in the response is inaudible. It is simply that they are not audible by their presence, but by their absence. Large areas of absent response are noticeable over time. Portions of the music are missing. The most noticeable aspect is the change in tone of a particular instrument or voice as it moves through the scale. Some notes stand out, others are lost and others have unexpected colorations of their harmonic structure. Reducing the peaks down to the median level will prevent them from standing out, but it will do nothing to restore the missing pieces caused by cancellation. These fine cancellations are still a concern for us, but they are not wide enough for us to attempt treatment with tonal solutions, i.e. filters.

Echo Perception

Tonal quality, spaciousness and discrete echoes do not come from separate planets. They are caused by the same mechanism: the presence of the original direct sound and imperfect copies that arrive late or with modified frequency response. There is no hard line between these perception characterizations. Like every other aspect of our audio waveform behavior, the line is frequency-dependent.

There is nothing in acoustical physics that shows a defining line between early arrivals that are heard as tonal change and those that are perceived as echoes. If measured in the time domain, any separation of time is seen as a discrete separate impulse. Adding more time offset simply spreads the impulses apart. In the frequency domain we see a continuous narrowing of the comb filtering. In either case nothing informs us directly of changes into spaciousness or discrete echoes. The dividing line is in our head.

We know where the tonal change comes from: early arrival summation. We know where the spatial perception comes from: middle arrival summation. We know where discrete echoes come from: late arrival summation. How do we define late? Late is when we hear it as an echo and we have now created a circular argument.

Since we know that this is excitation-signal-dependent, we need to look at how the signal type affects summation. The key difference is the role of summation duration. Continuous signals have extended summation duration time, and are experienced as coloration. Transient signals have the minimum duration and are the first to move over to the other side. Since high frequencies have the shortest periods, their transient peaks will have the shortest duration. They will be the first to cross the line.

Since the time offset aspect of summation operates as a sliding frequency response filter in our tonal perception it is fair game to apply a sliding scale for echo perception. Late for 10 kHz is a very different matter than late for 100 Hz. An arrival that is 5 ms late will cause comb filtering at 10 kHz that is 1/50th of an octave wide (an offset of 50 wavelengths). This is far beyond our critical bandwidth (an offset of six wavelengths) for tonal variation. What is it then? It has to be something. At the least it is a *potential* echo. If the signal is transient in nature we will hear it as an echo. We are certainly not going to try to equalize it, so for our purposes, it's an echo. What is happening at 100 Hz? 100 Hz? Do you read me? We regret to inform you that 100 Hz (half wavelength offset) has undergone the ultimate tonal variation: cancellation. Its neighbor, 200 Hz, received an octave-wide tonal boost in the process. The sliding scale of tonal to echo is as plain as the logarithmic teeth of our comb filter. Where the teeth are wide we hear tone, when they get too narrow we have echo potential.

One of the common numbers given for the echo perception threshold is 60 ms. Such delay would create filtering of approximately 1/24th octave in the range of 400 Hz. The time threshold numbers spring from research that used the voice as the source. 400 Hz is in the heart of the human voice range, so it makes perfect sense that we would perceive the voice as having crossed over the echo perception line. But the vast majority of a bass guitar is below 400 Hz, and the high hat is miles above this. What if we were to reverse the perspective of our threshold standard from time offset to its frequency response effects? What if we go with 1/24th octave (24 wavelengths late) as the threshold instead of the fixed time of 60 ms? This would translate to a sliding time scale with 2 ms at 12 kHz and 240 ms at 100 Hz. This would be significant since this would give us a single threshold from which to discern between time domain and frequency domain solutions. Optimization options for arrivals under the 1/24th octave threshold would include the possibility of equalization, whereas those beyond would be restricted to delay or other solutions. Does it hold up? Though I have no research laboratory to back up this claim, my own research and experience support it. I have listened to summed highly transient signals and continuously adjusted the time offset. The transition to echo happens gradually, beginning with only the highest frequencies and steadily moving more and more of the lower spectrum over to the other side. The transient peak showed perceptible separation of the highest frequencies in less than 2ms. This is consistent with 1/24th octave expectations. As the time is increased to 10 ms the splitting was extremely apparent in the mid-range and yet sounded only stretched in the lows. By 25 ms the signals sounded like they were strongly separated. But we know they weren't. How could the 30 Hz wavelength have separated from itself when it had not even completed a single cycle?

This does not necessarily conflict with the mainstream audiology community findings. Their accepted numbers are based on speech transmission and their 60 ms findings are consistent with the mid-range perception transition. Our work in system optimization goes beyond the range of speech.

Spatial Perception

There are times when we can precisely pinpoint a single source of sound. At the other extreme are times when we can pinpoint multiple sound sources in the form of echoes. In between these extremes is space, the final frontier. The gray world between is that of spatial perception. The sound source is experienced as "big," "fat," "wide," or any variety of terms that mean one thing: our ears are receiving conflicting information about the sound source location and duration (Everest, 1994, pp. 295-301). We know from our discussion of localization that the sound image will favor the direction of the first arrival, unless this is strongly dominated by level offset. That does not mean that such a perception is equivalent to hearing single source. Far from it. Consider the following. Precedence effect movement of the sonic image has already run half its course with just 2.5ms of time offset. At 10kHz this is an offset of 25 wavelengths, enough to move it over to the echo perception side. At 100 Hz it is laughably insignificant offset of 1/4 wavelength. Who is driving the spatial ship? It's the high end, of course. The perceived result is that the high-frequency image has moved the farthest toward the early location. The mid-range and low frequencies are stretched across the horizon between the sources, since they are separated by fewer wavelengths. The composite image is smeared across the horizon.

We could say that low frequencies have more spatial "elasticity" than their HF counterparts. The image stretches between such sources easily in contrast to the highs which snap apart quickly. By the time we add enough time offset to move the low-frequency image over to the early side we will have long ago crossed the line into echo perception at the higher frequencies. Obviously this won't work.

The spatial perception zone is the area where the time offset is low enough to move the image at some frequencies without being perceived as discrete echoes at others. Signals with lots of transient high-frequency content have the smallest window between the tonal and echo perception worlds. Signals with filtered high frequencies (and therefore reduced transient response) can arrive later, and add to the spatial experience under our echo perception sonar. Symphony hall acoustic design requires mastery of the sequence of arrivals such that later arrivals have lesser amounts of high-frequency content. The goal is maximizing the spatial perception in the listener's experience, and minimizing the echo perception.

Perspectives: The goal of any acoustic measurement system should be to help you understand what your gear is doing and what you are hearing. The more you see the measurement as an explanation of what you are hearing and the way that your system is affecting a signal, the more effective the measurement system will be in helping you optimize a system!

Sam Berkow

Perception Zone Detection

Is there any way we can tell where we are with an analyzer? Fortunately, yes. The answer can be found in a high-resolution frequency response. The tonal perception zone is found where the summation sources are a small number of wavelengths apart. If one follows critical band theory then we can set the threshold at 1/6th octave. Therefore the tonal zone is found when the combing is spaced by 1/6th octave or less. The echo perception zone surfaces when we have exceeded the ear's tonal resolution, which experimentally we have found at 1/24th octave (24 wavelengths). The area between is the spatial zone. An analyzer with 1/24th octave, or better, logarithmic frequency display can provide us this information. Such a response is shown in Fig. 3.20.

Stereo Perception

Introduction

Stereo music reproduction is one great life's great audio pleasures. Sound is spread over the horizon as each instrument occupies a unique place creating a rich musical panorama. Stereo reproduction is unquestionably more "fun" to listen to than a comparable monaural reproduction. That is, of course, provided that we are in the listening area where the stereo panoramic effect actually occurs. For those listeners outside of the preferred seating area, the mono system is at least as good, if not better than the stereo.

How does stereo work? What are its limits? Are there any side-effects? How does this fit into our system designs for large spaces? To begin to answer these questions we must first understand how stereo is perceived. Stereo is a state of mind. It happens in our heads. It is a by-product of our binaural localization and the confluence of our dual-channel listening system.

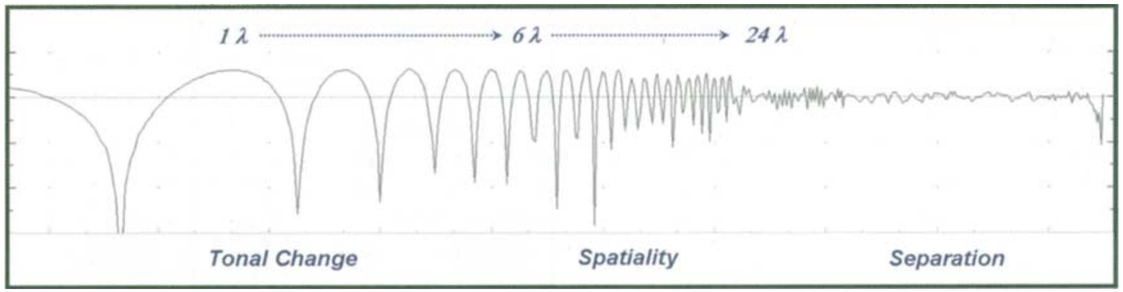


Figure 3 20 The role of wavelength offset in the perception of summation signals. Small offsets create a perception of tonal change while large offsets create the perception of separate sources. The wavelength offset at a particular frequency is a function of the time offset

Panoramic Field

We all know that the best seat in the house for stereo is the center, and that as we move off-center the quality of the effect diminishes. But why? Stereo perception is a based first on the ILD (level offset) system of our binaural hearing. We saw previously how a single channel of audio signal could be routed to two speakers at different levels and moved from side to side. The image can be placed anywhere on the horizon between the two speakers, the width of which is the **panoramic field.** Placement of the signal in the field is controlled by a panoramic potentiometer (hence the name "pan pot") on the mix console. This control maintains a constant overall level with a variable level offset between the left and right channels.

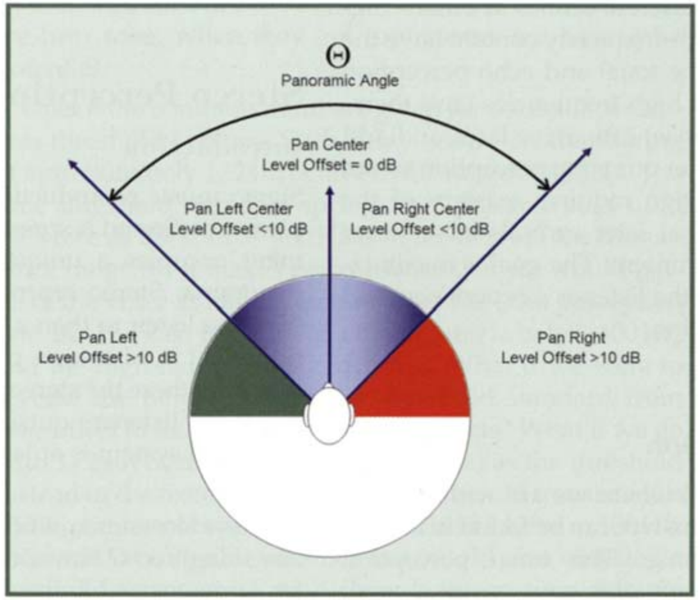
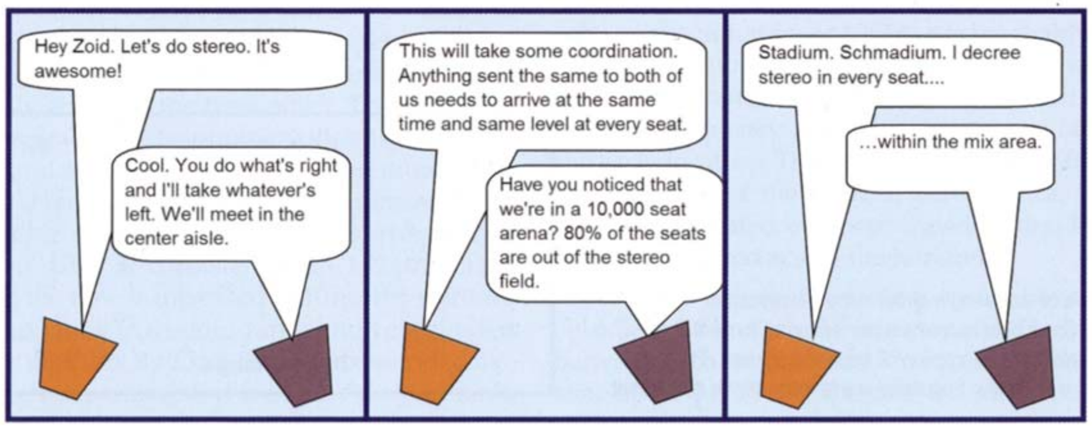


Figure 3 21 The panoramic field is controlled by the relative level settings of the pan pot

Distance Factor

As we approach a stereo pair, the angle from center to the individual speakers increases. The panoramic field expands proportionally. The inverse occurs as we move away. It is generally accepted that the optimal stereo panoramic field angle is around 30 degrees from center. This is easily accomplished at home or in a studio where the speakers are targeted to a single point, the "sweet spot." Overly wide stereo fields of greater panoramic width tend to have insufficient image stability in the center area, while smaller fields will have a compressed range of positions. The compressed field translates to "not as fun." Wide-angled fields have the smallest percentage of area where the stereo perception remains strong. Narrowed fields have a larger percentage of stereo area, but with a lesser

Trap 'n Zoid by 6o6



Perspectives: Delay is the most underestimated consideration in source panning and control. I work in surround situations and unconventional spaces, and the illusion of panning really requires close attention to the real effects of system timing for it to come off. In the end nothing substitutes for the ability to listen to a panned effect from many different locations in the house, and little by little the conventions of a technical production process are changing to make time available for these difficult effects to be programmed successfully.

Martin Carillo

Perspectives: Know your client. I once spent one and a half days placing and aligning both main and near-field monitors in the control room of a large studio complex. The studio room was well equipped with a grand piano, a very spacious drum booth, several vocal booths, etc., and was big enough for a sixty-piece orchestra. I remounted and rewired outboard equipment into lower profile racks to avoid reflections near the mix position and even time-aligned the near fields with the mains so that engineers could fade from one to the other seamlessly.

effect. In a concert setting we have an extended width and depth of coverage. Therefore it is a certainty that the width of the panoramic field will vary over distance and lateral position in the hall. Sadly we can not fulfill the marketing department's promise of stereo in every seat.

Off-Center Effects

From the center position our perception of the level panning is the most accurate, due to the fact that the relative time (ITD) between the sources is neutralized at zero. As we move off the center axis we find ourselves affected by both the time offset and level offset mechanisms. A small move to the left brings us closer to that side raising the time offset in favor of imaging to the left. The result is that a signal panned to the center now appears left of center and one that is panned to right center moves to the center. As we move further off-center, the discrepancy between the electrical panoramic placement and the acoustically perceived placement increases, until we reach the point where we are 5 ms closer to one side than the other. At this point all of the image steering capability of the precedence effect has run its course. Any further image movement will have to be done by the brute force of dominant level offsets. By the time 30 ms of time offset has accrued we begin to venture into the zone of strongly perceived echoes. Our stereo field effects have run their course from "fun" to negative as we move off-center.

With this lateral movement a number of factors conspire to reduce the quality of the stereo experience. The first is the angular compression of the panoramic field. At center we have equal angles between our position and the left and right speakers. As we move off-axis the angles of the stereo field are reduced and become asymmetrical. For our example we will consider a stereo field that comprises a total of 60 degrees (30 degrees for each speaker to center). We will begin with the listening position 30 degrees off-center. A leftward move reduces the overall angle between the sources by geometric triangulation. The angular spread between left and center widens from the listener perspective, while the angular spread between center and right compresses.

However, the leftward listening position shifts where a center-panned image occurs and causes the left channel's portion of the stereo field to be reduced. The right side does the reverse: narrower angle of coverage spread and wider angle of placement field. As we move far off-axis the center panned material moves far to the left and is indistinguishable from material panned to the left. This compresses the left's portion of the stereo horizon to a sliver. By contrast the right channel occupies the rest of the horizon line. A large amount of panoramic movement from center to left will result in a few degrees of movement. Panning from center to right will have very little effect until the signal is panned almost exclusively to the right, at which point the image jump quickly outward towards the far speaker.

Mixed Signals

Stereo panning can be managed for more than a single signal. Mixed signals may be used, with each channel having their own level panning. The result is that different channels can occupy different positions in the panoramic field simultaneously.

Stereo Scaling

Having listened to stereo at home for all of our lives it seems self-evident that such a scheme is transferable to the concert venue. Left and right channels can be placed in the same arrangement as at home. As long as the speaker arrays have lots of horizontal coverage overlap the stereo image will behave similarly to our living room, right? Unfortunately not. The distance factor dominates the stereo. Recall our previous discussion of the "Pythagorean problem" (see Fig. 2.21). The level offsets scale as a ratio of the distances. This leads us to view a large space as having the same stereo level distribution as our living room. This is true. The time offsets are a different story. They accrue linearly, one millisecond at a time. In time-offset terms the large space provides us with no quantity discount, and no ratio scaling. When we are off-center, we are out of time. Period.

The client was amiable enough but implied that I shouldn't worry if things weren't perfect. He had a very relaxed attitude considering I was aligning a studio that was clearly represented a major investment. As was my normal practice, I aligned the monitors for a wide "sweet area" and conducted listening tests using CD tracks that I knew well; then, once satisfied, approached the client for a selection of secondary master tapes so that we could fine align the system to his particular wishes.

The first track was a well-known TV advertisement for frozen peas followed by one for washing powder and another for toothpaste! The client then explained that most mixes would be finalized using a typical domestic TV as a reference monitor. The main and near-field monitors would be used to check for hum, clicks and high frequency crosstalk though ...

Jim Cousins

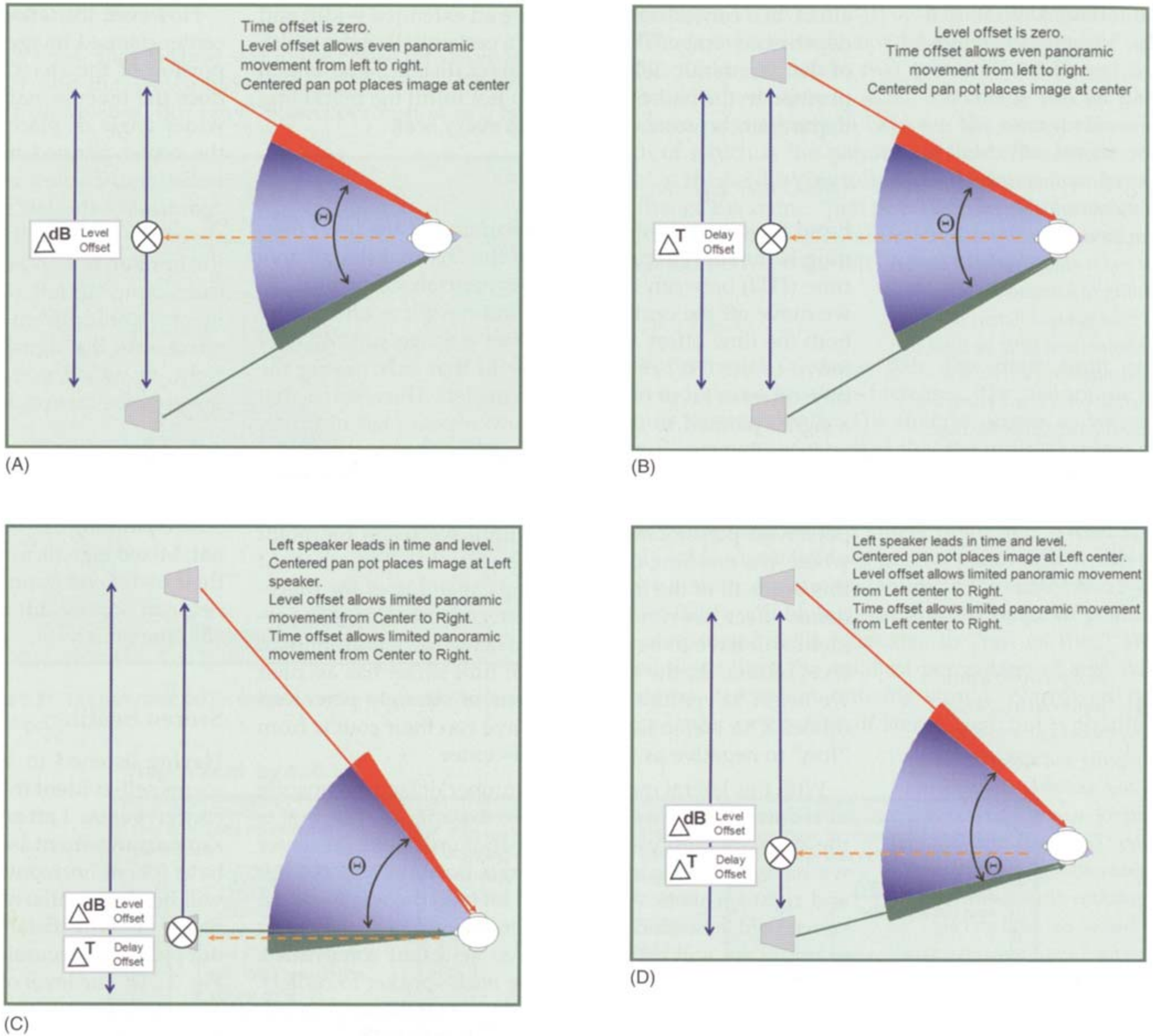


Figure 3.22 Stereo panoramic field over angle (A) level panning from center; (B) delay panning from center; (C) level and delay panning from side; (D) level and delay panning from side/center

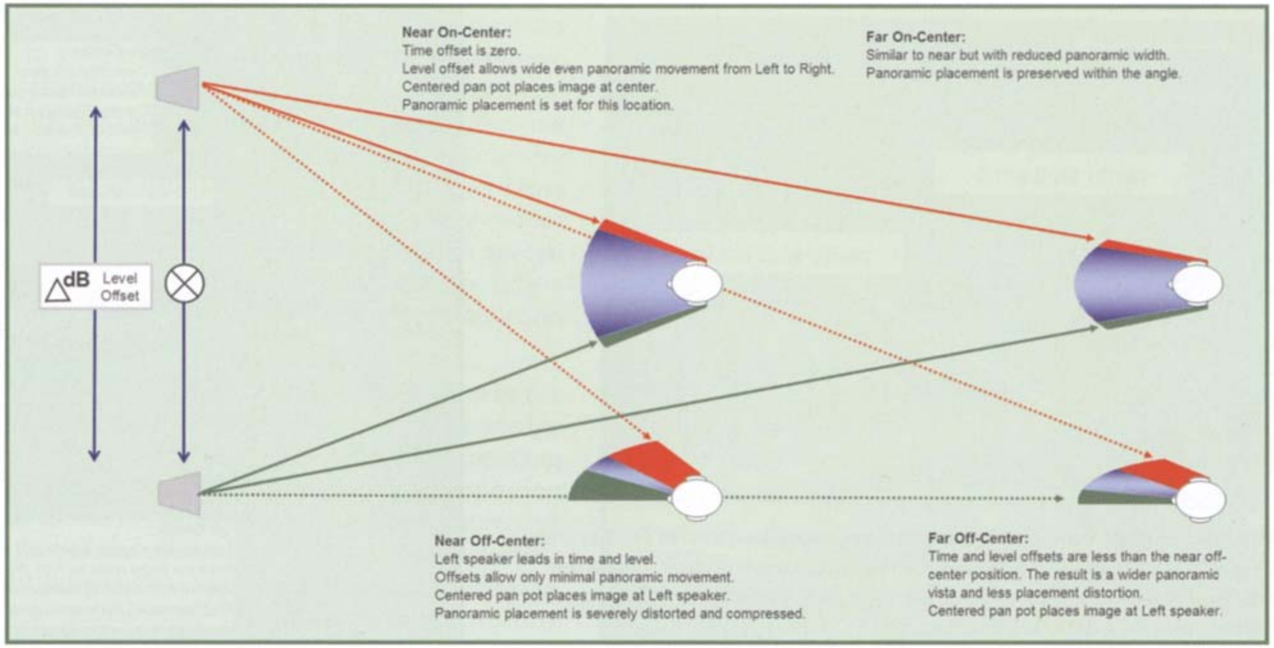


Figure 3.23 Stereo panoramic field. Multiple locations for a system over distance and angle

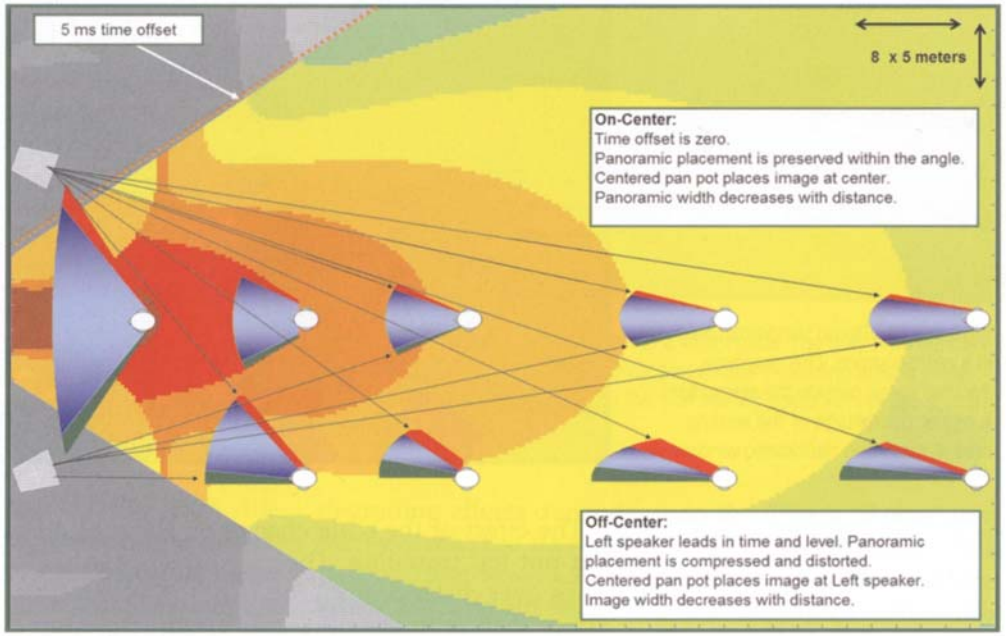


Figure 3.24 Stereo panoramic field in the home. The gray area denotes areas outside the stereo field

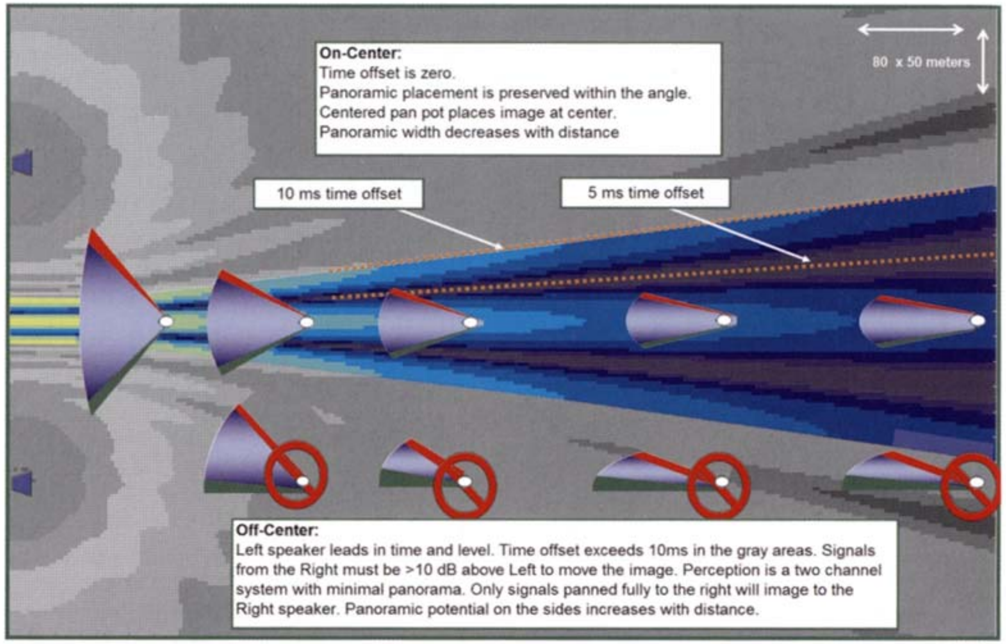


Figure 3.25 Stereo panoramic field on a concert scale. The gray area denotes areas outside the stereo field. The same dimensional relationship is maintained as in the previous figure, but the distances are 10x larger. Note the reduction of the proportion of seating inside the panoramic field area

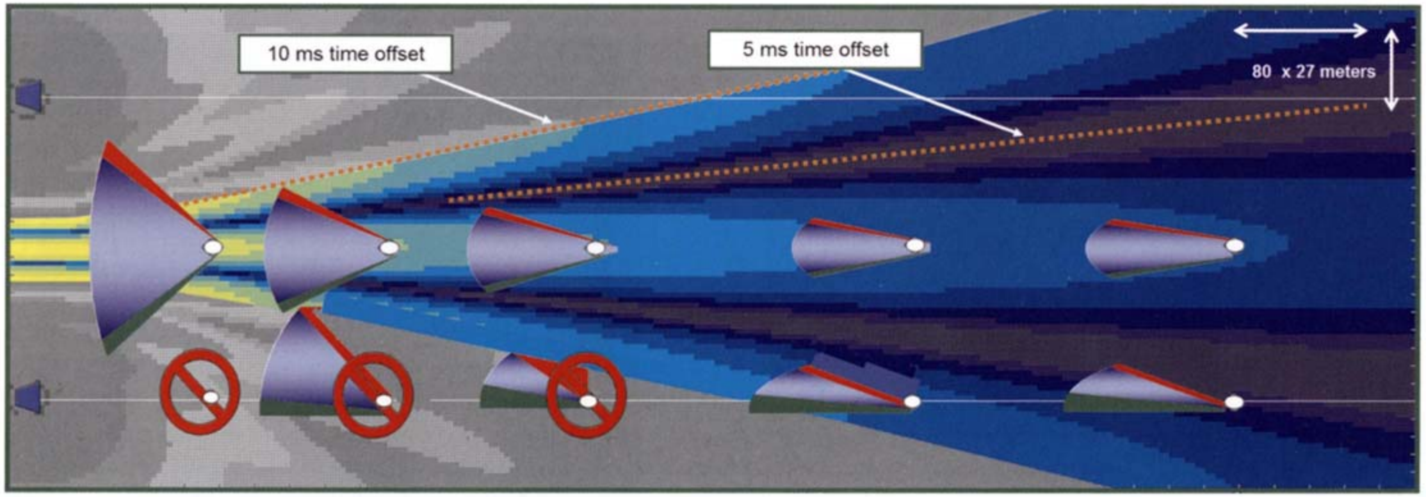


Figure 3.26 Stereo panoramic f
area is inside the panoramic window

The effect of the scale change on the center is obvious but not too troubling. The stereo field gradually diminishes over distance until the width is so small we hardly notice it. It is in the off-center areas where the outcome is not as expected. The culprit is not a few tenths of a dB of level offset. It is time. Our binaural localization system runs out of gas after 5ms, i.e. the localization needs more than 10dB of level dominance to move the image if it is more than 5 ms late. We do not have to go far off-center in an arena to get 5 ms closer to one side. After we have vacated the central zone the system becomes two monaural channels. Only the items panned exclusively to the one side will be heard from distinct locations. Everything else will appear at the nearest speaker location. For most arena concert listeners, the positive side of the stereo experience is all the excitement of the occasional tom drum or special effect appearing on the opposite side. The negative side is one of decreased intelligibility, increased tonal distortion and increased echo perception.

The shape of the stereo field on a large-scale system is a wedge that expands over distance. The width of the wedge is determined by the speaker spacing. As the spacing widens, the wedge narrows, leaving a smaller proportion of listeners in the field. As the spacing narrows, the wedge widens, but the panoramic width narrows, thereby creating a less than thrilling stereo experience. Here are our choices for large-scale spaces: wide panoramic stereo for a tiny minority of seats or narrow panoramic stereo for a minority of seats. Bear in mind that all but the center line seats will have distorted panoramic placements, due to the time and level offsets. The proportion of listeners in the wedge will depend upon the room shape. Long and narrow venues yield the highest. Wide rooms yield the lowest.

Perspectives: Spatial sound is a fascinating thing to play with, even though it has the so called "sweet spots." Even so I think that one must continue to work with audio in this fashion. Certainly the scenic and lighting designer's rarely consider the audience seating off-center. Looking at the visual reference (the stage) from different angles requires the sound to be as unique as is the visual input.

Jonathan Deans

Stereo Side-Effects

To determine whether stereo is worth the trouble requires an examination of the possible side-effects. If the effect enhances it for some, and does not degrade the experience for others, then we have a clear directive: go for it. But the side-effects are substantial. They are the products of summation due to the source displacement of the left and right sides. Any signal that is not panned exclusively to one side will create a stable summation in the space. Every seat that is off-center will be affected differently due to the changing time and level offsets between the competing sources. For those listeners closer to center the effect will be the tonal variations of comb filtering. Those on the sides will be subjected to discrete echoes. The extent of the degrading effects depends upon the degree of overlap in the systems. The damage done by stereo is proportional to the extent to which the level overlap occurs in places where the time offsets are too large to keep us in the stereo perception window.

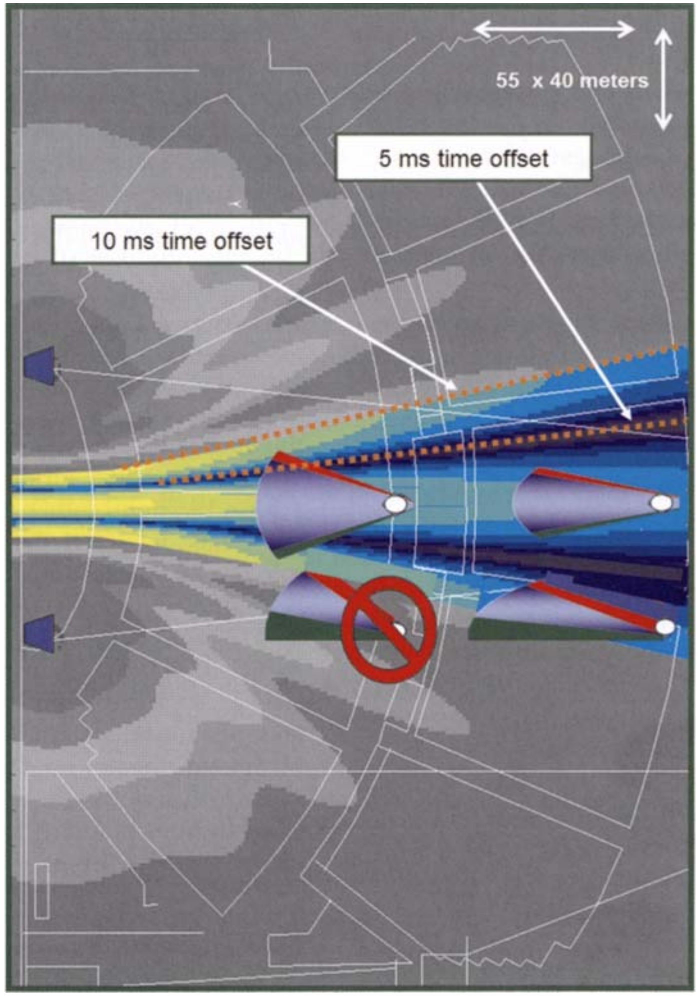


Figure 3.27 Stereo panoramic field in a wide space. The gray area denotes areas outside the stereo field. A lower percentage of seating area is in the panoramic field. Attempts to cover the side seating areas with stereo sound will result in perceptible echoes

Amplified Sound Detection

Is it live? Or is it Memorex™ ? This oft-quoted commercial ran for years on the US television stations. Jazz great Ella Fitzgerald would sing inside a vocal booth while a speaker blared her voice at a wine glass and ... crash! The wine glass shatters. We were to conclude that since the wine glass could not tell the difference between live sound and a cassette tape then we could be fooled as well. It would be nice if things were so simple.

One of our roles as audio engineers is to suppress the clues that can potentially break the illusion that listeners are in the direct sound field of the performers. We want the audience to focus on the artist, not the loudspeakers. We must consider all of the perception clues that tell us a sound system, rather than a magical breach of physics laws, has delivered the sound from distant performers into our lap. There are a large number of factors that must be managed to create the illusion. It takes only one exposure to break it. One hint might be the hand held mic. With sufficient effort it is possible to make audience members who lack professional experience, oblivious to the presence of a sound system. A perfect case in point is the sound of musical theater on Broadway. The existence of sound systems has not been noticed by the theater critics. How else can we explain the fact that every theatrical discipline except sound receives annual Tony Awards?

Distortion

All forms of distortion give us clues that a speaker is involved. Harmonic distortion is the least obtrusive, since it has the "musical" quality. Intermodulation distortion is far more noticeable. The by-products of this distortion are not harmonically related and are detected even at very low levels. Distortion in digital audio circuits is also unrelated harmonically and therefore is to be avoided.

Clipping occurs when an electronic stage is run past its voltage limits. Clipping is audible as momentary bursts of both harmonic and intermodulation distortion followed by undistorted periods when the signal is within limits. Operating the system within its limits is, however, a mix engineering issue, and there is no system design or optimization strategy that is immune from this possibility. Our design goals will be to provide sufficient headroom in the system to minimize the time spent in overload.

Compression

One means of preventing the destruction of loudspeaker components is the use of compressors and limiters. Both devices reduce the voltage gain through the system after a threshold level, based on the power limits of the drivers, has been exceeded. The sonic effect is to dynamically flatten the response, so that the loudest passages do not cause excessive clipping. This is far less audible than gross clipping but when taken to excess will be noticeable as an unnatural dynamic structure. The worst case is an overly worked compressor that "breathes," i.e. has audible clamp and release actions. Again, this is an operational issue.

Frequency Response Coloration

The frequency response has many clues for us in the identification of sources. If a clarinet does not sound like a clarinet, then we are set to wondering how this could be happening. Could a strange room reflection cause that? How many drinks did I have? Ah yes! There is a sound system involved!

A frequency response that has substantial peaks and dips will attract the attention of the ear. We have sonic expectations in our heads. Our memory contains a map of what a particular instrument or voice should, in our opinion, sound like. A secondary memory map contains what that instrument sounded like just seconds ago. If sound is consistently unnatural, or unnaturally inconsistent we are prone to suspect a sound system is involved.

False Perspective

A violin is a violin. But a violin placed on my shoulder has a very different sound from one that is 30 m distant. We carry more than a memory map of the violin in our

Perspectives: One of the things I have learned over the years is that bad sounding gigs get complaints but great sounding gigs only silence. So if the next day review in the local press doesn't mention sound it means everything was OK.

Miguel Lourtie

Perspectives: It is apparent that the general audience finds it hard to recognize sound, when the sound fits the visual cue. When sound is out of place, it is sensed as an effect or as an error.

Jonathan Deans

heads — we also carry a sonic perspective. A distant violin should not sound like a close one, nor should a close one sound far away. When this occurs we are sending the clues of false perspective.

When we reinforce a sound source we are distorting the natural sonic perspective. By adding level and extending the high-frequency response, a distant source is perceived as near. There is good reason to do this. It brings the audience closer to the performance. This also allows performers to reach the audience without having to bellow their voices in the manner of the classical theater actors and opera singers. For theatrical performers the advent of sound systems has allowed a much greater range of realistic speech and emotion to be available. A song can come down to a whisper, but still be heard. In short, with only "natural" sound the actors must speak in an unnatural way in order to project to the rear. With "unnatural" sound (that's us) the performers can speak in a natural way.

The process has its limits. We want the audience to suspend disbelief, as they would do while immersed in a movie plot. It is critical not to push the envelope too far. We do not want audience members to feel the actors are directly at their ears. Distant sound sources have reduced direct to reverberant ratios, large amounts of low-frequency room reflection summation and large amounts of high-frequency air loss. They are also low in level. Close sounds have a high ratio of direct to reverberant sound, a minimal amount of low-frequency summation from the room reflections and minimal loss of high frequencies. They are also much louder. We will need to be mindful of the expected natural qualities in order to remain undetected while moving the perspective.

As we move closer the level increases, the high-frequency response increases, the low frequencies smooth out and the room falls away. All of these qualities can be achieved by our sound system. The "whispering in your ear" quality is prevented by never allowing the range above 10 kHz to peak above the main body of the response. This extreme presence of HF energy on top of the signal occurs naturally only in the nearest field. Therefore high-frequency extension in our speakers aimed to the distant locations must be capable of reaching the back without overheating the near seats.

It is also possible to reverse the perspective and make near sources appear distant. This is not desirable in most cases. Nearby off-axis areas have a similar response to distant areas. As we move off-axis, the experience is one of becoming more distant. The lack of high-frequency coverage mimics the air loss and pushes our perspective back.

A second way to create false perspective is by overly aggressive low-frequency reduction in the far field. If the low frequencies are flattened down too aggressively at great distances the response too closely resembles a near field response to remain plausible.

If the low frequencies are reduced to a level under the mid-range, a "telephone" tonal quality arises. There is nothing in nature, short of talking through a tube, which reduces both highs and lows, while leaving the mid-range intact. If our system exhibits telephonic qualities, the illusion is broken, the tone is undesirable and the audience is liable to pick up and answer their cell phones.

Dual Perspective

Another way to expose the speakers is dual perspective. We cannot be both near to and far from a sound source at the same time. We cannot be simultaneously in a small room and a cathedral. If the high frequencies appear to come from a nearby source and the lows from far away, we have a case of dual perspective. The most common mistake in this regard is in the case of under balcony delays. Aggressive low-cut filters on the delays attempts to reduce low-frequency summation in the hopes of better intelligibility. Main systems and remote systems do not combine evenly over frequency. In the mid and upper ranges the combined perspective is near field, since the delays dominate. In the low frequencies the distant mains are dominant. This will always happen to some extent but the removal of the lows from the delays serves only to exacerbate the difference. This creates an unnatural dual perspective of both near and far field experiences.

Perspectives: In the intimate theater it's rare that a substitute for the real thing will do the trick. I've found that when you can't have the real phone ring or the TV onstage the audience won't maintain their suspension of disbelief but when you do have that level of verisimilitude then there is a certain amount of the metaphysical in the realm of that suspension that one has to make up for. Capturing an audience's belief in the action onstage demands a certain amount of astonishment and surprise around the edges of" the reality. The line between the sonic "reality" or naturalism, and the sonic possibility of the unreal or the dreamlike reality of real artifice can pose some of the most interesting problems for source, and speaker placement. And therein some of the greatest need for accurate loudspeaker delay, to facilitate convincing transitions from the world of the actors to the world of the audience.

Martin Carillo

Another version of dual perspective occurs in multiway speaker systems, such as a two-way speaker where the low driver and high driver have very different directional properties. As the directionality changes, the perspective moves with it. The increased directionality of the highs brings the listener's perspective forward while the lows move them back. It is natural for this to occur gradually in a room. It is not natural for it to happen in a single note as might happen if steep filters are used in the frequency divider.

A secondary form of this arises when speakers covering different frequency ranges are displaced. This is typical for subwoofers, which may be on the ground while the rest of the system is flown. It is possible to find a split sonic image with low and high frequencies coming from different areas. This is most audible when the displacement is large and/or the spectral crossover is steep. See Chapter 2 for more on the subject of crossover audibility.

Sonic Image Distortion

We have already discussed the mechanics of sonic imaging. If the sound appears to be coming from speakers instead of the natural source, this in an obvious case of false perspective. Even if the sound is a perfect copy of the original it will not fool anybody if it does not appear to be coming from the performer. Our speakers will need to be strategically selected and placed, with carefully set timing and level to produce plausible imaging results.

Video projection at the rear of the stage will need to be delayed to be synchronized with the sound from the speakers (Photo Dave Lawler)

Synchronization

Many large-scale events include video projection in order to enhance the experience for distant attendees. In such cases there is no longer any pretense that we are viewing the actual performer, therefore we will not need to hide the presence of the sound system. There are, however, the potentially disturbing effects of unsynchronized sound and video. The image from the video screens will arrive ahead of the sound, unless video delay is added. Lack of synchronization of video and audio decreases the speech intelligibility as our subconscious lip reading conflicts with the audio. Besides — it is just plain annoying!

Microphone Reception

Introduction

Our ears are not the only receptors of sound. Acoustical signals can be captured and converted into electronic signals by microphone transducers. Microphones have a great variety of uses in our systems, most of which are capturing the emission of sound sources that will be transmitted by our speakers. For signals of acoustic origin the microphone is the point of initial entry into our sound system and it can be said that our principal effort is to recreate the sonic experience at that microphone position to our listeners. In that sense the microphone is our surrogate ear, and its faithfulness to our ear's response would seem to be a defining factor. The reality, however, is not that simple.

Comparing Microphones to Our Ears

Microphones differ from our aural system in very distinct ways. Our application on stage has strong factors that favor microphones that are far different from our ears. First, our ears are a binaural system separated by our head. To recreate this experience requires a pair of mics, placed on an artificial head shape. This practice is known as "dummy head" or binaural recording. Recordings made in this fashion are particularly fascinating, but extremely limited. The chief limitation is that the recordings must be played back on an analogous speaker system: headphones. This limits us to the 3-D glasses, virtual-reality type of application and special rooms with a precise listening location. Every few years "a major breakthrough" occurs as someone announces the impossible: that they have a special processor that can provide this experience to an entire room.

The practical compromise for recreating the human hearing experience is a simple stereo recording pair. The mics are placed in close proximity and splayed apart. The individual mics will have increased directional control over that of our hearing system but will otherwise create a familiar experience. For this to work, the mics must be placed sufficiently distant for the stage sources to have mixed themselves together. Such a placement strategy is pretty much useless for sound reinforcement miking due to leakage. Mic placements that mimic our hearing are out for the stage. What will work?

Microphone directional control is of extreme importance in sound reinforcement applications. All acoustical stage sources will find their way to all mics at differing level and delay offsets. When the microphone signals are mixed together in the console the resulting electronic summation will have comb filter distortions of the source responses. There are several defenses against these comb filter effects. Close miking of the sources, with highly directional mics, affords a large degree of sonic isolation. This does, however, create a false perspective issue, as described above, where the listener is transported to positions such as inside the kick drum. To create a realistic perspective will require modification of the response at the mic or in the mix console.

A second defense is sonic isolation on stage, by use of baffles, or politely asking the musicians to confine their levels to the local area. All musicians want to hear themselves and some musicians actually wish to hear the other musicians. This presents an additional complication: stage monitors. These sound sources introduce an additional leakage path that leaves us wanting to maximize directional control in the mics.

We can conclude then that the stage area of popular music group is an acoustical combat zone. The mix engineer's approach is to isolate the sources as much as possible, with the hopes of reassembling the pieces in the mixer. Microphones that simulate the human hearing mechanism are not a suitable choice for this task.

Measurement Microphones

The primary mission of microphones for system optimization is to gather response samples at different locations in the room in order to monitor the transmission. These samplings will need to contain both the direct sound and the reverberant field as our listening experience is affected by both. The omnidirectional mic, aimed in the direction of the sound source, is a reasonable approximation to the response of a single ear. The pinna has a limited degree of high-frequency front/back directionality as does the omnidirectional mic. Cardioid mics, by contrast, are far more directional than the ear and change their low-frequency response over distance (the proximity effect). The role of our ears in vertical and horizontal localization will be lost, but the summation frequency response is a close approximation to the perceived tonal response. Therefore, the surrogate ears for our transmission monitoring will be single omnidirectional mics placed at strategic positions in the hall.

The criteria for measurement mics go beyond extremely flat frequency response alone. The mics must be low distortion, and have dynamic capability sufficient to measure the system transient peaks without distortion. The self noise must be below the ambient noise level in our space. Unfortunately this is rarely an issue, as all but the best recording studios have so much ambient noise that our microphone self noise is academic.

Omnidirectional Microphones

Measurement mic specifications:

- Frequency response: 27Hz to 18kHz ± 1.5dB
- Omnidirectional
- Free field
- THD < 1 per cent
- Max SPL without overload > 140 dB SPL
- Self noise < 30 dB SPL (A weighted).

Cardioid Microphones

There are experienced users in the field of system optimization that use multiple cardioid microphones for testing. There are several important considerations worth bearing in mind when substituting cardioid mics for this application.

Omni vs. cardioid measurement microphones:

- The cardioid microphones will give superior coherence readings, reduced frequency response ripple and generally easier to read data.
- The cardioid mics are more directional than the human ear, and therefore the optimistic readings may be misleading. This is somewhat mitigated by the fact that the omni mics are less directional than the human hearing, thus giving a pessimistic response.
- The cardioid mic must be very precisely aimed at the sound source for accurate results. This requires vigilant attention to prevent false conclusions regarding speaker high-frequency response.
- There is an additional complication when measuring the interaction between two sources. If the sources originate from different angles to the mic they cannot be both simultaneously on-axis. If the mic is placed on-axis to one it rejects the other, if placed at the center axis between the sources then the error is spread to both.
- The proximity effect, while well known for its effect in the near field of cardioid mics, continues in the far field. As the mic moves away from the source, the low-frequency response continues to roll off. Distant free-field measurements will show a reduction for each doubling distance. Indoors there is a natural tendency for low-frequency response to rise in the far field due to strong early reflections. This audible effect will be removed from the measured data by the proximity effect of the cardioid mic.

References

Duda, R. O. (1998). Sound Localization Research, San Jose State University, http://www-engr.sjsu.edu/~duda/Duda.Research.frameset html

Everest, R A. (1994). *Master Handbook of Acoustics,* TAB Books.

Robinson, D. W. and R. S. Dadson, A re-determination of the equal-loudness relations for pure tones, British Journal of Applied Psychology, 7 (1956) 166-181.

Section 2:
Design





4

Evaluation

> ***Evaluate*** *v.t. ascertain amount of; find numerical expression for; appraise; assess; to consider or examine something in order to judge its value, quality, importance, extent, or condition*
>
> Concise Oxford Dictionary

Introduction

Sound engineers are from Mars, acousticians are from Venus

The above caption refers to a popular book by John Gray that focuses on the challenges that we humans face in communicating our needs with other people that do not share our perspective. In Gray's book the parties are men and women respectively. Neither is right nor wrong, but both have very different approaches to life. The key to success, in Gray's view, is to maintain your separation, but understand the other's perspective. Then the relationship can meet the needs of both parties as effectively as possible. If the parties fail to understand each other's perspective a cycle of conflict and blame assessment degrades the potential for a satisfying union. So it can be with sound engineers and acousticians. However, it does not have to be so.

A symphony hall acoustician is informed that there is trouble in his recently designed concert hall whenever the sound system is used for popular and jazz music artists. It was known from the start of the project that this hall would feature such a variety of artists and the owners had been assured that the acoustics would be perfect. The hall has been a major success for the symphony so there was no need to further evaluate the architectural acoustics. The problem must be the sound system: the 40-element speaker system, designed for the space by audio engineers. The acoustician proposes a solution: a single omnidirectional point source speaker system, a dodecahedron, is to be hoisted above the stage and replace the existing speakers. This is the type of sound source that acousticians use to simulate the radiation of sound from a symphonic stage. The acoustician was perfectly correct in seeing that the transmission characteristics of the 40-element speaker system would not excite the room in the same way as does a symphony on stage. The omnidirectional speaker much more closely resembles the emission characteristics of natural sound from the stage.

Was this the solution? Of course not. The result was persistent feedback and unintelligible sound. The previous sound system alone could not be the solution either. Both approaches suffered the same fatal flaw: a mismatch of the emission/transmission system to the reception area. The natural sound transmission model works best when the emitting source is directly coupled to the reception area (the room) which is shaped and conditioned to act as the transmission system. The amplified sound transmission model works best when the sound source is uncoupled from the room and where the room acoustics provide minimal modification of the transmission. Amplified sound transmission into "perfect" symphony acoustics is as mismatched as would be the symphony outdoors without so much as a band shell. These are the ultimate impedance mismatches.

Natural Sound vs. Amplified Sound

Which is better, natural sound or amplified sound?

Most people can answer this question without the slightest hesitation: natural sound, of course. The conventional wisdom is that the "unnatural sound" of speaker systems is something that we put up with in those cases where natural sound is not feasible. Natural sound, however, would be the first choice. But in actual practice the opposite is the case. The principal market for exclusively "natural" sound is limited to extremely small venues or those cases where "unnatural" sound is forbidden by tradition. Natural sound can only thrive when the program material, the sound source and the acoustic space are in perfect harmony, e.g. symphony music in a symphony hall, opera in an opera house or spoken word in a small conference room. Take any of these acoustic signals out of their scale-matched, form-fitted environments and their vulnerability becomes immediately apparent.

If natural sound is so superior, then why has it lost 99 per cent of the market share in the last century? If it is the first choice, then why do so few choose it? Wouldn't we be shocked if 25 people were gathered together to hear a speech and there were no loudspeakers? What would be the answer if we asked people around the world to name the most famous classical music event they heard in their lifetime? The most likely answer would be "The Three Tenors" concerts, which were performed in stadiums through gigantic sound systems. Certainly none could argue that seeing the three famous tenors singing without a sound system in one of the world's great opera houses would be far preferable as a patron. The stadium venues were chosen for their superior gross revenues, not their acoustics. Nothing personal. Business is business. An optimized loudspeaker can beat optimized natural sound 99 times out of 100. If you are a sound engineer and don't believe this, I would suggest you rethink your career choice.

Consider the following: we have been handed the sound design contract for the Rogers and Hammerstein's musical *Carousel.* The production will be performed at the 1600-seat Majestic Theater which is highly rated for its excellent acoustics. This is the very same Broadway stage of *Carousel's* 1945 debut. The orchestra will have the same assembly of instruments in the same pit. Costumes and sets will be done in keeping with the traditional look of the show. There will be no ocean of noise coming from moving lights. This will be the revival to end all revivals. How about the sound system? If we believe that natural sound is best, then we would be obliged to resign the project. There is no need for an amplification system. The original show was composed and staged for natural sound. How could we presume to improve upon this? The following problems arise: the director hates it, the performers hate it, the audience hates it, and the critics hate it. The show closes in three nights and we will never work on Broadway again. Why? What has changed? Our expectations. Audiences today do not expect natural sound. We expect *magic* sound. Sound that arrives at our ears without the slightest effort required on our parts. We are sonic "couch potatoes." Get used to it. This is not just a fad. This is not just the audience either. The performers also want magic sound. They want to whisper, and yet hear themselves on stage and be heard at the back. That takes strong magic, but this is our job. We will not get the job done if we don't get realistic about the prominence of "unnatural sound."

Why does this matter? Am I just trying to provoke a fight? No. Acousticians, please put down your acoustic pistols. Our work toward the optimized design cannot move forward if we are working under an inapplicable construct. Submitting to the collective group-think of natural sound superiority does nothing to advance our position. When it comes to the sonic experience of audience members at a pop music concert, the sound system performance will be the decisive factor, as long as the acousticians (or whoever was responsible for the building design) have not created a hall totally unsuitable for us. A successful outcome is much

Perspectives: Most of my work has been doing what I call "critical events," those events where very good accuracy and/or intelligibility is required, like classical music or the spoken word. My theory is that if what people hear at the beginning of an event sounds as they expect it to sound; they will be quicker to forget they are listening to a reinforcement system and connect with the event itself. 1 call the goal "not offending the listeners" and over the years it seems to have worked well for me.

Alexander Yuill-Thornton II (Thorny)

more likely if we can express our acoustic requirements realistically, i.e. based on our need for amplified sound.

Contrasting Emission, Transmission and Reception Models

Natural sound transmission requires three aspects to be in harmony: the sound source emission, the stage transmission and the hall transmission. Only then can the listeners obtain satisfactory reception. The sound begins at the individual musicians on stage and travels directly to the audience reception area. It is accompanied by reflections from the stage and the house which mixes the sound together and provides tonal enhancement and increased loudness. The active transmission support of the stage and house are absolutely required for both the performers and the audience. It is as if the audience is on the stage, and the musicians are in the house. There is no acoustical separation between them. All aural reception for musicians and listeners occurs inside the sound transmitter: the room. There is a continuity of the sound over the space and over time. The room has a sound that goes beyond a particular composer or conductor. The room defines the sound and the sound defines the room.

Amplified sound applications also require three-part harmony. But it is a very different tune. Each musical emission source is captured by nearby receptors (microphones) before leaving the stage. The receptors enable the isolation of the sources from the audience, and from each other. These are individually electronically transmitted to the mixing console where the blending and tonal content of the sources is managed. The mixed sources are then sent to distinct and separate locations. Some are sent on stage, in whatever form is most suitable for the musicians to supplement their local needs. The signals are also sent to the audience via house speaker systems that provide controlled coverage to localized zones. The separation of these transmissions into isolated zones allows for each area of the audience to have matched sound, even though they are different distances from the stage. In amplified sound applications there is not *one* unified source of house sound, but rather locally distributed versions of the sound. On stage they are separated in order to fulfill the unique needs of the performers. In the house they are separated to fulfill the common needs of audience members that are placed at unique locations.

The separation of all stage emission sources allows us to capture everything that happens on stage and transmit it anywhere. The separation of all stage and house transmission sources allows us to send unmatched signals to unmatched areas to provide matched results for the listeners. Recall that with natural sound, unity of experience is achieved by a *total lack* of separation. By contrast, with amplified sound, this is achieved by an *abundance* of separation. In amplified sound there is a continuity of the sound over the space for a given performance but this sonic character does not persist over time. The local arena does not have the same sound for a rock concert on Tuesday, basketball on Thursday or the circus that plays over the weekend. The sound system may change from the house PA to a touring rig, the operator may change from night to night or simply the mix is changed. Any of these will be perceived by our audience as changing "the sound." The room is not defined by the sound and the sound is not defined by the room.

Natural emission/transmission/reception features:

- Individual sources are mixed on stage.
- Emission originates from multiple sources in one general location: the stage.
- Transmission support by the room is required.
- Separation between the emission sources and the transmission medium is prohibited. Musician, stage and house are all one.
- Reception becomes uniform in level and tone by lack of separation.

Amplified sound emission/transmission/reception features:

- Individual emission sources are isolated on stage, and the stage is isolated from the main transmission system.
- Transmission originates from multiple locations: the stage, the stage monitors, house main speakers, house auxiliary speakers.

Perspectives: One of my greatest ever audio experiences mis to hear the performers at Preservation Hall in New Orleans perform, and I went back every night of my trip there. The irony is that it's not a reinforced show — there isn't a speaker or amplifier in the room — but I learned more about what reinforcement and recordings should sound like from those nights than from a thousand CDs and hundreds of live concerts. The best players in the world, mixing themselves to each other five feet away from you. It would be impossible to recreate that sound without an experience of perfect acoustic sound.

Martin Carillo

- Transmission support by the room is an optional enhancement, not a requirement.
- Separation between the emission and transmission sources is required. Musician, stage and house are all separate.
- Reception becomes uniform in level and tone by virtue of separation.

Notice that the role of the room and stage acoustics were the prominent factors in the natural scenario and not even mentioned in reference to the sound reinforcement. This is not to say that these factors will not have decisive effects on the outcome. They most certainly will. But our needs regarding the room are so different that our thinking in this regard must be substantially restructured. The room is not required for level or tonal uniformity of our signal. In fact, strong early reflections, the centerpiece of classical architectural acoustics, are highly counterproductive to that effort in our case. The room acoustics will serve us best as a spatial enhancement effect rather than a power transmission aid. For our purposes we will need only low-level, diffuse reflections as a spatial enhancement. Otherwise we would prefer to handle the matter ourselves and ask the room to politely step aside. In the classical world the rooms have a sound. In the sound reinforcement world neither the room nor the speakers have a sound. The artists have a sound. When a comedian is booked at one of the world's great symphony halls, people do not come for the sound. They come for the jokes. The management will have to come to terms with the reality that the "perfect acoustics" of their symphony hall are imperfect for amplified sound applications. They need a sound system that can transmit into the complex irregular shapes of the hall and they need absorption. Lots of it. If not, there are going to be a lot of refund requests.

Transmission Path Differences

The decisive difference between the natural sound model and the speaker system model is found in analyzing their respective transmission paths. The natural sound path begins with emission from the instruments. These begin their transmission on stage and maintain a continuous path that ends with reception in the audience. The room's role is transmission enhancement. The speaker system model begins in the same way but the emission immediately arrives at a reception device: the microphone, which delivers a version of the emission source to the sound system mix console. The combined microphone signals of all stage sources are sent to a separate transmission device, the speaker, which carries out the delivery to the audience. The extent to which the sound system enjoys exclusive broadcast rights to the audience will depend upon how much of the original emissions leak from the stage into the audience area. For the moment we will consider the stage levels to be low enough to be considered negligible in the room. This breach in the transmission line created by the open receptors (microphones) allows for two forms of leakage between the transmission systems. The first is re-entry of the speaker transmission back through the mic. In all cases the re-entry is later than the original source, and combines with the summation related frequency response effects described in Chapter 2. In the most extreme case the transmission line leakage into the receivers that drive it becomes so great that the system becomes unstable, resulting in feedback. The second form of leakage is duplicate entry. This results from a single on-stage emission source arriving at two open mics. The separate arrivals are not synchronized in time and again the usual summation effects occur. The hazards of feedback and multiple open mics are well known to audio engineers so it may seem strange to characterize them with terms such as re-entry or duplicate entry summation. These terms are used here to illustrate that their *perceived* effect on the system response is the same as the effects of speaker/speaker interaction and speaker/room summation discussed previously in Chapter 2. The tonal distortion from 1 ms comb filtering will have the same frequency content whether it comes from the summation of two speakers in the room or two microphones at the console.

There is a multiplication factor at work here which will have very audible consequences: any summations that are brought in to the sound system through the microphones are present in the *source* signal to the transmission system. When the time offset signals from two summed microphones leave our speakers and hit the wall, it is the equivalent of two reflections of the original signal. Even if our speaker system is in an anechoic chamber the summation of a single source to multiple microphones will create the perception of reflected sound, albeit without any spatial aspect.

There are five principal avenues of summation in the relationships between the stage and the audience: source/room (the interaction of the stage emission sources with the room/stage), source/speaker (the interaction of the stage emission sources with their amplified signal), mic/mic (the summation of open microphone leakage at the console), speaker/speaker and speaker/room. All cause the same perceived tonal coloration effects. The echo and spatial perception effects have the same timing considerations but differ in their directional aspects. Each mechanism requires different management strategies. Our listening experience consists of, for lack of a better term, the summation of the summations. Discerning which of these mechanisms is responsible for a particular sonic artifact will take great skill in either/both our perception and analysis of measurements.

The natural sound transmission has only one of the five avenues of summation: source to room interaction. The amplified sound system has them all. The totality of the summation effects must be comparable for both natural sound and amplified sound in order for listeners to perceive an equivalent experience. The natural sound model provides it all with the room reflections. Since the amplified sound model will have additional sources, it must decrease the role of the room in proportion to the other summation avenues.

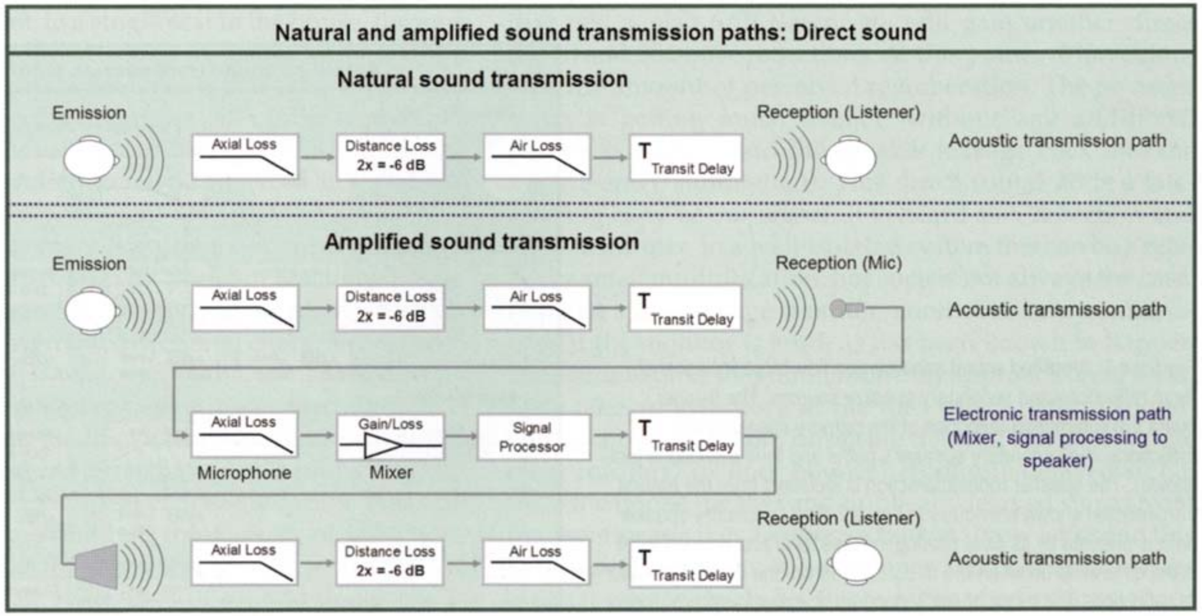


Figure 4.1 Direct sound transmission flow block for "natural" sound and "amplified" sound. The natural sound path is potentially filtered by the axial control of the source and also by the HF air loss. Level is lost at the rate of 6dB per doubling distance along the transmission path to the listener. The amplified sound has two acoustic paths with similar features. In the middle is the electronic path, which has the ability to compensate for the effects of the acoustic transmission

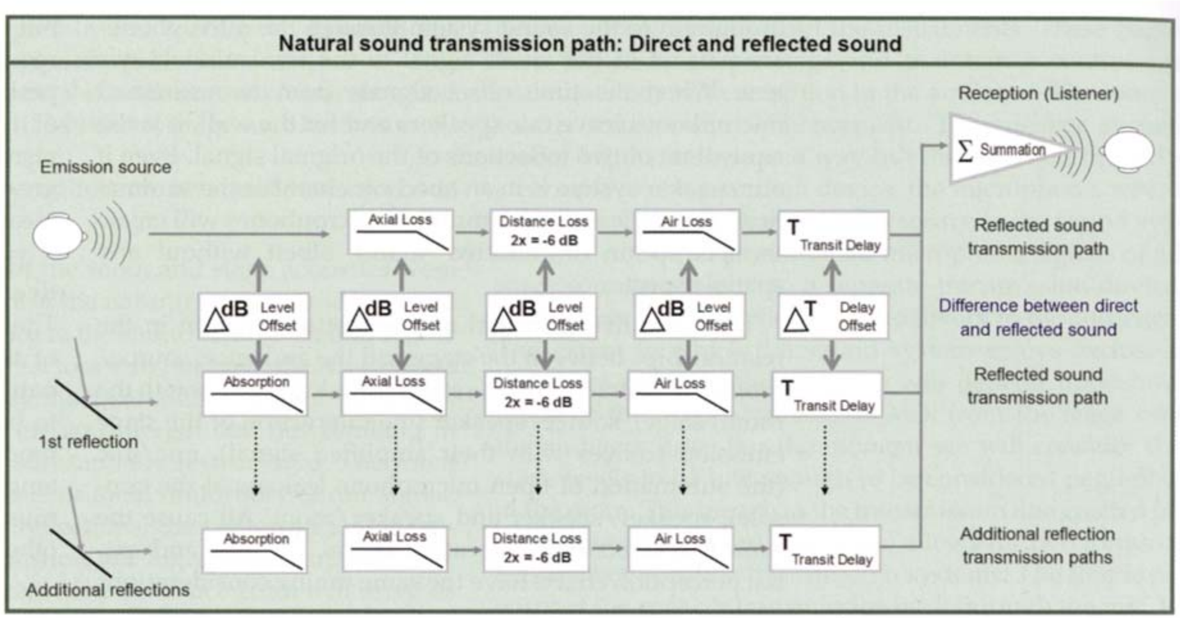


Figure **4.2** Natural sound transmission flow block inclusive of room reflections. The listener hears the summation response of the direct sound path and reflections. The reflections have distinct axial and HF air filter functions, as well as distance loss and transit time. It is the difference between the direct sound and the reflections in each of these categories that will be the decisive factor in the listener's experience

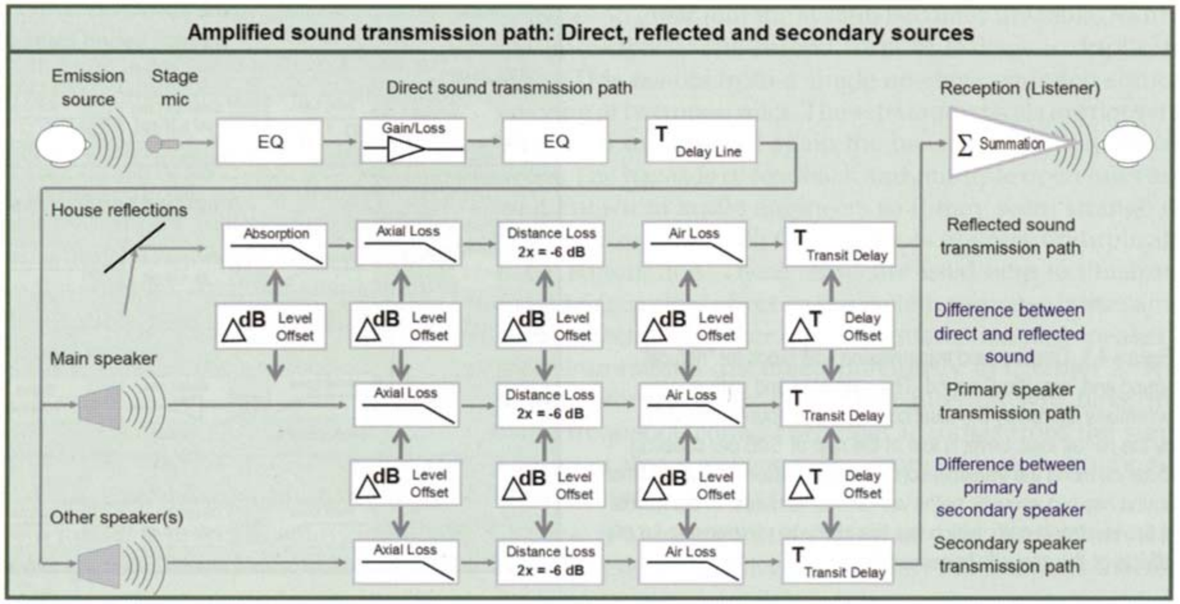


Figure 4.3 Amplified sound transmission flow block inclusive of room reflections and secondary speaker sources. The listener hears the summation response of the primary speaker path, its reflections, the secondary speaker's paths and their reflections (not shown). The speaker room/interaction is indistinct from the natural transmission shown previously in Fig. 4.2. The secondary speaker path is affected by its axial filtering, relative level and time offset. It is the difference between the direct sound from the primary speaker, its reflections and those of the secondary speaker in each of these categories that will be the decisive factor in the listener's experience

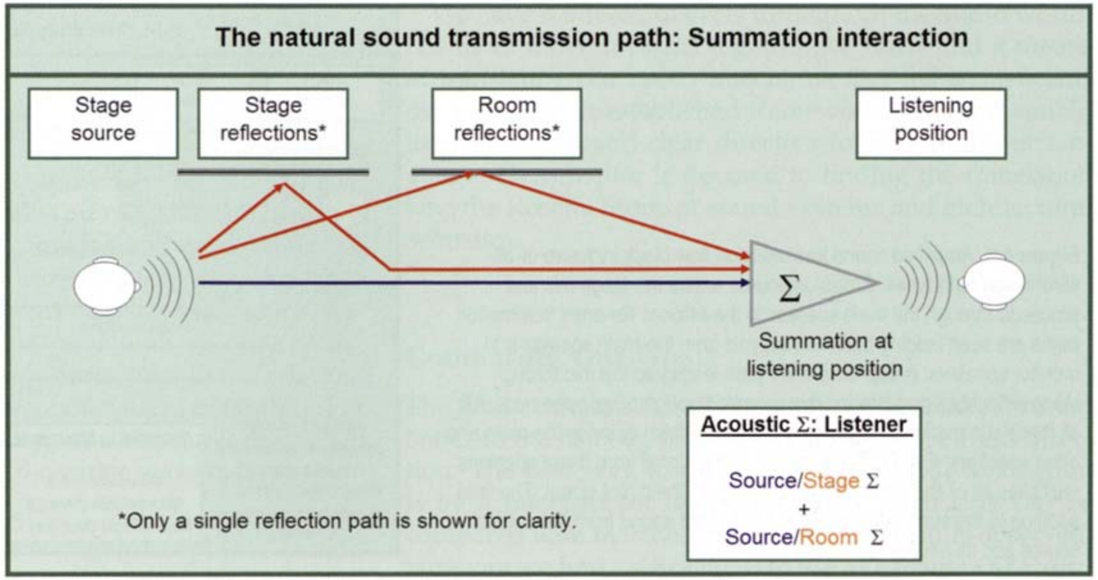


Figure 4.4 Natural sound transmission flow block diagram inclusive of all summation path types. Direct sound from the emission source is summed with stage and house reflections. No other paths exist for the single source, although it may be accompanied by other sources that duplicate its signal, e.g. multiple violins or choir members

Let's follow the transmission path of a single emitter on stage, the snare drum, to a single seat in the house, through the natural and amplified sound systems. In the natural model, the drum is struck and the impulse travels directly to the listener and indirectly (reflected) off the walls. Let's arbitrarily attribute a number of 100 reflection paths in the duration of time required to fall 60 dB. Now let's place an open mic close to the snare and feed it into a speaker. To simplify our example we will isolate the snare drum so that its natural sound emission path does not reach the house. The listener only hears the speaker. The speaker, which for simplicity will be given the directional characteristic such that it also creates 100 reflection paths, transmits directly to the listener. So far we are perfectly matched. If there is a floor reflection into the mic (a duplicate entry) then this signal will be transmitted as well. We now have two "direct sound" arrivals from the speaker. If the reflection was 6 dB down we can assume that half the number of reflections would be heard over the same time period as before. We are up to 150 audible reflections. If an additional mic is placed on stage (a second form of duplicate entry), the snare drum will also enter it. If the snare drum signal entering that mic is also 6 dB down, we will gain another direct signal and 50 house reflections. At this point we have doubled the amount of perceived reverberation. The *perceived* room is getting much livelier, without any additional plaster. Now let's add the speaker leakage back into the mic (re-entry summation). This direct sound adds a late-arriving copy of our signal. It is heard as a reflection and adds to the mix. In a well-isolated system this can be a relatively small multiplication, but such is not always the case. Now let's add a stage monitor, another re-entry summation. If the monitor is loud, as has been known to happen from time to time, the summation may approach unity level. Let's take an inventory at the mic: direct natural sound, natural floor reflection, duplicate delayed signal from the other mic (and its floor bounce), duplicate delayed signal from the house speaker and duplicate delayed signal from the monitor speaker. Getting dizzy yet? Now let's send this to the speaker and out to our listener. It is not hard to imagine that we can create 5X the number of reflections at our listening position without any effort at all. Now imagine

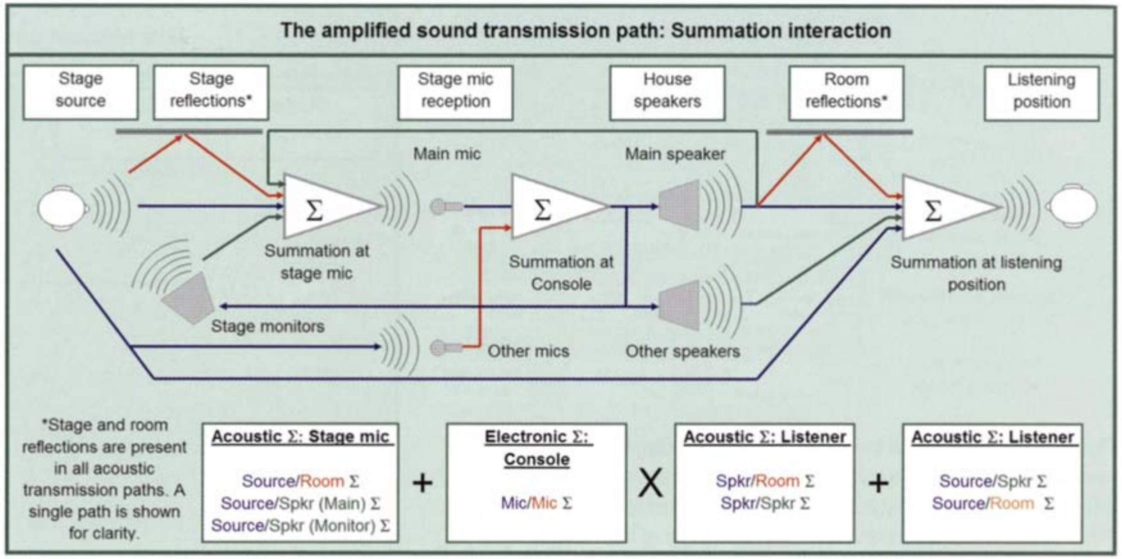


Figure 4.5 Amplified sound transmission flow block inclusive of all summation path types. Emission source enters the stage mic and proceeds through the main speaker to the listener. Re-entry summation paths are seen leading back into the mic from the main speaker and monitor speakers. A duplicate entry path leads into the mic through stage reflections and into the mix console through other open mics. All of these summations are present in the waveform *entering* the main and other speakers. Each room reflection of the signal from these speakers multiplies all of the summations contained in the direct signal. The final addition to the summation is the natural direct sound from the original source and its natural reflections

more open mics, more stage reflection leakage from side and rear walls and more stage monitors and we are potentially drowning in reverberation. Only a fraction of this is actually generated by the main speaker's excitation of the room. The only thing left to do now is for the band mixer to add some electronic reverberation!

The openings in our transmission line change the entire perspective about the room acoustics. Every re-entry and duplication at the mics will serve as a multiplier of the room's reflection pattern. In natural acoustics reflections can only be accrued by *addition.* In amplified sound they are accrued by both addition and multiplication.

Equally troubling is the fact that once the re-entry or duplicate leakage has summed into our transmission line it cannot be removed. The original emission signal entering the mic may bear only vague tonal resemblance to the waveform traveling down the wire due to the combing effects of the leakage. These pre-transmission defects in the signal will be plain for all to hear since they are embedded into the input signal traveling to our speakers and will be distributed to all seats.

I had the personal experience of optimizing the large-scale sound reinforcement systems for opera megastar Luciano Pavarotti. One might think that the biggest challenge to the achievement of "perceived" natural sound would be the acoustical properties of sports arenas. The far greater challenge was the stage monitor leakage into the singer's mic. When the monitor levels rose too high we were left to transmit a vocal channel that was in critical condition before it left the stage.

The implications of this should be obvious by now. Every effort should be made on our part to reduce the leakage into our mics. Whenever possible, we will use directional mics, close placement, noise gates, acoustic baffles and absorbing devices, in-ear monitors and directional main speakers. At the end of the day we will still have lots of leakage. Our mixed transmission signal for the speakers is already preconditioned with an elaborate reflection pattern, which will be multiplied by the walls. The conclusion is that our amplified sound system will need much less of the room reflections in order to create the same *perceived* final response as the natural system.

Relating the Worlds of Acousticians and Audio Engineers

Let's examine the relationship of the disciplines of acousticians and audio engineers. In the modern world we are doing pop music concerts in symphony halls and Pavarotti concerts in sports arenas. The success of our work is dependent upon theirs, and likewise their work is increasingly dependent upon ours. Almost every listening space designed or remodeled in the future will have a sound system. The public will evaluate the results on their perception of the *combined* effects of our efforts. Functional relationships are built upon mutual respect and understanding. For disciplines to interconnect they must be familiar enough with each other's roles to find creative ways to implement mutual solutions.

Architectural acoustics is a highly respected profession. Sound system engineering is not. Any doubts on this matter can be clarified at a cocktail party by alternately introducing yourself as an "acoustician" and as a "sound person." (We really can't even call ourselves "engineers" since there is no such recognized degree or licensing.)

Respect is not freely given. It is earned. The acousticians were in the room at least 400 years before us. The first step to gaining their respect is to clarify our needs in terms that they can relate to. This begins with a frank assessment of our position. What we will find is that while our roles are separated, our goals are unified. As illustrated above, the means to achieve them are almost polar opposites, resulting in a substantial bridge to cross. Acousticians have a long-established language for evaluating the sonic experience. Historically, many of us have deferred to the language of classical acoustics and tried our best to fit our audio planet in the acoustician's universe. Many of the terms that give the clearest directives for them leave us blank. The difference between a reverb time of 1.5 and 1.7 seconds tells an acoustician which of the two halls is more suitable for opera. The distinction is meaningless to audio engineers. Neither hall is suitable for us. In either case we will be in a fully defensive design mode to keep the direct sound from our speakers off the walls as much as possible.

Perspectives: Ears are what need to he satisfied. They are the final check on what has been done. No matter how good the measurements look, if it sounds bad something is wrong.

Alexander Yuill-Thornton II (Thorny)

We have scattered dialects throughout the audio world. For us to move forward together we must find a means of translating our needs into terms that the acousticians can fit into their established framework and yet maintain their meaning, and clear directive for action, in our language. This chapter is devoted to finding the translation key: the Rosetta Stone of sound systems and architectural acoustics.

Comparing Our Goals

The final evaluator is the listener. If we can make it sound better to the listener, then we are going in the right direction. The best way to make any side-by-side evaluation is by a matched set of specific goals. Each side can be compared with matched metrics and the results observed. How can we find a set of goals to use as a reference? Fortunately, one was provided for us by the esteemed acoustician Leo Beranek, when in 1962 he published *Music Acoustics & Architecture.* This groundbreaking book originated a set of evaluative criteria for concert hall performance. Beranek visited halls, measured their acoustic performance and interviewed conductors and critics. The statistical data of 54 concert halls was compiled and a framework created which scored the hall performance in different categories. Concert halls which were constructed hundreds of years ago were evaluated side by side with those of the recent past, by the same measures. The conclusion at the time was that no single criteria could be found which would make a hall great, while certain key trends were required for success. The key factors were prioritized by degree of effect, since no two halls had identical parameters.

Eighteen categories of subjective perception were considered. Some of the categories were given numerical scoring weights while others were found to be too interdependent to be assigned discrete values. These subjective measures were then compared with objective measures such as the room volume, reverb time, etc. The study correlated the subjective with the objective data and the result was a comprehensive assessment of the hall's physical parameters to the listener's experience. From that point forward architectural acoustics moved ahead on a foundation of scientific data. It continues to refine the process further, even now.

All of the eighteen categories remain relevant to our current-day listening experience, whether it is natural sound or through speakers. This is not surprising, since little has changed in the way of human anatomy in the last century. Therefore these metrics can be employed by audio engineers as well. If we can achieve the same subjective effect to a blindfolded listener, then we can all agree we have achieved an equivalent sonic experience. The objective means for acousticians and audio engineers to achieve matched subjective results are very different. This is the core of the issue. If we can relate these different means to achieving our common goals we will have found the translation guide we are seeking.

Let's begin by meeting Beranek's subjective parameters (Beranek, 1962, pp. 61-71). These parameters are paraphrased and stripped of any specific reference to the symphonic experience. What follows are the 18 attributes described in terms that are independent of musical genre, venue or sound transmission method.

1. Intimacy: this refers to the sonic perspective of the listeners. The desire is for listeners to experience a feeling of proximity to the music, as if we were listening in a small room. A lack of intimacy corresponds to feeling of distance and separation as in a large room.
2. Liveness: this is experienced as a fullness of tone in the mid-range and high frequencies.
3. Warmth: a fullness of tone in the low frequencies.
4. Loudness of the direct sound: the desire is for the loudness to be appropriately scaled to the musical content. If too loud the experience is unpleasant, if too low the experience lacks the desired impact.
5. Loudness of the reverberant sound: the desired effect is for the reverberation to have the appropriate mix of level and duration to provide additional loudness to the direct signal. An insufficient quantity will result in the overall experience lacking loudness, while excessive amounts will cause the opposite effect.
6. Definition, clarity: a clear and distinct sound.
7. Brilliance: bright, clear ringing sound, rich in harmonics.
8. Diffusion: the spatial aspect of reverberation is found here. Diffusion of the sound creates the experience of sound arriving from all directions.
9. Balance: this factor evaluates the relative levels of the instruments and voice. Good balance entails the instruments heard in their proper level perspectives. Poor balance is found when some instruments are favored over others.
10. Blend: good blending is perceived as a harmonious mix of the instruments.
11. Ensemble: this concerns how well the musicians can hear themselves. Good ensemble is obtained when the musicians can hear themselves well.
12. Immediacy of response: this is a measure of how well the musician's feel about the responsiveness of the sound. The goal is for the musicians to feel the sound and be able to adapt quickly enough to their changes so as not to disrupt them.
13. Texture: the fine grain of the listening experience. Texture is described in similar terms to the sense of touch. Music with fine texture has a richness and complexity to its outer surface.
14. Freedom from echo: the desired effect is that we do not hear discrete echoes.
15. Freedom from noise: the lowest amount of noise is desired.
16. Dynamic range: this is the range between the maximum level and the noise. The maximum level would be limited by comfort levels and the minimum by the ambient noise.
17. Tonal quality: rich tonal quality is free from the distortions of peaks and dips in the response over frequency. Poor tonal quality has an uneven frequency response that may cause certain notes to be lost and others to be unintentionally accentuated.
18. Uniformity: the extent to which we can create a similar experience for all listeners in the hall.

The first matter to consider in cross-referencing the perceptions of natural acoustics and speaker systems is the applicability of the acoustician's standard objective measures. Rooms are given single numbers to describe their prominent characteristics. These include volume, surface area, absorption coefficient, reverb time, initial time delay gap, bass ratio and others. Rooms are viewed as a single entity, with uniformity enforced by the overwhelming complexity of the reflection patterns. The lone exceptions to the singular referencing are under balcony spaces, which are seen as distinct adjoining spaces.

There are limits to the meaning of the numbers. These statistics are only applicable to describing the room as it reacts to a particular type of sound propagation (omnidirectional) from a single specified location (the stage). If the orchestra was moved to the side of the proscenium and had baffles added which restricted the propagation pattern to 30 degrees, it is obvious that only the physical properties of the hall such as volume and surface area would remain constant. The acoustic properties would need to be fully reevaluated. An example reverb time for a room could be more fully stated as 1.9 seconds with omnidirectional propagation from the stage. This qualification is self-evident to acousticians, but not so with audio engineers, who often tend to think of the reverb time as a fixed entity in spite of the fact that our sound system has direct control over decisive parameters in the reverb time equation. The relevance of this distinction should also be self-evident.

The speaker systems we use do not propagate omnidirectionally or symmetrically in some cases. They are not located on the stage and are not restricted to a single position. If we use highly directional speakers that focus the direct sound or add remote speakers that increase proximity to the listeners we will increase the direct-to-reverberant ratio. This technique reduces the reverb time in the speaker's on-axis area from what would be found with an omnidirectional propagation source. The decay will fall 60 dB more quickly because the loudest portions of direct sound propagation are more locally focused and less energy falls on the reflective surfaces. The off-axis areas see the reverse effect from their perspective. The reverberation time is extended, because the direct sound propagation is weak compared to the reflections. The room surface properties have not changed, but the combined acoustic properties of the room and the propagation source have been changed. This does not exempt speakers from the room's acoustic qualities; it simply adds the complexity of multiple perspectives. It means that we have to approach the room as it exists from the perspective of each individual speaker, and resist all-in-one numerical ratings that assume a singular point of omnidirectional transmission. Our under balcony speakers see a very different room than that of the mains. Therefore a separate evaluation is required. Recall that the same separation was noted for the natural acoustics model, lending precedence to separation in our realm.

The second reason requiring evaluation of each speaker/room relatio assigned to sections of the room, not the entirety. The natural acoustic model assumes the stage source to be omnipresent, i.e. responsible to reach every seat in the house. Our speakers divide the responsibility of coverage into different zones, in essence slicing the room into a series of adjoining rooms, each of which will be separately evaluated. The shapes, quantities and locations of the acoustical partitions are under our control, and are one of our most critical design decisions. They are the product of the propagation properties, location and relative level of the speaker subsystems. If the symphony was divided up and sent into the house to play identical parts in the different audience areas the same rules would apply.

An example should bring the point home: a frontfill speaker located on the center of the stage lip. No other speaker in our design more closely approximates the natural source. This speaker sees the room from the stage floor perspective and in common practice has a wide propagation pattern. Its direct sound can most likely reach every seat in the hall since all seats have line of sight to the stage lip. And yet the room acoustics are less relevant for this speaker than any other. This is because its transmission will travel no more than four seats deep and to the sides in the audience before it has met its neighboring speakers and hands over the coverage responsibilities to them. This is the acoustic partition. The partitioning does not prevent the frontfill speaker from reaching the walls and exciting the full range of reflection patterns. It does, however, make the reflections irrelevant by burying them miles deep under the stronger direct sound of the other systems. This brings us to the first conclusion that all evaluations of loudspeaker systems in the hall must view the room as a partitioned entity, with distinct local interaction with the speakers. Second, we conclude that the overall uniformity of experience in the hall will be a direct result of our ability to match the partitioned areas to each other, and to minimize the disturbance in the transition areas. This fundamental principal of the optimized design stands in stark contrast to the natural acoustic model.

The natural and amplified sound models seek a uniformity of experience, our common goal. The means to achieve such uniformity are often diametrically opposed. The natural model uses the summation properties of the room reflections to create uniformity. Uniformity of level over distance is accomplished by coupling zone, combing zone and combining zone summation (see Chapter 2). Uniformity of frequency response is accomplished by combing and combining zone summation of such massive density and complexity that the ear is unable to distinguish the frequency response details at one seat from another. Isolation zone summation is employed to create the gradual decay of the reverberation tail evenly over the space.

The speaker system model creates uniformity over distance by a variety of methods that rely principally upon isolation zone summation. The use of directional sources and the partitioning into isolated zones allow us to compensate for the differences in distance by selective level settings. Uniformity of frequency response is also accomplished by isolation zone summation and by minimizing the tonal variations over frequency and position that would result from combing zone summation. Isolation zone summation is likewise employed to create the gradual decay of the reverberation tail evenly over the space, our shared parameter. The speaker system model uses the coupling zone summation very sparingly. It is useful only in those parts of the frequency range (mostly the lows) and those locations (the partitions) where it can be employed with minimal effects from the combing zone. The acoustic partitions must be properly joined as phase-aligned spatial crossovers, thereby minimizing the transitional disturbance.

The itemized comparison of the 18 perception parameters is shown in Fig. 4.6. This chart details the differences in means required to approach the shared goals. An examination of the chart reveals the following trends:

- Freedom from echoes, noise and the maximization of dynamic range produced strong agreement.
- Diffusion is an architectural feature whose role is identical in both domains. Strong agreement was found in the means of achieving diffusion and the recognition of its value.
- Warmth and brilliance fall into the category of equalization. The means to achieve the equalization are vastly different but not oppositional.
- Balance, blend and texture fall into the category of mix engineering. The means to achieve the mix are vastly different but not in the least oppositional.
- The loudness of the direct sound falls into the category of source placement. The means to achieve equal loudness are different but not oppositional.
- Liveness, loudness of the reverberation and clarity fall into the category of decay-related architectural acoustics. These are somewhat opposed in that the room reverberation needs of the natural model are mandatory, while amplified sound has electronic enhancements available. The conflict arises when the reflection levels are too high for the amplified sound system.
- The response/attack on stage has somewhat conflicting factors. The outcome for amplified sound will be a compromise position determined primarily by the artist.
- Ensemble on stage calls for diametrically opposed approaches. The natural system requires the maximum coupling of the stage sources to each other and the house. Our systems require the maximum isolation.
- Intimacy is achieved by oppositional means. The strong reflections that create the feeling of an intimate space

Subjective attributes of musical-acoustic quality: A comparison of natural and amplified sound					
Attribute	Natural Acoustics (paraphrased from Beranek)		Comparison	Amplified Sound	
	Description	Transmission/Summation/Reception	Means	Transmission/Summation/Reception	Description
Intimacy	The feeling of being in an intimate space with the musicians is achieved by ensuring a pattern of strong early reflections within the window known as the initial time delay gap. The early arrivals should be within 5dB level offset and 20 ms time offset of the direct sound.	Strong reflection summations are needed for acoustic power addition. Multiple arrivals are required to provide sufficient comb filter density so that the peaks and dips are beyond the tonal resolution of the ear. This is aided by the fact that each seat is a different distance from every instrument. Therefore the summation frequency response signature does not repeat itself for any two instruments at any two seats.	Strongly Opposed	Strong reflection summations are not needed for acoustic power addition. Since the speaker is a single directional source the reflection pattern lacks complexity, and the combing summation will be perceived as a tonal distortion. The sounds of multiple instruments are combined in the mixer and share a common path from the speaker to the listener. Therefore the comb signature repeats itself for all instruments in that path.	Power requirements are satisfied by using high power speakers or overlapping speaker arrays. Intimacy is achieved by separation of the sources into local zones with controlled directionality. The proximity to the sources and a high ratio of direct sound provides a close sonic perspective to the performers.
Liveness	The optimal reverb time is program material dependent. It can be considered optimal when the reverberation is long enough to connect the music together without becoming so long that it restricts its ability to change. Symphonic music ranges from 1.5 to 2.2 seconds while opera ranges from 1.5 to 1.7 seconds.	Reverberation is highly dense Combining Zone and Isolation Zone summation. The density neutralizes the tonal signature and provides a desirable sense of spaciousness. Each seat in the hall has a unique reverberation pattern, but all are beyond the tonal signature range and therefore perceived similarly.	Somewhat opposed	The separation of the room into distinct speaker coverage areas breaks the unifying element of the room reverberation. The directional nature of the speakers and their local distribution increases their isolation from the room, thereby reducing the reflection density. Therefore, the reverberation is likely to be substantially less uniform and the probability of perceptible echoes rises substantially.	The room reverberation is best when it meets our minimum needs and principally provides spatial enhancement. The bulk of the reverberation can be supplied electronically, which has the following advantages: it can be applied selectively for different program materials and for different instruments. This provides the mixer with an electronic control for "variable electro-acoustics".
Warmth	Warmth is achieved by ensuring that the reverberation time in the low frequencies is around 25% longer than that of the midrange and highs.	Reflective surfaces should be of sufficient density and hardness to reflect low frequencies and ensure that their reverb time is long enough. Plaster and thick wood are the recommended materials.	Somewhat opposed	Coupled speaker (and subwoofer) arrays and nearby surfaces provide conditions for uniform low frequency summation. More distant surfaces will also increase the low frequency power, but at the expense of lost uniformity and, most importantly, reduced transient decay. This reduction results in a loss of "tightness" in the LF range.	Warmth can be achieved by having sufficient low frequency power capability, directional control of the array and equalization without need of the LF reverb time being in excess of the midrange.
Loudness of direct sound	Sufficient loudness is maintained by arranging the seating as close to the conductor as possible. 60 ft is seen as the benchmark.	Direct sound level is a function of transmission path length from stage to listener	Different but not conflicting	Direct sound level at the listener is a function of transmission path length, drive level of the speaker, and its axial orientation.	Directional speakers in separate locations allows us to maintain constant level over the space. High power capability brings listeners to any virtual distance.
Loudness of the reverberant sound	The level of the reverberant sound must scale with the venue size. Small rooms should have lower reverb times, while larger rooms need longer decay.	Strong extremely dense Combing Zone, Combining Zone and Isolation Zone summation is required.	Somewhat opposed	The ratio of the direct to reverberant sound should be kept as high as possible. This is achieved by minimal Combing Zone and Combining Zone summation. Isolation zone summation is accepted within limits for its ability to provide the sense of spaciousness.	The scaling of direct sound to reverberation need not be a fixed ratio, but may remain variable to accommodate program material changes. The appropriate ratio for speech or fast paced music is highly in favor of direct sound, while a slow ballad requires high levels of reverb. With minimal amounts from the room, we can add reverberation as needed.
Definition, clarity	Clarity is achieved by the optimal mix of intimacy, liveliness and loudness	A mix of direct sound, dense Combining Zone and Isolation Zone summation is required.	Somewhat opposed	Clarity is the result of a high ratio of Coupling and Isolation zone summation, and a low ratio of the Combing Zone.	Clarity is achieved by directional speakers with distinct coverage zones and carefully managed partitions.
Brilliance	Brilliance is achieved when the sound is intimate and the reverb time in the high frequency is properly balanced to the midrange.	A mix of direct sound, dense Combing Zone and Isolation Zone summation is required.	Different but not conflicting	Distant speakers will have high frequency boost to replace transmission distance HF air loss.	Separated directional speakers allow for even amounts of high frequency distribution through the hall.
Diffusion	Reflections provide a richer reverberation character if the surfaces are diffuse. This scattering provides a more gradual and dense reverberation tail.	The summation arrivals become more randomized, leading to more texture, a steady decay character and less risk of echo perception.	Strongly Agreed	The summation arrivals become more randomized, leading to more texture, a steady decay character and less risk of echo perception.	Reflections provide a richer reverberation character if the surfaces are diffuse. This scattering provides a more gradual and dense reverberation tail.

Figure 4 6 Comparison of natural and amplified sound models based on Beranek's evaluative criteria for subjective evaluation of concert hall sound. The same subjective criteria are used, since they are based upon our listening experience, not the mode of transmission. The means of optimizing transmission of that experience is compared and contrasted between the natural and amplified systems. The extent to which the means to achieve a comparable experience are in agreement or opposition is detailed. The pivotal role of summation is examined in each case, and is the most cited area of conflict between the approaches

Subjective attributes of musical-acoustic quality: A comparison of natural and amplified sound

Attribute	Natural Acoustics (paraphrased from Beranek)		Comparison	Amplified Sound	
	Description	Transmission/Summation/Reception	Means	Transmission/Summation/Reception	Description
Balance	The room must transmit all instruments in proper perspective in the mix. The stage and house reflections must have sufficient strength and density for uniform level and frequency response to be achieved. e.g. if the room lacks warmth the basses will not be balanced in the mix.	Each instrument has distinct transmission level, frequency range and point of origin. Balancing of the instruments is in the acoustic space.	Different but not conflicting	The role of the speaker system is transmission of premixed signals originating at the mix console outputs.	Stage sources are isolated and balanced in the mix electronically. The room does not provide a mixing function unless the sources are transmitted from separate speakers.
Blend	A harmonious blend is achieved by the positioning of the instruments. The spacing between the instruments on stage is a key factor.	Each instrument has distinct transmission point origin. Mixing of the instruments is in the acoustic space.	Different but not conflicting	The role of the speaker system is transmission of premixed signals originating at the mix console outputs.	Instruments are blended by separation into multiple transmission channels (left, right, etc.). Isolated electrical signals are blended in the acoustic space.
Ensemble:	A sense of ensemble on stage requires the conditions of acoustical intimacy described above. Musicians must feel like they are playing together in a small room rather than spread over a wide area, unable to maintain contact. This is achieved strong early reflections that come within an initial time delay gap under 20 ms.	Strong and dense Combing, Combining and Isolation Zone summation. The sound of each instrument must spread across the stage to the locations of other musicans. Instruments shall not be isolated from each other.	Strongly Opposed	Stage monitor speakers in proximity to microphones have the opportunity to reenter the transmission path. This results in Combing Zone summation and possible noise generation (feedback). Stage monitors also provide leakage into the house and disturb the acoustic partitioning of the house speaker system zones.	Ensemble on stage is achieved by isolating the individual sound sources, combining electronically in the mixer and returning the signal via individually adjustable stage monitor systems.
Response, attack	Musicians will feel the room is responsive dynamically if the loudness of the reverberation on stage is correctly scaled for them. If the reverb time is too long the room restricts their dynamic changes. If too low the music lacks a continuity between notes.	The stage is directly coupled to the house. How the transmission character of the house returns to the stage will be the decisive factor here.	Somewhat opposed	Monitors are the principal reference source on stage, yet the sound in the house will have an appreciable effect on their experience. Two principal sources dominate: the sidelobes of the sound system which can have objectionable off axis tonal quality, and the reflections from the rear wall.	Intimacy on stage is achieved by individually adjustable stage monitor systems, and isolation from the main transmission system. This is challenging because the artists desire to connect with the audience. A compromise position is reached that provides limited contact and limited isolation.
Texture	Fine texture is achieved by carefully spaced and sequenced reflection patterns.	This is achieved by a steady transition of the low level reflections through the Isolation Zone. The texture is enhanced by high reflection density in the Isolation Zone.	Different but not conflicting	This acoustic texture can be augmented electronically with the Isolation Zone summation introduced of outboard reverberation. Electronic texture is thereby added to the acoustic.	The electronic reverberation can be applied separately to different input channels.
Freedom from echo	Freedom from echo is achieved by preventing single reflections to stand out above the decay pattern or focus points to occur from the confluence of multiple reflections.	Reflections must transition steadily towards isolation. If the relative level of reflections rises after isolation the possibility of echo perception rises.	Strongly Agreed	Echo perception risk is greater in sound systems because of lower reflection density, and higher sound levels. Concentrated beams of directional speakers enhance chances of focused reflections.	Small number of separated focused sources raises the risk of perceptible echoes.
Freedom from noise	Freedom from noise is achieved by isolating the transmission from noise sources.		Strongly Agreed		All noise sources of the acoustic model are present. Additions include electronic noise, feedback.
Dynamic Range	Dynamic range is maximized by having the strongest summation of the direct sound and early reflections and the minimum noise.	The maximum acoustic addition requires both Coupling and Combing Zone summation.	Different but not conflicting	The maximum acoustic addition requires both Coupling and Combing Zone summation.	The maximum side is limited by two factors: power capability of the system, gain before feedback. Minimim side is limited by noise.
Tonal Distortion	Tonal distortion is prevented by not allowing absorption to occur at selected frequency ranges, or for sympathetic vibrations or resonances that add/subtract frequency content to the music.	Tonal Distortion is prevented by dense complex multiple summations of different relative time and level. If a single reflection is too dominant it will cause cancellations and excessive additions.	Strongly Opposed	Tonal distortion is minimized by focusing on Coupling Zone and Isolation Zone summations. Combing Zone is avoided to the maximum extent possible.	Tonal distortions unique to electroacoustic systems include harmonic distortion, compression and sonic image distortion.
Uniformity	The hall is free of "dead spots" or areas where the sound quality is distinct from the main areas. The maximum similarity of sonic experience is the desired outcome.	Uniformity is achieved by coupling all seating areas to the transmission system. Uniformity is achieved by saturating the frequency response with super dense Combing, Combining and Isolation Zone summations.	Strongly Opposed	Uniformity is achieved by isolating the speakers from each other and the room. The frequency response is maintained by Coupling Zone and Isolation Zone summations. Combing Zone is avoided to the maximum extent possible.	Uniformity is achieved by dividing the reception area into isolated transmission zones.

Figure 4.6 *(Continued)*

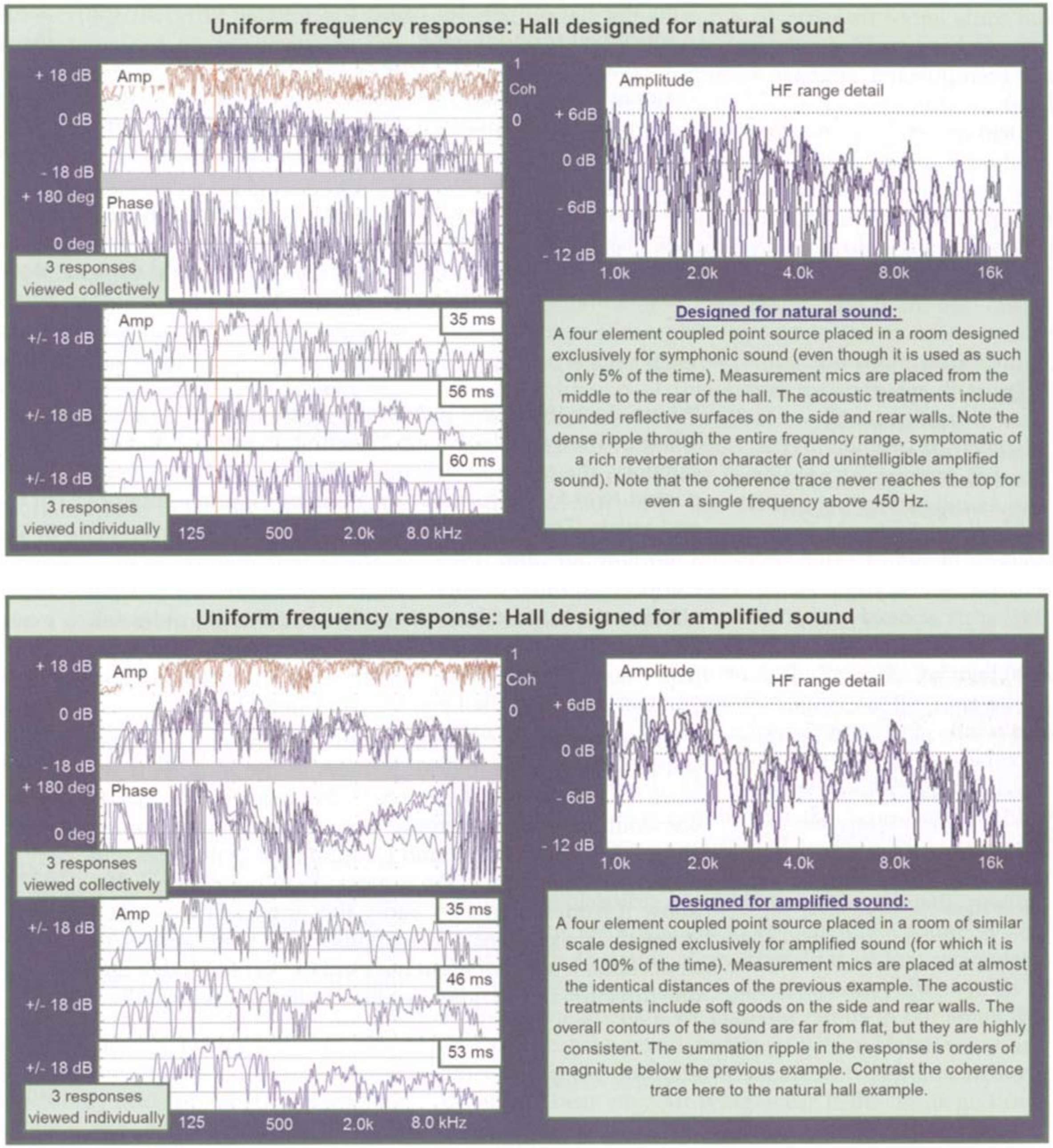


Figure 4.7 Comparison of frequency responses of speakers ir

in the natural model cause the reverse effect in the amplified sound model.

- Freedom from tonal variance is also strongly oppositional. Tonal variance is suppressed in the natural system by reflection saturation. Excessive reflections are the cause of strong tonal variation with speakers.
- Uniformity is the most oppositional of all categories. The unified nature of the natural transmission system mandates that uniformity be achieved by reflection saturation. The separated nature of amplified sound mandates that the reflections be minimized.

Perspectives: Amplified sound in a symphony space is the perfect example of not being able to have your cake and it eat it too. By its nature a large hall has a longer reverb time than you would ever want in a "theatre" space or "events" hall. The longer reverb times of 1.6 to 2.4 are the best friends of an orchestra on stage. But put them in the orchestra pit and amplify vocals for a theater performance and you will be in for the fight of your life. The vocals never achieve the edge and clarity that the audience is expecting as if they were in a movie theater or Broadway house with a lower reverb time.

Kevin Mochel

The Middle Ground

Where does this leave us? Where is the middle ground? What are the mutual solutions? The first strongly conflicting category, the sense of ensemble on stage, has long been passed over to the scope of the monitor system operators and artists. That was easy.

We are left with three categories that require creative solutions: intimacy, tonal variance and uniformity. They all have the same source of conflict and the two sides fall into the same camps in all cases. If we can solve one, we have them all. The disagreement is on strong reflections. Natural sound needs them. We can't stand them.

The worst of all possible solutions is the one that presents itself as most logical: to meet in the middle. If we were to give the orchestra half the reflections it needs we would still have far more than we want for amplified sound. If the middle ground is found here, both sides will be equally unhappy. This is more than hypothetical. We know it from experience. Sadly this approach was taken in the design of a great many halls in the 1960s and 1970s in particular. Most of them have either been remodeled in one direction or the other or fallen to the wrecking ball.

Variable Acoustics

The win/win solution is found in the form of variable acoustics. No middle ground need be taken. Simply design the hall with features that modify the reflection structure to accommodate the program material. This approach is now the standard for modern symphony hall designs. The lack of profitability of live symphonic music would likely prevent the construction of another symphony hall if not for the donations of its patrons. Even so this generosity has it limits and as a result the expectation for these halls is that they will book events which appeal to audiences beyond the classical realm. Almost every one of these requires a sound system. The room designs must allow for rapid reconfiguration of the hall for acoustic properties appropriate for pipe organ, chamber music, symphony, and opera and finally to our sound reinforcement applications. Curtains might drop from winches, wall panels rotate from hardwood (reflective) to soft goods (absorption), and reverberation chambers are opened or sealed. This has the potential to be the *optimized* acoustic design, one that is able to fit the management's needs to fill the hall and also fit the artistic needs of all participants.

To be fully optimized for all types of program material will require large range of reverberation control. In such cases the top priority client is almost always the symphony, even if their percentage of the bookings is small. The question then becomes one of how much soft goods to put into the room for the amplified sound performances. This has major economic impact on both construction and operational costs. As a result the halls are most often left with far more lively acoustics than is optimal for the sound system. It is possible to build a hall that has uncomfortably dead acoustics, even for amplified sound. It is, however, extremely unlikely that a hall with optimal acoustics for a symphony can be converted in a single labor call to one that is too dead for us. In short: in a classical hall we'll take as much absorption as the owners are willing to buy us.

Hybrids: Combining Natural and Amplified Sound Transmission

What about the case of a hybrid application which uses both natural and amplified sound transmission? This merger of transmission methods is what is truly meant by the term sound "reinforcement." It is the real middle ground. In such cases all parties need to be aware of the vulnerability of the situation. The extent to which the musicians can create a self-mixed sound on stage will be the decisive factor. This is a delicate balance. The levels on stage must be carefully controlled so that the speaker system truly plays a supplemental role. If a single instrument overpowers the others from the stage, the mixer will be obliged to raise the rest of the instruments in the sound system to keep pace. Things can quickly unravel from there. The sound system will need favorable locations and time alignment to provide plausible sonic imaging for the reinforced instruments. In essence, the sound system will need to be joined to the natural sound as a phase-aligned crossover: meeting in time and in level in the house. If either system falls far ahead in either of these parameters, the union will be lost.

Another hybrid approach is often used in musical theater. A highly directional system is employed for vocals only, where the sense of spatiality has less priority than vocal intelligibility. A separate stereo system comprising wide coverage speakers transmits the music mix, often blending with the direct sound from the orchestra pit. In this case the room reverberation is willfully added to the mix and provides enhanced spatial feeling.

Variable Electro-Acoustics: Artificial Reverberation

We have extensively discussed the analogous actions of speaker/room and speaker/speaker summation in previous chapters. Reflections can be modeled as "virtual speakers" and exhibit similar sonic effects. The reverse can also be true: speakers can be modeled as "virtual surfaces" and thereby create the sound of a reverberant space.

Perspectives: Using a focused system and VRAS (variable room acoustic system)II am able to widen the pleasure zone, for more of my audience. Aurally speaking that is.

Jonathan Deans

Artificial reverberation is far advanced from the simple electronic reverb found among the outboard front of house equipment, which only adds reverberation into the speakers that are carrying our direct sound. While this simple approach attaches a reverb tail that mimics the decay character of a real room, it does little to convince us that we are actually in a reverberant room, since our localization perception is unaffected. The simple reverb fails to gives us the experience of spatial envelopment of a decayed sound field arriving from all directions. Artificial reverberation requires a complex array of many distributed microphones and speakers which create a multidirectional, diffuse, spatially decayed field in the room. The microphones receive their signal from the sound system (or from an acoustic source) and recirculate the sound into the main system and/or spatially distributed, dedicated "reverberation source" speakers. The multiple microphones create a complex form of isolation zone re-entry summation with staggered timing and level relationships. The reproduction of this signal by spatially distributed loudspeakers creates the effect of virtual reflective walls in locations where there may actually be absorptive walls, or no walls at all. There is nothing to prevent this signal from re-entry into our distributed microphones, and therefore the recirculation continues in a manner consistent with the multiple reflection paths found in a reverberant space. This approach has its limits of stability and credibility. Any re-entry summation system is subjective to instability, which in the worst case is runaway feedback. Therefore we must carefully limit the recirculation levels. The credibility is another more subtle issue which relates to our perception of impossible acoustic effects that clue us to the presence of a sound system (previously discussed in Chapter 3). Our eyes have sized up the room and created an expectation that the acoustical qualities will be in a range proportional to the space. Excessive amounts of artificial reverberation will push the listener beyond the plausible expectations of the acoustic qualities of the space, leading them to suspect that a sound system is involved. An artificial reverberation system recirculates all sound sources in the room, including the sounds of audience members, if we clap our hands in our small theater and hear the reverberation of the Taj Mahal we would have to ask the question: what are you going to believe? What you hear, or your lying eyes? If the lights go down or we close our eyes, the range of plausible effect rises.

These systems were originally created during the time period when multipurpose halls attempted to navigate the middle ground of architectural acoustics. Halls were built with insufficient reverb and the artificial systems were added to supplement the natural sound performances. The reverberation system would then be disabled when the sound system was used. This could even be done within a single performance cycle such as in a house of worship: when the organ plays and/or the choir sings, the artificial reverberation is operational, but not for the spoken word. The early versions of artificial reverberation were as unsuccessful as the early multipurpose halls. The digital age, where we have affordable multichannel processors providing cross-matrixing of inputs to outputs and a variety of adjustable functions, has ushered in an advanced generation of artificial reverberation systems that make this option vastly more attractive and with much improved feedback stability.

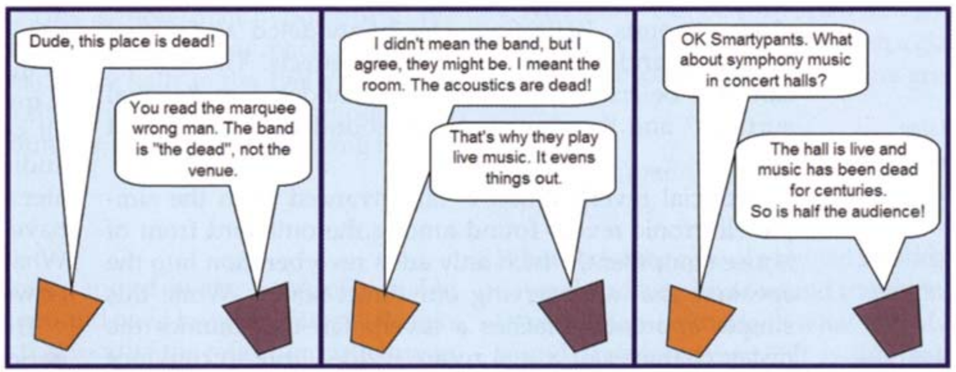
Perspectives: Certainly creating an infectious room helps the audience's response. Having a speaker system that can he controlled by a console which must never stop the creative choices with its limitations. Then applying a simulation of acoustical reverberation which has real-time cueable parameters, adds up to creating design choices of the actual room's space, as well as the designer's job of layering sounds.

Jonathan Deans

While artificial reverberation is still used in such applications currently, there is another option that is of keen interest to us: the use of artificial reverberation as a spatial envelopment effect for our sound system. This form of "variable electro-acoustics" allows us to selectively invoke the reverberation character of our "room" on demand. The sound system can be installed in a "dead" room and yet have the spatial effects of a "live" space when that is preferred. Advanced installations and highly creative sound designers have used this to great effect, creating entirely portable "soundscapes" that can be erected in the room and linked to the events unfolding on stage. The presence of artificial reverberation relieves the room of the diffusion and spatiality requirements of our sound system. A fully uniform direct field response provided by our speakers can be supplemented by the diffuse spatial field of the artificial reverb system. Visualize a wire frame drawing of a theater. As long as we can hang enough reverberation speakers all over the wire frame we've got ourselves pretty much "perfect acoustics." If only there was a way to keep the rain out.

The affordable technology of artificial reverberation redefines the boundary line between architectural acoustics and audio engineering. Spatial enhancement and a rich decay character are the most sought after acoustical properties in amplified sound applications. Now even those characteristics can be undersized in the room acoustics and "fixed in the mix." The acoustician's emphasis on architectural solutions has naturally led them to favor solutions inside their scope of expertise and control. Likewise the audio engineer. The hall built for natural acoustics alone has a high potential to reduce the options of audio engineers to damage control and emergency mode operation, unless we bring in truckloads of drapes. The hall with the acoustic properties of a pillow will allow us to mix like we are outdoors on a wind-free day, but leaves the audience with the feeling that all sound is coming at them, rather than around them. Artificial reverberation

Trap 'n Zoid by 6o6



moves the line towards erring in favor excessive absorption, rather than excessive reflections. It is a very substantial migration of the scope of work toward the audio engineering side, so don't expect this to be welcomed with open arms. Both parties are still very much in the game, however, as the creation of a plausible reverberation character in a given room still requires the expertise of those who know what that means: the acousticians. This technology opens up huge possibilities, and its acoustical success will be based upon cooperation between the audio and architectural sides. Artificial reverberation opens a second potential avenue for the optimized sound design to meet the optimized acoustic design.

Moving Forward

It is possible that a modern-day sound engineer came to this career because of a lifelong fascination with acoustical physics. It is equally possible that someone in architectural acoustics has always dreamed of stuffing rooms full of fiberglass. It is far more likely that members of both of these disciplines share a common root to their inspiration: music. For us the music has always been directly linked to our speakers and electronics. For the acousticians the music has been directly linked to the room. The modern audio engineer who thinks they can bluff their way through life without knowledge of acoustical physics is every bit the fool that many acousticians suspect us to be. Creative theories of sound propagation that will impress the lead guitarist will not persuade the lead acoustician. Likewise, the modern acoustician who throws up a dodecahedron speaker and thinks we will be impressed is just as exposed.

We happily acknowledge that there are designers on the architectural electro-acoustics side who fight the battles to achieve a workable system in the political, monetary and aesthetic context of the construction project. Their competence and the extent to which they reach their goals determines the converse extent of problems that the visiting engineers have to fix. The audio system design always follows the architectural in the design process. We cannot assess the sound system needs until the space has been defined. We're all in the soup together, so it will be better for all parties if we can anticipate each other's needs and perspectives to find the best solutions so we can all share the credit rather than pass around the blame.

First let's prioritize the goals of the architectural electro-acoustical designer so that the work of those that follow is more satisfactorily achieved.

The following list, provided by Dave Clark, sets forth the areas which should be defined in sufficient detail for us to evaluate how we can accommodate the sound requirements:

- An acoustical management plan (noise, vibration, room acoustics, variable acoustics)
- An accommodation zone for the main loudspeakers
- Shapes and sizes of attached volumes; accommodations for the loudspeakers therein
- Finishes for the various surfaces.

If these elements are covered, then the sound system can move forward toward optimization.

The following list, provided by Sam Berkow, sets forth some of the common goals our disciplines include:

- Tonal balance: relative ratio of decay rates between mid/low and mid/high frequencies.
- Freedom from potentially disturbing reflections (for both audience and performers).
- Uniformity of diffuse energy: this refers to under balcony spaces and other geometric "segmented volumes" within a space that can "hold" diffuse energy. This is a big deal in larger venues, such as arenas and stadiums.
- Appropriate levels of ambient noise: this is not an issue for rock systems, but can be a *huge* deal for speech systems, particularly in worship spaces.
- Uniformity of sound field: how different is the experience for various seating areas.
- Clarity and intelligibility: high direct-to-reverberant ratio, high coherence, etc.

We know what we need, but we cannot expect acousticians to read our minds. So we will make it clear by creating a base specification for the architectural acoustic qualities we desire in the hall.

So what do we want/need from the room architecture in cases where there is absolute certainty that the events will be amplified? Let's simplify things by considering the polar opposite of the symphony hall model. Instead of all of the sound propagating naturally from the stage, we have none. Every performer is using in-ear monitors, and the band is fully isolated. We can mute the PA just like a commercial on television. This represents the most extreme case, but a very common one. The discotheque, sports venue announcement and music effects, and many pop music concerts, are just a few examples where this model is valid. Next let's consider that we have a rock and roll band on stage with drums, guitars, etc., which leak off the stage and join the sound system in the house. Does this change the big picture of what we want in the architecture? Not likely.

Those applications where the natural sound and amplified sound coexist in relative parity are the middle ground. We will never be able to reach the middle if we don't know where the opposite ends are found.

Now let's explore an "amplified-speaker-centric" view of the qualities we look for in a performance space when the design meetings begin, or the doors swing open on the truck.

1. Room shape: we accept the fact that sound is only one factor in the overall room shape. The room shape of a basketball arena is set by its primary purpose and the sound will be forced to adapt to that shape. Our main systems will be designed to cover the majority of the macro shape and the fill systems will cover the remainder. Smaller-scale aspects of the room shape may also need the attention of fill systems. Careful consideration must be given at all scales as geometric complexity is added to the interior shape. It is understood that all seats will require line of sight to the stage. There are also "line of sound" considerations and the distinction is very relevant. In the natural acoustic model, line of sight equals line of sound, since the sound actually propagates from where we see it: the musicians. In the amplified model the sound will propagate from speakers which we can fairly well bet are *not* located on stage. Line of sight most definitely does not equal line of sound and yet the relationship is important. As sonic ventriloquists, we will need to make the listener *perceive* that the line of sound is linked to their line of sight, i.e. a plausible sonic image. This means that it will not be enough simply to find a line of sound to a location; it must come from a direction that gives us a chance to place the sonic image in the right location. Even seemingly minor shape intrusions into the macro shape have the potential to create a large impact to the sound design whenever it forces us to subdivide the system. A large deep balcony (and under balcony) is an example of a room shape function that will undoubtedly require coverage subdivision. But even a small under balcony area only a single row deep can require a dedicated system if it is found in an unfavorable location with respect to our main speakers. The "devil is in the details" as the old saying goes.
2. Matching the shape to the purpose: if stereo is the prime focus the room must be shaped in a way that is conducive to large-scale stereo. If the room is a wide fan shape or an "in the round" configuration, stereo imaging will be hard to realize and attempts to create it may cause degradation of the intelligibility. On the other hand, those two shapes are very well suited for mono. If we are doing musical theater with the expectation of undetected reinforcement, we will need a means to keep the image low. The list is endless but the theme is the same: consider how the sound is going to be transmitted inside the shape for the primary purposes for which the room is designed.
3. Under balcony spaces: if the under balcony spaces have sufficient height clearance and are not excessively deep we will not have to use under balcony speakers. The under balcony opening must be a minimum clear height for any rows of more than 3.3m (lift) and the need for this height increases with depth. The finishes of the under balcony space must address LF resonances.
4. Side seating areas: side areas on the floor may require some fill speakers but are otherwise straightforward.

Upper level side seats can be some of the most expensive seats in the house for sound. In particular the side seating boxes that wrap all the way around are the worst. First they are only one or two seats deep and so the walls are very close behind. The focus angle from a center speaker turns the side walls into rear walls from our point of view, so they need to be deadened. From a side speaker point of view we have near boxes that are very close to the speaker and these boxes will be overpowered if we try to reach beyond them. From a side or center cluster point of view they may be blocked on the lower (and perhaps upper) levels even when only one or two rows deep. Few people, lots of speakers, bad sonic imaging, high reflection hazards.

5. Low-frequency absorption/diffusion: There are few metrics more prized by audio engineers than "tight low end." The other prized attribute is "warmth." How can we achieve these? "Tightness" is primarily a function of transient response whereas "warmth" is primarily tonal. A reproduced impulse must rise and fall without lingering excessively in the space. How can people dance if they can't find the beat? We need the very-low-frequency decay to be fast enough to give space for the next kick drum impact, which will be arriving shortly. Installing as much low-frequency absorption as is reasonably possible works very favorably for us. We can obtain the warmth required by scaling the quantities of subwoofers appropriately and tonal balancing through relative level and equalization. The virtually unlimited direct sound power capability available to us from our speakers greatly relieves the room of its responsibilities to supplement the LF range with reflections. Our speaker system will likely have far less directional control in the LF range than MF and HF ranges. Since higher amounts of LF direct sound will impact the walls we can expect that the reflection will already be spectrally tilted in favor of the LF range, even if the room absorption were spectrally balanced.
6. Mid-frequency absorption/diffusion: mid-range pattern control in the speaker array is more manageable than low frequencies. Mid-range absorption will be welcomed particularly in areas near the speakers and in on axis areas such as inward and rear walls. Diffusion is second best and can provide some spatial enhancement.

High-frequency absorption/diffusion: this is the most forgiving on one hand, since we have the best directional control of the speakers. This is the least forgiving on the other hand, because the temporal threshold into discrete echo perception is the lowest. We have serious issues with things like off-beat high-hat cymbals appearing where they don't belong. Care must be taken that glass, plaster, concrete and metal surfaces are not positioned in such a way that they will reflect sound from the speaker positions into the audience or onto the stage. The rear surfaces of the hall are very challenging for us. We have to aim the speakers to the back wall, in order to get the same level from front to back. We don't want strong reflections coming back to the audience. We especially don't want them to refocus on the stage where it will disturb the performers and get back into the mics. Therefore these surfaces should be highly absorptive. Complexity in the rear wall will help to provide some diffusion, which will reduce the probability of discrete echo perception.

Ceilings: From the speaker system point of view, the ceiling has a very important function: it keeps the water out. Otherwise our policy with respect to ceilings is simple: avoidance. There are two principal areas of concern for us here: ceiling surfaces near to the speaker source location, and those near the final destination. Our preference for nearby surfaces is that they splay away from our angle of propagation so that we do not fold a nearby reflection into our direct sound path. Nearby surfaces with parallel or inward angles will benefit from absorption. On the destination side we encounter an analogous condition. An inward surface angle brings strong reflections down into the seats that can tolerate them the least: the most distant. Since we will need to reach the rear of the room it is inevitable that ceiling reflections will become an increasing factor at the rear. Downward tilts exacerbate the situation.

Sidewalls: again we can benefit from an outward splay angle. This provides the most favorable angular orientation for speakers located at either the stage sides or above center. The direct sound can be tailored to place the axial edge along the side walls and minimize the reflections. In addition, the reflected energy is guided first into the absorptive back wall.

10. The floor and seating areas: a symphony propagates from the stage level, not from ten meters above it as might be the case for our speaker system. The impact of floor reflections is very different from the natural model since sections of the speaker system are focused directly onto the floor surfaces at steep angles. In cases where we are focusing concentrated energy onto the floor surface rather than grazing along its surface we will need to consider the potential benefits of absorption, i.e. carpet. Now I hate carpet as much as the next guy, so let's keep it to a minimum. For starters we need only consider the aisles, since the seating area will have human absorbers. The question of covering the aisles depends upon their orientation to the speakers. The "ray of light" test will provide the answer: if we can stand in the aisle and follow its line to the speaker, that aisle is a good candidate for treatment, e.g. center cluster, fan-shaped aisles that lead to the stage.
11. Hidden speaker locations: to quote an old song, "Don't fence me in". We prefer not to have our speakers stuck in recesses, behind beams, HVAC ducts, lighting instruments, catwalks and more. This seems obvious, but we would not bring it up if it were not a problem we commonly see. Recessed locations, if required, must be acoustically transparent and must be open for the full angular coverage of the speaker/array. Common mistakes in recessed spaces are lack of any contingency capability, sight line intrusion, lack of absorption in the open cavity and rattles and buzzes in the local hardware. Very often these spaces are constructed so tightly around the speaker(s) that the focus angle cannot be modified. Give us breathing room, please. Why is it that it is perfectly OK to see 100 lighting instruments leaking light out of their sides and backs and we have to hide our black speakers? Speaker recesses are certain to add cost, and usually degrade the sound. Aesthetic considerations, however, should not take priority over operational requirements.
12. Main speaker locations: if plausible sonic imaging is desired we will need speaker locations that can keep the image low and central. For the floor seating level we will need proscenium side locations to keep the image low and frontfills in the stage lip for both vertical and horizontal localization. For upstairs seating areas a central source above the stage will be best for horizontal localization, but side locations (at a lower elevation) will be best for the vertical. Compromise may be required here, or both can be used.
13. Under balcony speaker locations: give us room to hang our speakers under the balcony so that they can be aimed toward the last row of seats. We cannot use downward-facing recessed ceiling speakers. Directly under the balcony lip is very rarely the best location. Most often it will be further back than this, so surface mounting or an angled opening will be needed. The role of the under balcony speaker is to improve the ratio of direct to reflected sound. Therefore, reflective surfaces near the speaker will be a concern. It is understood that such speakers will maintain a low profile to the ceiling surface. Absorption in the local ceiling area in front of the speaker would be beneficial. If the speakers are recessed, provide the maximum clearance for focus adjustment.
14. Frontfill speaker locations: this a case where recessed mounting is practical and preferable. Maximum height is important since the range of the frontfill will be limited to a single row if they are too low.
15. Mix position: lighting requirements don't change from night to night because the singer is hoarse or we have a stand-in for the lead actor. Go ahead and put them in the rafters behind glass. Sound changes minute-to-minute, night-to-night and we will not be calling cues to the follow spots over our headset. We need to be in the house, where we can hear the show and see the performers.
16. The stage: this is a case where we can get lots of bang for our buck with variable acoustics. Movable drapes or panel absorbers can deaden the stage for a rock

show, but leave it lively for the hybrid sound reinforcement applications. The stage floor should be free of resonance so that stage-stacked speakers do not excite a particular range. The ceiling above the stage should not be hard and inclined to focus the sound outward or downward like a bandshell.

17. Surface complexity: we can tolerate a higher amount of reflections if they are complex and diffuse. The high frequencies are highly controlled in our speakers, so small areas of glass and steel will not trouble us as long as they are diffuse and randomized.
18. Balcony fronts: these are just like back walls, only closer, and are more prone to focusing back to the stage. An upward angle is desired to route the first reflection up to the ceiling or back wall. Worst case is a flat, hard, curved surface. Center clusters will focus it back to the stage, while side arrays create a phantom on the opposite side.

Audio engineers and acousticians can work together and achieve mutually rewarding results. A critical examination of the role that natural sound will play in the space is required. If all events are to use a sound system, then all acoustic treatment should realistically factor in our requirements. If the venue requires both transmission types, it will need both types of rooms: variable transmission, variable acoustics. It will also need an optimized sound system, which now becomes the focus of our discussion.

By the way, the dodecahedron speaker lasted one listening test and now there are curtains added whenever the speaker system is used. Optimized sound system and optimized acoustics.

Reference

Beranek, L. L. (1962), *Music, Acoustics & Architecture,* John Wiley & Sons

5

predict *v.t. foretell, prophesy*

Concise Oxford Dictionary

Introduction

Prediction is the cornerstone of the design process. Measurement is the cornerstone of the optimization process. As designers, our role is to predict system performance so as to create the highest probability for successful optimization. Our optimization role is to act on the measured data and enact modifications and adjustments as required. This may prove or disprove the predictions. Our credibility as designers will depend upon how closely the aligned system meets expectations. These expectations were based on what our predictions indicated were possible. If the optimization stage is successful and requires little more than fine-tuning, our prediction will be deemed credible. If the optimization is a failure or requires costly and time-consuming changes, our credibility as designers is tarnished.

There is a circular relationship here: measurement and prediction. After all, where does the data that we use for prediction come from? Measurement. How do the systems we measure come to be installed? Prediction. A step forward in either of these disciplines is a step forward for both. Each time a system is measured, its underlying theory is tested. The successful parts are proved and those areas where the measured data disproves the prediction provide an impetus to re-examine the prediction data in the hopes of a better understanding. With each increase in understanding the prediction process will be elevated, and the measurement/optimization side can push the investigations and calibrations to a higher level. This is an ongoing cycle of discovery.

The lessons can only be learned if the prediction and measurement tools share a common language and system of quantifying results. We cannot correlate the predictions of tarot cards with the measured response of an acoustic analyzer. We must have prediction tools that depict the sound propagation as we will measure it. Likewise, we will need measurement tools that depict sound as we experience it and as we know it to exist. There is no benefit in either case to portraying the response in rosy terms. The devil is in the details here. As we have seen in Section 1, there is plenty of detail to be seen and heard.

Sound system predictions are primarily concerned with four principal responses:

- the free field transmission characteristics of the speaker
- the effects of the transmission medium (air)
- summation effects of multiple speakers
- summation effects of the room and speakers.

Whether we are using the most advanced 3-D computer-aided design (CAD) program and acoustic modeling or pencil on a napkin, the process is fundamentally the same. The nature of all of the above responses has been covered in Section 1 of this text. It is now time to move them beyond abstraction and apply these principles to our practical application. The first essential links will be the drawings.

Drawings

It is easy to make predictions of speaker performance in the abstract. It is difficult to make a living doing this. What we need is an application, otherwise known as a venue. The venue for a permanent installation is a single structure. The system design can be precisely fitted to the shape of the hall. For touring systems the venue is variable over time. Touring systems are designed to work in a particular scale. This might be theaters, arenas, outdoor amphitheaters or others. A touring system design will need to be flexible enough to accommodate the various differences in venue shape.

In either case we will need documentation of the shapes and scale of the rooms.

2-D Drawing Types

Design drawings come in 2-D and 3-D forms. Two-dimensional drawings are the norm. The minimum requirements are plan (horizontal) and longitudinal section and cross-section (vertical) perspectives.

Plan view: this is floor plan of a particular level as seen from above. If the level has sloped surfaces or areas with different floor heights, these will be marked. This data will be used to predict the horizontal coverage requirements.

Longitudinal section view: this is a vertical rendering of the interior features along the length of the room as seen from inside. The typical viewpoint is the room center, looking toward one of the side walls. This is assumed unless otherwise specified. Floor and ceiling, front and rear wall features are shown as they exist at points between the view perspective and the side wall. The side wall surface features can be seen provided they are not obscured by closer surfaces between the view perspective and the wall. A suspended lighting fixture located between the center line and side walls will appear in front of the wall details. A single drawing is all that is required in cases of side-to-side symmetrical rooms. This drawing will be used to predict the vertical coverage requirements for speakers with a front-to-back (or back-to-front) orientation.

Cross-section view: this shares common features with the longitudinal but is oriented to view the front (or back) rather than the side. In most cases the front and back will not be symmetrical. A typical theater would have a front cross-section view which shows the stage cutout and a rear view which shows the arrangement of the seating and back wall. This will be used to predict the vertical coverage requirements for speakers with a side-to-side orientation.

Elevation: an elevation is a vertical rendering of the wall surface. The view may be from the interior or exterior. Floor and ceiling, and intersecting side walls are shown as they exist at the point of intersection or attachment to the wall. The elevation is of limited use in prediction. For simple rooms the section and interior elevation views will differ only slightly. More complex shapes will bring about differences. For example, a room with a flat balcony front will show the same location for the balcony rail on both the section and elevation views. The elevation drawing would show a typical curved balcony only at a single point: the intersection at the side wall. The section view would show both balcony depths at the center and at the side wall. Elevation drawings are helpful in cases where speakers are to be recessed in openings in the wall.

Reflected ceiling plan: the reflected ceiling plan is like a plan view of the ceiling. The perspective is unique, and at first confusing, but is in fact quite logical. The view perspective is as if we looked down at the floor, which was covered with a mirror. We see the ceiling as if looking up, but the horizontal orientation is reversed. Why is this preferable? The reason is that the reflected ceiling plan can be laid directly over the floor plan and everything lines up. If an object appears at the same location in both drawings then

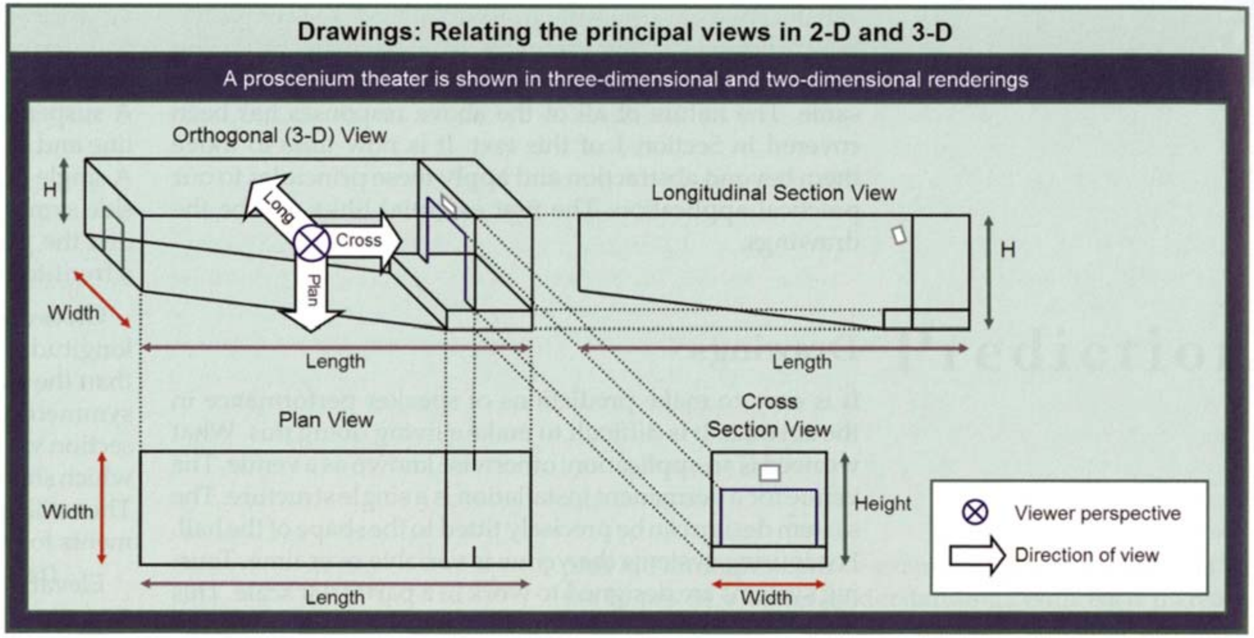


Figure 5.1 Relationship of the principal views

one is directly above the other. If we use a ceiling plan that uses the normal inside perspective, objects that are vertically aligned will be seen on opposite sides of the drawings.

Bathroom plan: most jobs give us either no plans at all or too many. I could paper the walls of my house with toilet layouts, although I have never yet designed a system for such a challenging space.

3-D Drawing Types

The room can also be represented in three-dimensional CAD drawings. The advantages of 3-D CAD renderings are quite obvious. They can duplicate any of the 2-D perspectives and add orthogonal views that provide improved clarity. This is especially useful for complex surfaces or perspectives at angles not matched to the 2-D renderings. Another advantage is that 3-D renderings look a million times cooler than any 2-D drawing. The biggest disadvantage is that the operation of 3-D CAD programs is beyond the capability of all but a very small minority of audio professionals. The next is that, even with a skilled operator, these drawings are extremely time-consuming and very expensive. This moves 3-D CAD out of the budget realm for many sound system designers and clients. For these reasons we will limit our discussion to 2-D drawings.

2-D Drawings in a 3-D World

It is worth taking a moment to consider the vulnerabilities of 2-D drawings in a 3-D world. The 2-D viewpoint is made up of vertical and horizontal slices. The vulnerability of this approach arises whenever we are concerned with the response of speakers that radiate at angles other than these flat planes. Sound does not radiate outward in a flat disk shape like the rings of Saturn. It radiates spherically. Therefore, action occurs at all axial planes rather than just the chosen two.

This does not, however, render our 2-D renderings useless. A predictable error factor arises which can be compensated for fairly accurately. There are two planes to consider: the propagation plane (that of the speaker) and the drawing plane (that of the drawing). The error factor is lowest when the two planes are matched and correspondingly rises as the planes diverge. There are two forms of divergence, distance and angle. We will use examples to illustrate the two forms.

Distance divergence is easily visualized by the ceiling speaker placed at the building center line. When viewed in section (or cross-section) the propagation plane is a match to the drawing plane. The propagation distance from the speaker to the floor at any location can be found without error, and the relative levels estimated. When viewed in plan the propagation and drawing planes are not matched. The propagation plane is actually matched to the reflected ceiling plan. It is the relative levels skimming across the ceiling that are represented. Meanwhile down on the floor, the seat directly under the speaker appears as the sound source with rings of radiation dropping by 6 dB with each doubling distance. This is a far cry from the situation of the floor and the error increases in proportion to the distance between the planes. An accurate plan rendering of the relative sound levels on the floor will need to be estimated from the section data.

Angular divergence has similar properties and added complexity. A representative example of this is found in comparison of a frontfill speaker and a center cluster. The center cluster is directly overhead of the frontfill, giving them identical placement in plan view and different placements in section and cross-section. The section view reveals the speakers with propagation and drawing planes matched. The plan view plane is very close to the frontfill propagation plane and a far distance from the center cluster's plane. The estimates of relative level will be reasonably accurate for the frontfill, which will drop off at the inverse square law rate over the seating area. Estimates for the cluster would show much higher differences in level than would actually occur. The reason is that the plane divergence is not compensated. At the front row the propagation and drawing angles would diverge by nearly 90 degrees, creating a maximum of exaggeration in

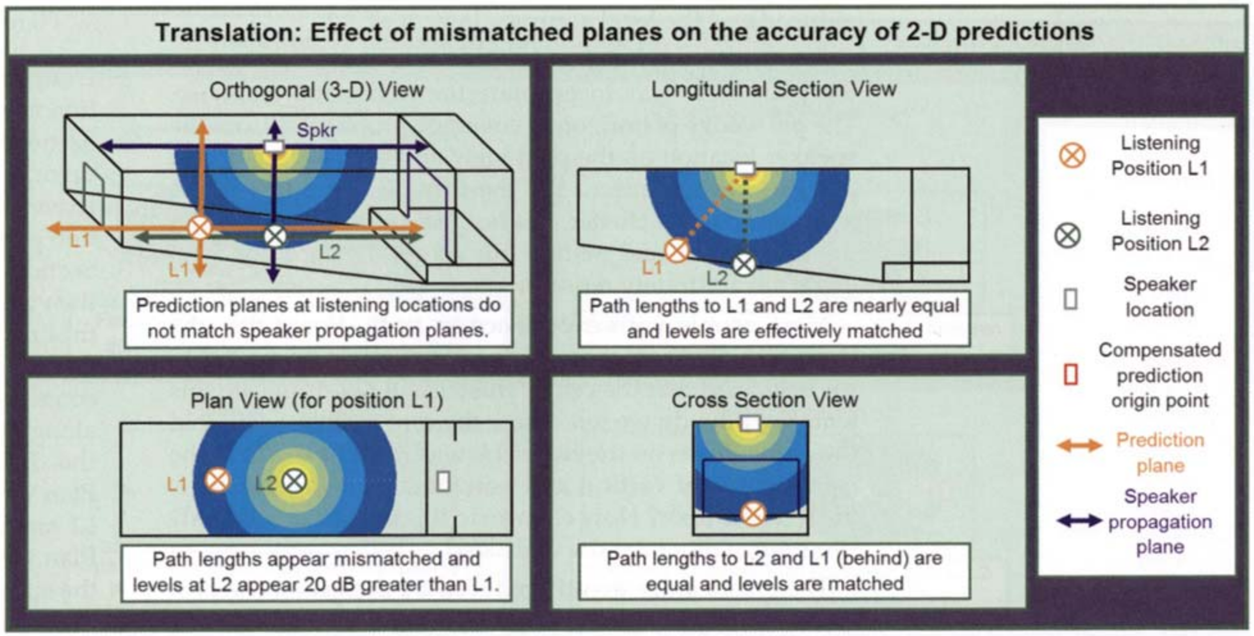


Figure 5.2 The effect of mismatched propagation and pr horizontal propagation is the ceiling. The vertical position of horizontal plane of the prediction is the floor. This mismatch causes levels to appear to vary widely on the floor

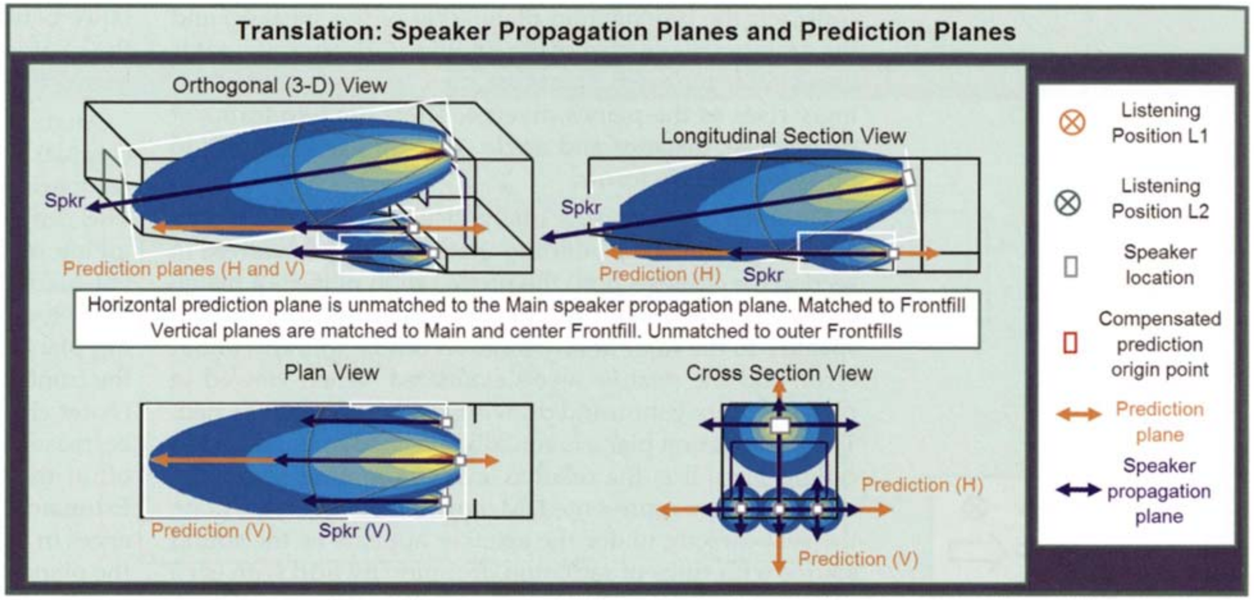


Figure 5 3 Angular relationship of the speaker propagation and prediction planes. The angular orientation of the main speaker's horizontal propagation is downward. The angular orientation of the horizontal plane of the prediction is flat. This mismatch causes levels to appear to vary widely on the floor in the plan view. The propagation plane and prediction planes of center frontfill speaker are matched in both level and angle for both vertical and horizontal planes. The outer frontfills are matched only in the horizontal planes

level. As we move back in the hall the angular difference is reduced and the level accuracy improves.

A secondary issue arises. The whole reason to look at the plan view was to estimate the horizontal response. The pie wedge of horizontal coverage propagates from the speaker location on the plan view and denotes what is in and out of the pattern. For the frontfill our rendering is accurate. For the cluster it is not, because the seats are in the other plane and neither our actual distance nor axial angle are accurately represented.

Most people with experience in audio know that the front fill will overpower the near rows and never get to the back, and that the center cluster will not duplicate this feat. But how do we tell where the center cluster will join the game down on the floor? How can we find where the combination of vertical and horizontal coverage patterns meet on the floor? How can we do this without a 3-D CAD program, a month and a whiz kid?

The error can be greatly reduced by the geometric process of triangulation. The error in the plan view rendering will be reduced using what we know from the section view. Here is how it works (Figs 5.4 and 5.5).

1. Designate a listening position in both the plan and section views. Let's call it "L1."
2. Section view: draw a vertical line (A) from the speaker up or down to the height of L1. This distance is the difference between the horizontal drawing and propagation planes.

   Section view: draw a horizontal line (B) from the L1 to the vertical line. This distance is the difference between the cross-section drawing and propagation planes.

   Section view: complete the triangle with a line (C) that connects the speaker to L1. This is the distance traveled along the propagation plane that arrives at L1. Measure this distance.

   Plan view: the horizontal distance from the speaker to L1 matches the length of line B.

   Plan view: draw a line of length C from the L1 through the speaker position and mark it "S1." This is the compensated speaker location for position L1.

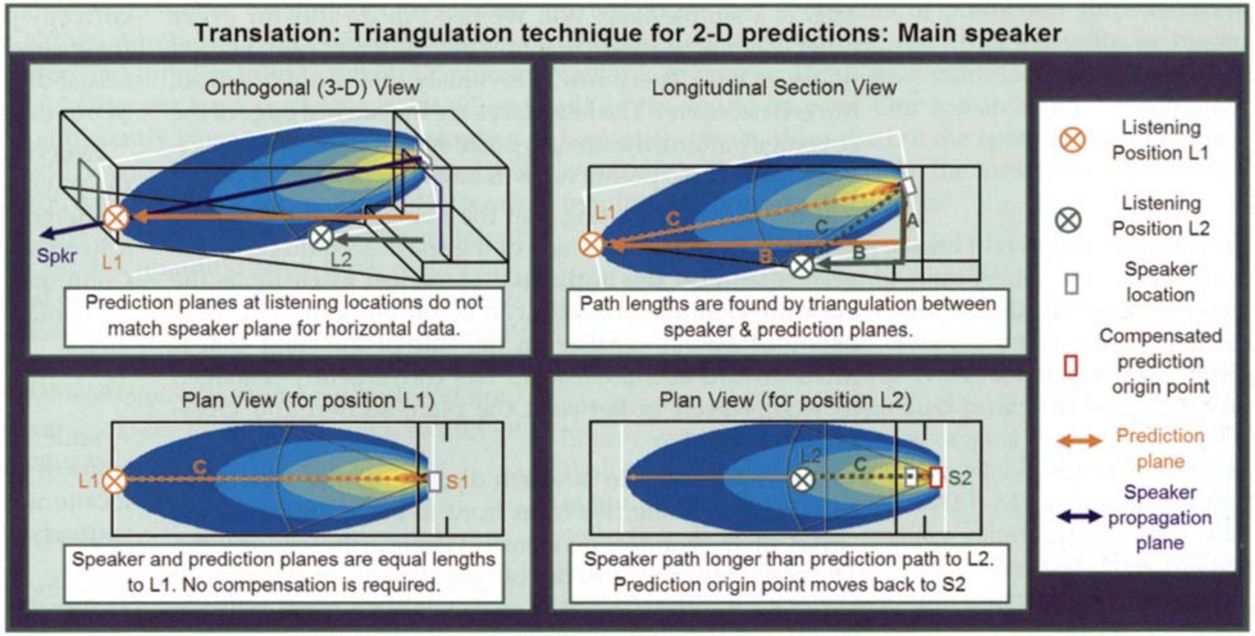


Figure 5.4 The use of triangulation to compensate for the planar mismatch in the main speaker

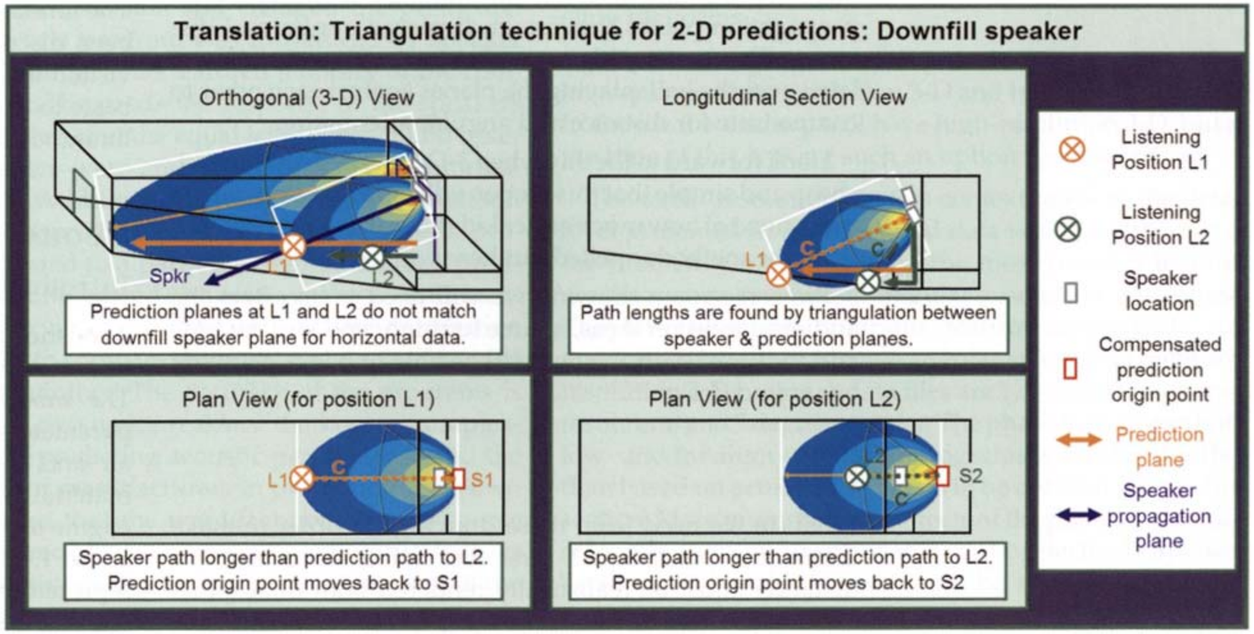


Figure 5.5 The use of triangulation to compensate for the planar mismatch in the downfill speaker

This is a simple task. Will we need to do this for every other listening position that falls on a different propagation plane, namely every row? Obviously this is not practical, nor is it necessary. The key places are the nearest edge of the coverage pattern, the on-axis position and the farthest edge. With these three points we will have the data we need.

It is also worth noting that the planes are interchangeable. The propagation plane of the speaker could be on the vertical and off the horizontal plane just as easily as the opposite example above. It can be off on both the propagation planes. A speaker on the side of the stage that is tilted inward and downward has conveniently managed to wedge itself in between the plan, section and cross-section planes!

One more scenario is worth discussing: the rake. Most halls have a gentle rise from front to back, known as the "rake" of the seating. This is easily visible in the section view. The plan view may also denote the changes in elevation at various locations. The rake makes it even more unlikely that the propagation and drawing planes will match. The reason that this is worth noting is that there are huge advantages in our sound design to positioning our speakers with propagation planes that are at displaced angles and distances. This is a very big part of how we will create uniform level over the hall, playing the planes against each other to compensate for distance and angular attenuation.

I look forward to the day when 3-D CAD design is so fast, cheap and simple that this section will be interesting only in the context of how much easier kids have it these days. Until then, the methods spelled out here will get the job done.

From the room drawing we will need to ascertain the audience location areas, source location, and surface type and location. With this we will be able to begin planning our sound design.

Scale

What do we need from these drawings in order to make reasonable predictions? First and foremost: scale. The distances must be realistically rendered both in absolute and relative terms. The overall shape of the room must be correctly proportioned, as this will be the decisive factor for speaker position, array type and coverage pattern. The actual distances must be realistic as these will be the basis of our estimates of how much power is required.

Every drawing has its scale shown somewhere, so that the physical size of the drawing can be related to what is rendered. Paper drawings, such as blueprints, will have a fixed scale, and can be read directly with a scale ruler. Computer drawings can be viewed at different scales dynamically as the zoom in or out of the drawings.

Scale Rulers

A scale ruler is used to ascertain distances. The dimensions of the audience shapes, the distance from speaker locations to audience areas, the distance between speaker sources are examples of uses for the scale ruler.

Protractors

The basic protractor can be used to provide an initial estimate of the angular aspects of the coverage requirements. The limitations of the protractor to our design efforts have been discussed previously in Chapter 1 and will also be visited in Chapter 6. There is still plenty of work for this basic tool, provided that we maintain awareness of these limitations.

Acoustic Modeling Programs

Introduction

Acoustic modeling programs seek to provide advance information about the performance of speakers in a space. The modeling is based on the measured performance parameters of loudspeakers, acoustic transmission through air and the measured reflective properties of building materials. Those properties are fed into a mathematical engine that simulates the interaction of the speakers in the space. Even with a perfect set of equations, and a perfect representation of the room dimensions, the accuracy of such modeling is something which can be approached and never reached. This is a matter of resolution. Our modeling is based on a finite number of data points to describe an entity that can be measured in infinite detail. The principal question is whether we have reached a point where the resolution is sufficiently close to our ear's perception capabilities. We will not need to factor in those parts of the wall that took two coats of paint if the acoustic effects are not audible. Our best hope is to have enough information to select speaker model, quantity, array type, placement, focus angle, signal processing allotment and acoustic treatment. As we have seen in the previous section, the properties of transmission, summation and reception will play the critical role in these choices. We can say with certainty that our prediction program must have a high degree of accuracy in these regards.

A Very Brief History

The traditional tools of prediction have been the scale ruler, the protractor and the manufacturer's data sheet. Coverage areas are mapped out on blueprints. In the 1980s, commercially available computer-based acoustic modeling programs became available and changed the prediction landscape. The first program was introduced by speaker manufacturer Bose Inc. and featured a library of the company's speakers. The predicted response of the speakers in a room is displayed in the equal level contour and other formats. Designers were anxious to obtain this information technology. Additional manufacturers followed suit and an era of proprietary programs followed. Each program was dedicated to a particular manufacturer's products. They were unique and required a major commitment in time, training and cash. The lack of standardization made it difficult to compare the predicted performance of one product to another. The accuracy of the programs is questionable for two major reasons: the known complexity of the task of predicting acoustic performance and the vested interests of manufacturers in promoting their own products. The fact that the manufacturers are the source of the performance data creates a credibility issue in regard to the motives, objectivity and scientific basis of the prediction programs. It seems obvious, in retrospect, that it would be best if creators of prediction software were disinterested third parties, not affiliated with, or biased toward, any particular speaker manufacturer. We are, after all, in the sound *business.* Our research and development is market-driven. Who else but the speaker manufacturers would be willing to undertake the massive time and costs required to support such efforts?

Concerns about creating a level playing field arose over time and calls for standardization began. Standardization would allow designers to access data libraries of competing manufacturers and compare and contrast speaker performance on equal terms. The first and foremost of the speaker manufacturer originated programs to open up its library was EASE™. This was originally designed by Dr Wolfgang Ahnert and is now available through speaker manufacturer Renkus-Heinz. EASE has a large functional collection of competing speaker manufacturer data and as a result presently enjoys a wide user base. New modeling programs have emerged from independent companies that are not aligned to any particular manufacturer. Some of these companies have agreed on a data format so that manufacturers can have their data used by multiple modeling platforms.

Overall, the programs fall into two basic categories: low- and medium-resolution 3-D and high-resolution 2-D. In the ideal world we will have high-resolution 3-D, but at the time of this writing such an option has not arrived.

The term "resolution" in this context refers to the detail level presented in the measured data which is the basis of the predictive model. Among the most relevant features are frequency resolution, angular resolution and phase. The low-resolution programs feature octave-wide frequency spans with 10 degree angular spans. The medium-resolution 3-D program data files are 1/3 octave frequency resolution and 5 degree angular. The phase response in both low- and medium-resolution programs is idealized rather than based on actual data. This will be detailed later in this section. More important, the effects of the phase on acoustical summation are an option that can be selectively invoked by the user. Would that we could be so lucky in the real world to be able to hear systems interact in such a way!

There is one widely available program with high-resolution data. This is Meyer Sound's MAPP Online. This contrasts with the other programs in that the calculations are not done in the user's computer, but rather at a central server. Scenarios are constructed in the host computer and sent to the server via the Internet for calculation and returned as graphics moments later. This allows for unprecedented detail in the simulations and is obtainable on virtually any computer. There are two very serious limitations. First is that only the company's own speakers are in the data library, and second, there are only two slices of 2-D data. The program does, however, incorporate measured phase and has extremely high resolution: 1/24th octave, 1 degree angular.

Speaker Data Files

Acoustic modeling begins with the speaker data file. Speaker performance parameters are compiled into a table of values which were obtained in a reflection-free environment, such as an anechoic chamber. From this core data the direct sound transmission of the speaker will be simulated in free field. The coverage pattern over frequency and the sound pressure level over distance are some of the many features which can be observed.

The detail level of the predicted response is limited by the resolution of the original data. The predicted response will show a coverage pattern as a continuous line encircling the speaker. The actual "known" data points from which the polar representations are derived come from the core speaker data. The continuous nature of the line comes from the mathematical process of **interpolation,** which "connects the dots" between known data points. Interpolation is the mathematical equivalent of a "leap of faith," assuming that the route traversed between the two known points does not deviate from expectations. Interpolation is a fact of life in acoustic data. Without it we will be looking at a series of dots on the screen. It is only the degree of interpolation that concerns us. If the polar plot is based on just four data points taken at 90 degree intervals, the extent of the interpolation will leave us with serious concerns about what actually happened between the known points.

Resolution

As mentioned briefly above, a key parameter in the makeup of the speaker data file and prediction processing is its two forms of resolution: frequency and angular. Both of these forms must be sufficiently high for us to see the details of the summation interactions. Both of these have two forms: acquisition and display. The resolution of the acquired measurement data will be a limiting factor in the processing and display of the speaker behavior. The modeling program display will use interpolation between the data points to draw lines between the known points. The accuracy of the program decreases as the span of interpolation increases.

Frequency Resolution

The most common prediction response is the equal level contour. This is a polar representation, typically separated by colors. Each plot represents a frequency range; in essence, a slice of the speaker's frequency response spread over the space. A full characterization is derived from the complete series of layers that describe each range. The quantity of layers we use to slice the response is the frequency resolution. The display resolution is limited to no more than the acquisition resolution. The display resolution can be less than the acquisition by adding the layers together to get a composite display over the desired range. The slices can be octave, 1/3 octave or any user selected percentage bandwidth, up to the acquisition limit.

A well-designed single low-order speaker in free field should have a relatively stable response shape over frequency. This is the most forgiving case for low-resolution data. Increased resolution will show details, but not significant differences. As speaker order increases, the rate of response shape change over frequency rises. In such cases low-resolution data will have lower resemblance to its high-resolution counterparts. An illustrative example is shown in Fig. 5.6.

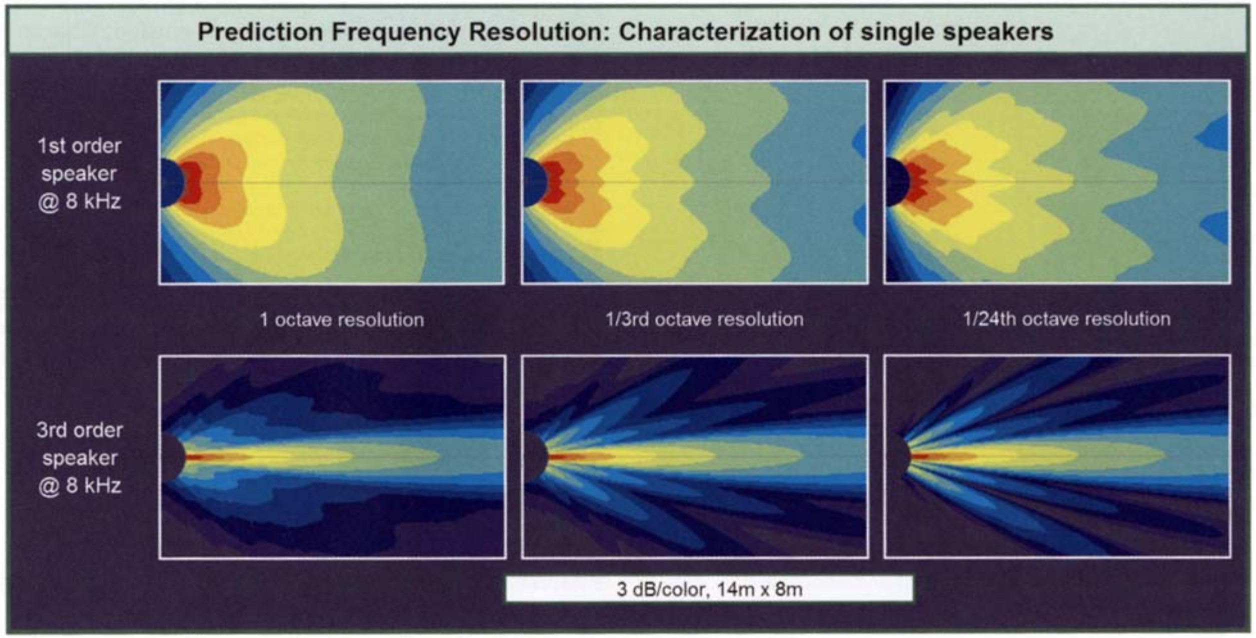


Figure 5.6 Frequency resolution issues related to characterization of a single speaker. Top first-order speaker shows only minor differences between low and high resolution data. Bottom third-order speaker shows substantial differences in side lobe details

Low-resolution data is related to the family of high-resolution data that comprises the same bandwidth. For example, octave bandwidth data spans virtually the same range as the three 1/3rd octave spans in that range. 6th octave resolution logically follows as double the number of slices and so on. Therefore, if we look at three 1/3rd octave slices, we can see the family members that, combined together, comprise the octave response. Once again we see (Fig. 5.7) that low-resolution data can remain representative, as long as the parts which make it up are closely related. As the components diverge, as in the case of a high-order speaker (Fig. 5.8) the low-resolution response loses it resemblance to its component parts.

When summation occurs the low-resolution data will smooth over the combing and remove it from view. In such instances we can create a false sense of equivalence between design strategies that have very different amounts of combing zone summation. Two examples are shown in Figs 5.9 and 5.10. The high resolution of the frequency axis gives us sufficient detail to view features that will be noticeable to our hearing, while the low-resolution predictions remove the evidence.

Angular Resolution

Polar data is a series of points evenly spaced around a circle or sphere. The spacing, in degrees, is the **angular resolution.** The source data for the program is acquired by measurement of a device at a given resolution. The accuracy of the angular representation of the displayed data is proportional to the source resolution. Once again the data will include some degree of interpolation in order to create smooth continuous curves.

The accuracy of the characterization will depend upon whether there are sufficient data points to track the angular rate of change in the measured device. Speakers with a low rate of change can be accurately characterized with

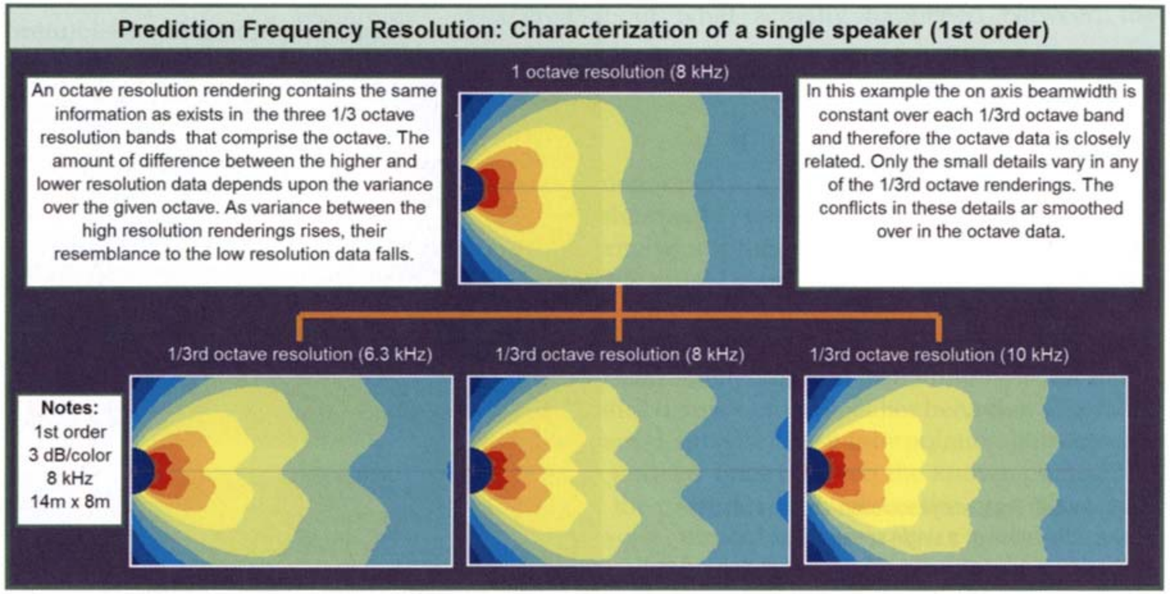


Figure 5.7 First-order speaker at octave resolution compared to The differences are confined to the details

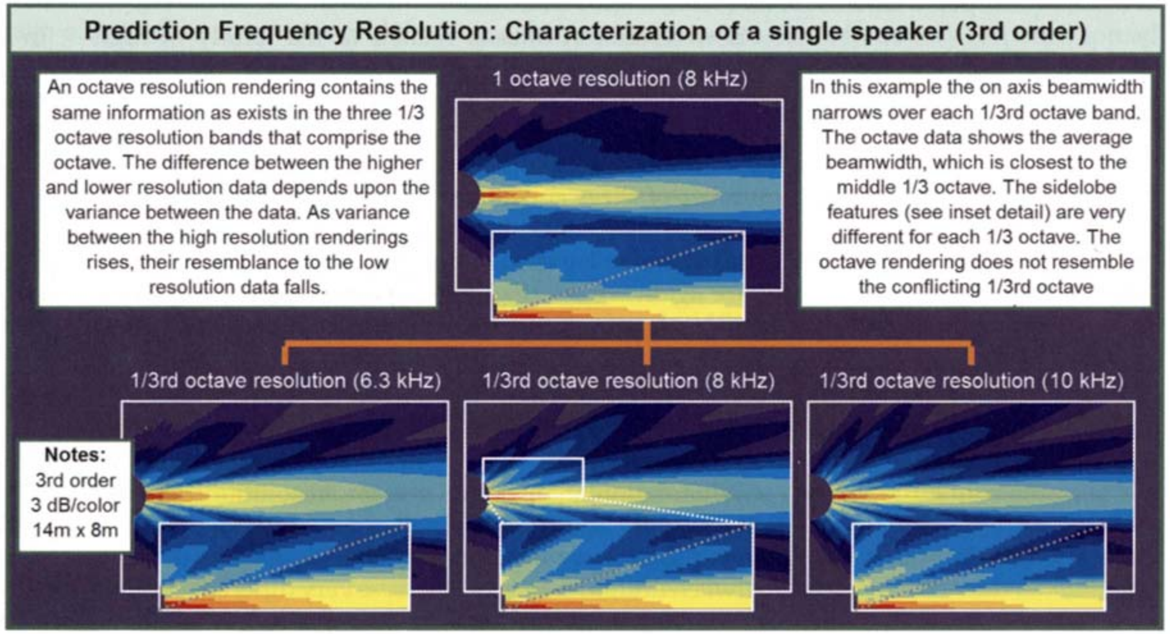


Figure 5.8 Third-order speaker at octave resolution compared to the 1/3rd octave components that comprise the same frequency range. The differences are found in the side lobe details and in the beamwidth

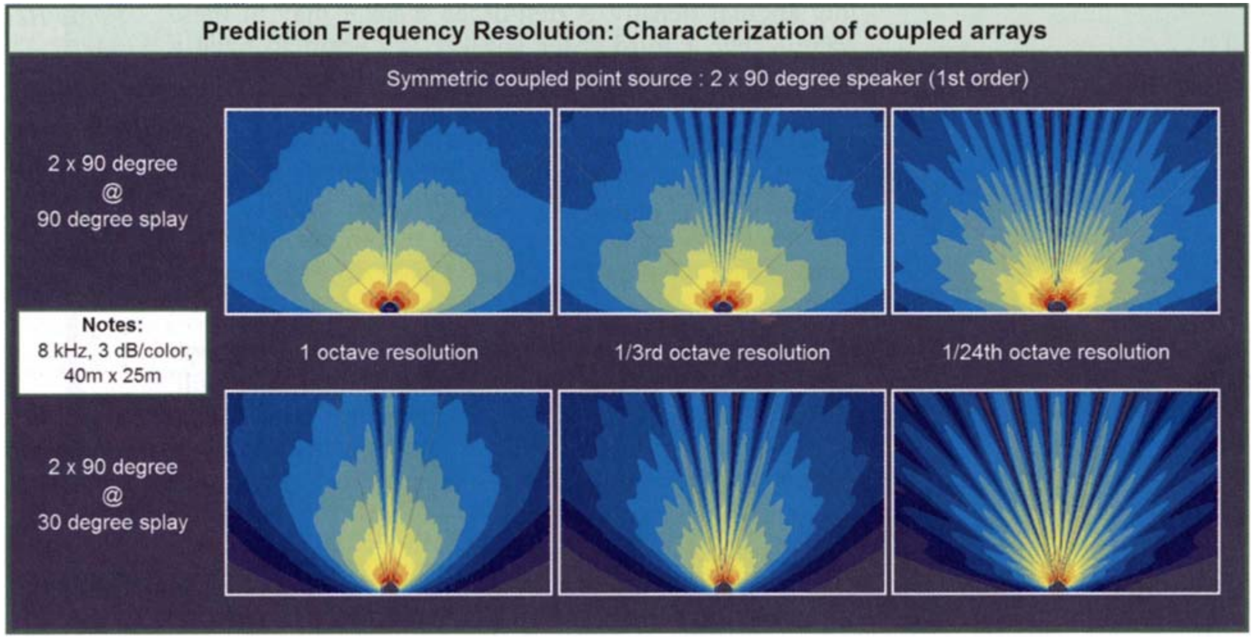


Figure 5.9 Frequency resolution issues related to characterization of summation in coupled speaker arrays. Top non-overlapping array has minimal comb filter summation. This is best revealed in the high-resolution rendering, where the combing zone and isolation zone interaction can be clearly differentiated. Bottom the overlapping array appears to have a narrowed shape but equivalent combing in the low-resolution prediction. The high-resolution panel shows the combing zone interaction to be dominant over most of the coverage area

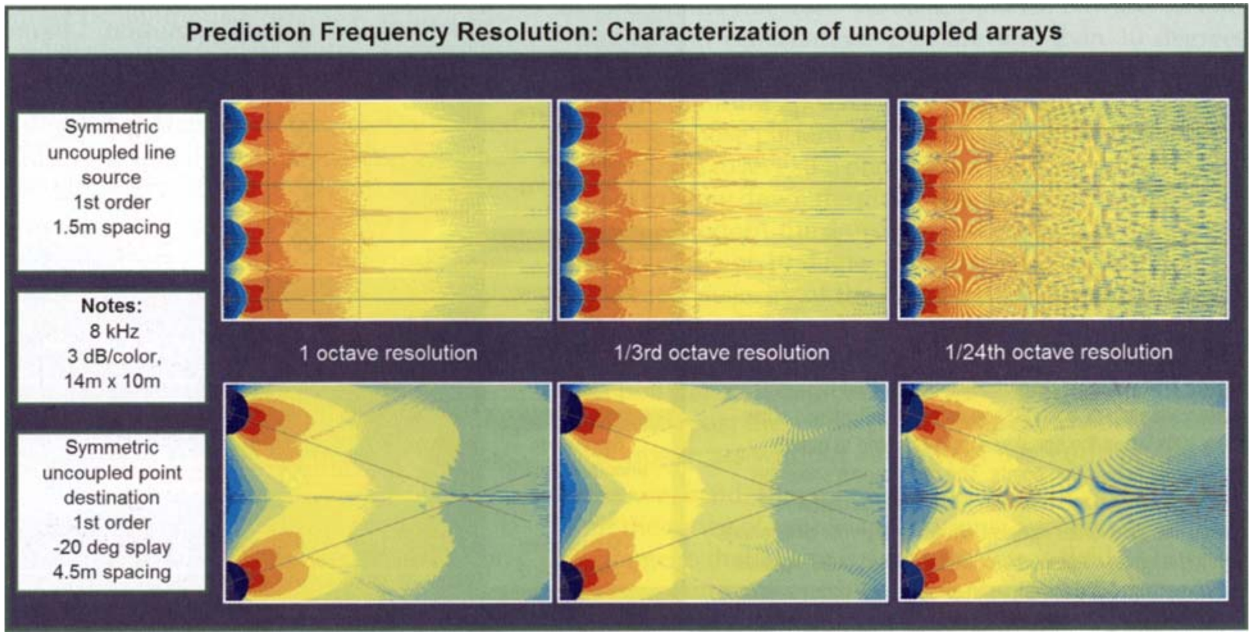


Figure 5.10 Frequency resolution issues related to characterization of summation in uncoupled speaker arrays. The lower resolution renderings do not reveal the high level of ripple variance from the speaker interaction. With low-resolution predictions the overlap areas appear to be the most uniform. The high-resolution data reveal the overlap area to be the least uniform

low angular density. A first-order speaker may fit these criteria, but a third-order speaker is certain to have a high rate of change over a small angular area. The first-order speaker is not exempt, however. Many first-order speakers have horns with substantial ripple variance spread over the coverage angle (seen as lobes and fingers). Low angular resolution, especially when coupled with low-resolution frequency acquisition, can create a false equivalence between speakers of vastly different audible quality.

Low-resolution data, either angular or frequency, will reduce the investigative quality of our predictions. It will, however, provide much more optimistic views of the response, as combing is removed by angular smoothing, frequency smoothing, or both. We will benefit by throwing every bit of resolution available to our calculation and display. Using low resolution provides a video solution for an audio problem.

Phase Response

If we have no phase data, the response of a single speaker can be viewed without significant risk. Summation characterizations without phase data violate the second law of thermodynamics. Energy is created (the positive side of summation) without a compensating loss (the cancellation side). The result is a response that we cannot possibly observe with our ears or our analyzer. The role of phase response in summation modeling will be covered shortly.

Common Loudspeaker Format

In an effort to standardize the data available for acoustic modeling a format has been created which allows any manufacturer to submit data that can be used for several programs. In keeping with the audio industry rule — that two standards are always better than one — we have two formats: common loudspeaker formats 1 and 2 (CLF1 and

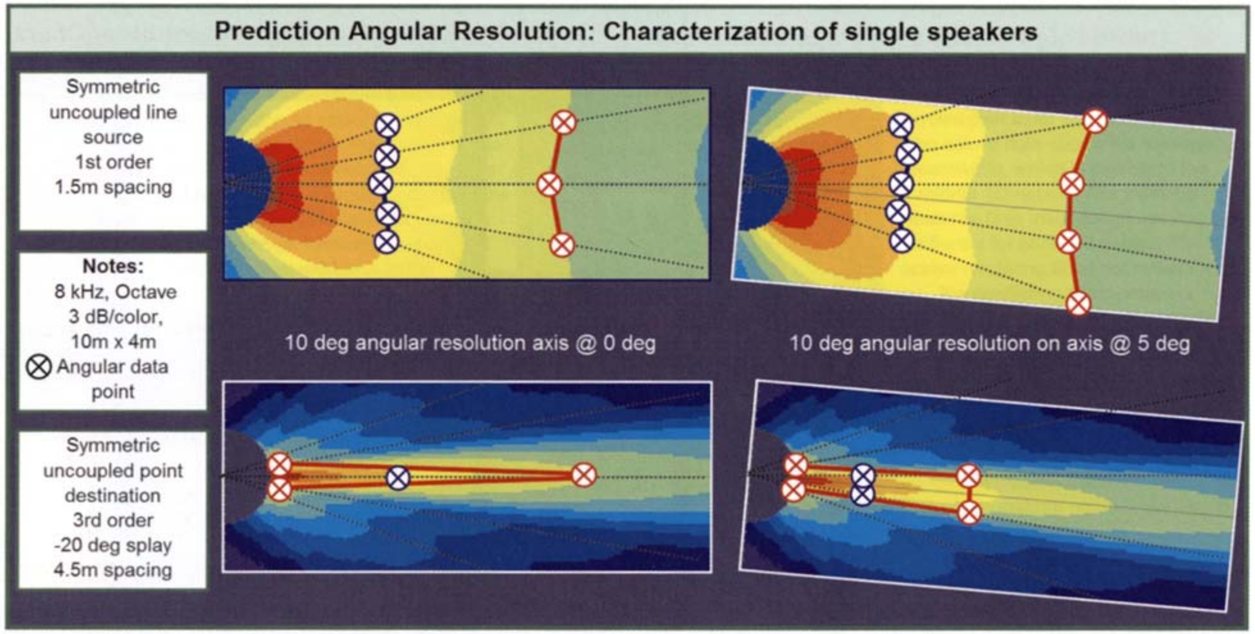


Figure 5.11 Angular resolution issues related to characterization of a single speaker. The left panels show the results when the speaker axis is oriented to match the angular acquisition points. The right panels show the speaker axis oriented between the angular acquisition points. Top first-order speaker shows only minor differences with the angular displacement. Bottom third-order speaker shows substantial differences in shape. High-order speakers require high angular resolution for accurate characterization

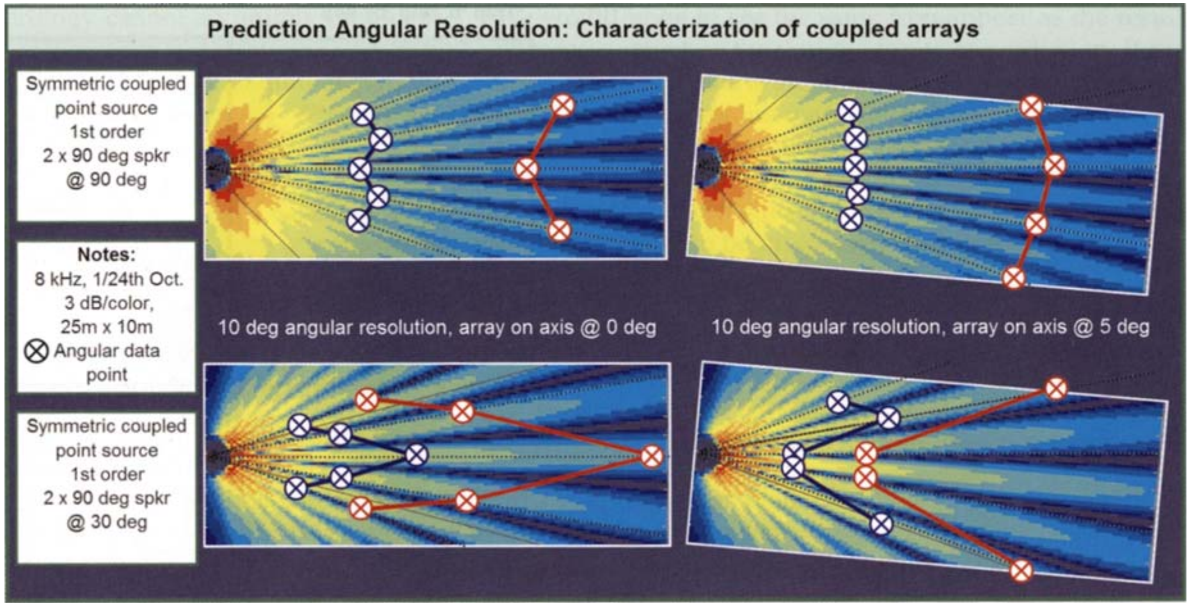


Figure 5.12 Angular resolution issues related to characterization of summation in a coupled speaker array. The right panels show the array axis oriented between the angular acquisition points. Top the non-overlapping array shows only minor differences with the angular displacement. Bottom the overlapping array shows substantial differences in shape. Overlapping arrays require high angular resolution for accurate characterization

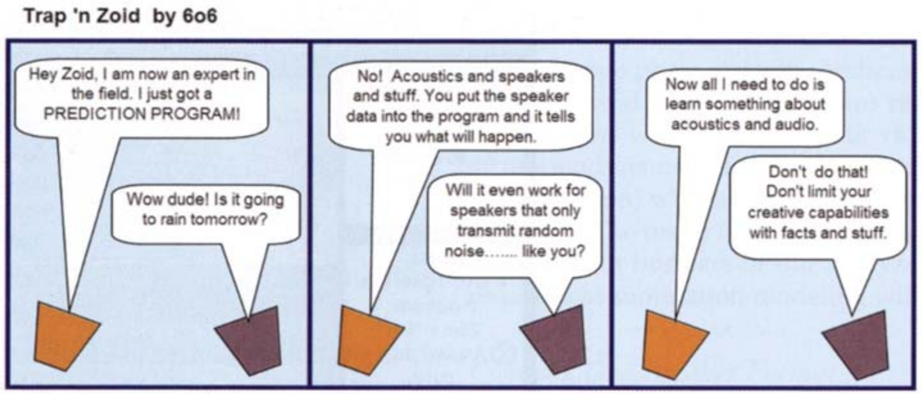
CLF2). The first one is low resolution (octave, 10 degrees) and the second is medium resolution (1/3rd octave, 5 degrees). The data is acquired by an impulse response so the phase response could ostensibly be accessed by the programs.

There are a great many commendable features to CLE The first and foremost is that fantasy data cannot be entered into the system by marketing departments. The data must be certified as having been actually measured within certain standard conditions and approved facilities.

CLF1 and CLF2 can be used to create 3-D spherical modeling. In the best case the speaker is measured as a full sphere. In the worst case the spherical data is interpolated from the horizontal and vertical planes. Notably the data files inform the user when spherical interpolation is taking place.

Acoustic prediction as a whole stands to benefit from the implementation of a standard. Anything that levels the playing field between competitors has positive potential. This benefit is only fully realized if the standard is sufficiently rigorous to provide the key data that differentiates the performance of systems and designs. CLF1 is of such low resolution that we must be very cautious about any evaluations made with it. The features of wide coverage (first-order) speakers will be discernible as the number of data points can trace the basic pattern. Beware. Details of lobes and cancellations that are less than 10 degrees wide or less than an octave span will be glossed over. A far inferior-sounding speaker may appear on equal footing with one whose pattern is more uniform. The intention of a standard is to level the playing field, but we must be careful not to plow down the goal posts at the same time. Consider the modern third-order speakers which commonly have patterns of 10 degrees or less. (This description fits the greater majority of the speakers which are commonly referred to in our industry as "line array" boxes.) That is ±5 degrees from the on axis point. CLF1 finds its first off-axis data point at ±10 degrees, where the speaker has already fallen well past the -6 dB point. How can we design unity class crossovers without visibility of the —6 dB point?

Octave band frequency resolution will also severely limit the range of usable applications. Specifically it limits the range to that of single speakers in free field. Why? The reason is summation which, as we have learned, requires detailed frequency response resolution for us to properly



interpret the asymmetrical peak and dip action and the linearly spaced combing. Octave resolution levels the playing field again as it plows through the combing and leaves us unable to see the difference between an optimized design and one with massive combing.

CLF2, comparable (although not compatible) with EASE, has higher resolution. The change is one of incremental scale. The frequency resolution is 3X and the angular resolution is 2X that of CLF1. This is an obvious improvement, but is it enough? Based on my experience, the answer is no. These improvements of resolution are still not sufficient to overcome the obstacles described above for CLF1.

Perspectives: For 20 years, I have been struggling with the art, the alchemy, the voodoo of system design, and how it relates to final tuning in the room. The hard part is to communicate my needs to the client and the architectural team without any historical data (or even a room to measure). How do you talk about invisible waves in an environment? A good dose of experience, prediction software, and audio analysis done in real-time is my recipe to design, prove, and improve all my audio system.

Francois Bergeron

Other Formats

The final format we will discuss in the prediction arena is MAPP Online. This is a 2-D only prediction that is extremely high resolution with complex data. MAPP uses two slices, vertical and horizontal, with speaker data measured at over 1/24 octave, 1 degree angular resolution. This program also utilizes the actual measured phase response. Therefore different speaker models can be combined with realistic summations. MAPP is so hugely superior in its speaker summation accuracy that nothing else that I have seen even comes close. There are, however, three severe limitations at this time: the speaker data library has only Meyer Sound speakers, the room renderings can only create a single square or rectangular shape, and all data is 2-D. These limitations obviously preclude its use as a universal tool.

Modern sound designs are reliant on highly directional speakers that are arrayed in large numbers at overlapping angles. Minute changes of 1 and 2 degrees have very demonstrable effects on the array performance. Low-resolution, phase-free data cannot provide the critical details needed to make the fully informed decisions about such systems. My crystal ball (the old-fashioned form of prediction) says the future holds a standard format with high-resolution complex data. Could that be CLF3?

Acoustic Transmission Properties

Direct Sound Transmission

The direct sound transmission is assumed to have the standard free-field loss rate of —6 dB per doubling of distance. Most programs assume the transmission loss to be uniform over all frequencies, although we know that this is never the case due to the HF air absorption loss. In order to accurately compensate for the frequency response changes over distance we must have information regarding the temperature and humidity in the space. Some programs include the environmental factors with user-adjustable temperature and humidity. Even if such environmental compensation has perfect mathematical accuracy, it is still of limited use practically. For indoor applications we can use a nominal "room temperature" and standard humidity but even these cannot be guaranteed to occur. For outdoor applications the situation is wildly dynamic. Professionals in the field of meteorology cannot accurately predict the weather, so it is unrealistic for us to attempt to factor this in to design decisions beyond the nominal standard conditions.

Ray-Tracing Model

The paths of the sound propagation are typically computed by the ray-tracing method. The sound moves outward from the source in a straight line as would rays of light from the sun. In the free field, the sound continues outward and steadily loses level. The relative level values for each ray are adjusted to mimic the coverage pattern shape of the particular speaker.

If the ray strikes a surface it is reflected at the angle of incidence as with light off a mirror. The rays and reflections from additional speakers will intersect and pass through those of the original source. This is the typical approximation of transmission and summation properties of sound available in modeling programs. This approach is sufficient for us to gain huge amounts of advance information about the potential performance of a system, and to greatly enhance our chances for a smooth and successful calibration process. During that final stage, the system's response will be measured, and the results will hopefully resemble the prediction model in the most important parameters. We cannot rely on the prediction model to be perfect, nor do we need it to be. It will be useful, however, to bear in mind where we should expect the results to match prediction, and where we need to go it alone and rely on the on-site measured data exclusively. To make the point let's consider the weather-related effects. Even if the prediction program can factor in these effects, it cannot predict the weather. Ongoing monitoring and compensation of weather-related effects will require measurement.

A series of figures begins here which illustrates the expected differences we would encounter in the field between the ray-tracing prediction model and a high-resolution measured frequency response. (Analyzers will be discussed in Chapter 8.) The summation zone frequency response icons introduced in Chapter 2 reappear here. We would hope to see the same icon appear as the result of measurement or the various forms of prediction. If we see a difference, this is significant, since these summation icons are the milestone markers in our understanding of the system response. A change in the representative icon between the measured and predicted responses would be a source of concern since it may lead our design process off-course, only to be discovered during the optimization process.

Each figure contains the expected measured result and three versions of the ray-tracing model-predicted results: high-resolution, low-resolution, and phase-free. As we will see, the highest correlation will always be the highest resolution, and the lowest is the "no phase" response.

Refraction

The refractive properties of sound pose a particularly difficult challenge for acoustic modeling. Acoustic refraction, like its counterpart in the physics of light transmission, is the bending of a sound transmission as it passes through layers of media. In practical terms this will likely only be a substantive effect for us in large-scale outdoor applications. Thermal layers in the atmosphere can cause a bending upwards (cool air over warm air) or downwards (warm over cool). Wind conditions can also cause highly variable refractive effects that bend the sound on a moment-to-moment basis. For a single speaker source the sound can noticeably change its response, sounding as if we are moving slightly off-axis and returning. Since the transmission is being bent that is exactly what is happening: sound that was directed away from us is bent into our location. The effect is much more noticeable on coupled speaker arrays, particularly those with a high percentage of overlap. The refraction effectively performs a changing angular orientation between our position and the cluster. The result is movement through the volatile comb filter interaction of the summation which shifts its filter frequency, a process known a "flanging." The anticipation of weather conditions is beyond the abilities of our prediction modeling and we will simply have to deal with this as an operational issue.

Figure 5.13 Comparison of the ray-tracing prediction models with the expected high-resolution measured response. Top direct sound path on axis response. Middle filtered reflection. Bottom unfiltered reflection

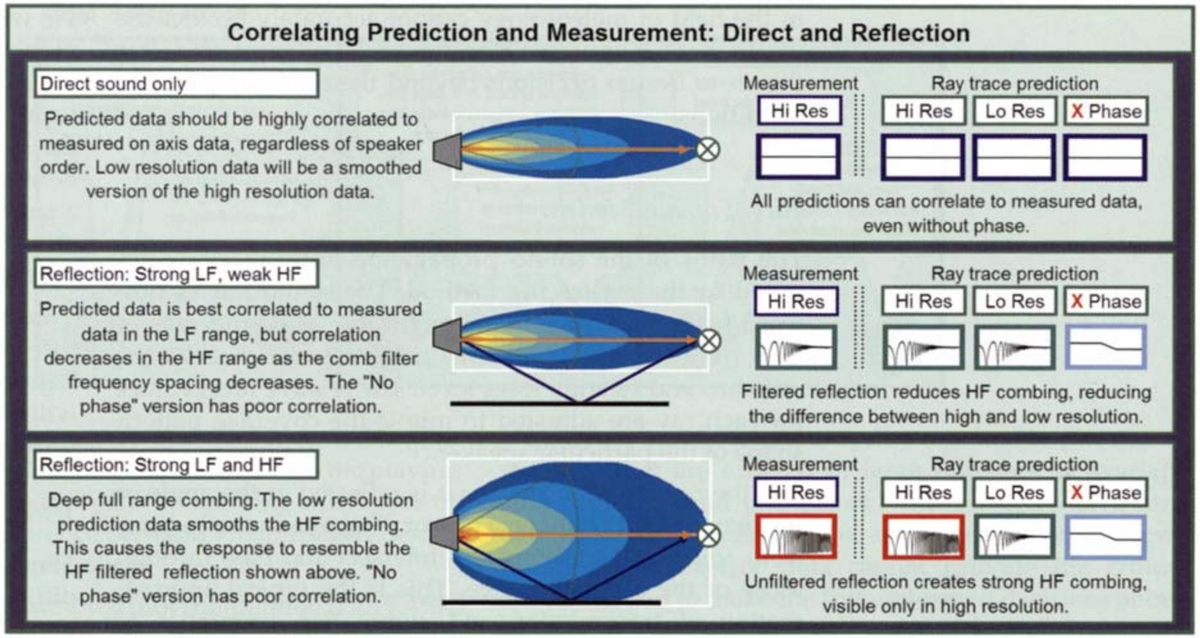
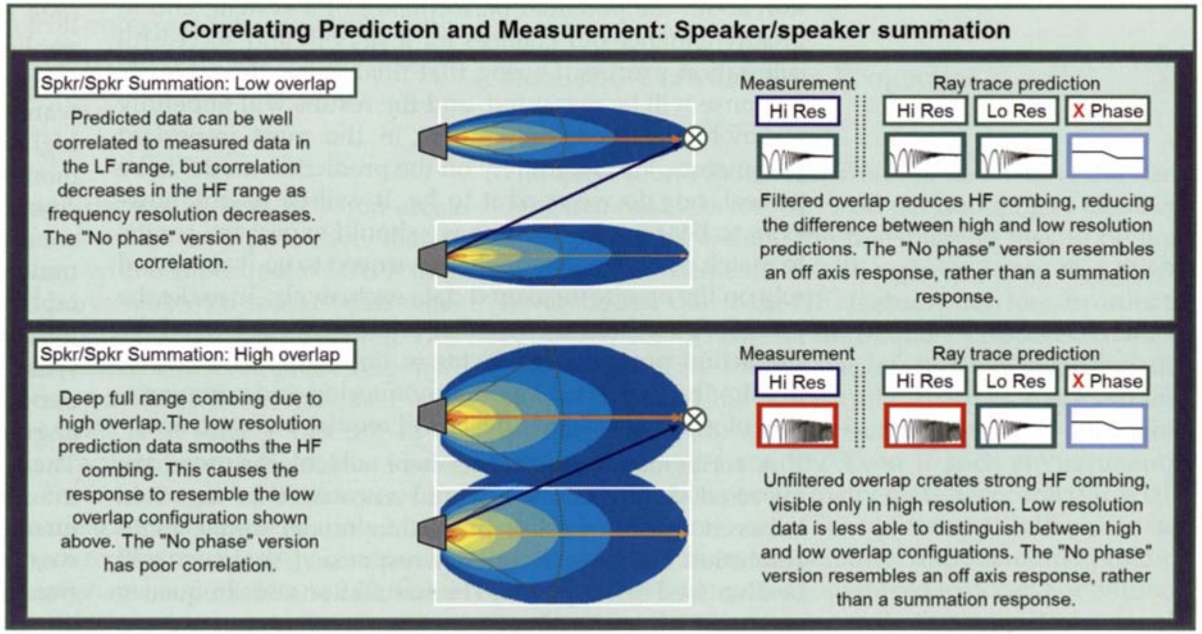


Figure 5.14 Comparison of the ray-tracing prediction models with the expected high-resolution measured response. Top speaker interaction with low percentage of overlap. Bottom speaker interaction with high percentage of overlap



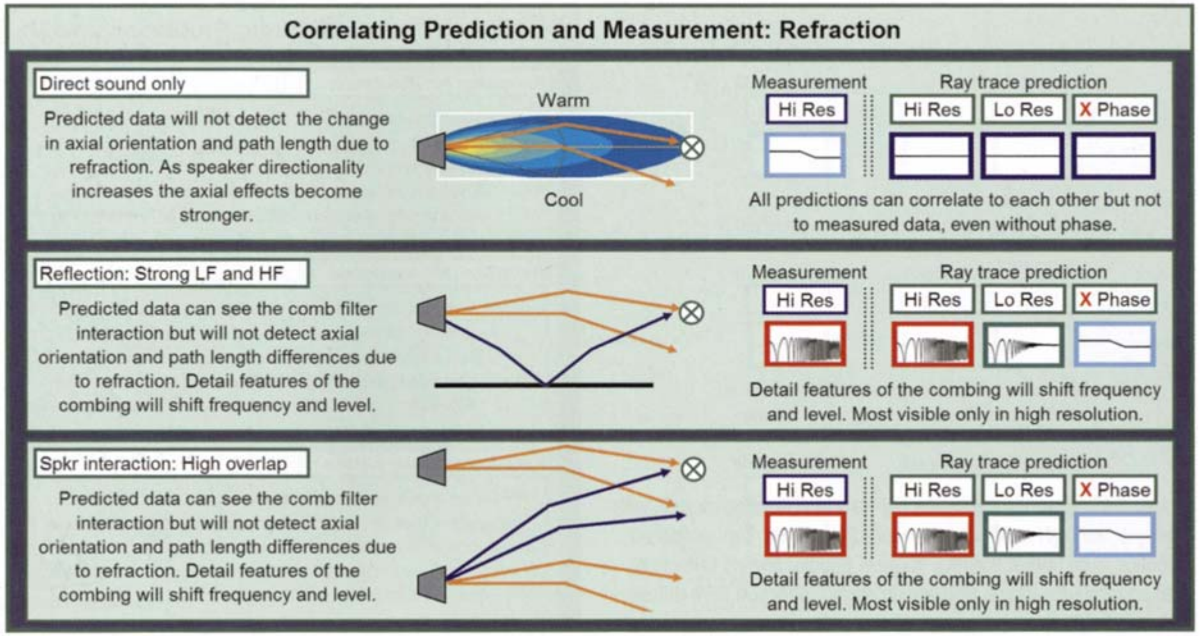


Figure 5.15 Comparison of the ray-tracing prediction models with the expected high resolution measured response. Top refraction of the direct sound path. Middle refractive effects on unfiltered reflection. Bottom refraction of speaker/speaker summation

Diffusion

For a surface to reflect all frequencies perfectly evenly it must be large enough for the largest wavelengths to bounce off it like a mirror. Complex surfaces reflect the sound in a manner that scatters the sound in different directions over frequency. This type of reflection is termed diffusion and differs from the simplistic ray-tracing-type model. Surfaces which contain raised and lowered areas of various sizes and angles present a variable face to sound waves. Because the sound waves vary in wavelength the differing elevations have different effects over frequency. An example of a diffusive surface is the series of statues along the side walls of the Boston Symphony Hall. These statues reflect the sound in a variety of directions, and change over frequency. Engineered diffusive surfaces are commercially available. These surfaces are designed to specific ratios of dimensions that create a controlled series of diffusive effects over frequency. One might say that these surfaces seek to create a uniform field of non-uniformity. Diffusive surfaces are extremely complex to model for the very reasons that make them so desirable acoustically: a complex scattering effect that differs at every frequency. The properties of diffusion are so complex that we should be prepared to encounter them as they are measured in the field, rather than rely on their accurate characterization in prediction.

Diffraction

If there were a brick wall that was two meters tall between us, it would be impossible for you to see me. And yet we could carry on a conversation without too much difficulty. Diffraction at the top edge of the walls allows us to hear each other. If there is an open window in our room we will source all sound as coming from the window, even though we have our doubts that a cat could be howling from the ledge of our fourth-story apartment. Diffraction of the sound through this opening creates a virtual source and the sound is retransmitted from there. It is only when we stick our heads outside the window, and move to the

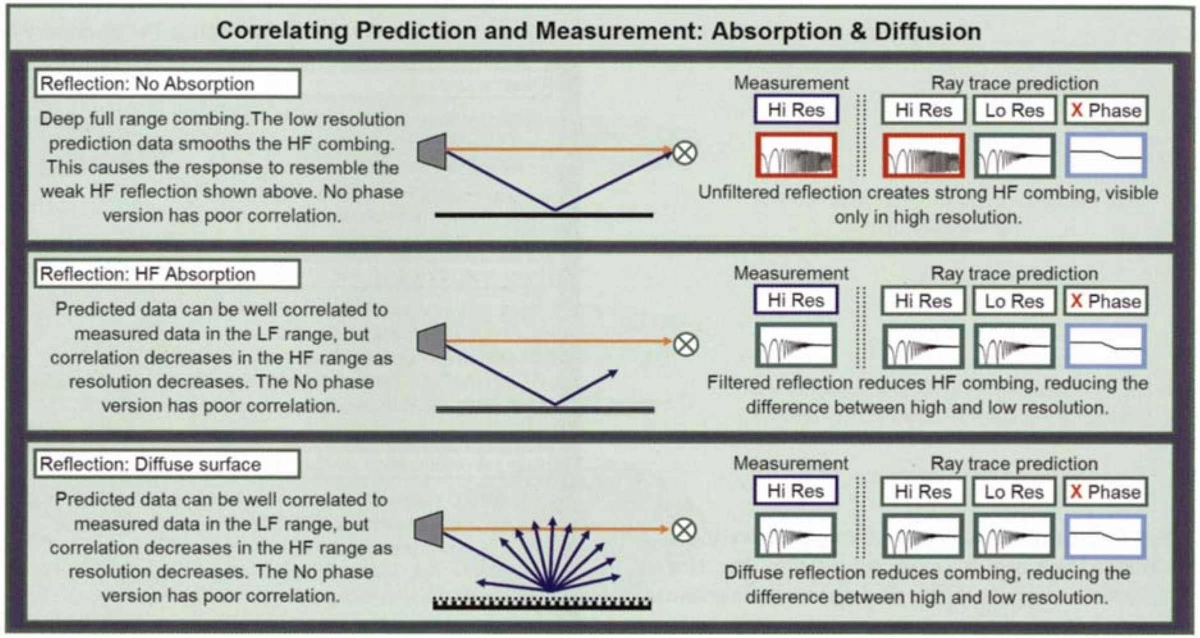


Figure 5.16 Comparison of the ray-tracing prediction models with the expected high-resolution measured response. Top unfiltered reflection from non-absorptive surface. Middle filtered reflection from absorptive surface. Bottom unfiltered reflection from diffuse surface

other side of the diffractor, that we will localize the offending cat on the ground.

The properties of diffraction are complex and beyond the scope of this book. The most basic properties are the relationship of the opening and/or barrier size relative to the wavelength. Small barriers (compared to the wavelength) will be passed by with minimal effect. Large barriers (compared to wavelength) will reflect a substantial portion and leave a shadow area of low-level sound behind the barrier. Any physical barrier will have varying diffractive properties over frequency since it presents a different ratio of size/wavelength over frequency. A small pillar will block only the very high frequencies while others work their way around it. Standing behind a pillar that is 30 meters in diameter will be nice and quiet. Prediction programs have a difficult time modeling diffraction effects. Every obstacle in the transmission path that is not a solid continuous surface will need to have different reflective properties over frequency. Any openings in a surface will need to be modeled as secondary sound sources. Fortunately this loss of accuracy is not as critical for our application as it would be for analysis of sound isolation. Our application does not often call for concert sound to be transmitted through window openings. Likewise we would not be expected to tune our system from behind a pillar or barrier.

On the other hand, it is not unusual for us to encounter structural steel in front of our speaker clusters. In such cases we will encounter high-frequency shadowing, while the low frequencies freely move past the steel. The ray-tracing model will either see this as a full range shadow or as non-existent. In practice the final position of our speakers is subject to change as we fine-tune the focus angle. Time spent predicting the exact placement and frequency range of the sound shadow would be better spent petitioning for a creative structural solution that removes the obstruction. The properties of diffraction are so complex that we should be prepared to encounter them as they are measured in the field, rather than rely on their accurate characterization in prediction.

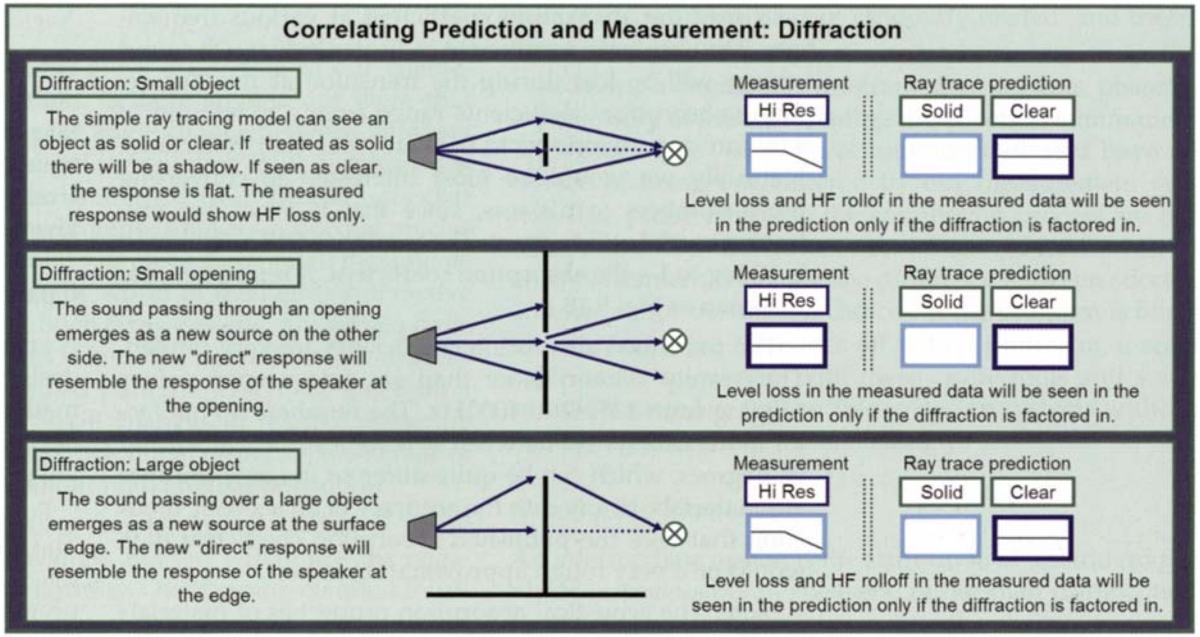


Figure 5,17 Comparison of the ray-tracing prediction models with the expected high-resolution measured response. Top diffraction around a small object. Ray-tracing model may either characterize the object as solid or clear. Middle diffraction through a small opening. Bottom diffraction around a large object such as over a wall

Perspectives: The fact that good sound is so rare today shows us that we have a long way to go on the education front. Most bad systems today are caused by designer or operator error: poor speaker coverage, poor time alignment and/or poor gain structure. Good equipment and powerful analysis tools are available and affordable, so there is little excuse for these problems any more.

John Huntington

Resonance

Another acoustic property of rooms is resonance. Resonance is the result of room dimensions that fall into a spacing relationship with a particular set of wavelengths. The result is that these wavelengths can become so highly efficient that they sustain long after others have decayed. Resonant cavities inside the room can present a problem for the decay characteristic, again by prolonging certain frequency ranges. External resonant chambers coupled to the room can provide effective absorption of low and mid-range frequencies. These specially designed cavities are known as "Helmholtz resonators" after the German scientist who quantified their effects. The effects of resonance would be very challenging to account for in modeling programs. If they are accounted for in the programs, to what extent are they practically useful for the design of sound systems? What decisions would we change regarding our sound system if this is omitted? None. The resonant properties relevant to our application will show up when the building is done, our speakers are in place and we are ready to tune.

Materia] Absorption Properties

The ray-tracing method is the standard for reflection modeling. A reflection bounces off a 100 per cent reflective wall and continues onward with only its additional air transmission loss. If we are fortunate enough to have surfaces with some absorption this will need to be incorporated into the model.

Absorption Coefficient

The acoustic properties of building materials are compiled into value tables, much as we had in the case of the speakers. Where do the values come from? Primarily these values are sourced from the manufacturers of building materials. Their published numbers are the results of standardized testing performed for the building trade. These tests find the **absorption coefficient** at various frequencies. The absorption coefficient indicates how much sound energy will be lost during the transition at the surface. These absorption coefficients range from a maximum of 1.00 (an open window) to 0.00 (100 per cent reflection). Naturally we would be most interested in converting these numbers to dB loss, since that is what our ray-tracing model will support. The number of dB attenuation is 10 log X 1 — the absorption coefficient. A reference table is shown as Fig. 5.18.

The published absorption coefficients are very limited. They rarely contain more than six octave span values ranging from 125 Hz to 4000 Hz. The numbers do not factor in the differences between striking a surface at 9.0 and 90 degrees, which can be quite different in some (but not all) materials. In any case the accuracy of an acoustic modeling that uses the published absorption coefficient data would be a very rough approximation at best.

Note: The acoustical absorption properties of materials are a complex subject that fills volumes in its own right. Readers wishing more details are advised to seek out the many books on this subject.

Absorption Level Loss Reference	
Loss (dB) = 10 x log (1-α)	
Absorption	
Value (dB)	Coefficient (α)
-0.10	0.02
-0.25	0.06
-0.5	0.10
-1.0	0.20
-1.5	0.30
-2.0	0.37
-3.0	0.50
-4.0	0.60
-5.0	0.69
-6.0	0.75
-7.0	0.80
-8.0	0.84
-9.0	0.88
-10.0	0.90
Total	1.00

Figure 5.18 Absorption coefficient level reference (approximate). The above reference chart converts the absorption coefficient (a) into dB loss for a single reflection. The ray-tracing model may computations like this simulate the loss in the reflected ray as it exits the reflective/absorptive surface

Surface Detail Omissions

Let's say we have valid material data parameters for all of our surfaces. How much of an error or omission does it take to affect our data? Not much. Our modeling will only have the major features of the room. Details of building structures will be omitted for practical reasons. Another layer of error is added by the omissions.

Characterizing Summation

A single speaker in free field would be the only system design that would not require characterization of summation. Neglecting the fact that anechoic chamber music never became popular, why would we even need a modeling program for such an application? Point and shoot.

For all other applications the summation is omnipresent. The importance of managing summation to the success of the optimized design cannot be overstated. The modeling program is of extremely limited use to us if it does not provide an accurate view of complex summation. This requires high-resolution frequency response and phase.

Frequency Resolution

The importance of frequency resolution to our characterization of summation was covered in Chapter 2 but bears repeating. Octave band resolution was shown to be capable of characterizing the summation interaction of signals which were no more than one wavelength apart. As the signals move further apart the peaks and dips become narrower and will be smoothed over. It is important to visualize this critical limitation. Octave resolution at 1 kHz limits us to path differences of one foot (300 mm) before we start losing a detail. How are we to characterize the summation of a complex speaker array if the details of the interaction are smoothed over? How can we characterize speaker/room summation if paths longer than a single wavelength are smoothed? The fact that these summation effects are clearly audible as tonal distortion (see Chapter 3) should give us pause in consideration of any data of such low resolution. One third octave resolution expands the vista to three wavelengths, but this still makes for a very limited view.

Angular Resolution

Angular resolution also plays a key role to our characterization of summation. Is there anyone out there that believes a 5 degree splay angle change between speakers is negligible? If our angular resolution is 10 degrees, such changes will likely be smoothed over. Modern speaker systems have machined rigging frames supporting precise splay angle changes of 0.5 degrees. Such small increments matter a great deal in the world of the highly interactive third-order speakers. Such array designs can contain large quantities of speakers within the span of a few degrees and can create combined patterns which are less than a single degree wide. Small changes in relative angle are critical to the combined directional response, the details of which will be lost without high angular resolution.

A first-order speaker can be aimed within 5 or 10 degrees of its optimal angle and few will notice any negative effects on the uniformity. Overlapping coupled third-order arrays that are aimed with absolute precision can have the highest levels of uniformity imaginable. An error of a few degrees throws it all away.

Phase Response

Without phase the summation is characterized as purely additive. No cancellation. Even polarity reversals or delay line settings don't matter! The relative levels are added up by the 20 log dB addition formula (Fig. 1.10). Each added speaker has the potential to make the combined system louder at all frequencies at all locations. The representation is derived from the amplitude summation as if we are always in the coupling zone no matter how much phase delay there is between the arrivals. The impossibility of this occurring is detailed in Chapter 2. There may be applications where a phase-free addition model is a valid characterization of the perceived response. Those applications would need to be ones where the phase relationship between the summed sources is random, such as a series of uncorrected sources. The interactions we are concerned with are correlated summations (copies of the source signal) such as reflections and other speakers. For these interactions the phase component is not randomly related, and therefore cannot be discounted.

In short, a prediction program that neglects phase is capable only of misrepresenting the effects of summation. The "phase-free" equal level contours indicated have no chance of being encountered by our measurement system. The phase effects to the summation process are not merely an additional feature that we can choose to view as an enhancement. They are the difference between success and failure for our design choices. If the program is blind to summation we are better off with a protractor, a scale ruler and a calculator. With these crude tools and some common sense, we will be able to design systems without neglecting phase, albeit painstakingly.

Quasi-phase Response

An intermediate option is the attribution of a standard flat phase responses on all speakers, rather than using actual measured responses. In this scenario the phase is assumed to be flat at all frequencies and the relative phase between two sources is incorporated, based on the relative time offsets. For example, the sources are 1 ms apart, the response shown is as if it were an electrical summation with the comb filtering that results with 1 kHz phase cycling. This is a huge improvement over the phase-free computation. If the design incorporates matched speaker models, the predicted summation response will approach measured reality. However, the limitation is found when we use different speaker models. Since each model will have unique phase responses over frequency the measured summation response will vary with the differences between the models.

Another limitation to the quasi-phase response is its simulation of the phase response as originating from a single source point. A measured phase response of a two-way speaker would reveal what we can easily confirm with our eyes, that the high and low frequencies come from different physical locations. This means that even the summation of two identical model speakers will not be accurate in the acoustical crossover range in the plane where the HF and LF drivers are displaced. This is the confluence of the spectral and spatial divider as discussed previously in Fig. 2.34.

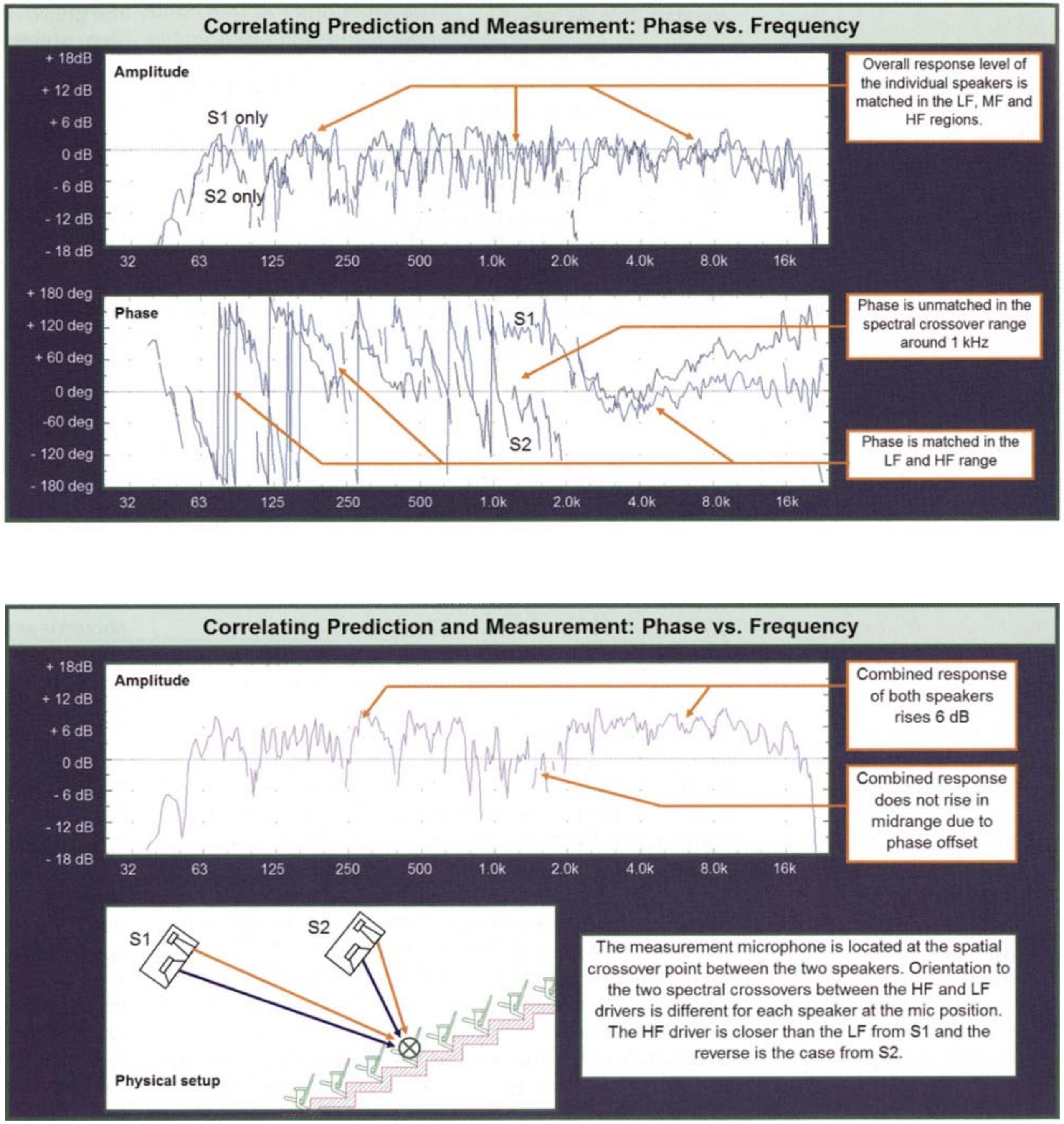


Figure 5.19 Effects of phase over frequency over angle in two-way speakers. The angular orientation to a source affects the spatial crossover orientation to the spectral crossover. Top and middle amplitude and phase responses over frequency of the individual speakers at the spatial crossover position. Bottom amplitude response of the combined speakers and the physical setup

Applications

Acoustic modeling is a design tool. Its role is to help us with design-related decisions. The utility of these programs relates directly to how accurately the differences in various options are shown so that informed decisions can be made. Finally we will see how well these decisions stand up under the scrutiny of the alignment process.

It will be great to have extremely high-resolution 3-D data with measured phase data. At the current time no available program offers this. In the mean time we have to work with what is available. Each of the programs has its strengths and weaknesses. If we can't have it all, how do we prioritize?

First and foremost, the end goal is a successful optimization, which will be guided by the data we view on our analyzer. Our chances of success are greatest if the data seen in the prediction stage reflects what we will find in the optimization stage. This correlation must be prioritized far above considerations of presentation quality to boardroom executives. The best choice is the prediction that yields the wisest selection of speaker model, position, array types, signal processing resources and acoustic treatment. Matters that can be adjusted on site, such as equalization, fine focus angle and relative level will be left to the calibration stage.

What can we conclude from the above features of acoustic modeling programs? First we must concede that the task is extremely complex. Layer upon layer of acoustical effects are added to the direct sound leaving our speaker. It is undeniable that the accuracy of predictive data goes down with every additional step away from our sound source. It is also certain that the addition of each and every surface and speaker further degrades the accuracy. The odds are stacked very hard against the room side of the equation. The acoustical mechanisms are far more complex and the data available for computation are far less complete.

The highest accuracy is found in the direct field response of a single speaker. If the speaker characterization is of sufficient detail, we can clearly see its nearby direct field coverage pattern, and level, in the space. Second place goes to more distant renderings of the same speaker, having factored in the direct sound propagation loss and perhaps the air-transmission-related frequency response losses. Third place goes to the direct field response of multiple speakers. The summation effects of multiple speakers are far less complex than that of rooms. Last place goes to speaker/room summation which we have little hope of finding accurate renderings beyond the ray tracing of the first few reflections.

Bearing in mind that our mission here is to design speaker systems, not symphony halls, the above order of accuracy is the best we can hope for. Our priorities are in line with the prediction system capabilities. We are far more concerned about the direct field response of the speakers precisely because those are the parameters most under our control during the alignment process. The choice of speaker model, the aim angle, the array type and the distribution of signal processing will all be based primarily on the direct field response. If the system cannot produce a minimum variance sound distribution in the free field it is folly to rely on the room to create it for us. The room will only degrade our uniformity. In those places where strong reflections will occur we will advocate for the reflective surfaces to be treated. But it would not be prudent to expect that the program will be capable of highly accurate characterizations as to how effective our measures will be. As discussed in the previous chapter, we will usually work toward minimum excitation of the room and maximum absorption. We seek maximum control of the direct field, and then move into damage control mode in the room. Anything unsolved in the direct field will likely remain so in the room.

For our application, the modeling of the room acoustics should not be considered on equal footing with the modeling of the direct sound of an array. The accuracy of the array rendering can be extremely high, while the room cannot be. The most effective solution to acoustic problems in the room can be seen clearly by accurate rendering of the array, even if the room is not modeled. The solution for the array will never be found if its own rendering is not accurate, no matter how much detail we have in the rendering of the room. Which comes first, the direct sound or the reflection? If the direct sound is not accurately characterized how can we expect that the room reflection characterization is meaningful?

There are no simple answers here, and many reasonable questions. Among them are these:

What is the problem with the inclusion of room reflections in our computations of uniformity? Isn't a modeling of questionable accuracy better than none at all?

Consider the following:

- If the resolution of the speaker data is poor, how can the reflection data be any better?
- If the speakers combine additively in the program without phase factored in, how are we going to implement this in a physical world in which phase is not optional?
- If the required coverage angle is 40 degrees for a seating area with curtains on the side wall, what is the required coverage angle if the curtains are removed?
- If we need a 132 dB SPL speaker for a seating area with curtains on the side wall, how many dB SPL would we need if the curtains are removed?
- If a coupled point source array were the best choice with the curtains, would it still be the best choice without?
- If there is a 10 X 20 meter wall of glass at the back of the hall, how accurate does the modeling need to be to let us know this will need acoustic treatment?
- In the above scenario would it not be most helpful to have pinpoint accuracy of the direct sound characterization, so that we can steer away from the glass?
- When we see prediction maps that show even SPL at every seat, doesn't it seem strange that we have never heard such a thing in real life?
- If the environment is expected to be hot and humid does this mean we should change the speaker model or focus angle?
- Does knowing that the room has a reverb time of 2.3 seconds tell me what kind of array to use, where to place it, focus angle or model?

Conclusions

The reader will draw their own conclusions as to which, if any, modeling program to use for their design work. As the author, I cannot even consider a means to move the discussion forward towards design and optimization without a phase response. There is nothing I could recommend that would have a better chance of success, nothing that could indicate the viability of one approach over another. I say this because my perspective always includes the way the designed system will look under the microscope of the optimization process. That viewpoint is always high resolution and always includes phase. I have never measured a system that in any way behaved like a phase-free rendering. I have, however, seen systems that behaved very precisely like that of the high-resolution, phase-inclusive renderings found in MAPP Online.

Consider this a confession of author bias or a limitation of my vision if you will. But suffice to say I lack the ability to visualize sound in the manner that is shown on low-resolution phase-free data. Until such time as high-resolution prediction moves into 3-D (and I learn to operate it) I will struggle with the limitations of 2-D vision over the errors and omissions of low-resolution phase-free data.

We are now ready to move on to the application of the prediction process to our goals of the optimized design. In the next chapter we will conduct a thorough search for speaker system designs that will provide the maximum uniformity of sonic experience over the listening area. The phase-inclusive, high-resolution prediction model will guide us toward this goal.

A central theme as we continue here will be the fact that every element of our speaker system will play an individual role in the combined response of the system. That individual identity will always be maintained in the physical world. The optimization process will deal with the interactions of elements on a case-by-case basis. If the prediction program does not have sufficient resolution for us to identify the individual contributions we will not be able to anticipate the response we will find in the room with our analyzer. Low resolution and absentee phase both cause an identity crisis to combined response renderings. Two elements go in and appear to be transformed into a unified whole. A microphone in the local area will reveal an entirely different story.

The prediction renderings in this book are exclusively from MAPP Online. The astute reader will realize that this means that all of the prediction characterizations are of Meyer Sound speakers. Those who wish to exploit this as a means to promote, denigrate, or emulate Meyer speakers will likely be disappointed. The principles of physics that these speakers operate under are impervious to marketing departments. The design principles that underlie the approaches advocated here are universalized to be inclusive of speakers that fall into the classification system (first to third order) described in Chapter 2, and the power classifications found later in Chapter 7. The precise extent to which systems behave as described here is dependent primarily on how closely they conform to the classification specifications. Each and every make and model of speaker will have unique characteristics, and never in this text will reference be made to the particulars of any one. The models of other manufacturers can be viewed in the context of where they fall in the classifications which can be found by comparing their published specifications. The classification of many models will be easily seen while others may fall in the gray areas between. In every instance, the design and optimization principals still apply. In any case, there is nothing promoted here — positive or negative — that the reader should view as an exclusive attribute of Meyer speakers. The role of advocacy for makes of models of speakers is respectfully left to the manufacturers.

Variation

> ***variation*** *n.* ***1.*** *varying, departure from a former or normal condition or action or amount or from a standard or type, extent of this*
>
> ***variance*** *n.* ***1.*** *disagreement, difference of opinion, dispute, lack of harmony*
>
> Concise Oxford Dictionary

Introduction

In the beginning there was silence. Then there were speakers: single speakers, column speakers, ceiling speakers. Then there were clusters of horns and a few woofers. Then speakers got legs and learned to walk. Horns and woofers and subwoofers were each placed in boxes and stacked on stage. Then speakers got wings and learned to fly. They have been up in the air ever since. In the course of speaker evolution there has been an adaptation in favor of two particular traits in the audio gene pool: increased power and decreased variation. The natural selection of the marketplace has seen to it that those speakers which are dominant in one of these traits survive and continue to adapt. The others have become extinct. If some single speaker species were to attain dominance in both categories it would undoubtedly assume its position at the top. These two dominant traits have often been rivals and for the most part, the brawn of the power has been selected in favor of the brains of minimum variance. But this begs the question: can we find a way to select for power without sacrificing uniformity, and vice versa. Is there a middle ground?

If the question is limited to a single speaker element, the answer is yes. We can substitute an alternative speaker with similar parameters, yet higher power and not increase the variance. If we are to gain the power by using multiple speakers the answer is a qualified no. Power addition by multiple elements will increase variance in all but a very small fraction of circumstances. In the vast majority of cases they are mutually exclusive, but fortunately, there is a middle ground. The purpose of this section is to show the roadmap to the middle ground. Once we know how to get there, we can evaluate the tradeoffs involved in those instances where we must move off the middle ground in favor of one side or the other.

Given multiple speakers we can select for maximum power or minimum variance by the manner in which they are arrayed. To select for maximum power we run every element at the same level and aim them all at a single location, for example, the mix position. Each element can be selectively delayed to arrive perfectly in phase for the maximum single point summation. The maximum concentration of power into this single point ensures the maximum variation from all others. The reverse occurs if we array the speakers for the minimum overlap between the elements, and adjust the relative levels to correspond to the shape of the room. In this case the sound is spread evenly throughout the space but a minimum amount of power summation is achieved. The middle ground is found where we can harness the power of summation constructively, with a minimal compromise in variance. Power and uniformity can be brought together on a limited basis. Massive amounts of power can be harnessed and wide areas covered, but this can only be achieved by careful coordination between the parties. That is what this chapter is all about.

Now back to our Neanderthal past. In the olden days there were two exclusive audio worlds: installation world and touring world. Installation world was primarily concerned with uniformity and had the attention of the manufacturers on their quest. The standard installation featured horns arrayed as a point source and custom fit to the shape of the intended coverage area. The systems were successsful by the standards of the day but lacked the power capability required by the emerging touring industry. Touring folks wanted power, power and more power. The touring people took the available tools and built piles upon piles of horns and woofers into ever larger stacks on stage. The logistics of piling up speakers in the time allotted forced the innovation of the "all-in-one" box, the most well-known example of this being the Clair Brothers' S-4. This one size fits all approach allowed touring system sound design to become a simple matter of scale. A small venue was an eight-per-side venue, while a large venue was fifty per side, and so on. Alternately, the same venues could be done with proportionally larger numbers of smaller boxes or a smaller number of larger boxes. This was the era of the "wall of sound." The speakers were primarily wide coverage units with 100 per cent coverage pattern overlap. Maximum addition came at a cost: maximum variance. Usually a small portion of the speakers were angled slightly outward as a courtesy to the side seating area, creating a horizontal version of the letter "J." When flying hardware became sufficiently practical, these arrays became airborne and the listeners in the first rows were given some relief from the full frontal blast, a noticeable improvement in level uniformity. The vertical alignment could then be modified to also take on the "J" shape of the horizontal arrangement.

Enter the trapezoid. In the early 1980s a fundamental evolution occurred in favor of uniformity: the trapezoidal-shaped speaker cabinet, made popular by Meyer Sound. The cabinet was narrowed at the rear which allowed for speakers to be arrayed at an angle, and yet keep the driver displacements to a minimum. This resembled the point source horn array of the installation market, but was in the all-in-one portable speaker format usable for touring. The touring industry now had point source tools to provide greater uniformity and the installation market had tools which could provide serious power. This was an important step in the progression toward shared installation and touring products that persists to the present.

Nonetheless the wall of sound was far from finished. The trapezoidal shape alone does not create a point source array; it merely provides a physical aid. Widespread overlapping of the coverage patterns continued as standard practice in applications requiring extreme power. While some touring systems employed minimum overlap point source techniques and greatly enhanced uniformity, others configured the speakers in the familiar J shape and continued with business as usual. The most common version was a compromise which used the point source in the horizontal plane and the J in the vertical plane.

Meanwhile in the medium-sized venues of the musical theater world, an alternate approach emerged. Variance was reduced to a minimum by strategic separation of the point source approach into its uncoupled form. The systems were a complex weave of interconnected sources with small dedicated zones of coverage. The needs for both uniformity and power were satisfied by keeping the speakers close to the separate audience zones and selectively setting their level and delay as required.

The touring industry moved toward its current direction with the introduction of the V-Dosc line array system by L'Acoustic in the 1990s. This system had extremely narrow high-frequency vertical coverage and extremely wide full-range horizontal coverage. With this innovation, one single speaker could handle a majority of the horizontal coverage needs, removing the comb filtering that occurred when the wall of sound or overlapping point sources were used. The narrow slicing of the vertical plane allowed for a pinpoint beam of sound to be thrown to the back of the hall, or far beyond into the next county The key innovation here was the conformance of the speaker element's properties of symmetry to the listening area shape. Audiences are usually spread in wide symmetric horizontal shapes, and narrow asymmetric vertical shapes. This type of speaker element was adapted in favor of the dominant audience shape.

The wall of sound was officially dead and has now been replaced by the "line of sound." The scalar factor persists, however. A small venue is an 8 box/side venue and a big venue is a 50 box/side venue. Still, the same venues can be done with a proportionally larger number of smaller boxes or a smaller number of larger boxes. The line of sound has vastly superior horizontal coverage uniformity over the wall of sound. The horizontal plane is now literally "seamless" in that there are no acoustical crossovers to reduce uniformity. The degree to which this compares to the optimized point source will depend on how well the room shape conforms to the fixed horizontal shape of the element.

Unfortunately, there are two key strategies from the wall of sound that persist in the common implementations of the line of sound. The first is the belief that significant level reductions of any element in the array will have negative effects upon the overall performance of the array. There is truth to the assertion that reducing the level of one element in the array will reduce the overall level of the array. But what use is a more powerful array if its power is concentrated in the wrong place? If we are to create a uniform response in the room, this cherished belief of the SPL Preservation Society will need to be abandoned. After all, the paying public does not care about how loud it is at the mix position. They care about *where they are sitting.* It is only through the spreading of uniform level, not concentration of level, that the public gets the benefit of our modern speaker arrays.

The standard current-day line of sound is implemented in three basic shapes: the vertical line, the arc (point source) and the hybrid of the two forms, the "J." Each of these will be examined in detail in this chapter and its performance evaluated in terms of uniformity and power. In the end we will see that comparable performance can be obtained from "line array" speakers and "point source array" speakers when the minimum variance strategies are employed. We will see that there are no such things as "line array" speakers, and that it is the configuration of an array that gives it its name, not the nature of the individual elements. We will see that both the modern so-called "line array" speakers and the old-fashioned "point source" speakers are equally evolved, and equally equipped to maintain survival. They will thrive in different environments and we are all better off for having more options available to us. Our job will be to ensure that they are put in the right places and tuned in the correct manner.

Our first mission will be to find the road to minimum variance. This is much more complex than the road to power, which is principally a matter of scale. Minimum variation can only be achieved with a constant focus on the spatial geometry of speaker coverage. Rooms come in an unlimited variety of shapes. Speaker coverage patterns come in a substantial but limited number of shapes. The shape that we seek is the rarest form: that which can remain substantially constant over frequency. Rooms do not change their size or shape over frequency. If we are to fill the room evenly with full-range sound we will need to keep the coverage shape as constant as we can in spite of the 600:1 range in the size of the wavelengths we are shaping. This is an impossible task, but the quest continues to reach the closest practical approximation with the current technology. My twenty-five-year search has yielded precisely three shapes that can be reliably created and provide satisfactory results over a substantial space. It is not a large number, but it is enough. It is possible to partition the coverage responsibilities of any room into these shapes, which can then be patched together with spatial acoustic crossovers. These shapes are entirely scalable, making them equally as applicable to stadiums or home theaters. They are manufacturer-independent, to the degree that the criteria of the individual element speaker coverage shape are met. These criteria will be described in the conclusion of this chapter.

The Minimum Variance Principles

Variation Defined

Our pursuit is a single experience for all members of the audience. This is a spatial pursuit. Everything will be framed in terms of the distribution of the sound over the space. Since perfection in this regard is impossible we need to define clear terms that will indicate how closely we approach this unobtainable, yet ultimately desirable goal. The term we will use is variation, or variance, which we will define in terms that relate to our previous discussions. The variation that we seek to minimize herein has three principal components: the spatial distribution of level, the spatial distribution of the frequency response, and the spatial distribution and amount of frequency response ripple. These are the major players in the game, with other related aspects that follow on their coat tails. A final category, variation in sonic image placement, will be handled in the next chapter.

- Level variance — variation in sound level. Differences in the overall level over the space. This is evaluated as dB, e.g. it is 6 dB down under the balcony, compared to the 24th row.
- Spectral variance — the difference in relative level over frequency between locations, e.g. it is 6 dB down in the 8 kHz range under the balcony, compared to the flat response in the 24th row.
- Ripple variance — differences in the extent of the summation related peaks and dips in the response. This is characterized by the summation zones described in Chapter 2, e.g. there is ±12 dB of ripple in the mid-range under the balcony, compared to ±3 dB in the 24th row.

There are other related categories that follow the main definitions above. Intelligibility loss is most closely related to the amount of ripple variance, but high-frequency roll-off (spectral variance) and even overall level loss (level variance) would also cause degradation. Therefore, we can conclude that the minimum variance strategies for the principal categories have a high probability of yielding high intelligibility. Another related category would be dynamic range: differences in the maximum level capability or the noise floor over location. In order for level variance to remain minimized while the system is in operation, the maximum capabilities of the speaker elements must be matched in their respective areas of coverage, e.g. the under balcony seating area has a matched maximum level to the main floor. To achieve this we must ensure that the speaker models assigned to cover different areas will have maximum capabilities that are scaled appropriately for their range of operation. The under balcony speaker will be a lower-power system than the mains, but it will be capable of matching the mains in the confined range of coverage where they meet. The scaling of speakers to their range of coverage will be discussed in the next chapter, and therefore we will assume that the scaling is appropriate for the duration of this chapter. We will also make another assumption here: that all speakers have a flat spectral response on-axis. Therefore, our discussions of variation over position or frequency all originate from the same source.

Our main focus will be on the three primary factors, leaving the related factors to follow as noted. Both the causes and effects of these mechanisms are complex and highly interrelated, and will not easily yield their secrets. The rewards are worth the effort, though. When our coverage shape matches our seating area, matched in relative level and spectral balance with minimum amounts of ripple, we will have achieved our goal of minimum variance.

Causes and Standard Progressions of Variance

Since we will be examining variance in three categories, we will discuss their respective sources individually. For example, the causes of level variance are distinct from those that cause ripple. Each form of variance has standard progressions of effect. Knowledge of these progressions will allow us to anticipate them in our design and to identify them in the field during optimization. Such knowledge will also be the key to reducing variance by playing one progression against the other so that their combined effects offset the individual ones.

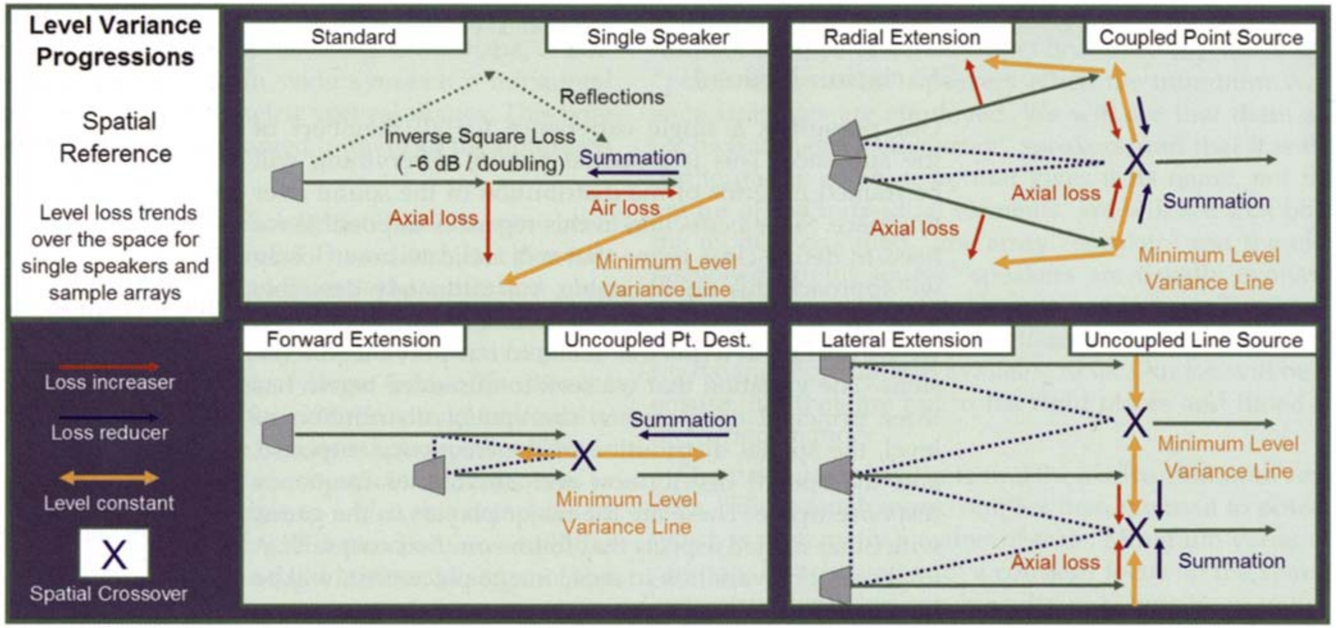


Figure 6.1 The level variance spatial progressions

Level Variance

The standard level variance progressions for single speakers and some representative arrays are shown in Fig. 6.1 .

Single Speaker

Level variance is a simple matter of proximity and axial orientation. If we move further away the level falls. There is no cost-free way to get around the inverse square law and the 6 dB loss of direct sound per doubling of distance. The loss rate will decrease when the room is added to the equation, but such summation-based changes will add spectral and ripple variance. The overall level will also drop as we move off-axis, but the loss rate will not be uniform over frequency. Off-axis response will be a combination of level and spectral variance. The primary option for management of level variance in a single speaker is tilt. The focus angle is adjusted to aim the speaker at the most distant coverage point. This offsets the two level variance factors, distance and axial. The level at distant seats will have more of the propagation distance loss, while the closer seats will have more of the axial loss. Tilt, in a variety of forms, will be a key factor in variance management.

Multiple Speakers

There are three directions for extending the minimum level variance area beyond the limits of a single speaker element: forward, radial and lateral.

Forward extension is accomplished by adding speakers in the on-axis plane of the original source. These are delayed, of course, so that the forward radiations are synchronized. The level of the forward speaker can be set to join the mains and create a combined level that is matched to an area between the sources, thereby temporarily stemming the loss of level over distance. The decisive forces are offsetting amounts of distance loss and summation addition. Forward extension accrues incrementally as devices are added and is limited only by our budgets.

The second avenue is radial extension. The additional sources are aimed to intersect at the coverage edges and fused together to create a wider angle. The decisive forces are offsetting amounts of axial loss and summation addition. The upper limit of radial extension, for obvious reasons, is 360 degrees.

The third direction is lateral. Uncoupled secondary sources are added to the side of the original and a line of equal level is created that links the sources. The equal level line of the combined speakers is found at the distance of the unity class crossover point, i.e. where the —6 dB points in the coverage pattern edges meet. The decisive forces are again offsetting amounts of axial loss and summation addition. Lateral extension accrues incrementally as devices are added and again, is limited only by our budgets.

The three extension mechanisms can be employed separately or in combination. Delay speakers under a rounded balcony are an example of forward, radial and lateral extensions working together, in the combination of the mains and the uncoupled point source array. The intersection points are found ahead of the elements, their position being a function of forward distance, lateral spacing and splay angle.

Spectral Variance

Everything we just discussed regarding level variance applies in this case. The difference is that we must apply it separately for each frequency range. No problem. Just add 600 layers of complexity. No speaker holds the same coverage shape over frequency so we will have our work cut out for us. We can look at this as changes in coverage pattern shape over frequency (spatial over the spectrum), or as changes in frequency response over the coverage area (spectral over the space). We will use the term spectral variance, since all of our variance forms can be considered as "over the space."

For a single speaker this value is directly related to the speaker order. High-order speakers have the highest spectral variance, since their coverage shape is the most variable over frequency. They are, not coincidentally, the least well suited for single-speaker applications.

Spectral variation for a single full-range speaker is a fixed parameter. Equalization, level or delay will not change the coverage pattern over frequency. Arrays, however, can be constructed with combined coverage shapes that differ from the individual components. A coupled point source array comprised of high-order speakers can create a minimum spectral variance combined shape by the spreading of the isolated high frequencies and the narrowing of the overlapping low frequencies. A very simplified example: a pair of second-order speakers are arrayed at their unity splay angle of 40 degrees. The high-frequency coverage is spread to 80 degrees. The low-frequency coverage angle for the individual speakers was far wider than the 40 degree splay angle. The resulting overlap couples at the center and narrows the low-frequency coverage. The combined shape of the array has lower spectral variance than the individual elements.

Spectral Tilt

The example cited above allowed us to reduce the level variance over the space. At the same time we were also changing the spectral response. How? When the array is constructed the high frequencies are isolated by the unity splay angle, but the low frequencies overlap. The result is summation addition in the lows that is not matched by a comparable addition in the highs. The spectral response tilts upward in the low frequencies but is unchanged in the highs. This is **spectral tilt.**

Once again we are talking about tilt. The previous discussion mentioned the role of tilt, i.e. the aiming of the speaker, in the management of level variance. This is spatial tilt, and compensates for proximity. The second form of tilt in our systems is spectral tilt, i.e. the overall shape of the frequency response.

For example: the difference between on- and off-axis response of a single speaker is not simply level. In fact only the highs have been significantly reduced by the axial loss. The primary difference between the two locations is the amount of spectral tilt. *The difference in spectral tilt between the two locations is the spectral variance.* Let's distill this down one more time to make the distinction clear.

Comparing spectral tilt and spectral variation:

- Two points with matched flat responses: no spectral tilt and no spectral variance.

• Two points with matched HF rolloffs: spectral tilt but no spectral variance.
• Two points with matched LF boosts: spectral tilt but no spectral variance.
• One point has HF rolloff: (spectral tilt), the other is flat (no tilt): spectral variance.
• One point has LF boost (spectral tilt), and one is flat (no tilt): spectral variance.

Our goal is to minimize spectral variance, not necessarily spectral tilt. In fact many clients prefer fairly large amounts of spectral tilt. Our goal is to ensure that the desired tilt is found in all areas of the listening space, i.e. minimum spectral variance.

There are two transmission-related effects that cause spectral tilt: axial loss and air loss. As we move off-center we engage the axial filters and the high-frequency range begins its descent. A secondary high-frequency filtering effect is in the transmission air loss. The result is that even the on-axis response cannot remain flat over distance. In contrast to the spectral variance of a single speaker, spectral tilt is not a fixed entity. It can be modified with equalization.

There are also two summation-related effects that cause this spectral tilt: reflections and the previously mentioned combination with other speakers. When we add the room reflections to the equation we bring in summation effects which will decrease the loss rate over distance at some frequencies, principally the lows. Once again we will see spectral tilt in favor of the lows.

Some representative samples of summation-based mechanisms are shown in Fig. 6.2, where trend lines of the spectral tilting effects are compared. There is a consistent trend in the progressions in favor of more low-frequency content over the highs. It is this common trait that will be exploited to create a consistent spectral tilt. Since all of the progressions lead to tilting, the spectral variance is reduced by matching the tilts rather than by futile efforts to stop the progression. In the end the tilt can be leveled by equalization to the extent desired. A frequency response does not need to be flat to be considered a minimum spectral variance.

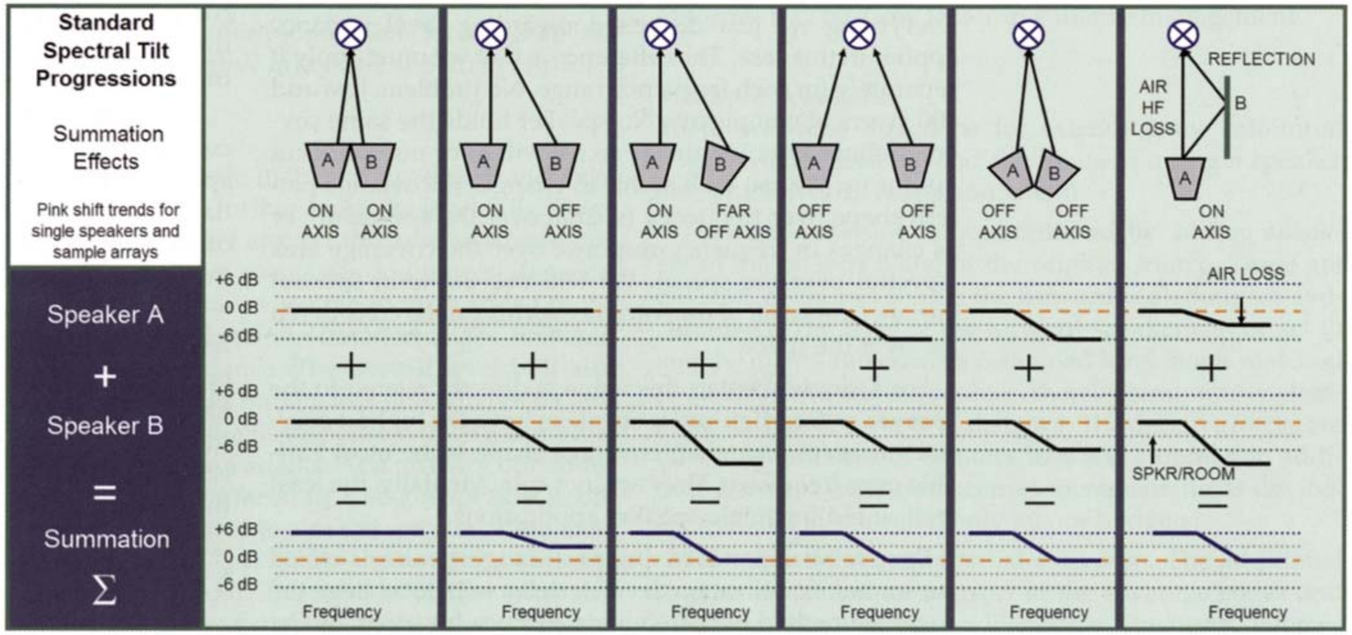


Figure 6.2 Summation-related spectral variance progressions

The decisive factor is consistency. Responses that are tilted in a similar manner are just as matched as flat ones.

Introducing "Pink Shift"

In the early days of acoustics much of the research was done by scientists who shared interests in astronomy. We will divert for a moment to utilize an astronomical concept as an analogous feature in our variance discussion. The concept is "red shift," which is used by astronomers to measure the distance between remote objects in the universe. The red shift is a measure of the spectral change due to the Doppler effect, which changes the perceived frequency of light between moving sources. The distance to the object is proportional to the amount of red shift. Our tie-in to acoustics is not related to the acoustical version of the Doppler effect, as we will not be covering the challenges of doing concerts in which the speakers are mounted on moving trains. The concept is linked to perceptions of distance to the sound source. We are all familiar with white and "pink" noise. Pink noise is filtered white noise (equal energy per frequency) with a steady 3 dB reduction per octave. This creates equal energy per octave, and balances the noise spectrum for our logarithmic hearing. The frequency response tilting related to air transmission loss and the frequency response progressions just described could be viewed as "pink shift" added to the response. As the spectral tilt increases the "pink shift" rises. In natural acoustic transmission this pink shift is directly related to sonic source distance. The farther we are from a source the greater the degree of tilting due to HF air loss and LF summation in the room. Our internal sonar system estimates source distance by factoring in our expectations of "pink shift." One of the ways the ear detects the presence of loudspeakers is by false perspective (discussed previously in Chapter 3) which occurs when we extend the HF response to such an extent that the sonic perspective is too close to be credible. The "pink shift" does not match our visual expectations. Another false perspective occurs when we are close to an off-axis speaker. The situation is reversed, with the pink shift being greater than we would expect from a close source.

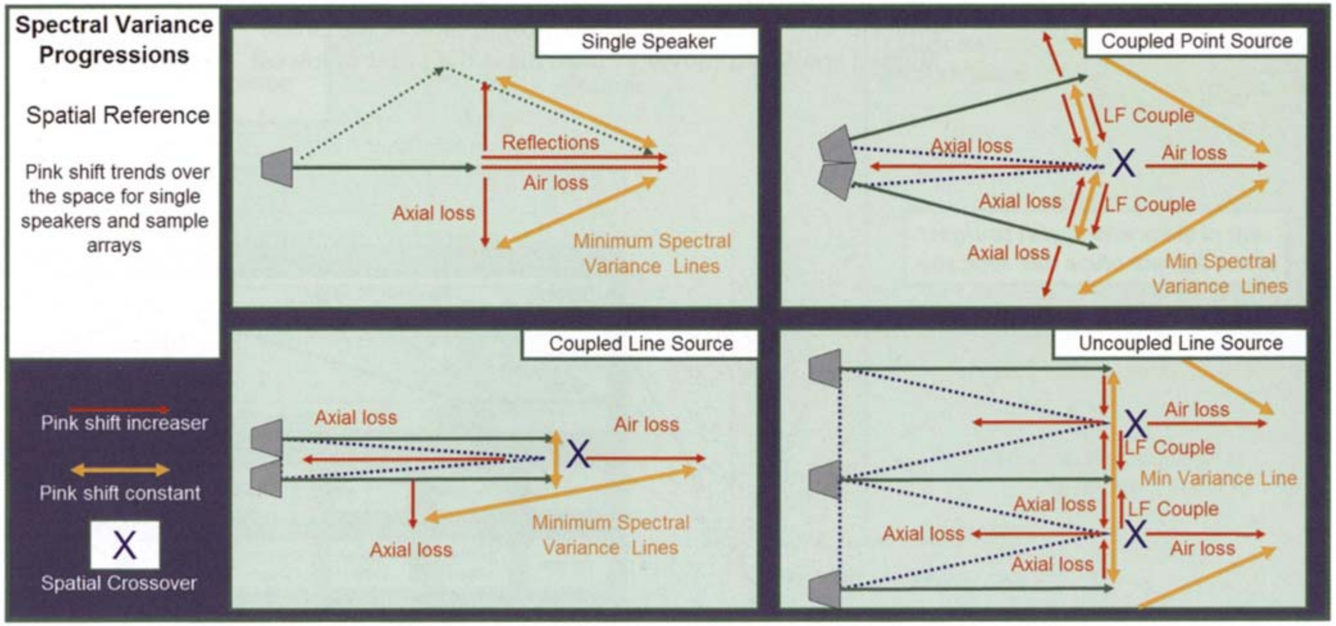


Figure 6.3 Spectral variance spatial progressions

Perspectives: When we tune a sound system, perhaps we should try to avoid over-equalization as much as possible during the process of dealing with the MF and LF ranges. Meanwhile, could a gentle downward curve of frequency (amplitude response) be a recommendable measure to take? The issue of downward curve has some uncertain factors, though. The degree of this curve depends on the distance among SP system and measurement microphones, how the extent of the area to cover stands (including reverberation, shape of a venue and the setting conditions of the speaker system) and so on.

Akira Masu

Summation effects are also a prime source of pink shift in the system. Speaker/speaker summation almost always shifts the response toward pink because low frequencies have more overlap than highs. Speaker/room summation will also move things in this direction because absorption is usually more effective as frequency rises.

The convenience of expressing the spectral tilt in this way is that we now have a term for the same response shape regardless of whether it resulted from LF summation, HF loss or both. The major trends in the distribution of pink shift over the space are shown in Fig. 6.3. As in the level variance scenario we can expand upon the area covered by a single speaker with forward, radial and lateral extensions.

Ripple Variance

The sources of frequency response ripple were explored in depth in Chapter 2. The summation zones are the markers on our ripple variance map. The standard ripple variance progression is shown in Fig. 6.4. The cycle is extendible and repeatable for multiple elements but follows a familiar pattern. The center point in the progression is the phase-aligned spatial crossover point. This is the point of lowest variance in the progression, and yet it is the center of the area of the highest rate of change in ripple variance. The coupling point is analogous to the point of lowest wind speed in a hurricane: the eye of the storm. The area just outside the eye, however, contains the highest variations in wind speed. That is the nature of this spatial crossover area, and we hope to calm the storm outside of this area by isolating the systems as quickly as possible.

The ripple variance progression can be found in all forms of speaker arrays. Any two speakers in a room will eventually meet somewhere, at least in the low frequencies, regardless of their spacing, relative level or angular orientation. Our focus is on full-range spatial crossover transitions so we will limit our discussion to those that fall within the coverage edges of the elements. The transitional points will fall in different areas for each array configuration and therein we find the key to managing ripple variance. Some array configurations confine the variance to a small percentage of their coverage. Others fall into the combing zone and never climb out again.

The primary indicators are source displacement and overlap. The least variance occurs when both of these are low. The highest variance occurs when they are both high.

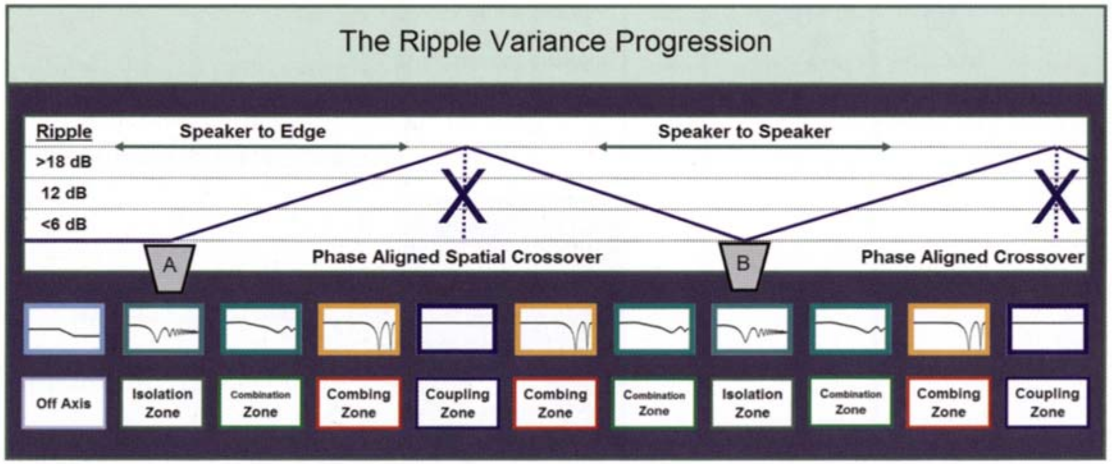


Figure 6.4 The ripple variance progression

The ripple variance progression:

- Frequency response ripple: as the spatial crossover is approached, the ripple variance increases. As overlap increases the ripple variance increases.
- From the spatial crossover point: coupling zone to combing zone to combining zone to isolation zone.
- From the isolation zone: to coverage edge, or to combining to combing to coupling to the next crossover point.

This cycle repeats for each speaker transition until the coverage edge is reached. Each location where a spatial crossover is found will need to be phase-aligned. Phase-aligned spatial crossovers cannot eliminate ripple variance over the space. They are simply the best means to contain it.

The progression rate is not equal over frequency. The first to enter (and the first to leave) each of the zonal progressions are the high frequencies. The finalists in both cases are the lows. It is possible, and quite common, to run the high frequencies from the spatial crossover to isolation before the lows have ever left the coupling zone. The coupled point source array is the classic example of this. Therefore, our roadmap for the variance progression will need to factor in frequency, as the transitional milestones are not evenly spaced.

Ripple Variance Geometry

The concept of triangulation as a means to visualize summation over the space was introduced in Chapter 2. Since ripple summation is one of the primary forms of variance, this concept will be reintroduced and applied directly to visualizing variance trends. Refer to Fig. 6.5. The four triangle types give strong indication of the spatial behavior of ripple variance. The isosceles triangle, the representative of the summation coupling zone, is the only path that can obtain zero ripple variance. Forward movement along the isosceles triangle center will be time-coincident and therefore free of ripple. This happy scenario is also an inescapable trap. Any position off the center line plunges us into the regions of highest variance, since we fall out of time but lack isolation. This is the summation combing zone and is represented by the right and acute triangles. The

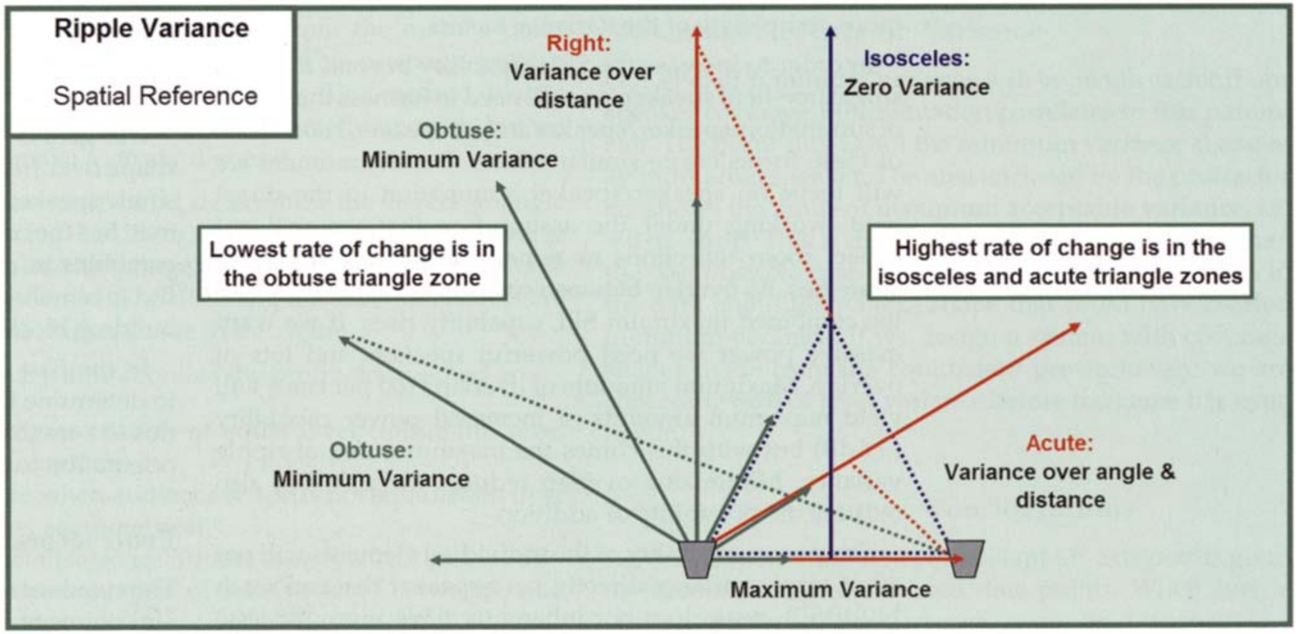


Figure 6.5 Ripple variance spatial reference

acute triangle area is enclosed in the isosceles perimeter and has the highest variant rate possible with changes in orientation (distance or angle) to the sources. The right triangle area is found outside the isosceles perimeter but will exhibit variance with changes in distance from the sources.

The final triangle type, the obtuse, is our representative of combining zone and isolation zone summation. The obtuse triangle opens the door to minimum variance. We will still have to walk through it by providing angular isolation between the elements so that on-axis orientation to one element corresponds to off-axis orientation of others. The obtuse angle direction is not free from variance but does have the best prospects for the minimum variance over an extended area.

Minimum Variance vs. Maximum SPL

The introduction for this chapter set the stage for discussion of the compromise between minimum variance and maximum SPL capability. The SPL side of the equation is a simple one to represent, especially compared to the enormous complexity of the variance factors.

In order to increase the SPL capability beyond that of a single free-field speaker, we will need to harness the power of summation: speaker/speaker and/or speaker/room. Both of these forms create similar effects. For the moment we will focus on speaker/speaker summation in the direct field, working under the assumption that we will not expect room reflections to remedy flaws in our design strategies. As overlap between speaker elements increases the combined maximum SPL capability rises. If we want massive power we need powerful speakers and lots of overlap. Maximum amounts of overlap (100 per cent) will yield maximum amounts of increased power capability (+6 dB) but with this comes the maximum risk of ripple variance. Minimizing overlap reduces the risk, but also reduces the possibility of addition.

The power capability of the individual elements will not affect ripple variance directly, i.e. a speaker that can reach 140 dB SPL peak does not inherently have more (or less) ripple than one with 6 dB less power capability. Assuming a properly scaled system (a subject which will be covered in the next chapter) we can continue without concern about this at the moment.

If we are going to gamble with overlap we will need to know how much ripple variance we are risking in exchange for the potential power gain. The reader is advised that for brevity this chapter will not be providing ongoing commentary about the amount of power that is gained each time overlap occurs. When we see overlap we will know that power is gained, and we will see the price. When we see isolation we will know that power is neither gained nor lost.

With that said we are now ready to move forward to the details of searching for minimum variance.

The Speaker/Room Link: Aspect Ratio

Rooms come in shapes. Speaker coverage comes in shapes. How can we relate them? This is not a simple matter. Room shapes have unlimited variety, but are constant over frequency. Speaker coverage shapes have limited variety but are highly variable over frequency.

Our goal is to fill the room shape with an even coverage shape over frequency. There are two ways to do this: use a single speaker with constant coverage over frequency that matches the room shape, or build an array of speakers that combines to create constant coverage over frequency in the given shape. In either case the single speaker is the building block and we will begin from there.

In the first chapter we discussed the various methods to determine the coverage pattern. It is now time to apply this to a method for selecting the appropriate speaker and orientation to create a minimum variance response.

Protractors and Pizza

The standard method of determining coverage pattern and aim points of a single speaker is the "protractor method,"

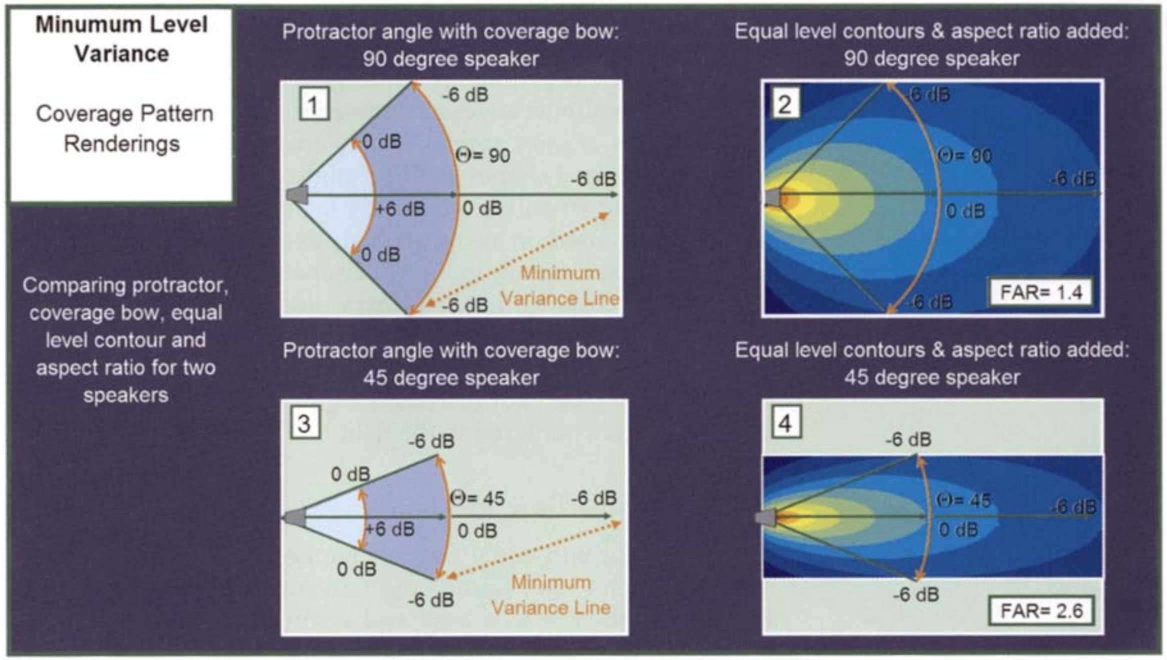


Figure 6.6 The relationship of different coverage pattern renderings and minimum level variance. (1) Coverage angle determination by the protractor method for a 90 degree speaker. The on axis point is found at a given distance and an arc traced until the level has dropped by 6dB. The coverage angle is the angle between the two edges. Doubling distance loss shown as the dark shaded area extends the coverage outward. The relationship between the on- and off-axis points is maintained but all points drop by 6 dB. The arc line traces the line of maximum acceptable variance. The "coverage bow" is also shown here. Double distance on-axis is equal in level to single distance off-axis. A line between these two points would trace the line of minimum level variance. (2) The equal level contour shows the lines of zero variance. The aspect ratio is shown as a rectangle with a length vs. width ratio of 1.4 (90 degree coverage angle). (3) Protractor and coverage bow renderings of a 45 degree speaker. (4) Equal level contours and aspect ratio for a 45 degree speaker

which traces an equidistant arc from the on-axis point until it reaches the -6 dB point. The coverage pattern shape comes out like a slice of pizza. This method, like its analogous food, is fine if we need something quick, but we should not make a steady diet of it.

The protractor method determines the coverage angle as shown in Fig. 6.6. If we add the workings of the inverse square law into the situation we can see that the shape continues outward. The doubling distance loss effects are shown as the darker shade in the figure.

The primary limitations of the protractor method are:

- lack of representation of equal level (minimum level variance) contours
- inaccurate when audience shape is not equidistant over the arc, i.e. asymmetrical
- lack of detail as to rolloff rate between O dB and —6 dB
- lack of representation of relative level between array components.

Maximum Acceptable Variance

Our goal of minimum variance will be made easier if our speaker coverage representation correlates to this parameter. The pizza slice is not the minimum variance shape of an individual speaker. The area enclosed by the protractor arc shows the range of **maximum acceptable variance,** i.e. +0 to -6 dB. This is the worst acceptable case, rather than the best possible case. The protractor method leaves us in the dark about the pattern shape that could have created minimum variance. If we design a system with coverage shapes made from the equidistant protractor arc we are conceding 6 dB of level variance before the game has even begun.

Asymmetric Coverage Considerations

An on-axis point and two equidistant off-axis points gives us three meaningfully related data points. When such a shape is presented for coverage, it can be described as a symmetrical coverage shape. This is often applicable for horizontal coverage shapes where the audience is spread evenly over an arc, square or rectangle. Vertical coverage, by contrast, is almost always asymmetrical, i.e. the on- and off-axis areas are not spread evenly from center. Wildly inaccurate vertical shapes and aim angles result from conclusions made when the inherently symmetrical protractor method is used on asymmetrical coverage shapes. For a single speaker, minimum level variance results from the offsetting of distance loss and axial loss. Moving closer to a source while simultaneously moving off-axis creates a stalemate of gain vs. loss. We must be able to see these two factors at play. Failure to account for this is a common design error encountered in the field.

The Coverage Bow Method

The protractor and equal contour methods are related in a shape which I have termed the "coverage bow," named for its resemblance to a bow and arrow. The speaker is located where the hand grips the string. The bow is our familiar protractor arc. The tip of the arrow represents a distant on-axis point, precisely twice as far as the bow. We will call this —6 dB due to the doubling distance loss. The hand that grips the bow is the on-axis reference. This would be O dB. The tips of the bow represent the off-axis level, at the standard distance point. These positions are 6 dB down, the product of axial loss. If we draw a curve between the arrow tip (—6 dB) and the bow tips (—6 dB) we have traced the shape of minimum level variance. The coverage bow is scalable. Each time the size is doubled the level drops 6 dB, but the proportions remain. As the coverage narrows the length of the arrow increases and the tension on the bow bends the arms back, reducing the width. This is an apt analogy since increased tension extends the range of the bow, making it a "long throw system." Increased directionality in our speaker will have the same effect.

The coverage bow links the locations most closely related to level and spectral variance: far on-axis and near off-axis. As we walk along the contour line between these two points we will be matched in both level and frequency response. Obviously this is where we want our audience to be.

The coverage bow is limited in that the line between our known points is an interpolated estimate rather than being based on more precise data: actual measurements along the arc. The bow can be combined with the equal level contours, giving us both an angular and equilevel rendering of the response over the space.

The coverage bow method is superior to the protractor for analyzing the coverage requirements for asymmetrical areas since it factors in the relationship of distance and axial losses. The equal level contours take the process further. Both are shown in Fig. 6.7 where they are applied to an asymmetrical coverage area. The coverage bow and equal level contours trace along the desired coverage shape, and take into account the relative distance. The result is minimum level variance across the listening area.

The Forward Aspect Ratio

Once we have obtained the equal level contours we can simplify the characterization of the coverage pattern of a speaker into its proportions of length and width. This is termed the aspect ratio, an expression that is commonly used by architects to characterize the shape of a room. Since the choice of optimal speaker coverage pattern is related to the shape of the room, the aspect ratio is a logical choice for this application. The aspect ratio is found by creating a rectangle that encloses four points: the speaker, the on-axis point at a given distance and the off-axis point at half that distance. The aspect ratio indicates the dimensional outlines of the shape of the minimum variance area for a given speaker.

For our purposes we consider speakers to be projecting forward, therefore, we will focus on the area in front of the speaker, reworking our term into the **forward aspect ratio** (FAR). This is defined as a single level contour's length from the speaker forward compared to its width. The street term for this is "throw." Long throw speakers get to the back of the hall, while short throw speakers play to the front. How is this quantified? Where is the break point? How far does a speaker throw anyway? Infinity actually, unless

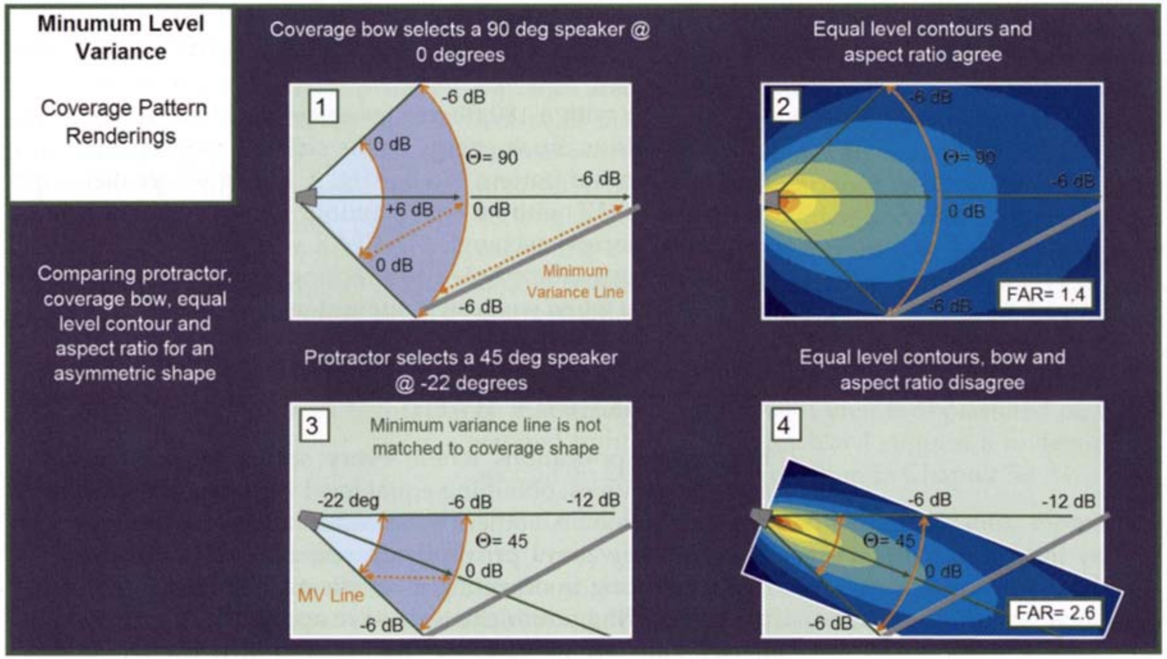


Figure 6.7 The relationship of different coverage pattern renderings and minimum level variance in asymmetric applications. The listening area (gray line) is found approaching the off-axis edge of the speaker. The protractor method (lower panels) would advise us to steer the speaker downward and reduce the coverage angle. Note the level variance of more that 10dB. The minimum variance method (upper panels) uses asymmetrical aiming (spatial tilt) to compensate for the speaker's asymmetrical orientation to the listening area. The level variance is less than 3dB in the minimum variance scenario. Note also that the spectral variance will be high in the lower scenario. The center seats are both near and on-axis whereas the most distant seats are also the most off-axis. The upper seats will be severely pink-shifted as compared to the on-axis area. In the minimum variance scenario (upper) the speaker/room summation will pink-shift the far on-axis response in a manner similar to the near off-axis response. This is a representative example of trading variance for level. The down-tilted scenario will be more powerful, yet will have higher spectral variance

Coverage Angle	Aspect Ratio
360	0.50
270	0.75
180	1.00
150	1.04
120	1.15
90	1.41
80	1.56
70	1.74
60	2.0
50	2.3
45	2.6
40	3.0
35	3.3
30	3.9
25	4.6
20	5.8
15	7.7
10	11.5
7.5	15.4
5	23.0

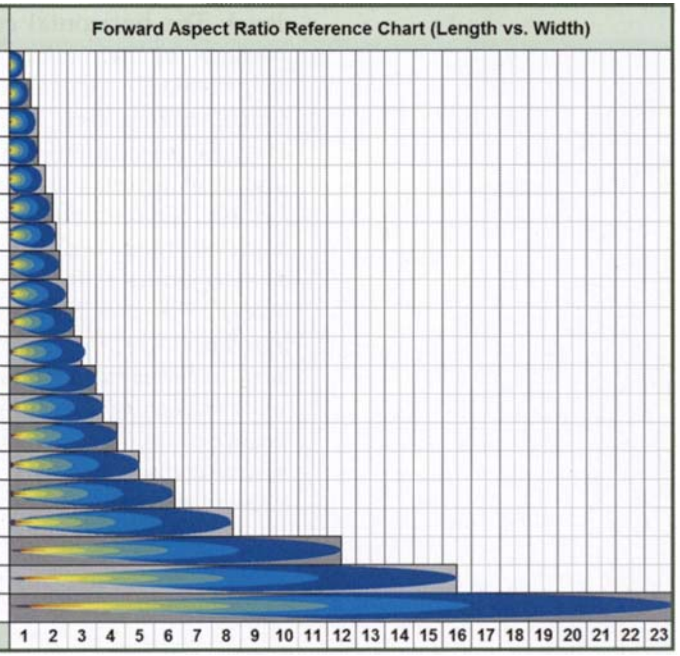


Figure 6.8 The relationship of coverage angle and aspect ratio for typical values

something stops it or it runs out of air. Aspect ratio gives us a scalable shape that we can work with for design and comparison.

A speaker with a 180 degree coverage angle has a FAR of 1. It sends as much energy to the sides as to the front. Its equal contour pattern is a square. A 10 degree speaker has a FAR of 12, with an equal contour pattern that is an extremely narrow rectangle. The FAR will help us put a square peg in a square hole and a rectangular peg in a rectangular hole when we design our systems.

The Proximity Ratio

For those applications where every seat is equidistant from our sources, obtaining equal level is a simple task of point and shoot. All others will need an alternate strategy. That strategy plays proximity in one plane against the other, resulting in offsetting level effects. A simple example is the flying center cluster. As we approach the cluster horizontally we move away from it vertically. The result is offsetting level effects that help us maintain a constant level. The horizontal coverage would not change if we were to set the cluster down on the stage floor. Our vertical orientation becomes flattened, and no longer provides an offsetting effect. The level variation would be the standard inverse square law drop-off from front to back. These different approaches illustrate the **proximity ratio,** the difference between the closest and farthest seats in the coverage of the speaker(s). If all seats are equidistant in a plane, the proximity ratio is 1. As the number climbs in one plane it must be offset in the other. This has practical limits and there are other considerations such as sonic image. In the end, whatever proximity difference remains creates an asymmetry level requirement for which our source must compensate. This can be achieved in at least one of two ways: the speaker coverage pattern must be oriented to compensate for the asymmetry, or multiple speakers must be employed to create an asymmetrical coverage shape in proportion to the listeners.

The proximity ratio can be expressed as a number or in dB (20hX log of the ratio); e.g. a ratio of 2 signifies that there is 6 dB of asymmetry that will need to be compensated for zero level variance.

The proximity ratio has implications regarding the control of spectral variance. While it is easy to design speakers and arrays with sufficient directional control to compensate the HF range levels for high ratios, this task is difficult to maintain in the low frequencies. This can result in large-scale spectral variance due to excess pink shift in the nearby areas.

The question remains how best to determine the need for supplemental fill in the gap coverage areas close to the speaker source. The answer depends upon whether or not we have listeners in those areas. If the listening area extends deep into the near corners of our aspect ratio rectangle, the response will be beyond our pink shift limits. Fresh on-axis sound from fill speakers will be required to reverse the pink shift.

The proximity ratio will help us to determine when to call for sound reinforcements (pun intended) in the form of downfill, sidefill, infill, etc. If the proximity ratio is 1 it is obvious that we would not need supplementary speakers. With a proximity ratio of 2 (6 dB) we can expect to offset that doubling of distance with 6 dB of axial loss in the near areas, as we have seen with the coverage bow. The pink shift from the axial loss in the near off-axis area may be comparable to that found from HF air loss and room reflections in the distant area. But this parity has its limits. A proximity ratio of 4 will require 12 dB of pink shift compensation between the two locations in order to minimize the spectral variance. This will require very strong directional control over an extended frequency range, a considerable challenge for an array, and virtually impossible for a single speaker. As the proximity ratio rises the speaker system must respond with increased directional control over the full bandwidth in order to minimize spectral variance.

Proximity ratio is a measure of the asymmetry of the coverage shape. The solution to the spectral variance created by a high proximity ratio is asymmetry in our speaker system. We will need to add array elements at lower levels to bring on-axis sound into the near areas. The on-axis infusion adds a mix of non-pink-shifted signal into the near coverage, and stems the tide of combined pink shift there. This addition raises the HF range in the local area, with minimal local LF effects and with negligible effects in the distant seats. The net result is a reduction in spectral variance.

Coverage areas with proximity ratios of 2 or greater can benefit from the addition of a fill subsystem. The auxiliary system is "layered" in under the main speaker. Each succeeding integer change in the proximity ratio is indicative of the potential benefit of adding another layer. This is another round in the ongoing battle to fight asymmetry (in throw distance) with asymmetry (in coverage shape).

Minimum Level Variance

The most basic shape of minimum level variance can be seen in the aspect ratio rectangle. The relationship is shown in Fig. 6.9. As the coverage angle narrows, the rectangle elongates. Provided that the orientation of the speaker is symmetrical to the rectangle, the aspect ratio, and hence the coverage angle, is the same. The minimum variance area can be found as a solid shape and as a line. The solid shape is the standard for symmetric applications such as is often found in the horizontal plane. The optimal area for minimum level variance is the back half of the rectangle, the entirety of which does not endure changes exceeding a 6dB range. The front half of the rectangle is highly variable, containing areas of high levels on-axis and extreme low levels in the front corners. In practical terms these areas are avoided by offsetting our orientation in the opposing plane, i.e. the speaker is lifted vertically so that no one is seated with their face plastered on to the cabinet front. (If this is unclear I suggest a review of the section on 2-D to 3-D conversion in Chapter 5.)

The line of minimum level variance is most representative for asymmetric applications, such as typical vertical coverage, and extends from the on-axis rear to the off-axis mid-point. Minimum variance in our 3-D world will occur

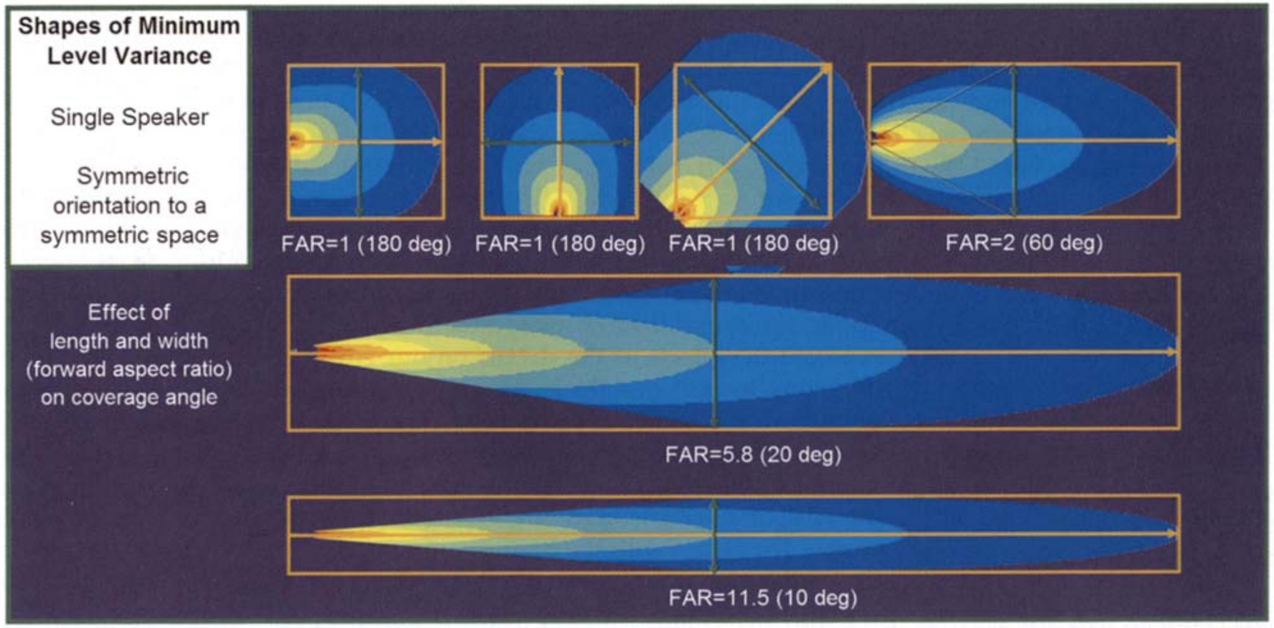


Figure **6.9** Various examples of the relationship of the equal level contours, coverage bow and aspect ratio rectangles

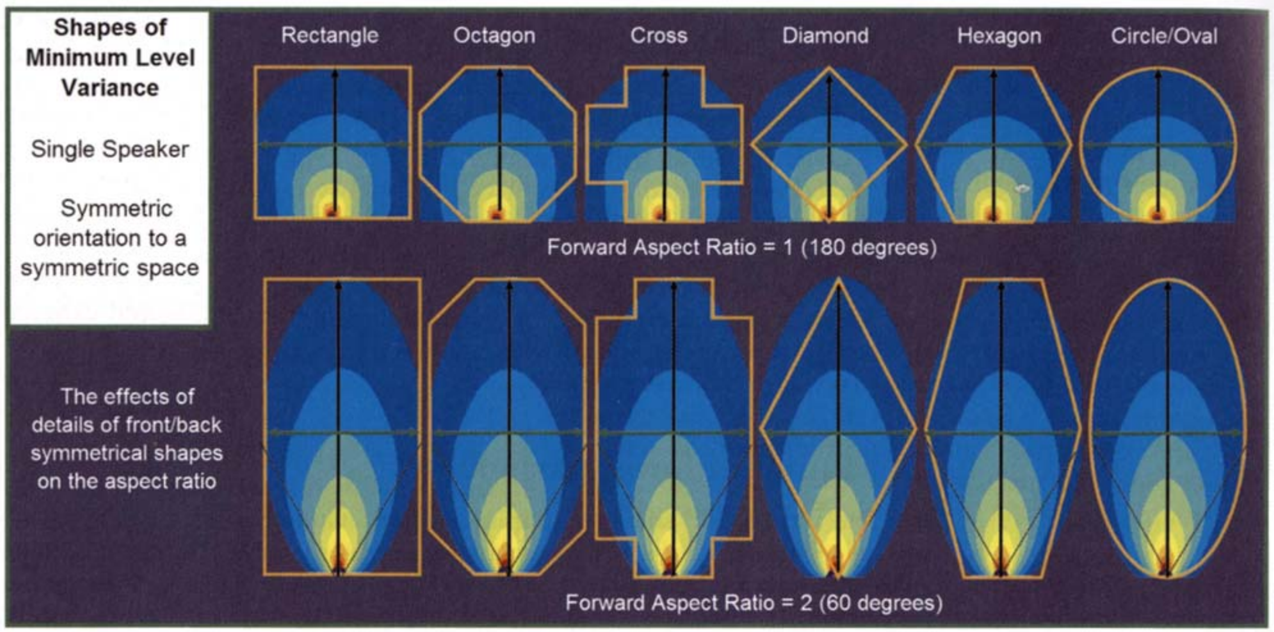


Figure 6.10 The rectangular shape of the aspect ratio compared to other front/back and left/right symmetrical shapes

when we have the line of minimum variance covering the vertical plane at the same time as the solid shape of minimum variance covers the horizontal plane.

Single Speakers

Fully Symmetric Shapes

Since not all listening spaces are rectangles it is necessary to consider other basic shapes as shown in Fig. 6.10. Here we see various left/right (top and bottom for vertical applications) and front/back symmetrical shapes and can evaluate the effects of the details on our coverage angle decisions. In each case the lengths and widths at the mid-points are the same, which renders them equivalent in aspect ratio to the rectangle. The coverage angle to achieve minimum level variance remains constant, although the amount of overflow and underflow changes for each shape. The shape with the highest amount of overflow is the diamond, but this is also the shape where the coverage angle presented by the shape is exactly equal to that of speaker! It can also be seen that there is a tracking relationship of these shapes to the square (top row) and the rectangle (bottom row), a trend that continues with longer aspect ratios.

Front/Back Asymmetric Shapes

There are also shapes that are left/right (top and bottom for vertical applications) symmetrical but asymmetrical from front to back. Three of these shapes are shown in Fig. 6.11 in both orientations. In all cases we maintain a constant on-axis length, but the mid-point width changes for each shape. The decisive factor in determining aspect ratio is on-axis length by mid-point width, and therefore the coverage angle changes in this case. The details of the shape determine the amount of overflow and underflow, which is proportional to the ratio of length and mid-point width. The aspect ratio is the same whether the shape is wider in the far field or reversed. The overflow areas are also symmetrically opposite but an important point must be made here: the differences in axial orientation and source

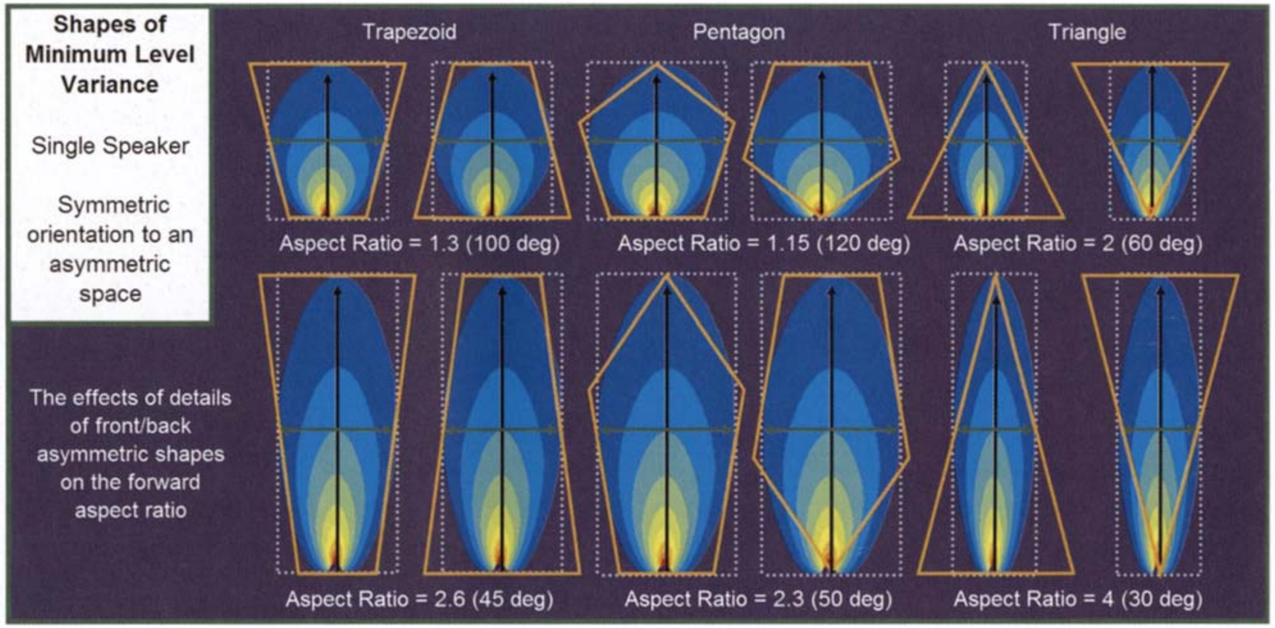


Figure **6**.11 The rectangular shape of the aspect ratio compared to other shapes that are front/back asymmetrical and left/right symmetrical

proximity means that these areas will have far different responses. An underflow (or overflow) area at the top of these shapes is still within the coverage angle and most likely only a few dB different in level from the on-axis response. The symmetrical opposite at the bottom is off-axis and has rapid doubling distance loss. Therefore there is no single speaker solution for this type of coverage shape. The wide bottom, narrow top configuration will require a main system (as shown) and two sidefill systems to fill in the remaining shapes.

Left/Right and Front/Back Asymmetry

Another form of asymmetry arises when the shape is oriented differently to the left and right (top and bottom for vertical applications) of the speaker. This occurs when an off-center origination point for the source is aimed into a symmetric shape or with a center origination into an asymmetric shape. These scenarios are shown in Fig. 6.12. The key to maintaining minimum level variance is to match symmetry with symmetry, or to compensate asymmetry with a complementary asymmetry. The speaker placed in the corner of a rectangle must be aimed at the opposite corner to maintain symmetric balance. A speaker placed in the corner of a rhombus is also aimed at the opposite corner. In this case we have asymmetric coverage meeting asymmetric orientation. The left and right sides will not be matched but the rear areas will. A speaker that is placed in the corner on the side of a rhombus (an asymmetric left/right presentation to the speaker) cannot achieve minimum variance. If the speaker is asymmetrically oriented it will be unmatched on the sides. If it is oriented symmetrically the coverage along the rear will vary in level.

The Maximum Acceptable Variance Method

An alternate to coverage angle determination is the maximum acceptable variance method, which seeks to find the coverage angle where the —6 dB points fall on the listening area edge. The degree of contrast this method has to the minimum variance method depends upon the coverage area shape. The difference between the two methods can

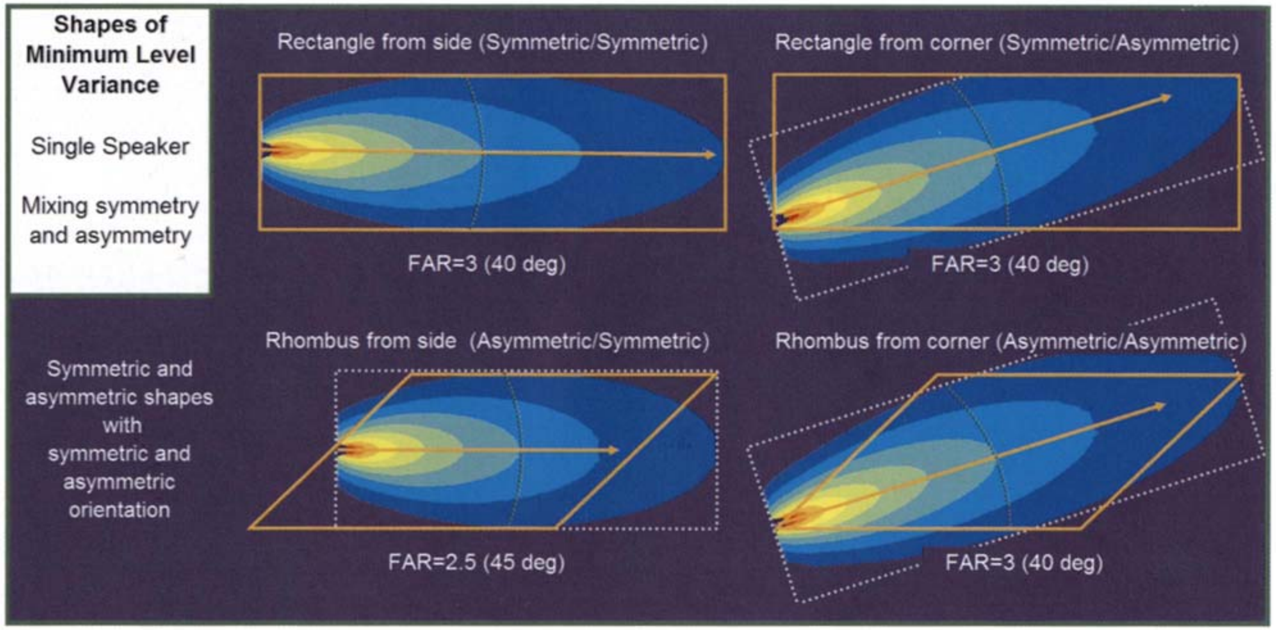


Figure 6.12 Evaluation of the aspect ratio for shapes and orientations that are front/back and side/side asymmetrical

Perspectives:

1. *Point the speakers at the audience. Firing at walls does not help.*
2. *Try to get all frequencies at the same level and time to all members of the audience.*
3. *Listen — but not only to the loudspeakers. It's the performance which counts.*
4. *Less is more. Especially if half of it is out of phase.*
5. *Don't say "flat" in front of producers and band managers.*

Thomas Züllich

vary from a difference of 2:1 to none at all. The decisive factor is the choice of the width. The maximum acceptable variance method would use the end-point width as the reference point in contrast to the mid-point width used by the minimum variance method. In the rectangle this change reduces the coverage angle by half as shown in Fig. 6.13. The result of this angular reduction is that the mid-point side areas are now —6 dB from the far on-axis response. This contrasts to the O dB difference of the minimum variance method. This disadvantage is countered by the reduced level of reflections if the side edge of the shape happens to be a surface. As a result we will keep both methods in mind, knowing that they represent the maximum and minimum coverage angles suitable for the job. The minimum variance is the preferred method but will need to give way in cases where the ripple variance caused by reflections would be worse than the level variance caused by the angular reduction. Also note that the maximum acceptable variance method increases the size of our near side underlap areas. These areas will need supplemental coverage.

In the horizontal world of speaker coverage we seek to fill open containers with a field of sound. These were represented by the open shapes shown in the previous figures. In the vertical world we lay a line of sound atop a solid surface. The result is that the vertical coverage is inherently asymmetrical and, with rare exception, we will be using little more than half of the speaker's coverage pattern. Nevertheless the aspect ratio is still the key parameter for evaluating coverage. However, we will modify our approach to accommodate the key geometric representative of the vertical plane: the right triangle. The triangle is the essential shape for vertical coverage because it best serves our second most important sensory organ: our eyes. The audience area is seated along the hypotenuse to allow viewing of the performance, providing a triangular shape to the speaker's perspective. For those applications where performances are viewed from flat fields we must bear in mind that the speakers are lifted high enough above the audience so that the same perspective is created. Readers wishing to replicate the flat field are advised to tilt this book to the right until the desired effect is achieved.

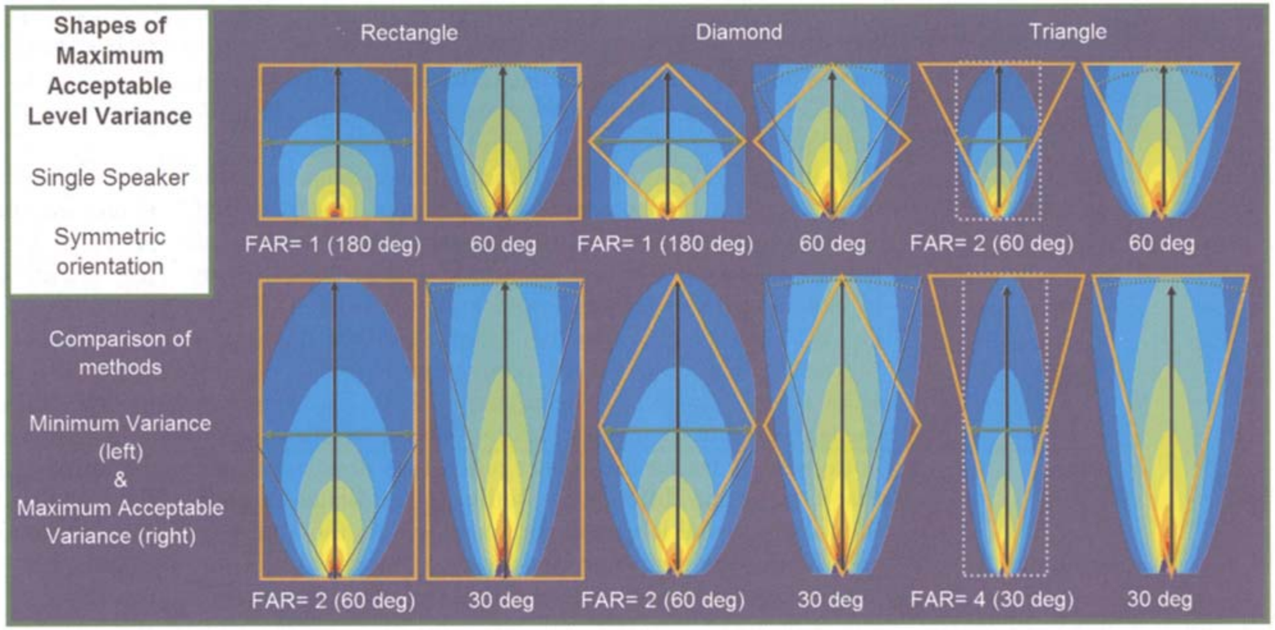


Figure **6,13** The maximum acceptable variance method results in a speaker coverage that is 50 per cent of the minimum variance method

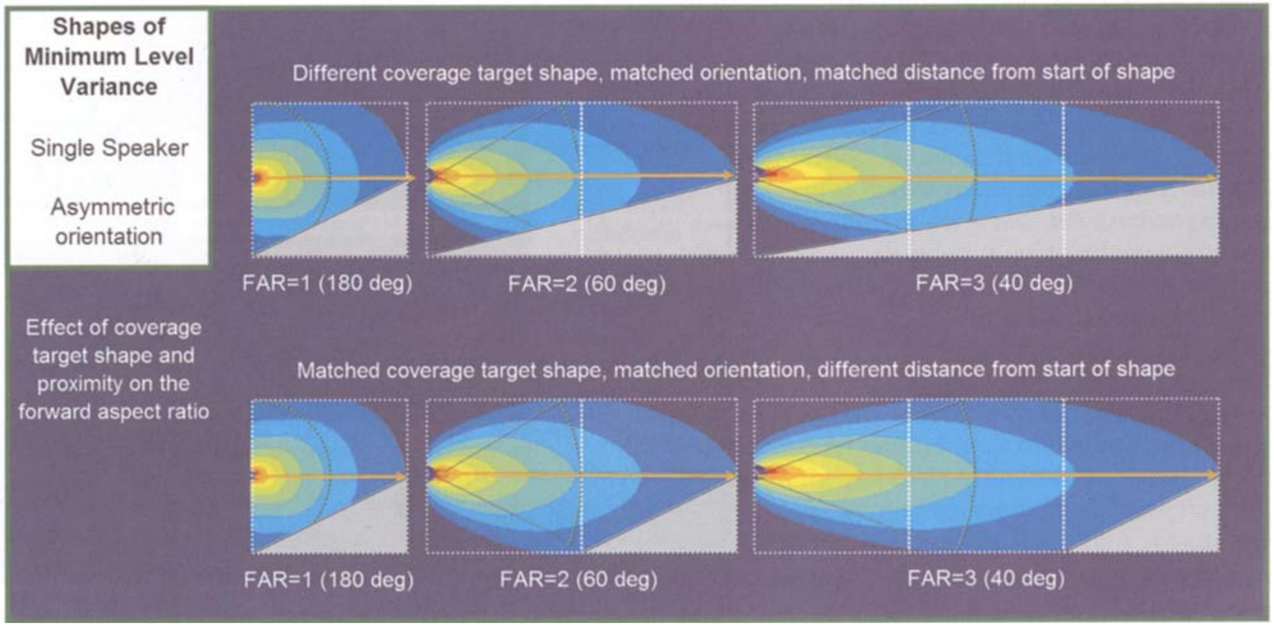


Figure 6.14 Minimum level variance shapes for a single speaker with an asymmetric line of coverage

Our concern when laying down lines of coverage is shaping the minimum variance levels at the point of impact on the surface (the hypotenuse). The length of the aspect ratio is found by orienting the speaker toward the farthest point (as we have done previously). The width is found by the same method as before: a line is drawn from the mid-length point to the bottom of the coverage, and extended symmetrically above. There are three primary variables at play here and we will once again do our best to view them in isolation. The first are length and width as shown in Fig. 6.14. As we move further from the shape the aspect ratio rises for a given height. Notice that neither the slope angle nor the depth of its origin are decisive. The length to the last seat and the height difference between the top and bottom seats frame the aspect ratio. This does not mean that it makes no difference whether the audience extends the full or partial length of the coverage. As the audience area fills in the near area, the single speaker solution no longer applies and the need for supplemental downfill speakers arises. The coverage shape of the main speaker is, however, unchanged.

Orientation to the Shape

The third factor is the angular orientation between the source and surface. The coverage angle required for a given shape at a given distance will change its aspect ratio as the orientation angle moves. A representative scenario is found in Fig. 6.15 in which the coverage angles range from 32 to 180 degrees to accommodate the different orientations. In most cases the aspect ratio is found by drawing a box that includes the on-axis point (center of the length) at the farthest seat and the nearest seat being the rectangle outer edge. There is one notable exception at the upper right. In this case the speaker angle is equal to the hypotenuse angle, and therefore presents a symmetrical flat surface. Here, the aim point should logically be the center.

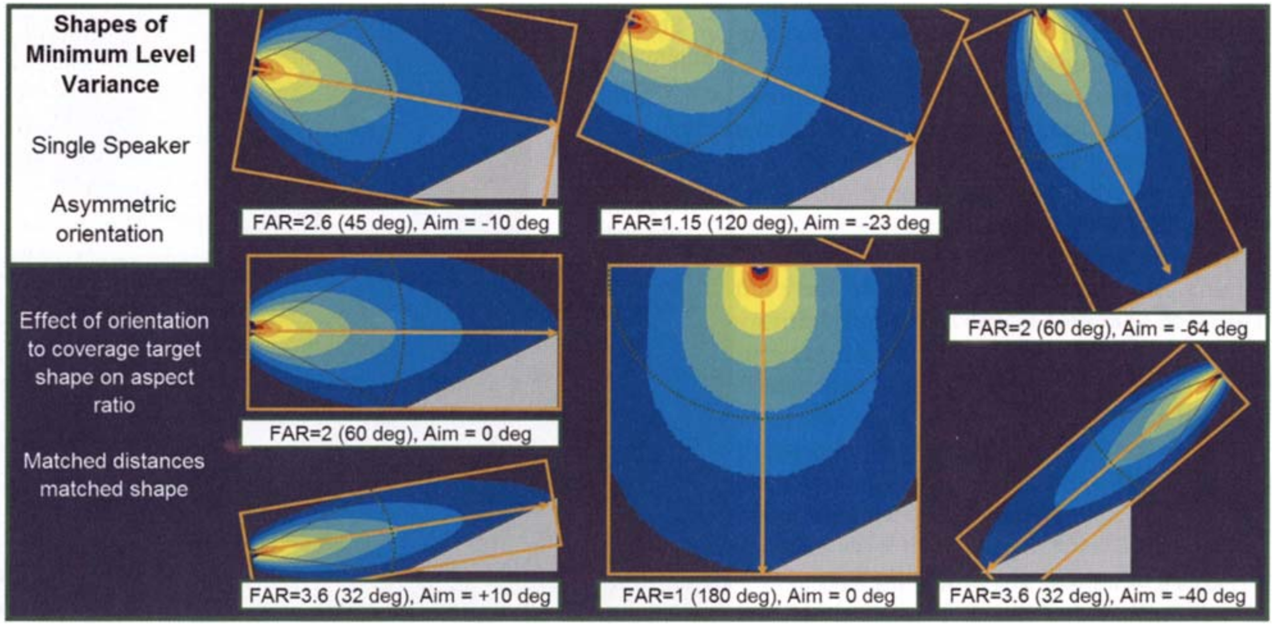
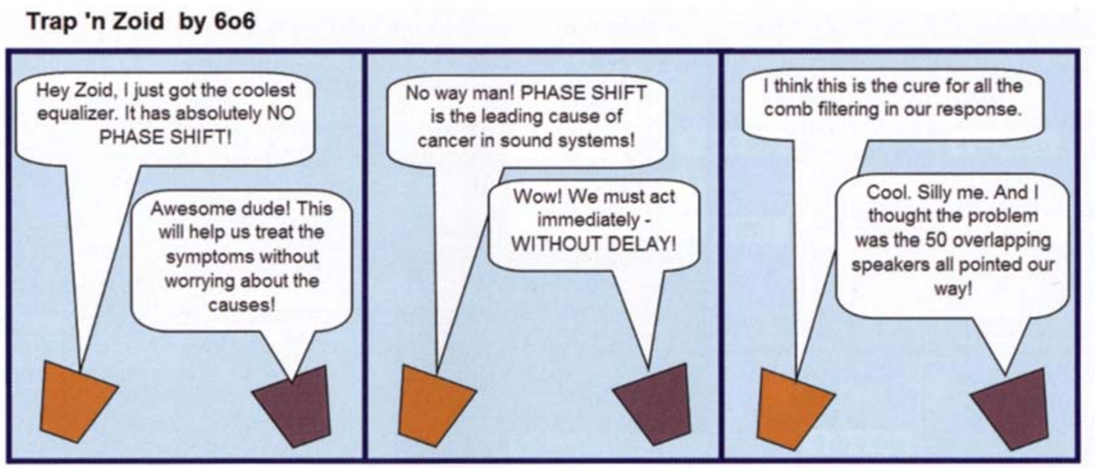


Figure 6.15 Minimum level variance shapes for a single speaker with an asymmetric line of coverage



Coupled Speaker Arrays

It is now time to see how these principles apply to speaker arrays. What follows is a series of scenarios for various arrays. Each contains unique characteristics and serves to illustrate one or more of the mechanisms at work. In most cases there are renderings of both the individual aspect ratio icons and acoustical prediction, in order to illustrate the different contributions to the combined effect. The upper portions of the figures show aspect ratio icons placed on top of each other to make a combined shape. The lower panels show the acoustic prediction of the same speaker at a matched scale.

Coupled Line Source

We will begin our comparisons as usual with the coupled line source array as shown in Fig. 6.16. The two scenarios chosen to represent this are arrays comprised of third- and second-order elements. The third-order system (left panel) was chosen because such systems are widely in use. The second-order system is shown for comparison purposes because most of the other array scenarios in this series will use that model. We begin with the third-order system. The difference between the combined aspect ratio icons and the individual component parts is seen as a stretching of the rounded front of the on-axis area into a flat line. The coverage angle appears stretched by the line length (or height) but no change in angle is indicated. The stretching creates a line of minimum variance that is approximately as long as the array length (or height). If we were to double the line length, the coverage line would also double. This line can not be quantified as a coverage angle, but rather a coverage length. This is a challenging concept for audio engineers as we are accustomed to visualizing coverage in angular terms. A look at the aspect ratio rendering shows the vertical line of minimum variance moving steadily away from the speakers and never changing its height. Meanwhile the rounded outside edges that mark the coverage angle continue to expand outward in scale with the doubling distance. In the prediction panel below we see that the vertical line of minimum variance appears but the rounded coverage angle edges do not. This array has no coverage angle as shown, and will not develop one until it has reached far enough way to have found the top of our old friend, the parallel pyramid (Chapter 2). At the point where it gains a coverage angle it will lose its fixed width minimum variance line. At this point we can see that our aspect ratio combined rendering tells only part of the story. We will get to the rest of the story shortly.

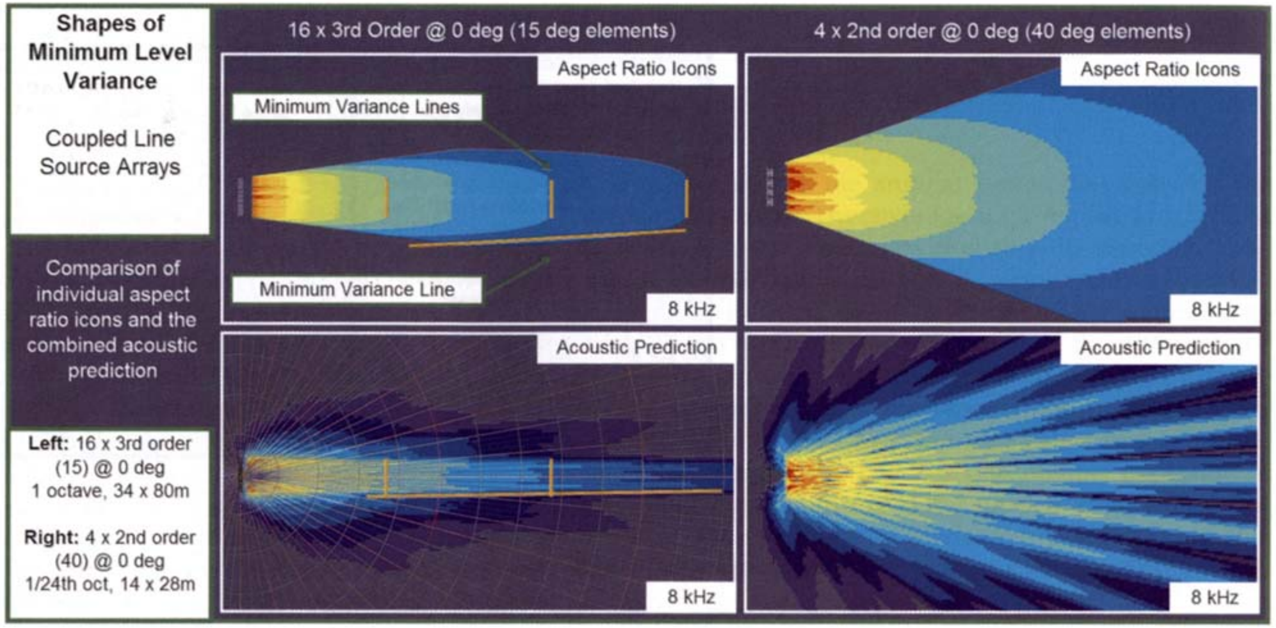


Figure 6.16 Minimum level variance shapes for the symmetric coupled line source array. Left 16-element third-order speaker array with 0 degree splay angle and constant level. Right a four-element second-order speaker array with 0 degree splay angle and constant level

The next item worth noting is the effect of the coupling on the angled minimum variance line found on the underside of the array. The upward angle seen in the icons has almost leveled off in the combined prediction. This is another challenging concept, a line of minimum variance that appears tangential to the array orientation. The hopes for minimum variance in a tangential array rest on its ability to overcome three extreme challenges: to find an audience with a tangential orientation or to maintain the same coverage width over frequency. The evidence presented as we move through this section suggests that this is not feasible, barring acoustic transmission technologies beyond the scope of my understanding. When we investigate the pursuit of minimum spectral variance, the vulnerability of this approach should become clear.

The second-order system shows a higher correlation in one regard to the individual response in that the overall shape of the coverage angle is matched. This is due to the fact that coupling zone summation does not occur because of the large wavelength displacement. The array does not focus the high frequencies into a concentrated beam as did the third-order system. Instead the patterns pass through each other, creating a scattered appearance. The result is high ripple variance (combing zone summation) in the combined response which is not represented in the aspect ratio icons. The excessive ripple variance renders our search for minimum level variance here academic. In the top panels the aspect ratio icons failed to incorporate the beam behavior of the overlapped speakers. In the lower panels the icons failed to account for the combing behavior of the overlapped speakers. These icons have their limits. Let's try another approach.

Instead of stacking the aspect ratio icons side by side (like the speakers) let's try stacking them end to end (like the speaker patterns). The combined shape of a fully coupled line source can be visualized as a forward extension by means of simple additions of the aspect ratio shapes. This relationship is shown in Fig. 6.17 where successive doublings show the forward extension. A single element provides the base aspect ratio shape. Each time a speaker

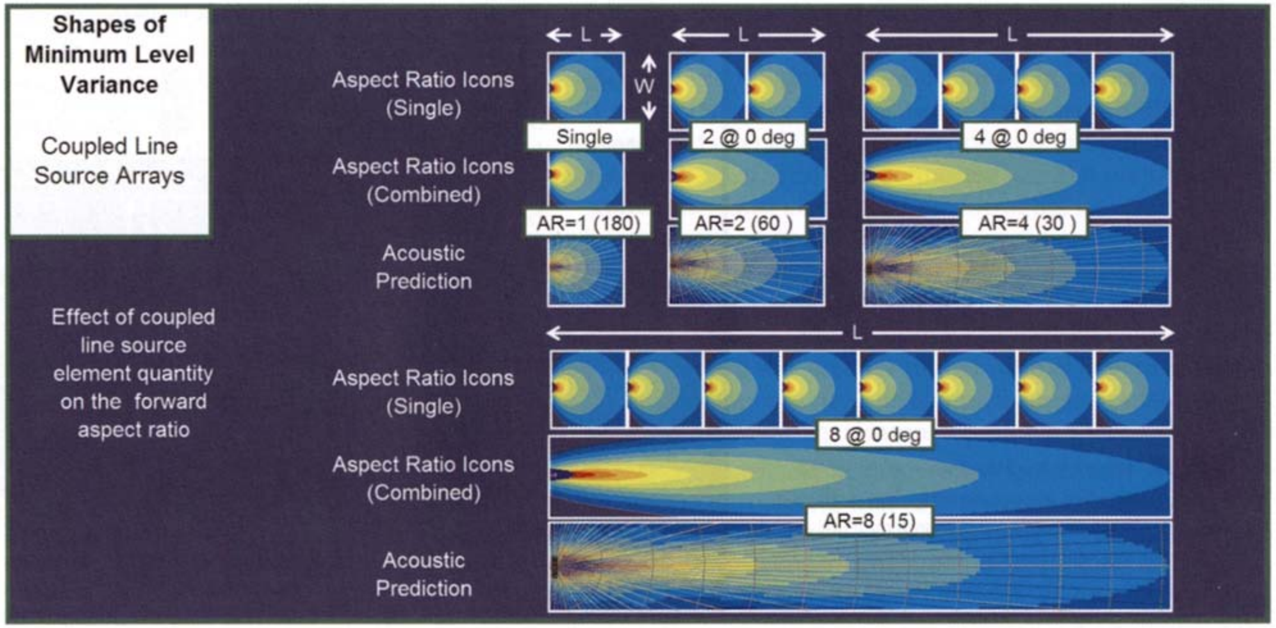


Figure 6.17 The effects of quantity on the combined aspect ratio for the symmetric fully coupled line source array. The elements are stacked along the width. In this example each element has a 180 degree pattern. Successive quantity doublings cause successive doublings in the ratio of length to width. The combined aspect ratio matches the shape of the individual element aspect ratios placed in a forward line. Is this a coincidence?

is added to the coupled line source the combined pattern is narrowed by the pattern overlap. The individual shape of each additional array element is added in front of the previous one, thereby creating a combined aspect ratio which is, in length, a simple multiple of the original element. Each time the quantity of elements is doubled the FAR doubles, and the on-axis power capability adds 6 dB. In each case three renderings are seen, the individual element blocked as they are stacked forward, the combined AR icon and an acoustic prediction of the given quantity of elements.

Note: This effect will not occur until the array is fully coupled; i.e. the individual patterns of all elements have overlapped. Remember that coupled arrays are within partial wavelengths of each other, and uncoupled arrays are multiple wavelengths apart. The wavelength displacements must be small enough for the beams to concentrate.

The combined shape of the uncoupled line source can be seen as a lateral extension by means of simple stacking of the aspect ratio shapes. The combined aspect ratio becomes wider while not becoming longer and therefore lower in number. The combined FAR is found by dividing the single element FAR by the number of elements. A single 360 element provides the base aspect ratio shape for our example scenarios found in Fig. 6.18. Its FAR value is 0.5, the lowest possible for a single element. Four sequential element quantity doublings follow from left to right that each add up to approximately the same combined FAR. In each case the coverage angle and spacing are reduced proportionally as the quantity increases, thereby preserving the combined shape.

At the 50 per cent point of the coverage length we can find a line of minimum variance which connects from the on-axis point of each element though the spatial crossover of the adjoining elements. This line extends to the on-axis points of the outermost elements and then falls off to the point of maximum acceptable variance, —6 dB. The uncoupled array cannot be characterized as having a combined coverage angle, but rather a defined spatial shape. Our combined coverage fills the same shape as the single

255

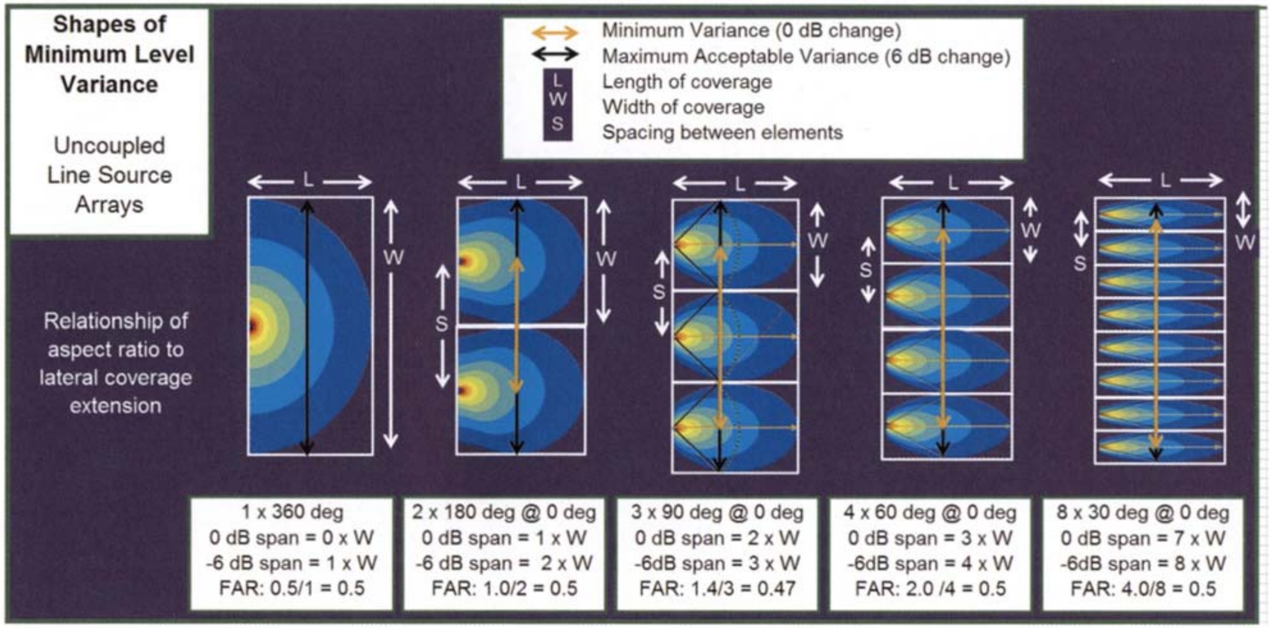


Figure 6.18 Minimum level variance shapes for the symmetric uncoupled line source array. The combined aspect ratio of an uncoupled line source matches the shape of the individual element aspect ratios placed in a lateral line. Note the contrast to the coupled line source shown in Fig. 6.17

360 degree speaker, and fills in greater amounts of the rectangle as the slices become smaller.

It is reasonable at this juncture to ponder how the coupled and uncoupled versions of the line source array could provide such different responses. In the coupled version the aspect ratios extend their length, while in the uncoupled version they extend their width. The difference is attributed to their position in the parallel pyramid. The fully coupled response is the top of the pyramid, while the fully uncoupled is the bottom. In reality, the uncoupled array is simply a coupled array waiting to happen unless the displacement is too high to allow a coherent beam to form (as with the second-order speakers in Fig. 6.16). We just have to move forward the distance required for all of the elements to have overlap, at which point the array is "coupled" and can be described by a coverage angle. A coupled array, conversely, is an uncoupled array that has already congealed. Even the most compact array imaginable has some point in its near field where we have not yet found the unity class crossover. It is uncoupled at that point. The key link here is wavelength. The transition point from coupling to uncoupling is wavelength, and therefore frequency-dependent. There are several reasons to split this hair. The first is that the bottom two floors of the parallel pyramid are one of the minimum variance shapes that we will be using in our designs. The second is that the transition zone between these two worlds is a poorly understood area in modern speaker array design. Long, coupled line source arrays and hybrid "J" arrays are often designed for rooms with large amounts of the listening area inside the highly volatile and variant space between the uncoupled and coupled worlds. The behavior of arrays in the process of transforming between these opposing forces of beam spreading (at the bottom) and beam concentration (at the top) is very challenging to characterize. The array cannot be defined by either a constant width or angle in this transition zone. If an array can not be defined, our chances of finding a repeatable alignment strategy that can provide minimum variance are poor. Any doubts about the difficulty that this zone poses can be alleviated by glancing at the published design data from different manufacturers of "line array" systems. Those expecting to find

Perspectives: Some of my Cirque du Soleil designs have been touring since the early 1990s. Increasing our tuning and analysis capabilities allowed us to clearly see how and where initial designs were flawed. Each time we set up and re-tuned the systems, we were able to make changes that increased the intelligibility and overall audio quality. The skills required from the audio teams today are totally different than 1.5 years ago. The understanding of the laws of physics as they apply to complex systems and the introduction of source-independent analyzers revolutionized system tuning for live shows. Of course, the audience also keeps raising their expectations.

Francois Bergeron

a clear and straightforward answer to "what is the coverage angle?" are in for a surprise.

Before we move on let's go back and take a look at the left panels of Fig. 6.16. The reason that the aspect ratio icons did not match the predicted response is that we used the uncoupled method (spreading the icons) rather than the coupled method (stacking the icons). The predicted response, meanwhile, is working its way up the pyramid.

Coupled Point Source

The addition of angular separation complicates the relationship of the individual aspect ratios to the combined shape. The forms of addition are no longer exclusively beam concentration, but rather beam spreading. The combined shape is expressed as a coverage angle, in degrees, rather than the rectangular shape of the aspect ratio. The shape of minimum level variance is now an arc. Our pizza has finally arrived.

The coupled point source spreads energy in radial form. There is no simple relationship between the combined aspect ratio and the splayed elements in the couple point source (see Fig. 6.19). The individual aspect ratio rectangles are spread along the arc like a fan of playing cards. If the unity splay angle is used, the gaps are filled in by the shared energy and a radial line of minimum level variance is created. The line connects from the on-axis center of the outermost elements. Outside of the last elements the line of maximum acceptable variance continues until the -6 dB point is reached. An array for a given angular span of minimum variance can comprise a small number of wide elements or a large number of narrow elements. The latter scenario will create a higher ratio of coverage within the minimum variance line compared to the outside edges framed within the maximum acceptable span.

In the scenario above (shown in Fig. 6.19) there are a few matters worth noting that are not immediately apparent. The first matter concerns the difference between the minimum variance line (0 dB) and the maximum acceptable variance line (0 to -6 dB). The upper right portion of the figure shows a 180 degree single speaker, and its standard square aspect ratio shape. The radial line of maximum

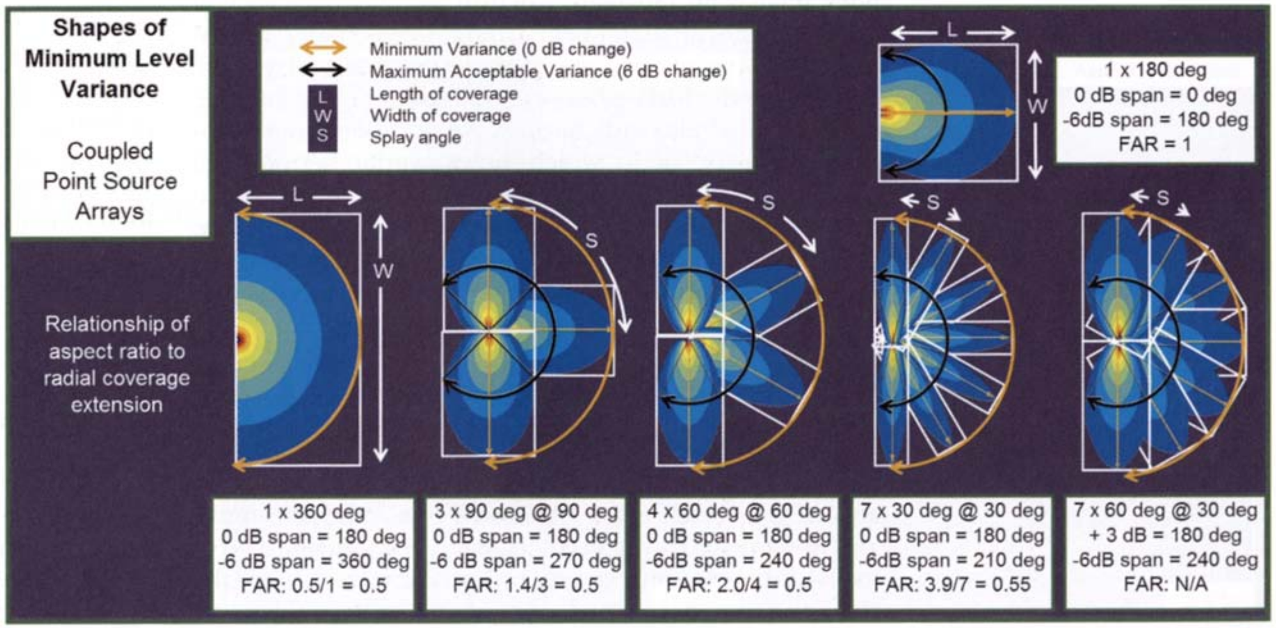


Figure 6.19 The effects of quantity on the combined aspect ratio for the symmetric coupled point source array

acceptable variance spans 180 degrees, and yet the radial line for 0 dB spans 0 degrees. Why? Because a radial arc drawn with the speaker at the center would trace a line that immediately shows level decline as it leaves on-axis until the —6 dB point is reached. By contrast let's view the 360 degree pattern shown on the lower left. This pattern has a forward aspect ratio of 0.5, since half as much energy is going forward as is going to the combined sides. This speaker has a 180 degree span on the 0 dB line and a 360 degree span on the 0 to -6 dB line. Now we will view the combined arrays and compare and contrast them to the two singles. In each of the four cases the combined minimum variance line has a span of 180 degrees. The maximum acceptable variance span, however, changes with each array. The outside half of the outermost elements extends this line beyond the 0 dB line, until -6 dB is reached. The difference between the two spans is the total coverage angle of one element. The lesson here is that an array made of larger quantities of narrow elements will have a sharper edge than one made of fewer wide elements. The edge is revealed as the difference between the minimum variance and maximum acceptable variance lines. If we want 180 degrees of 0 dB level variance we will need to create an array with a 180 degrees span between the on-axis points of the outermost elements. If we are concerned with leakage we will find advantage in keeping the individual elements narrow. An additional note concerns the last scenario, which shows overlap between the elements. While the overlap will add power capability to the front of the array and cause increased ripple variance, this will not change the ratio of the minimum variance and maximum acceptable variance spans.

The second notable feature is the relationship of the individual and combined aspect ratios. There are two factors which will make the combined FAR unrepresentative for these arrays: coverage that extends rearward and overlap.

The combined aspect ratio for all of our arrays settles around 0.5, which is the FAR value of a 360 degree individual element. This is no coincidence. The radial spread of the multiple elements creates an equal level arc that is the same shape as that found in the 360 degree single element. The symmetric coupled point source is essentially recreating sections of a 360 degree shape. The extent to which the shape is filled is based on the individual elements, splay angle and quantity, but once we have exceeded 180 degrees of 0 dB coverage the FAR can not fall further below its 0.5 value. Four sequential element quantity doublings follow from left to right which each add up to approximately the same combined FAR, yet with varying amounts of extension behind the array center.

When there is angular overlap, the combined shape will behave as something in between the coupled line source and coupled point source. As overlap increases the pattern narrows and extension moves forward. As overlap decreases the pattern extends radially. This dance will be a major ongoing theme in this chapter.

Symmetric Coupled Point Source

The symmetric coupled point source can provide curved lines of equal level radiating outward from the virtual point source. This is the most intuitive of the array types. The combined response will strongly resemble the spread of the individual aspect ratio patterns as long as the overlap is low (see Fig. 6.20). As overlap increases, the combined pattern will lose its resemblance to the individual responses as the pattern bulges forward at the center and potentially develops side lobes. The left panel of Fig. 6.20 shows a second-order array with minimum overlap, while the right panel shows a large third-order array with greater than 50 per cent overlap. In both cases the combined predicted response holds the spread shape of the individual components.

Asymmetric Coupled Point Source

An asymmetric coupled point source can also create curved and diagonal contours. The methods used to create this effect are shown in two scenarios in Fig. 6.21. The left panels show a log level taper and constant splay angle (50 per cent overlap) among matched components. Such an array aims the on-axis point of each succeeding element at the —6 dB point of the one above. Each layer is also tapered by —6 dB resulting in a (—6dB) + (—6 dB) meeting point in

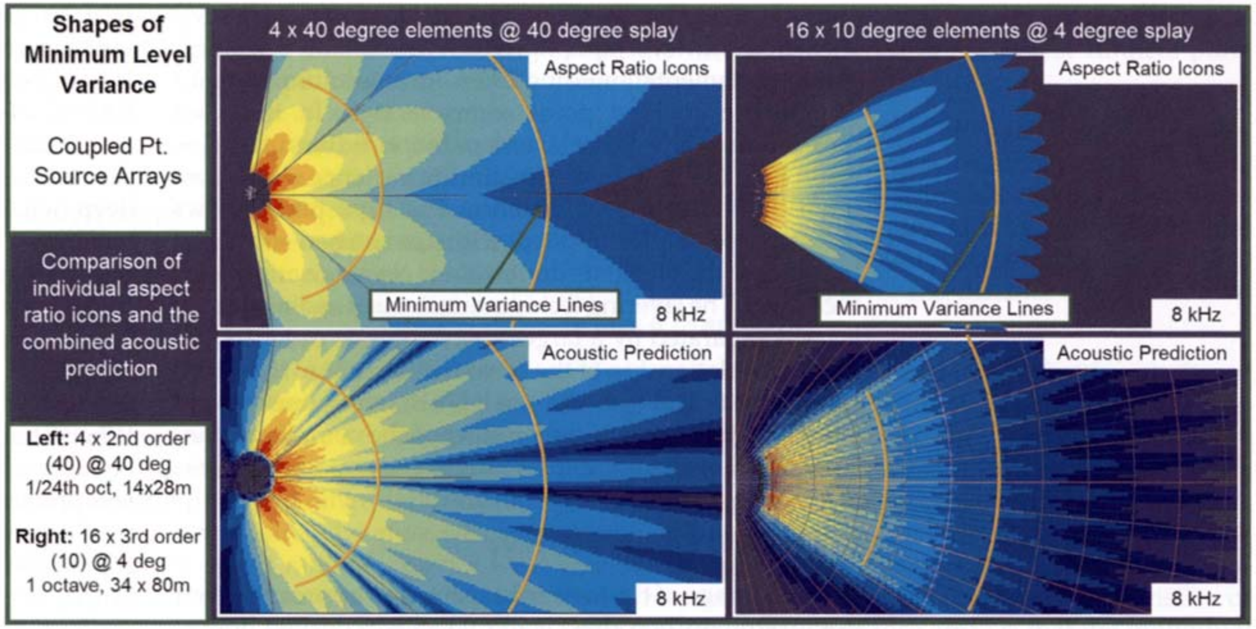


Figure 6.20 Minimum level variance shapes for the symmetric coupled point source array. Left constant speaker order, unity splay angle, and constant level. Right constant speaker order, 60 per cent overlap splay angle, and constant level

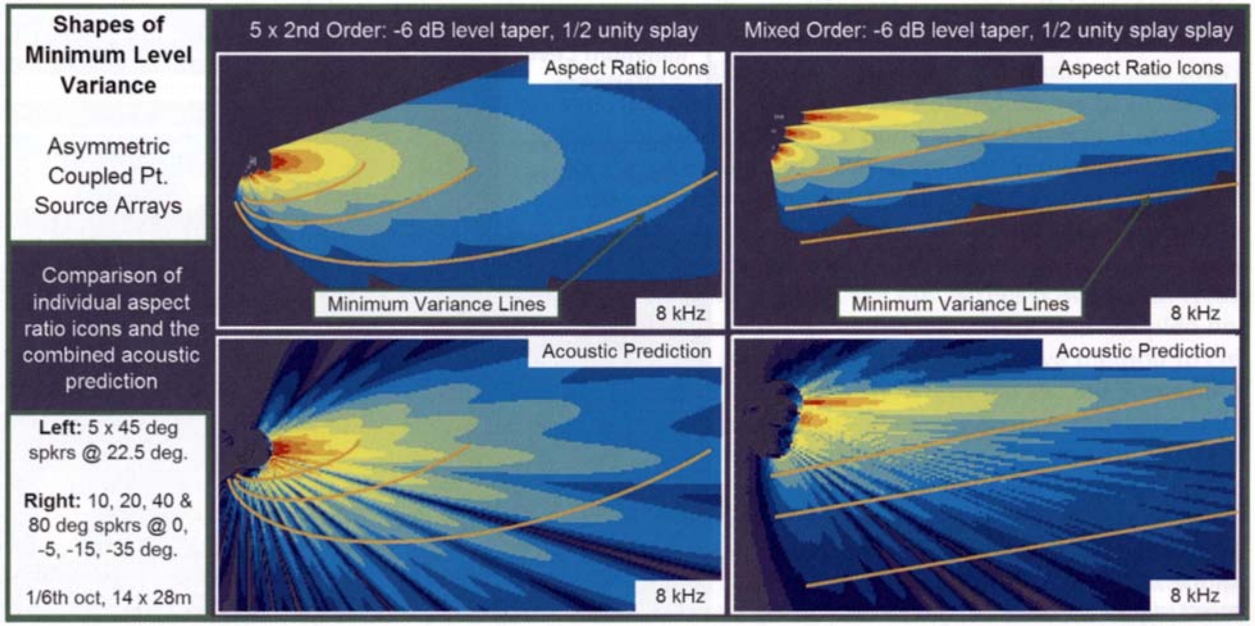


Figure 6.21 Minimum level variance shapes for the asymmetric coupled point source array. Left constant speaker order and splay angle, with tapered level. Right mixed speaker order with splay angle tapered to provide a constant overlap ratio, with tapered level

front of each lower speaker. This technique of partial overlap mixed with level tapering, which we will call "layering," is the fundamental building block for minimum variance asymmetric coupled point source arrays. In this case the maximum amount of layer separation, 6 dB, is used. The result is a curved region of minimum variance which continues indefinitely over distance. The right panel shows the same principles applied with unmatched elements. In this case the elements double their coverage angle with each layer. The layering technique is applied again but each layer is separated by a larger splay angle in order to preserve the relationship of the on-axis aim point to the —6 dB edge of the unit above. The result is a diagonal line of equal level. The difference is the product of the complementary asymmetry of changing splay angles and coverage angles.

Uncoupled Speaker Arrays

The advantage of the forward aspect ratio approach to our design strategy becomes clearer when we consider how to manage a space that is wider than it is deep. Once we go past 180 degrees of forward coverage we can't go wider without going backwards. 360 degrees of coverage does indeed give us coverage twice as wide as 180 degrees (the FAR for 360 degrees is 0.5 since half the energy goes forward) but the energy going to the rear could be a serious problem. If the shape is more than twice as wide as it is deep, what then? 370 degrees? 720 degrees? The best way to deal with coverage shapes which are wider than they are deep is by using multiple sources, i.e. uncoupled arrays.

Uncoupled Line Source

Symmetric Uncoupled Line Source

When the aspect ratio of a desired coverage area falls below a value of one, the single speaker solution becomes less preferable. For simple rectangular shapes the uncoupled line source is capable of extending coverage laterally. Determining the range of minimum variance for such arrays is a simple matter since it is directly related to the aspect ratio. Simply put, the minimum variance coverage zone begins at the 50 per cent point of the aspect ratio

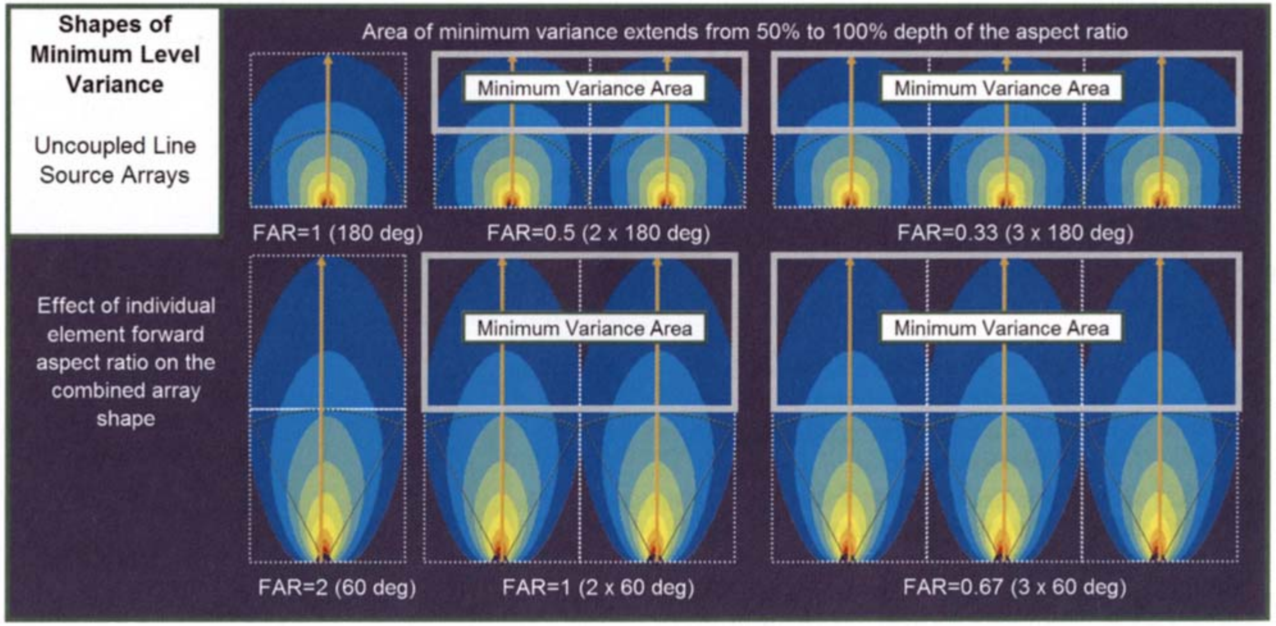


Figure 6.22 Minimum level variance shapes for the symmetric uncoupled line source array. The combined aspect ratio is a multiple of the individual element FARs

length and ends at the 100 per cent point. This is a scalable result. Let's look at Fig. 6.22 and use the data as an example. The upper series uses 180 degree speakers with an FAR of 1. The shape, a square, will be given an arbitrary dimension of 10 m. The speakers are spaced 10 m apart and so the coverage begins (the unity crossover point) at a distance of 5 m forward and overlaps with additional elements at the 10m point. The distance where the overlap from additional speakers joins the party (triple coverage or more) is considered the end of the minimum variance response area. If we change the distance between the elements to 1 m or 100 m the beginning and end points remain the 50 per cent and 100 per cent points of the aspect ratio depth. Now let's move to the lower series where a 60 degree element is used. The same distance is found between the elements but the start point, end point and range depth are all doubled, precisely as is the aspect ratio. Therefore we can conclude that if we know the aspect ratio of a device we can determine how far to space the elements to begin coverage. Likewise, we can predict the desirable end point to the coverage, beyond which the signal will become highly ripple-variant. This tells us when we should use other speakers which can provide lower variance in the local area. If we know the spacing and where the coverage is desired, we can determine the required aspect ratio.

Asymmetric Uncoupled Line Source

Minimum level variance can be spread asymmetrically over a limited area with uncoupled sources if level, position and orientation are carefully coordinated. The principle involved is one of offsetting effects, in this case distance related loss and electronic gain. Consider a race between unequal runners. If they all share the same starting point the outcome is assured. The faster runners will take an early lead and never relinquish it. They will run a proportionally larger distance in a given amount of time. Now consider what will happen if the starting points are staggered such that the slower runners have a proportional head start. The final outcome is still certain to be victory for the fastest runner, at least eventually. At the beginning of the race the slowest runners will lead. There will be one moment where everyone comes together as the head start compensation has elapsed for all parties. From this point the race is restarted and the fastest runners plow ahead.

This is how sound sources at unequal drive levels interact. The louder source always wins. If they are to be made equal for any position forward of the origin of the louder speaker, we must handicap the race by moving the quiet speaker forward by an amount proportional to its level offset. The relationship of level and distance offsets is shown in Figs 6.23 and 6.24.

Two examples of distance and level offsetting are found in Fig. 6.25. These are constructed to illustrate the process, rather than any particular practical application. Each of the figures from this series contains views of the individual responses as aspect ratio renderings, and their combined predicted response. This gives insight into the role that the individual elements play in creating the combined response. One of the consistent characteristics of most minimum level variance areas is that the individual response shapes remain recognizable. This will be the case in the series which follows. The first scenario shows a logarithmically spaced speaker series with log tapered levels. Log spacing in this case means that the distance from each sequential speaker is half the distance of the previous. It is also half the level (—6dB). Once we have reached past the unity crossover point we see an equal level contour angling up and away from the louder speakers. The angle remains constant because the spacing and level tapering are changing at the same rate. If the level taper rate and spacing are not complementary the contour will bend into a curved line. We saw earlier how we can create a straight contour line with a symmetrical uncoupled line source array (see Fig. 6.22). An alternative method to create a straight contour is shown in the second scenario of Fig. 6.25. The method involves staggering the origins of the array elements and tapering the levels accordingly. Once again we have log spacing and log level tapering but this time the spacing is forward. The levels all match at the unity crossover line, the vertical line. Because the origins are staggered the doubling distance for each source is different. As quickly as the speakers come together as

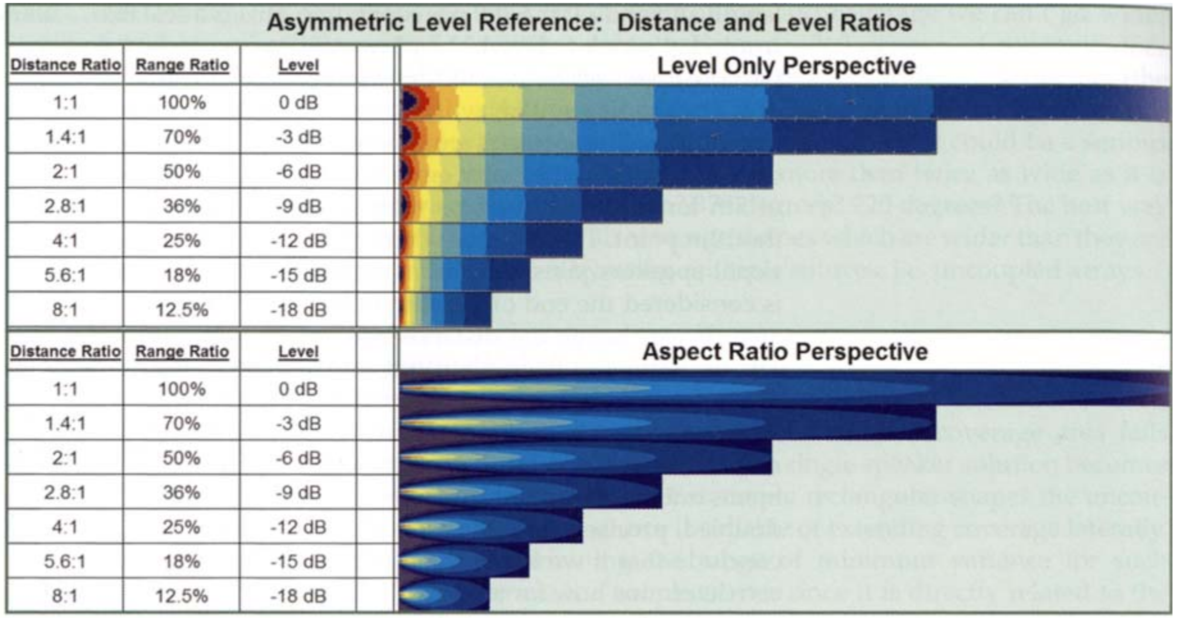


Figure 6.23 Level and distance ratios of sources with matched origins, including the aspect ratio scaling

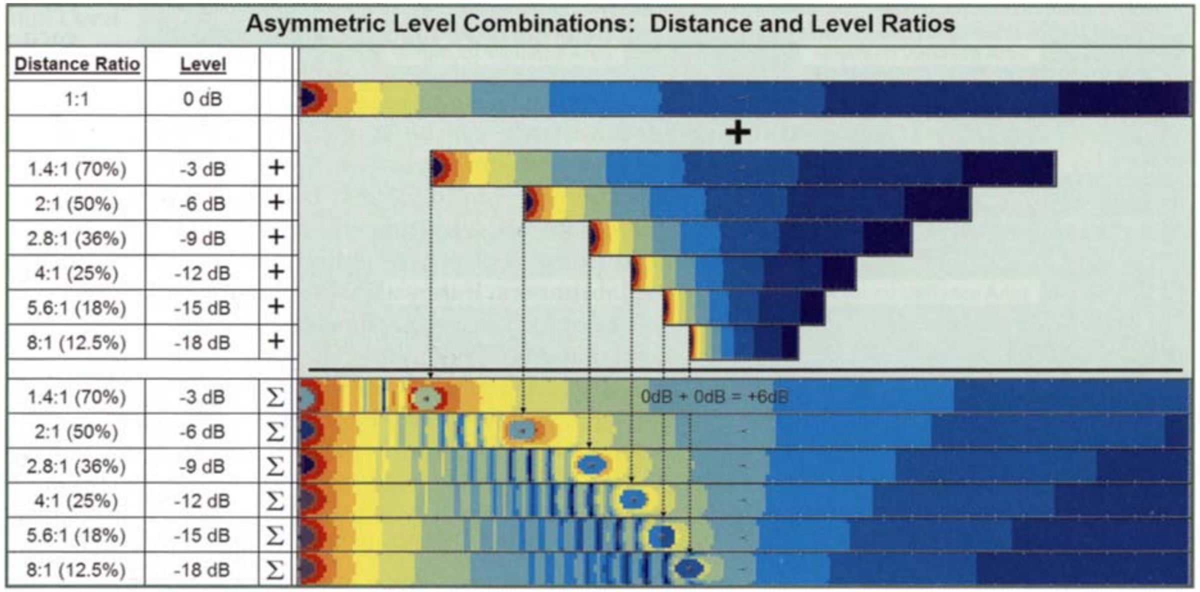


Figure 6.24 Level and distance ratios of sources with unmatched origins, including the summed levels

a straight line, the contour begins a tilt in favor of the long-distance speakers. Both of these scenarios, like all uncoupled arrays, can only hold together for a limited area. Prior to the optimal range the response contains gap areas. Beyond the prime area the longer-range speakers dominate the soundscape.

Uncoupled Point Source

We have seen that the uncoupled line source has a standard range of minimum variance that runs from half the aspect ratio length to the full length as shown in Fig. 6.26. Beyond this range the level variance remains low but the excessive ripple variance renders this response undesirable. The uncoupled point source adds angular isolation which has the potential to expand the usable range to a great distance. The angular separation moves the start of the minimum variance zone outward to some degree, but pushes the end of the zone out to a much greater degree. For this reason the uncoupled point source is often a superior choice to uncoupled line source. A reference chart for calculation of minimum variance start and stop points for the uncoupled line source and point source can be found in Chapter 7.

Uncoupled Point Destination

Minimum level variance has no greater challenge than the uncoupled point destination array. Two representative applications are shown in Fig. 6.27. The left panel shows the symmetric version, which is almost entirely incapable of minimum level variance due to high ripple variance from the combination of overlap and displacement. Because the angular offset reduces rather than increases isolation, the low variance range is the smallest of the symmetric uncoupled types. The majority of the areas where the combined result will provide a zone of constant level are those where the systems have not yet combined: the isolation zone. In short, this array cannot be used as an array, but rather as two individual responses with a usable area located between the individual elements and the train wreck overlap zone. This array type is recognized in practical applications as the infill array, often used as horizontal fill in the early central seating area.

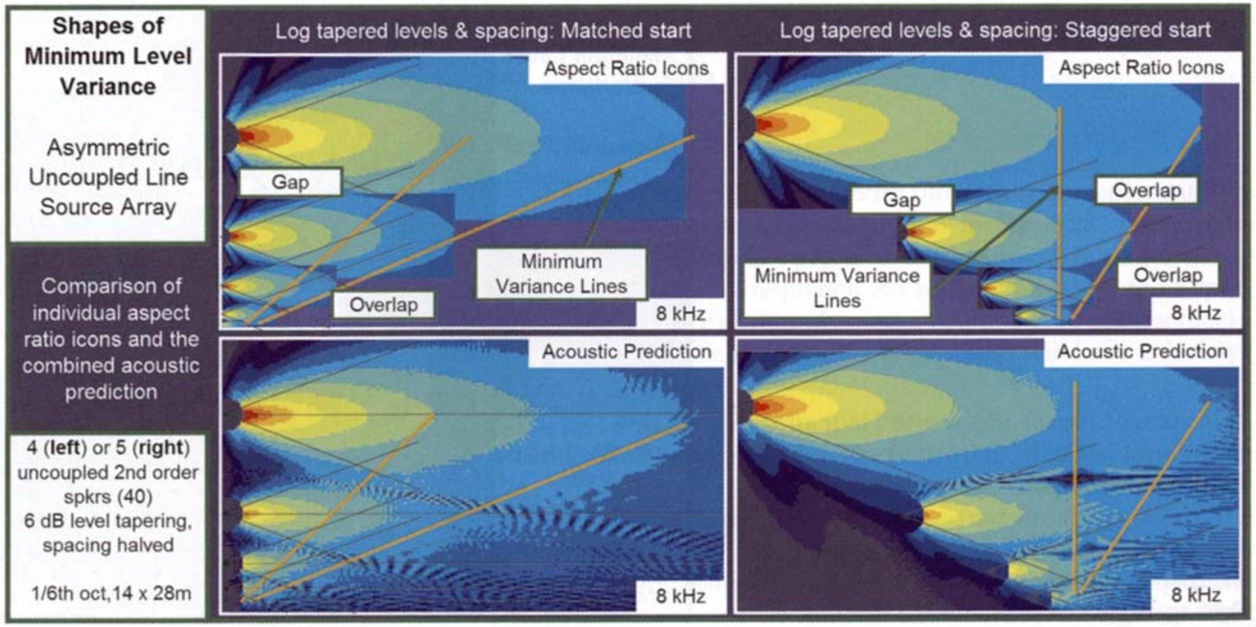


Figure 6.25 Minimum level variance shapes for the asymmetric uncoupled line source array. Left five-element second-order speaker array with 0 degree splay angle, matched line of origin, log lateral spacing and log tapered levels. Right four-element second-order speaker array with 0 degree splay angle, log staggered line of origin, log lateral spacing and log tapered levels

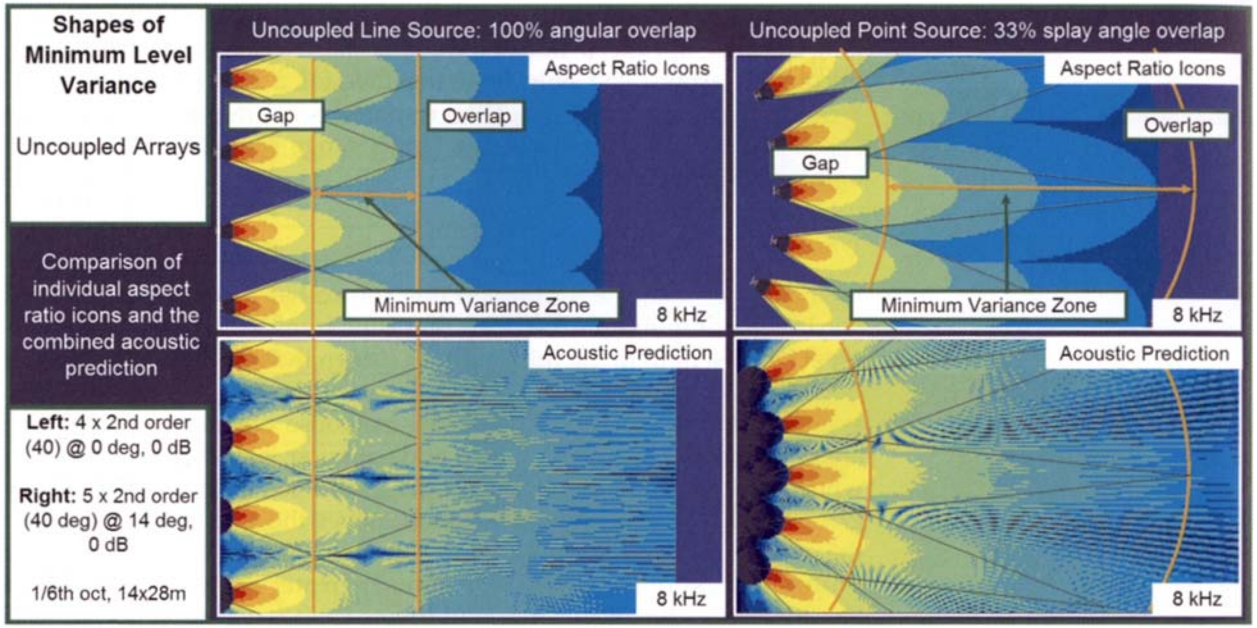


Figure 6.26 Minimum level variance shapes for the symmetric uncoupled line and point source array. Left four-element second-order speaker array with 0 degree splay angle, matched line of origin, constant lateral spacing and levels. Right a five-element second-order speaker array with 14 degree splay angle, constant lateral spacing and levels. Note the depth extension of the minimum variance zone

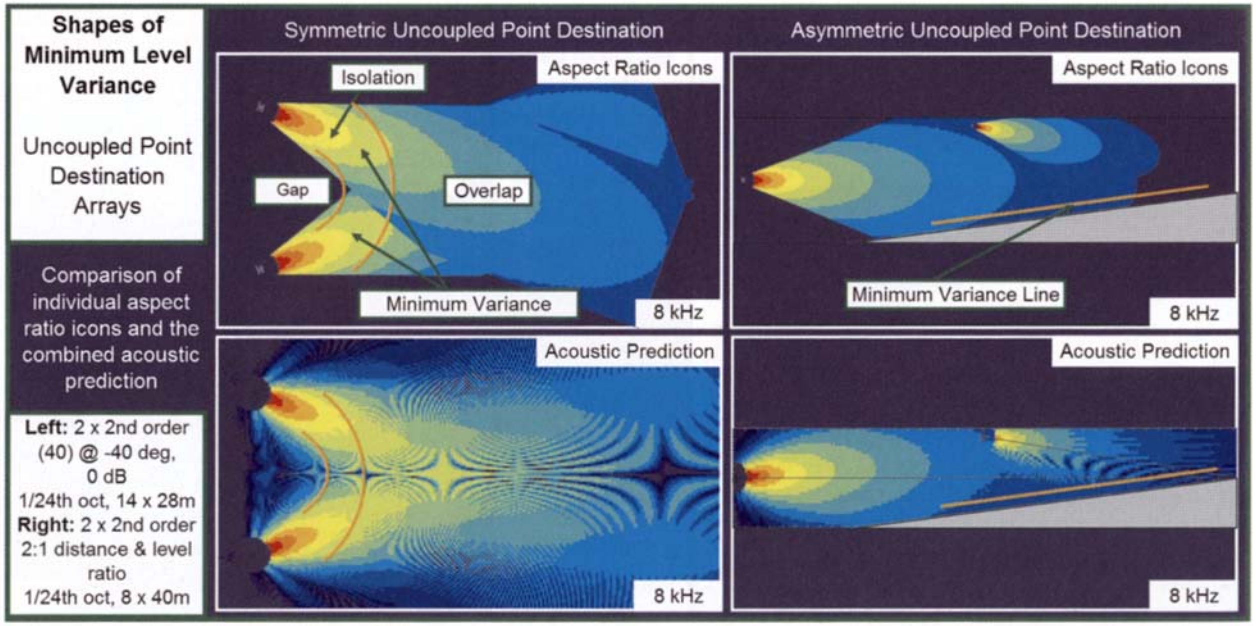


Figure 6.27 Minimum level variance shapes for the uncoupled point destination array. Left symmetric version. Right asymmetric version

The right panel shows the asymmetric version, recognizable as the vertical plane of the "delay speaker." Assuming the systems are synchronized, we can observe the forward range extension provided by the delay speaker. The combined levels reduce the rate of distance loss, thereby creating a line of equal level that is on-axis to the elements. This is the only means of providing level extension which does not rely on the offsetting effects of distance and axial loss rates. When used in combination with those effects, the loss rate can be reduced to the lowest amount practical. Such is the case shown here where the on-axis responses meet and increase the level 6 dB. The locations before the meeting point are both closer and more off-axis to both elements, thereby providing a level stalemate. The range of the line of equal level will depend upon the distance/level ratio as discussed previously with regard to the asymmetric uncoupled line source.

Minimum Spectral Variance

We now move into the third part of the variance equation: frequency. Up to this point we have seen how ripple and level variance are minimized over the space. The final piece of the puzzle will be extending the zones of minimum level variance over frequency. Since both of the previous two factors will change over frequency their effects will be closely interrelated to this third category. The complexity of our task has grown by a few thousand layers.

Relating Aspect Ratio, Beamwidth and Speaker Order

All renderings of coverage must qualify the frequency range. Aspect ratio is no exception. The beamwidth parameter, the display of coverage pattern vs. frequency, was introduced in the first chapter. It is now time to apply this. Two speaker models may be specified at 30 degrees. One speaker may maintain a constant 30 degrees over the range from 2kHz to 16kHz. The other may range from 60 degrees to 15 degrees through that range but "average out" to 30 degrees over the three octaves in question. These two 30 degree speakers will give very different results, both as single units and in arrays. The aspect ratio over frequency is derived from the beamwidth over frequency. This conversion gives us the shapes that the speaker creates over frequency. Since the room does not change its shape over frequency we have a measure of speaker conformity to the room. We might learn that a speaker is a perfect fit at 10kHz but is 10 times too wide at 1 kHz.

Arrays can be characterized as beam combinations. When we combine two elements the result can be one or all of three outcomes: beam concentration, beam spreading, or simply passing through each other. A typical array will have a combination of beam concentration and spreading effects over frequency. Beam concentration results from coupling zone and combining zone summation. Beam spreading requires some degree of isolation as will occur with combining zone and isolation zone summation. Both beam effects may occur simultaneously, with HF spreading and LF concentration being common. Beam pass-through occurs when the wavelengths are small and the pattern overlap is high. The beams are unable to either focus into a coherent forward beam (concentration) or join together in a radial or lateral extension (beam spreading). The pass through beams may create ripple variance fingers and lobes (see Figs 2.85 and 6.16) or simply pass by and go on to isolation.

Single Speakers

Let's observe the beamwidth plots for some representative single speaker elements shown in Fig. 6.28. These plots show the sloped nature of the beamwidth for all single speakers over frequency. As solitary elements their spectral variance is a simple matter of slope. As building blocks for arrays the shape of the beamwidth will have a decisive strategic effect.

There are two principal beamwidth shapes that are suitable for array building. The first is the "plateau" type which flattens out in the high frequencies. The plateau type can be seen here as the first- and second-order speakers. The flattened beamwidth area provides a clearly defined unity splay angle, and consistent aspect ratio shape. This qualifies these speakers as the best element for single-speaker

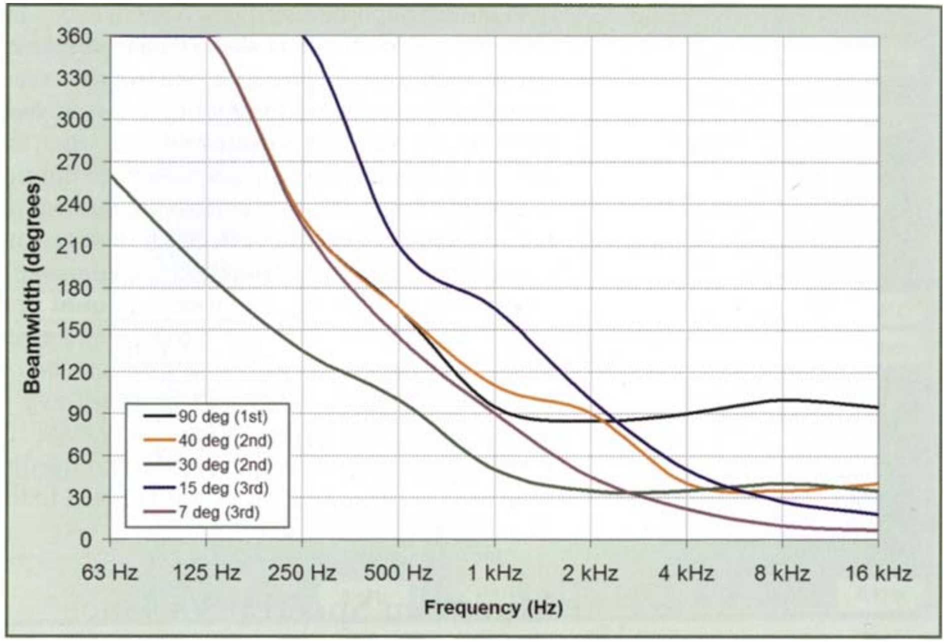


Figure 6.28 Beamwidth over frequency plots for the first-, second- and third-order single elements used throughout this section

applications and minimally overlapped arrays. The highly sloped third-order elements have no consistent unity splay angle, and therefore will only be used in highly overlapped coupled-array configurations.

The 90 degree first-order speaker is a small-format front-loaded enclosure and shows a steady narrowing from the low-frequency range until 1 kHz. From this point upward the coverage angle is constant within a few degrees. The unity splay angle for this speaker would be approximately 90 degrees, a designation that will hold true for the four top octaves. The behavior of an array constructed at the unity splay angle will be primarily characterized by beam spreading over the plateau range.

There are two second-order speaker types described here. The first is a 40 degree unit similar to the first-order system, but with a different HF horn. The beamwidth in this case does not flatten until 4 kHz, so the frequency range with unity spatial crossover capability is reduced to just two octaves. The beam concentration behavior will be a more dominant factor overall for arrays built with this element for two reasons: the reduced angular isolation (due to a smaller unity splay angle) and reduced frequency range of isolation will both increase the percentage of overlap. The 30 degree second-order system is a large-format horn-loaded enclosure and therefore is capable of expanding the frequency range of directional control. The unity splay angle of 30 degrees is held until 2 kHz (three octaves) and the slope of the beamwidth rise is less severe than the front-loaded systems. The result is that the transition into beam concentration will be more gradual, which will afford a larger range where the spreading and concentrating forces offset each other to create a constant combined beamwidth.

Let's take a moment to consider the beamwidth plateau, where it comes from and the nature of its limits. It seems logical at first glance to design a speaker system with a constant beamwidth over frequency. There are many trade names for this, built into product names such as "constant directivity" and "constant Q," to name a few. Such a speaker would allow us create radial arrays that are assembled like slices of pizza. If we could do so, there would be a single unity splay angle for all frequencies, no gaps and no overlap. This is fine in some circumstances, but not all. The lack of overlap means there is no power addition. Therefore if we need more power at any frequencies, we will need to reduce the splay angle, which will overlap all frequencies and introduce ripple variance.

There are physical realities that prevent practical speaker systems from achieving constant beamwidth over their full range. The 300:1 range in wavelength that must be managed by a full-range speaker (60-18 kHz) will be quite a challenge to control. In order to maintain constant beamwidth as frequency falls, we will need increasing horn size to manage the wavelengths.

The angular position of the plateau is also important. A long plateau can be managed at a wide angle, even by a small cabinet, as in the case of our 90 degree first-order cabinet. To maintain a plateau when starting from a narrow angle is orders of magnitude more difficult.

In simple terms, the frequency range of the beamwidth plateau relates directly to the enclosure size, particularly the depth. The two second-order systems shown here are very different sizes. The ability of the larger system to achieve a large beamwidth plateau even with a narrow 30 degree angle is all about size. A much deeper HF horn and horn-loaded LF driver extend the plateau to four octaves, equivalent to the first-order system.

The ultimate challenge would be to make a constant beamwidth over those four octaves at a very narrow angle; this can be done successfully with a parabolic reflector type speaker design. Such systems are large and impractical for all but a small range of applications.

The third-order systems lack the flattened beamwidth feature found in the other orders, or at the very least, have a severely reduced portion. The dominant feature is a steady narrowing of coverage. This precludes the unity splay angle from having anything more than a minute range of isolated beam-spreading behavior when employed in point source arrays. Overlap increases as frequency falls. The result is a gradual blending of the beam factors. Rising frequency tends toward beam spreading; falling frequency tends towards beam concentration.

The physical size of the typical third-order speaker is as small as possible. The driving force in the engineering of such systems is reduced wavelength displacement, and therefore minimal ripple variance. The compromise parameter is beamwidth slope, which cannot be maintained in the miniature cabinet geometry. The creation of a constant beamwidth will have to come from the combinations of multiple elements.

The keys to flattening the combined beamwidth (another way of saying minimum spectral variance) are splay angle and quantity. First-order systems will flatten with a minimum quantity of units — even a single unit is reasonably flattened. Third-order units must have substantial quantities in order for the highs to spread far enough to meet the concentrating lows.

Speaker order classification can also be applied to the aspect ratio where it reveals the differences in coverage pattern or "throw" over frequency. This gives us a picture of the shape that will be created over frequency for each type of speaker. The applicability of the speaker order as selection criteria becomes immediately apparent upon conversion to aspect ratio over frequency. The first-order speakers maintain a highly consistent shape. If the desired coverage was an aspect ratio of 1.4 the entire frequency range could be contained with only minimal overflow in the LF range. This is the quality that makes first-order systems the best choice as single speakers and for small arrays. Their ability to fill a simple symmetrical shape evenly over frequency makes the first-order speaker the leader in horizontal coverage applications. By contrast a coverage area that is appropriate for the HF range of a third-order speaker would find massive overflow at all frequencies below that. This factor precludes the single third-order speaker from most applications. The third-order speaker will thrive as quantities increase. The foremost application will be in the creation of asymmetrical arrays, which makes this type of speaker the leader in

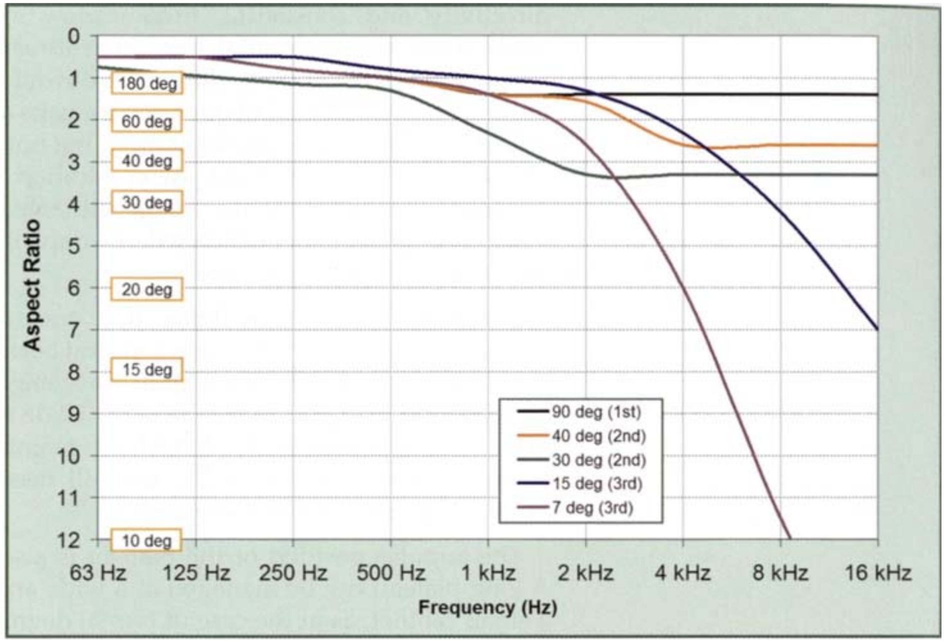


Figure 6 29 Aspect ratio over frequency plots for the first-, second- and third-order single elements used throughout this section

vertical applications. In the middle ground are the second-order systems, which have applications in both planes, but in limited quantities.

Now we will put this to work on our family of speaker arrays.

Coupled Line Source Array

The third-order speaker is most often chosen as the base element for coupled line sources. The futility of attempting to reduce the spectral variance of the third-order speaker in a coupled line source is shown in Fig. 6.30. Successive quantity doublings are shown and the narrowing of the combined beamwidth can be seen at all frequencies as quantity rises. This has all the aspects of a military arms race, with each increase in quantity pushing the entire response closer to 0 degrees coverage. This chart may give the impression of increased flattening of the beamwidth with quantity since the beamwidth of the sixteen-box version is visually more horizontal than the lower quantities. This must not be confused with beamwidth plateau response described earlier. This is an illusion due to the vertical scaling of the beamwidth plot. The coverage angle is *proportionally* equivalent over frequency in all cases, i.e. if the 1 kHz response is a factor of 12 wider than the 8 kHz response for a single box that relationship will be maintained at any quantity.

The illusion breaks down when we rescale the coverage of the same array into the aspect ratio over frequency as shown in Fig. 6.31. The aspect ratio scaling reveals the beam concentration of the parallel pyramid. As quantity increases the beam continuously narrows and the shape of the coverage extends forward towards infinity. The low-frequency range of the coupled line source will never get any closer to the shape of the mid-range and HF range, no matter what the quantity. A beamwidth change of 1 degree seems like a small matter, and this is reflected in the beamwidth vertical scaling. But a one degree change between 90 and 91 degrees is a very different matter than between 1 and 2 degrees! The aspect ratio scaling reveals this.

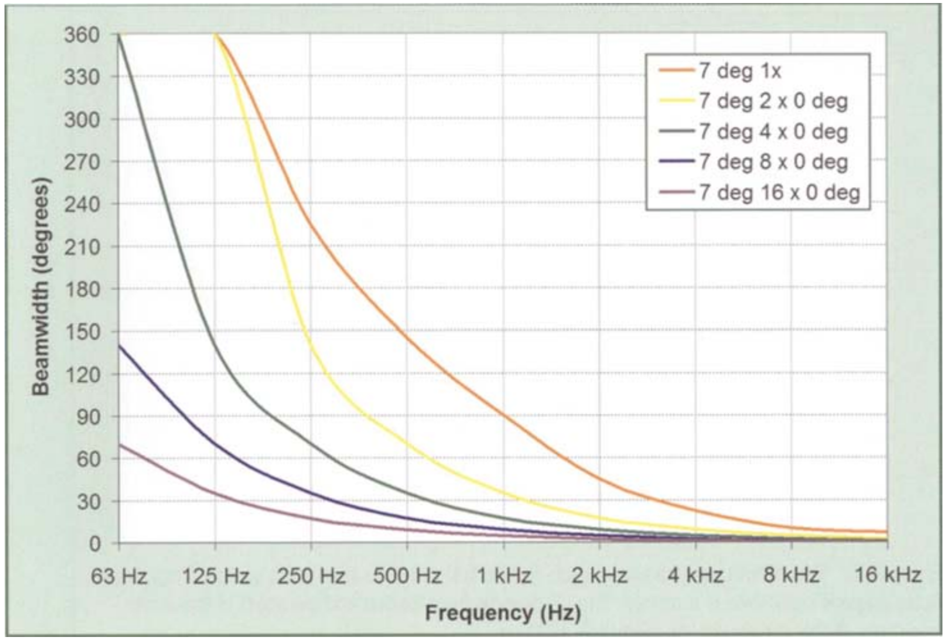


Figure 6.30 Beamwidth over frequency plots for the third-order coupled line source array

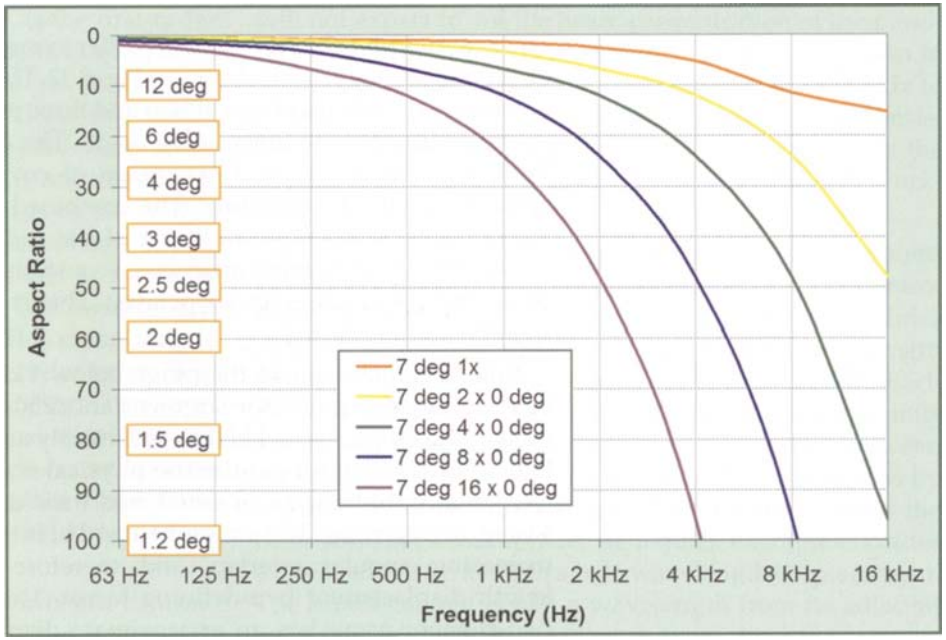


Figure 6.31 Aspect ratio over frequency plots for the third-order coupled line source array

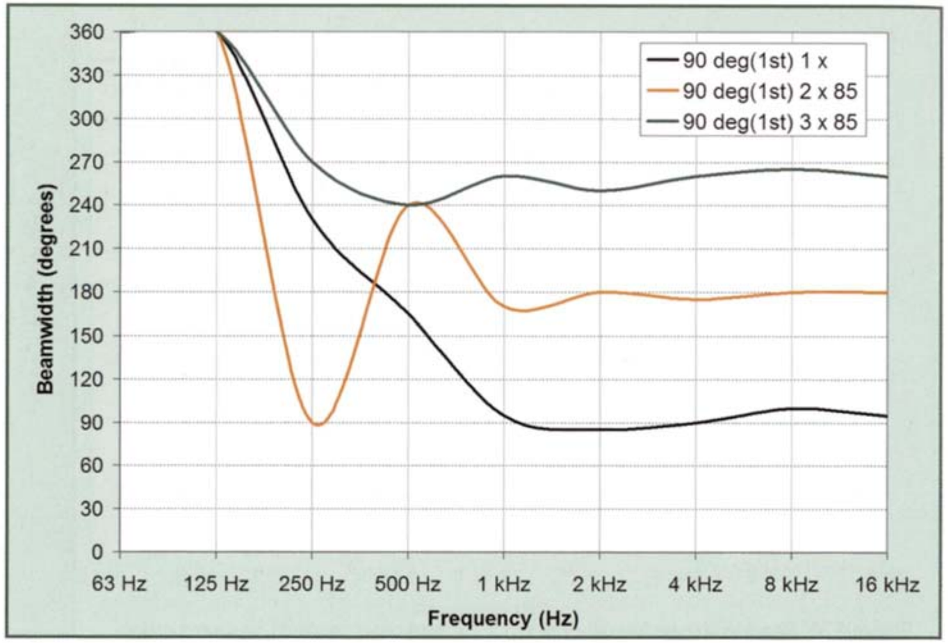


Figure 6.32 Beamwidth over frequency plots for the first-order coupled point source array. Splay angle is unity, level is matched. The LF driver is front-loaded and the spectral crossover frequency of the speaker is approximately 1200 Hz

Coupled Point Source Array

The beamwidth behavior of a first-order symmetrical coupled point source array is shown in Fig. 6.32. The single unit is compared with quantities of two and three units arrayed at approximately the unity splay angle. The range above 1 kHz exhibits the typical expansion of coverage angle achieved by beam spreading. The combined beamwidth above 1 kHz is a simple multiple of the individual elements. The three-element array has a constant beamwidth of around 260 degrees for six octaves. This is a minimum spectral variance array.

Note: The behavior of the range below 1 kHz merits a brief explanation due to the large-variance 250 Hz to 500 Hz range. The range above 1 kHz is sufficiently angularly isolated that its behavior requires no physical scaling, i.e. the elements could be large or small, and their displacement scaled accordingly. The range below 1 kHz has steadily increasing angular overlap, and therefore the wavelength displacement is a defining factor. This particular two-element array has an approximate displacement of one wave-length, X (0.69m) at 500Hz (0.5X at 250Hz). The type of beam-width variance seen here will occur as a matter of course as the frequency falls under 1 lambda displacement, with the narrowest point being reached at 0.5 lambda. (See the discussions of wavelength displacement in Chapter 2 and Figs 2.24 and 2.25.) The scale of the speaker elements will affect the frequency range, with larger displacements moving the affected range upward. As more elements are added the interaction effects take on a more complex nature since every element has wavelength displacement relationships to every other element. The interactions below the isolated frequency range should be viewed as trends, with the understanding that the family of effects will be found in different frequency ranges on different models. In short, we can say with some certainty that two coupled first-order speakers splayed at 90° will become narrow (at 0.5X) but we cannot say that this will necessarily occur at 250 Hz.

The same principles are applied for the second-order system shown in Fig. 6.33. The physical size and driver

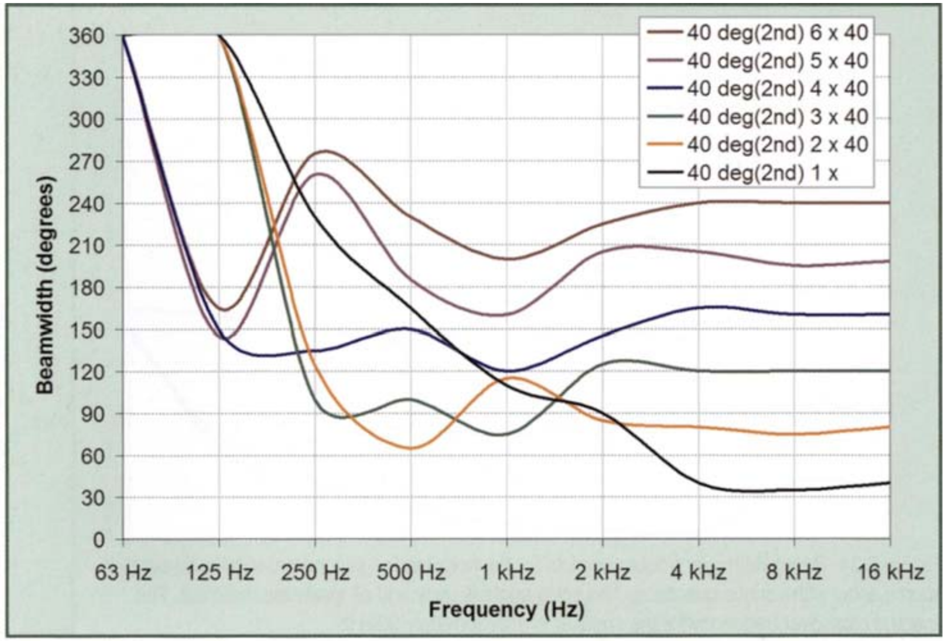


Figure 6.33 Beamwidth over frequency plots for the second-order coupled point source array. Splay angle is unity, level is matched. The LF driver is front-loaded and the spectral crossover frequency of the speaker is approximately 1200 Hz

complement of the elements is the same as the first-order units described in Fig. 6.32, so the disparities can be attributed to HF horn differences and splay angle. In this case the beam spreading range is limited to 4 kHz and above. In the isolated range the beam spreads by the unity splay angle (in this case 40 degrees) with each additional element. There are several noteworthy trends in the ranges below 4 kHz where the quantity of elements has a strong effect on the beamwidth variance. The first is found in the two-element scenario, where a large beamwidth variance occurs between 500 Hz and 1 kHz. This characteristic signature matches the two-element first-order array (Fig. 6.32) except that it appears one octave higher. The rise in frequency is due to the reduced physical displacement which results from the smaller splay angle (same spacing at the rear, but closer at the front). Notice that this volatile range shows substantial variance from the spread beam area for all quantities of elements.

The next trend is one toward coherent beam concentration. Since this is not a coupled line source array we will not expect to see the beam concentration of the parallel pyramid in full force, but rather a reduced version in the ranges of maximum overlap. A survey of the 63 Hz to 125 Hz range shows a combined response that is consistently narrower than the single-element response. As the quantity rises the beam narrows, the unmistakable workings of beam concentration.

The range between 125 Hz and 4 kHz is in a battleground between two competing and offsetting factors: angular isolation and wavelength displacement. The physical realities of multiple element clusters include layered wavelength displacements as the number of elements increases. Each element is displaced the same amount from adjacent units but the interaction does not stop there. It continues between separated units as well, although the effect is reduced by angular isolation. This layering of interactions creates the complex weave found here. As a quick example consider a five-element array with a half-wavelength displacement. The center cabinet is half a wavelength from its adjacent neighbors, but a full wavelength from the outside units.

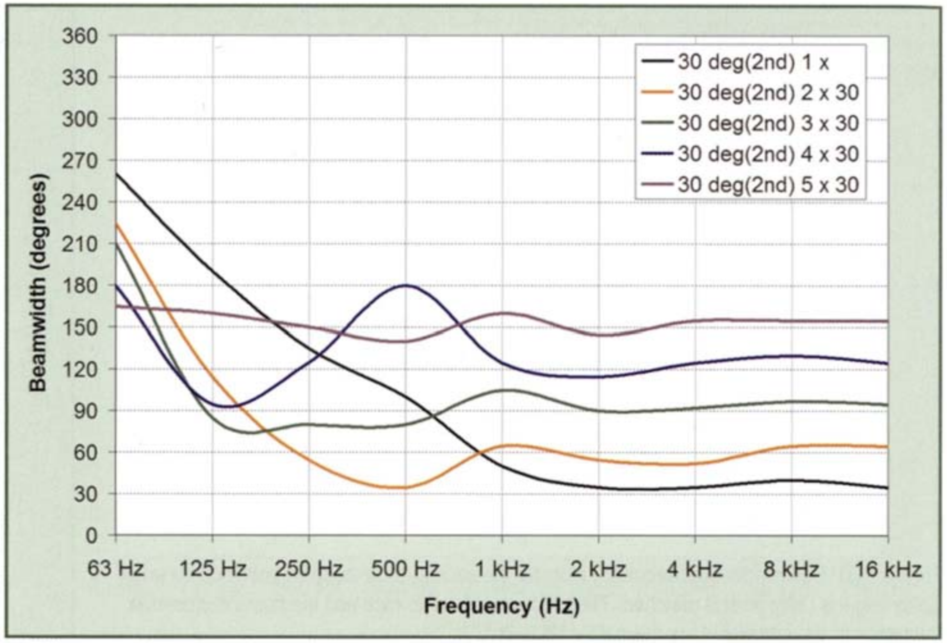


Figure 6.34 Beamwidth over frequency plots for the horn-loaded second-order coupled point source array with various quantities. The splay angle is unity and all levels are matched. The spectral crossover frequency for the speaker is approximately 900 Hz

The outermost units are two full wavelengths apart. If the elements have a wide pattern in the given frequency range there is going to be a complex weave of interactions.

The 250 Hz to 500 Hz range provides a glimpse into the competing factors. At quantities of two to four elements the combined beamwidth is narrower than a single unit. Beam concentration is the dominant factor here. With five and six units the angular spread is so wide (200 to 240 degrees) that the outermost elements have achieved sufficient isolation from each other for beam spreading to take place in addition to the ongoing beam concentration of the central elements. The result is a combined response that is wider than the single element.

An alternative second-order system is shown in Fig. 6.34. The elements are large-format horn-loaded systems with 30 degree nominal pattern above 2 kHz. The rate of low-frequency beamwidth expansion is noticeably lower than the previous examples, due to the horn loading of the low-frequency range. The increased low-frequency control provides a higher degree of isolation than we had with the front-loaded second-order speaker in spite of the reduced splay angle. The results are a noticeably more uniform series of beamwidth responses over quantity. As quantity increases the beam-spreading and beam-concentration forces push the response in opposite directions. The HF response is continually widened, while the LF response narrows, resulting in a more consistent combined beamwidth than the single unit from which it comprises. The mid-range area sees both concentration (quantities <3) and spreading (quantities >3). The five-element array is notable as an example of the minimum spectral variance. The offsetting array factors have met in the 160 degree range over the entire frequency range of the device.

Next we will examine the effect of angular overlap for a fixed quantity of elements. In this case (Fig. 6.35) we will reuse the five-element horn-loaded second-order array just featured. The effects of increasing angular overlap are compared to the response using the unity splay angle of 30 degrees. A splay angle of 22.5 degrees gives us 75 per cent isolation and 25 per cent overlap. The combined result

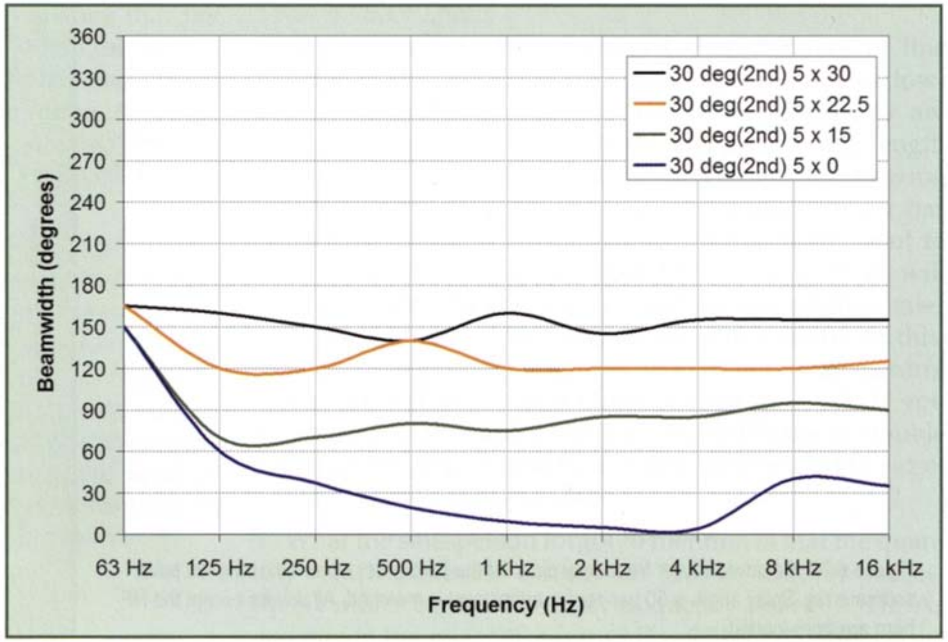


Figure 6.35 Beamwidth over frequency plots for the horn-loaded second-order coupled point source array at various splay angles. All levels are matched

approximates a 25 per cent reduction in beamwidth over the full range. The accompanying effects of this overlap cannot be seen in the beamwidth, but include increased ripple variance and, on the positive side, increased power addition. When the overlap is increased to 50 per cent (15 degrees splay) an undesirable effect arises. The reduced angular isolation leaves the multiple unit beam concentration in the low mids without the offsetting effects of beam spreading. The result is excessive beam concentration (the parallel pyramid), which creates a low/mid response that is substantially *narrower* than the HF response. In the outermost 10 degrees (per side) of the coverage there will be full extension of the HF response but the mid-range and low mids will be reduced by more than 6 dB. This creates one of our most undesirable and easily detectible tonal responses: the telephone. As a general rule it is critical that the beamwidth of an array (or a single element) should never be narrower in the mids and lows than the HF range. In addition to the telephone quality of this array we have increased ripple variance. But at least it's really loud! We now have an array capable of making us want to leave the room.

Finally we reach the 100 per cent overlap point that comes with a 0 degree splay angle. This, of course, has converted our point source array to a coupled line source. The beamwidth can be seen to collapse inward, as we would expect from the parallel pyramid, and this continues until we reach the point where beam concentration breaks down. The breakdown is due to the wavelength displacement being so high that the individual responses pass through each other without forming a single beam. For this frequency range the array behavior can no longer be classified as coupled and therefore the beam model breaks down. The result is that the combined beamwidth returns to that of an individual element, albeit with massive ripple variance. The beamwidth changes by a 10:1 factor in the octave between 4 and 8 kHz, which removes all hope of minimum variance performance for this array. It is noteworthy that this type of approach was the most common array type for concert sound before the advent of

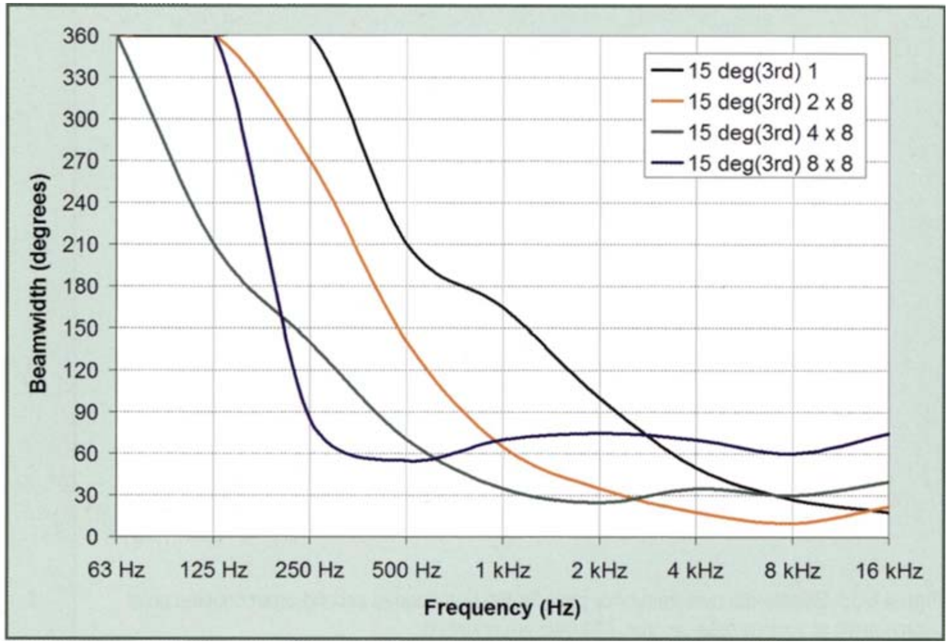


Figure 6.38 Beamwidth over frequency plots for the third-order symmetric coupled point source array. Splay angle is 50 per cent overlap, level is matched. All drivers except the HF horn are front-loaded

third-order speakers in the 1990s. Power, yes. Minimum variance? No.

We will move next to third-order speakers. Naturally the splay angle is reduced to a small fraction of those used previously for first- and second-order systems. In this case the frequency range where the unity splay angle holds is the lowest ever: practically none. The response of a single element is compared with successive doublings up to eight units. There are several important trends. We begin with the beam-spreading behavior at 8 kHz. Notice that the response of a single unit (15 degrees) is wider than that of two (8 degrees), and equal to that of four. This is consistent with the fact that the chosen angle (the manufacturer's maximum splay with the standard rigging frames) is less than the unity splay angle. As the quantity increases the outermost units overlap less with each other and the beam begins to spread. As frequency falls the overlap becomes stronger and beam concentration becomes increasingly dominant. The result is a gradual flattening of the combined beamwidth, in sharp contrast to the steady downward slope of the single element. This is the key design feature of the third-order speaker. A point source array comprising steadily narrowing elements will cause steadily offsetting effects among the beam behaviors. For a given splay angle, as the individual beams narrow (frequency rising) the combined beam spreads due to isolation. As the individual beams widen (frequency falling) the combined beam concentrates due to overlap. The result is a push-pull effect that flattens the beamwidth over larger frequency ranges as the quantity increases. Observation of the response over quantity reveals an approximate doubling of the flattened area with each doubling of quantity. Two units yields flattening down to 2 kHz, four units extend it down to 1 kHz and eight units flatten the beamwidth all the way to 500 Hz.

Speaker Array Methods

Our goal is to find the speaker array methods that will lead to minimum spectral variance. We are looking to create

Perspectives: If I could get a flat frequency and phase response at every seat in the house, I could just lay down and die, having achieved an entire career's quest.

Fred Gilpin

a series of equal level contours over frequency that lay over each other like floors of a building. When this occurs we have achieved our goal. If the response is spectrally tilted, and it will be, it will be tilted everywhere. As stated before, this is "to dream the impossible dream," but we will see that some approaches will get us a whole lot closer than others.

Our investigation will once again use the familiar pattern of isolating the principal parameters and viewing their contribution to the overall response by a series of steps such as successive doublings. It is neither necessary nor practical to analyze every frequency band in succession. Each scenario is presented as a composite of four frequency ranges, ranging from 125 Hz to 8 kHz in two octave increments. This "four-way" approach provides a sufficient level of detail to clearly identify the trends, and helps this book to be somewhat shorter than Tolstoy's *War and Peace.*

Coupled Arrays

There are four coupled speaker array types that will be analyzed for their spectral variance behavior. They include the line source, the symmetric point source, the asymmetric point source and the hybrid combination of the line and point source, also known as the "J" array. All four of the arrays have beam concentration and beam-spreading properties, and all four differ in how these are applied. Two of these array types will pass the test of minimum variance capability, and become candidates for optimized design.

Coupled Line Source

The coupled line source contains an unlimited quantity of speakers at the most limited number of angles: 1. The defining factors are the individual elements, the spacing and the overall line length. We will isolate each of these factors and observe their independent contributions to the overall response character. The coupled line source has three variables that will be investigated for their effects upon spectral variance: speaker order, quantity and level asymmetry.

Speaker Order Effects

One of the most often discussed characteristics of line sources is the length of the array. The amount of low-frequency directional control and power capability are closely linked to the length of the array. If the line length is less than half of the radiated wavelength the steering effects are minor. Once the half-wavelength barrier has been crossed the steering becomes strong. If we want to extend the low-frequency control by an octave, we will need to double the line length. It did not take speaker sales personnel long to comprehend the implications of this. What self-respecting engineer would not want control down through the lowest range of their system? "If you want that octave under control, you will have to double the quantity. Sure would be a shame to have 125 Hz out of control. Thanks for your order."

What the salesperson forgot to mention is that the quantity which achieves "control" at 125 Hz achieves successively more "control" at every frequency above 125 Hz. As we saw in the quantity races of Fig. 6.31 we do indeed get more control as quantity *and* line length increase. We also saw that the amount of control at 125 Hz was no closer to the amount of control at 500 Hz when we extended the line. We simply narrowed both ranges.

The line length sets the limit of LF steering, but that is about it. A 1/2 lambda line made up of four elements will have an almost identical coverage angle to one made of eight of the same elements (at twice the maximum dB SPL). As soon as the wavelengths are smaller than the line length the picture changes entirely. The much more dominant performance factors are the displacement between devices *within* the line. As the wavelength decreases the distinction between four elements and eight elements changes from a non-factor to the dominant factor. As line length increases, we are extending the foundation of the parallel pyramid. The quantity of elements will define the stacking height of the pyramid. Once the displacement between the elements reaches 1 lambda, the parallel pyramid is in total control, and line length merely sets the starting width of the stack.

Let's now examine the line length independently and in so doing also learn about the effects of speaker order

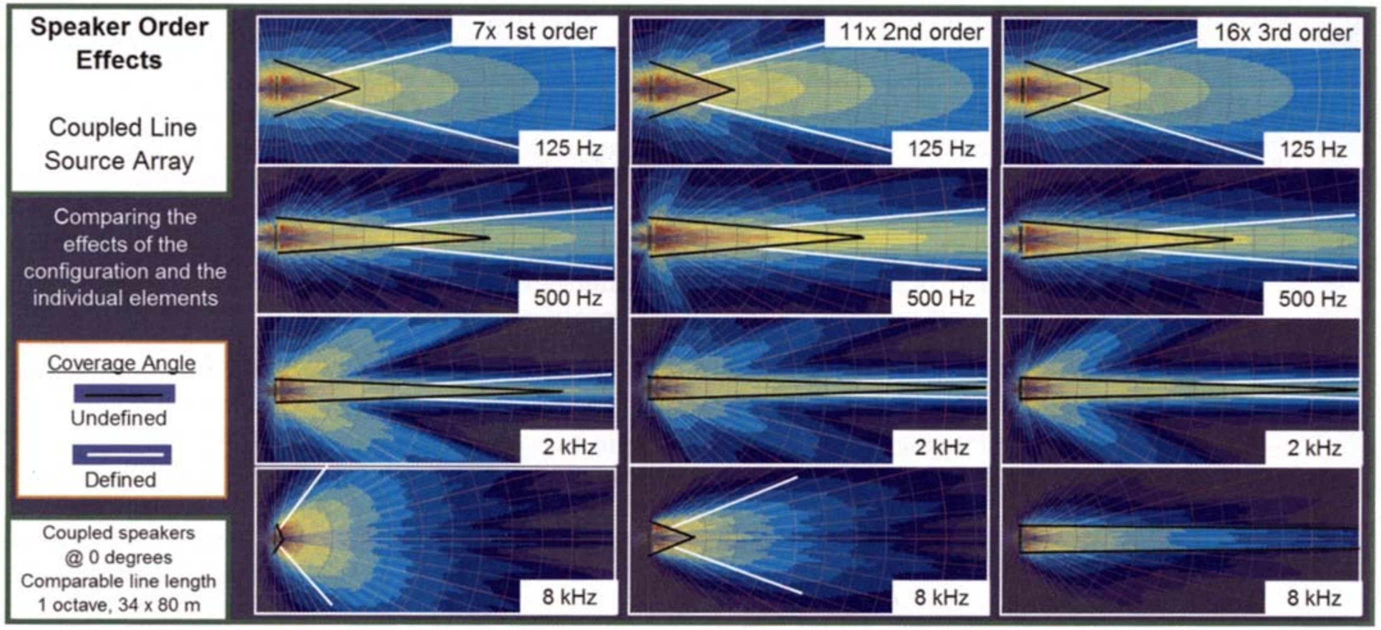


Figure 6.37 Speaker order effects over frequency for the coupled line source array

in the coupled line source. Figure 6.37 contains three scenarios for comparison. All contain the approximately identical line length, although they are made of different elements and quantities. A smaller number of first-order systems are compared with successively larger numbers of second- and third-order elements. The results show nearly perfect consistency in the 125 Hz and 500 Hz ranges. Clearly the line length, not the nature of the individual elements, is the dominant force here. The angular overlap of the elements is so prevalent that the pyramid construction process progresses at the same rate, since the overall space at the pyramid floor (the line length) is matched. As we move into the 2 kHz range, we note that the forward beam shape remains matched but the level of the side lobe wings is highly variable. The first-order speaker in this case has the widest response at 2 kHz and as a result shows the strongest side lobes escaping the forward beam control. The first-order system is showing the symptoms of beam pass-through. As we reach 8 kHz we can see the systems separate out into distinct responses. The third-order system continues the well-known pattern of the parallel pyramid. In this case line length is still important since it will influence the distance traveled before the assembly is complete. The first-order speaker HF response is not dominated by line length. The tiny high-frequency wavelengths with their 90 degree patterns pass through each other as uncoupled sources. The combined response is a tattered and torn version of 90 degree coverage with gross ripple variance levels due to deep combing. (The extent of the variance is smoothed over by the octave resolution of this plot.) The second-order system behaves in a similar fashion, albeit restricted to a 40 degree combat zone.

We can conclude the following. Line length is the dominant factor in low-frequency beam concentration in the coupled line source. The beam concentration can only continue into the HF range if the beamwidth decreases in proportion to the wavelength displacement of the elements. Only the third-order systems with a constantly narrowing beamwidth slope meet these criteria.

Quantity Effects

The next factor to investigate is quantity, beginning with two units and doubling in four steps to 16. The results are shown in Fig. 6.38. The following trends emerge.

1. The coverage pattern narrows as quantity rises. The coverage pattern is consistently reduced by 50 per cent with each quantity doubling.
2. The ratio of spectral variance is unchanged by the quantity. All frequencies are narrowed by the quantity increase. The difference between the four frequency ranges persists. There is no quantity where the narrowing of the lows will catch up to the narrowing of the higher ranges.
3. The beam concentration (the parallel pyramid) is the primary driving force in the interactions.
4. The larger quantities give the appearance of reduced spectral variance if the range of viewing does not extend far enough for coupling to occur over the full frequency range. This results from the low- and low/mid-frequency ranges having run the full course of the pyramid while the high-mid and high-frequency ranges have not yet reached the summit.

Let's look at how this stacks up on our scorecard. If we have achieved minimum spectral variance we will see the coverage pattern shapes lie on top of each other as frequency rises. This is clearly not the case. As frequency rises the coverage narrows. There are two candidates for lines of minimum spectral variance. We can traverse a unity crossover line along the width (height) of the array. This line is present in all of the arrays, at all frequencies. The distance of the line from the source varies with frequency, and the width (height) varies with quantity. As we increase our distance, this line steadily shrinks away until we reach the pyramid summit. The largest quantity at the highest frequency will be the last to reach the top of the pyramid. Once we have reached the top, the behavior of the array takes on the character of a single element. This is the characteristic of the coupled line source that is its major selling point. The fact that this transition into a coverage angle occurs at a different distance as frequency rises is not,

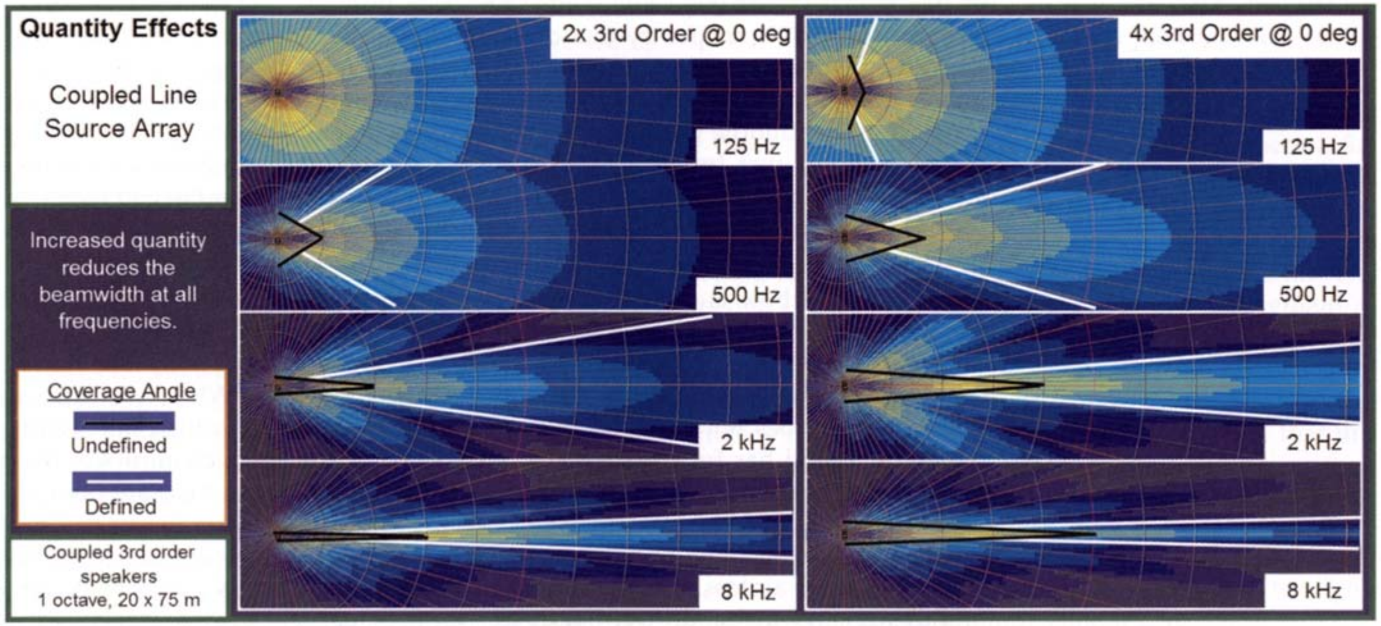


Figure 6 38(a) Quantity effects over frequency for the coupled line source array

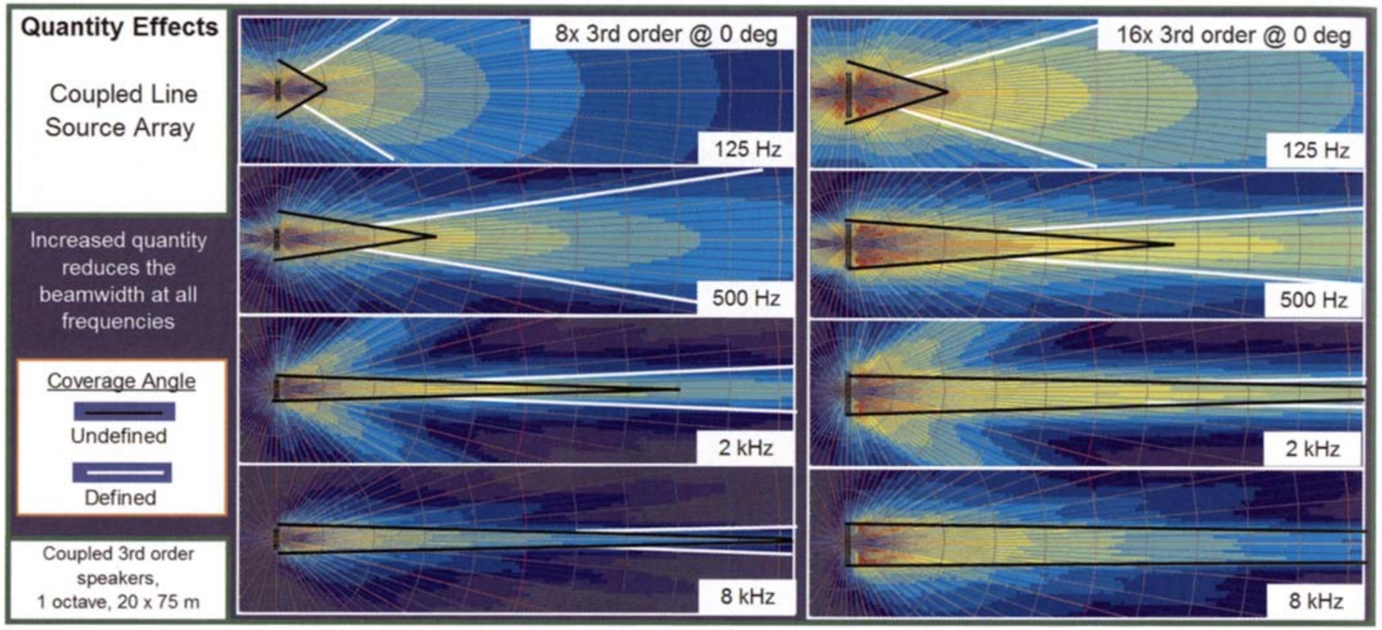


Figure 6.38(b)

however, widely understood or acknowledged. The result of this is that the coupled line source proceeds through a morphology over frequency that can extend over huge distances. At the beginning of the process we have a flat line of minimum variance, a "wall of sound" extending forward. In the middle ground the system is in a state of metamorphosis over frequency and distance. At the far end we have a defined coverage angle, which means the second type of minimum variance line can now be found: the coverage bow.

The coverage bow is pulled tighter as frequency rises, and therefore the off-axis near point moves continually. The result is that the paths which connect our points of equality continually change over frequency. The high-frequency range does not reach its full extension until the parallel pyramid has fully assembled. Since 8kHz is the most directional of the ranges shown, it takes the longest distance to top out. When the top is reached the response is flat at the pyramid summit if we temporarily neglect (or equalize) the HF air loss. It is lonely at the top, and we will find there is no way but down from there.

Let's follow three paths: directly back toward the source, radially moving off-axis and the coverage bow path of minimum variance. Any movement closer to the source will begin an HF decline. The path back toward the center of the line source is a steady pink shifting of the response. The reason is our movement inside the pyramid causes us to fall off-axis to some of the elements. As we move closer we move into the gap crossover areas of the HF elements and their contribution to the summed response is reduced. Lower frequencies, due to their wider patterns, are still fully overlapped and their addition is at the maximum. The closer we get, a progressively fewer number of HF elements remain in overlap mode and the result is steady pink shifting.

When we add air absorption effects to this equation the pink shifting provides a partial offsetting effect which improves the situation. The top of the pyramid, the most distant point, has the most HF air loss. As we move closer the air loss decreases at the same time that the axial loss increases (as just described). How closely these factors compensate for each other will be application- and weather-dependent.

Now let's get back to the top and move radially. We won't have to go far. The -6 dB point will be first reached at the highest frequency. All others will have less than -6 dB at this point and we will have pink-shifted as expected. Additional movement will cause successively lower frequencies to fall under the 6 dB line. It is important to note that we can not quantify the coverage pattern as an angle until the pyramid has run its full course. The coverage of an unassembled pyramid can only be expressed in terms of area, not angle, since it will not hold an angle over distance until all of the contributing elements have reached unity or overlap crossover status. The spread HF response of an unassembled pyramid is often confused with angular spreading and leads to the belief that the line source can create a uniform frequency spread when given a sufficient number of elements. This can be seen in Fig. 6.38 where the sixteen-unit array appears to have less variance over frequency than the lower quantities. Such an effect can be heightened if we restrict our viewing to lesser distances. The continued variance over frequency becomes immediately apparent when we extend the vista until the pyramid has fully assembled in all cases. It is justifiable to question why it matters whether we have a consistent angular coverage. If we have created the shape we want, does it matter if the beams will not fully sum until they have bored holes through our walls and gone across the street? The reason it matters is that when we are inside the parallel pyramid we are in the acute angle zone of ripple summation. We are in the area of highest rate of change of position; i.e. highest ripple variance. Smoothed prediction plots will gloss over this critical factor. Minimum variance design will require success in all three of our primary categories: level, frequency response and ripple.

We have one more path yet to follow. This would be our coverage bow path. To begin that journey we will have to travel outward to double the distance from the source to the pyramid peak. Remember that it is at the peak that our line source becomes "a single speaker." We cannot expect inverse square law behavior until the speaker is unified. Now we can find the coverage bow line of minimum variance by the standard method: double distance on-axis to single distance (the peak distance) off-axis. What we will find are coverage angles that relate to the original element, divided by the quantity. These effects are shown in Fig. 6.39.

Before moving on let's take a moment to consider what type of element would be required to create a minimum variance coupled line source array. We have seen that the continuous sloping beamwidth type (third-order) will not pass the test. We have also seen that arrays of wide angle plateau beamwidth units (see Fig. 6.37) will fall apart with ripple variance in the HF range. The displacement is too large and the angular overlap too high. We will need to minimize the displacement and stabilize the angular overlap. The element will ideally be as small as physically possible and stacked for minimal displacement.

There are two directions we can go. We can seek the "wall of sound" line of minimum variance, or the coverage bow. Both would require a constant beamwidth over frequency. The coverage bow will be the same shape over angle and over distance for all frequencies if all the beamwidth is constant. What angle would we choose? If the angle is wide we will be scattered by ripple variance. If the angle is narrow we will need a very large enclosure to control the mid-range and lower frequencies. This physical reality will likely create too much displacement for effective coupling. This is not terribly promising.

Now let's try to create a "line of sound," where a flat plane of even level radiates from the line of speakers. What would the angle be this? Not applicable. The coverage would have to be defined as a width, and that width would be equal to the displacement between the line elements. The pattern would move forward and remain as a coverage width, never overlapping into the parallel pyramid. Let's visualize the analogous response of a spectral crossover with infinite dB/octave divisions of the frequency response. Neglect for a moment the fact that this is not possible in acoustic propagation at any frequency, much less over a wide band. Consider that such a scheme would only spread the sound over the line, offering no acoustic summation, and therefore no power addition. To achieve power addition we are going to need to live with overlap. We will need to find a better way.

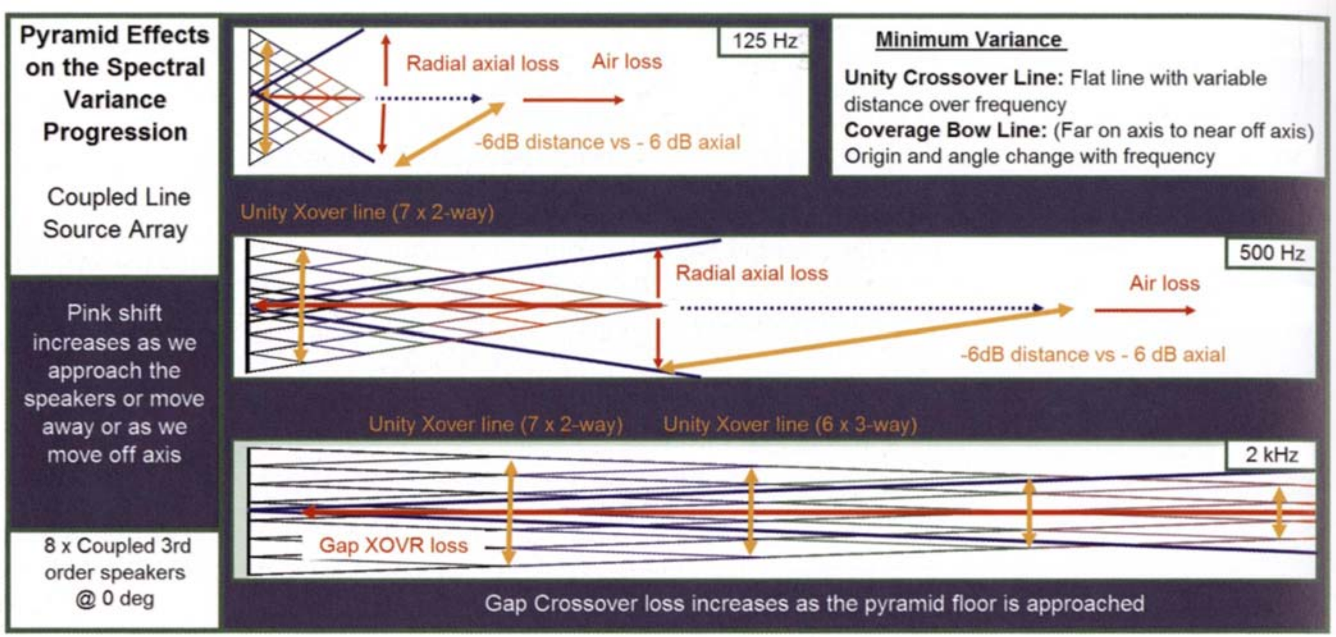


Figure 6.39 Spectral variance progressions in the coupled line source

There are many "line array" systems in use that claim in various manners to merge the behavior of a continuous line of elements into a single device. These systems are able to create a line of minimum variance for a limited frequency range for a limited distance. The listener and the system optimization engineer will likely find themselves in the metamorphic zone of transition between dimensional and angularly defined radiation. The listener doesn't care, but to the optimization engineer the distinction is relevant. Optimization strategies require a definable and stable coverage pattern. A system in transition can only be tuned for a point in space. These systems can made loud and powerful. The modern "line of sound" is a huge improvement from the old style "wall of sound" since the interference is limited in one plane and greatly reduced in the other. But we need not settle so quickly for power over minimum variance. The modern third-order speaker can still fulfill that promise, as we will soon see.

We will now set our course instead on finding minimum variance arrays using the types of speaker we know to exist: low-order speakers of the plateau beamwidth type and high-order speakers of the continuous slope type.

Asymmetric Level Effects

There is only one option for asymmetry in the coupled line source array: variable level. The elements charged with the closest coverage can be level-tapered to compensate for their proximity. This is, of course, a misnomer. Since all speakers have the same orientation there is no separation into near and far coverage. The results of such level tapering are, not surprisingly, ineffective at achieving any progress toward our low-variance goals. The redundant orientation of the elements dictates the coverage will overlap and the level reductions simply skew the geometry of the pyramid summation. In the gap crossover regions the level tapering causes an asymmetrical redirection of the response in favor of the louder elements. The response has the appearance of the minimum variance shape in the HF response for a limited distance.

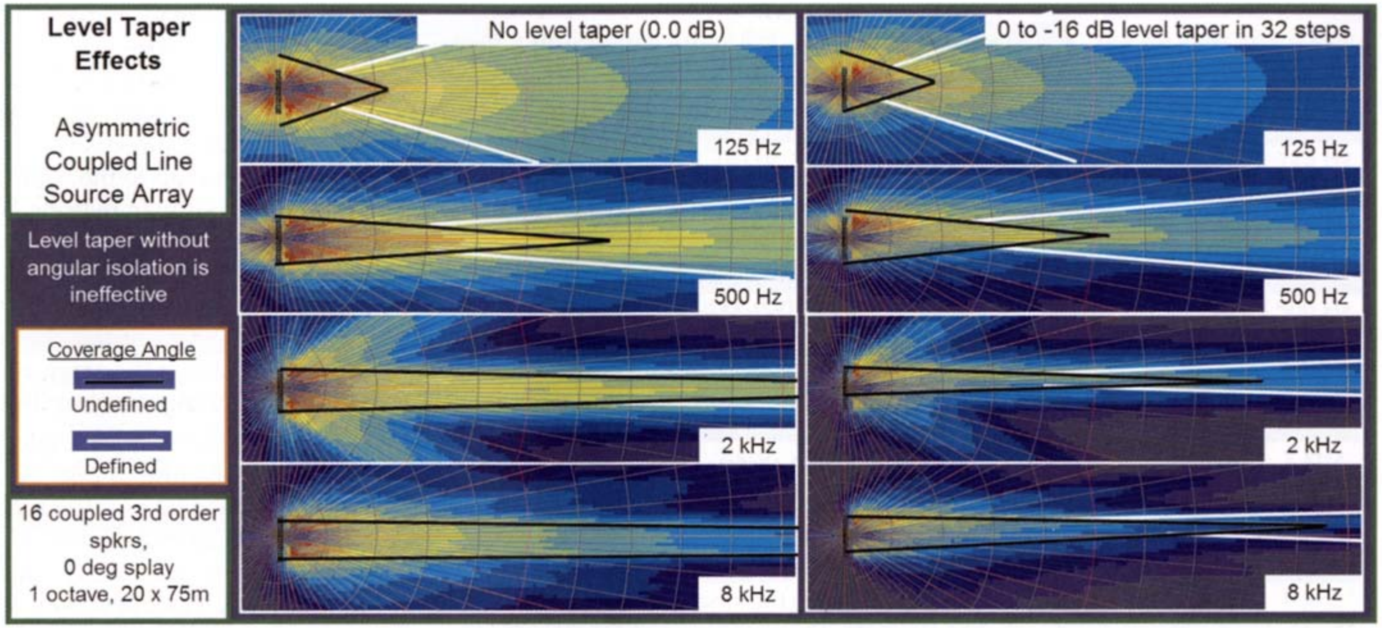


Figure 6.40 Asymmetric level effects over frequency for the coupled line source array

After all of the crossovers have overlapped the response resumes the beam concentration behavior we have found in all line source arrays. The level tapering effectively reduces the number of elements contributing to the pyramid and as result the behavior mimics an array of shorter length (fewer elements). This trend can be seen clearly when the degree of symmetry increases as frequency falls. An important trend worth noting is the position of the pyramid peak over frequency. Notice the upward movement of the on-axis point as frequency rises. This is due to the increase in isolation as frequency rises, resulting in a rising beam center composed of fewer overlapping elements. As fewer elements are involved the center of the pyramid rises. An insurmountable challenge is created for finding minimum variance in the frequency response shape. If we take the high-frequency on-axis far point as a starting place we are already off the center axis in the other ranges. We will have to cross through on-axis beams at lower frequencies to get to our near off-axis milepost. The frequency response variance through this transition will not be the smooth and steady pink shift that we are hoping for.

The other categories of variance fare no better than their symmetrical counterparts. Ripple will still be high due to overlap inside the acute triangle zone. Level tapering reduces the overall power capability of the array, without providing significant improvements in variance minimization. That makes this a poor tradeoff.

Coupled Point Source

The symmetric version of the coupled point source has two variables that will be investigated for their effects upon spectral variance: splay angle and quantity. The relationship between quantity and speaker order will be seen indirectly. The asymmetric versions will include the effects of unmatched speaker order, level and splay angle as both independent factors and in combination. The scalar factor, i.e. the ability to duplicate the same effects in smaller (or larger) spaces using the same principles of asymmetry, will also be seen.

Splay Angle Effects

The first variable we will explore will be splay angle. We will begin where we left off and add progressively larger bends to the line source we have been examining. The presence of an angle, even a small one, creates an opportunity for reduced spectral variance that cannot be found in the coupled line source. Isolation, or at least some tendency toward isolation, is added to the summation equation and we will have the chance of reversing the relentless pattern narrowing. Figure 6.41 shows four different scenarios with successive splay angle doublings. In all cases we are using sixteen of the third-order speaker elements with identical drive levels. What is revealed is a gradual emergence from the familiar parallel beam concentration behavior toward a new dominant factor: beam spreading. We will see that these two behaviors are the governors of the low- and high-frequency ranges respectively, and they will meet at some point in between.

We begin with the smaller angle increment of 0.5 degrees per element. The combined angle spread is then 8 degrees (16 X 0.5 degrees). The low-frequency response (125Hz) is virtually unchanged from the coupled line source (see Fig. 6.38 for reference). This is not surprising since neither an individual spread of 0.5 degrees nor a combined spread of 8 degrees can be expected to have an appreciable effect upon elements with 300 degree individual patterns. As we move up to the 500Hz range the change is also negligible, for similar reasons. The 2kHz response begins to show some substantive change. The pyramid convergence is being slowed by the introduction of the splay angle allowing the expansion of the beam. The 8kHz response reveals a system at the crossroads. The response is frozen at the width of the array, unable to converge to the center (beam concentration) or diverge away from it (beam spreading).

As the splay angle doubles to 1 degree we see the beam-spreading action become more pervasive. The affected range moves down in frequency and the amount of angular spread opens up. The 16 degree spread is sufficient to reverse the beam concentration in the HF response. The coverage is clearly defined as an angular area with sharp and distinct edges. The coverage shape created by beam spreading is distinct from one created by beam

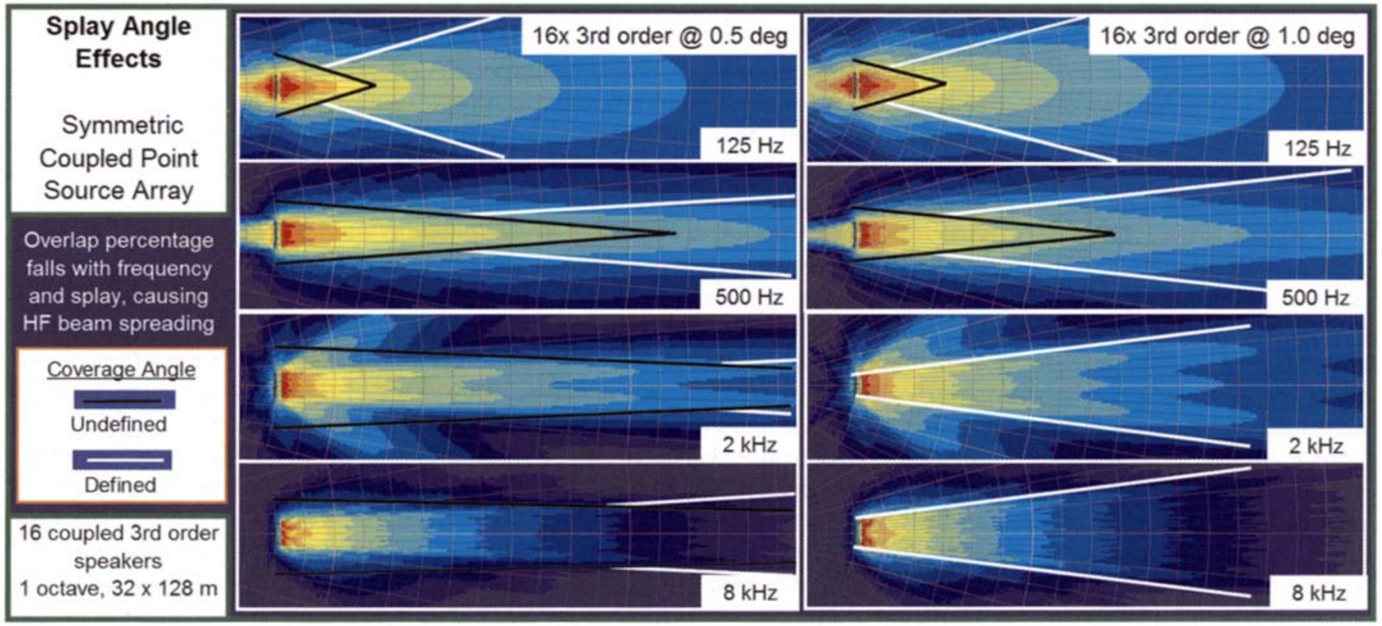


Figure 6 41(a) Splay angle effects over frequency for the coupled point source array

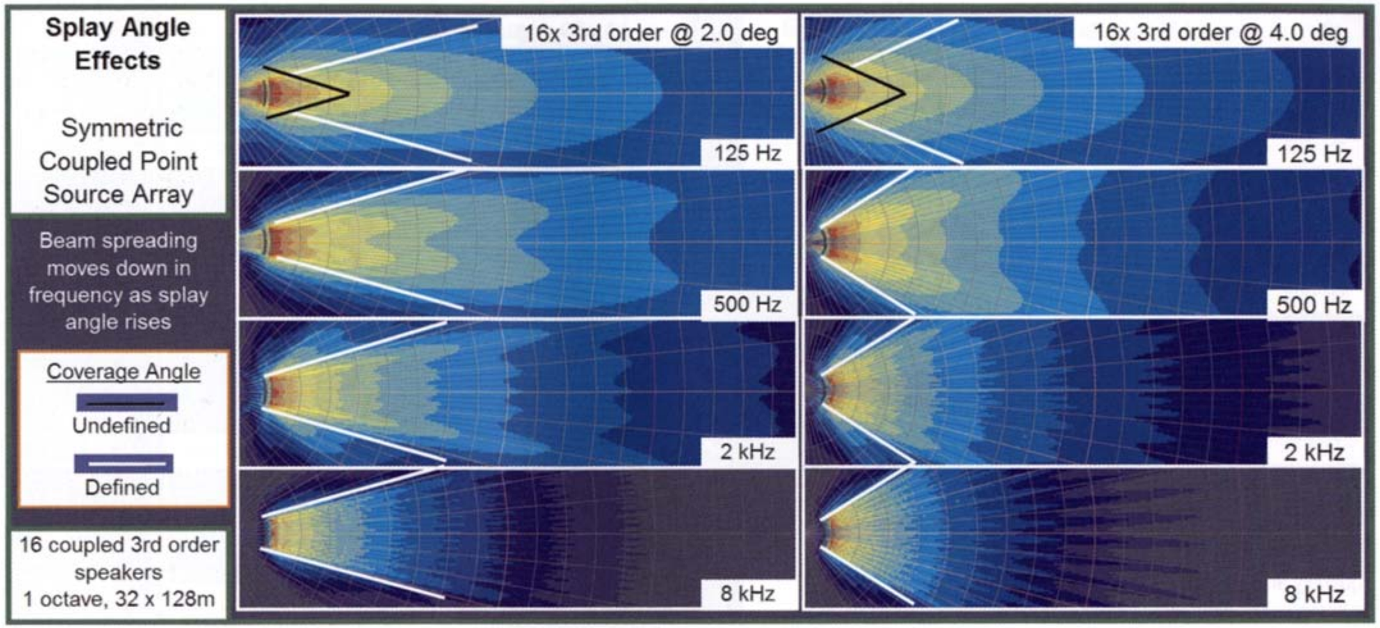


Figure 6.41(b)

concentration, even if they are described as having the same coverage angle. Since coverage angle is defined by comparing on-axis to the —6 dB point the distinction is missed. The beam spread shape holds its 0 dB value over the arc and then finally drops sharply to -6 dB and more. The beam concentration version begins its decline as soon as the on-axis point is left. Only the 0 dB and -6 dB points can be expected to match when such pattern shapes are compared.

Let's double the angle again. Now we have 2 degree splays creating a full spread of 32 degrees (Fig. 6.41(b)). The 500 Hz response is now finding itself in the crossroads, while the upper ranges show highly developed beam spreading. The final doubling (4 degrees) leaves only the 125 Hz response left in beam concentration mode. We have, however, reached a significant milestone: the coverage pattern over this eight-octave span is nearly perfectly matched. All four responses show —6 dB points approximately 64 degrees apart. The HF response has the most clearly defined edges. The LF response has the beam concentration shape but still reaches —6 dB at the same spot.

These four arrays can now be evaluated with our minimum variance criteria. They all show lower spectral variance than we found in the coupled line source arrays. The spreading of the highs stemmed the tide of endless narrowing formerly found when all elements were in beam concentration mode at all frequencies. The minimum spectral variance is found in the widest of the four arrays. An added bonus is that this will have the lowest ripple variance due to reduced overlap. We have found a minimum variance array: the symmetrical point source.

The success of this array type is not restricted to third-order speakers, but rather can be applied to all orders. Refer to Fig. 6.42. Here we see four different recipes for 90 degree coverage, ranging from a single first-order speaker to a six-element coupled point source of second-order speakers. In all cases the HF response falls in the 90 degree range. As coverage angle decreases the quantity required to fill the area rises. As quantity rises the low-frequency response narrows and the combing increases. A reduction in spectral variance is accompanied by a rise in ripple variance. Such are the tradeoffs we will be facing.

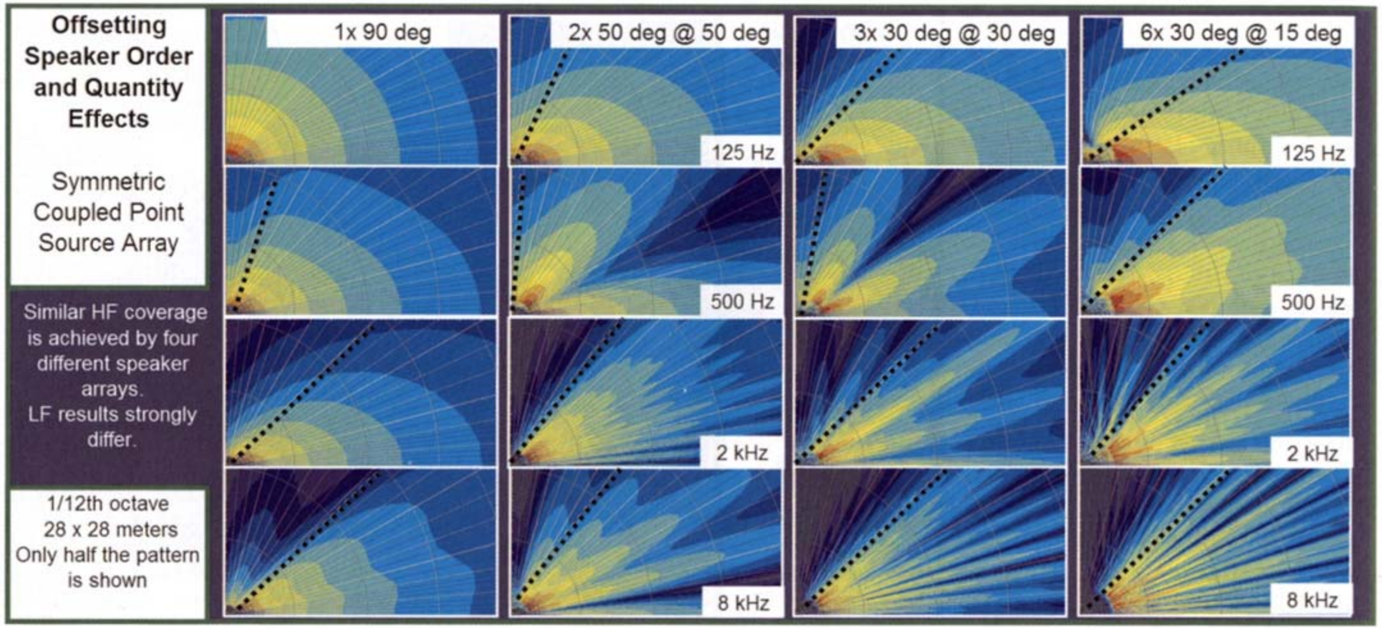


Figure 6.42 Offsetting speaker order and quantity effects over frequency for the symmetric coupled point source array

Asymmetric Level Effects

The fully symmetric coupled point source creates a radial fan shape of minimum variance. We will now investigate the effects of level asymmetry. Refer to Fig. 6.43. Our example 32 unit array has a constant 1 degree splay angle. The level is continuously tapered so that each cabinet receives a successively lower drive level. Note: such a continuous taper is not typical of the practical implementation in the field but is used here to provide a clear illustration of the trends.

An 8 dB level taper is obtained by 31 successive 0.25 dB level drops. Naturally we have an asymmetrical reshaping of the coverage pattern. Notice the difference between the two lower-frequency ranges and the two high-frequency ranges. The high-frequency ranges have been sculpted into a shape that clearly shows the successive level tapering. As level falls the belly of the array is sucked in by an increment amount in scale with our expectations regarding level and distance (see Fig. 6.24). This behavior takes on this malleable form because we have some measure of angular isolation. Contrast the matched 8 kHz and 2 kHz response here to the uninspiring efforts at level tapering explored previously for the line source array (Fig. 6.40). The low-frequency ranges, however, bear a much closer resemblance to the line source scenario because the 1 degree angular splay does not provide sufficient isolation to break out of beam concentration mode. Therefore, these ranges act more like a shortened line source array and we again see the beam center rising with frequency. If the beam remained centered on the array geometric axis it would conflict with the direction in which equal level is spread in the upper frequency ranges. The position where the beam would focus without level tapering is indicated by the dotted blue line. The desirability of lifting the beam is immediately apparent, as is the fact that the 500 Hz range is lifted to an extent that it resembles the HF responses. The excess overlap of the LF response makes for heavier lifting at 125 Hz and the effects are minimal.

As the level taper increases, the sculpted shape of the beam spread response bends further outward. The result

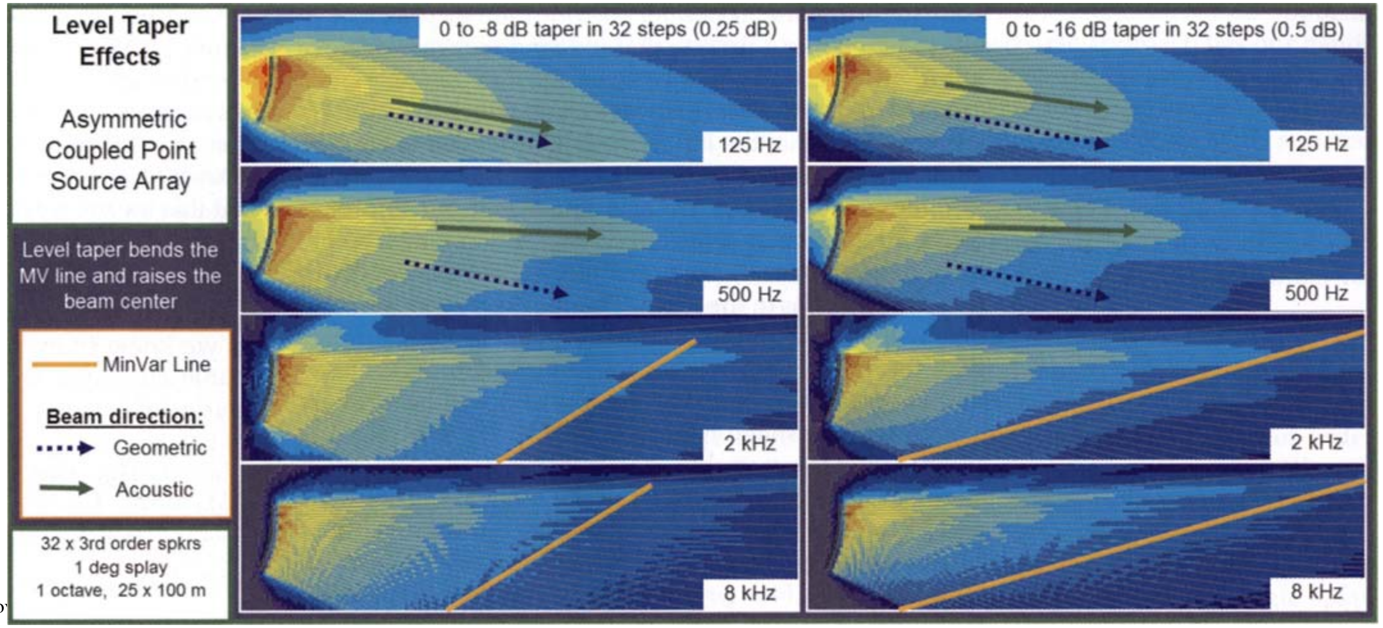


Figure 6.43 Level asymmetry effects o

is that the angle of the line of equal level is bent downward (Fig. 6.43, right side). If we visualize this array as a section view we can see that the increased level taper makes it more suitable for the lesser rake angle in the audience seating. An unambiguous steering control has been revealed that allows us to form-fit the response over a tilted surface. The steering is highly effective when we have angular isolation and less so when the overlap is high. Therefore, the ease of steering is generally inversely proportional to frequency. The 16 dB tapering scenario also reveals a slightly larger degree of beam lifting in the lower frequencies (125 Hz in this example). The extent of the change does little more than keep up with the increasing slope of the HF responses. No significant ground is gained or lost in the race toward steering all frequencies together.

It is worth noting that the low-frequency ranges are less "controlled" as a result of the level tapering. With all levels at full the beam is the narrowest, which is the property typically synonymous with the normally desirable parameter of "control." If we remove the level taper, the increase in control comes at the cost of placing the beam center in the wrong position. Less control is the consequence of level tapering, and the reduced apparent line length centers the beam on the level dominant array elements: the upper units. The loss of control comes with the bonus of steering the beam in the direction of the minimum variance frequency response shape.

Now let's examine this array in terms of our variance evaluation parameters: level, spectral and ripple. In terms of level we have a clearly defined line of equal level that is modifiable by the taper amount. The shape is held over frequency and shows many of the trends recently explored in the symmetrical point source. The minimum variance is found when beam spreading and beam concentration meet. In this asymmetrical case, however, the coverage is no longer definable as an angle, but rather as a shape. The beam-spreading response clearly delineates the shape and can be somewhat precisely tailored to fit the space. The task for the beam concentration range is upward response steering toward the direction of the spread.

It is worthwhile to ponder the question of priorities at this juncture. Let's assume that we cannot always achieve minimum spectral variance over the full range. Which range would be the best to focus on and which to let go. We know already that the HF range will be easiest to control, while the LF range will usually be overlapped and therefore more challenging. We also know that our efforts to control the LF range by beam concentration may backfire on us by over-steering the upper ranges. The answer to our prioritization lies in the room. The speaker/room summation will be a far more dominant factor in the LF range than the HF (unless we have a ***really*** bad room). Therefore we have to expect that the low end will be strongly affected. If our shape is not a match in the low end it is still possible to get usable energy from the speaker/room summation. Beyond the LF band the room's contributions will be too late to provide much usable addition, without costly combing. Therefore we must prioritize downward with frequency. First priority is to match the high mids to the highs. Then we add the low mids. If we make it all the way to the lows we can break out the champagne.

Next let's look at frequency response shape over listening position. Can we make a run from on-axis far to off-axis near? Yes we can, in one direction. The other side has no coverage anyway! An equal level line runs directly along this path, and the frequency response will be quite consistent as we move along the hypotenuse line. Recall that this means we have high prospects for matched spectral tilt. An equalization solution can be employed which will benefit a substantial area.

The final consideration for this array is ripple variance. As we know from our in-depth study of summation, the spatial crossover areas will exhibit substantial ripple in their overlap areas. In this case there are 31 asymmetrical spatial crossovers, each of which will have some measure of volatility. This particular example features a 1 degree splay angle so the amount of overlap is substantial. There are, however, two significant factors which reduce the ripple:

1. The angular splay and level tapering reduce the amount of double, triple and quadruple overlap, etc., as would be found in the symmetric coupled line source array.

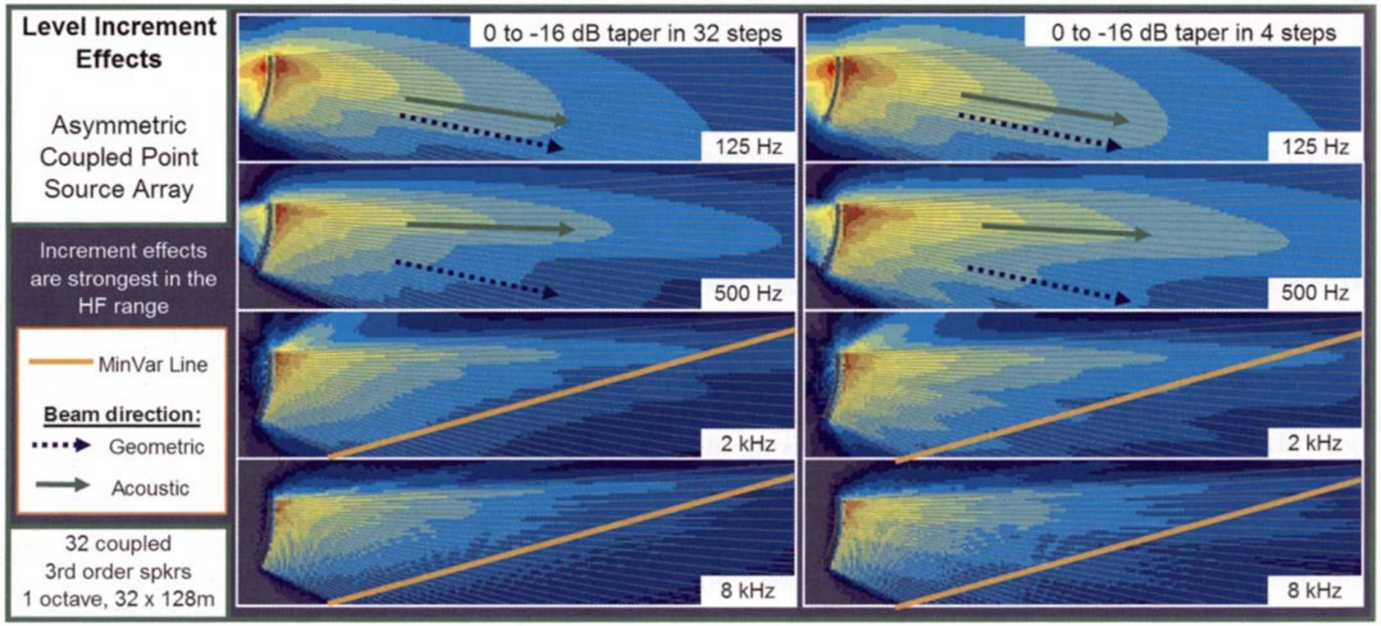


Figure 6.44 Effect of number of level increments over frequency for the asymmetrical coupled point source array

We are no longer in the acute triangle area of summation variance, but rather the obtuse triangle that leads toward isolation.

2. The time offsets can be low (provided the boxes are small and tightly configured) due to the minimal displacements between the overlapping sources.

We can conclude that the level tapered asymmetrical coupled point source has substantial potential for minimum variance performance in all three categories.

Level Increment Effects

How much effect does the number of level taper increments have on the outcome? If we get 16 dB of level taper, does it matter whether it's done in thirty-two steps or two? We can immediately guess that there will be some effect, but how much and at what frequency? The answer is as incremental as the question. The finer we slice the response the smaller will be the tatters on the edges. The example shown in Fig. 6.44 shows the response with thirty-two and four increments respectively. The differences are confined almost entirely to the high-frequency range, the area which has the most isolation between elements. Since isolation and individualized response control go hand in hand, this outcome is not surprising. There is no mistaking the effects, however. The four-step scenario leaves telltale fingerprints in the 8 kHz response. The key concept here is that the incremental level changes are a direct cause of spectral variance. The HF shape becomes unmatched to those whose edges are smoothed over by greater amounts of overlap.

The squaring of the HF response shape will rise as element isolation increases. If the splay angle were opened and isolation increased, we would expect to see more obvious break points and for the effect to become stronger at lower frequencies. An array of a first-order system requires the lowest ratio of elements to increments; while a third-order system is the most capable of grouping multiple elements on a single channel of processing. There is a practical trade-off between level increments and spectral variance. Fine slicing has signal processing and amplification costs, more steps in the optimization process and increased opportunity for wiring errors. The coarse approach ties our hands in terms of tapering the level of the array to create the overall desired shape (minimum level variance) without creating the spectral variance that comes with squared response edges.

From the minimum variance point of view we have a simple paradigm: when elements have fixed angles and asymmetric coverage distances they require asymmetrical drive levels to achieve matched levels.

Speaker Order Effects

We have previously viewed the effect of speaker order on the behavior of line source arrays (Fig. 6.37). We will now briefly visit the same concept for asymmetrical coupled point source arrays by viewing second- and third-order arrays with nearly matched parameters (Fig. 6.45). The overall length of the arrays, the overall angle splay and the level tapering (both overall level and number of increments) are matched. The individual coverage and quantity of elements are different, of course, with the third-order units having a larger quantity of narrower elements. The responses reveal substantial areas of similarity and some important differences. The extent of the similarities strongly suggests that array configuration is a far more decisive factor than the nature of the individual elements. Arrays composed of different elements with matched configuration have greater similarity than matched elements configured differently.

All four frequency ranges show the same basic shape. Both arrays would meet the basic minimum spectral variance criteria as found in the last example (Fig. 6.44). Beam spreading in the upper frequency ranges and beam steering in the lower ranges is found in both arrays.

The high-frequency responses of the two arrays have a similar overall shape in their principal area of coverage, as shown by the orange line. One notable difference is in the HF coverage above the top element in the array. The second-order system, with its wider elements, has more leakage above the array top.

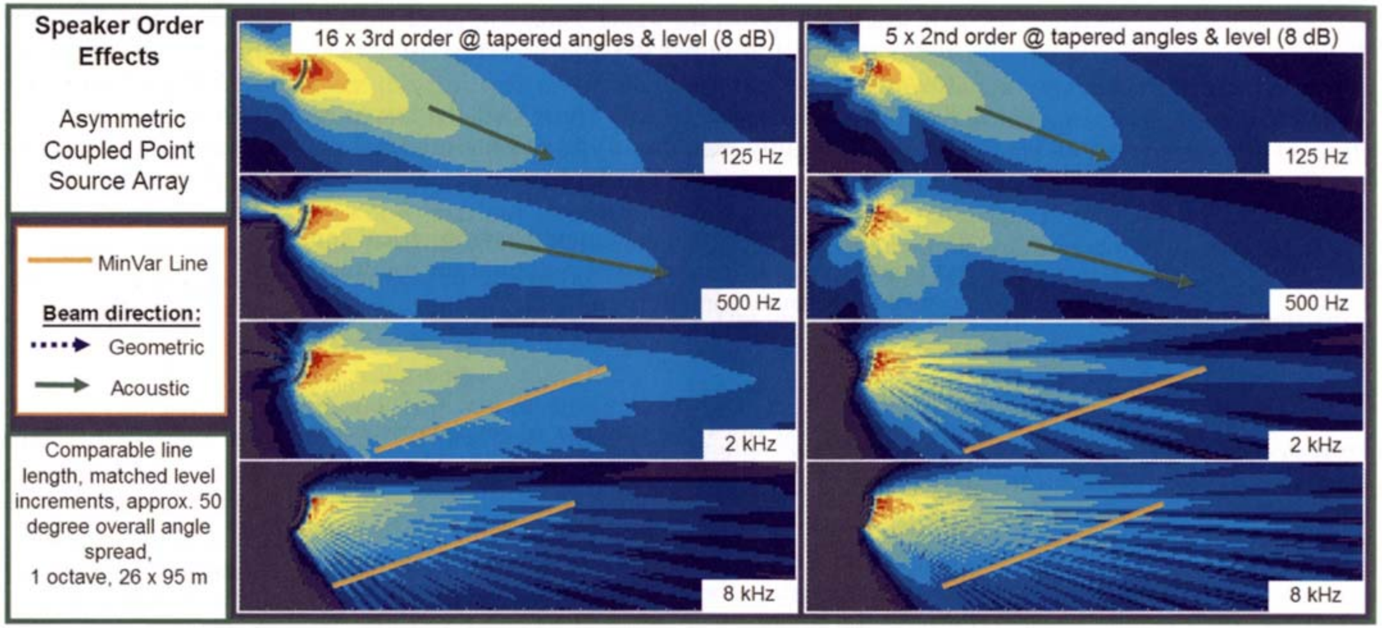


Figure 6.45 Speaker order effects over frequency for the asymmetric coupled point source array

The low and low-mid responses show the same focus point for the main beams. This is expected due to their comparable line length and level taper. The ripple variance is higher in the second-order system due to the larger displacements between the smaller numbers of elements.

What can we conclude from this? Minimum spectral variance can be achieved with this configuration regardless of speaker order, provided offsetting quantities are applied. As speaker order increases we will have the potential advantage of increased power, due to the use of more elements employed to create a given shape.

Combined Asymmetric Splay Angle and Level Effects

There are two principal mechanisms that can steer the line of minimum variance on the diagonal: level asymmetry and angular asymmetry. The two effects can be used in tandem to produce a combined effect which exceeds either of their individual efforts. The individual effects of level tapering and angle tapering are shown separately and combined in Fig. 6.46. These different mechanisms can create very similar shaping effects over the full spectrum. The extent of the similarity will, of course, depend upon the amount of asymmetry in either category. The combined effect indicates that these two mechanisms can be used in tandem, to create highly asymmetric arrays with minimum spectral variance.

Combined Asymmetric Speaker Order, Splay Angle and Level Effects

Another form of the asymmetric point source features unmatched elements. There are, of course, unlimited possibilities in this regard. Our representative example (Fig. 6.47) will be another version of the "layered" array, where we put together successive halving of coverage angle, level and splay angle. The result is a curved line of minimum variance in the familiar fashion of the asymmetric point source. This illustrates the wide variety of options available to us to create this basic shape. As previously noted, the isolated areas of the frequency response are the most responsive to asymmetric control techniques.

Perspectives: Starting with a blank palette is a wonderful thing for the mix artist — this is essentially what a comprehensive optimization allows you to do — wipe the room clean.

George Douglas

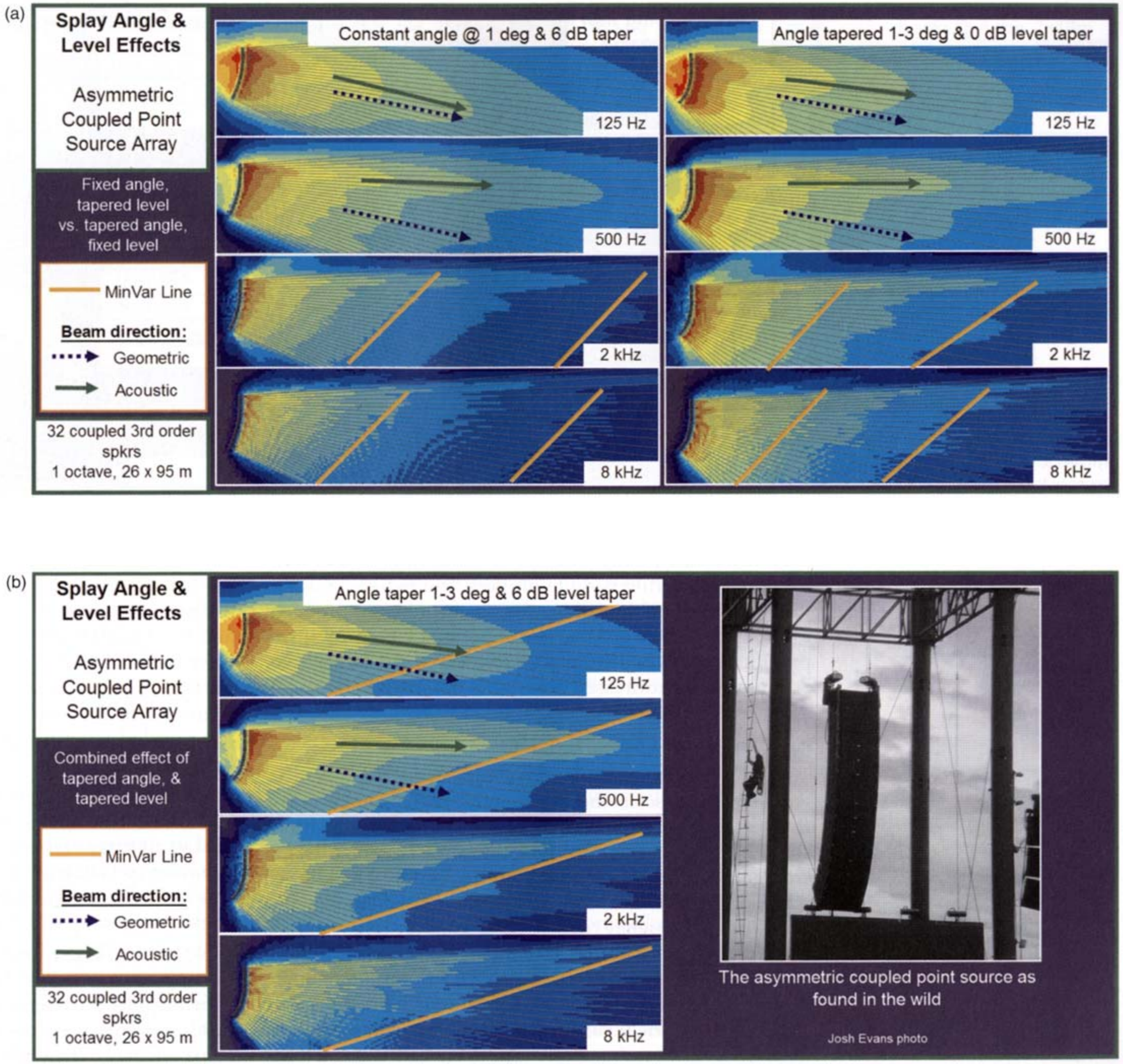


Figure 6.46(a, b) Separate and combined asymmetric level and angle effects over frequency for the coupled point source array

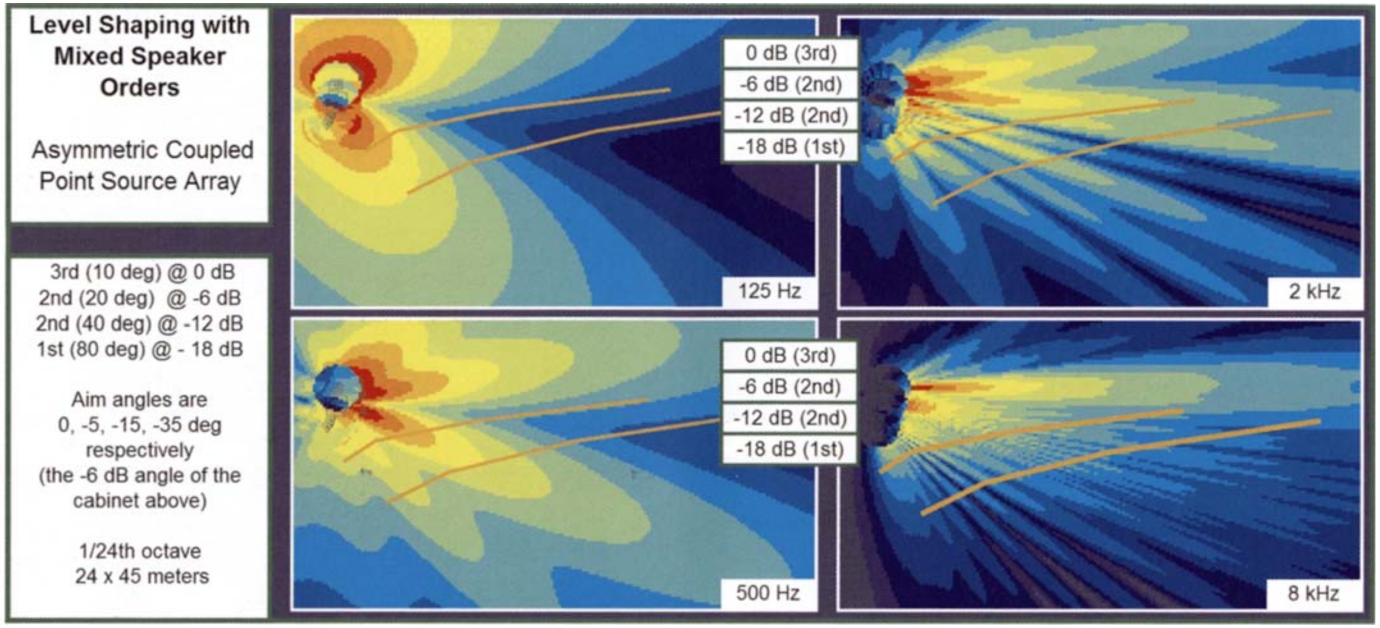


Figure 6.47 Separate and combined asymmetric level and angle effects over frequency for the coupled point source array

Note that the example here is of a small array and therefore the LF steering is quite limited. The reduced scale is the reason behind this, rather than some inherent property in mixing speaker orders. The superior LF control of the previous arrays is attributable to their much larger size. Scaling will be addressed next.

Scalar Effects

The next parameter under investigation is the extent to which we can scale the asymmetric point source array. If we maintain the same overall angle and level relationships from top to bottom will the same coverage shape emerge? Will the shape hold over frequency? The answer is a qualified yes, with the highest scalar correlation being achieved as frequency rises. For our example (Fig. 6.48) we will employ our familiar parameter doubling technique and observe the trends. In this case we will simultaneously double and half parameters as required to create scaled conditions. When quantity is halved, the splay angle is doubled so that the overall angular spread is preserved. As quantity is halved the modeled space is halved as well so that all dimensional relationships are preserved. The unchanged parameters include the individual elements, and the overall continuous level taper from top to bottom. Note that the smaller arrays maintain a line length in proportion to the graph size, but this is smaller in absolute terms. Since the frequency (and therefore wavelength) is not rescaled we can expect to see changes related to the line length.

The first trend to note is found in the ranges from 2 kHz and above. The overall coverage shape is functionally identical in all three cases. We have immediate validation that behavior of isolated array elements is scalable. All three arrays show similar beam spreading behavior over the 32 degree angle and 8 dB level taper.

Note: the 8 kHz response of the small-scale array gives the appearance of having increased ripple variance. This is principally a by-product of the rescaling of the prediction program's resolution, rather than solely an acoustic response. Since the large array is predicted in a 4x larger

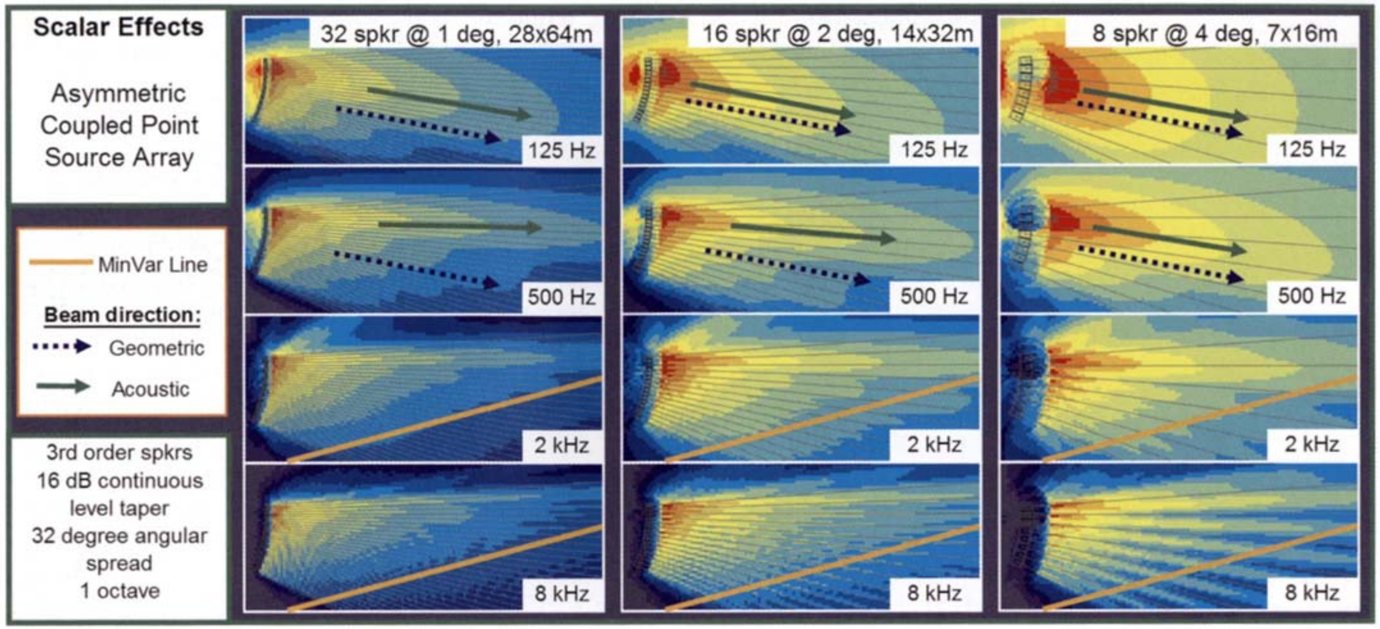


Figure 6.48 Scalar effects over frequency for the asymmetric coupled point source array

space the 8 kHz wavelengths are 1/4 scale. The result is the visual smoothing of the ripple variance, not the acoustical elimination of it. Bear in mind that spectral variance is the primary focus of this section and for this reason the prediction resolution is limited to emphasize the spectral trends.

Now let's move down to the 500 Hz range. Beam concentration is the dominant mechanism here, and therefore the actual line length, not its scalar ratio, will determine the coverage pattern in this range. In all three cases the position of the geometric beam center is the same. The capability to steer the response above the geometric center point is related to the line length and the extent of the asymmetry. The upward beam steering for the small-scale array is far less than the larger arrays due to the reduced line length. A similar trend is found in the 125 Hz range, with only the largest array showing significant lift capability.

We can conclude then that the scalar quantity is primarily a power capability and low-frequency steering issue. As quantity increases the power capability rises, albeit at the cost of increased ripple variance. This is as fair a trade as we can expect. As the quantities decrease we can expect to find increased spectral variance since the LF shape will have less conformity to the HF shape.

Hybrid Line/Point Source

Multiple unit arrays can be configured as a hybrid of both line source and point source systems. This is common practice in many applications utilizing third-order speakers that are marketed as "line array" devices. Most modern system designs do not place all of the array elements in a straight vertical line. At the very least designers feel compelled to point one or two enclosures at the bottom of the array toward the early row of seats. This is the "J" array referred to in the opening of this chapter.

The hybrid array is inherently asymmetric. Therefore our discussion will proceed immediately to investigating that factor. The upper section of the array will act as a symmetric coupled line source, while the lower section will act as some version of the coupled point source. The key parameter is the transition point between the arrays which is inherently an asymmetric spatial crossover. The position and degree of asymmetry of the spatial crossover will be the result of the relative quantities of elements in the two array types.

As usual we will conduct a series of doubling scenarios to isolate the effects. In this case we will modify the proportion of line source and point source components. The overall element quantity remains 16 in all cases and the orientation of the top and bottom elements remains fixed. The line source section is presented only in symmetrical form, since we have previously shown the lack of merit in level tapering this array type (see Fig. 6.40). The point source section is shown in both symmetric and combined level plus angle asymmetric forms.

The line source components comprise two, four and eight of the sixteen total elements in the different scenarios. The results are shown in Fig. 6.49. We will begin with eight elements of each type. In this case the high-frequency ranges clearly show the two array types as coexisting in separate worlds. The eight line source elements project a perfectly focused beam that we recognize as standard beam concentration behavior, while the bottom eight point source units give us the standard HF beam spreading behavior. The resulting shape is that of a gun pointed at the rear of the hall. (We have modernized things and replaced the coverage bow with the coverage gun.) We are unlikely to encounter a room with an audience spread along this shape.

The reason that these two array types don't mix is that the beam concentration behavior steers the energy away from any components that are inclined toward beam spreading. The concentration increases the on-axis level and narrows the beam and takes it out of reach of the spreading components. The energy from the spread beam elements is unable to merge with the slope of the parallel pyramid.

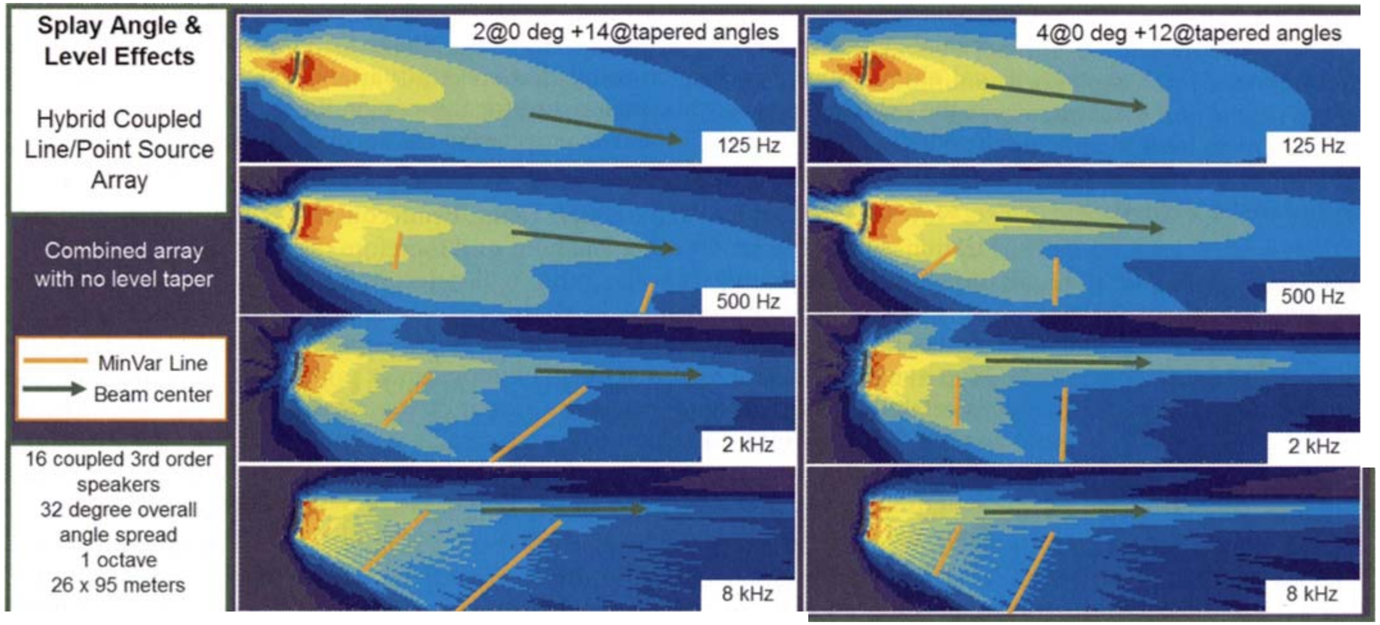


Figure 6.49 Level and splay angle asymmetry effects over frequency for the hybrid coupled line/point source array

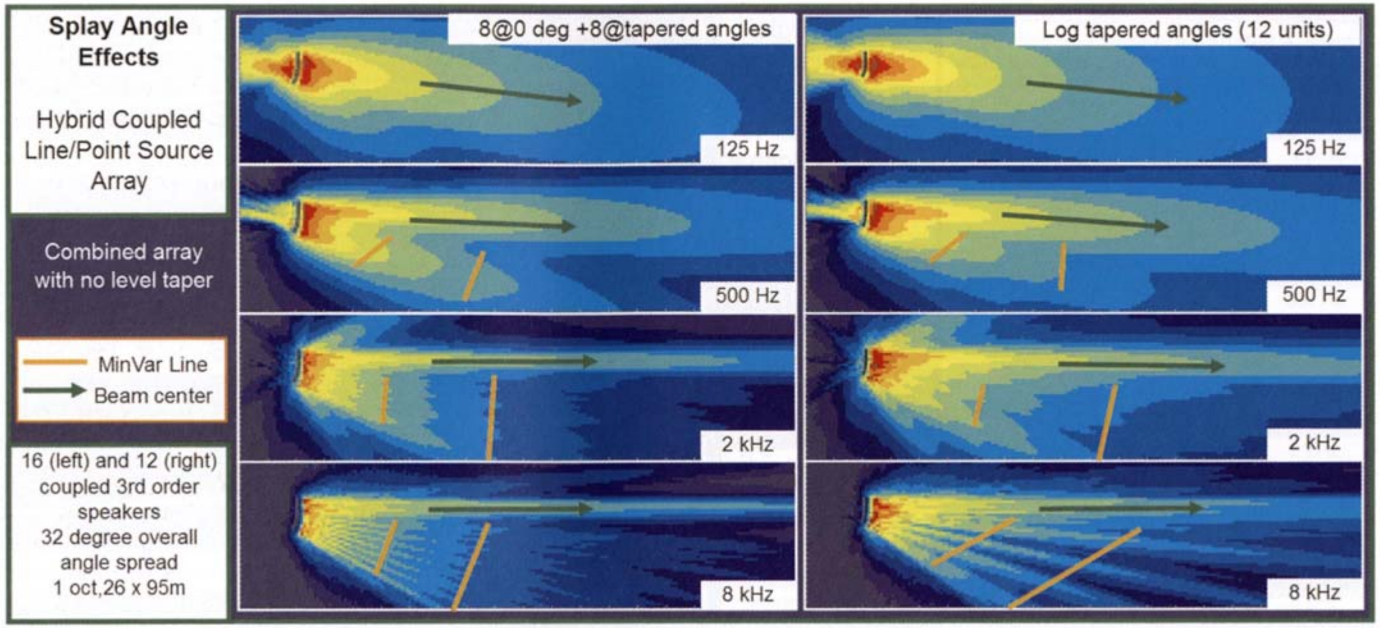


Figure 6.49 *(Continued)*

As we move down in frequency we find that the increasing overlap eventually causes the beam concentration to dominate in both array types. The combined response then takes the shape of the typical asymmetric point source.

We are trying to mix two shapes that cannot possibly be married together over distance: a variable pyramid and a fixed arc (see Fig. 6.50). Let's start with the pyramid. The line source, until it has completed the parallel pyramid, does not have a defined coverage angle but rather a defined physical width. The distance where the summit is reached, and the coverage angle that emerges, is hugely variable over frequency. The amount of spectral variance we have over a given width inside the pyramid will change over distance. It is never constant. Now on to the arc. The coupled point source is defined by a coverage angle. We don't have to go far at all for the angle to become rock solid over distance. It is fixed over unlimited distance and if well-designed, it is fixed over frequency. The amount of spectral variance we have across the coverage angle will be stable over distance. Now how in the world are we going to put these things together to create a combined response shape that is stable over distance over frequency? I do not have an answer to that.

Let's continue. Let's try reducing the quantity of line source units. This reduces the on-axis addition and retards the narrowing of the line source beam. The effect is so minimal as to be almost negligible. The HF responses show a slight thickening in the merger area between the beam concentrating and spreading zones. We still have a loaded gun aimed at the audience. Even if we reduce the line source proportion to just two units, the gun shape persists. As few as two elements at zero degrees provide a beam concentration which puts it out of reach of neighboring beam-spreading elements.

One additional scenario shown here is a "log" angle taper, a practice recommended by one of the manufacturers of "line array" speakers. This is actually a form of the asymmetrical point source, since no angles are 0 degrees. The initial angles are so minute (0.2 degrees) that the isolation is negligible and results are closer to the hybrid arrays than to the asymmetrical point sources covered earlier. The gun barrel remains in view for the HF responses.

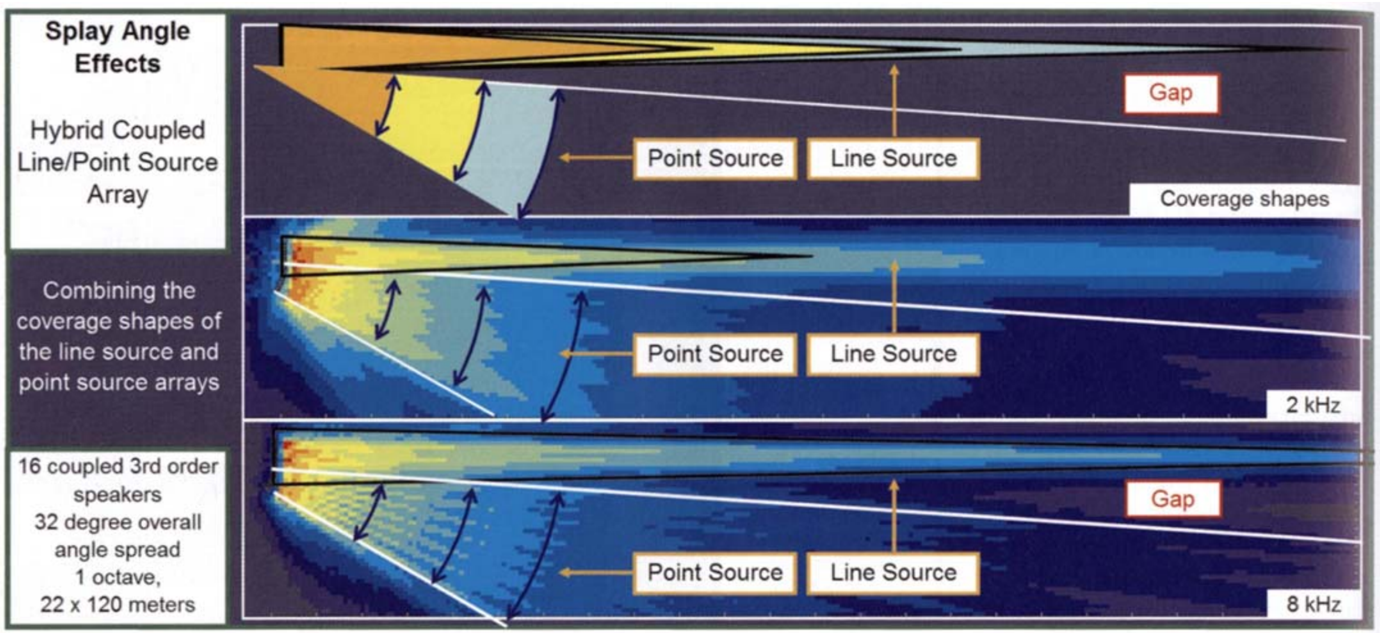


Figure 6.50 Coverage shapes for the components of the hybrid coupled line/point source array

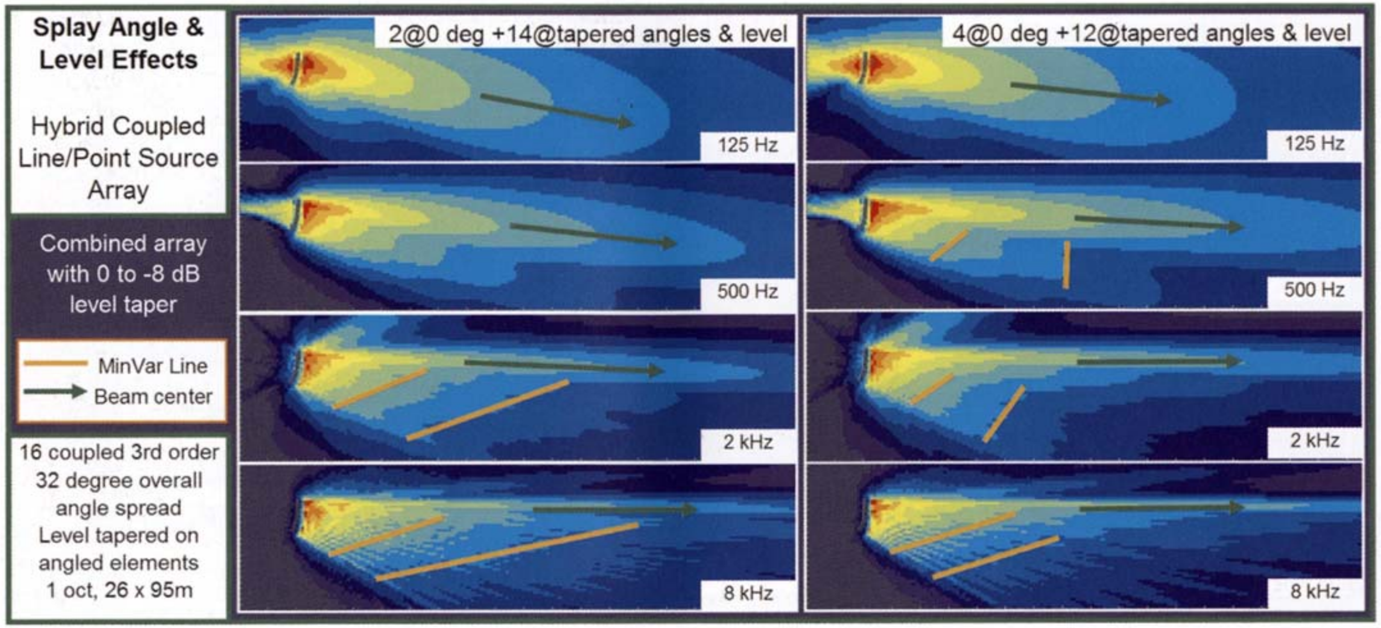
Asymmetric Splay Angle and Level Effects
Before leaving the hybrid arrays let's take a moment to consider the effects of level tapering as shown in Fig. 6.51. Since we have previously shown the lack of merit in level tapering the coupled line source, we will limit the tapering to the point source elements. The level taper in all cases is gradual and reaches a total of 8 dB at the bottom. The results do not provide any evidence of improvement in the separation between the line and point source systems. Only the HF range in the area covered exclusively by point source elements shows any substantive change. The difference in shaping does create a combined shape that is much closer to one we might find in the field. The gunpoint beam, however, would still blow a hole through the back of the hall.

Uncoupled Arrays

Our study of coupled arrays revealed that the creation of consistent level contours over frequency is extremely challenging. In even the best of circumstances we can only hope to provide a reasonable approximation over the maximum frequency range and listening area. This, however, is a "walk in the park" compared to the hazards we will meet with uncoupled arrays. A review of Chapter 2 will remind us of the spatial properties of summation, which will affect all frequencies differently as displacement becomes a dominant parameter. When substantial displacement and overlap are found together, all hope of minimum spectral variance goes with it. There are a limited number of directions we can go where we can hold the forces at bay for a defined area. This section cannot run through all of the possible iterations, but instead will focus on the directions where limited success is possible.

Uncoupled arrays have three directions in which we can create a reasonably consistent shape: lateral, radial and forward. Example applications will reveal the trends in all three of these directions.

Uncoupled Line Source
We begin with the symmetric version of the uncoupled line source as shown in Fig. 6.52. The minimum variance area



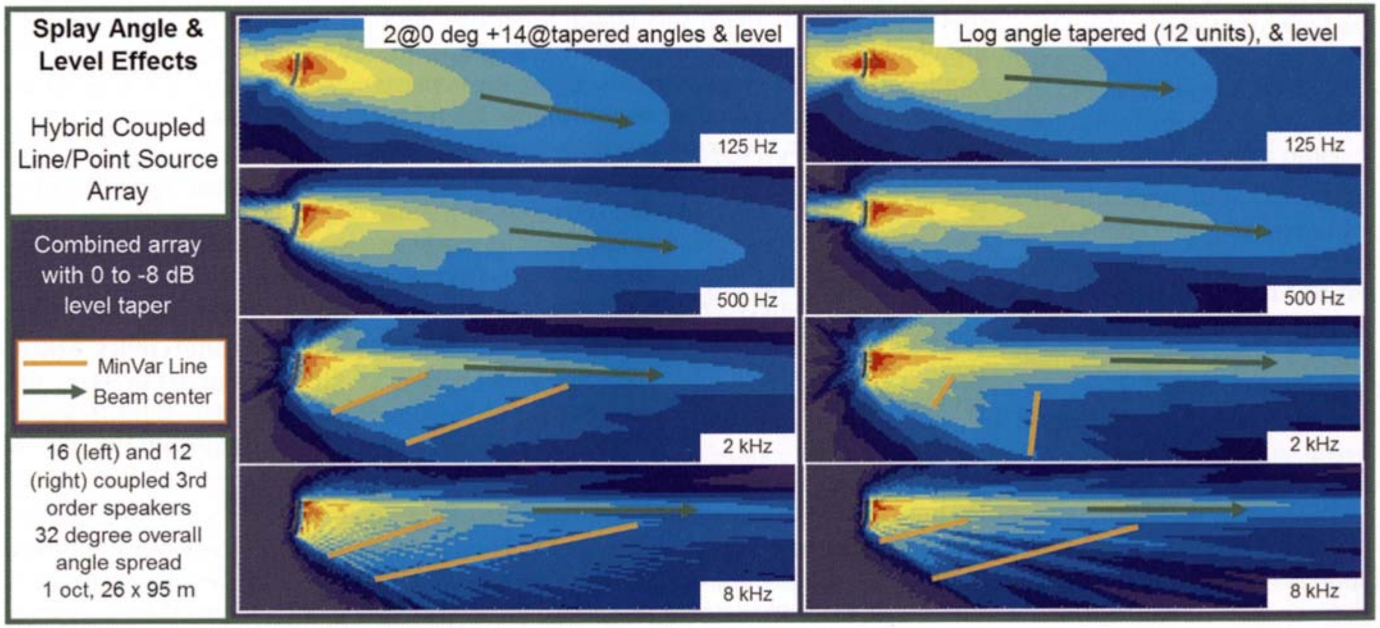


Figure 6.51 Level and splay angle asymmetry effects over frequency for the hybrid coupled line/point source array

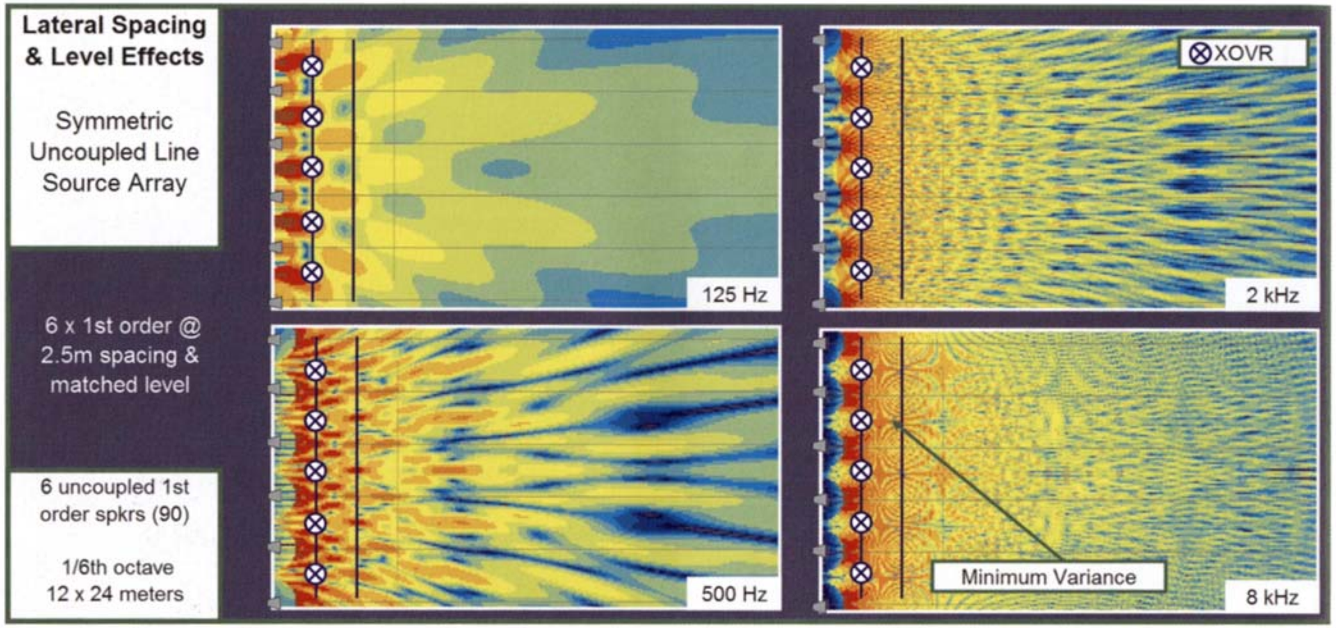


Figure 6.52 Coverage shape over frequency for the uncoupled line source array

runs from 50 per cent to 100 per cent of the aspect ratio as discussed previously in this chapter. Because there is no angular isolation, our coverage range depth is limited by ripple variance. Success is measured by the consistency of the start and stop points for the coverage. Therefore a consistent shape for the array requires a consistent shape in the original elements, which favors the first-order speaker. As speaker order increases, the start and stop areas move forward only in the highs, leaving the mid-range and lows behind. As spacing widens the start and stop points will scale proportionately. Frequency does not scale, so the ripple variance will differ with displacement. The effects of these variables are shown in the figures in Chapter 2 and for brevity are omitted here.

Offsetting Effects of Level, Spacing, and Speaker Order

We move forward now with an example of shape shifting with the uncoupled line source. The variables are relative level, spacing and speaker order. There are an infinite number of ways to mix these three variables. For brevity and illustration we will show how the offsetting of these variables can be used to create an asymmetric combined shape. This is a variation of the layering technique we applied previously to the asymmetric coupled point source.

We begin with a series of first-order elements with asymmetric log spacing and level as shown in Fig. 6.53. Log spacing in this case is a halving of the displacement with each successive element. The offsetting log level change is 6 dB per element. The lower elements are delayed so as to create a series of phase-aligned spatial crossovers marked as blue circles. The result is an angled series of equal level contours favoring the louder elements similar to the asymmetric point source. The shape is fairly consistent over frequency because of the low order of the elements. The low-frequency ripple variance is due to the lack of isolation among the uncoupled elements.

Next we can view the same relationship with second-order speakers as shown in Fig. 6.54. The spacing is reduced proportionally due to the higher aspect ratio of

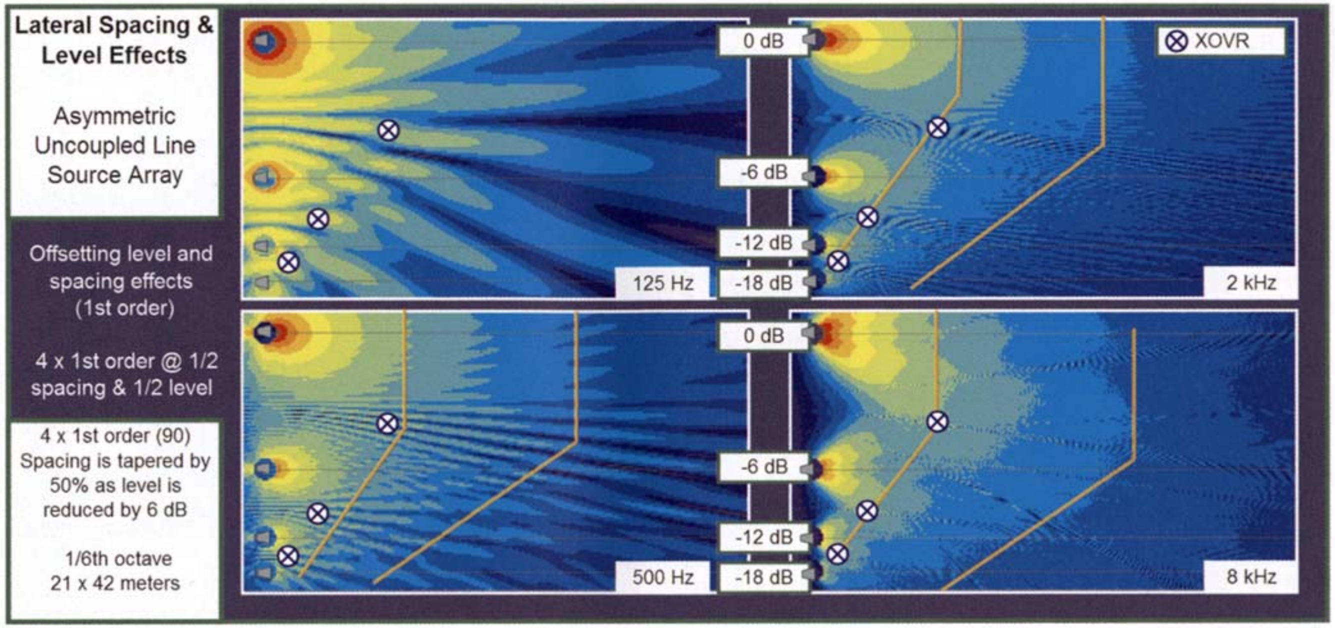


Figure 6.53 Offsetting level and spacing angle asymmetry effects over frequency for the uncoupled line source array (first-order speakers)

the elements. The offsetting log proportions of spacing and level remain but the physical displacement changes. The resulting shape shows a forward extension of the high-frequency coverage as would be expected. The resulting shape is less stable over frequency than its first-order counterpart due to larger amounts of overlap below the directional HF range. The increase in angular isolation does not offset the increased proximity between the elements below the HF range. The expanding beamwidth over frequency causes a steady upward bend in the combined shape as frequency falls.

The key concept here is one of compromise. We gain a limited shape extension but pay a price in spectral variance. The third-order system (not shown) is proportionally less stable over frequency as would be expected.

A third scenario (Fig. 6.55) introduces asymmetry in the element coverage angle by using speakers of mixed orders. In this case the aspect ratio for each element is approximately double the previous, which replaces the log spacing as the offset parameter against the log level taper. The spacing is linear. The resulting shape is a curve that holds its shape though the highs and mids. Once again we have isolated beam spreading in the highs. The low-mids and lows are too far displaced to create coherent beam concentration. What is seen is the ripple variance that results from the uncoupled beams passing through each other.

We can also travel forward at the same time. This gives us a variable origination point that becomes part of our log offset equation. We will begin with a first-order speaker element which has log spacing in both the lateral and forward directions and the usual log level taper (Fig. 6.56). The crossovers are phase aligned, resulting in a flat line of matched level in front of the array. This is an important but illusory milestone, as it can not be maintained over distance due to asymmetrical doubling distance losses between the near and distant sources (see Fig. 6.25). The transition can be made gradual enough that the shape provides a viable target for merging forward and laterally displaced systems. As frequency falls the early lines of equal level bend inward toward the louder elements.

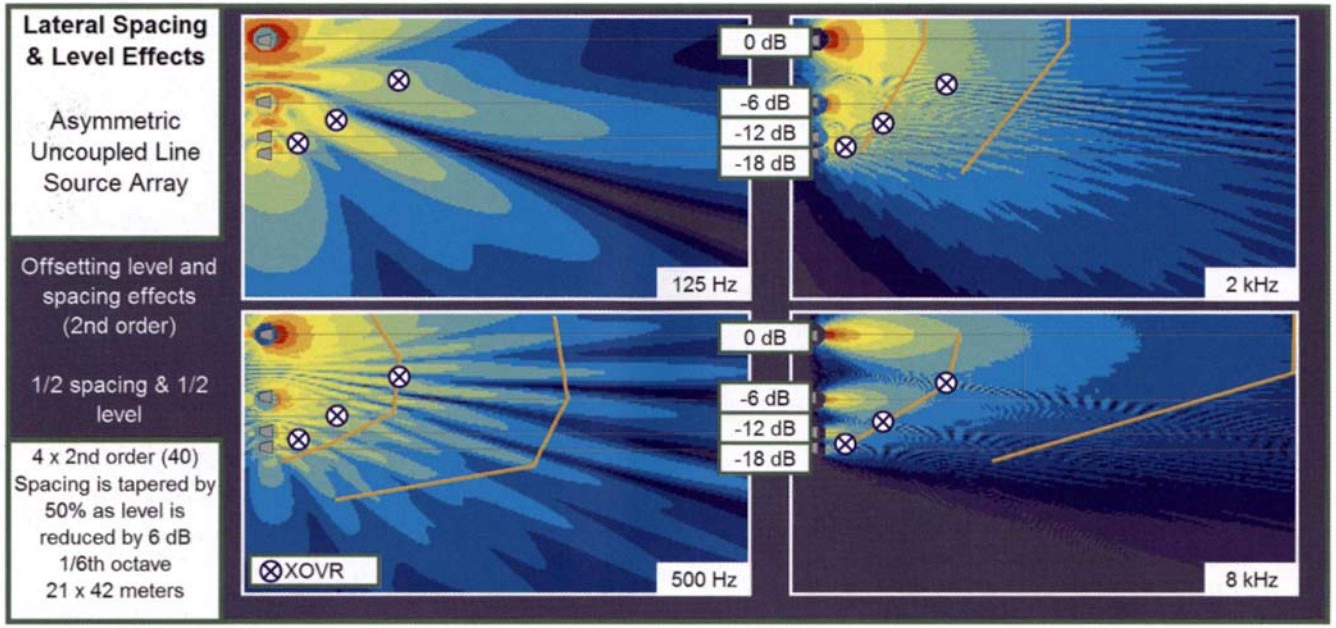


Figure 6.54 Offsetting level and spacing a
speakers)

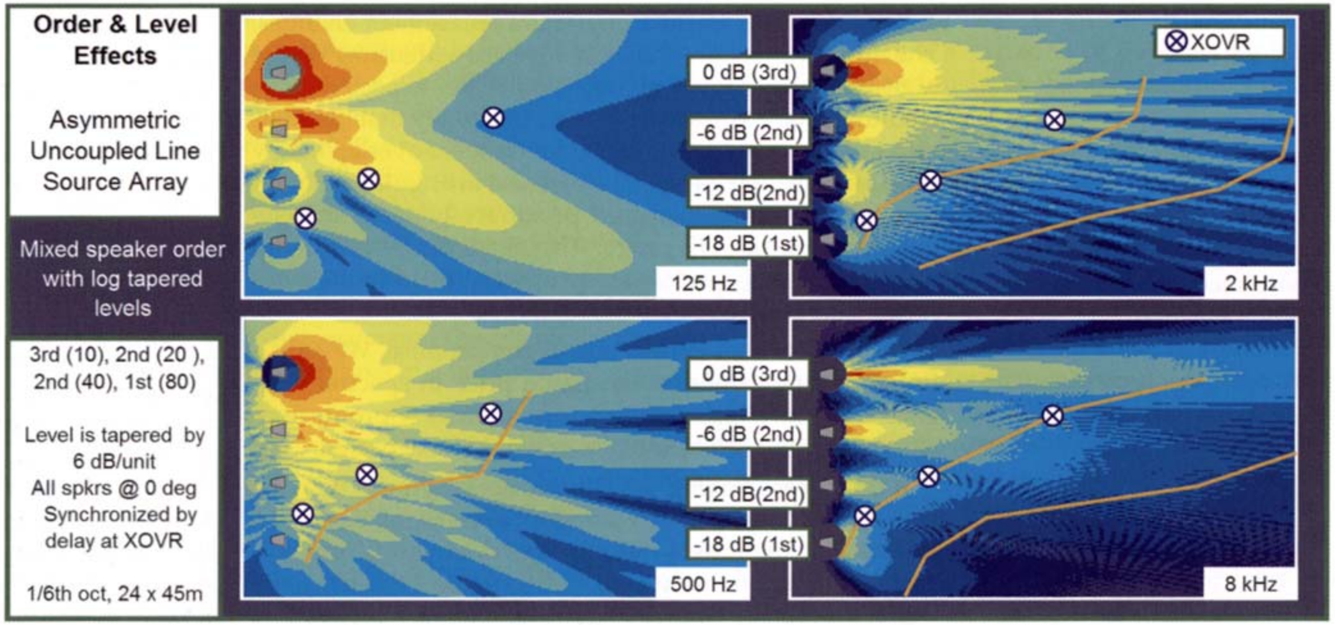


Figure 6.55 Offsetting level and spacing angle asymmetry effects over frequency for the uncoupled line source array (mixed order speakers)

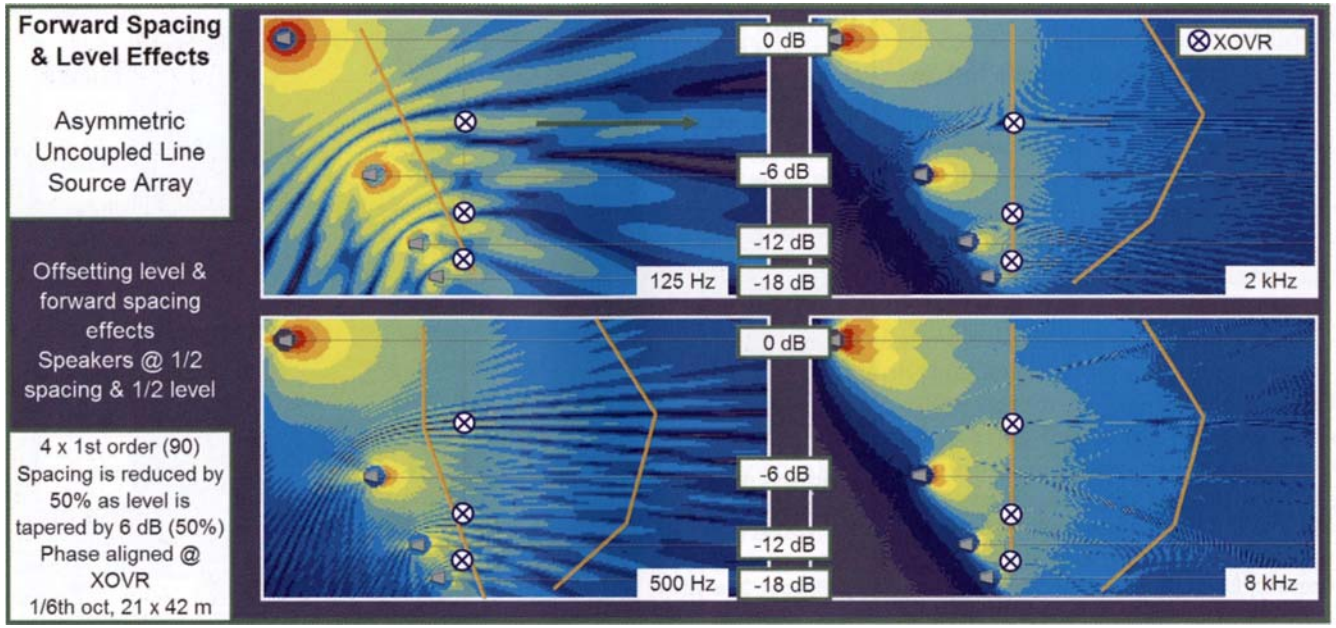


Figure 6.56 Offsetting level, point of origin and spacing angle asymmetry effects over frequency for the uncoupled line source array (first-order speakers)

Offsetting Effects of Origin, Level and Spacing
We can also view the same type of scenario with second-order speakers as shown in Fig. 6.57. Consistent with previous trends we find that the usable area and frequency range are both reduced proportionally.

It may be difficult to grasp the practical application of these scenarios. It would be highly unusual to design a system with such a series of sequenced relationships. The intent is to show the mechanisms that control these interactions and present them as optional tools. By doubling and tripling the effects, the trends pop out. The obvious conclusion is that any uncoupled array form is range-limited in two categories: distance and frequency range. Distance buries us with ripple variance. As frequency falls we lack isolation even at close range and the ripple arises again.

It would all be well and good to conclude that use of such arrays should be abandoned in favor of the powerful and low-variance coupled arrays. But we can't. The place where these asymmetrical uncoupled line and point sources come into play is when we get to the big picture. Step back and look at these figures and replace the individual elements with coupled arrays. At the top are the upper mains on the top of the proscenium side wall. Next are the lower mains that are flown three meters off the deck. Then we have the front fills. The combination of these coupled arrays will behave as the asymmetric uncoupled arrays found here. The big picture is a weave of the little ones.

Uncoupled Symmetric Point Source
The radial expansion method is very well known to us from previous discussions in both Chapter 2 and the level variance section here. The uncoupled point source array has an arc of minimum level variance whose start and stop points are determined by displacement and angular overlap. The response of a second-order system is shown in Fig. 6.58. The enclosed area shows a high degree of level conformity over frequency, albeit not immune to ripple variance. Even the response in the low-frequency range shows a higher degree of correlation than the previous

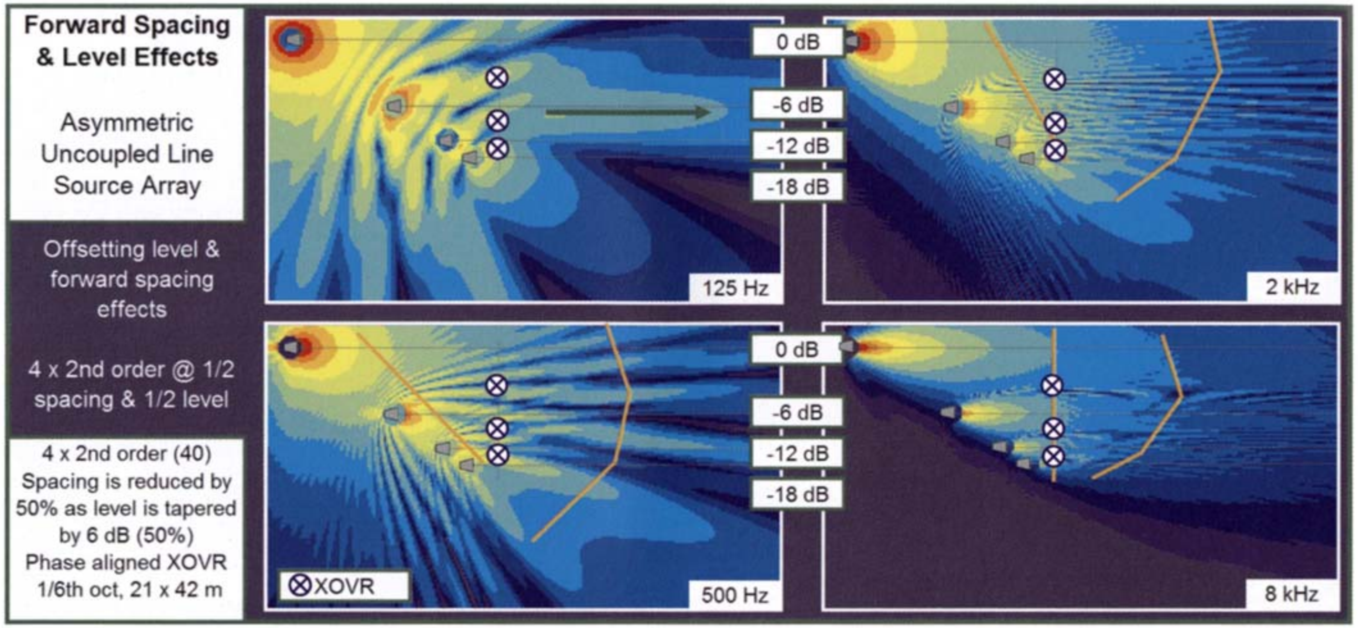


Figure 6.57 Offsetting level, point of origin and spacing angle asymmetry effects over frequency for the uncoupled line source array (second-order speakers)

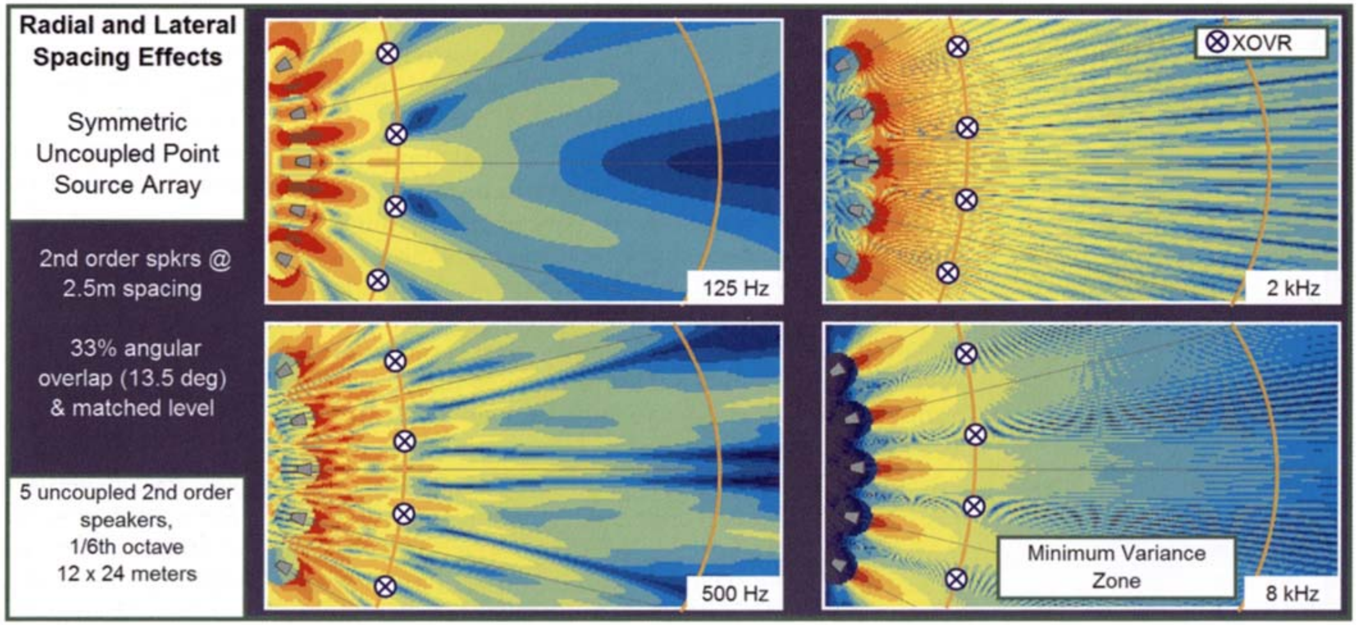


Figure 6.58 Coverage shape over frequency for the uncoupled point source array

scenarios. The uncoupled point source can be made asymmetrical in a variety of forms: level, angle and speaker order. The trends of these actions are largely consistent with those we have just studied.

Uncoupled Asymmetric Point Destination

Sources that are forward-displaced and operated at different levels will have a limited depth of interaction. The decisive factors are the distance and level ratios. At some point along the forward line the two sources will have matched levels, this being the spatial crossover point. Continuing in a forward direction away from both sources will cause each of the sources to lose level. The rate of level loss will be asymmetrical since it is based on the different doubling distances from the sources. The rate at which the more distant speaker becomes dominant in level increases with the distance ratio. The size of the affected area is inversely proportional to the level ratio. The relationship of these two factors was shown previously in Fig. 6.24.

In our example scenario (Fig. 6.59) the spatial crossover point is created at the same distance in all cases. The variable factors, distance ratio and level ratio are matched in order to create a matched combined response at the listening position. The combined level at this position, not coincidentally, is matched to the mid-point in the depth of coverage from the primary source.

Let's begin with a distance ratio of 2:1. This might be 20 m and 10 m respectively, or any scalar equivalent. This distance ratio from a single source would create 6 dB of level difference by the inverse square law, and that will be duplicated by our adding 6 dB of attenuation to the forward speaker. The result is they both arrive at the meeting point at equal level and combine for 6 dB of addition. Forward of the meeting point we see the loss rate over distance is greater for the forward speaker. The result is that the forward speaker slowly drops out of the combination equation as we move onward. The remaining iterations move in doubling distance ratios and 6 dB incremental level reductions. The result is a constant level at the spatial crossover point but a steady reduction of the range of interaction both before and after the spatial crossover point. The high ratios decrease the ill effects of "backflow" from the remote speaker, but also reduce its range of additive forward power.

The typical delay speaker scenario is an inward angled uncoupled point destination array. The role of the forward speaker is to improve the ratio of direct sound over the reflected and to reduce the propagation level loss due to inverse square law. Three different delay scenarios are shown in Fig. 6.59 with varying ratios of distance and level. These scenarios add some realism to the previous discussion which focused on pure level, with no angular component or ripple variance considerations. The three combined scenarios are compared to a baseline of the response with no delay speaker. In the three combined cases the speakers create a phase-aligned crossover at the same point. The distance ratio, the angular relationship and the relative level are all adjusted to create a matched meeting point. Both the main and delay speaker are second-order.

We will begin with discussion of the low-frequency response. In all cases the level in the spatial crossover area is raised by 6 dB from the baseline (top row), which is the sonic equivalent of moving the listeners forward to the center of the room. The 2:1 scenario is able to substantially arrest the level loss at the rear of the room but this comes at the cost of ripple variance in the local area of the delay speaker. The ripple is highest between the sources, especially near the back of the delay speaker. This is due to the time offset between the sources, which are set to be phase-aligned at the spatial crossover. This "backflow" area behind the delay speaker is a recognizable feature of our triangulation discussion: it is on the straight line of maximum ripple variance between two sources (see Fig. 6.5). The 4:1 (-12 dB) and 8:1 (-18 dB) scenarios have less forward addition and proportionally less backflow.

A secondary factor is the angular orientation of the speakers in its own right. Figure 6.60 shows the effects of various angular orientations with the same distance ratio. Whereas the level ratios just discussed create an area of equal level, each angular orientation creates a different line of equal time. The assumption is that the delay has

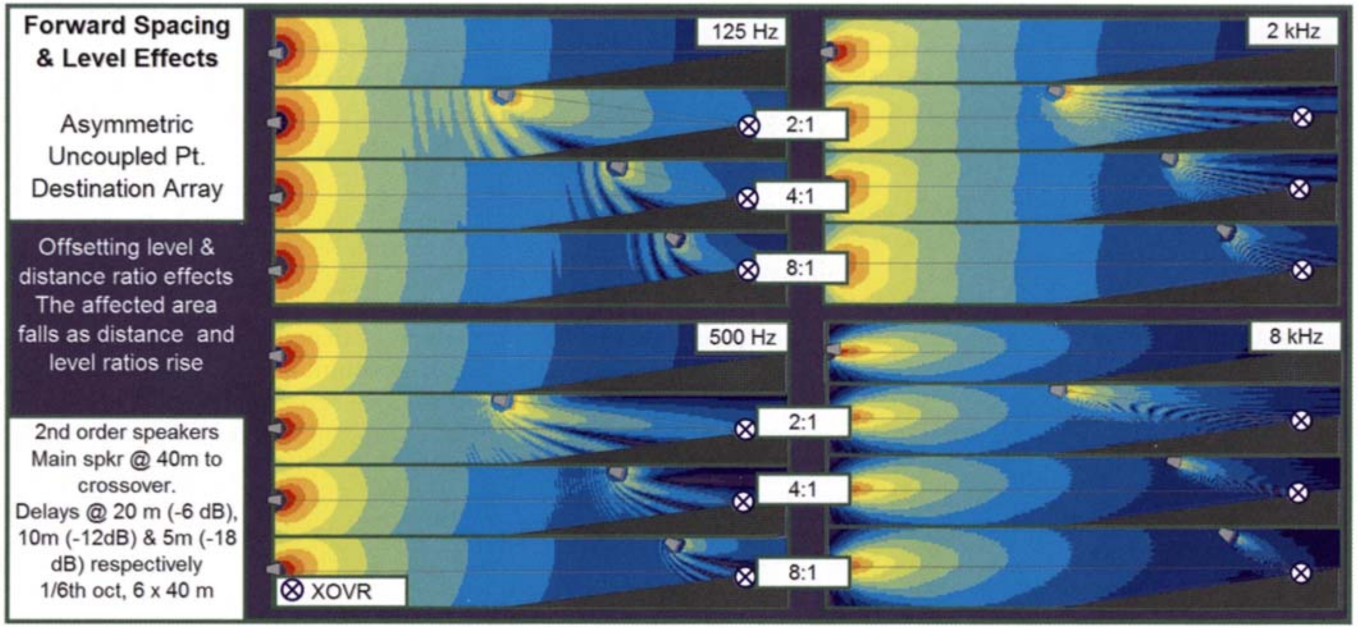


Figure 6.59 Level and distance ratios applied to the asymmetric uncoupled point destination array

been synchronized at the spatial crossover. The equal time contours are marked "sync" on the drawing. This line indicates the locations that remain equitemporal between the speakers. Additional lines can be seen that show the incremental time offsets of 1 ms up to 5 ms apart. The 5 ms limit indicates a range where the comb filtering effects would extend down to 100 Hz, low enough to degrade the entire working range of most delays.

In some cases the delay speakers lead the main and in some cases, vice versa. There are several important trends to note. The first is the fact that as the angular offset increases, the size of the usable area (inside the 5 ms limits) decreases. A second trend line lies in the relationship of time offset and level offset. Note that the area behind the spatial crossover (from the main speaker viewpoint) shows the delay speaker ahead of the mains. Recall that the doubling distance loss rates are asymmetric and therefore the level offset favors the main speaker. What we have here are offsetting factors in the battle for sonic image (see Chapter 3). As we move back in the hall the image stays constant due to the offsetting factors. As we move forward of the spatial crossover we find the main leading in time, and (until we move off-axis) the delay leading in level. Again the competing offsets create an image stalemate.

The mid-range plots show the effects of increased directional control as frequency rises. The range of interaction is reduced at the same time as the ripple sensitivity (due to shorter wavelengths) is increased. The result is decreased backflow ripple but we are subject to increased ripple in the forward interactive area.

Note also the effects of the changing angular relationship between the sources as the distance ratio changes (Fig. 6.59 again). The inward orientation angle of the point destination array increases as the delay speaker moves back in the hall. This causes increased ripple variance over the space that is consistent with our findings regarding the angular affects of point destination summation (see Figs 2.81-2.88). Finally we reach the 8 kHz range where the directional isolation has reached its maximum. Backflow is negligible and the range of addition is small. Such small additions, however, are progress toward our goals. We have improved

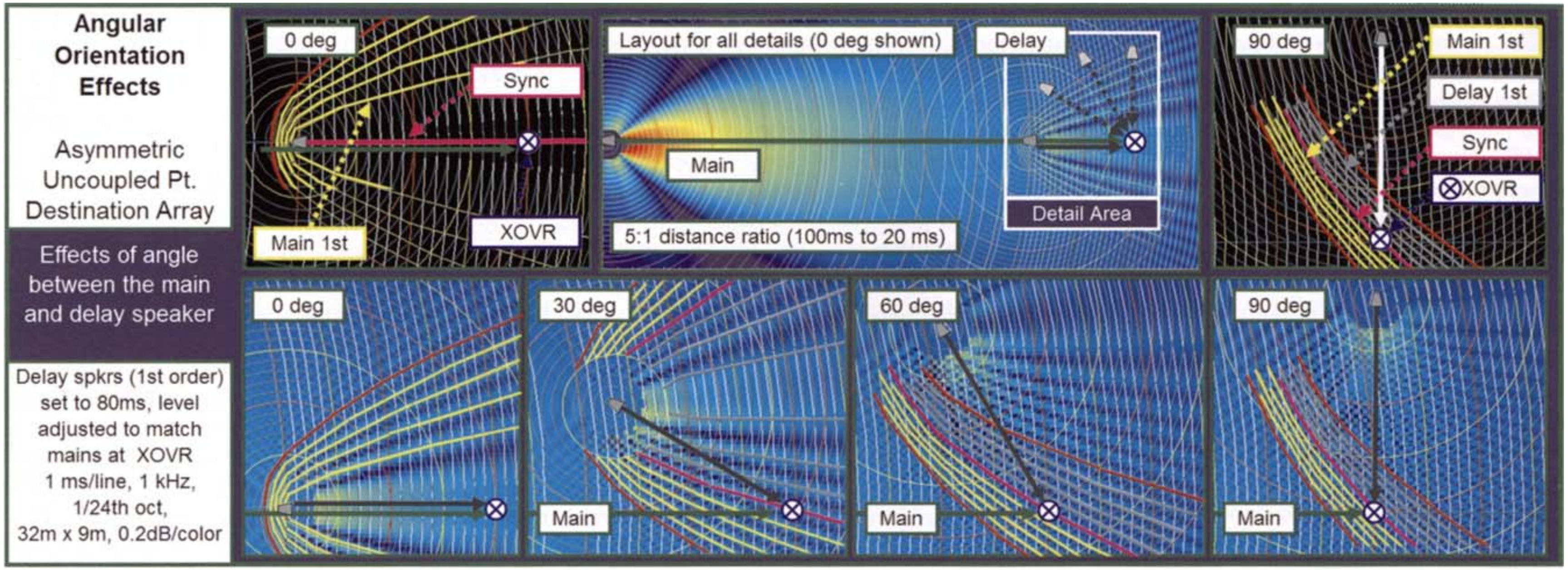


Figure 4.60 Angular orientation effects applied to the asymmetric uncoupled point destination array

the front/back level uniformity and decreased the direct to reverberant ratio (another form of reduced ripple variance) in the rear. The latter statement is true despite the variance levels revealed here. The dominant force in ripple variance at the rear of the hall would be room reflections, and our refreshed direct sound, even with the ripple shown, has a high prospect for overall ripple reduction. Finally we have reduced variance in the frequency response shape, since we can expect that the main speaker response will have HF air loss, which will be restored by the forward speaker.

Minimum Ripple Variance

The subject of ripple variance has already been covered in depth in Chapter 2. The spatial distribution of the ripple variance should now, hopefully, be further clarified by the work done in this chapter. In summary we can conclude the following: minimum ripple variance results from angular or level offset isolation (or both) mixed with source displacement. As displacement rises, our need for isolation increases.

Subwoofer Arrays

Subwoofers present a unique chance to design arrays outside of the rule set found for full-range speakers. Subwoofer arrays have two unique features that open the door of opportunity. The first is that subwoofers are in separate enclosures that cover only a three-octave range (approximate). The second is that the large wavelengths within the subwoofer range are able to diffract around neighboring objects, most notably other subwoofers. We would not think for a moment of placing a full-range speaker facing directly into the back of another such speaker (DJs excepted). With subwoofers this is an option, and a very useful and practical one at that.

Perspectives: The end-fire subwoofer arrangement proved to be extremely efficient for cancellation behind the array. Each sub was separated by 38" with corresponding delay placed on all but the rearmost sub which had no delay relative ***to*** *the other three pairs. While standing out front gain increased as each element was turned on. Standing in back revealed successive amounts of cancellation resulting* ***in*** *a very "clean" stage missing the usual low-****frequency*** *buildup. This pleasantly surprised* ***not*** *only the monitor mixer but the performers.*

Dave (Db Dave) Dennison

Individual subwoofers come in two basic flavors, omnidirectional and cardioid. The omnidirectional version, like its microphone counterpart, is not truly omnidirectional, especially as frequency rises. This is important as we will soon see. The cardioid versions consist of front and rear firing drivers that use phase offset to create coupling zone summation at the front and cancellation zone summation at the rear. These are engineered products and if done well can produce a substantial rear rejection over a wide frequency range. The advantages of cardioid steering are self-evident. Consult the manufacturers for details.

Directional arrays can be created with combinations of individual omnidirectional units. There are, as usual, various options available and we will, as usual, isolate their effects and see where it leads. We will limit the discussion to coupled arrays. Suffice to say that, with the notable exception of the cardioid subwoofer, the behavior of uncoupled subwoofer arrays fails the minimum variance test in the first degree. There is simply too much overlap. Let's take a look at the coupled options.

1. forward arrays: end-fire arrays, dual element in-line
2. lateral arrays: the coupled line source
3. radial extension: the coupled point source.

These are familiar iterations that we have studied before. We will begin with the forward arrays.

End Fire

Multiple subwoofers can be placed in a forward line, one behind the other at a predetermined spacing. Delay is added to all but the last of the speakers in a timed sequence that synchronizes the forward speakers to the wavefront passing over it from the units behind. This is the same concept as we would use to time a series of delay towers in a stadium, but on a miniature scale. The small scale gives us sufficient proximity to provide repeated coupling zone summation in the front and to stack up cancellations in the rear. The spacing/delay relationship must be carried out fairly precisely. The most common version uses a fixed spacing and delay series. Each forward speaker contains

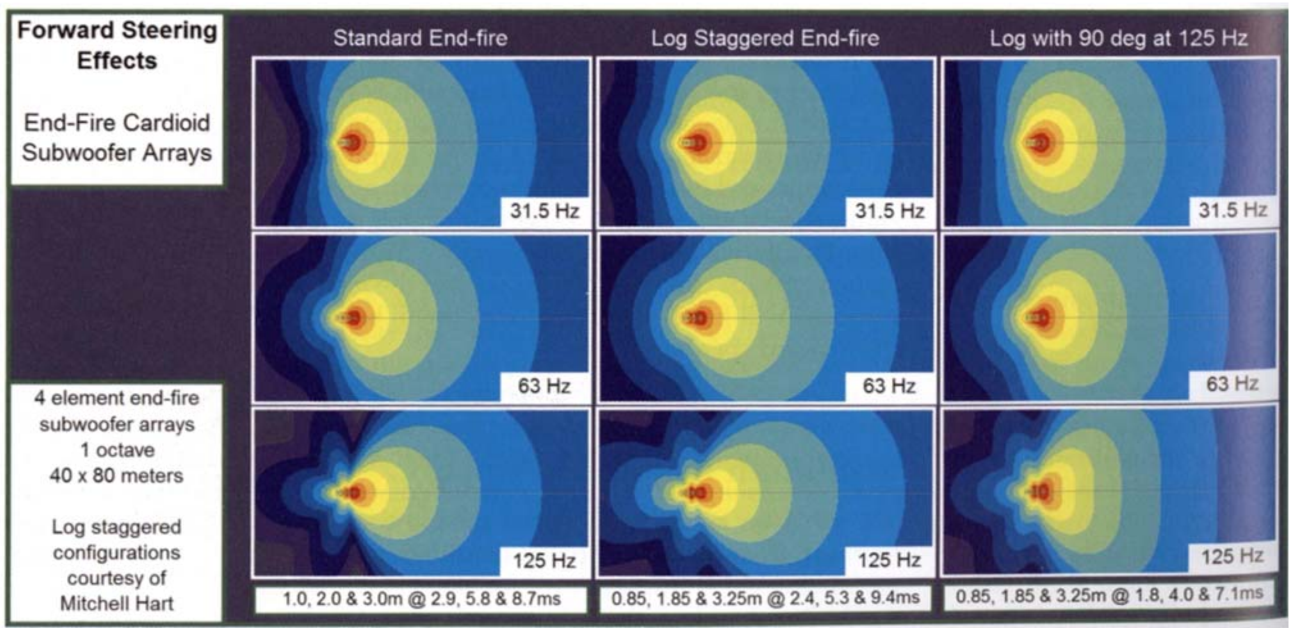


Figure 6.61 End-fire subwoofer array configurations

Perspectives: The end-fire subwoofer arrangement proved to be extremely efficient for cancellation behind the array. Each sub was separated by 38" with corresponding delay placed on all but the rearmost sub which had no delay relative to the other three pairs. While standing out front gain increased as each element was turned on. Standing in back revealed successive amounts of cancellation resulting in a very "clean" stage missing the usual low-frequency buildup. This pleasantly surprised not only the monitor mixer but the performers.

Dave (Db Dave) Dennison

Individual subwoofers come in two basic flavors, omnidirectional and cardioid. The omnidirectional version, like its microphone counterpart, is not truly omnidirectional, especially as frequency rises. This is important as we will soon see. The cardioid versions consist of front and rear firing drivers that use phase offset to create coupling zone summation at the front and cancellation zone summation at the rear. These are engineered products and if done well can produce a substantial rear rejection over a wide frequency range. The advantages of cardioid steering are self-evident. Consult the manufacturers for details.

Directional arrays can be created with combinations of individual omnidirectional units. There are, as usual, various options available and we will, as usual, isolate their effects and see where it leads. We will limit the discussion to coupled arrays. Suffice to say that, with the notable exception of the cardioid subwoofer, the behavior of uncoupled subwoofer arrays fails the minimum variance test in the first degree. There is simply too much overlap. Let's take a look at the coupled options.

1. forward arrays: end-fire arrays, dual element in-line
2. lateral arrays: the coupled line source
3. radial extension: the coupled point source.

These are familiar iterations that we have studied before. We will begin with the forward arrays.

End Fire

Multiple subwoofers can be placed in a forward line, one behind the other at a predetermined spacing. Delay is added to all but the last of the speakers in a timed sequence that synchronizes the forward speakers to the wavefront passing over it from the units behind. This is the same concept as we would use to time a series of delay towers in a stadium, but on a miniature scale. The small scale gives us sufficient proximity to provide repeated coupling zone summation in the front and to stack up cancellations in the rear. The spacing/delay relationship must be carried out fairly precisely. The most common version uses a fixed spacing and delay series. Each forward speaker contains

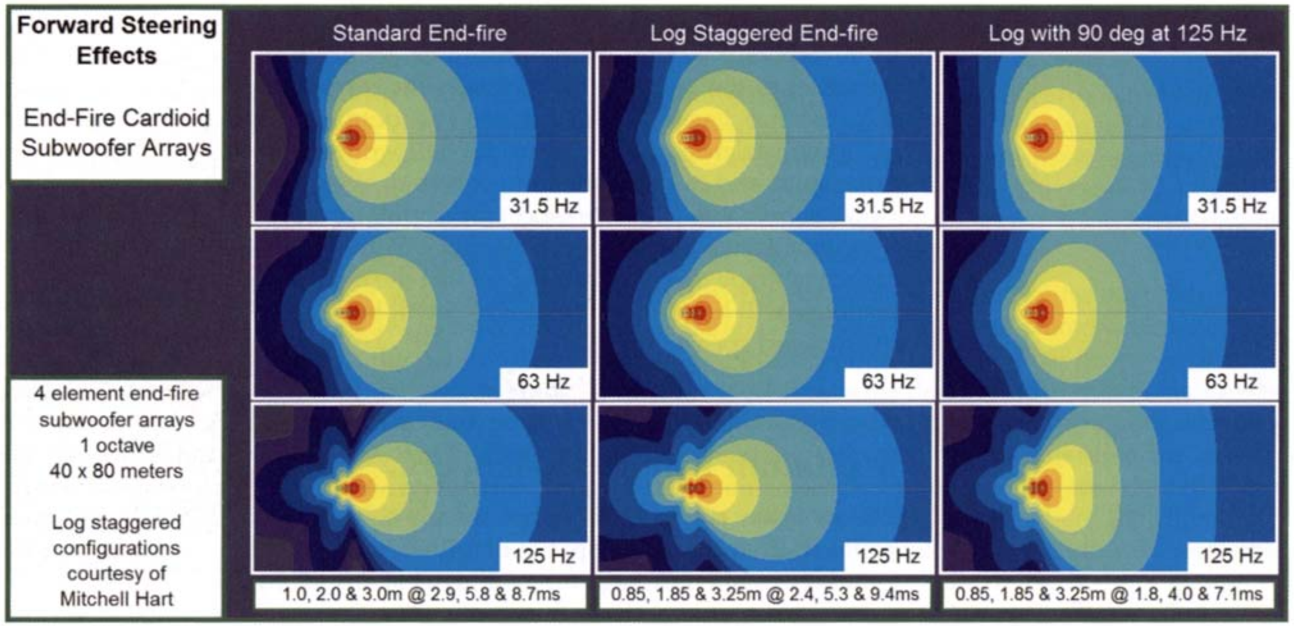


Figure 6.61 End-fire subwoofer array configurations

Perspectives: As a war story to do with steering of low frequencies, in the early 1990s John Meyer, Jamie Anderson, and myself experimented with steering subwoofer patterns. We (Grateful Dead) were playing large stadiums and trying to get thorax-rattling (make your pants legs flap in the breeze) frequencies to the upper deck at the far end of the venues. We had sixteen Meyer Sound 650's stacked on end on either side of the left/right PA stacks. That's thirty-two eighteen-inch speakers straight up in a column on either side. Phil Lesh (the bass player) freaked and couldn't deal with the sublevel on stage. In order to overcome this issue, we first decoupled the PA wings from the stage by making them free-standing. That only had a minor effect, so John came up with an idea to apply some noise-canceling theory. Using a SIM FFT analyzer, we measured the frequency response at Phil's position on stage and stored it in the analyzer. Then we flew two 650R2's in a column on either side of the stage as sidefill. Next, using a phase-correct parametric equalizer we exactly matched the low frequency of the sidefills to the stored response of the mains. Then, while looking at the phase response, we introduced delay on the sidefill speakers until the phase traces were an exact match. Finally, we swapped the polarity of the sidefills. This was a great

additional delay comparable to its forward advantage in the propagation race. There are other versions such as those with log staggered spacings and others. There is a lot of innovation going on in this field at present, so the reader is advised to investigate this as it progresses.

The end-fire array is not 100 per cent efficient, i.e. the maximum SPL of a quantity of subwoofers coupled together in the standard fashion would be greater. This cost in efficiency can be weighed against the huge potental benefits of the rear rejection: vastly improved low-frequency isolation on stage, and reduced speaker/room ripple variance in the house (see Fig. 6.61). The end-fire array is attributed to Harry Olson.

Two-Element In-line Technique

A smaller-scale version that creates a large-scale result is found in the two-element in-line technique. This consists of just two speakers spaced 1/4 wavelength apart in a forward line with a delay and polarity orientation that provides on the order of 28 dB of front to back level ratio. This method is more practical than the end-fire array which requires an extended depth, and secure area for implementation in a performance space. While this array configuration has been attributed historically to George Augsberger it was introduced to me by Mauricio Ramirez and its implementation has gained widespread popularity under his practical guidance. This method is shown in Fig. 6.62.

An important note about the forward extension subwoofer arrays. These can be used in combination with the linear and radial extensions. Of course, the subwoofer array might look like a graveyard by the time we are done.

Inverted Stack

A cardioid system can be assembled from individual speakers in a fashion similar to the fully engineered models. Some units in a subwoofer stack are reversed in orientation and polarity and cancel the back wave, while adding to the front. This is not a simple matter; the voltage

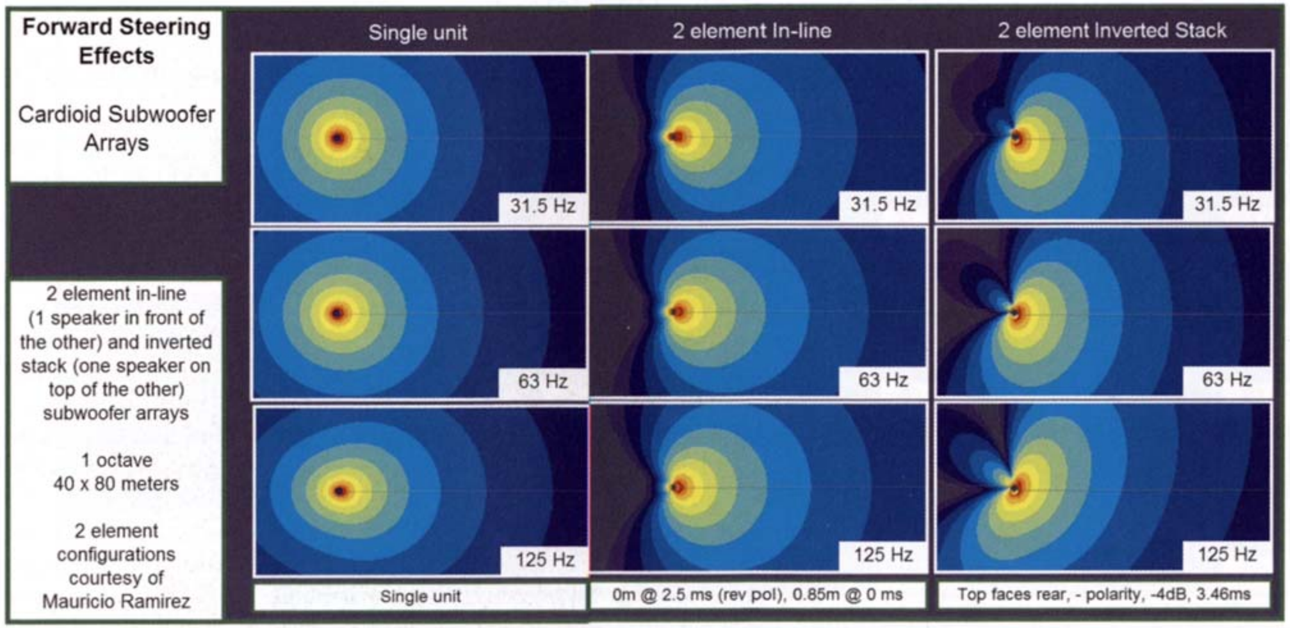


Figure 6.62 Alternative cardioid type subwoofer array configurations

range, the array will begin to uncouple at the upper end and ripple variance will become severe. It is possible to selectively move the lowest frequencies independently. The tool is an all-pass filter, which is able to provide phase delay over a limited band of frequencies without affecting others. This can be set to delay the range around 30 Hz, allowing for additional delay there without oversteering the 125 Hz range. When the delay at 30 Hz reaches a comparable number of wavelengths to the 125 Hz delay, the angles will match. This would be a 4:1 ratio in delay (just like the wavelength). How practical is this? It will cost a lot of time and money, because multiple channels of signal processing will be required to sequentially taper the delay. Additional wiring and set-up time will be needed to ensure that the signal feeds are going to the correct speakers. Then there is the calibration time. Can it work? Yes. Is there another way to do this. Yes. Aim the line of subwoofers in the direction you want the beam to go.

Essentially what the practice of delay-induced beam steering represents is a purposely misaligned spatial crossover. The position where the level offset is 0 dB is no longer the position of 0 ms time offset. While it moves the direction of the main beam, we have made no progress toward flattening the beamwidth of the subwoofers.

Beam Steering by Level Tapering

We can consider the possibility of level tapering. This was already discussed and the results shown in Fig. 6.63. The level tapering of devices without angular separation does little more than reduce the power and widen the pattern, but gets us no further toward matching the LF beamwidth throughout its range. This amounts to "shortening of the line array." We could try the combination of level and delay tapering. Will two ineffective mechanisms combine to make one effective one? Unfortunately no. Without

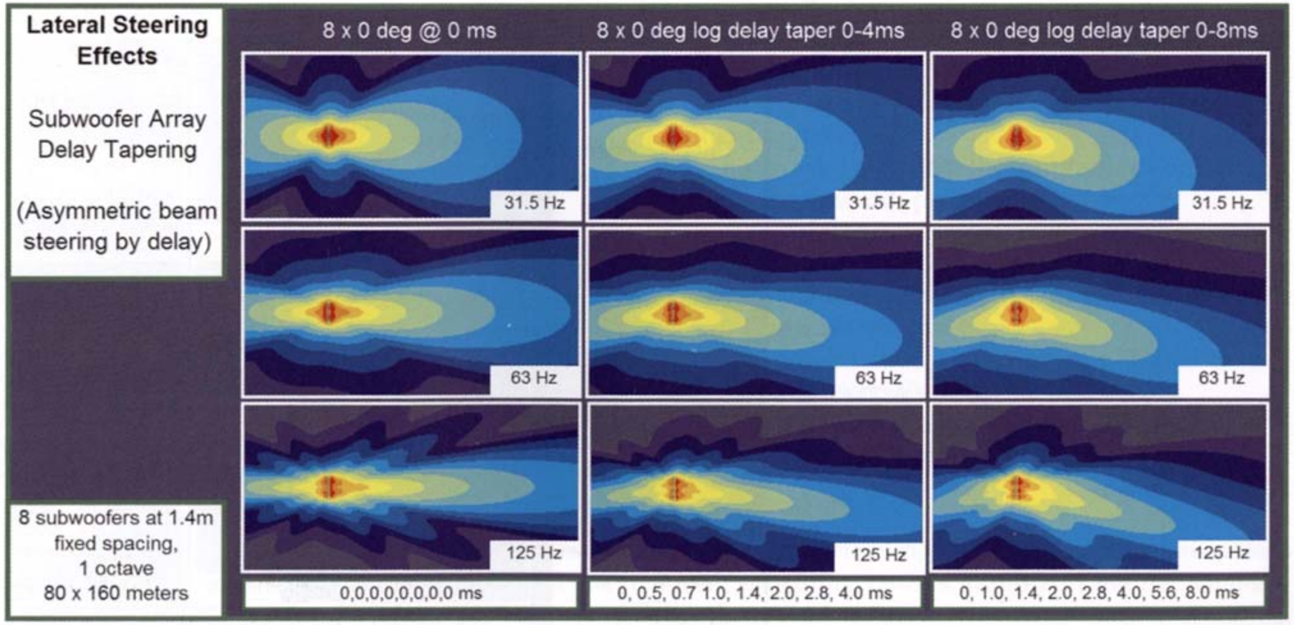


Figure 6.63 Beam-steering subwoofer array configurations

angular separation we lack an effective means to alter the combined beamwidth.

Coupled and **Uncoupled** Point **Source**

A third option is the creation of a coupled point source by introducing a splay angle between the elements. The behavior of the coupled point source is well known to us. We can take individual elements with beamwidths that narrow over frequency and combine them into an array with constant beamwidth. We can aim it where we want and can thereby fulfill both of our subwoofer missions.

Let's take a moment to consider why this is not common practice. The first answer lies in the practical world. It can be very difficult to find the space required to curve an array that is taking up valuable seating space on the floor. Stage fronts are flat, the security perimeter is flat, and on it goes. If we want the space set aside for us to do this, we are going to have to make a strong case for it. Secondly there is a widespread belief that the subwoofer response is de facto "too wide" and therefore every effort must be taken to narrow the pattern. This is only the case with fairly small quantities of subwoofers. By the time we've laid down a line of four or five typical subwoofers on their sides, the tables may have turned, at least at the top end of the subwoofer range. Longer lines exacerbate the situation. The third factor comes from the SPL Preservation Society. Anything that steers energy away from the center of the hall will reduce the dB SPL (at the center of the hall). If we splay the array outward, a level reduction at the center is a certainty. That is how we can reduce the double-stacked addition at the center. If the overriding concern is the dB SPL at the mix position then the outward splay is counterproductive, and they will have to live with the consequences: mixing from the spectrally non-representative pyramid peak.

Our treatment of subwoofer arrays is by no means comprehensive here. This area is moving forward rapidly,

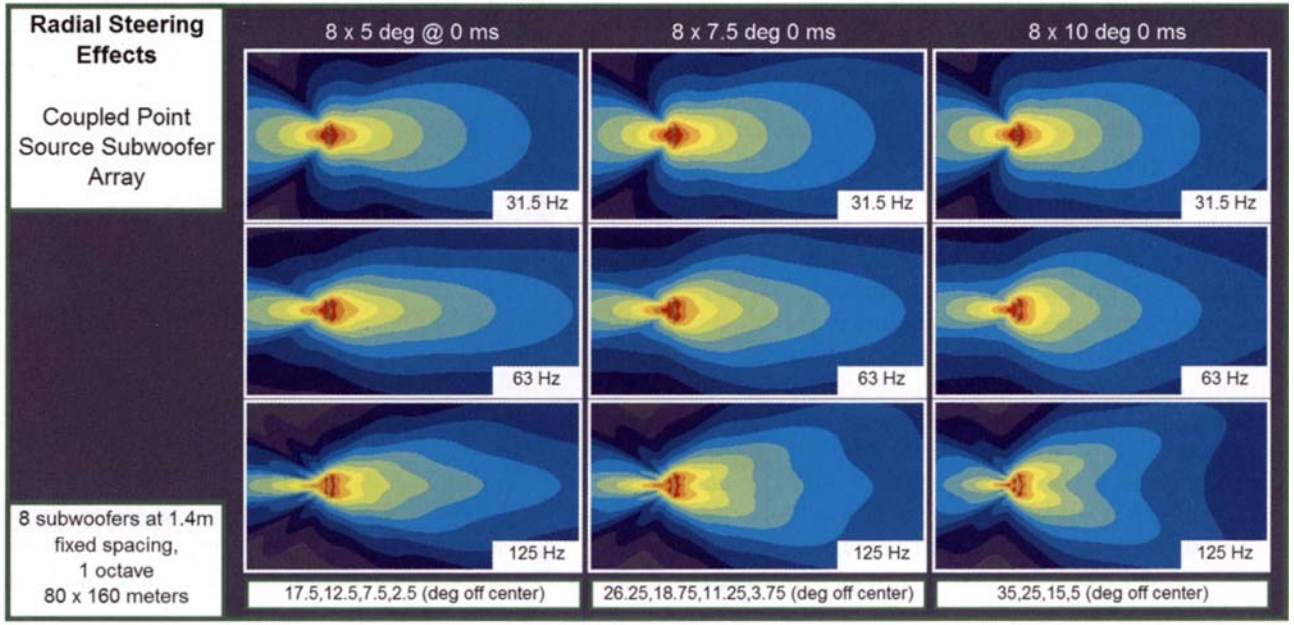


Figure 6.64 Radial subwoofer array configurations

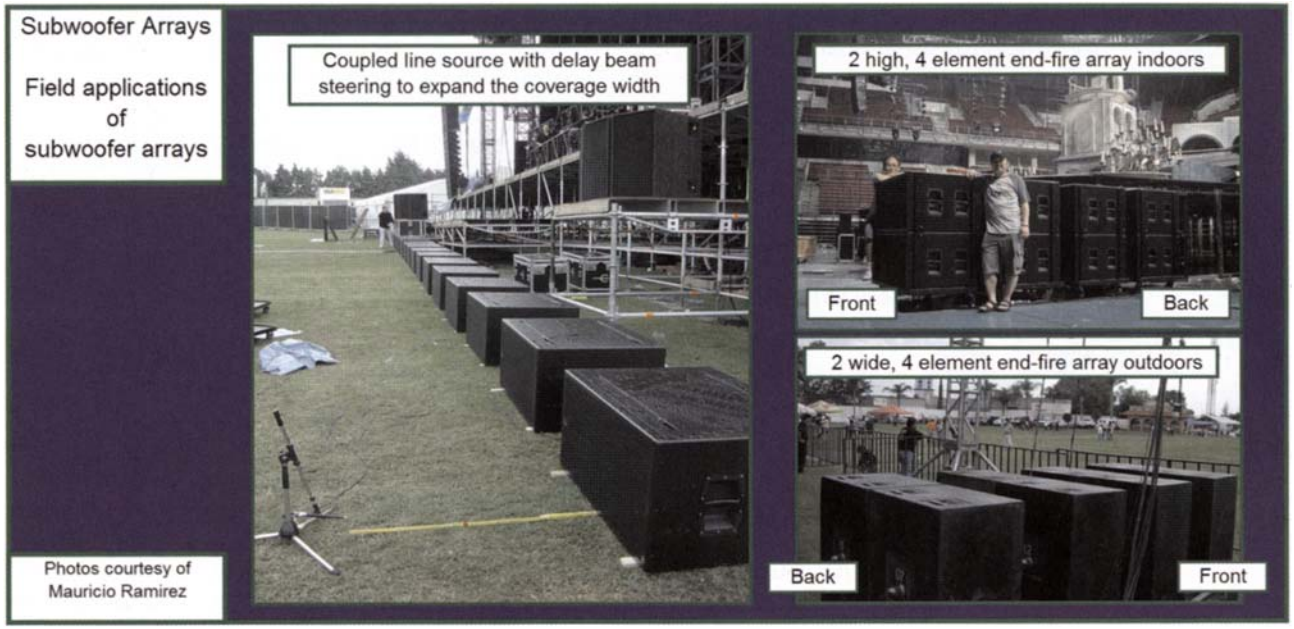


Figure 6.65 Subwoofer configuration photos

due in large part to the increased understanding in how phase affects speaker arrays. This knowledge, coupled with advanced prediction and measurement capability, give this area a lot of promise for the future.

Conclusion

At the beginning of this chapter we discussed the relationship between variance and power. We have run through an exhaustive, albeit incomplete, series of scenarios to expose the principal mechanisms at play here.

Speaker Order and Beamwidth

Regarding speaker order we can conclude the following.

First-order speakers:

- The most suitable for single-element applications with a low proximity ratio.
- The least suitable for angular overlap.
- The least suitable for combination for power.
- The beamwidth must not reverse direction as frequency rises. Once the plateau section of the beamwidth has been reached the coverage must remain constant (not widen at higher frequencies).
- Should be arrayed at or near the unity splay angle found along their beamwidth plateau.

Second-order speakers:

- The most suitable for single-element applications with a medium proximity ratio.
- Suitable for combination for power on a limited basis.
- No beamwidth reversals (as in the first-order).
- Should be arrayed at or near the unity splay angle found along their beamwidth plateau.

Third-order speakers:

- The least suitable for single-element applications unless a severe proximity ratio is found and the two-element option is not viable.
- The most suitable for angular overlap.

- Suitable for combination for maximum power addition.
- A constant downward beamwidth slope with no reversals as frequency rises.
- Should be arrayed at an angle greater than zero degrees (relative).

Maximum Power vs. Minimum. Variance

Regarding power capability we can conclude the following:

- Coupled arrays can increase concentrated power capability beyond a single element while maintaining minimum variance over long distances.
- The coupled point source has the highest ratio of increased power to variance over the widest frequency span and spatial area.
- The coupled line source can provide unlimited power but is incapable of minimum spectral variance over the space.
- Asymmetric level tapering will decrease the overall power capability at the geometric center of the array but provide the opportunity for minimum variance over an asymmetric space.
- Uncoupled arrays can increase distributed power capability beyond a single element while maintaining minimum variance, but only over a very limited distance.

Minimum Variance Coverage Shapes

The introduction also promised to reveal the limited number of minimum variance shapes (independent of the speaker/room interaction).

And the winners are . . .

- The rectangle (forward version): as found in the individual speaker.
- The arc: as found in the symmetric point source. The coupled version is range-limited only by spectral variance due to air absorption in the HF range, and by the power loss over distance. The uncoupled version is range-limited by overlap-induced ripple variance.
- The right triangle: as found in the asymmetric coupled point source.

Runners up include the various irregular shapes that can be maintained for small areas in the junction of asymmetric uncoupled systems. These include forward delay systems among others.

The Minimum Variance Menu

This yields a series of options: the building blocks for our systems, and subsystems. We have a menu to choose from that is capable of satisfying our need to fill the acoustic space. The first choice, our appetizer, is the single speaker.

Single Speakers

Minimum variance principles for single speakers:

- Position the speakers with the lowest possible proximity ratio to the audience in at least one plane. The limits to this are sonic imaging, as in the case of a high center cluster, or practical realities such as sightlines. Frontfills are a case of high proximity ratio due to sight lines. (This parameter is valid for all array types and will not be repeated for brevity.)
- Use speakers that have minimum beamwidth ratio over frequency (first- or second-order).
- Match the properties of symmetry (symmetric orientation for symmetric shapes, asymmetric orientation for asymmetric shapes).
- Match the speaker aspect ratio to audience coverage shape.
- If the speaker coverage pattern overlap onto room surfaces is too strong, then use an array instead.

Coupled Arrays

The next section of the menu is the main course: the coupled arrays. This will fill up the bulk of the space. The menu is shown in Fig. 6.66(a).

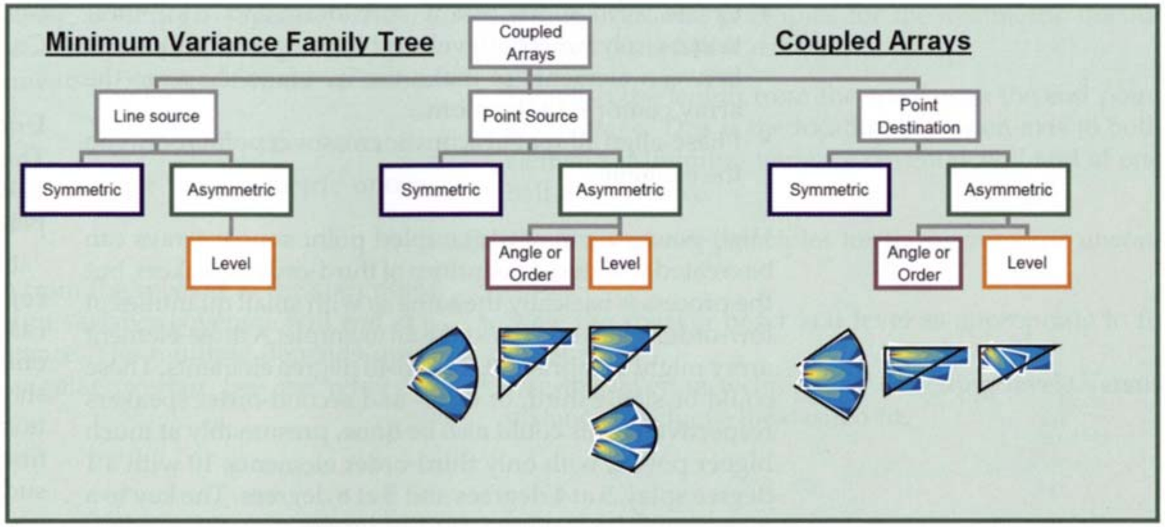


Figure 6.66(a) The minimum menu for coupled arrays

Minimum Variance Principles for the symmetric point source array:

- Fill the arc angle required to fit the shape of the audience area. The number of elements required will depend upon the element speaker order. Low-order speakers must have minimal overlap. High-order systems can be overlapped for maximum power addition.
- Beam spreading will shape the high-frequency coverage.
- Beam concentration will shape the low-frequency coverage.
- Apply the above principles until the shapes have met in the mid-range.

This combined array can now be viewed as a single element, with its speaker order defined by its combined shape, rather than individual elements. For example, an array of six third-order elements splayed at 2 degrees will combine to become a single third-order element of approximately 12 degrees. If splayed at 4 degrees they would combine to make a second-order element of 24 degrees. These two arrays can then be combined as elements in another array such as the asymmetric point source. This is how we create manageable high-power arrays out of large quantities of third-order enclosures. Individual elements are combined into symmetric sections that can then be assembled into the right triangle shape of the asymmetric coupled point source. The same principles will apply whether the asymmetric point source is made up of individual or composite elements.

Minimum variance principles for the coupled asymmetric point source array:

- Fill the triangle required to fit the shape of the audience area. The number of elements required is proportional to the proximity ratio and the speaker order. As asymmetry increases the number of elements rises. As a general rule, the number of elements should be one less than the proximity ratio. (A three-element array will cover a 4:1 proximity ratio.)
- Begin with the longest distance. This element must have the highest order of the elements, since it will have the longest throw.
- Continue to subdivide the space by layering the additional elements under the first until the shape is filled. As proximity increases, lower-order and lower-power elements can be used to provide comparable level over the area.

- Create asymmetric beam spreading and concentration by asymmetric level and/or angle relationships between elements to make the coverage shape of the array conform to the room.
- Phase-align all spatial acoustic crossover points between the elements.

High-power asymmetric coupled point source arrays can be created with large quantities of third-order speakers, but the process is basically the same as with small quantities of low-order speakers. Let's take an example: A three-element array might comprise 10,20 and 40 degree elements. These could be single third, or third- and second-order speakers respectively. This could also be done, presumably at much higher power, with only third-order elements: 10 with a 1 degree splay, 5 at 4 degrees and 5 at 8 degrees. The key to a manageable high-power third-order design is the subdivision into symmetrical subsystems that can be calibrated as single-array elements. An array with 15 different angles might seem like a great way to custom-fit the array into the space. But every time we introduce asymmetry we have created a unique spatial crossover, an unknown in the minimum variance equation. Each unknown will require several measurements during our calibration procedure. Complexity must be introduced only when warranted.

Uncoupled Arrays

The final menu selection will be the side orders — uncoupled systems to fill in the gaps in coverage, at those places where the mains are not the best fit.

Uncoupled arrays can be made of a single speaker, or combinations of previously combined arrays. In either case, these are the array elements. Uncoupled arrays can only maintain minimum variance over a limited range, and should be of the first or second speaker order. An exception to this is the asymmetric combination of a low level first- or second-order element with a third-order system, such as might be found with a main and delay.

Minimum variance principles for the symmetric uncoupled line source arrays:

- Establish the length from the speaker to the start point of coverage. Minimum variance coverage will end at twice this distance. This is the length figure for the aspect ratio. The spacing is found as the aspect ratio width

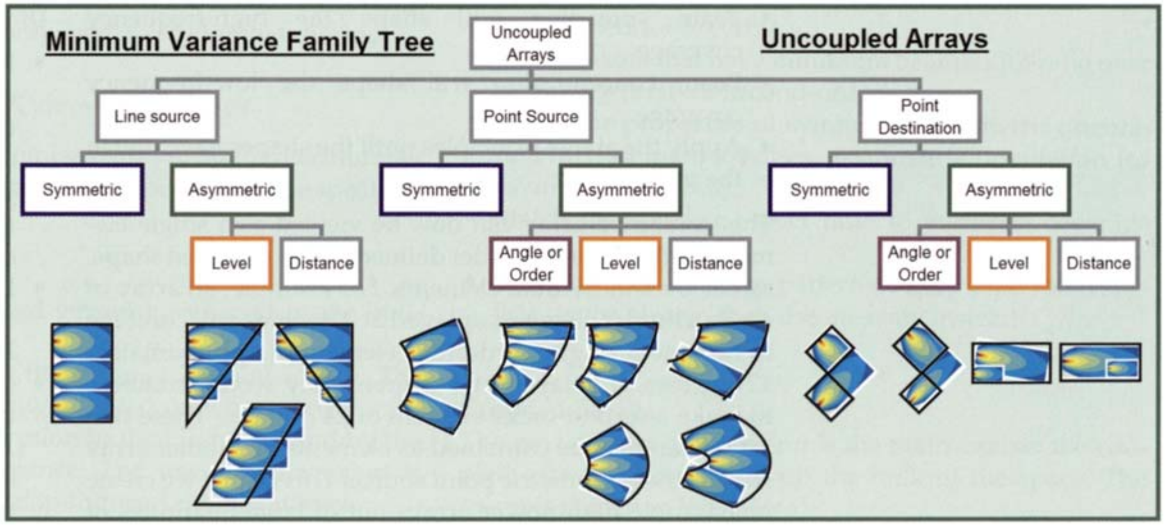


Figure **6.66**(b) The minimum variance family tree for uncoupled arrays for the given length, e.g. a 90 degree speaker (FAR of 1.4:1) needs to cover from 3.5 to 7 meters. The spacing between the elements is 5 meters. See the reference chart in Fig. 7.20.

Minimum variance principles for the symmetric uncoupled point source arrays:

- Establish the length from the speaker to the start point of coverage. Minimum variance coverage will end at a multiple of this distance. The multiple depends upon the percentage of angular overlap. See the reference chart in Fig. 7.21.

Minimum variance principles for the symmetric uncoupled point destination arrays:

- Establish the length from the speaker to the end point of coverage. This is the location that is on-axis to both elements. Minimum variance coverage will end at one half this distance.

Minimum variance principles for the asymmetric uncoupled arrays:

- Scale the speaker order and level as appropriate to fit the shape.
- Use the layering technique of offsetting levels versus distance to achieve the desired fit.

7

Specification

> ***specification*** *n. specifying or being specified; specified detail, esp. (in pl) detailed description of construction, workmanship, materials, etc., of work (to be) undertaken, prepared by architect, engineer, etc.*
>
> ***design*** *n.* **1.** *mental plan; scheme of attack.* **2.** *purpose; end in view; adaptation of means to ends.* **3.** *preliminary sketch for picture, plan of building, machine, etc.; arrangement of lines, drawing, etc.* **4.** *established from of a product; general idea, construction from parts*
>
> Concise Oxford Dictionary

Introduction

It is now time to put a sound system together in order to fill a space with sound. We will need to move beyond the simplicity of speaker order and define speaker coverage with more precision so that we can shape the sound to the shape of the room. The sound system will be a carefully constructed assemblage of parts, each of which has a defined role to play. The definition of such roles will be essential for the on-site deconstruction and reconstruction process that will follow: optimization.

Specification is the culmination of the design process. We have considered the transmission path and the challenges that we are likely to encounter en route. We have learned how the sound system will be perceived by the audience members. We understand the role that the room acoustics will play. We know how to read the plans for the space and we will use our power of prediction to choose the best locations, the array types, the speaker elements and the signal processing that drives them. We understand the shapes that are available to us for making minimum variance designs. In short, every chapter of this book has provided information which leads us to this point of ultimate decision: specification.

The Principles of Specification

Specification Defined

The end result of the design process will be a series of documents that will be used to acquire and install a sound system. The list will include drawings of the room with speaker locations, flow block diagrams, equipment lists and more. There are potentially endless levels of documentation which can be supplied for a particular installation. The specification may include conduit size, terminal strip model number and labeling, and more. For our purposes here, we will limit the specification scope to those items related to the optimized design as outlined in this book. These include the main components of the speaker system that will have critical effects on minimum variance, rather than the details such as cabling. This is not to say that these other details are less important. They are, however, well covered in other texts that focus on the nuts and bolts of system installation.

The specification arises from the particular needs of a client. The specifications are the final answers to a series of questions that flow between the designer and the client. We find out what the client wants and what they are willing to do to get it. We will need input from the client as to their expectations and they will need input from us as to what is feasible within the framework of optimized design guidelines. It is not expected that clients will tailor their expectations to what is realistically achievable. They may advocate for particular features which will make optimized design impossible. There may be architectural features that will degrade uniformity in some areas. There are certain to be cost issues.

Goals and Challenges

The goals and challenges of specification are an accumulation of those found in the previous chapters.

Transmission and summation goals of the optimized design:

- Minimum spectral variance over the space
- Minimum level variance over the space
- Minimum ripple variance over the space
- Maximum power addition over the space.

Reception goals of the optimized design:

- Spatial separation as desired by the artist. This could be stereo, surround and other multichannel strategies. Spatial separation is achieved by discrete electronic signal paths to speakers in unique locations.
- Credible tonal and spatial relationship to the source. The goal is for the listener to focus on the source, not the speakers. For the speakers to move to the background in the listener's mind requires realistic tonal quality and sonic image placement.
- Appropriate frequency bandwidth. This is related to the credible tonal relationship. Systems that will transmit voice only can remain credible without full-range extension down to 40 Hz. All other applications require full-range extension, which typically means the addition of subwoofers.
- Sufficient loudness. We must meet the power capability expectations of the listeners for the given program material.
- Desirable acoustics: tonal balance. Presence of an appropriately scaled diffuse reverberation field. Freedom from discrete echoes, resonances or other acoustical deficiencies.
- Clarity: we must have a sufficiently high direct-to-reverberant ratio for listeners to perceive the signal clearly.

Challenges to the optimized design:

- Unfavorable architectural acoustics for amplified sound. These qualities were discussed at length in Chapter 4.
- Acoustic path obstruction. Anything blocking all or part of the direct sound path from the speaker to the listener. These include a huge variety of possibilities such as scrims, set pieces, lighting instruments and steel, to name just a few.
- Poor speaker positions and focus angle: there are unlimited varieties of poor speaker positions. Any time we think we've seen it all, we find something worse. The position may create increased variance, echo perception, misplaced sonic image, low gain before feedback, or other complications. A desirable location and/or focus angle for one category may degrade another. This is an area of constant compromise.
- Insufficient coverage: gaps in coverage due to lack of resources and/or poor positions and/or focus angles.
- Excessive coverage: overlaps in coverage due to excess resources and/or poor positions and/or focus angles.
- Insufficient system subdivision: signal processing and amplifier channels must be available in sufficient quantity to provide spectral and spatial division as required for minimum level variance. Independent signal processing channels will be required to minimize spectral variance (equalization) and ripple variance (level and delay).

Specific Questions

In the end we have a client, a room, a tour, something that needs a system. Our design will be tailored to the particular needs of the client and the space. To find the right answers we need to ask the right questions. Here are some of the relevant questions to the design process.

Question 1: How many channels?

The sound system design is made up of a combination of channels of sound transmission. Each channel fills its reception space with unique program material. The reception space, the listening area, is only a subset of the room, since the listeners are not filling the whole space. In the horizontal plane the listening area and the room plan will have close correlation. In the vertical plane the listening area is confined to the lines on which the seats are placed.

A single channel (monaural) system fills the entire listening area. A stereo system contains two channels, each of which fills their own areas. They may each fill the entire listening space or they may share only a reduced portion where the stereo panoramic imaging is viable. In such cases, the left channel coverage area is a subset of the overall listening area, as is the right channel.

The number of channels is limited only by our imagination and someone else's budget.

Question 2: Where is the channel's desired sound image?

Each channel has some role to play in the production. The role of the left channel in a stereo system is self-explanatory. Other channels need a location as well. The mains, surrounds, and other sound sources all need a home. We need to know where the client wants the sound image to appear. Only then can we select the best positions to achieve this.

Subsystems share the same basic sound image location, although it may not be exactly the same. For example, a center cluster and its under balcony delay subsystem both share a common desired sound image at center stage. The practical realities of where the delay speakers can be placed might prevent some of the speakers from aligning exactly to the same image point. The general goal is shared, even if it cannot be implemented precisely.

Question 3: What are the specific practical limitations?

Every hall has its own set of peculiar circumstances. At a particular symphony hall the system needed to be able to disappear without a trace in less time than the minimum union call. Knowing this in advance prevented a potential waste of time on solutions that would not fit these criteria.

Some examples of job specific limitations include:

1. Landmark status concerns: no mounting on the plaster fascia will be allowed.
2. Invisibility: all speakers must be hidden behind scrims.
3. Sight lines: no speakers allowed below the proscenium.
4. Scenic elements: speaker locations must dodge set pieces and structural members.
5. Must be Brand X speakers: this theater is sponsored by the X Corporation, or the client is the dealer of Brand X.
6. Must be portable: multipurpose halls and roadhouses may need to remove and replace the house system on a regular basis.
7. Must fit in the truck: touring companies seem to have strong opinions about this one.
8. Must be easily rigged: same as above.
9. Must be dummy-proof: same as above (X10).
10. Must not cause water to pour into the cruise ship (true story).
11. Must be a line array even if it is the wrong tool for the job.

Question 4: Where are open microphones in the channel's coverage area?

We must consider the implications of microphone locations with respect to our speaker location. In most cases the microphone positions take practical priority over speaker locations. Fine-tuning of positions may be negotiable, but for the most part we know where the main mics will be. Our speaker placement must be mindful of providing sufficient isolation to provide the acoustic gain required for satisfactory level and stability in the house. This is an inexact science, and involves ongoing operational aspects. Any system with an open mic can be sent into feedback eventually, if turned up loud enough. The principal considerations are microphone and speaker placement, microphone and speaker directional control (especially the rear sector), distance from sources to the mic and the levels expected in the house. This is, of course, evaluated on a speaker-by-speaker basis (see Chapter 4).

Question 5: What is the channel's program material?

It is tempting to consider program material strictly in global terms. If the application is rock and roll we know that we need a full-range high-power system. While this is true of the left and right main channels, it is not necessarily so for the frontfills, which may be limited to voice only. The program material must be evaluated on a channel-by-channel basis.

Program material provides us with information in two principal categories: frequency range and expected level. The frequency range factor can be boiled down to the type of question we might hear at a fast food franchise: "Would you like subs with that?"

The expected level factor relates directly to SPL. The SPL needs, however, cannot be determined by a single simple number. There are very critical aspects regarding the distribution of SPL over frequency for particular program materials. The SPL needs of almost all forms of popular music are tilted strongly in favor of the lows and low midrange frequencies. This is fortunate since our ability to provide substantial power addition while maintaining minimum variance is also heavily stacked toward the low end. The minimum variance techniques of HF isolation and LF overlap create a combined SPL that will vary over frequency. We should never be surprised at the lack of applicability of a single number parameter in audio. After all, the frequency range encompassed by any single number is usually 600:1.

The program material informs us as to what type of array options are available to us. If the power requirements are high (such as pop music) then coupled arrays must be selected in favor of uncoupled distribution. If power is king, then bow to the king and move on. Minimum variance is rendered academic if a system is run to overload. Reference charts for SPL by program material are found in Fig. 1.14.

A secondary consideration is the issue of stereo. If the predominant program material is pop music, every effort needs to be made to provide stereo in those areas that can perceive it. No amount of monaural perfection will satisfy someone who really wanted stereo. The key to success is to minimize the overlap to those areas where stereo perception has a reasonable probability of occurring (see Chapter 3).

Question 6: What is the channel's range of coverage?

How far does the sound need to go? What is the maximum distance we will expect this system to throw? This distance will be used to estimate the power requirements of the speakers in the given channel. The program material provides the SPL target. Now we will need to deliver that level over the shape of the space in spite of the distance-related losses. A reference chart for SPL loss over distance by program material is found in Fig. 1.41.

Question 7: What is the channel's coverage area shape?

The shape of the coverage area could be anything: a simple symmetrical shape, a simple asymmetrical shape or something complex and multifaceted. We have seen the minimum variance shapes we can create with single speakers and arrays (Chapter 6). If the coverage shape fits one of our minimum variance shapes, the task is relatively simple. If not, we will subdivide the space into portions which can be individually covered and merged into a minimum variance whole.

Question 8: What are the acoustic properties of the channel's coverage area?

This is related but not equivalent to the acoustical properties of the room in general. Each signal channel has a distinct coverage area which may incorporate only a portion of the overall space. This coverage area may then be subdivided into subsystems for the control of specific zones of coverage. The acoustical conditions for that channel are evaluated from the perspective of the subsystems that play the dominant roles in the area. The main system will require a separate evaluation of the room acoustics from its perspective. Areas covered by fill speakers contain a mix of the mains as well. Their local acoustic qualities will need to be evaluated both separately and in combination. An illustrative example is the under balcony delay. It is the unfavorable acoustics seen by the main speaker that create the need for the delay speaker. The delay speaker shares its coverage area with the mains but sees the room surfaces from a very different perspective, one that is much more favorable to minimum variance. The combined results will include the combined acoustic effects.

Question 9: What is the budget?

I hope you didn't think that I know any better than you how to get the client to tell us how much they are willing to spend. Sadly, this often the first issue discussed by the client, yet without knowing the answers to Questions 1-8 this cannot be answered!

Perspectives: It's not a lightical, it's a musical.

Martin Carillo

Specific Answers

The design answers run from global to local decisions. Most design answers are separated along system channels: the left, right, surrounds, etc. Each of these will have separate speakers, signal processing and acoustic considerations. They will each in their own way comply with the practical concerns and share the common acoustic space. They will each draw from the common resources, most notably the budget.

Answer 1: Speaker Locations

Each signal channel has a basic location dictated by the practical concerns and desired source location. Subsystems and array elements for that channel are oriented to create the same sound source location. For example, the surround left channel speakers will be distributed from front to back along the left side of the hall at the upper and lower levels. From all perspectives the sound source will appear at the extreme left side. Another example would be a vocal channel that is intended to source to the singers onstage. Such a system will require multiple locations in order to preserve the sonic image. The back half of the hall can be well served by a central cluster, since the vertical sonic image distortion decreases as the floor level rises. The front areas are better served by infill and frontfill speakers at a low vertical position.

Answer 2: Speaker Array Types

Once we know the location we can look at the listening area from this perspective. This will inform us as to what type of array will best fill the shape. Long and narrow shapes will use coupled arrays. Wide and shallow spaces will use uncoupled arrays. The proximity ratio will help us to evaluate the degree of asymmetry that must be overcome to achieve minimum variance. Both the horizontal and vertical plane coverage shapes will be evaluated. Another indicator of array type is program material. High power program will require coupled arrays even for shapes that could otherwise be covered by a single speaker.

Answer 3: Speaker Models

Once we know the type of array we can discern the models of speaker that will comprise it. The required power and coverage angle will be the primary drivers in this decision. The coverage shape of coupled arrays can achieve a given power capability by being filled with a single high-power speaker, or multiple units of lower power. For uncoupled arrays the choice of model will depend upon the power requirements and operating distance. Longer distances will use a narrower speaker and/or wide spacing. Shorter throw distances will use wide speakers and/or closer spacings.

Answer 4: Speaker Quantities

The quantity of speakers is interrelated with the model choice. A single 90 degree speaker may meet the coverage

Perspectives: Know your client. I once spent one and a half days placing and aligning both main and near-field monitors in the control room of a large studio complex. The studio room was well equipped with a grand piano, a very spacious drum booth, several vocal booths, etc., and was big enough for a sixty-piece orchestra. I re-mounted and rewired outboard equipment into lower profile racks to avoid reflections near the. mix position and even time-aligned the near fields with the mains so that engineers could fade from one to the other seamlessly.

The client was amiable enough but implied that I shouldn 't worry if things weren't perfect. He had a very relaxed attitude considering I was aligning a studio that was clearly represented a major investment. As was my normal practice, I aligned the monitors for a wide "sweet area" and conducted listening tests using CD tracks that I knew well then, once satisfied, approached the client for a selection of secondary master tapes so that we could fine-align the system to his particular wishes.

The first track was a well-known TV advertisement for frozen peas followed by one for washing powder and another for toothpaste! The client then explained that most mixes would be finalised using a typical domestic TV as a reference monitor. The main and near-field monitors would be used to check for hum, clicks and high-frequency crosstalk though ...

Jim Cousins

requirement but fail the power needs. A 3 X 30 degree symmetric coupled point source array may be sufficient in both categories. For unity splay angle arrays the quantity required is the overall angle divided by the individual element coverage angles. If additional power is required we will need to add more elements and overlap the coverage. For overlapped arrays the required quantity is the coverage angle of the shape divided by the splay angle.

In the case of uncoupled arrays the quantity is dictated by the element aspect ratio and the coverage depth and width.

Answer 5: Speaker Angles

There are two levels of decision-making here. First is the orientation of the array. Second is the orientation of the individual elements in the array. Both horizontal and vertical angles will be considered independently.

Answer 6: Acoustic Treatment

Requests for acoustic treatment will be determined by the extent of the threat posed by the untreated surfaces. The most pressing concerns are surfaces that are on-axis to speakers at angles approaching 90 degrees (the back wall, for example). Reflective surfaces that return on-axis energy toward the center or front of the hall or the stage are the top candidates for absorptive treatment. The need for treatment of such surfaces is proportional to the surface area and reflectivity.

Answer 7: Signal-Processing Channels

Speakers that transmit a unique signal will require unique signal processing. This could be the result of different signals channels such as stereo. It could also be for unique level setting, equalization or delay settings for subsystem components. A general rule is this: any form of acoustical asymmetry between elements requires some form of compensatory asymmetry to match the responses. If something is different in the physical world — different model, different angle, different throw distance, etc. — there should be separate processing.

Answer 8: Cost

A small fraction of the video projection system.

Compromise

The process of sound system design is always one of compromise, but that does not mean we need to compromise our principles. There is no job where we will have unlimited budget and placement options. It is our job to inform the client of the quality implications of the limits placed on our design. Then all parties can seek to reach a suitable compromise or a creative solution based on the weight of the other factors involved. There will be times when we have to make decisions that will have both positive and negative effects. How can we make these choices? Let's take some guidance from others.

TANSTAAFL

Science fiction author Robert Heinlein created the TANSTAAFL principle in his 1966 novel *The Moon Is a Harsh Mistress.* This term (pronounced "t-ahn-st-ah-ful") is an abbreviation for "There ain't no such thing as a free lunch." The concept is informative in all walks of life and no less applicable to our decision-making process in system specification. Every decision we make comes with a cost, some form of tradeoff. This theme will be central to this chapter, as we will attempt to maintain a focus on the nature of the tradeoffs we are making. Let's begin with a simple illustration: the overlap of two speakers in a point source array. Increased overlap yields higher power capability for narrower combined coverage angle and higher variance. Decreased overlap yields minimal power addition, expanded coverage angle and minimum variance. We must evaluate which is more important, and this is often a difficult choice. If the power capability is too low the system will be run into hard overload. This is a far worse result than

Perspectives: I am not sure there is any "best" way to do anything. All solutions seem to be compromises in some way. The trick is to minimize the compromises in the total system, and that takes in all the elements in a performance or event space. Sound is only one part of a much larger system.

Alexander Yuill-Thornton II (Thorny)

Perspectives: While it is nice to have a "perfect" system, with the best of everything, this is not always possible, or maybe necessary. Large systems in large spaces are hindered by long reverberation times and slap reflections. Spending lots of time auditioning very fine microphone preamps or other processing equipment might be better spent on refining the basic system design or doing a better job of directing the speakers and optimizing the system. Toys are nice, but can be easily swamped by the nature of the beast.

Alexander Yuill-Thornton II (Thorny)

frequency response ripple variance due to overlap-induced comb filtering.

Sales and marketing departments are trained to suppress our consideration of TANSTAAFL. Each day an assortment of free offers are presented to us. All have strings attached. We must not lose sight of this guiding principle. We must be realistic about the consequences of the choices we make. If it seems too good to be true. It is.

Acoustic Triage

The medical community has developed a resource allocation and prioritization system that is implemented under emergency conditions. This system is known as "triage" and distributes resources based on a combination of probability of success and degree of need. The triage system seeks to prevent excessive allocation of resources for lost causes, and prioritizes those with minor problems at the bottom of the list. There is no point in using up the entire blood bank and all of the surgical facilities for a single doomed patient when these precious resources can be used to save hundreds who still have a fighting chance. On the other end it would be imprudent to treat those with bruised toes ahead of those who are in critical yet recoverable condition.

In our acoustic triage model we seek to identify areas that will require special care. We will seek to create and spread our resources for maximum benefit to the largest areas. We seek to prevent solutions for one area that create problems for others. If this is the case, we will use the principles of triage to help our decision-making. An example: there is a small distant balcony area. This can be covered by upward aiming of the speakers in the main cluster. This will result in excellent imaging for the balcony seats. The presence of a 10 meter tall mirrored wall directly above the top row gives rise to concerns about reflections that will disturb the rest of the room and the stage. Triage principles do not require us to abandon these seats but rather to seek an alternative solution. Over-balcony fill speakers can cover this area with minimal risk to the surrounding areas. The imaging will suffer upstairs but this might be considered a reasonable tradeoff.

The principles of minimum variance, TANSTAAFL and triage are the philosophical underpinnings of our specification strategies. They combine to create an approach that seeks the maximum benefit for the majority, is realistic about the tradeoffs required and is committed to allocation of the resources to the areas where tangible benefits can be assured.

Channel/System Types

There are several roles that channels can play. The most straightforward is the role of a main system such as the center cluster which reinforces speech. Stereo mains in a concert-setting play a similar role with the added dimension of horizontal panorama for some listeners. In either case this system is charged with the principal responsibility of coverage for the majority of the program material. These will be termed **main systems.**

A very different part is played by speakers fed by the surround channels. They might be used only sparingly for a spatial enhancement or special effect, hence the term **effects system.** Such speakers provide a strictly auxiliary role, independent of the main system. Their relationship to the main systems at any given moment is entirely under artistic jurisdiction. It is assumed that the relationship between the main and effects systems (or effects and other effects) will fail the stable summation criteria outlined in Chapter 2. As a result, we will design and optimize these types of channels as unrelated roommates that share the common space.

The design requirements for each channel will be assessed separately. The needs of the left surround speakers are distinct from those of the left mains, or any other channel. The components dedicated to the transmission of a particular channel comprise a **system.** A system can have a single speaker or any number above that. There is an important distinction between systems and **subsystems.** A system can comprise multiple subsystems driven by the same source channel. The job of the system is to fill the space. Not all spaces are amenable to optimal coverage

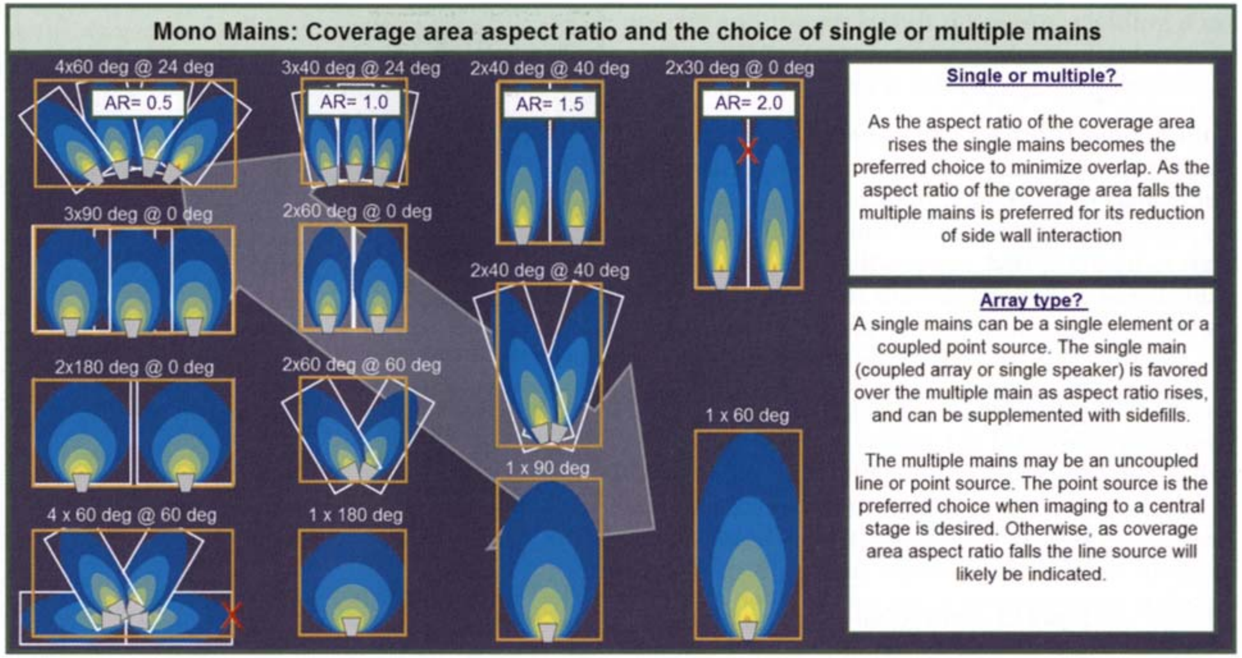


Figure 7.1 Mono system channel design considerations

by a single speaker or a fully symmetric array alone. Such spaces require subdivision of the coverage into separate speakers with adjustable parameters. Subsystems share the space, and the goals of the particular source channel.

Mono

A monaural system is charged with filling the entire listening space with the primary transmission signal. Straightforward and simple, this mainstay system can include an unlimited number of related subsystems. All subsystems share a common desired sonic image location and relative level.

General design principles for monaural systems (with stage source):

- Center cluster (coupled point source) is the standard. Alternative is the dual mono system located on the left and right of the stage.
- Related subsystems should reconcile to a shared approximate point source to cluster(s) and/or performers.
- Subsystems should be scaled as appropriate for their relative throw distance.
- Subdivide when appropriate in order to level compensate for different distances, to minimize reflections, improve sonic image or reach obstructed areas.

Stereo

Stereo systems are a hybridization of two separate channels and two matched channels. The degree of separation of the input signals is an open variable under artistic control. The stable summation criteria cannot be met under these circumstances and therefore the design and optimization must proceed with the channels seen as unrelated. Each channel of a stereo system can include an unlimited number of related subsystems. The benefits of stereo can be evaluated by the TANSTAAFL and triage principles.

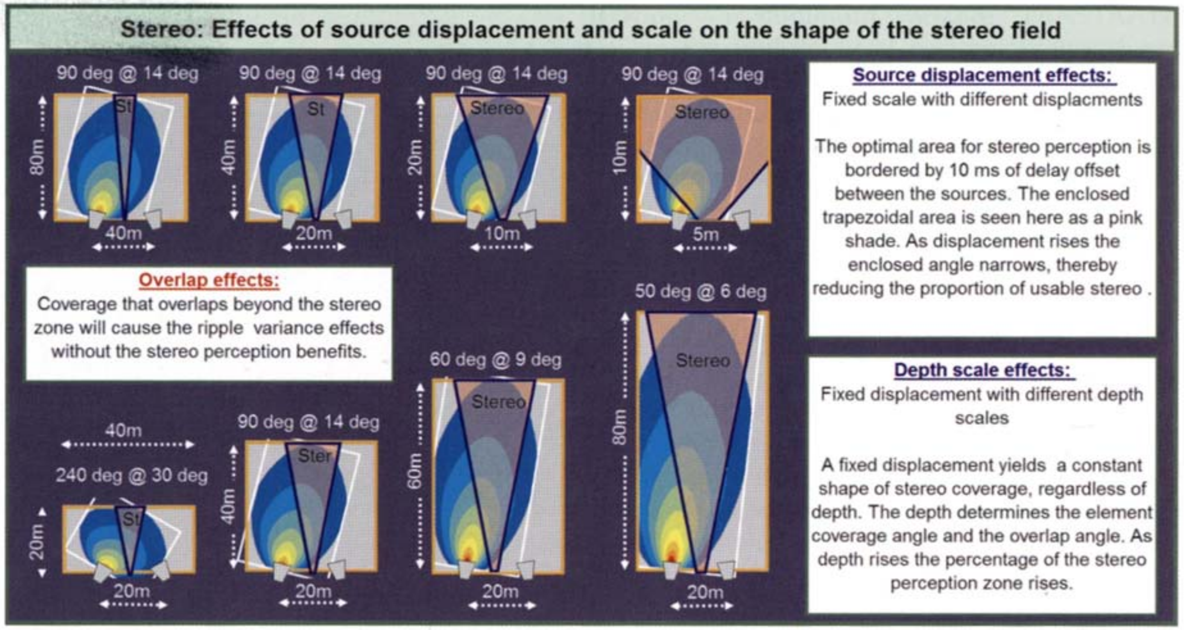


Figure 7.2 Stereo system design considerations

Stereo requires overlap from separated sources. This is certain to increase variance. After all, stereo is a form of desirable variance. The benefits to the central area must outweigh the costs on the sides. The triage principle allows us to evaluate when to stop the bleeding on the sides and let them live with a single clear channel.

General design principles for stereo systems:

- Horizontal overlap is required in the center seating area.
- Be mindful of the non-scalable time offset considerations regarding stereo imaging. Keep the stereo overlap in the zones within a 10 ms time offset as much as possible.
- Stereo auxiliary systems are rarely worth the effort and expense and loss of coherence. Examples include stereo under balcony delays, alternating left/right schemes in the round etc.
- If the stereo spacing is too narrow it is not worth the trouble. Examples of this include center clusters that are split into stereo (yes, people have done this !).

Surround

The role of surround systems is to provide perimeter sound sourcing. Left and right surrounds contrast to their stereo counterparts just as headphone stereo perception differs from our living room. When listening to stereo in a room, the panoramic horizon remains in front of us at all times, even if the signal is panned to just one side. In surround sound, like headphones, a left side signal is truly arriving from that direction. Surrounds can be used to purposefully move our attention to a perimeter location or to envelop listeners inside of a fully multidimensional soundscape.

The surrounds are typically spaced along the side and rear walls as an uncoupled line source array in the horizontal plane. In multilevel halls we will find additional sets of surrounds on each level to cover path obstructed areas.

From the design point of view there are several notable aspects. Since surrounds are located on the side and rear walls, it is an absolute certainty that we will have a very high proximity ratio between the nearby seats and those at the center and opposite side. In our ideal world we can reach deep into the hall and provide localization to a large area without overpowering and annoying the nearby seats. It will be to our advantage to move the speakers up as high as practical, thus reducing the proximity ratio. This has its limits, sometimes physical ones such as ceilings and balconies. Other limits include excessive vertical image lift and reverberation. In many cases we will employ excess pink shift to our advantage. When practical limitations leave us with far too much proximity ratio we can use pink-shift-induced "false sonic perspective" to give us more breathing room with the local listeners (see Chapter 3). The technique is to purposely aim the speakers so that the nearby seats are out of the coverage pattern. Additional help can be found in a highly directional speaker, provided it is aimed in the direction of the seats at the opposite side of the hall. This is one application where a third-order speaker is well suited to solo coverage. Why? The excess pink shift in the local area gives the listeners the perception they are more distant from the source. As the distance increases we move more on-axis and the pink shift decreases, yielding a net offset of effects. If the sound is going to be too loud in some locations, it is preferable for it to have the HF rolled off.

It is important to remember that each surround channel is charged with covering the full space, not to subdivide coverage as we would do with related subsystems. Surround speakers will provide the most uniform *experience* if they are aimed far across the space. Left is aimed at right and the rears are aimed at the front row. The worst thing we can do is aim the left surrounds down to cover the left and the right to cover the right, etc.

The reason a typical surround sound array is an uncoupled line source along the wall is due to the high proximity to listeners on the sides and to provide a sound image that appears along the wall. A coupled point source would be a poor choice here since it would provide pin-point imaging, rather than a general direction. The range limits of the uncoupled point source leave us to consider a design dilemma. TANSTAAFL: if the spacing and number of

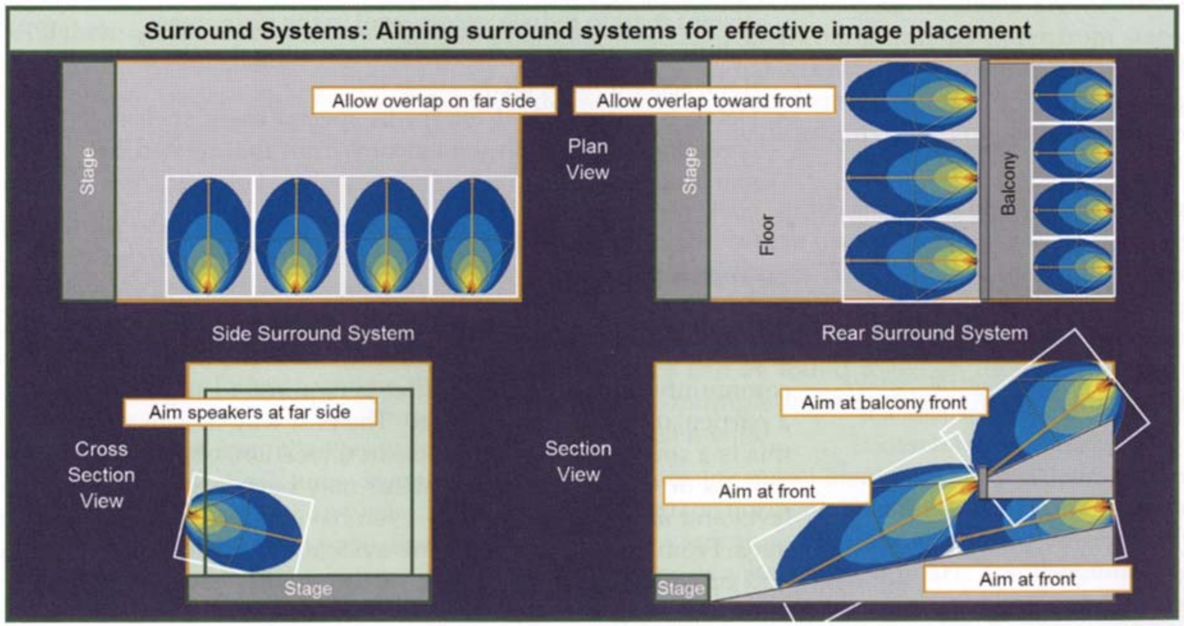


Figure 7.3 Surround channel design considerations

elements is sufficient to cover the nearest listeners, it is certain to be highly overlapped on the opposite side. If we design for minimum overlap on the opposite side, large portions of the hall will be in the gap coverage zone. Triage: design the coverage range to end at the midpoint of the listening shape. (The unity crossover line is found at half of that distance.) Those listeners beyond the middle will have excess overlap (ripple variance). Those listeners closer than the unity spatial crossover will have excess pink shift (spectral variance).

Each seating level will require local surround sources.

General design principles for surrounds:

- Uncoupled line source comprising first-order elements is the standard choice.
- Subdivide when appropriate in order to level compensate for different distances.
- Minimize proximity ratio as much as practical.
- Space the speakers (horizontal) by the aspect ratio method for half of the coverage depth. Expect overlap coverage in the distant areas.
- Aim speakers (vertical) by the asymmetric method (most distant seat) to reduce excess level in the near areas.
- Subdivide coverage for different seating levels (first floor, second floor, etc.).
- Delay is not required except in special cases such as speakers mounted on the balcony front that related to rear wall surrounds.

Perspectives: These days line arrays are very fashionable. People are using them for everything. Even in situations when they are the worst of tools. Don't get me wrong. I too like line arrays, but to use them in old Italian (horseshoe) theaters where the requirement is short throw and extended vertical coverage is at least a questionable choice. These places usually require less powerful crystal clear multipoint designs.

Miguel Lourtie

Source Effects

It is sometimes desirable to localize to an exact point in the room for a moment. This is standard fare in the theatrical community, where a sound source is required to augment a particular cue on (or off) stage. The best way to achieve this is a source speaker at the desired location, often hidden inside set pieces. Such systems may have any power level and array type, and may even have related subsystems. From our point of view this system will be designed and optimized as a stand-alone entity which covers the entire listening area.

System Subdivision

The term subdivision here connotes the creation of subsystems that carry copies of a shared original source waveform. The purpose of subdivision is to tailor the response for the particular coverage shape. If this is a match for the shape created by the speaker system when all elements are identically driven, there is no reason to bear the additional expense of subdivision. If the array coverage shape is not compliant to the listening area shape we will need to introduce separate level controls.

System subdivision concerns four types of differentiation: speaker type, level, delay and equalization. If different speaker models covering the same frequency range are used, it is mandatory that the remaining parameters be separated. If matched speakers are driven at different levels, then delay and equalization should also be separated. Equalization is the last of the parameters to separate, and is rarely found in isolation.

System subdivision is a countermeasure to asymmetry and complexity. How much asymmetry is required before subdivision is required? Let's be practical and apply TANSTAAFL. Every subdivision has the costs of the material, installation, ongoing maintenance, and calibration time. It also opens up the door for human error. A good practical guideline is the 1.4 rule, also known as 3 dB. If two same model speakers have throw distances that differ by more than a ratio of 1.4, subdivision is indicated. For example, a three-element cluster with outer elements that have a different throw distance than the center unit can be evaluated this way.

Once level asymmetry is introduced, delay and equalization follow suit. We have to assume that a speaker that is run at a different level is operating at a different range of coverage length. Why else would we change the level? This is a different acoustical situation. Separate EQ is indicated. If the level is offset between two devices then the spatial crossover is not at the equidistant point between them. Delay is indicated, the amount of which is proportional to the level offset and physical displacement between the elements. If the level offset is small and the geometry of the

Perspectives: Anticipate optimization. As I design a system for a particular use, I think about how I will go about bringing it up and the possible equalization strategies 1 might use. More than once this process has changed some of the design decisions I have made. I feel that as a designer 1 am responsible for the suystem's design, how to best optimize it, and how to get the most out of it when it is used. When I don't manage to do all three, the system doesn't work as well as it might have.

Alexander Yuill-Thornton II (Thorny)

enclosures allows for extremely low displacement it may be deemed impractical. In such cases the delay required to phase-align the asymmetric spatial crossover is extremely minute and may not be worth the expense, trouble and (most importantly) risk of error. It is often said that "line array" speakers do not require delay tapering. Remember that no element can be termed an "array" until it is placed in combination with others. Since we have already eliminated the coupled line source from the minimum variance menu, there is little need to enter the delay debate there. Suffice to say that coupled speakers lacking some measure of acoustic subdivision (angular isolation) cannot reap much benefit from electronic subdivision in any form, whether it is level, EQ or delay. Our applications for the "line array" products are highly overlapped versions of the asymmetric coupled point source. These will have asymmetric spatial crossovers wherever asymmetric conditions are found along the way. In large arrays, gratuitous amounts of angular and level asymmetry may create a beast much more complicated to disentangle than we can handle. A practical strategy for arrays with high quantities of elements is described later in this chapter, where the role of delay tapering in that application will be described. To be on the safe side, if there is a level break, have EQ and delay at the ready.

Subsystem Types

The speaker system is a family of related subsystems, which can in turn be a family of related speakers and so on. A main array may be a symmetric coupled point source in the horizontal plane. The subsystem that covers the area below the mains, the downfill array, is also a symmetric coupled point source in the horizontal plane. Joined together, these two arrays will be an asymmetric coupled point source in the vertical plane. The list can go on from there as we add sidefills, infills, frontfills, delays, etc. Each subsystem has a classification as either a single speaker or an array, and is capable of creating a minimum variance shape as outlined in the previous chapter. Each subsystem in turn joins with its neighbors and forms newly combined arrays from the smaller ones. These second-generation arrays will need to be classified as well, and a new combined shape of minimum variance can be created. The process continues until all of the subsystems are woven together into a single entity.

Main Systems

Main systems are the principal building blocks. They will take the first and largest portion of the room shape. It will be the job of the fill systems to share the leftovers. Many systems will contain only a single main speaker, or array. The fill systems will join the mains to plug the gaps in coverage. There are some shapes which will require multiple mains that in turn will be joined together as uncoupled arrays.

Multiple mains are found when the room geometry or other physical factors favor an approach beyond a single source point. The arrays must be of roughly equal rank in terms of power capability and length of throw, to be considered multiple mains rather than a main and a fill system. These systems are most often seen in dual monaural horizontal arrangements where main systems are deployed flanking the stage. Such configurations differ from stereo in that the same signal is sent to both systems. This overall combined configuration is an uncoupled line source array, although it may be made up of a pair of coupled point source arrays.

The multiple main strategy is also often employed in the vertical plane, where upper and lower systems cover roughly equidistant areas. This configuration is fairly common in venues where over- and under balcony seating areas are roughly equidistant from the speakers. These two configurations can be found together, thereby creating a four-element multiple main system with two horizontal and vertical elements.

Another four (or more) element multiple main system is found when the horizontal "in the round" configuration is used (the vertical configuration requires zero gravity conditions and is much less popular). The listening area is sliced into coverage sections of roughly equal throw

Figure 7.4 Typical horizontal relationships of multiple main systems

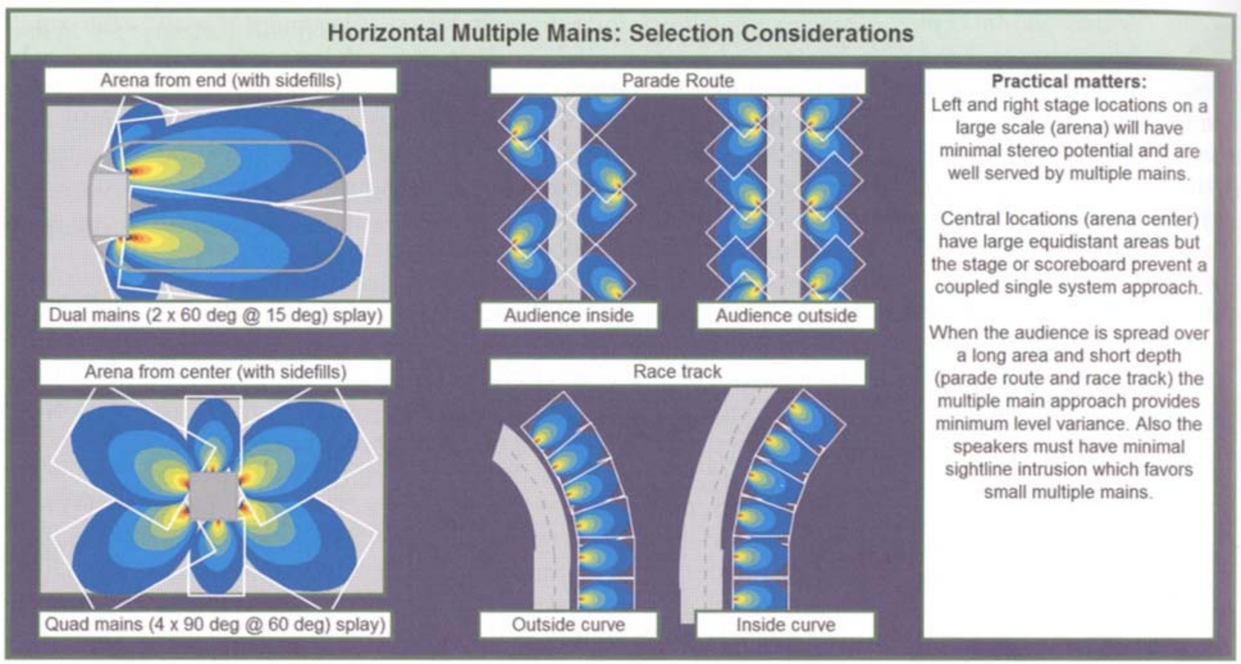
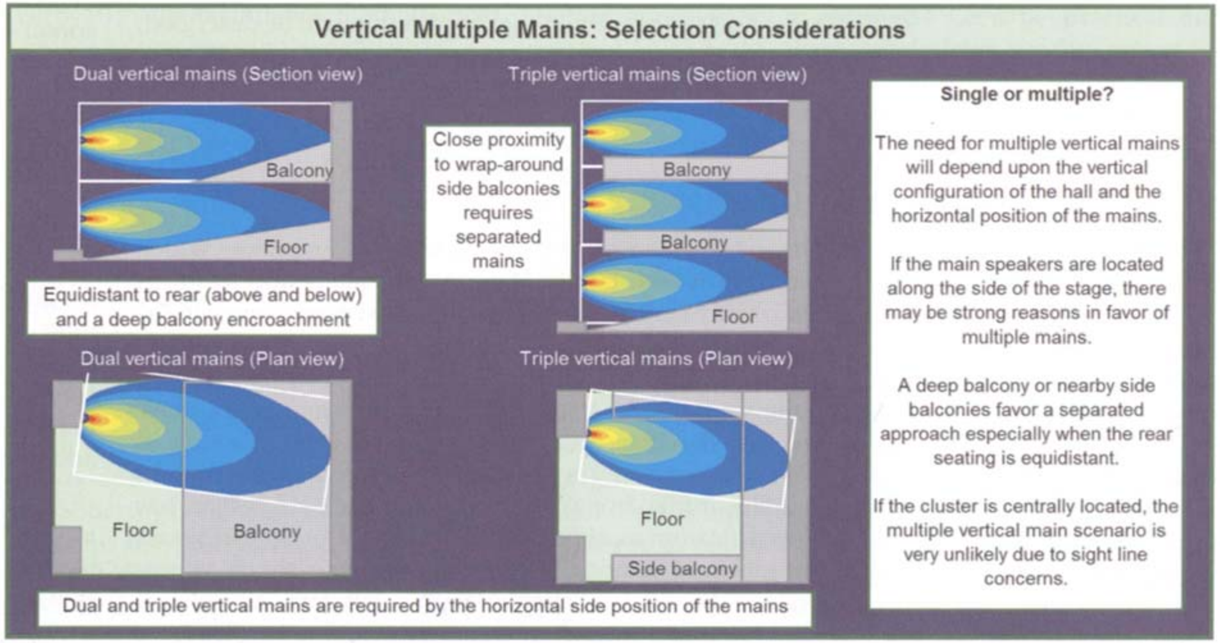


Figure 7.5 Typical vertical relationships of multiple main systems



Perspectives: The most important factor in optimization is the speaker placement. The mounting of the speaker is 80 per cent but is thought of as only 20 per cent. Naturally, this is best done in the system plan before it enters the site.

Kazuyuki Kudo, AST Inc.

length for each main system. The overall combined array is an uncoupled point source. A typical example of this would be arena scoreboard systems.

There is no limit to the number of multiple mains. If the venue shape requires lateral extensions with equidistant throw length, the number of mains can rise sharply. An example of this is a parade route, where the audience distance remains constant. The overall configuration is an uncoupled line source (on the straight areas) and uncoupled point sources and point destinations on the curves. Another example is found in racetracks and sports stadiums where the systems are located on the roof overhangs.

Sidefill

Sidefill subsystems provide horizontal radial extension to the main system. Sidefill systems may be directly coupled to the main system or uncoupled single elements or arrays. The typical sidefill system is an asymmetric component of the coupled point source in the horizontal plane. The common factor is that they have a significantly shorter throw distance than the mains, and therefore qualify as subsystems. The sidefill system will be assumed to be semi-isolated from the mains, with the degree of independence being inversely proportional to frequency.

Infill

Infill subsystems also provide horizontal radial extension to the main system(s). These differ from sidefills in that they are oriented toward the center rather than the outside. The infill systems form a symmetric point destination along the room center line. Infill coverage must be restricted to very limited areas in their separate local areas. The intersection of a center downfill with a pair of infills is a three-dimensional point destination array in the center seating area. This configuration is a mistake few designers make more than once.

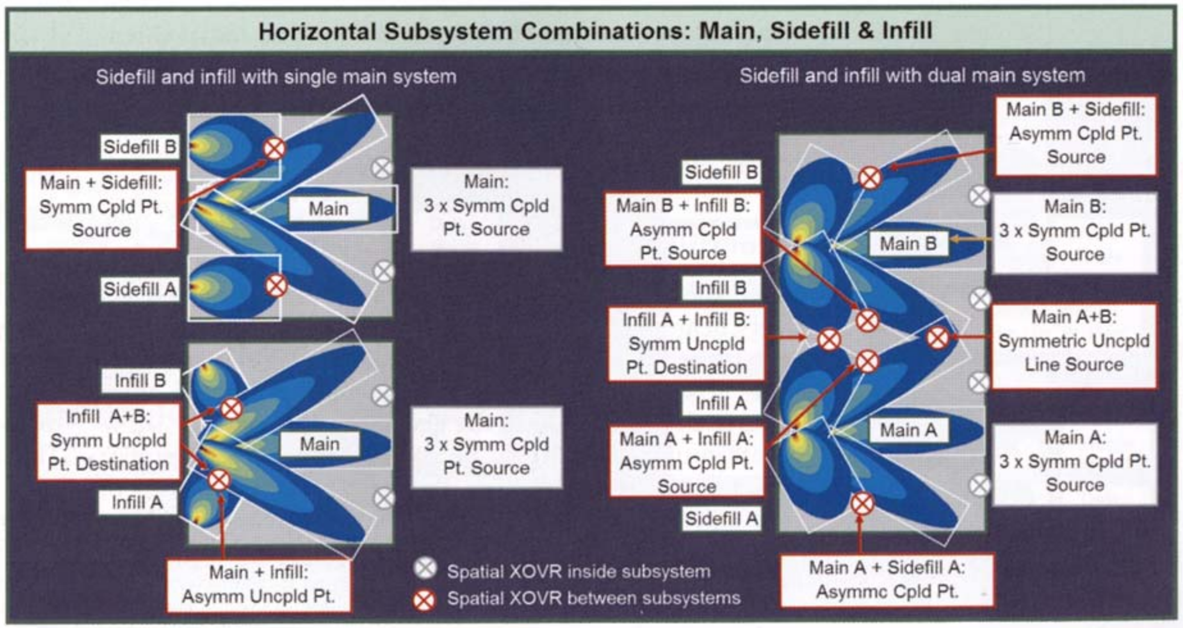


Figure 7.6 Typical horizontal relationships between main, sidefill and infill subsystems

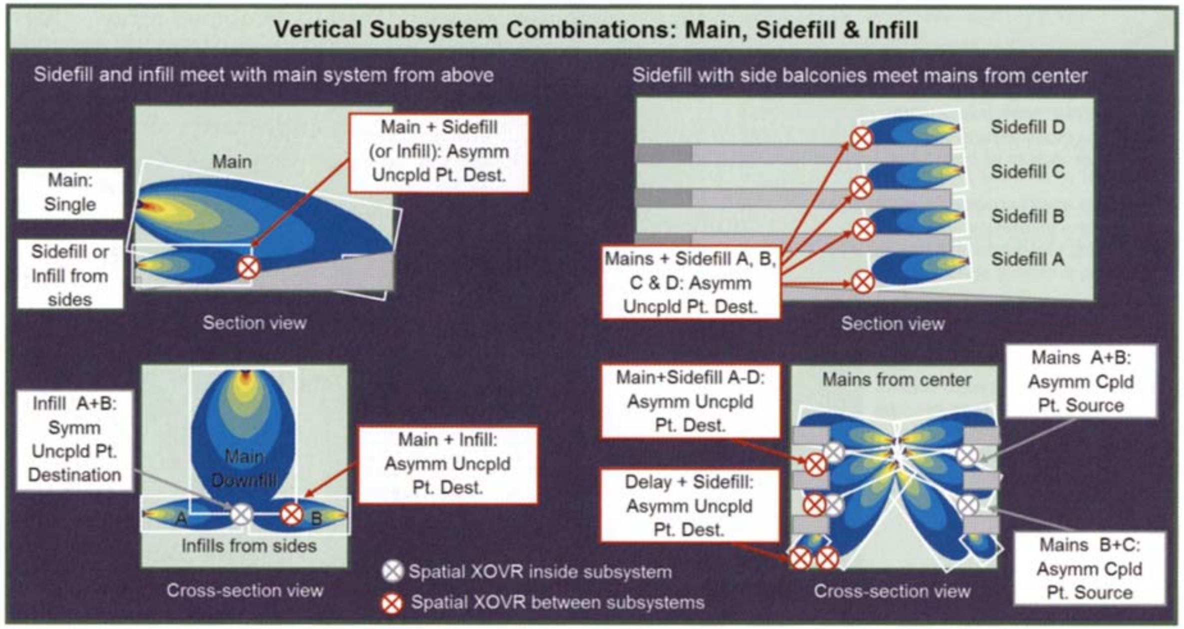


Figure 7.7 Typical vertical relationships between main, sidefill and infill subsystems

Downfill

The role of the downfill array is as a vertical radial extension of the main array. The typical downfill system is an asymmetric component of the coupled point source. It can be unmatched in any or all of the categories of speaker order, splay angle and level.

Frontfill

The role of the frontfill system is coverage of the seats very near to the stage. The speakers are usually located on the stage lip, either mounted into or laying on the stage. The array configuration is the uncoupled line or point source depending upon the geometry of the stage. These seats could alternatively be covered with fewer sources by downfill speakers from above or infill speakers from the sides. The frontfills join a flying main system as an uncoupled asymmetric point destination array in the vertical plane. They connect to an infill array as an uncoupled asymmetric point destination in the horizontal plane.

Advantages of the frontfill array:

1. Minimum horizontal and vertical image distortion
2. Minimum level, spectral and ripple variance
3. Minimum loss of gain before feedback
4. Minimum overlap into other subsystems.

The frontfill array is very limited in forward coverage depth as was discussed in Chapter 6.

Delays

Delay systems are forward extensions of the mains. Delay speakers are inherently uncoupled from the mains and combine as an asymmetric point destination array. The delay speaker may operate singly or as their own array. The most favorable array configuration is that which places the elements along an equidistant arc from the main speaker (an uncoupled point source array). This is not always practical due to the restrictions in available placement options. Another complication arises in the case of placing delays for a left/right dual mono or stereo

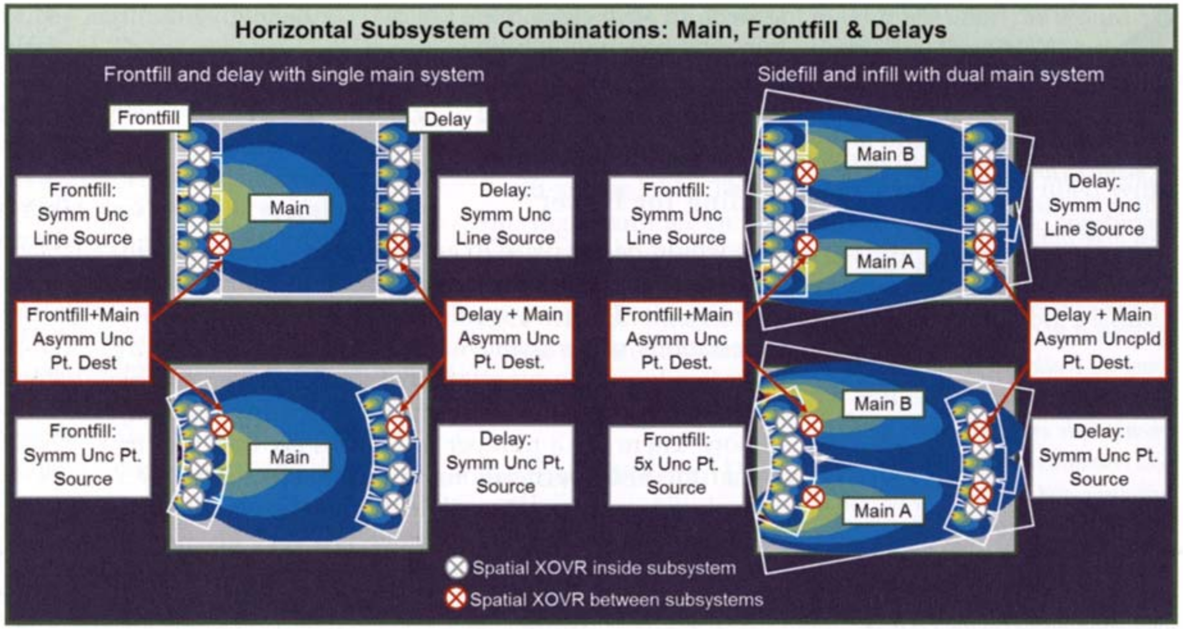


Figure **7.8** Typical horizontal relationships between main, frontfill and de

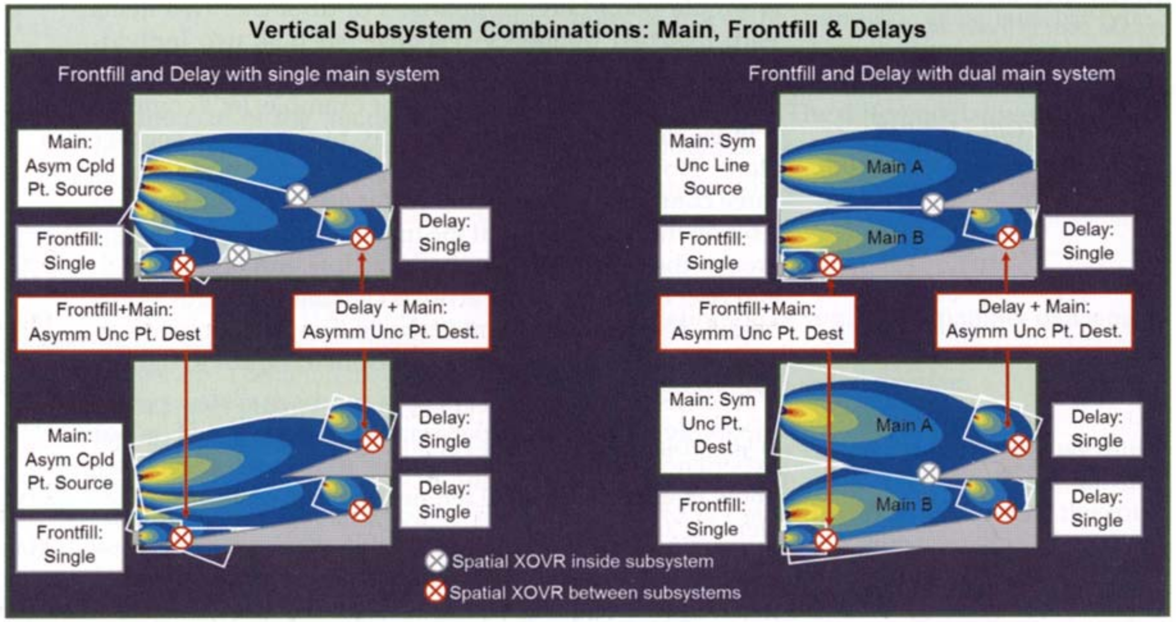


Figure **7.9** Typical vertical relationships between main, frontfill and delay s

Perspectives: Dialing in a system is like breaking in a new ride. A driver can affect how the performance of a new machine will finally settle in. So many factors affect large-system calibration that buyers/users need to be thinking in terms of months or years, not weeks. Some professionals with professional equipment can certainly get it way *down the track in a power burst of several days. However, after that, when people are in the room, and new resonances show up in the structure, and the actual operating volume falls into the pocket, and real noise floors are present,* then *it's time to get tweaky. Let the compromises begin! There is nothing like having relationships with other professionals, pro-test gear, freedom to experiment, and the trust of leadership, all affecting the final outcome towards excellence.*

Chris Gille; JeffPelletier Willow Creek Community Church

system. In such cases there are two conflicting concentric lines and no single solution.

The Scalable Design

Scaling for Power

This would be the place to go through a series of calculations based on speaker sensitivity, maximum SPL, coupling, quantity, etc. Then we would compare the numbers with expectations of program material and the prorated distance loss less some expectation of speaker/room summation. But this is not how it's done by me, or any others that I know. There are a number of reasons for this. The first is that the SPL specifications from manufacturers are so convoluted and conditional that model-to-model comparison based on spec sheets alone is a very iffy process. The next is that that behavior of arrays is complex and interdependent, and the combined SPL is highly variable over frequency. dB SPL ratings provide a classification for speaker elements, but the relevance to combined arrays is so fraught with variable conditions that its use is minimal. A simple point here: comparison of dB SPL of two items with matched frequency response, such as two individual elements, tells us something concrete. Comparison of unmatched elements does not. For example: let's compare a three-element point source array to the center speaker alone. How do we compare this pink-shifted spectrum of three combined elements to the flat original? The low end rises on the center axis but the high does not. In the on-axis area of the outer elements both the low end and the high end rise. If the SPL of the array is characterized only for its on-axis response, the numbers seem disappointing. Meanwhile entire listening areas went from darkness to light.

It seems to me that most people approach system power capability with tangible personal experience based on our experience of real devices, rather than spec sheets. We know that this venue will require a certain scale of speaker element and quantity.

I call this the Goldilocks sound design approach, and use it myself: Papa Bear, Mama Bear and Baby Bear. The sound design is approached by first defining the power scale of the largest and longest throw system as the Papa Bear. The smaller systems are scaled from there by evaluation of the relative distances involved. If the downfills only need to go half the distance of the mains they can be scaled down by 6 dB. If the under balcony systems go 25 per cent of the way, then 12 dB down in dB SPL scaling will be appropriate.

It would be a waste of time for me to preach to anyone what is the right dB SPL for pop music or a house of worship. Who would follow it? We know the products we know, and that begins the scaling process. I realize that this sounds terribly unscientific but this is a time for realism. No serious professional is going to specify a main array full of elements they have never heard, based on specifications posted here or by the manufacturer. Let's be real!

This is one of the most important decisions about the sound system design, but the answer will not be found in a book. There is no substitute for real field experience. Our experience of sound power requires a personal encounter. We can specify cable from a piece of paper. Speakers must be heard.

The scaling factor can be evaluated at two levels:

- What is the power scale required to do the job with minimal overlap between the elements?
- What is the power scale required to do the job with larger amounts of overlap between the elements?

Naturally the latter of these can be done with elements of a smaller power scale, but we may pay a price in variance. In any case, once the main system power scale is defined, the rest will follow by range scaling. We will consider the matter of power scaling closed then. Everything that follows from here will reflect on power scaling only in relative terms.

Overlapping for Power

Major power addition is available to us in the form of overlapped patterns. We know by the TANSTAAFL principle that we will have to pay with variance. Is there a best way or is this a case of "by any means necessary"? There are very clear-cut choices here. The answer lies in the beamwidth and speaker order. The plateau area of the beamwidth response is the most expensive (in ripple variance) form of overlap. The first-order speaker is the worst candidate for overlap. And every effort should be made to minimize this. The second-order speaker will typically have a smaller plateau and this gives us more room, but the ripple variance cost is still high. The third-order speaker is the easy winner here. In fact the third-order speaker needs overlap to expand its coverage beyond a minimal area. An overlapping coupled point source comprising constant slope beamwidth elements has a *fixed angular offset* that creates a *variable overlap percentage* over frequency. As frequency falls the overlap percentage rises. The rising overlap is met by expanding wavelength, which keeps the ripple variance under control. It is a delicate dance, but as long as the displacement is kept small, the opportunity is presented for large-scale power addition which increases as frequency falls.

Scaling the Shape

We have studied the shapes that can be created by speakers to fill the room with minimum variance coverage. The coverage shapes are scalable to the size of the venue. If the coverage shape calls for an asymmetric point source, the array must have the appropriate power scaling for the application. For a given program material the power needs rise with the venue size. For a given venue size the power needs rise or fall with the program material requirements. The shape that must be created is the same. The larger-scale version will use higher power, higher-order elements and contain greater amounts of overlap. The smaller-scale and lower-power systems will use fewer elements, lower speaker orders, and less overlap, to create the same basic shape.

The scalable design is used to fill a comparable space. We can fill 90 degrees of symmetric coverage with a single first-order speaker or ninety third-order systems splayed 1 degree apart. The shape is comparable but the amount of power could be orders of magnitude apart, as would be the budget, the rigging requirements, the ripple variance and a host of other tradeoffs. The final decision will rest on the best compromise.

All of the array shapes shown in the minimum variance menu are scalable. The coupled arrays can be filled with varying quantities and percentage overlap. The combination of one array with another is simply a second generation built upon the previous. Coupled arrays at different locations will combine as uncoupled arrays in the big picture. This is more of our scalable paradigm.

Array Design Procedures

Main System Design

If a single speaker will suffice, there is nothing more required here than to follow the single element guidelines for symmetric and asymmetric aiming. The coverage pattern is determined by the aspect ratio.

If such a scheme is not satisfactory we will resort to a coupled point source of the symmetric or asymmetric type. There are two versions of each we will consider: minimum variance (with maximum power) and maximum power (with minimum variance). The difference between the two approaches is one single parameter: angular overlap.

The coupled point source is chosen for three primary reasons:

1. Radial extension beyond the capability of a single unit.
2. Radial edge definition beyond the capability of a single unit.
3. Power addition through radial overlap.

Once we cross the line to the coupled point source we must abandon the simplicity of the aspect ratio for the radial arc. There is no "correct" element with which to start the coupled point source. We must start with something and then the process begins as we see what is left to cover. After the next piece is added we re-evaluate again until the shape is filled. The process can be restarted until a satisfactory assemblage of puzzle pieces is found.

Let's begin with the symmetric version.

Symmetric Coupled Point Source

There are three modes of overlap behavior that characterize the symmetric point source array over quantity. The 100 per cent overlap model is, of course, not a point source at all but rather a line source. In any case it is the ultimate extreme limit, and its behavior is characterized by forward aspect ratio (FAR) multiplication. This is the maximum on-axis power addition, as well. We cannot tolerate the interactions on the coupled line source for reasons that have been previously discussed. We must have some degree of angular isolation. The outer extreme on the other end is the unity splay. In this case the behavior is characterized as coverage angle multiplication. This has the least on-axis power addition, but we should not forget that we are spreading power addition across the arc. The middle ground behavior is the partial overlap configuration which is characterized as splay angle multiplication. This is the middle ground in on-axis power as well.

How angular overlap affects combined coverage angle:

1. 0 per cent overlap: Combined coverage = Quantity X the element coverage angle
2. 5-95 per cent overlap: Combined coverage = Quantity X the splay angle between the elements
3. 100 per cent overlap: Combined coverage = Quantity X the forward aspect ratio

Let's say we measure the coverage area and we need 40 degrees of coverage. What are our options? There are countless ways to create a 40 degree coverage pattern. The following example illustrates the role of overlap in the computation of combined coverage angle.

Symmetric Coupled Point Source Design Reference

Overlap	dB Addition			Coverage Addition		
Angle (%)	Qty = 2	Qty = 4	Qty = 8	Qty = 2	Qty = 4	Qty = 8
100%	6	12	18	2xFAR(1/2 Coverage)	4 x FAR	8 x FAR
75%	5	5.5	7.5	2 x Splay Angle	4 x Splay Angle	8 x Splay Angle
50%	4.5	4.5	4.5	2 x Splay Angle	4 x Splay Angle	8 x Splay Angle
25%	3	3	3	2 x Splay Angle	4 x Splay Angle	8 x Splay Angle
0%	0	0	0	2 x Coverage Angle	4 x Coverage Angle	8 x Coverage Angle

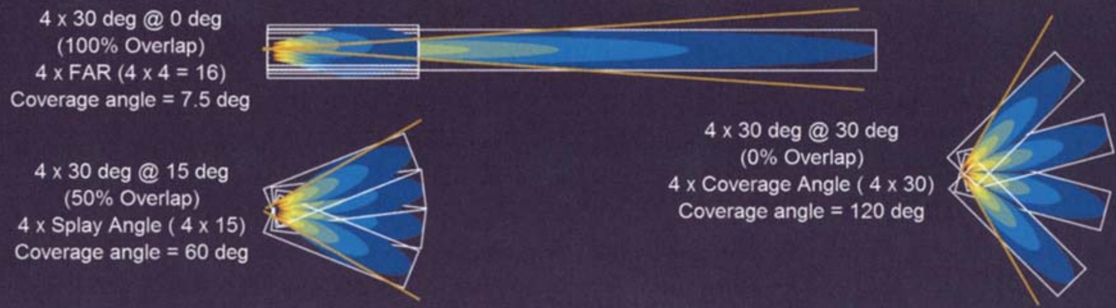


Figure 7.10 Design reference for the symmetric coupled point sou

Combined coverage angle of 40 degrees:

- 1 X 40 degree speaker (FAR 3)
- 2 X 20 degree speaker @ 0 per cent overlap (coverage of 20 degrees X 2 = 40 degrees)
- 2 X 27 degree speaker @ 25 per cent overlap (splay of 20 degrees X 2 = 40 degrees)
- 2 X 40 degree speaker @ 50 per cent overlap (splay of 20 degrees X 2 = 40 degrees)
- 4 X 40 degree speaker @ 75 per cent overlap (splay of 10 degrees X 4 = 40 degrees)
- 8 X 20 degree speaker @ 75 per cent overlap (splay of 5 degrees X 8 = 40 degrees)
- 2 X 80 degree speaker @ 100 per cent overlap (FAR 1.5 + FAR 1.5 = 3)

All of the above scenarios have TANSTAAFL and triage considerations. Overlap is never free of ripple variance. There is always a tradeoff of power and ripple. As overlap increases, the key to minimum ripple variance is low displacement. The third-order speaker is our best choice in high overlap designs. Low overlap designs favor the first- and second-order plateau beamwidth with its extended isolation zone.

Coupled point source design procedure:

1. Define the coverage shape as an arc from the speaker array location.
2. Select an element with an aspect ratio that will fit inside the shape
3. Place additional elements at the unity splay angle to provide radial extension until the shape is filled. The combined coverage angle equals the individual coverage angle multiplied by the quantity.

If power addition is required, then reduce the splay angle to less than unity and add units until the shape is filled. The combined coverage angle equals the splay angle multiplied by the quantity.

Asymmetric Coupled Point Source

The asymmetric coupled point source is chosen for the same reasons as the symmetric point source but fits into a different shape. The power addition of the asymmetric point source is self-scaling since the entire rationale for this array is that it fits into a variable distance shape. The levels are tapered as appropriate to scale the array to the distances presented. The unity splay angle must then be compensated as appropriate for the level offset factor.

There is no "correct" element with which to start the asymmetric coupled point source. We must start with something and then the process begins as we see what is left to cover. After the next piece is added we re-evaluate again until the shape is filled. The process can be restarted until a satisfactory assemblage of puzzle pieces is found.

It is a simple matter to find the unity splay in a symmetric array. A pair of 30 degree speakers will be splayed 30 degrees apart. The equation is the symmetric unity splay equation:

$$(\text{Coverage}^1 + \text{Coverage}^2)/2 = \text{unity splay*}$$

*when levels are matched

Therefore

$$(30 \text{ degrees} + 30 \text{ degrees})/2 = 30 \text{ degrees}$$

The spatial crossover would be at the geometric and level mid-point: 15 degrees off-axis to either element. What about if we want to merge a 30 degree and 60 degree speaker? The same equation applies, as long as the levels are matched.

$$(30 \text{ degrees} + 60 \text{ degrees})/2 = 45 \text{ degrees}$$

With a 45 degree splay the elements will meet at the same spatial crossover point: 15 degrees off-axis from the 30 degree element. It would be met there by the —6 dB edge of the wider element, 30 degrees off-axis to its center. The spatial crossover is the level center, but not the geometric center.

The equation must be modified if the levels are offset between the elements. There is no best splay angle between two speakers, until we know the distance/level relationship. If we were to take two matched elements and turn one down 6 dB, the supposed unity splay angle will not provide unity results. The geometric center will find —6 dB from one element and —12 dB from the other. What is the unity splay angle then? A change of 6 dB is a level difference of 50 per cent. The splay angle will need to be adjusted by the same ratio to bring back unity performance through the spatial crossover. A reduction of 6 dB would reduce the range of the speaker by half its distance. That will be the decisive number.

The compensated unity splay equation:

$$((\text{Coverage}^1 + \text{Coverage}^2)/2)\ X\ (\text{Range}^2/\text{Range}^1) = \text{Compensated unity splay*}$$

*assumes that levels are set in proportion to the distance

Here is an example of two 30 degrees speakers with one covering half the distance of the other (—6 dB)

$$((30 \text{ degrees} + 30 \text{ degrees})/2)\ x\ (0.5/1) = \text{Compensated unity splay}$$

$$((60 \text{ degrees})/2)\ X\ (0.5) = \text{Compensated unity splay}$$

$$30 \text{ degrees } X\ 0.5 = 15 \text{ degrees}$$

Here is an example of a 30 degree speaker joined with a 60 degree element that is covering 70 per cent of the range of the other (-3 dB)

$$((30 \text{ degrees} + 60 \text{ degrees})/2)\ X\ (0.7/1) = \text{Compensated unity splay}$$

$$((90 \text{ degrees})/2)\ X\ (0.7) = \text{Compensated unity splay}$$

$$45 \text{ degrees x } 0.7 = 31.5 \text{ degrees}$$

Asymmetric coupled point source design procedure:

1. Define the coverage shape and speaker location.
2. Select the first element and aim it at the most distant location within the shape. This element will have the narrowest coverage angle.
3. The next step is to define the distance at which the transition into the next element is expected to occur (the on-axis location of the next element). For example, the first element is 30 degrees, draw a line 30 degrees off-axis and define the range. If this distance is less than the first element, then a compensated unity splay angle will be required.
4. Select an additional element. The combined coverage of the two elements is used to calculate the symmetric unity splay angle. The level offset provides the compensation percentage.
5. Position the second speaker at the compensated angle and appropriately scaled range. If the coverage angle or range difference between the elements is very large it may be necessary to adjust the compensation. This is

Asymmetric Coupled Pt. Source Design Reference

Elements	Unity Splay	Compensated Unity Splay Angle		
Combined Coverage Angle	Symmetric XOVR	Range Ratio	Level Offset	Asymmetric XOVR
(deg)	(deg)	(%)	(dB)	(deg)
90	45	100%	0 dB	-----
90	45	90%	-1 dB	40.50
90	45	80%	-2 dB	36.00
90	45	70%	-3 dB	31.50
90	45	63%	-4 dB	28.35
90	45	56%	-5 dB	25.20
90	45	50%	-6 dB	22.50
90	45	44%	-7 dB	19.80
90	45	40%	-8 dB	18.00
90	45	35%	-9 dB	15.75
90	45	32%	-10 dB	14.40
90	45	25%	-12 dB	11.25

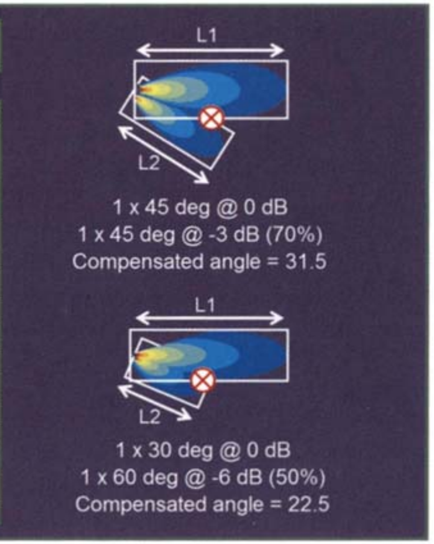


Figure 7.11 Design reference for the asymmetric coupled point source. The computations for a 90 degree combined coverage angle are shown. Other angles can be substituted and prorated by the same percentage rates

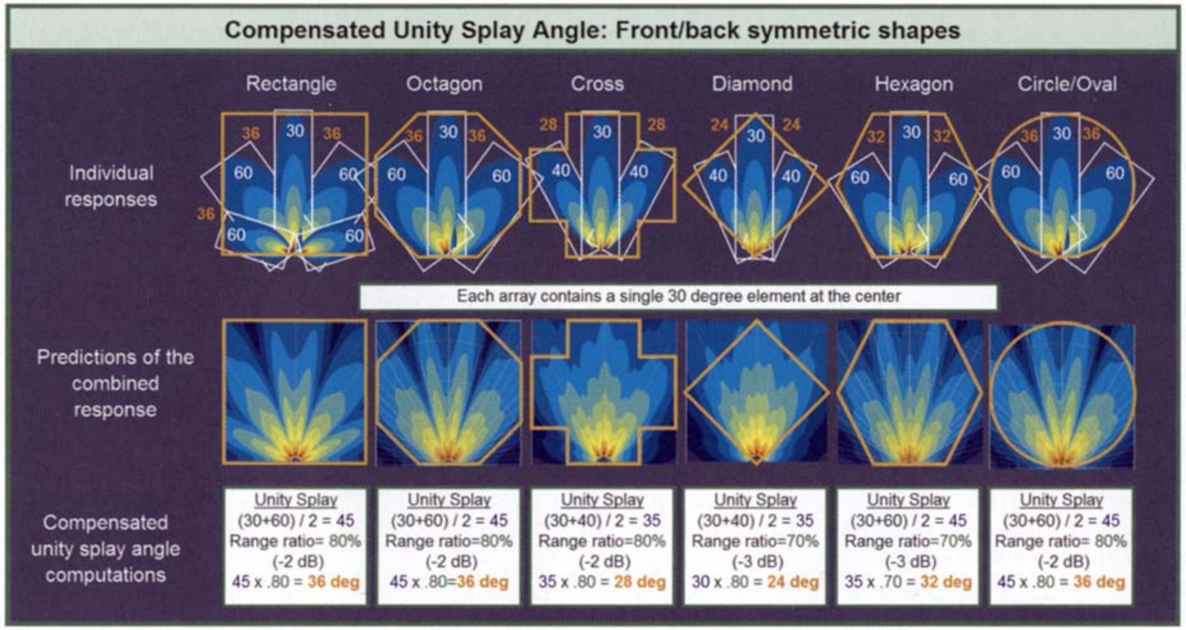


Figure 7.12 Design examples for the asymmetric coupled point sourc ... ening a

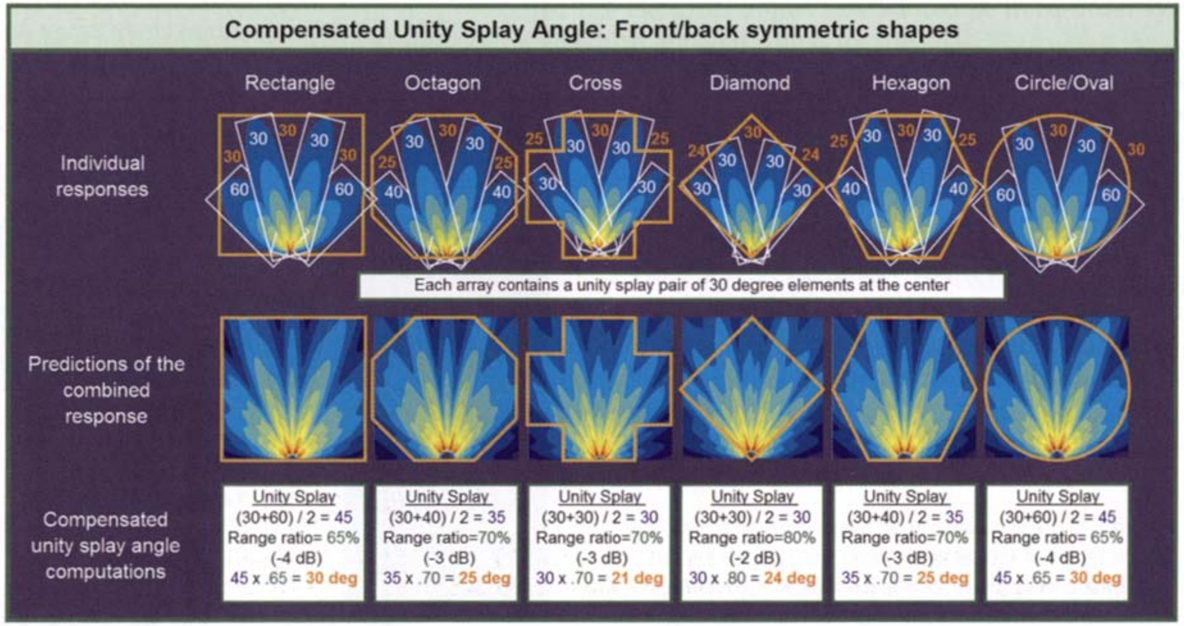


Figure 7.13 Design examples for the asymmetric coupled point source. In each case the inner elements are symmetric unity splay while t ... ıpe

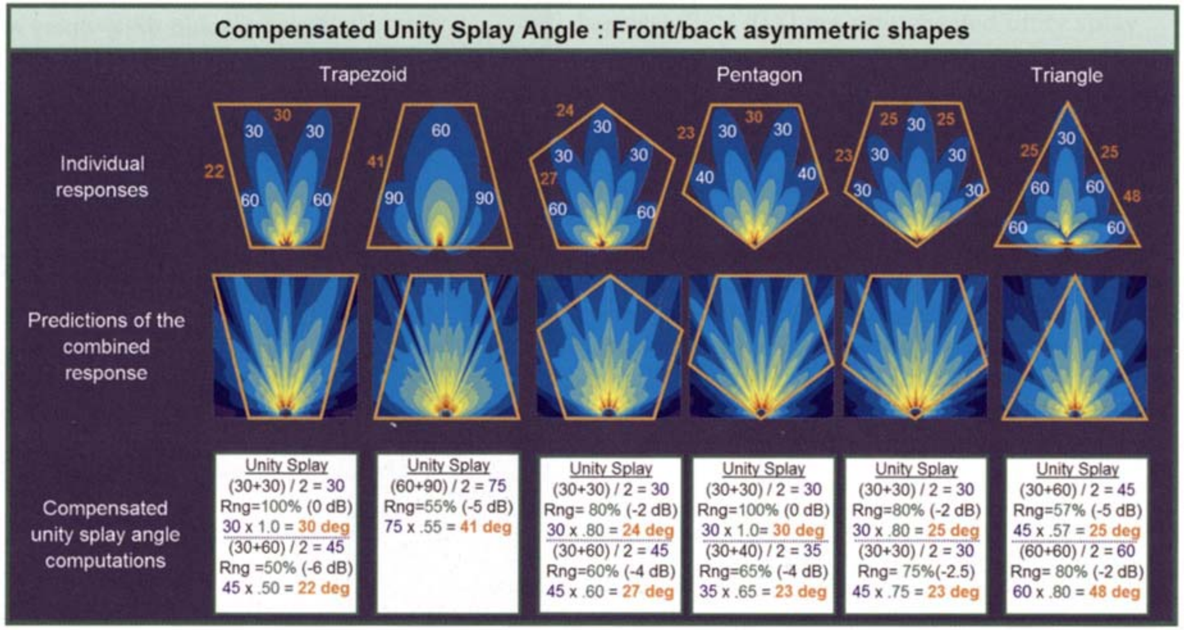


Figure 7.14 Design examples for the asymmetric coupled point source. In each case the shapes are front/back asymmetric. The outer elements are designed with compensated unity splay angles that cu

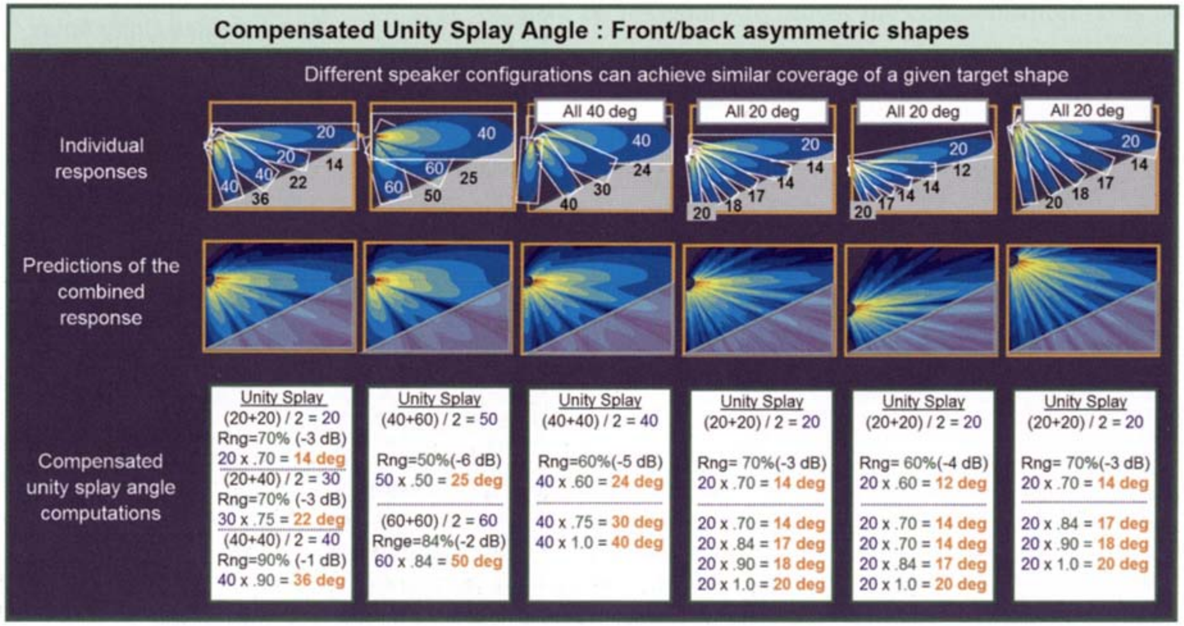


Figure 7.15 Design examples for the asymmetric coupled point source, typical of vertical applications. In each case the same shape is pre

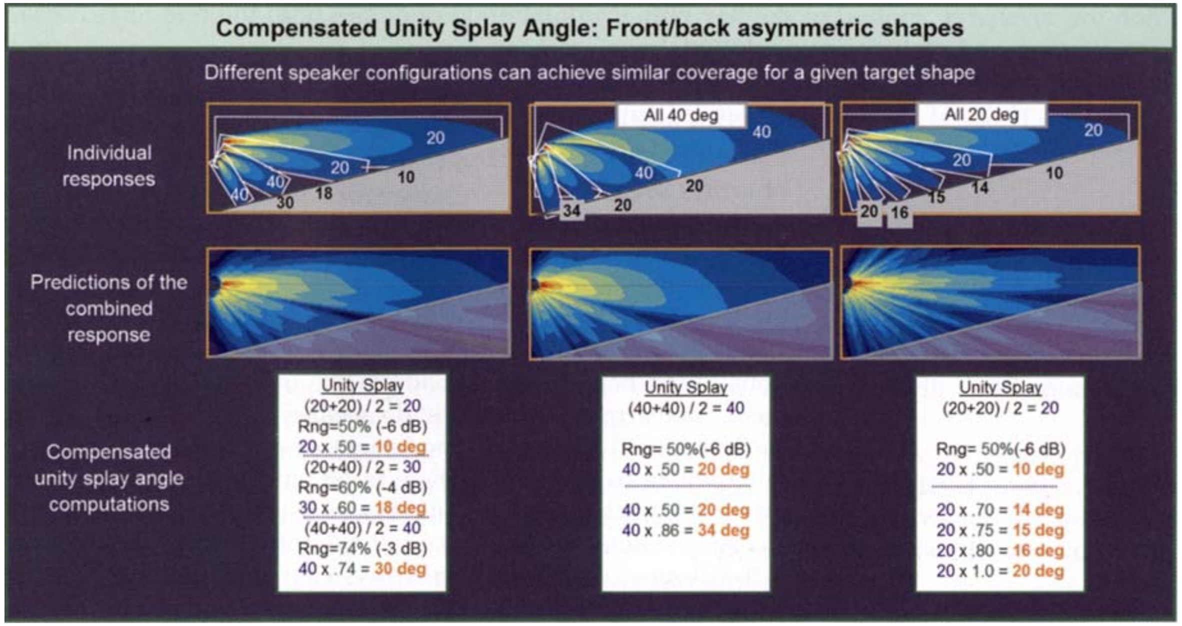


Figure 7.16 Design examples for the asymmetric coupled point source, typical of vertical applications. In each case the same shape is presented for coverage. Different combinations of elements and positions all create a minimum variance coverage of the shape

done by reassessing the range ratio and adjusting the compensation until the best fit is found.

6. If power addition is required, then reduce the splay angle to less than unity and add units until the shape is filled.
7. Continue the process with the third element being aligned to the second element and so forth.

Recall in Figs 6.10-6.13 where we saw that the details in the building shape had only a minimal effect on the single speaker solutions. The upgrade to a coupled array opens up a huge variety of possibilities available to customize our array to the room shape (Figs 7.12 - 7.14).

Asymmetric-Composite Coupled Point Source

An asymmetric coupled point source can be constructed in a modular form from pre-combined symmetric coupled point sources. This is the means by which manageable high-power arrays can be designed with substantial overlap. The distance/level layering principles of the asymmetric coupled point source provide the outer structure, which combines highly overlapped symmetric subsystems. The elements in this case should be third-order systems, which are well suited for the task.

The composite array provides a practical and definable character to the gazillion element "line array" that is the current mainstay of the industry. Let's take a moment and consider the situation, since this is indeed the most popular type of main system array.

First we don't have to go through the fact that this is not actually a "line array" since it is curved. We know why we need it curved. Next let's remember that we want the optimization engineer to have something they can disassemble and put together on site, to calibrate the final product to the space.

Now here is a rendering of how these arrays are designed in the modern age.

The design process:

1. Select the element based on power class and budget.
2. Divide the budget by the number of elements.

3. Play with every iteration of angles until the best fit is found.

If you're feeling insulted at the moment, please forgive me. This was my design process until I found a better way. I have run mics in a line from top to bottom of these types of arrays searching for ways to enact optimization solutions. Without defined zones, the optimization process is little more than the same kind of iterative guesswork that began in the design process. Even the experience of having tuned hundreds of asymmetric coupled point source arrays was no match for the complex overlapping interactions found here. Much is made in the industry of how these "line arrays" act as a single speaker. This fallacy is easily exposed by a small movement of a measurement microphone in the space. There is no single speaker in the world that has this form of rampant asymmetry which creates a different on-axis beam at every frequency. That is the destiny of coupled "line arrays" implemented with a host of different angles on adjacent elements and driven by a single processing channel. Even if we were to define it as a single speaker, we can not hope to calibrate it as a single speaker. Calibration requires a defined beginning, end and *(most* importantly) center of its coverage pattern. The center of a multi-element asymmetric array is inherently unstable over frequency. The coverage pattern of an individual element is not stable over frequency, leaving us with a variable percentage of overlap over frequency. The highs may spread out toward the ends but the lows are certain to concentrate towards the middle. The mid-range is pulled in both directions. None of them agree on the center. Without a defined center the process of optimization has lost its primary base of operations. How we can define "off-axis" if we can't find "on-axis"? Without a center, we will not be able to ascertain a nominal level setting, or an equalization that represents the middle ground of the coverage segment. In short, our hopes for a definable center point lie in symmetry, without which, we are left with optimization based on an arbitrary mic placement.

The use of a single processing channel would be logical, provided all elements on that channel are symmetrically arrayed. If they are not, the symmetrical solutions enacted by the signal processor will not have the expec results.

Does this mean we need sixteen different process channels to drive a sixteen-element array? Not likely. W building could possibly have a shape that would requ fifteen different splay angles and/or level adjustme between our cabinets? How can we make a sixteen-elem cluster into a manageable, predictable array? Break it i sections of symmetric coupled point source arrays. Ea has a top, middle and a bottom. We then have center axis areas for position adjustment, level setting, and equ ization. We even have a defined spatial crossover. Eve thing we need to define the system is in place. The segme will be designed to take on the character of a single speak which they can much more closely assume, due to th symmetry. The "composite speakers" that comprise asymmetric coupled point source would mimic the ch acteristics of a comparable array of single elements: longest throw goes to the narrowest speaker at the hi est drive level, and on down we go. How many subsyst breaks should there be? This is a practical as well as arc tectural issue. Asymmetry in the required coverage sha will need to be met with complementary asymmetry in array, but we cannot expect the system to conform to sm shape variations in the room. Each subsection will need be individually aligned and woven together with the c ers during the calibration stage. A practical assessment the benefits of each level of complexity should be carefu considered. In most cases four subsections should get job done.

The design process progresses in nearly identical fa ion. Define the longest section and subdivide. The bigg difference is that the segments are arc sections (since th comprise overlapped elements). Arc sections combine ferently than the rectangular shapes of the single elemen There is no unity angle compensation process, since the sections cannot encroach into each other's coverage. T spatial crossover is highly overlapped (as much as the e ments). Therefore, there will be no effort made to pha align in the crossover area. Just as we used the on-axis po of the individual elements as a level setting point, will do the same with the center point of each composite segment.

Before we move on, let's take a moment to address the issue of level segmentation. There is much talk about how much level tapering is healthy or proper for the system, and how large the incremental changes should be between adjacent cabinets. The argument that all speakers need to be turned up "to 11" is as old as the hills. It dates back to the wall of sound days and is immortalized in the movie *This is Spinal Tap.* The mentality is that turning down some elements will deprive the system of its vital power, the unspoken remainder of the sentence being, "at the mix position". The reality is that the lack of level tapering overpowers the front areas and oversteers the mid-range beam so that massive amounts of people get boom and sizzle and nothing in the middle. If a uniform experience is desired, the SPL Preservation Society will need to back off and let us taper the level in accordance with the hall geometry. If that is not enough power, then get a bigger system so that all of audience members get a share of the action. As for incremental taper size, the answer is simple. Look at the proximity ratio from top to bottom. Whatever we don't get from overlapped power concentration we will need to make up with level tapering. The increments can be few and large, or many and small. The only rule is that whatever is driven by one channel is a symmetric defined subsystem.

Principles of the asymmetric-composite coupled point source:

1. Symmetric subsystems are design to create a defined combined coverage angle. This is typically highly overlapped, so the coverage angle will be the quantity of elements X the splay angle.
2. The center line of each symmetric subsystem provides the distance offset parameter for level setting. Angular compensation is *not* required, as we are combining arc sections of multiple elements, rather than the rectangular shapes of a single speaker.
3. The number of asymmetric layers and the division between the layers can be modified until the best fit for the shape is achieved.

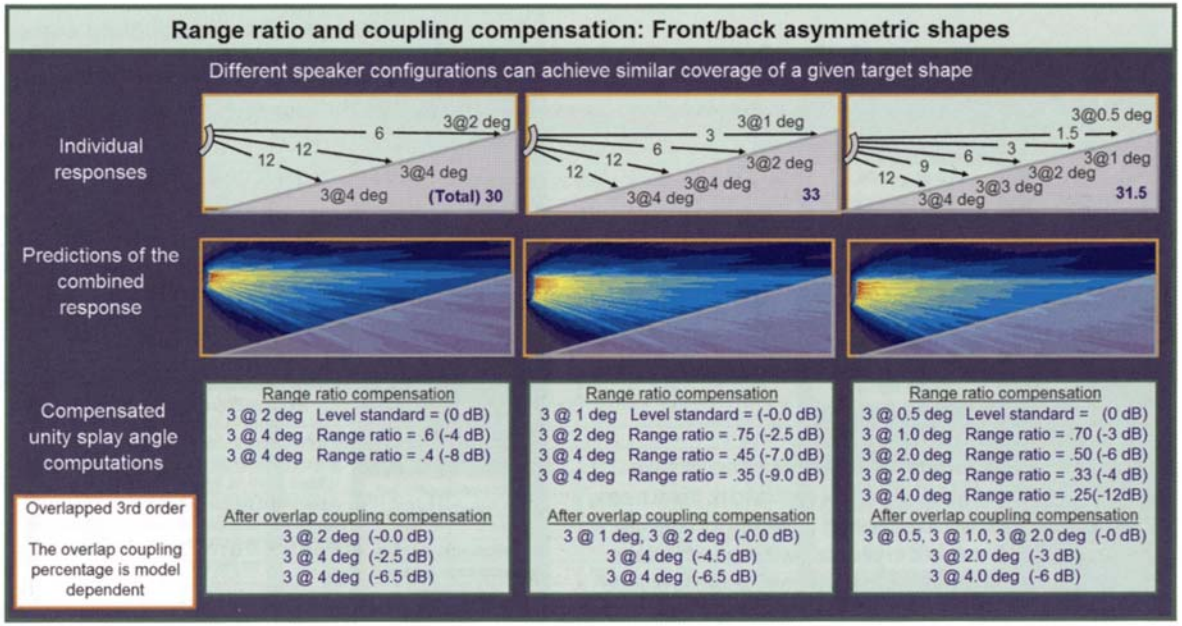


Figure 7.17 Design examples for the asymmetric composite coupled poi inimun

339

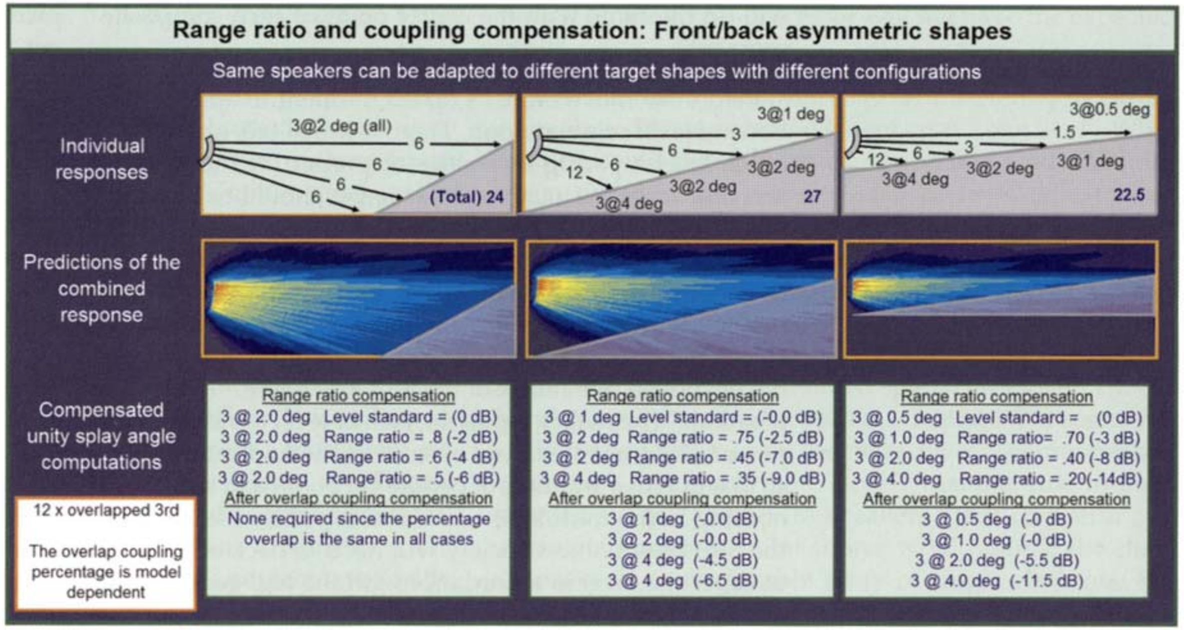


Figure 7.18 Design examples for the asymmetric composite coupled point source, typical of vertical applications. In each case a differe

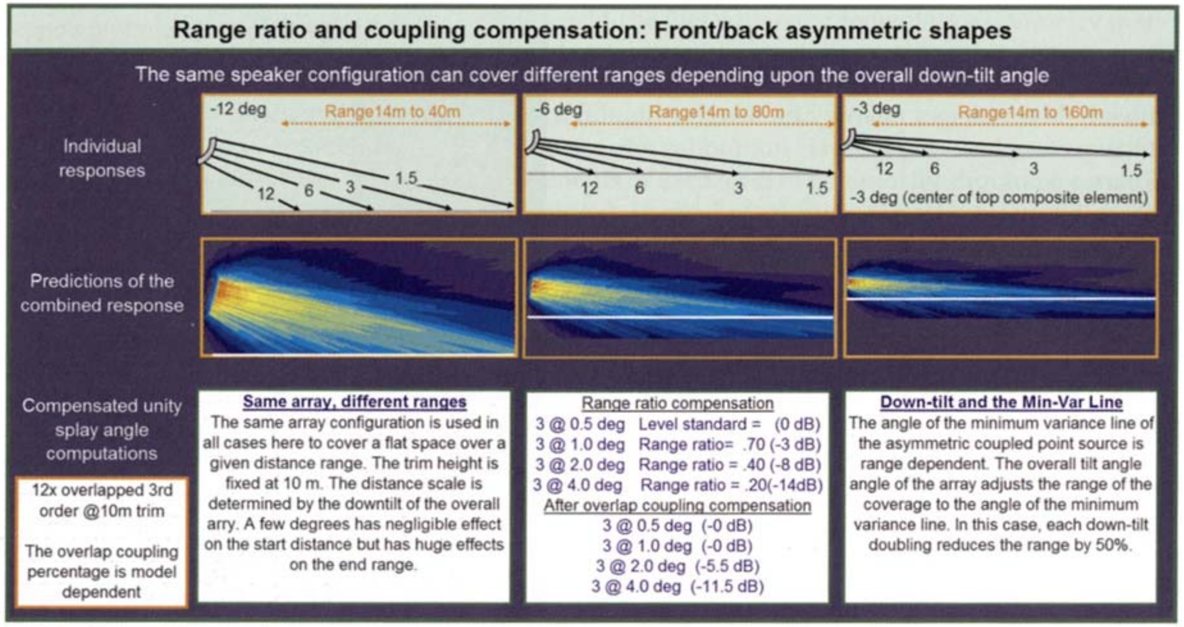


Figure 7.19 Design examples for the asymmetric composite coupled point source, typical of vertical applications. In each case a differe ran

Asymmetric-composite coupled point source design procedure:

1. Define the coverage shape and speaker location.
2. Select the most distant location, which is now defined as the coverage top (the vertical plane is assumed).
3. Select a lower limit that subdivides the coverage shape. The angle between the upper and lower limit will be the area covered by the first symmetric subsystem. This symmetric subsystem will have the narrowest coverage angle and should have the highest percentage of overlap between the elements. The distance from the source to the center of this coverage arc is the level setting distance.
4. Fill the selected arc angle with a sufficient number of overlapping elements whose combined splay angle equals the designated arc angle.
5. Subdivide the remaining space with a second arc segment, and repeat the process. The distance/level scalar is found at the center of each arc. The number of subdivisions is optional.
6. Repeat the process, each time decreasing the overlap percentage until the coverage shape is filled.

Symmetric Uncoupled Line Source

This is used for multiple mains and a variety of auxiliary systems such as frontfill, under balcony, etc. This is a range-limited system.

Symmetric uncoupled line source design procedure:

1. Define the coverage shape as a line from the speaker array location.
2. Define the line length and the minimum and maximum desired range.
3. Select an element and place it in the central area of the line.
4. Define the coverage by the aspect ratio method. The maximum range is the length.
5. The element-to-element spacing is found by stacking the aspect ratio rectangles along the line.
6. Place additional elements until the line length is filled.
7. Evaluate the minimum range of coverage for gaps. If the gaps are too large, the maximum range will need to be reduced and the element spacing reduced. Alternative

Symmetric Uncoupled Line Source Design Reference

Element		Known Maximum Length			Known Spacing		
Coverage (deg)	Aspect Ratio	Unity Length	Maximum Length	Width (Spacing)	Unity Length	Maximum Length	Width (Spacing)
180	1.00	0.50	1.00	1.00	0.50	1.00	1.00
150	1.04	0.50	1.00	0.96	0.52	1.04	1.00
120	1.15	0.50	1.00	0.87	0.58	1.15	1.00
90	1.41	0.50	1.00	0.71	0.71	1.41	1.00
80	1.56	0.50	1.00	0.64	0.78	1.56	1.00
70	1.74	0.50	1.00	0.57	0.87	1.74	1.00
60	2.0	0.50	1.00	0.50	1.00	2.00	1.00
50	2.3	0.50	1.00	0.43	1.15	2.30	1.00
45	2.6	0.50	1.00	0.38	1.31	2.61	1.00
40	3.0	0.50	1.00	0.33	1.50	3.00	1.00
35	3.3	0.50	1.00	0.30	1.67	3.33	1.00
30	3.9	0.50	1.00	0.26	1.93	3.86	1.00
25	4.6	0.50	1.00	0.22	2.31	4.62	1.00
20	5.8	0.50	1.00	0.17	2.88	5.76	1.00
15	7.7	0.50	1.00	0.13	3.83	7.66	1.00
10	11.5	0.50	1.00	0.09	5.75	11.50	1.00
7.5	15.4	0.50	1.00	0.06	7.70	15.40	1.00
5	23.00	0.50	1.00	0.04	11.50	23.00	1.00

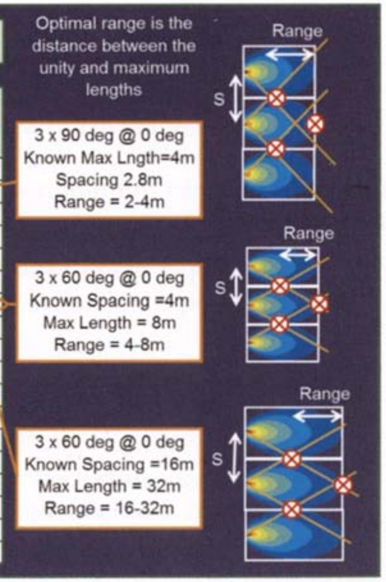


Figure 7.20 Design reference for the symmetric uncoupled line source

option is to use wider elements (which will also reduce the maximum range).
8. The maximum range is limited by the three-element overlap point, which is a function of the aspect ratio and the spacing. Consult the uncoupled array spacing reference (Fig. 7.20) for the applicable values.

Asymmetric Uncoupled Line Source

This is used for multiple mains and a variety of auxiliary systems such as frontfill, under balcony, etc. This is a range-limited system.

Asymmetric uncoupled line source design procedure:

1. Define the coverage shape as a curved line.
2. Define the maximum range for the element with the longest throw. If different models of speakers are being used, this should be the highest order element.
3. Select an element and place it with its aspect ratio scaled to the throw distance.
4. Place the next element and scale the length as appropriate for the given distance. The relative spacing and speaker order will determine the amount of overlap. A unity spatial crossover will require a compensated spacing/level/speaker order relationship.
5. Determine the relative distance range required of the second element (as compared to the first) and scale the power capability as appropriate.
6. Continue with each additional element having compensated spatial crossovers to the neighboring element until the desired shape is achieved. Shorter range elements may be power scaled as appropriate.

Symmetric Uncoupled Point Source

Symmetric uncoupled point source design procedure:

1. Define the coverage shape as a limited range arc.
2. Define the arc radius, length and the minimum and maximum desired range.
3. Select an element and place it in the central area of the arc.
4. The maximum range is limited by the three-element overlap point, which is a function of the aspect ratio, the angular offset and the spacing. Consult the uncoupled array spacing reference (Fig. 7.21) for the applicable values.
5. Place additional elements until the arc is filled.

Asymmetric Uncoupled Point Source

Asymmetric uncoupled point source design procedure:

1. Define the coverage shape as an asymmetric arc (a portion of an ellipse rather than a circle).
2. Define the maximum range for the element with the longest throw. If different models of speakers are being used, this should be the highest order element.
3. Select an element and place it with its aspect ratio scaled to the throw distance.
4. Place the next element and scale the length as appropriate for the given distance. The relative spacing and speaker order and angular offset will determine the amount of overlap. A unity spatial crossover will require a compensated spacing/level/speaker order relationship.
5. Determine the relative distance range required of the second element (as compared to the first) and scale the power capability as appropriate.
6. Continue with each additional element having compensated spatial crossovers to the neighboring element until the desired shape is achieved. Shorter-range elements may be power scaled as appropriate.

Symmetric Uncoupled Point Destination

Symmetric uncoupled point destination design procedure:

It is assumed that this is a fill system (e.g. infill) being combined with a previously defined main system. The

Symmetric Uncoupled Pt. Source Design Reference

Elements			Working Range		
Element Angle	Splay Angle	Overlap Angle (%)	Unity 1st XOVR	Maximum 2nd XOVR	Spacing Multiplier
90	0	100%	0.50	1.00	FAR
90	9	90%	0.57	1.13	FAR
90	15	83%	0.60	1.25	FAR
90	22.5	75%	0.63	1.50	FAR
90	30	67%	0.65	2.25	FAR
90	36	60%	0.67	4.00	FAR
90	40	55%	0.70	7.00	FAR
90	45	50%	0.80	Inf	FAR
90	50	45%	0.90	Inf	FAR
90	54	40%	0.95	Inf	FAR
90	0.37	37%	0.97	Inf	FAR
90	60	33%	1.00	Inf	FAR
90	67.5	30%	1.33	Inf	FAR
90	72	17%	2.00	Inf	FAR
90	80	11%	3.50	Inf	FAR
90	90	0%	Inf	Inf	FAR

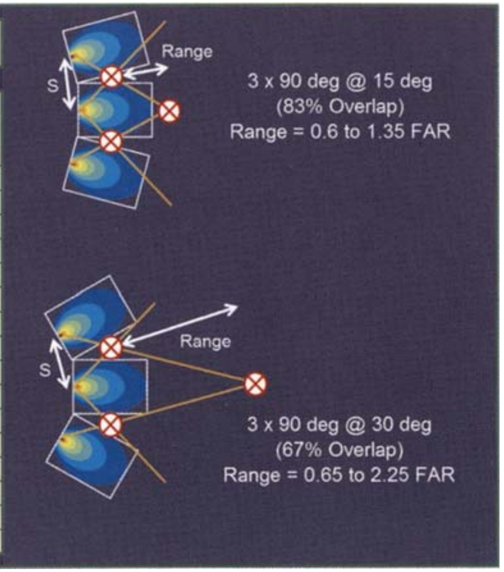


Figure 7,21 Design reference for the symmetric uncoupled point source

range of such systems is extremely limited and must be bordered by systems with superior means of minimum variance coverage.

Note: this is not to be confused with stereo which consists of separate channels of information. The center-panned signals of a stereo system will, however, behave as a symmetric uncoupled point destination.

1. Determine the target area of coverage extension to be provided by the fill system.
2. Orient the fill speaker for minimum variance coverage of the target area.
3. The intersection of the on-axis lines of the elements provides the forward aspect ratio scale length. The usable coverage of the array is limited to 50 per cent of the FAR length, due to ripple variance.

Determine the relative distance range required of the fill (as compared to the mains) and scale the power capability as appropriate.

Asymmetric Uncoupled Point Destination

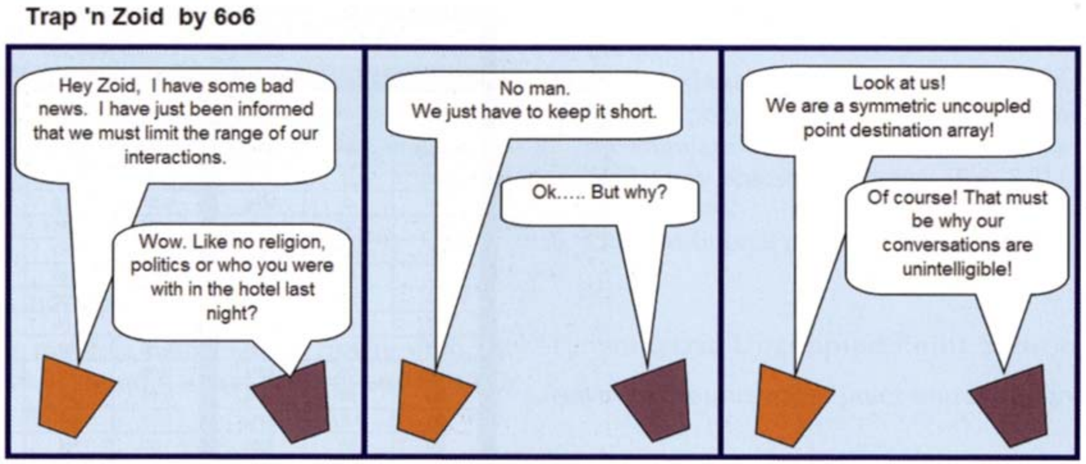
Asymmetric uncoupled point destination design procedure:

1. It is assumed that this is a fill system which is being combined with a previously defined main system.
2. Determine the target area of coverage extension to be provided by the fill system.
3. Orient the fill speaker for minimum variance coverage of the target area.

Determine the relative distance range required of the fill (as compared to the mains) and scale the power capability as appropriate.

The Diagonal Plane

Our sound system will not perform as paper thin slices in the vertical and diagonal planes. Back in Chapter 5 we discussed the issues of 2-D renderings in a 3-D world. It is now time to put these practices to work. The issue arises



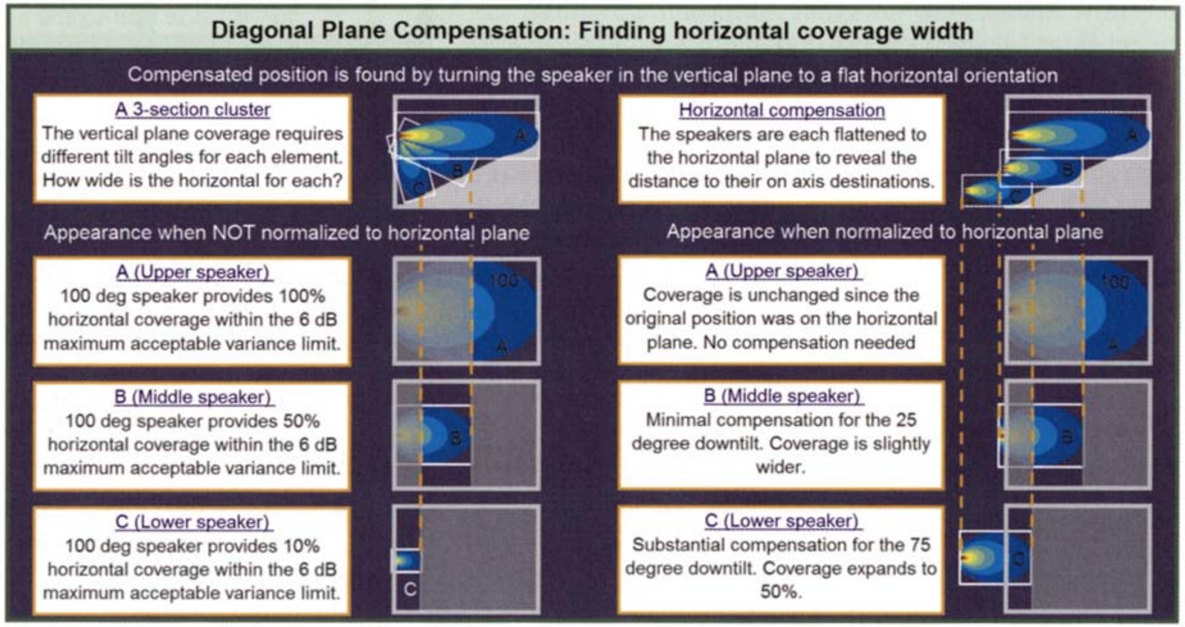


Figure 7.22 Diagonal plane design compensation

when the speakers are propagating along a plane different from our 2-D viewpoints. As the propagation planes move toward the diagonal of our rendering plane our need to compensate increases. The poster child for this is the asymmetric point source so we will use it as a starting example.

Let's place a four-element version as a center cluster in a rectangular "shoe box" type concert hall. The coverage requirement, expressed as a width (in meters/feet etc.) is the same for all sections. The horizontal coverage expressed as an angle is different for each. Each section

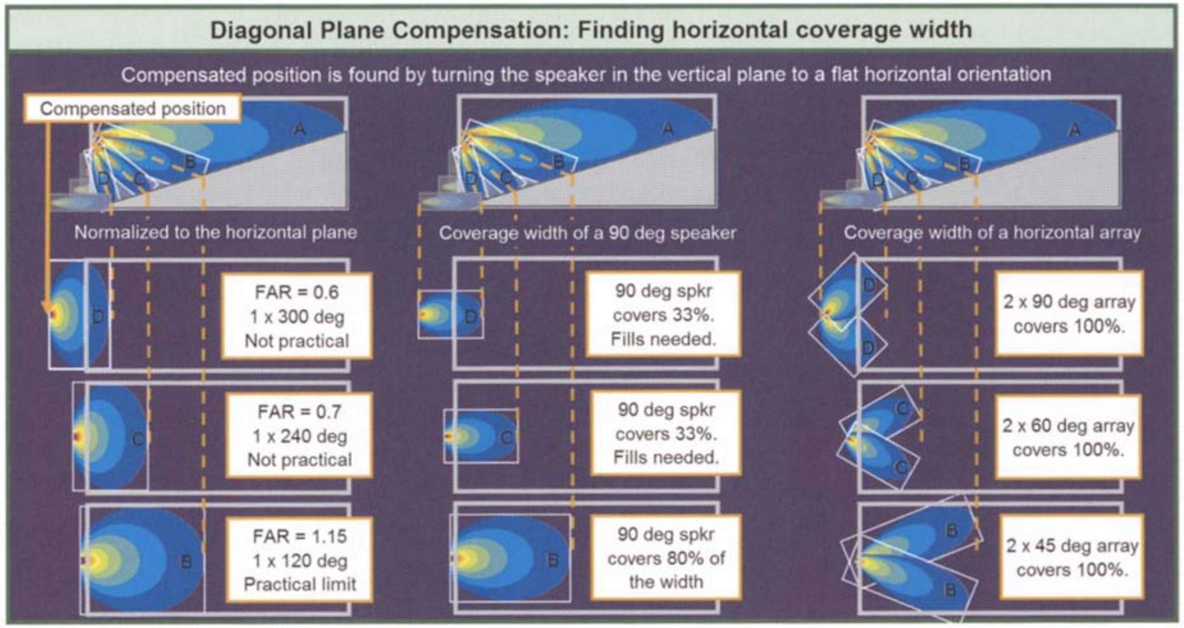


Figure 7.23 Diagonal plane design compensation. Alternative

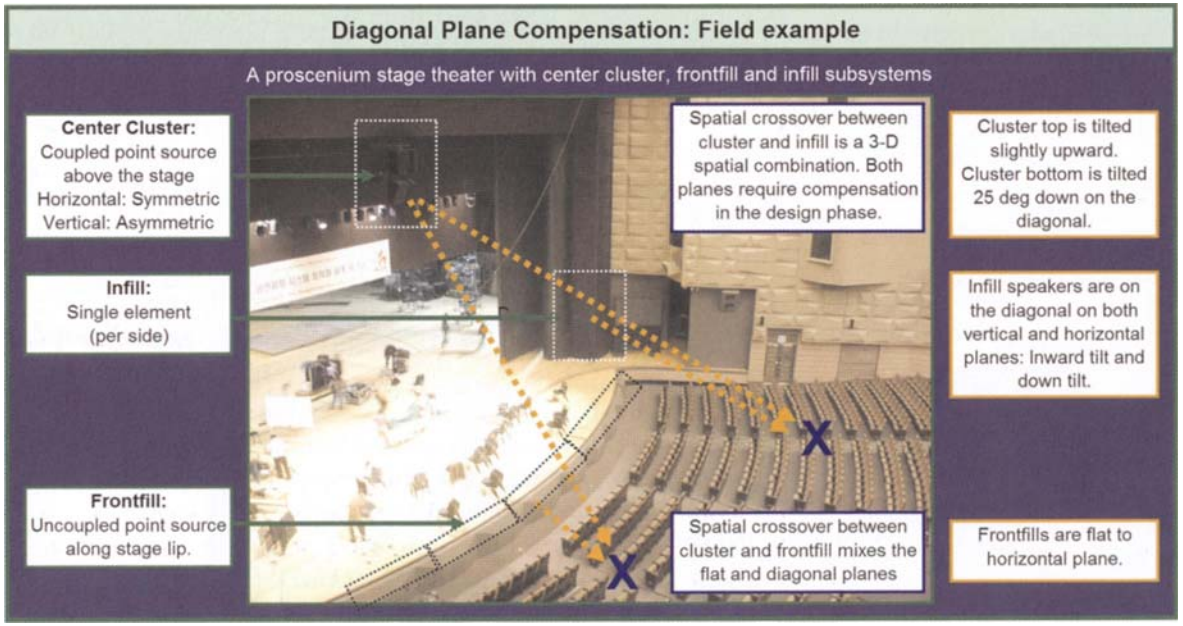


Figure 7.24 Diagonal plane design compensation. The implications of diagonal plane propagation are shown in this field example

sees a different aspect ratio since the length (range) of coverage is changing over the fixed width. As we move down the cluster from top to bottom we can expect to need a wider element at each incremental break.

The design calculation for this will require triangulation compensation of the speaker placement as described in Chapter 5. We will now put this to use on some example clusters (see Figs 7.22-7.24).

Balcony Battles

I don't like to admit defeat, but as the old saying goes, "It is better to run away and live to fight another day." There is one standard room shape for which there is often no single point solution: the "W." The "W" turned on it side is the shape that our system sees when faced with a deep balcony overhang. If viewed from above, the lower area suffers from path obstruction. We must shoot through the people in the front of the balcony to try to cover as much of the floor rear as we can see. If viewed from below, we must shoot through the faces of the people in the lower balcony to get to the back. If we start from the middle, we face a geometric puzzle that might require M.C. Escher to do our drawings. And once again the people at the balcony front take the full frontal assault.

It should not be surprising that this geometric double take would be the Waterloo of the coupled point source. The shape is a pair of vertically stacked right triangles. We know the shapes of minimum variance and there are none that allow the sound to throw deep, go short and then repeat. Such a shape can ONLY be done by an uncoupled pair of asymmetric coupled point sources.

The insurmountable obstacle is not the HF range. Even a relatively small amount of angular isolation will provide the beam steering needed to split the HF above and below the balcony front. It is the mid-range and below, intractably stuck in the center, that sends it crashing on to the front of the balcony.

There are three players in the balcony battle: mains, over-balcony delays and under balcony delays. The mains may be a single point or multiple main configuration. The question before us is to determine which combination of these will be the most effective. The first determination will usually be the mains, since the role of the others will be to supplement those areas not optimally covered by the mains. As we will see, however, the role of the fills may play a part in determining which of the mains scenarios will work best. There is a symbiotic relationship here. As usual the decisions will be tradeoffs.

- Single main: best for minimum ripple variance, worst for level and spectral variance.
- Multiple main: worst for ripple variance, best for minimum level and spectral variance.

The extent of the differences can be evaluated by rating the balcony in dB. This is done by observing the proximity ratio. Notice first that we will need two proximity ratio calculations: back to front upstairs (over the balcony and back to front under the balcony). It is this dual proximity that is the source of all the trouble, of course, since the coupled array is not capable of doubling back its response over its full frequency range. We are far, then near, then far again, then near again. The key to this evaluation is the amount of "comeback" required which we will quantify as the **return ratio.** If this is too high, then splitting the array is indicated.

The return ratio is the number of dB closer we find the front of the balcony to the farthest area angularly adjacent to it. Where does the sound go when it skims just above, or below, the balcony front? If it has to go twice as far, we will have a 6dB proximity challenge to overcome in a degree or two of angular change. Figure 7.25 shows the process of assessing the return ratio. The position of the main cluster will change the return ratio for a given shape. Mains positions above the balcony will use the balcony front and the last floor level seating with visual contact to the mains. Mains positions below the balcony will use the balcony front and the last upper level seating with visual contact to the mains. If the mains are centered at the level of the balcony we will use the balcony front and the last upper or lower level seat, whichever is farthest.

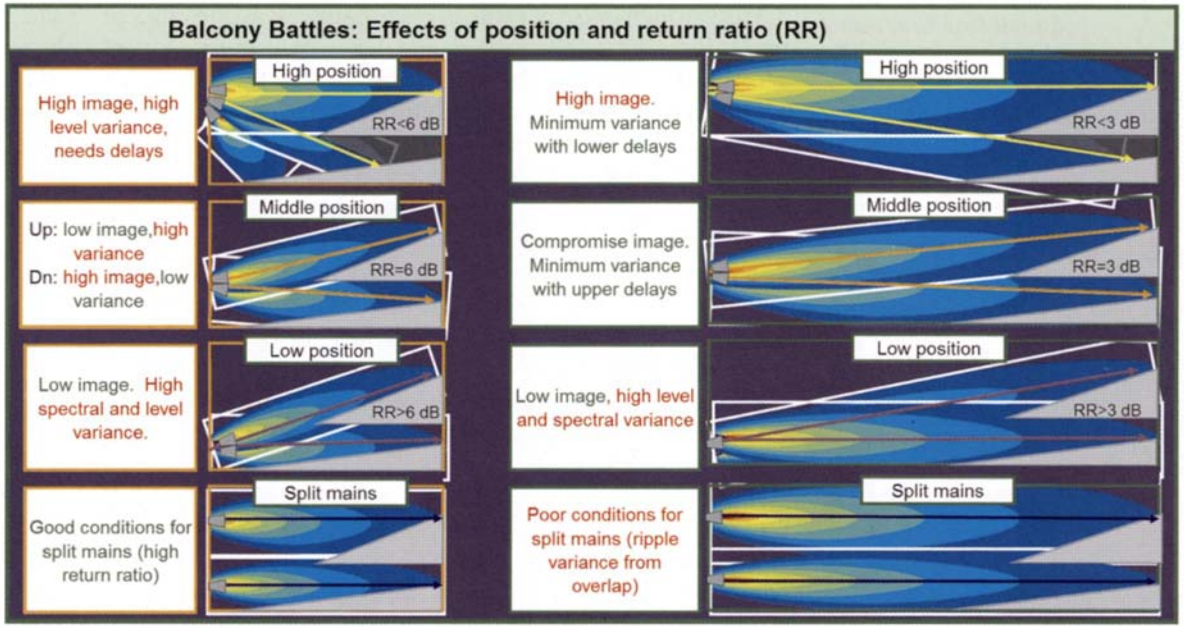


Figure 7.25 Balcony battles. As the return ratio rises, the use of a split mains configuration becomes preferable

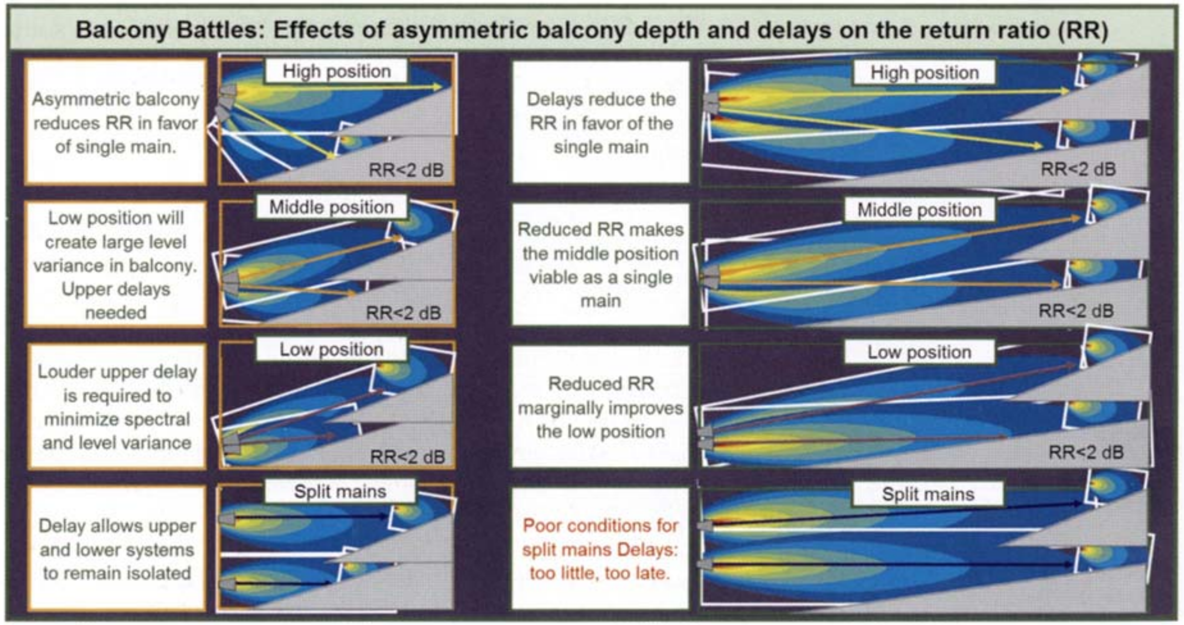


Figure 7.26 Balcony battles. An asymmetric balcony configuration can alter the return ratio, and affect the decision whether to split the mains or add delays

How much return ratio is acceptable is a question of tradeoffs. Obviously 6 dB is the limit of our maximum acceptable variance. The tradeoffs may be a host of practical matters as well as sound image and variance. In any case it is important to be realistic about the fact that we can make very slight returns on a coupled cluster. There are various methods out there to attempt this, such as "balcony bars" that provide some additional spacing or unplugging some of the elements in the array that are facing the balcony front. There is no arguing with the fact that a speaker facing directly into balcony fascia should be turned off if other speakers can still get the coverage where we need it. This is, however, a very poor solution in most cases. First, it is very unlikely that the balcony front runs along a line that is at the identical vertical orientation to any element in the array. This is a 3-D geometry issue. If the cluster is at the balcony level and the speaker is facing at zero degrees the probability is extremely high that the angle will be constant and the speaker should be deleted. If the cluster is above or below the balcony the probability of planar convergence is up there with a harmonic convergence of Jupiter, Saturn and Mars.

The other options are tapering the angles and levels down and then trying to put the genie back in the bottle by closing the angles and raising the levels. This will take care of the very high frequencies but the MF and LF range will not cooperate.

Here is where the fill speakers come in. First we need to evaluate when it is that fills are mandatory This is far more likely in the single mains scenario, since line of sound to the speakers will be cut off in many instances. The split mains should have clear lines of sound to all locations. The fill speakers have the ability to rescue the single mains by virtue of reducing the coverage range. Over-balcony speakers will reduce the upper range, which may allow a lower mains position with a better return ratio. Likewise under-balcony speakers will allow a central position to reach the distant upper areas without the need to return to pick up distant lower areas. The combination of both will ease the burden further and up the probability of a successful single coupled mains. This has its price however: TANSTAAFL. The virtues of low ripple variance in the single mains must be weighed against the high ripple variance likely from the interaction of uncoupled under- and over-balcony systems. The split mains has one uncoupled spatial crossover. If the cost of keeping the mains in one piece is two vertical uncoupled crossovers and piles of uncoupled crossovers in the horizontal, the deal does not look so sweet.

Next let's consider the conditions where a delay fill system, while not mandatory, is potentially useful. The following questions should be considered.

Balcony delay considerations:

1. Do we need delays at all?
2. Will the delays improve or degrade the coherence?
3. Will the delays improve or degrade the imaging?

The question of need arises from the shape of the under (or over) balcony space. It is not a simple matter of depth. For instance we cannot simply say that we need delays if the under balcony is 20 meters deep. This is not the case if it is also 20 meters tall ! Once again enter the aspect ratio. The depth to height ratio will be a key but not a solitary factor. As the aspect ratio rises the probability that we will need delays goes up but there are additional factors.

A big piece of the balcony battle is doubling distance. Distant sources (like our mains) reflect off the same room surfaces as the local source (our under balcony rescue team). The difference is doubling distance. The round trip from the back wall is a small proportion of additional distance as seen by the main system. The milliseconds add up, but the level loss is minimal, a troublesome combination. The perspective of the delay speaker would see the same number of milliseconds to take a round trip via the back wall but sees that as a much larger proportion of its path length. Once again level changes as a ratio while time changes as a difference. The result is that the reflections from the mains are much stronger than we would find from the delays, even though they are hitting the same surfaces. Therefore we need to factor in two aspect ratios: the mains to the back wall, and the under balcony speaker

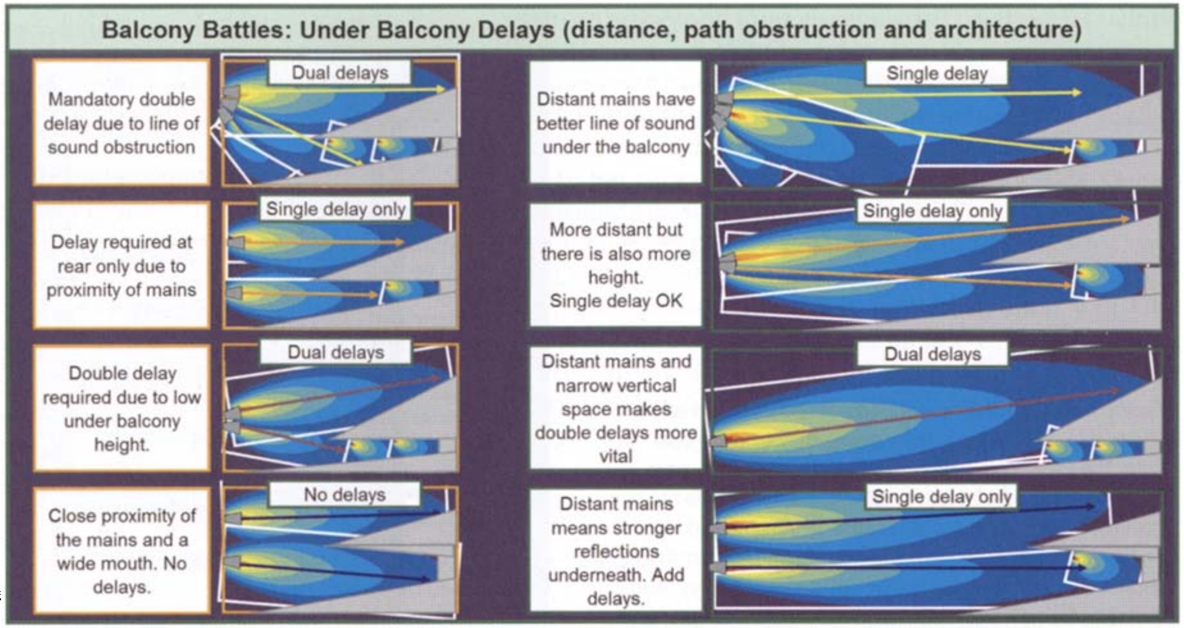


Figure 7.27 Balcony battles. Multiple factors must be considere

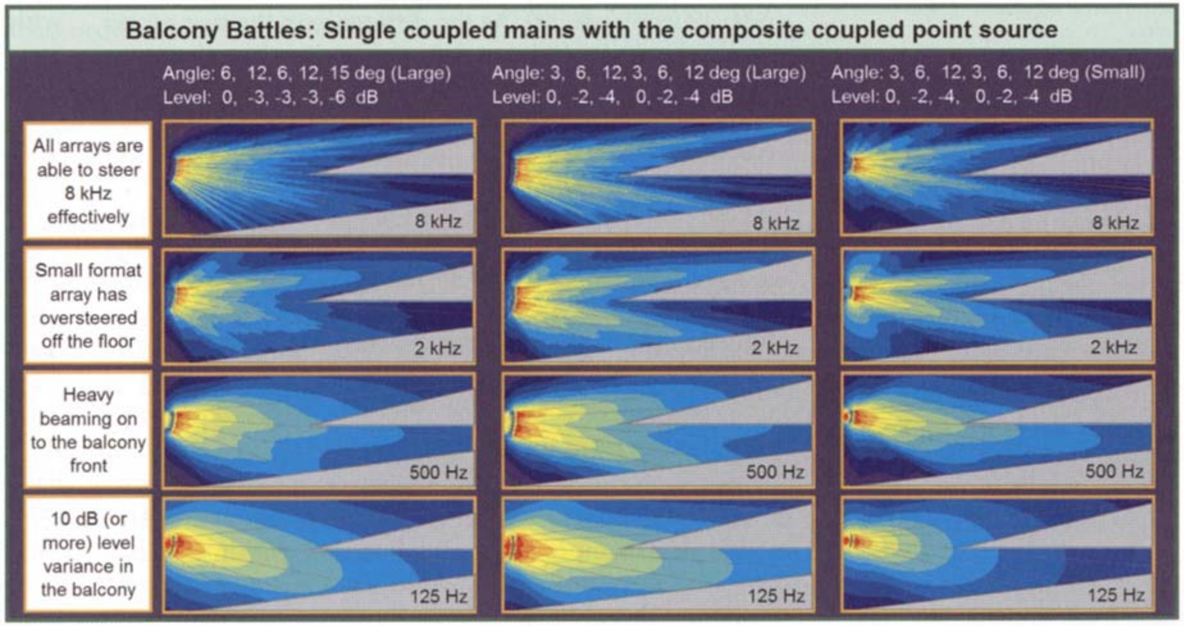


Figure 7.28 The coupled point source main system attempts to split itself and bend around the balcony front. This can only be maintained for a limited frequency range. As frequency falls the beaming at th

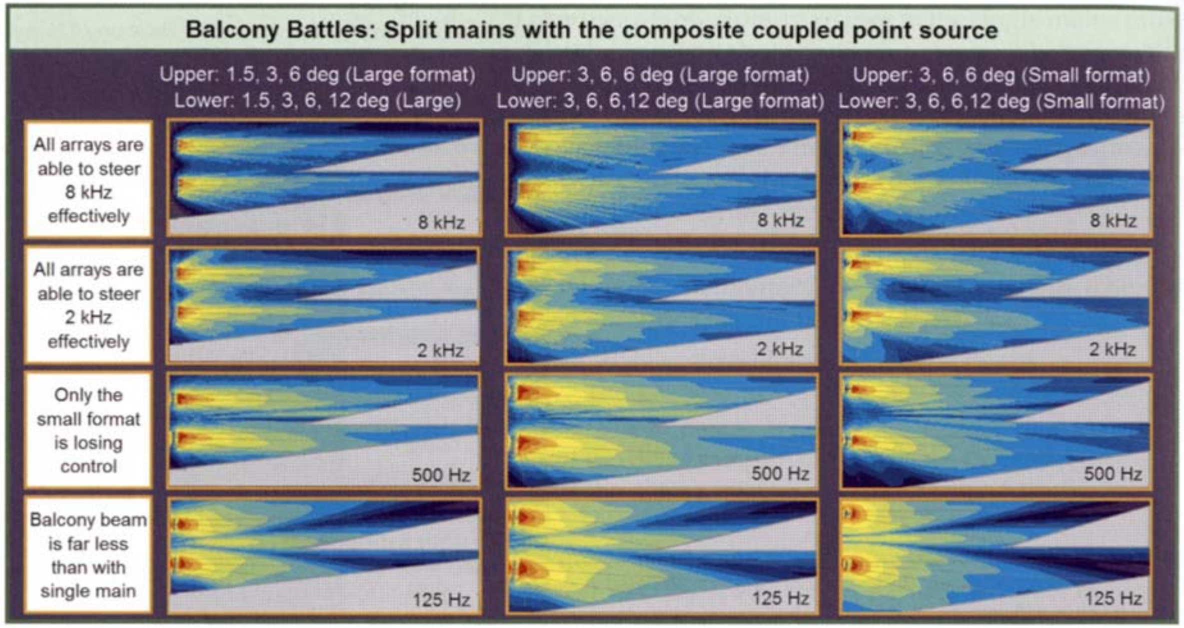


Figure 7.29 A split pair of coupled point source main systems attempts to split itself and cover above and below the balcony front. The separation can be maintained over a much wider frequency range. As frequency falls the beaming at the balcony front rises, but the extent of the rise is small compared to the single coupled array shown previously. The smaller length arrays show more center balcony overlap but still have far less beaming than the single coupled system

to the back wall. Both would use the same height (the under balcony height). As the difference in the two aspect ratios rises, the need for delays increases. So the same shape of under balcony would have a higher need for delays the farther away it gets from the mains. This is not because of the direct sound loss. It is because of the lack of loss in the reflected sound. There is often a misconception that the under balcony area has the HF range rolled off. So long as the line of sound is clear this is not the case. What we have instead is pink shift due to LF addition because the reflections are strong. Fig. 7.27 shows some examples of the process of determining the need for delays.

A third factor is the absorption (or lack thereof) of the under balcony surfaces. Predictably as absorption decreases our need for delays increases.

Now on to the question of whether the delays can help or not. It will do us no good to call the rescue team if they are not able to improve the situation. There is more to this than first appears. TANSAAFL. Delays are uncoupled arrays and there is no guarantee that their interaction with each other will not create as much ripple variance as that which they are trying to suppress. The potential for this increases when the delay speakers are too far forward of the position where they are needed, are spaced too closely, or have nearby reflective surfaces that degrade the signal. If the positions we are allowed to use are too distant, the mains will do better without any help from the delays.

Multichannel Sound

Much has been made here of the balancing act between concentrated power and minimum variance. As long as we continue working on the same basic assumptions, there are only incremental improvements possible as we find ways to minimize displacement in ever more powerful packages. There are, however, alternative strategies available, one in particular that has been tried long ago and mostly forgotten.

The conventional high-power sound system mixes the huge number of individual input channels into just two high-pressure pipes: left and right. The sound system must reproduce the entirety of the signal presented and uses huge amounts of acoustic overlap to achieve the required SPL. Because the signal is electronically pre-combined, the speaker system must be able to meet the power requirements of all of the channels, all of the time. This requires a lot of headroom, which is achieved by acoustical overlap, which costs us variance. TANSTAAFL and triage again.

But who says it is mandatory to do all of the combining in electronic form? Stereo, as we know, combines in the acoustic space, as do all forms of natural sound. Why limit ourselves to two channels? Why not have a separate speaker system for the vocals, the guitars or whatever?

The concept is the separation of the power and coverage requirements into a true multichannel system, an alternative form of isolation. We can take four channels of stage sources and combine them electronically into a highly overlapped four-speaker array. The signals are mixed electronically and summed acoustically. We could alternatively separate the four channels and send them individually to each speaker. The signals are mixed acoustically and no stable summation occurs. Is there a difference? Yes, and no. The total acoustic power available is the same (the laws of thermodynamics). The power available on a per/input channel basis is reduced, but the perceived loss in a musical context may be far less than one would assume. Have you ever considered how it is that a 50 W Marshall stack can keep up with our 10 kW sound system? Our sound system has to meet the power needs of the entire mixed waveform of the guitar and everybody else. If we muted every channel except the guitar, our PA will be able to blow away the Marshall stack, no problem.

Multichannel sound in the cinema is nothing new. A center vocal channel is used, along with stereo music and effects. This is done to minimize ripple variance in the voice channel, which would decrease intelligibility. We can do the same in our concert sound systems, in the form of a left, center, right configuration. Each channel must cover its required area, which for the center system is the entire room.

We can go beyond three channels. In fact, there is no limit. We can have a complete sound system for every instrument on stage and do all of the mixing in the room, the most similar configuration possible to natural sound transmission. This has been implemented, long ago, which brings us back to the 1974 Grateful Dead concert at the Des Moines Fairgrounds that was discussed in the Preface. This sound system had completely separate sets of speakers for each string instrument, the drum set and the vocals. The sound was mixed acoustically. Their implementation had fatal flaws, such as the minor issue of all of the speakers facing directly *into* the vocal mics, but the fundamental concept of multichannel separation is potentially viable.

There are a number of developments since 1974 that create a more favorable environment for multichannel mixing.

1. Affordable and flexible multichannel mixing consoles and signal processing.
2. Huge increases in stage source isolation that allow us access to single source channels entirely free of acoustic leakage from the other stage sources.
3. In-ear monitors providing reduced leakage from the stage in to the listening area.
4. Speaker systems that have extremely smooth, wide horizontal coverage and controlled asymmetric vertical coverage, hung with ease from one or two rigging points.

If we have forty-eight of the third-order speakers in the budget we have some choices ahead. We can go stereo (2 X 24) with lots of overlap. L / C / R (3 X 1 6) with more separation and less overlap or subdivide further and further as long as we can meet our coverage needs and not fall short of the power requirements of the individual channels.

This is not something to be entered into lightly. It requires careful forethought in regard to which channels are to be mixed electronically and which are to be mixed acoustically. The key to a successful multichannel implementation is electronic isolation of the source channels and the saving of their combination for the acoustic medium. If isolation is not achieved there will be a host of unexpected summation interactions between the related signals as they arrive from the different speakers. If source isolation can be achieved the possibilities are unlimited. Not only can we have unlimited power with minimum variance, we can start playing with spatial movement all around the space. Now the fun begins!

There is one final consideration in regard to the multichannel paradigm: the use of multichannel mixing precludes us from being able to obtain accurate stable acoustical analysis data of the system in its performance context. If we are mixing the sound acoustically, we will not be able to compare the combined signal to an electrical reference, and therefore we will be unable to perform ongoing optimization with an analyzer during performance.

The process of measuring the sound system will now become the focus of our study.

Section 3:
Optimization

Examination

examine *n.* **1.** *detailed inspection (of, into).* **2.** *testing of knowledge or ability*

examine *v.t.* **1.** *inquire into nature, condition etc., of, test by examination, theory, statement; formally question*

Concise Oxford Dictionary

Examination Defined

The installed audio system is not yet ready for operation. We cannot assume that everything is in perfect working order, nor can we expect that fine-tuning will not be required. We must test the theory behind the system design and question whether the installation has been carried out to meet our specifications. This process of fact-finding regarding the installed system will require examination, and the process of examination requires measurement tools. These are distinct from the prediction tools found in the previous section. Measurement tools examine what is, not what will be. The device to be examined must be physically present. Examination tools do not assume, and cannot be coerced. They do not favor a particular manufacturer or design approach. They will prove us to be brilliant designers and meticulous installers, or dreamers and pretenders, just the same.

Without examination, superstition and science are on the same footing. Examination puts an end to circular discussions about what might happen if we change something and or what might be the cause of some particular problem. The theories can be tested, the questions resolved and we can move forward having learned something.

Examination tools take a variety of forms, from the simple physical tools used to measure a speaker angle to the complex audio analyzer. Each has a role to play in the big picture. The most challenging to us is the audio analyzer, the diagnostic tool we will employ to monitor the variance in the electronic and acoustic signals as they move through the transmission path to the listener. We are listening to the transmission. A key role of the analyzer is to help us understand what we hear. These tools inform the designer and optimization engineer about the degree of level, spectral and ripple variance over the space and tell us how well the original signal has been preserved.

Fortunately we have a reference copy of the original state of the signal as it left the mix console and began its hazardous journey to the listener. This gives our audio measurement system a target. The measurement system will monitor the sound system for any changes to the original signal and report those to the operator. But like any diagnostic tool, its role is merely to show the symptoms to the trained observer, not to cure it. No disease has ever been cured by a thermometer or the advanced technologies of X-rays or magnetic resonance imaging (MRI). No sound system transmission error will be cured by an analyzer. A thermometer is easy to read but has a very limited, albeit useful, diagnostic scope. An MRI is immensely powerful, with extremely high resolution and the ability to produce data far beyond what our senses could discern directly. Such a system requires specific technical training to operate and advanced knowledge of anatomy to interpret the data well enough to make the correct diagnosis. Even more advanced training in medical science is required to create a successful treatment plan. This will require a doctor (and of course, an insurance agent).

We would all prefer that sound system measurement tools have the ease of operation of the thermometer and the power of the MRI, but this is not the case. The challenges faced by operators of modern audio systems are of such complexity that one-dimensional diagnostic tools will not yield sufficient data to detect the causes of errors. This, in turn, makes it unlikely that such errors can be successfully treated. The necessary analyzer is the equivalent of an MRI for our sound systems. It requires training to operate, even more so to interpret the data. With this tool we will be able to penetrate the invisibility of sound with startling clarity. It can be, but is not always, a pretty picture.

This section focuses on the tools of the sound system examination trade. These range from a folded piece of paper to the complex audio analyzer. Overall, the tools fall roughly into three categories: physical, electrical and acoustical. Some devices can be used for more than one purpose. Each of these tools has their uses and each has limitations, strengths and weaknesses.

Physical Measurement Tools

Inclinometer

This device reads the vertical angle of a device or surface. It is a simple mechanical gadget that uses gravity to position a rotating arrow to indicate the inclination (hence the name) on a fixed scale in degrees. Language destroyers like me have been known to refer to this as an "angle-ometer."

The role of the inclinometer is to quantify tilt angles for speakers. The final values for tilt angles will be derived by acoustic measurement but that does not negate the utility of the inclinometer. It is extremely useful for providing the initial angles, derived from drawings, etc. Another application is the copy and paste: after a speaker is measured and the optimal angle found, the inclinometer can ensure that the chosen angle is copied to symmetrically equivalent speakers.

Figure 8.1 Inclinometers used for speaker vertical angle adjustment and verification

Protractor

Horizontal angle determination is more challenging since gravity cannot be used to solve the difference. A protractor is able to measure the angle between surfaces. The limitation of protractors is their physical size. If they are too small their accuracy is poor. If they are too large they can be unwieldy in the tight confines of a speaker cluster.

Origami Angle Finder

This piece of test equipment is the only disposable and recyclable item in your optimization toolbox. This is, of course, the folded piece of paper, and is used to find the angle between the sides or centers of splayed speakers. Paper is manufactured with very reliable right angles. The

Figure 8.2 Protractor used for speaker horizontal angle adjustment and verification (courtesy of Starrett Corporation)

paper can repeatedly be folded in half to easily create 45, 22.5 and 11.25 degree angles. A little practice is required to master folding in thirds to create 60, 30 and 15 degrees. The paper is then placed between the speaker sides and the desired angle found. The folding is easier and the finished shape more symmetrical if the paper is first trimmed to a square shape. This is not a joke, and is often the only way to get an angle in the field. When we are leaning out of a Genie lift, trying to work our hands around the rigging frame and the speakers are jammed together, there is often no practical way to get another tool in there. Also, it is an absolute certainty that someone at the job site can furnish us with this tool.

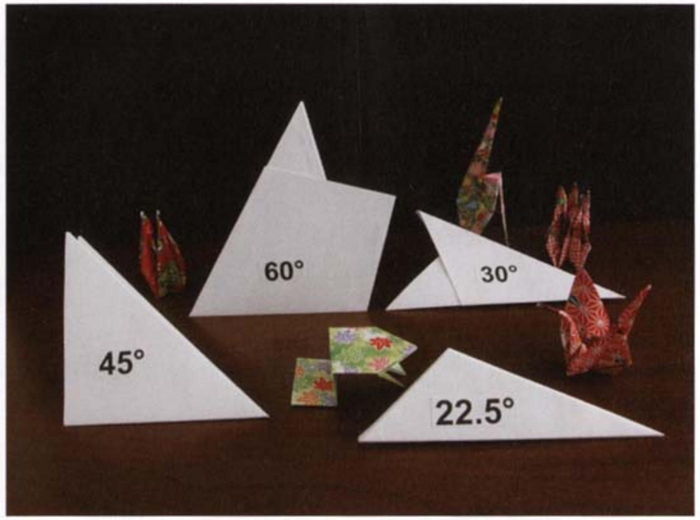


Figure 8.3 The origami angle finder

Laser Pointer

A reliable and accurate laser pointer can be used to aim speakers with pinpoint accuracy. Laser pointers manufactured for surveying or construction professionals are sufficiently accurate for our purposes. There are even some advanced lasers that are purpose-designed for audio. The hand-held pointer used for lectures does not have sufficient accuracy for our application. Here is a handy test for the presentation laser pointer: place it on a speaker and roll it along the top of the box. If the aim is accurate the laser dot will move in a horizontal line. If not, the device should not be used.

The pointer is placed on the speaker in a manner consistent with its on-axis orientation. It is often not physically possible to place the laser exactly in the center of the enclosure so a small displacement must be allowed for. It is worth noting that the displacement does not scale and is not angular. If the laser is 0.5 m above the cabinet center, then the on-axis point is 0.5 m below the spot where the laser dot is seen. The speaker can then be positioned to obtain the desired focus point. Like the inclinometer, the

Figure 8.4 Laser pointers. The upper unit is manufactured to tolerances suitable for optimization. The presentation quality device is not suitable

laser is very useful for the copy and paste of symmetrically opposite speaker aim settings. The laser can also be used to view the position of other points in the speaker response such as the off-axis point. This can be helpful in the evaluation of speaker focus angle.

There are advanced scanning lasers that create an adjustable width fan shape which can be set to the speaker's coverage angle. In this way the full extent of speaker coverage can be seen at once.

Another laser-based device is the laser distance or range finder. If this device is able to find its reflection from a surface, the device can determine the distance traveled. This can provide the approximate distances between the range finder and other objects in the room. Dark objects, such as black speaker enclosures, give a very weak reflection which limits the reliability of the reading. The laser device as a distance/delay finder is fine for approximations but should never be thought of as an option to replace the acoustic measurement of delay times. There are better tools for this as we shall soon see.

Thermometer

Since the speed of sound is affected by the temperature, it stands to reason that a thermometer will provide relevant information. The accuracy level of a typical thermometer should suffice for our purposes. Few optimization decisions will be based on the readout of temperature. For permanent installations the temperature at the time of the initial optimization is worth noting, particularly if it is not within the expected range of daily operation. In touring optimizations the temperature can be monitored on an ongoing basis and the changes in speed of sound anticipated and compensated.

Hygrometer

Since the humidity percentage will affect our high-frequency transmission this factor will also be worth monitoring for the same reasons as above. The **hygrometer** is an instrument that measures humidity. Humidity readings can be observed and changes in the high-frequency response can be expected and compensated.

Simple Audio Measurement

Tools

There are many types of electronic measurement tools available to the audio engineer. Most of these devices are principally concerned with the measure of voltage and resistance and, to a lesser extent, current. The analog audio signal is transmitted as voltage change over time. Therefore, characterizing this value will be our main focus.

Volt/ohm meter (VOM)

Volt/ohm meters (VOMs) are a mainstay of any installation or maintenance-oriented audio engineer. The VOM provides AC and DC voltage testing and continuity and short circuit detection. AC voltage and current measurements can be average measurements calibrated as RMS, or true RMS (root mean square). For sine waves this makes no difference but for music and noise, there will be a difference in the readings. The VOM's principal role is in the pre-verification stage of the system.

Polarity Tester

The polarity tester, also known as the "phase popper," uses a pulse generator and a receiver device. The generator drives the line and the receiver decodes the electrical (or acoustical) signal at the other end. The electrical readings can be reliable as long the response through the measured system is flat over the full frequency band. If not, the phase shift in the circuit can cause the receiver to incorrectly decode the impulse and give a false reading. The potential for error rises tremendously when the receiver is acoustical. The band-limited nature of all loudspeakers, even full-range ones, inherently causes phase delay over frequency. Each speaker in a multiway system must be tested individually, which potentially adds the phase shift of the electrical crossover into the readings. Add axial response and the room acoustics into the equation and we begin to see how heavily the deck is stacked against the simple green/red readout of the polarity tester. This device is perfectly fine for line testing, but should only be used as a preliminary check for speaker verification. There are far better tools for this as we shall see.

Listen Box

This is an audible audio test device. The box is a battery-powered miniature amplifier and speaker that listens to the signal anywhere in the path. The high input impedance minimizes loading of the line. This type of technology originated as a telephone line troubleshooting system. The phone system device uses an inductance coil to listen in on a line even through the insulation. The limited frequency range of an induction coil, while sufficient for telephone monitoring, is not full-range enough for professional audio. Therefore, the listen box needs to make at least an unbalanced connection to function. The listen box is an extremely fast and efficient tool for locating where your signal is, and isn't. They are low-cost and most useful in the pre-verification stage. A high impedance headphone amplifier and a set of headphones is a good substitute.

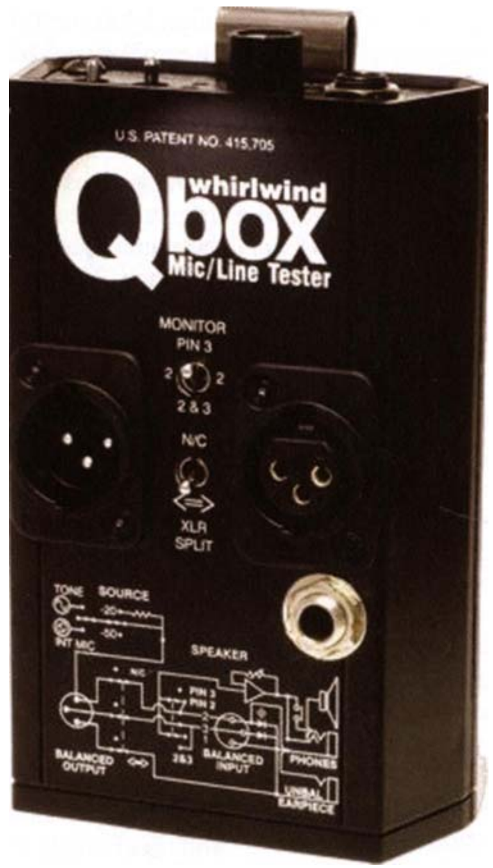


Figure 8.5 Listen box. Headphones or speaker can be plugged into the output. Any point in the signal path can then be aurally monitored (courtesy of Whirlwind USA)

Impedance Tester

An impedance tester differs from a VOM in how it measures audio lines. The VOM, discussed previously, can measure the DC resistance of a circuit. Audio signals are alternating

Perspectives: Technology makes it easier for some to ruin something that sounded great, and they can save time as they ruin it, while it makes it harder for others to decide what is the best way to do something when there are so many good options.

Martin Carillo

current (AC) and create a different type of load on the circuit, known as a reactive load. A reactive load consists of DC resistance and capacitive or inductive reactance. The combination of these is termed the impedance. Loudspeakers are rated by their impedance, such as 8ohms, but if measured by an ohmmeter might read *6ohms* (the DC resistance). The impedance tester will more accurately reflect the load as the driving source sees it. This test reveals the fact that the impedance is highly variable over frequency. For this reason all impedance test results must specify a frequency. In most cases the ohmmeter is sufficient to establish continuity or to detect short circuits. If transformers are in the signal path they will appear as short circuits to DC readings. The impedance tester will see the transformer impedance and give an accurate reading. DC blocking capacitors installed for HF driver protection will appear as an open circuit to a VOM, while an impedance meter will see through them to the driver. An impedance tester is highly recommended for 70 volt systems.

Oscilloscope

An oscilloscope is a waveform analysis device. Electrical signals are displayed as voltage over time. Both axes are independently adjustable, allowing for the monitoring of virtually any electrical signal. DC voltage, line voltage and audio signals can be viewed. The oscilloscope tracks the waveform, therefore the voltage levels are peak-to-peak representations of the AC signal (see Chapter 1). Oscilloscopes can monitor amplifier clipping, oscillations and much more. These are not usually used in system alignment, but are a useful troubleshooting and verification tool.

The role of the oscilloscope in system calibration was much greater in the past. Its ability to measure amplitude versus time allows for the analysis of phase relationships over frequency. Delay setting and even echo identification could be performed with an impulse fed into the system and monitored on the oscilloscope through a microphone. Fortunately there are easier ways to do this in the current era, as we will see.

Sound Level Meter

A sound level meter provides a single number to describe the dB SPL at a given location. The reading can cover a particular range of frequencies, and a range of time. The SPL readings are given particular subunit ratings that correspond to the user settings for frequency weighting and time constants. These units were described in Chapter 1. The sound level meter is an operational tool, not an optimization tool. There are no calibration parameters for which the sound level meter is not out-performed by others. There are simply no optimization parameters that can be answered by a single numerical reading such as found on the readout of a sound level meter. This includes level setting, for which the sound level meter is only suitable when setting symmetrically matched systems. The sound level meter is, however, the foremost tool for mix engineer bragging rights and outdoor concert law enforcement.

Real-Time Analyzer (RTA)

The real-time analyzer has a number of applications in which it is the *best* tool for system optimization: zero. There is nothing that the stand-alone RTA can do for us that cannot be done better by other analyzers. Signal path, noise floor, hum, THD, polarity, latency and frequency response analysis are the mainstays of the verification process. Equalization, delay setting, level setting, speaker focus, and room acoustic property analysis are the mainstays of the calibration process. The RTA is inferior in all of these categories to more advanced tools.

This is not to say that the functions of the RTA can not be put to use. It can serve as an introductory ear/eye training tool. We can listen to music and match what we hear with what we see. Feedback can be induced and identified to train us to be able to hear it before it gets out of control. But even such time-honored functions of the RTA can be replicated by modern FFT analyzers that duplicate the functionality of the RTA without being burdened by the RTA's inherent limitations. The RTA, like so many things, has been replaced by a computer.

In the past the RTA was the only spectrum analyzer in use for most sound reinforcement applications. It was affordable, small and easy to operate. These have now been replaced with dual-channel FFT analyzers, which are affordable, small, but require training to operate. The reason the FFT analyzer has taken over is engineers have learned that having highly accurate data is worth the trouble it takes to learn how to operate these systems. What is an RTA, and what is it that makes it so inferior?

The real-time analyzer is a bank of parallel log spaced band pass filters (octave and l/3rd octave are typical) at a standard set of center frequencies. The output of each filter is sent to a full-wave rectifier circuit that creates a waveform which represents the absolute value of the filtered waveform. The next step is integration, the process by which a DC value is derived from the absolute value. The integration time constant, derived from a resistor/capacitor (RC) circuit, can be set to fast (250ms) or slow (1 s). The end result is that the thirty-one l/3rd octave bands each contain an integrated value which represents the average value in that frequency range in the current stretch of "real time." The term "real time" connotes that the displayed values represent a continuous stream of time, with no gaps as might occur with an analyzer which takes samples of time. Unless the RTA is paused the data continually streams in time.

The RTA is designed to show a flat response when driven by a source containing equal energy per l/3rd octave, i.e. pink noise. Since pink noise has random components, the RTA requires averaging to settle down to a flat response. The RTA can make coarse frequency resolution evaluations of the system response. But l/3rd octave resolution is far too low to be considered for equalization use, as will be discussed later in this section. Low resolution data is only sufficient for measuring the noise floor of a device, or exploring gross response trends such as traffic noise in urban areas.

The crippling limitation of the RTA is its rendering of its middle name: time. It can not measure time, so it is useless for setting delays or identifying reflections. It has no phase response and therefore we cannot understand the nature of speaker summation, which, as we know, leaves us blind. It cannot discern between multiple arrivals within its integration time, and therefore cannot separate early arrivals from late. Thus it has no means of mimicking the ear's perception of tonal change vs. discrete echoes. Because it sees all of the energy in the room, regardless of timing, the low-frequency addition is greatly exaggerated above the ear's perception. The longer the reverberation time in the room the more the RTA low-frequency response hangs on. If the response is equalized to flat in a reverberant space, the system will sound as if the low frequencies are massively deficient. An RTA user only makes this mistake once. After that they learn not to trust the RTA. Unfortunately the failings of the RTA caused many engineers to close their minds to audio analyzers as a whole. "Analyzers? We don't need no stinking analyzers" became a familiar position taken by front of house mixers.

The RTA presents a one-dimensional reading of the complex questions of the sound system response: amplitude over frequency. This has led the users to assume that a one-dimensional solution is applicable, the modification of amplitude over frequency, otherwise known as equalization. This is the worst of the RTA's attributes. The mentality of equalization as the primary or solitary solution leads to an endless repetition of the same mistakes. Since the real nature of the problem is not revealed, the solution is not sought, much less proved. The RTA cannot bring the data needed to frame the questions of how our sound systems are reacting to themselves and the room. The sources of variance are not found, and the principles of minimum variance are not applied. In many applications the examination of the system is conducted only at a single location (the mix position) and therefore level, spectral and ripple variance over the space are not even considered in the equalization of the system. Hence the same mistakes are made night after night, year after year.

Phase, as we know, is the holder of the puppet strings that control our sound system's performance. Without knowledge of phase we are destined to be surprised at every turn. With an RTA, evidence of polarity is wiped out by the rectifier, and evidence of the phase is removed by

Perspectives: Before high-resolution analyzers we owned an RTA and took it to the gig for the "wow factor." People loved talking into the mic and seeing a bunch of LEDs bouncing up and down. So did I.

d. (Mack) mcbryde

the integrator. Once lost these cannot be recovered. This erasure deprives us of our best evidence toward finding the causes of variance.

This is not to say that there is no use for the RTA. There may be applications in related fields such as urban or industrial noise analysis where random, uncorrelated sources are to be measured in low resolution. It is doubtful, however, that even these fields of study would not find advanced analyzers to be superior. In our field of sound system optimization the RTA was never the *right* tool even when it was the only readily available tool. The final word is this: the advanced analyzers can mathematically duplicate the RTA computation and display should the need arise. The RTA, however, cannot return the favor.

Perspectives: Long ago I was privileged to go and see one of the alignment wizards work their dark magic on a sound system. In the back of a venue, hunched over a multi-thousand-dollar HP FFT analyzer, the alignment engineer would stare hard at a squiggly line containing information covering only a few octaves. They would switch between several screens to see the whole frequency response and also switch the measurement points so often that I was most amazed at how they could juggle all that info in their head at 3 a.m. working on several hours of sleep. Today, just 10 years later, I'm writing this in an airport bar on a laptop that runs analysis software an order of magnitude more powerful and easier to use than the tools those guys had available in the bad old days.

John Huntington

Complex Audio Measurement Tools

The audio signal is complex. The word complex has specific meaning here. It means that the audio signal comprises three defining components: frequency, magnitude and phase. The relevant limitation of the simple audio measurement tools discussed previously is that they characterize some but not all of these three parameters. Complex analyzers give us access to audio signals in a way that the simple analyzers can never do. This complexity, however, requires us to rise to the task of interpretation, a job that will take some time and effort to master.

This subject is so deep and vast that large volumes are dedicated to its exploration. These will be filled with integral mathematics, calculus and differential equations since that is the language of complex signal analysis. No attempt to duplicate these efforts will be conducted here. This saves me the embarrassment of exposing my lack of mathematical prowess and you the annoyance of attempting to understand equations that are performed deep inside the analyzer which we will never see. Let's return to our analogy of medical diagnostics. MRI machines are designed, manufactured or serviced by experts in electromagnetism, not medical doctors. The MRI image is interpreted by doctors, as an aid to their diagnostic skills. The doctors need to know only enough about the MRI acquisition process to ensure that they read the data correctly. Their primary emphasis should be on seeing the data in terms of what is expected, what is abnormal, and what course of action will best serve the patient. So it will be for us and the complex audio analyzer. We will focus our efforts in this chapter to become competent at diagnostic interpretation of the data. The next two chapters will provide treatment plans.

The approach here is one of minimal math and the use of analogies is employed fairly liberally. Readers who are savvy in the ways of Fourier transform equations are politely asked to allow me some liberty in simplification. The broad outlines are presented, with minimal emphasis on any features that are specific to a particular manufacturer or product.

Notice of author bias: in the interests of fairness, let it be known that I have been involved with the design of the Meyer Sound Laboratories Source Independent Measurement (SIM™) analysis system since its inception in 1984. It is my goal here to present the analyzer in the most generic terms possible, free of product-specific endorsement or agenda.

There are several different ways to capture the complex data, many different ways to process the captured data and uncountable ways to display it. Various commercial products are tailored toward specific tasks and markets such as noise control, academic and industrial research and vibration. These markets are much larger than professional audio, and as a result, there have been very few tools developed by the mainstream makers of complex audio analyzers which meet our specific needs. In large measure the professional audio industry has resorted to creating our own measurement tools, very much in keeping with most other aspects of our industry.

The complex analyzers differ from the RTA and other simple analyzers in that their data is based on a strictly mathematical process of classifying the captured waveform. In short, complex analyzers capture the signal as a sampled waveform known as the time record. The complex

Perspectives: I've definitely learned more from measuring with a dual domain analyzer than I have from all of the books I've read.

Fred Gilpin

data (real and imaginary numbers that lead us to amplitude and phase) is derived by mathematical computation of the time record information. This allows us to return to the original signal and do additional operations as desired. This is not the case with the RTA. As with the analogy we used earlier of MRI scan versus the thermometer, the complex analyzer is much more powerful and so it will take some serious training to utilize its power effectively.

The Fourier Transform

In the eighteenth century French mathematician Jean Baptist Fourier developed an equation that explained the complex nature of waveforms. The Fourier theorem distills an audio waveform into its base components: frequency, magnitude and phase. In its most simplified form it states that any complex waveform can be characterized as a combination of individual sine waves, with defined amplitude and phase components.

The Fourier transform is a mathematical application of that theorem and takes a waveform captured over a period of time and performs the component separation. This is termed a "transform" because it converts the amplitude vs. time waveform (as seen on the oscilloscope) into amplitude vs. frequency (as seen in the RTA) and phase vs. frequency (like neither). This simplifies in terminology to a conversion of "time domain" data into "frequency domain" data. The process can be applied in reverse as well, i.e. the frequency domain data can be converted to the time domain. If you have the time data, you can get the frequency data and vice versa. See Fig. 1.2 to review the concept of time and frequency domain conversion.

Analyzer Basics

The fast Fourier transform (FFT) is the practical implementation of the Fourier transform formula. This is the engine for the complex audio analyzers used in sound system optimization. It is extremely fortunate for audio engineers that we are not required to calculate FFT equations in order to use them. We only need to understand them well enough to interpret the data and act on it.

Before getting into the details, we will take an overview. In the process, we will define the terms which will be used to describe the FFT functions.

- **FFT,** the acronym for fast Fourier transform, describes the process of converting the time record data into frequency response data.
- **Sampling rate** is the clock frequency of the analog-to-digital conversion.
- **Nyquist frequency** is the highest frequency that can be captured at a given sample rate. This frequency is half the sampling rate.
- **Time record** (also called the time window) refers to the period of time over which a waveform is sampled, expressed in ms.
- **FFT lines** (bins) are the number of samples (division operations) of the time record.
- **Time bandwidth product** is the relationship between the length of the time record and the bandwidth. The relationship is reciprocal, therefore the combined value is always one. A short time record creates a wide bandwidth, while a long time record creates a narrow bandwidth.
- **Resolution** is the width of each evenly spaced FFT frequency *"line"* or *"bin"* which is calculated by dividing the *sampling rate* by the number of samples in the *time window.* There are always a power of two number of samples in a FFT *time window* (128, 256, 512, 1024, etc.). For example, a 1024 point FFT at 44.1 kHz yields a resolution of approximately 43 Hz.
- **Bandwidth** describes the frequency span of a filter function (in Hz).
- **Constant bandwidth** is a linear rendering of bandwidth, with each filter (or frequency spacing) having the same bandwidth expressed in Hz. The FFT calculates filters with constant bandwidth.
- **Percentage bandwidth** describes the frequency span of a filter function (in octaves).
- **Constant percentage bandwidth** is a logarithmic rendering of bandwidth, with each filter (or frequency spacing) having the same percentage bandwidth expressed

in octaves, e.g. 1/3rd octave. The RTA filters are constant percentage bandwidth.

- Fixed points per octave **(PPO), also known as** "constant Q" is a quasi-log expression of frequency resolution. PPO refers to the number of frequency divisions in an octave, regardless of the individual bandwidth or percentage bandwidth of the individual bins, e.g. 24 points per octave. This is created by selecting an octave-wide portion from multiple FFTs taken from different time records. Each individual octave is linear but the assembled set of FFTs are logarithmically related, hence the term "quasi-log."

The Time Record

We begin with a waveform sampled over a period of time. This period of time is termed the time record, and is analogous to the shutter speed on a camera. The time record is the period when our audio shutter is open. The time record length is an open variable for us. It is the first critical parameter of the analysis since it determines the lowest measurable frequency. A waveform must complete a full cycle to be fully characterized. Therefore, frequencies which have a longer period (low frequencies) require longer time records than high frequencies.

Once the time record is captured the FFT begins the process of sorting the waveform into frequency components. This process is a series of interrogations that can go on for as long as desired (short of infinity). The process amounts to this:

Question 1: Describe the part of this waveform that completed one cycle during the time record.

Magnitude value:. . . . . . . Phase value:.......

Question 2: Describe the part of this waveform that completed two cycles during the time record.

Magnitude value:. . . . . . . Phase value:.......

And on and on for as high as we want to go (just short of half of our digital sampling rate).

Each time the question is asked the data is reported for a different frequency, successively moving upward.

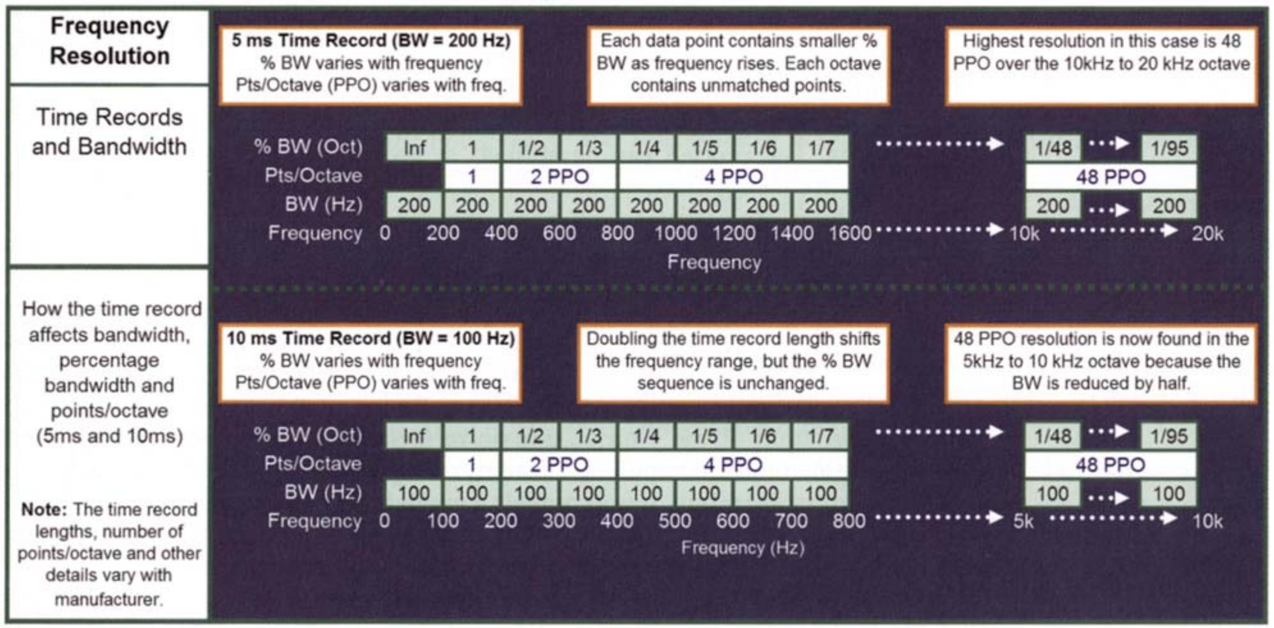


Figure 8.6 The relationship of the time record to frequency resolution in the FFT analyzer

The increment of frequency rise (in Hz) is the **bandwidth (frequency resolution)** of the measurement. The bandwidth is the reciprocal of the time record, another application of the familiar T = 1/F formula. It is important to note that the bandwidth is described in Hz, rather than the percentage of an octave that most audio engineers might expect. This is because the spacing is derived from successive divisions of time periods, and time is linear. Therefore the frequency resolution is linear (constant). This contrasts to the RTA which is based on analog (or digital copies of analog) filters and yields **percentage bandwidth** spacing. The term 1/3rd octave is a description of percentage bandwidth, which is a logarithmic rendering of bandwidth. The term bandwidth designates linear frequency spacing and is measured in Hz.

Linear and Log

The frequency-sorting system of the human hearing mechanism is mostly logarithmic, i.e. each doubling of frequency is perceived as an equal spacing. In a log frequency axis display the data is separated into evenly spaced increments of percentage bandwidth. This is in contrast to a linear expression of the frequency axis where constant bandwidth spacing is shown. In a linear display the high octaves are given visual preference, occupying screen real estate far out of proportion to their share of our hearing perception. The uppermost octave occupies the right half of the screen, and each octave below takes one half of the remaining screen. The low octaves are compressed to the left. Since the log display is obviously far closer to the way that we perceive sound, it might seem logical to disregard the linear scale and move on. However, there are some compelling reasons for us to understand the linear frequency axis. The foremost is that anything that affects time in our measurements affects frequency in a linear fashion. There is no log time, only linear. The frequency response effects of ripple variance are the result of time offset and therefore produce linear frequency axis combination effects (comb filtering). We will need to become bilingual in both log and linear in order to identify linear interaction in a log display (see Chapter 2).

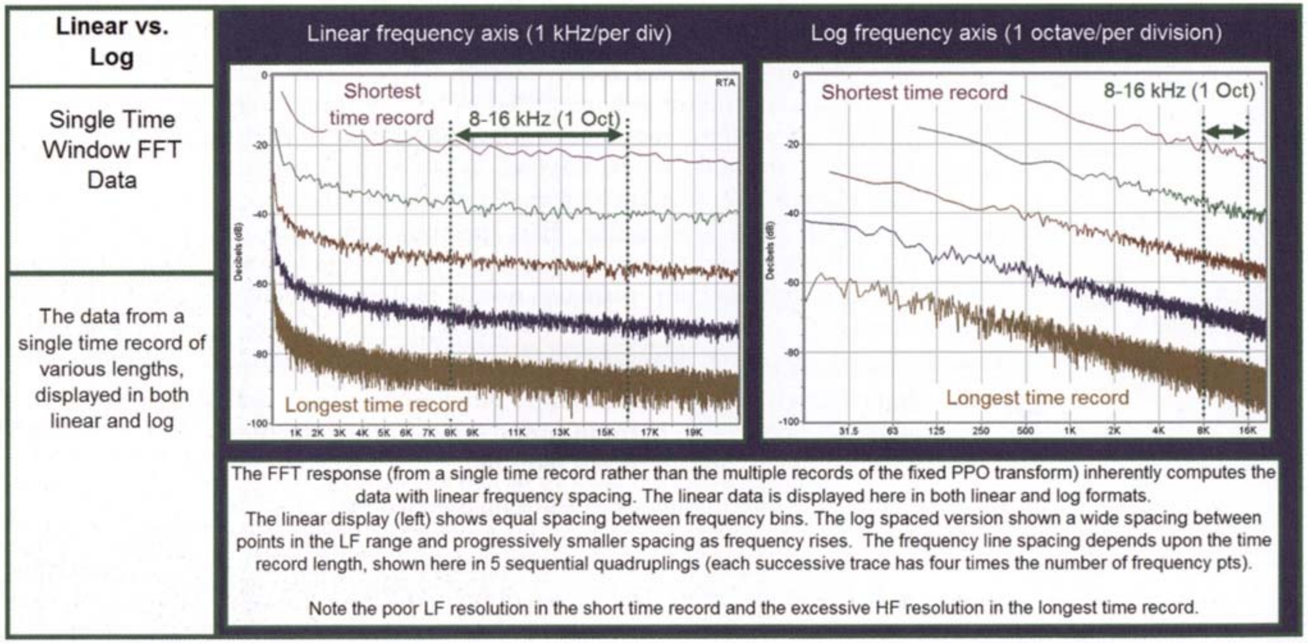


Figure 8.7 Frequency response measurement with various time record lengths, frequency resolution, (left) Linear frequency axis, (right) log frequency axis. As the time record increases the number of data points increases (upper trace of each set has the shortest time record) (courtesy of SIA-SMAART)

Perspectives: The use of dual-channel FFT analysis changed the way I thought about tuning a system. The optimization engineer leaves nothing to chance.

Tony Meola

Linear is not our first choice but it has its place, and is not in any danger of going out of style. Harmonic analysis, for example, a key aspect of speech and musical instrument research, uses linear as its first language. Harmonics are linearly related to the fundamental and are easily identified on a linear scale.

Frequency Resolution

There is no limit to the possible amount of frequency resolution. We can have 1 Hz resolution and have 20,000 data points to work with. Why not 0.1 Hz resolution? Is 1/3rd octave resolution enough?

The question of the appropriate amount of resolution depends upon the nature of what we need to detect with the measurements and how we are going to act upon the data. We don't have all day, so we will need the analyzer to have appropriate amounts of speed and responsiveness. Some actions require only broadband answers; others are focused on the details. If we have excess resolution we can always reduce it, so the defining factor will be the actions that require the highest. Recall the discussions of Chapter 3 regarding the resolution of the human hearing system. This will be the decisive factor. We have seen that 1/6th octave resolution is regarded as the "critical bandwidth" for perception of tonal variation. We also saw that frequency response ripple beyond 1/24th octave resolution (24 wavelengths late) pushed us over the threshold into echo perception rather than spaciousness or tonality. Therefore resolutions of around 1/24th octave will be sufficient for our primary needs in the frequency domain.

Some of our needs will be met with lower resolution, but none will require more than 1/24th for critical decision making. There is no harm in progressing further (such as 1/48th octave) but this is not required.

Here is the frequency resolution needed for specific tasks:

- **Noise floor analysis:** low resolution, 1/3rd octave is sufficient for general use. Higher resolution has the ability to detect low-level oscillations that might go unseen in low resolution.
- **Off-axis** response: medium resolution (1/6th) is sufficient for general use and suitable for characterization of spectral variance (pink shift). Higher resolution has the ability to detect ripple variance that might go unseen in low resolution.
- Ripple variance: high resolution (1/24th) is required.
- **Hum:** high resolution is required because the hum spikes are linearly spaced multiples of the line frequency. Low resolution will make the level of the higher hum spikes appear at reduced level due to smoothing.
- **On-axis** frequency response: high resolution (1/24th) is required. Resolution must be suitable for characterization of level, ripple and spectral variance. Higher resolution has the ability to detect ripple variance that might go unseen in low resolution.
- Spatial crossover frequency response: high resolution (1/24th) is required for the same reasons as the on-axis response.

Since very high resolution will be required to monitor summation ripple variance this will be the definitive category. Peaks or dips in the frequency response which are larger than the resolution, those whose bandwidths are larger than the analyzer bandwidth, will be seen clearly. Those that are smaller will be seen less clearly.

What exactly do we need to see? To fully characterize a peak or dip we must define a minimum of three features: center frequency, maximum deviation (boost or cut) and percentage bandwidth. To find these will require at the very least three frequency data points: the center frequency and the points above and below where the response has returned to unity. If we have additional resolution it can be put to use in filling in the shape of the curves as they connect between these points. With this in mind we can begin a consideration of how much is required to ensure the accuracy we need.

The analyzer divides the frequency spectrum into data points which for brevity we will term "bins." Each bin has a center frequency and meets the neighboring bin so that any data that falls between them is shared equally. The center frequency of the peak (or dip) under test may match the bin center or it could be somewhere between two bins. For our discussion we shall consider two possibilities: centered on one bin or centered between two. The bandwidth (BW) of the peak may be matched to the BW of the bin or it could be narrower or wider.

A sequence of scenarios with varying ratios of frequency resolution to the measured bandwidth is shown in Fig. 8.8. As the ratio increases the analyzer gains increasing immunity for the effects of peaks falling on bin centers and edges.

By the time the bin/peak (BW) ratio reaches 4 the peak shape and level can be clearly identified, regardless of centering. Since there will never be a guarantee that peaks and dips in the field will conform to our measurement bin centers, it is important that we maintain a reasonable immunity from bin center accuracy issues. This is done by "over sampling", in this case a 4:1 ratio. Therefore if we are continuing with the 1/6th octave critical bandwidth for equalization, we will need a 1/24th octave analyzer to ensure accuracy.

Recall the earlier discussion regarding the 1/3rd octave RTA and the matching 1/3rd octave graphic equalizer (Chapter 1). It should be clear now that such a seemingly appropriate matching will fare very poorly in the field.

Fixed Points per Octave (Constant Q Transform)

The ultimate tool for measurement would be one that has high resolution in complex form, but displays it in a fashion intuitive to our ears. That would be linearly based data, which includes the amplitude and phase values, with a constant high degree of frequency resolution per octave. This cannot be done with an RTA, even if it had filters as narrow as 1/24th octave, because the phase data is lost. Nor can this be done with a single FFT analyzer by making a log display, because the frequency resolution is still linear. It can, however, be done by taking multiple FFT

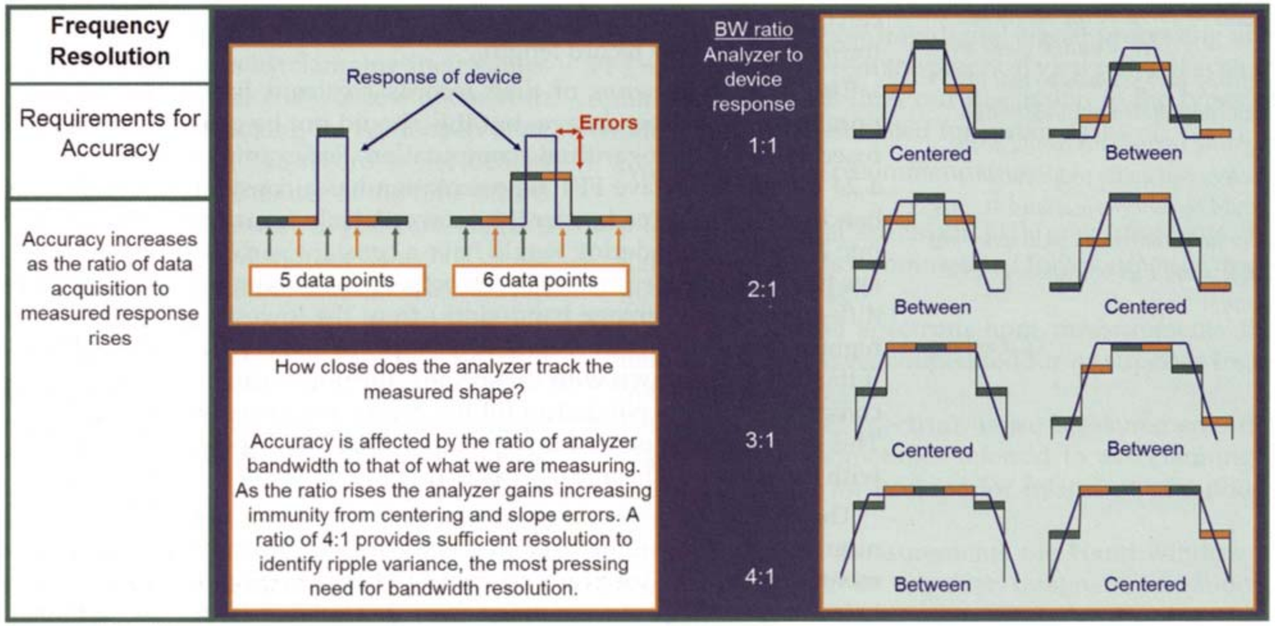


Figure 8.8 Bin/accuracy issues

Perspectives: The widespread use of the fix-point-per-octave analyzer has had a profound effect on the way that systems are optimized. There is a relationship between the nature of human hearing and a multitime-windowed transfer function that is much stronger than standard RTA or single-windowed transfer function techniques. As a result there is strong correlation between what we measure with a fixed point per octave FFT transfer function and what we hear, giving new meaning and clarity to all those "squiggly lines."

Sam Berkow

Perspectives: If you compare the diagnostic work of audio, video and lighting professionals to that of medical professionals, video and lighting technicians would be the doctors and audio technicians would be the veterinarians. It takes a lot more diagnostic skill when the subject can't let you know what's wrong.

TC Furlong

measurements in parallel and patching the responses together into a single "quasi-log" display. This is accomplished by taking a separate time record for each octave and using only a single octave of the data acquired.

The process begins at the high frequencies, where the shortest time record is used. An FFT is taken and a fixed number of data points are acquired for the upper octave. This becomes the base resolution, expressed in "points per octave" (PPO). We will not use the next octave down because it contains only half as many data points, due to the constant bandwidth of the linear data. This cuts the frequency resolution in half, and by the next octave it is cut in half again.

The challenge of the lost resolution is solved by keeping only our highest octave data and throwing away the rest. We then take another FFT from a time record of twice the length, which gives us matching frequency resolution an octave below the first. Each successive time record is twice as long as the previous, which optimizes the analyzer to achieve the same number of PPO for the octave below the previous. This process continues until the lowest frequencies desired have been measured. Note that the process can continue infinitely but reaches a practical limit due to the expanding time record length.

The sequenced series of time records creates a fixed number of points per octave, but this should not be confused with a true logarithmic computation. For example, a 24 points per octave FFT representation has a constant bandwidth that spreads over the octave. A truly logarithmic 24th octave rendering would have a constant percentage bandwidth for each slice of the octave. The maximum difference (in percentage bandwidth) from the lowest to highest frequency bin over an octave is approximately 2:1. If the display is drawn with log spacing, the points can be compressed and expanded to fill the octave very evenly. This combines the ease of interpretation of a log display, with the power of the linearly calculated FFT.

The fixed PPO computation includes a highly consistent number of wavelengths over frequency in the measurement window. The base resolution of 24 PPO corresponds to inclusion of the direct signal and those that follow within a 24-wavelength duration. Reflections or other summations which arrive beyond that duration will be seen as uncorrelated to the original signal due to their late arrival beyond the length of the acquisition time record. Therefore, we are seeing the same ratio of direct to early reflections over frequency. Contrast this to any single fixed time record, where we would see a different number of wavelengths for every frequency (more at the high end than the lows). For example, we will consider an analyzer with a single time record length of 24 ms. At 1 kHz the analyzer would see the direct signal and those copies (echoes and speaker interactions) that fall with a 24-wavelength duration. The time record length corresponds to the edge between tonal spaciousness and echo perception. The same time record would see 240 wavelengths at 10 kHz (perceived as an echo) and only two wavelengths at 80 Hz (perceived as a strong tonal variation). Our goal is to have the thresholds of tonal, spatial and echo perception appearing in a similar visual fashion over frequency. This requires the multiple time records of the fixed PPO transform.

Window Functions

The process of counting out frequencies from the time record goes along just fine as long as the frequency content is made up of exactly even multiples of the bandwidth. But what about frequencies that are fractions of the frequency resolution. If left untreated these frequencies will masquerade as bursts of noise and create substantive errors in the frequency response. The mathematical device used to correct this is called the **window function.** The errors caused by the fractional frequencies are reduced by the application of this function.

The fractional problem occurs because the FFT analyzer works under the assumption that the sampled waveform is repeated ad infinitum. While this is true of much of today's popular music, we can not count on it. A by-product of the removal of infinity in our calculations is the forcing of waveforms to appear to our calculations as if they have gone on for infinity. Therefore, the time records must have matching characteristics at the beginning and

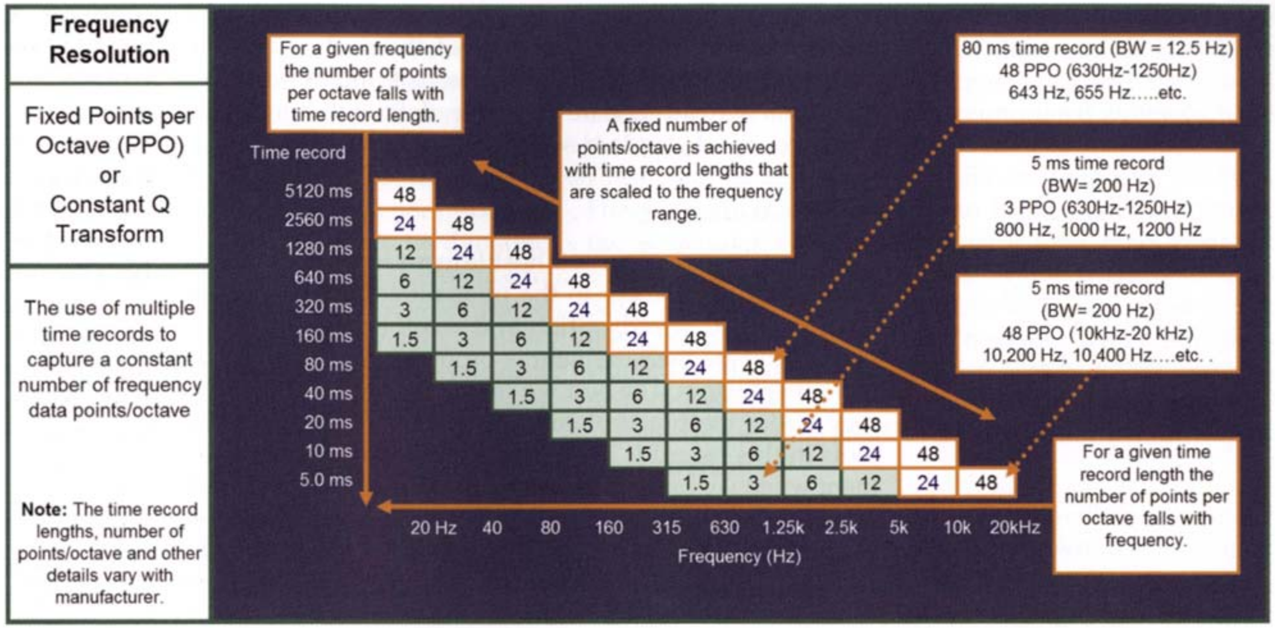


Figure 8.9 The fixed PPO (Constant Q) transfc

end points. The window function forces the time records to be matched at the end points by clamping their values to zero. The "window" is closed. The window opens both at the front and back end, reaching its fully open state in the middle. In essence, the time record is not treated equally. Events that occur in the center of the time record are favored over those in the ends. As you can imagine, this creates some errors in the response. Our choices are to live with these errors, instruct the musicians as to what frequencies they can play, or measure for infinity. We will take the errors. Fortunately they are a long way down if we manage them well.

There are many ways to open and close the window: fast, slow, rounded, triangular. The rate at which the window opens and reaches its peak creates a unique shape. These functions have been named for their shape or their inventor and are optimized for particular types of measurements. The subject of windows and their prospective benefits and errors is a topic that is covered in great detail in advanced texts regarding digital signal processing and FFT measurement. From the practical viewpoint of system optimization we will limit our discussion to the types of windows that are best suited for certain tasks.

Window function recommendations:

- For single-channel spectrum THD measurements the flat-top window is recommended for its superior treatment of sine waves.
- For single-channel spectrum hum measurements the flat-top window is recommended for its superior treatment of sine waves.
- For single-channel spectrum noise measurements the Harm window (sometimes referred to as "Hanning") is recommended for its superior treatment of random signals.
- For dual-channel measurements the Harm window is recommended for its superior treatment of random signals.

369

In most cases default window functions are appropriately pre selected by the manufacturer of the analyzer rather than left to the user. Users with special window function requirements can deviate from the default settings. For our needs most standard settings will be sufficient.

Signal Averaging

The acoustic environment is hostile to measurement. There are many disturbances to the transmitted signal such as ambient noise, reverberation and weather changes. Any single acoustic measurement is prone to error from these and other factors. The best we can do is to maximize accuracy under the practical conditions. One of the foremost tools in this regard is averaging. Why trust one measurement when we can look at a whole series of them and find out what the trends are? If we take a second measurement we have information that can't be found in the individuals viewed singularly. Not only do we get a mathematical average of the two, we get the extent of the differences between them: the deviation. Not only have we doubled our statistical reliability but we have a measure of stability that will serve to help us ascertain the credibility of the data. If we take even more averages we will increase our statistical reliability and also further clarify the range of deviation.

Why do these factors matter? The statistical reliability is how we immunize the signal from the errors caused by the noise. Noise will be added to the input signal by the time it reaches the output. Guaranteed. The relationship of the signal with the noise fails the stable summation criteria developed in Chapter 2. The two signals are not correlated. This causes an unstable output signal that, on a moment-to-moment basis, is non-linearly related to the original input signal. Since noise is random, its summation will eventually average itself out to zero due to the random phase relationship of the noise to the signal. Therefore if we average our signals sufficiently, the noise in the response sums to zero and its summation effect upon the signal at any given moment will be nullified.

Of secondary interest is the size of the deviation. If the noise is strong compared to the signal, the deviation will be large because the noise values will dominate the response. A larger number of averages are required to provide comparable reliability to one with a lower noise component. As the signal-to-noise ratio falls, the number of averages required to acquire a stable response rises. The summation effects of the noise are statistically reduced, but the presence of the noise contamination to our signal is not; an important distinction. Our aural perception of a noisy signal is not equivalent to one that is free of noise. Both the signal-to-noise ratio and the averaged signal character are relevant to our experience. Therefore we will use averaging to determine the signal by statistically eliminating the noise and use the signal-to-noise computation to track the amount of deviation. The measure of deviation will be coherence, which will be covered shortly.

Averaging Types

There are numerous mathematical possibilities for signal averaging. We cannot cover all of them and instead will focus the application of those schemes that are of primary interest to optimization. Further details on the math behind these and on additional schemes can be found in the publications of the manufacturer or in scientific journals. The schemes can be divided into a few main categories: the math behind the averaging, the weighting of the different samples, and how the current numbers are maintained.

There are two primary math options: RMS and vector averaging. RMS (root-mean-square) style averaging is used for single-channel spectrum RTA-type simulations and for impulse response averaging. RMS averaging is suitable for averaging of random signals. It is relatively poor at discriminating random from correlated signals, which means it is not our first choice for transfer function measurements. Vector averaging, by contrast, is highly sensitive to variations in either amplitude or phase. The term vector refers to the fact that this averaging scheme relies on the vector value between the amplitude and phase (real and imaginary) aspects of the signal.

Weighting schemes fall roughly into two types: weighted or unweighted. The term weighting refers to how much value each sample contributes to the final

Perspectives: Just about anyone can afford a quality measurement system today. When SIA-SMAART first started appearing in the field for live applications, misuse was rampant. I suppose it still is. So often I still see bad decisions based on bad data, or bad decisions based on a fundamental lack of understanding of the issues at hand. Still, the greatest challenge I see in the field is getting a useable measurement and then knowing how to interpret it. We still have plenty of people happily producing flat lines on a display without even considering whether or not the measurement is valid. Validation and verification — "are we measuring the right thing, and are we measuring it properly?" — ought to be the focus.

Doug Fowler

average. If all samples are treated equally the averaging is termed unweighted. This is the usual form of averaging and is assumed unless noted. If some samples are given extra levels of consideration the average is considered "weighted." The most common form of weighted averaging is exponential averaging, where the most recent samples are given higher weighting than older ones. Exponential averaging is much quicker to react to dynamic changes in a system, since it gives statistical preference to more recent samples. Unweighted averaging is slower to react since it factors old data with equal weight as new. The exponential has less stability but higher speed for a given number of samples than a comparable unweighted average.

The manner in which old samples are accommodated is the final parameter. There are three popular schemes. The first is the accumulator. In this type each new sample is added to the combined average on an ongoing basis. A first sample enters and its value becomes the original base value. When the next arrives the new values are added to the base and divided by two. This yields an average between the two values. When a third sample is taken it is added to the previous combined data and the sum is divided by three and so on. Accumulated averages have the advantage of being the most stable over the long run, since there are an unlimited number of samples possible. However, should a change occur in the system response, the accumulator could take a very long time to detect such change since it may have in it hundreds of samples of the outdated response. In those cases you must restart the accumulator to freshen the response with a new base value. The visual appearance of an accumulator resembles watching plaster set. At first the plaster is highly flexible and steadily turning solid, becoming steadily more impervious to change.

A second scheme is a fixed number with an automatic restart. Such a scheme is like a slow windshield wiper. The response builds up for the fixed number and then, swish, restart. This has its place when we are measuring systems where changes are expected, such as an equalizer

Signal Averaging

Speed and Stability

How the averaging scheme affects the data with variable response over time

FIFO (First in, first out)

New data arrives (sample # 1) that is higher in value than all previous. This will cause the average to rise. FIFO averaging continually updates. The average rises fastest when the number of averages is low. Stability rises as more averages are used.

Sample #	1	2	3	4	5	6	7	8
Value	120	115	110	105	100	95	90	85
Weighting	1	1	1	1	1	1	1	1
Weighted Value	120	115	110	105	100	95	90	85
Avg Value (8)	102.5 (820 / 8)							
Avg Value (4)	112.5 (450 / 4)							
Avg Value (2)	117.5	(235 / 2)						

Accumulate

The average is kept as a single lump value of unlimited accumulations. This scheme has variable speed and stability. At first it is fast and unstable, then becomes stable and slow.

Sample #	1	2-8
Value	120	700
Weighting	1	7
Weighted Value	120	100
Avg Value (All)	102.5 (820 / 8)	

Weighted (Most recent has highest value)

The new sample has the highest weighting. The average value moves rapidly toward the new value. Change is most rapid with fewer averages. Weighted averages are the fastest acting, but least stable.

Sample #	1	2	3	4	5	6	7	8
Value	120	115	110	105	100	95	90	85
Weighting	1	1/2	1/3	1/4	1/5	1/6	1/7	1/8
Weighted Value	120	58	36	26	20	16	13	10
Avg Value (8)	110 (300 / 2.7)							
Avg Value (4)	115.4 (240 / 2/1)							
Avg Value (2)	118.3	(178 / 1.5)						

Figure 8.10 Averaging types

or delay line. For stable system measurements the restart is an annoyance.

A third scheme is the "first in, first out" (FIFO) style. In this style the data flows through a pipeline of fixed length and new data is flushed through the averager in sequence. New data pushes out old data so that there is never a need for restarting. The data is always current.

Speed vs. Stability

Maximum stability of the measurement comes from the maximum number of averages. Maximum speed comes from the fewest. Electronic signals usually have the luxury of a high signal-to-noise ratio, compared to the noisy acoustic environment. Electronic measurements are able to utilize lower numbers of averages, thereby yielding higher speed without significant loss of stability. Acoustic measurements must sacrifice speed for maximum stability by using large numbers of averages.

Thresholding

An additional point of interest to us is the prospect of signals which contain varying signal levels over frequency and time, otherwise known as music. Well, *some* music, at least. If we are to maintain statistical equality we will need to measure loud signals with a smaller number of averages compared to quiet ones. When the music stops, what then? The answer to this comes in the implementation of a threshold, a device that monitors the input signal and checks for signal strength. Only frequencies with sufficient data sent to their input will be able to be accurately characterized at the output. If we know we are not going to get accurate data why bother to measure it? If the current signal is a flute solo, what do we think we are going to learn about the subwoofer response?

The device that accomplishes this is called an **amplitude threshold.** The threshold acts like a noise gate. For each frequency bin the input level is checked against the threshold. If it is above the bar, the signal is sent onward for transfer function analysis and into the averager. If not, the analyzer simply passes on the data and waits for the next sample. When the music stops the analyzer idles. Since there is nothing being sent to the system there is no reason to measure it and every reason not to.

Not all dual-channel FFT analyzers utilize this capability. Those that do use it enjoy a higher immunity from noise in the transfer function and therefore higher stability.

Single-Channel Spectrum Applications

The frequency response can now be viewed in high resolution. The response shown represents combined amplitude vs. frequency conversion of the waveforms obtained from the various time records. It contains a record of the energy over frequency and can be averaged from a series of successive time records to create a more statistically valid response. This response comes from the data streamed from a single signal channel. That channel can be measured for various operations such as spectral content, total harmonic distortion (THD), noise floor and others. We will refer to a single-channel response here as the **spectrum.**

The single-channel analyzer sees the incoming data and relates it to an internal standard of reference. The displayed values are the difference between the internal standard and the measured response. The internal standard may be user-adjustable, such as a particular voltage or dB SPL reference value, but it must be known. In other words, if we are measuring voltage at the output of a system we have an internal reference in our analyzer that is calibrated in volts. If we want to discern the voltage gain through the device we have two choices. We must stick to a standard, known input signal, or we will have to reset our calibration point to the actual level sent into the device and scale our display in relative terms. The single-channel analyzer is therefore source-dependent, i.e. it must operate with a source with known amplitude and phase characteristics over frequency, if it is to provide us with that knowledge of the device under test. This seemingly small point will loom very large as a limitation of the single-channel analyzer.

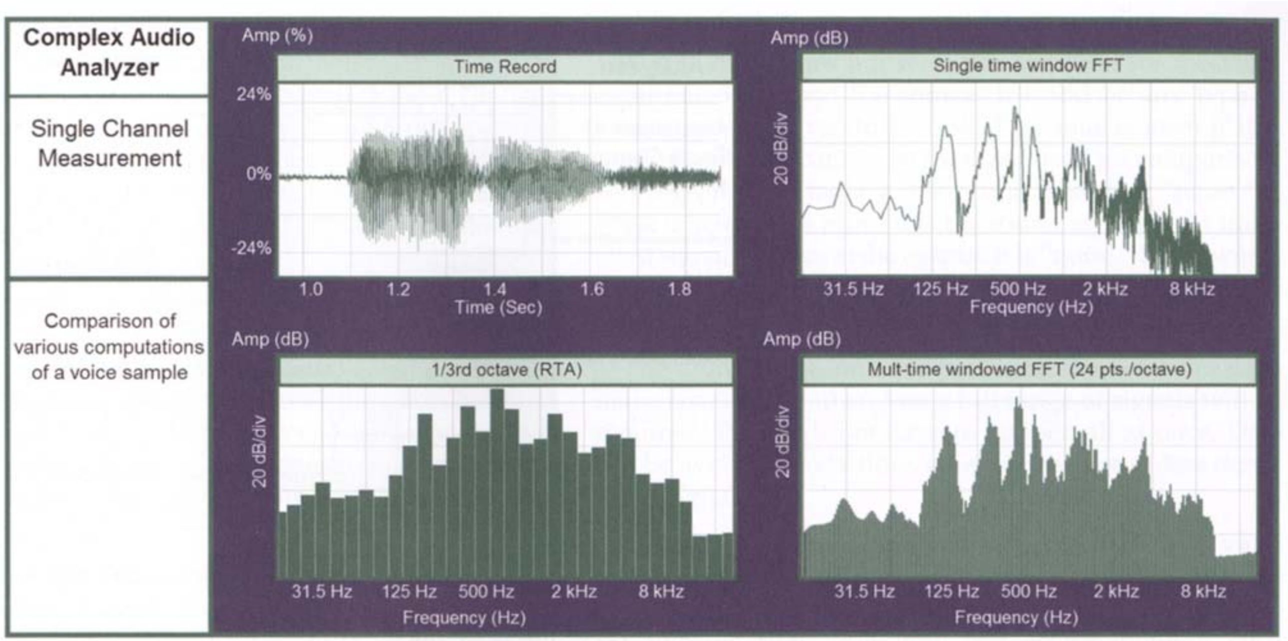


Figure 8.11 Single-channel spectrum of a voice measured with a multiple time-windowed, fixed points per octave FFT analyzer (24 PPO) (courtesy of SIA-SMAART)

The primary applications of single-channel measurements in the optimization process are:

1. the monitoring of source frequency content, level and range
2. total harmonic distortion (THD)
3. maximum input and output capability
4. hum and noise floor.

These functions fall almost exclusively into the verification category of the optimization process, and are covered in detail in Chapter 9.

Single-Channel FFT Limitations

Single-channel measurements have inherent limitations that affect their utility. Single-channel spectrum measurements must have known source signals in order to make conclusions regarding the response of the device under test. A music signal played through an equalizer would show the combined response of the music and the equalizer. What part is the music, and what part is the equalizer? We have two unknowns in our equation: unknown input (the music) vs. unknown output (the music and the equalizer).

Single-channel frequency response measurements must have pink noise or some other signal with a known frequency response. This source signal is then assumed to be the "known" in the calculation of the frequency response of the device under test and deviations from that "known" response at the output are observed. THD measurements require a low-distortion sine wave, while noise floor measurements require no source at all.

A second limitation involves the testing of systems with multiple components linked in series. This includes even the most simplistic of sound systems. If multiple devices are placed in series, the combined response can be found by driving the first device with the known source and measuring the output of the last device. However, with

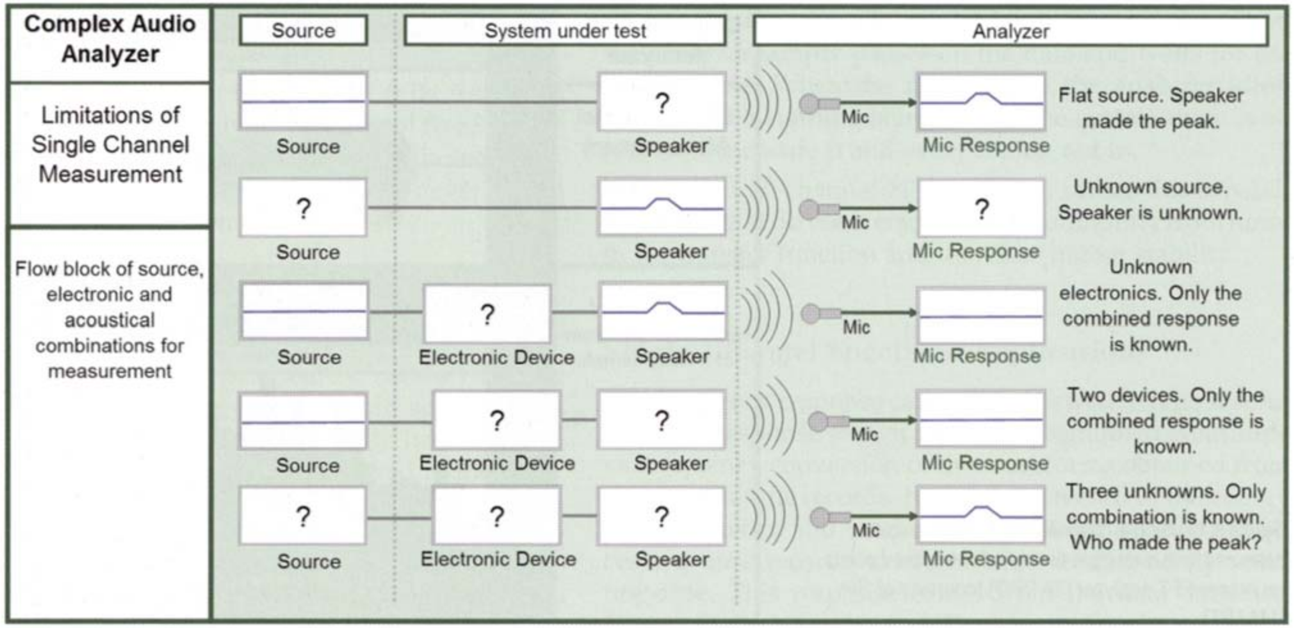


Figure 8.12 Limitations of single-point audio measurements

a single-point measurement system it is not possible to discern the individual effects of the components without breaking the signal chain or making assumptions. Only the first device is driven by a "known" source. All others are driven by a preceding device, each of which has its own effect on the signal. By the time we reach the output of the second device we again have two unknowns in our equation: unknown input vs. unknown output.

The spectrum response alone is will not be sufficient for our optimization needs since it lacks the relative phase data that we know will be critical to our decision-making. Relative phase is derived from the comparison of two phase responses: the input and the output.

To illustrate the difference consider the case of the Beatle's 1968 recording of *Abbey Road.* If we were to listen to this music currently it would be over one trillion degrees out of phase at 1 kHz from the original source material. The phase relationship to the original session, however, is not relevant to our perception since we will not hear a summation of the original music along with our CD recording. The original sound has long since died away. If we play this music through speakers, the phase response of the room reflections will have a very large effect on our perception, since we will be hearing both the original and reflected signals summed together. In this case the relative phase of the summed signals will be very relevant to the sound quality.

Transfer Function Measurement

The response of a device from input to output is its **transfer function.** The process of comparison measurement between input and output is termed transfer function measurement. This could be a passive device, such as a cable, attenuator or filter, or an active analog or digital circuit. The transfer function of a device is obtained by comparing its input signal to its output signal.

The transfer function of a hypothetical perfect transmission system would be zero. Zero change in level, zero delay, and zero noise at all frequencies. Any device that passes signal will have some deviation in transfer level and transfer time and will add some noise in the process. Transfer function measurement can detect these changes and has a variety of ways to display them for our analysis. Whereas the spectrum response measured the output in absolute terms, the transfer function will always use relative terms: relative amplitude, relative phase, relative time and signal-to-noise ratio.

The basic principle of the transfer function analysis is a dual-channel measurement where one channel is designated as "known" and the other as "unknown." The known channel becomes the standard and the differences between the two are attributed to the device(s) between the two points. This is usually the input and output of a device but could also be the output of two different devices such as microphones.

Perspectives Transfer function. I love my transfer function. Mr Fourier, you are a genius. Thousands of screaming little girls thank you. Again and again and again ...

Francois Bergeron

The known channel can also be termed the reference channel or input channel. The unknown can be termed the measurement channel or output channel. In any case the differences between these are the point of interest.

The signal driving the input can be any type of signal. This signal becomes our reference standard for measurement and is termed the **source.** It could be any type of music, speech or random noise. This source, even if it is random noise, becomes the standard used for comparison. It is important to make a distinction regarding "noise." If noise is sent to the input it is the source signal. If an unrelated signal appears at the output, it is "noise." In the world of transfer function measurement, we would be more accurate to call our favorite test signal "pink source."

The source must contain a particular frequency for us to make conclusions, therefore a full range of signals will be required. This does not have to happen all at once. Data can be averaged over time, allowing the use of less dense input signals.

Certain conditions must be met to obtain a valid transfer function measurement. These conditions can be approached, but never perfectly met; therefore practical considerations will limit our achievement of such conditions.

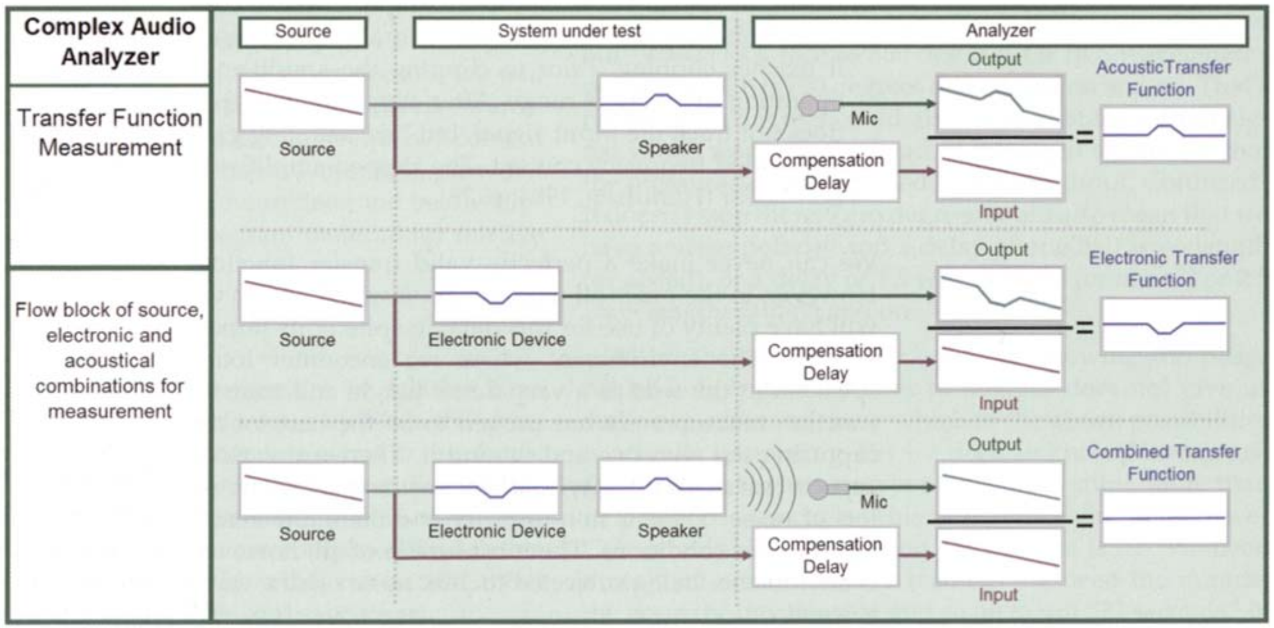


Figure 8.13 Two-point transfer function measure
points

1. **Stability:** the transfer function response of a device must be stable over the period of the time record. The device under test can not be changing its transfer level or its transfer time during the sample period. If this occurs, the comparison between the source and the output is not stable. An illustrative example: a valid transfer function measurement of a rapidly moving speaker is not possible (Herlufsen, 1984, p. 25).
2. **Time invariance:** the tested device must not change its response over time. In practical terms, this refers to a time period longer than our measurement acquisition time, and includes the period over which averaging is taking place. A speaker's response could be averaged over a 16 second period. This response will change with temperature. If the temperature is stable over those 16 seconds, the measurement is valid for that period of time (Herlufsen, 1984, p. 25).
3. **Linearity:** the device must be linear. In practical terms the output must be proportional to the input and be relatively free of distortion and noise. It may have gain, loss or delay at any given frequency, but those properties remain consistent regardless of the nature of the input signal. For example, an amplifier has a constant voltage gain that is applied to all signals until it reaches clipping. Prior to clipping the amplifier is operating in its linear range. After clipping the output does not track the input signal, but has changing gain and added frequency content. The clipped amplifier is non-linear (Herlufsen, 1984, p. 25).

We can never make a perfectly valid transfer function. However, in the practical world of professional audio we will have plenty of use for this data, in spite of its imperfections. The environment where we encounter loudspeakers in the wild is a very dense jungle and transfer function measurement has proven to be the best tool at capturing the response and taming it. There are various aspects that make this environment hostile. There is noise, lots of noise. There is non-linearity and there are changing weather conditions. The most hostile of all, however, is an audience being subjected to pink noise or sine wave sweeps.

The dual-channel FFT math has some important qualities that aid us in these environs.

1. **Source independence:** the ability to measure with the program material as the source means the analyzer can continue to provide a continual stream of accurate and meaningful data about the system even with an audience present. The value of this is inherently obvious.
2. **Non-intrusive access:** the two measurement points (input and output) can be fed from any two points in the system without interrupting the signal path.
3. **Complex frequency response:** the dual-channel method provides relative level, relative phase, relative time and signal relative to noise data. These measurements provide the key data for decision-making.
4. **Best fit for non-linear data:** non-linear elements are detectable because the analyzer has a copy of the original response to compare to the output. The analyzer is able to indicate frequencies where non-linear data is present in the measurement.
5. **Noise immunity:** all measurements contain noise. The dual-channel method is able to identify noise as it did the non-linear data above. It also can minimize the errors in the output response by averaging the response. Averaging the series of responses causes the variations due to noise to be minimized.

The above features are the keys to how the dual-channel FFT analyzer is successfully applied to system optimization. Now we will begin to view the analyzer in the context of our application.

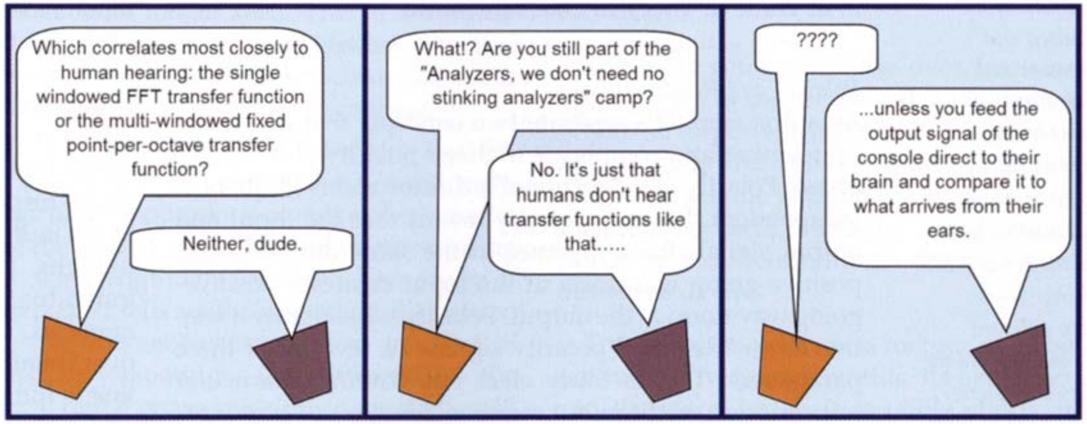
Frequency Responses

There is a large family of responses over frequency now available to us. Our efforts here will limit the scope to three main players: relative amplitude, relative phase and coherence, over frequency.

Relative Amplitude

Transfer function amplitude is a measure of relative level between the two channels over frequency. Since this is a

Trap 'n Zoid by 6o6



relative measurement it is most commonly expressed in dB. Unity gain is OdB with gains being positive numbers and losses expressed as negative. Transfer function amplitude can be modeled simplistically as division:

Output/Input = Transfer function amplitude

If the input rises and the output rises with it, we will see no change. If one changes and the other doesn't, we will see the change. Transfer function amplitude is source-independent. The source drive level or frequency content will not affect the outcome of the measurement, provided it is sufficient to rise above the noise floor and below clipping. This allows us to use program material for relative amplitude measurements.

Relative Phase

Transfer function phase is a measure of relative time between the two channels over frequency. Since the absolute phase of the individual channels is not in our focus, all of our references will be to relative phase. For brevity the term "phase" will be hereafter assumed to mean relative phase. Relative phase is more complex than relative amplitude. There is almost nothing that can be said about relative phase without qualifiers and exceptions.

Let's begin with the simplest one: zero degrees. Our first assumption would be that a reading of zero degrees would correlate to unity time (no time difference between input and output). Not necessarily. Unity time could be a reading of 180 degrees if there is a polarity reversal between the measured channels. Zero degrees could also mean one full wavelength delayed, which is the same point on our phase clock face as 360 degrees. Or two wavelengths delayed which is 720 degrees and on and on and on. That's not all. Zero degrees could also mean that the output signal is one wavelength ahead of the input signal; i.e. that the measured output arrived before the input. Confused? It doesn't stop there. Zero degrees could also mean that we have reverse polarity and a delay of one half wavelength (or ahead), or a delay of 1.5 wavelengths (or ahead) or 2.5 wavelengths and on and on.

All of these possibilities are real. Knowing the phase response only as a number of degrees does not give us sufficient *context* to know which of the above possibilities is actually happening. Yet we need to know. Having just the phase value as a number gives us nothing more than its position on a circle. This is like receiving the answer "22 seconds" to the question "What time is it?" We need more than a second hand reading. We need the minute, the hour, the day the year and so on to put "22 seconds" in

Perspectives *The term "phase" is one of the most misused and misunderstood terms in audio. Phase is simply "frequency-dependent delay"! Of course you "hear" phase it represents the time relationship between frequencies.... Measuring phase is a great tool that plays a critical role in system tuning. While manufacturers have different opinion regarding the importance of "relative" vs. absolute phase, systems that are tuned to be linear in phase typically sound their best!*

Sam Berkow

context. We will need to learn to read the phase response in its context. This will take some time.

Polarity and Relative Phase

As a first step let's separate two concepts that get mixed up together and complicate matters: polarity and relative phase. Polarity is a directional indicator and is frequency-independent. Positive polarity means that the input and output signals track together in the same direction, e.g. positive-going waveform at the input creates a positive-going waveform at the output. Polarity is inherently a relative term. "Reverse" polarity signifies a reversal of these parameters. This is fairly clear cut within a particular transmission medium but requires attention to standards when signals are transformed between mediums, e.g. they change state from electrical energy to magnetic or mechanical. In such cases the relative polarity is considered "normal" if the change follows current standards.

There is a phase component but no delay component to polarity. There is 180 degrees of phase shift at all frequencies, or none. We can run an auto race clockwise or counterclockwise without changing the outcome. This is not to say polarity is irrelevant. Speakers and racecars will all need to be running in the same direction or there are going to be head-on collisions.

Relative phase is frequency-dependent. Ninety degrees of relative phase has no meaning without frequency. For our purposes the context for phase comes from three factors:

- frequency: this tells us the time period for the particular range
- phase slope: the rate of phase angle change over a given frequency span
- phase slope direction: this will tell us who came first — the input or the output.

With these three parameters we can begin the conversion to phase delay, the measure of signal delay (in ms) over the frequency range of interest. Let's begin the search. The phase angle is impossible to miss since it must be somewhere on the 360 degree circle. Select the frequency of interest and read out the number of degrees. Voila! The inherent limitation of this data can be visualized with the analogy of a clock. If we are only shown the second hand we can only tell the time relative to the current minute, which is unknown. Until we know the minute, hour, date and year we will not be able to conclusively put the movement of the second hand into a larger temporal context. The phase response over frequency display connects the dots of the individual phase angles into a single line from lows to highs. The continuity of the line gives a context to the individual points, by linking them together with other frequencies. The line that connects any two phase data points over frequency is termed the **phase slope.** The slope of the line is indicative as to whether there is delay (and hence an offset in time) between the input and output channel in that frequency range. If the line is horizontal, the slope is flat, an indication of no time difference between the two points. If the line is tilted, a delay between the input and output channels is indicated. As the slope steepens, more phase delay is indicated. A vertical line would indicate infinite delay.

The slope direction (up or down) indicates which of the two channels is leading in time. Reading the frequency axis from left to right, a downward slope indicates that the output arrives after the input data. An upward slope indicates that the output arrives before the input data. You might be thinking this is a misprint, or that I am having a mental polarity reversal but I repeat, the output before the input. To get our head around this we must put on our transfer function thinking cap and remember that our perspective is *relative.* Any two points in a system can be compared. The temporal relationship between them is an open variable.

Wraparound

The phase response over frequency display contains an unexpected and potentially confusing visual feature that requires explanation. The vertical axis is a straight-line representation of a circular (0-360 degrees) function. How then does it display phase shifts greater than 360 degrees? When a trace reaches the edge of the display, where does

Perspectives The creative forces wanted to put movie screens in front of our speaker clusters. Naturally we were concerned about the audio quality effects so we asked for specs on different screens. The impressive manufacturer data included a phase response over frequency. Because of our experience with complex acoustical analyzers we could see that this was "Photoshop phase," i.e. a cut-and-paste copy of the amplitude trace. The data was disregarded entirely and we conducted our own measurements from the ground up.

Mike Shannon

it go? One option would be to simply crop the image and the phase response is no longer seen. This is not necessary or helpful. Instead, when the trace reaches the edge (360 degrees) we will draw a line (the wraparound) to the opposite edge (0 degrees) and continue from there.

The source of the confusion is that 0 degrees and 360 degrees appear to be the same points when viewed without context. Returning to our clock analogy, they are the same position of the second hand. The 360 degree reading is, however, one wavelength late (or early). Our study of summation (Chapter 2) tells us the important distinction between being synchronous as compared to an offset of one wavelength. The wraparound line differentiates the otherwise identical radial phase angle positions such as 10 degrees and 370 degrees. The wraparound provides *context* to the radial angle. A phase response over frequency that contains no wraparounds has confined all of its time offsets to within a single wavelength at all frequencies. A response with multiple wraparounds has some ranges that are more than a wavelength behind others. The conversion from radial phase to phase over frequency is illustrated in Fig. 8.14.

The most common phase over frequency display has zero degrees at the center and ± 180 degrees at the edges. When the response reaches -180 degrees the wraparound line jumps up and connects to the next point, $+179$ degrees, which is the same point in the circle as -181 degrees. The wraparound vertical line is a visual artifact and not indicative of a discontinuity in the phase response of the measured device.

The phase response is a merger zone between the linear and log frequency worlds. Its operation is time-based, and therefore, linear. Its audible effects are logarithmic. Therefore this is a good time for us to bring the two worlds together. The linear representation of a fixed delay over frequency shows perfectly even spacing between the wraparounds as shown in Fig. 8.14. This is due to the fact that a complete wraparound occurs with each multiple of the base frequency ($F = 1/T$), after another 360 degrees of

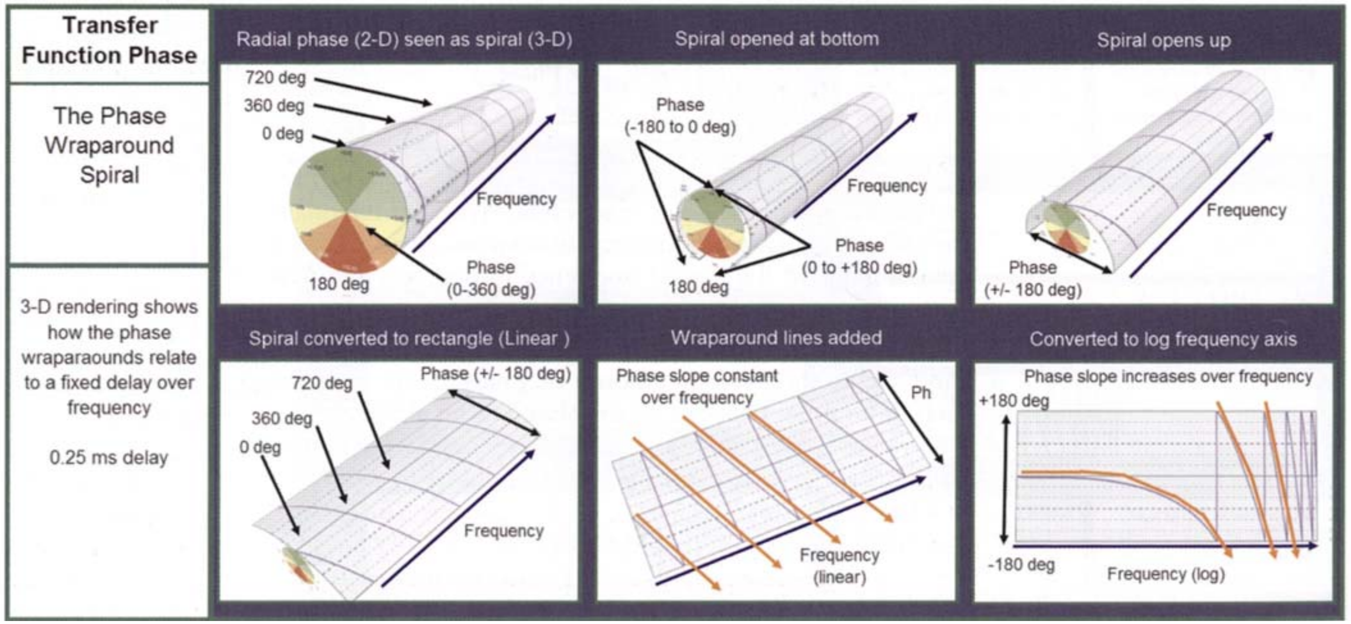


Figure 8.14 Phase response "unwound." (3-D artwork provided by Greg Linhares, courtesy of Meyer Sound)

phase shift has accrued. The beauty of the linear display is its clarity of time-related interactions over frequency. The phase slope stays absolutely constant, because the time delay is also constant. The downside is that it does not resemble how we hear it.

By contrast the log representation shows the phase response as having a different spacing between each wraparound. The phase slope becomes steeper as frequency rises. Also, the wraparound becomes increasingly compressed as frequency rises. Since our hearing mechanism is log we will have to learn how to read phase slope in log, although it is the more difficult of the two visually.

Phase Delay

Flat phase over frequency comes in two principal forms: 0 degrees and 180 degrees. Either way we have zero phase delay over frequency. The 180 degrees version is reverse polarity. Once we introduce a slope into the situation we need a formula that will decode the slope into time. That formula is a variation of the $T = 1/F$ which we have encountered before.

The phase delay formula:

$$T = \frac{\dfrac{\text{Phase}_{\text{HF}} - \text{Phase}_{\text{LF}}}{360}}{\text{Freq}_{\text{HF}} - \text{Freq}_{\text{LF}}}$$

where T is the phase delay in seconds, Phase_{HF} is the phase angle at the highest frequency in degrees, Phase_{LF} is the phase angle at the lowest frequency in degrees, Freq_{HF} is the highest frequency, and Freq_{LF} is the lowest frequency.

This formula can be applied to any range of frequencies, and yields the average amount of phase delay over the selected range. This can be simplified as an expression of the rate of phase change over a given frequency span as follows:

$$T = \frac{\dfrac{\text{Phase change}}{360}}{\text{Frequency change}}$$

We will now apply the formula to a transfer function phase trace as shown in Fig. 8.15. Application of the above formula will compute to 1 ms of phase delay at all frequencies.

Phase Slope

For a fixed amount of phase delay over frequency the phase slope will have a constantly increasing angle over frequency (a log display is assumed from now on). Each succeeding octave will have double the number of wraparounds, since it is double the frequency span. Each succeeding wraparound will be incrementally steeper than the previous, e.g. if the slope at the first wraparound is *x* degrees, then the next wrap will be *2x* degrees, followed by 3x, 4x and so on.

Unfortunately it is not possible to relate this to a particular angle, since the vertical to horizontal ratios of a graph are variable. The graph could be short and wide, tall and narrow, or square. Each of these would yield a different slope angle for the same delay, even though in all cases there is the same rate of phase change over frequency. There is no magic number, such as 1 ms of delay, which creates a 45 degree slope at a given frequency. This does not eliminate its usefulness. With a log frequency display the slope angle indicates the number of wavelengths of delay, rather than the time. The same slope angle at different frequencies indicates the same number of wavelengths.

Let's summarize the phase slope properties.

- **For a given frequency:** the slope angle is proportional to the number of wavelengths of delay.
- **For a given slope angle:** the amount of phase delay is inversely proportional to the frequency.
- **For a given phase delay:** the slope angle is proportional to the frequency.

The number of wavelengths of delay is proportional to frequency.

From a practical point of view the phase slope will inform us as to trends in the response of a device as shown in Fig. 8.16. This will become most useful when we move

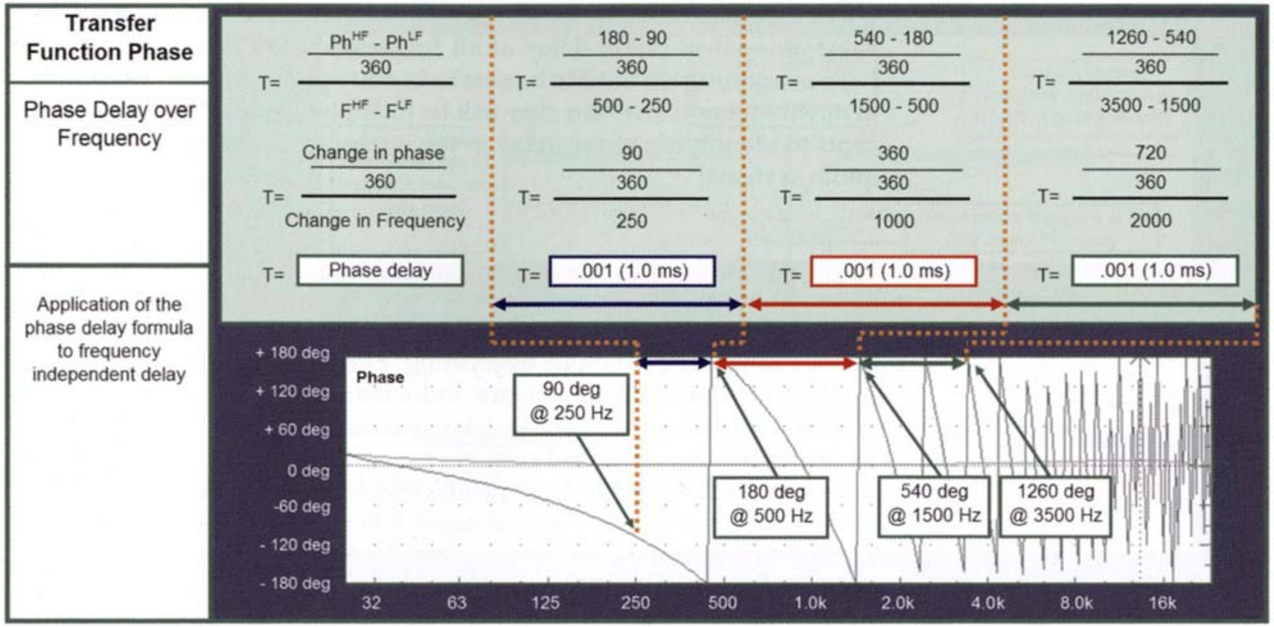


Figure 8.15 Example application of the phase delay formula for an electronic device with a fixed 1 ms delay over all frequencies

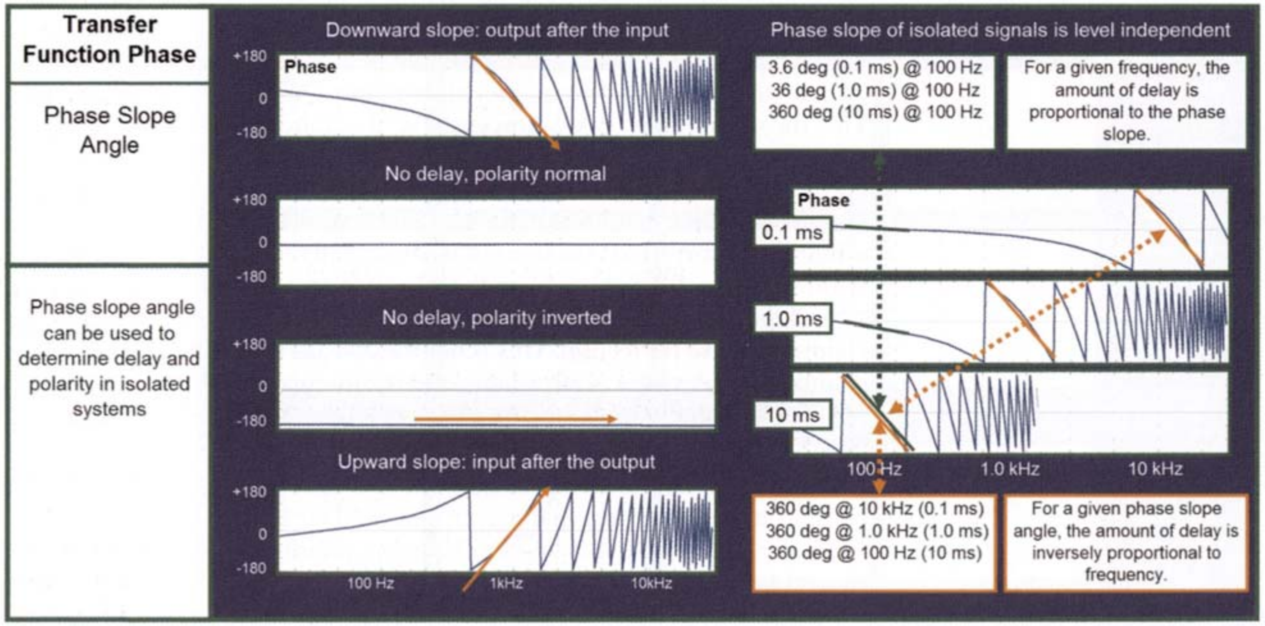


Figure 8.16 Reading the phase slope

on to the phase analysis of devices that do not have the same amount of phase delay at all frequencies. We now have everything we need to be able to identify phase delay at any frequency. The next step will be to apply these concepts to identifying phase delay in the imperfect world of audio systems.

Frequency-Dependent Delay

There are many cases where devices in our audio systems have different amounts of phase delay over their frequency range. For example: everything. Even a cable has phase delay due to its capacitance, inductance and length. Hopefully the amount of phase delay will be so small as to be negligible. As a practical rule, we can assume that any device which has a non-flat frequency response within the audible range will have phase delay differences over frequency. In addition there can be phase delay effects from out of band filters. AC coupling circuits operating below the audible range can have phase delay that reaches up into the audible region. Transient intermodulation filters (TIM) for analog circuits and anti-aliasing filters for digital located beyond 20 kHz can cause phase delay to reach down into the audible range as well. Frequency divider filters, equalization boost and cut, whether analog or digital, are also creators of frequency-dependent phase delay. The most dramatic case, however, is the loudspeaker. A single loudspeaker has a different phase delay over its entire operating range. A well-designed "phase corrected" system will have extended ranges of minimal phase shift, while "sales corrected" versions will have small ranges and wide variance. It has long been the expectation of many a sales department that we users lack the tools to investigate claims of phase perfection. This is not the case, as the dual-channel FFT analyzer is affordable and commonplace. The skills to assess the loudspeaker phase response will soon be yours.

Phase delay is a moving target, but one that we must be able to identify and manage in the field. It is due to the complexity of identifying loudspeaker phase delay in the wild that we have spent the time honing our skills on simple fixed delays in captivity. We are almost there, but we will continue to build up the complexity sequentially by first exploring frequency-dependent delay in electronic circuits.

Equalization Filters

The equalization filter was also introduced in Chapter 1. Again we investigate the filter parameter phase effects. The center frequency, filter topology and filter slope and magnitude of the cut and boost all have effects on the phase response.

For equalization filters:

1. The amount of phase delay is inversely proportional to the center frequency, i.e. a filter centered at 1 kHz will have twice the phase delay as one at 2 kHz.
2. The phase delay will have specific characteristics for each filter model type.
3. The amount of phase delay is inversely proportional to the bandwidth, i.e. narrow filters create increased delay.

The most common relationship is that the phase response slope will be the first derivative of the amplitude response. The derivative is a calculus term that describes the slope, or rate of change, of a function or curve. The first derivative is found by reading the tangent of the slope of the amplitude trace, which yields the phase trace. The tangent of a flat line (flat amplitude) is a horizontal line. The following example uses an equalization filter with cut at a particular frequency. When the amplitude trace bends downward the tangent line angles downward as well. As the amplitude slope increases the tangent turns downward and the phase slope increases. As the filter nears the bottom the amplitude loss rate decreases and the tangent tilt slows. The point where the rate of amplitude loss begins to slow corresponds to the turning point in the phase response. When the amplitude trace reaches its lowest point the tangent goes horizontal again and the phase response can be seen crossing through zero. As the amplitude rises back toward unity the tangent tilts upward and the phase response moves above the zero degrees line.

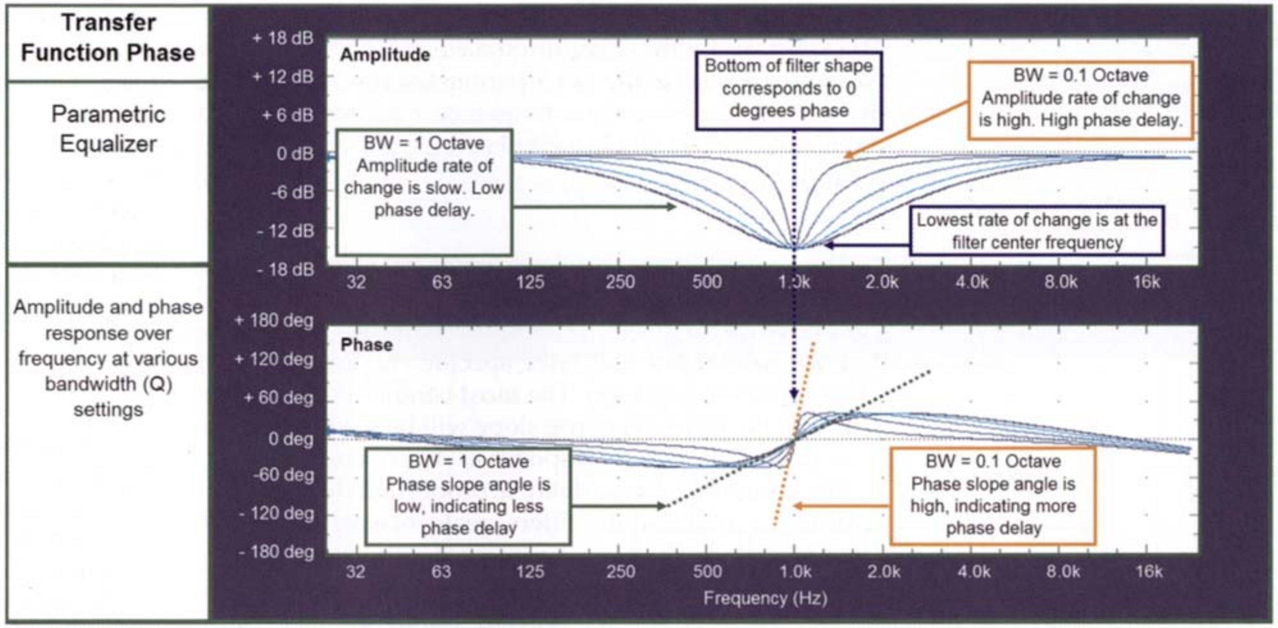


Figure 8.17 Phase slope angles for example filters used in parametric equalizers

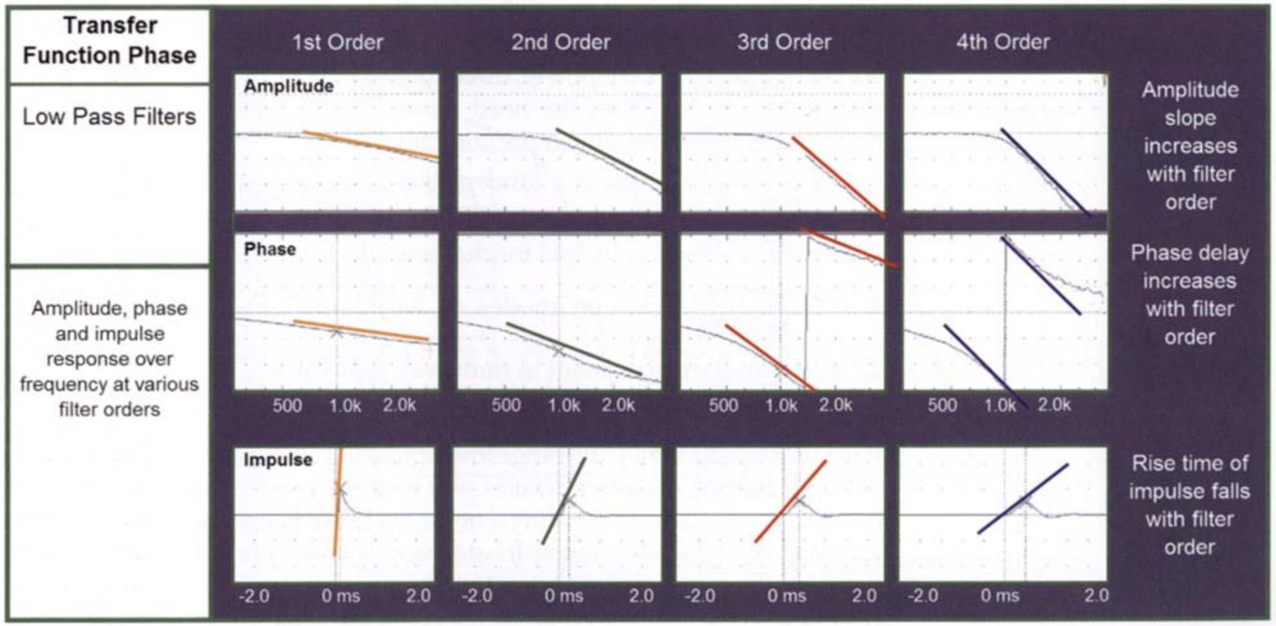


Figure 8.18 Phase slope angle for example low-pass filters used in spectral dividers

Frequency (Spectral) Dividers

The spectral divider was introduced in Chapter 1. Now we will investigate the filter parameter effects on phase over frequency. The corner frequency, filter topology and filter slope (order) all have effects on the phase response.

Effect of low and high pass frequency divider filters on phase delay:

1. The amount of phase delay is inversely proportional to the corner frequency, i.e. a filter that turns at 1 kHz will have twice the phase delay as one that turns at 2 kHz.
2. The phase delay will have specific characteristics for each filter model type. The most common relationship is that the phase response slope will be the first derivative of the amplitude response as described above.
3. The amount of phase delay is proportional to the filter order, i.e. higher-order filters create increased delay.

Loudspeakers

Loudspeakers create the ultimate challenge. They are mechanical devices attempting to produce wavelengths that vary in size by a factor of over 600:1. To produce all of these frequencies at the same level and same time would be a seemingly impossible task for a single speaker. Then, add in our need for high power, directional control and low distortion which creates the need for multiway systems and the mechanics become more complex.

The inevitable result is a loudspeaker response which has different amounts of phase delay over frequency.

There are several principal mechanisms for this:

1. the different radiation modes of individual speaker over frequency
2. mechanical displacement in multiway systems
3. crossover irregularities in multiway systems.

If we measure a single, full-range loudspeaker it is virtually assured that the high frequencies will lead the lows. This is due to the nature of the different modes of radiation of speakers. When the wavelength is smaller than the speaker diameter the sound radiates from the speaker with a piston-type motion. Over this range a well-designed speaker is capable of creating an even phase response over frequency. When the wavelength exceeds the piston size the nature of the radiation changes. One of the artifacts of that change is increased phase delay. The phase delay increases as the wavelength increases in comparison to the speaker size. This relationship is scalar, so even mid-range frequencies such as 1 kHz are large compared to a HF driver that is less than half its size. For subwoofers, the entire operating range contains wavelengths that are huge compared to the radiating devices. As a result, subwoofer phase response will show a steady increase of phase delay as frequency goes down.

In multiway systems three factors play parts in the phase response: physical displacement of the two drivers, the motion factors described above which act on the two drivers differently due to their different diameters, and the electronic response. These aspects of the phase response will affect how the acoustical crossover functions when the speakers are combined. For our current purpose it will suffice for us to observe the phase delay without venturing further into its causes.

Summation Effects on the Phase Slope

There is an important distinction that must be made from wraparound which is indicative of phase delay and which is found in a summed frequency response. If two (or more) time offset signals are combined the phase responses combine also. When the phase responses fall 180 degrees apart the disparity between the summed phase responses is at maximum. This causes a response that appears similar to a wraparound, but in this case represents a real acoustic result, rather than a visual artifact of our analyzer. The wraparound and summation phase disturbances can both be seen in Fig. 8.16.

The key distinction here is that the combined summation phase is a compromise value between two (or more) oppositional parties. The outcome of the interaction depends upon both the phase offset and level offset. If the levels are equal, the combined phase will fall halfway between the conflicting values, e.g. 90 degrees and 0 degrees will create 45 degrees in combination. This seems simple enough at

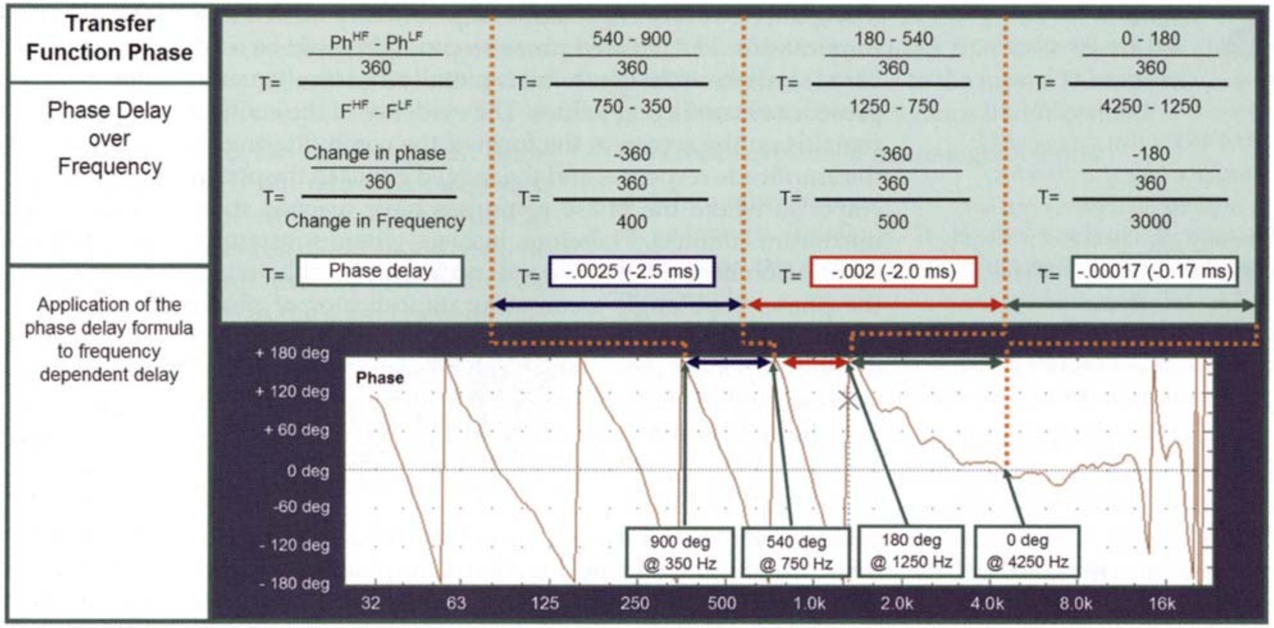


Figure 8.19 Phase response of an example loudspeaker with variable phase delay over frequency

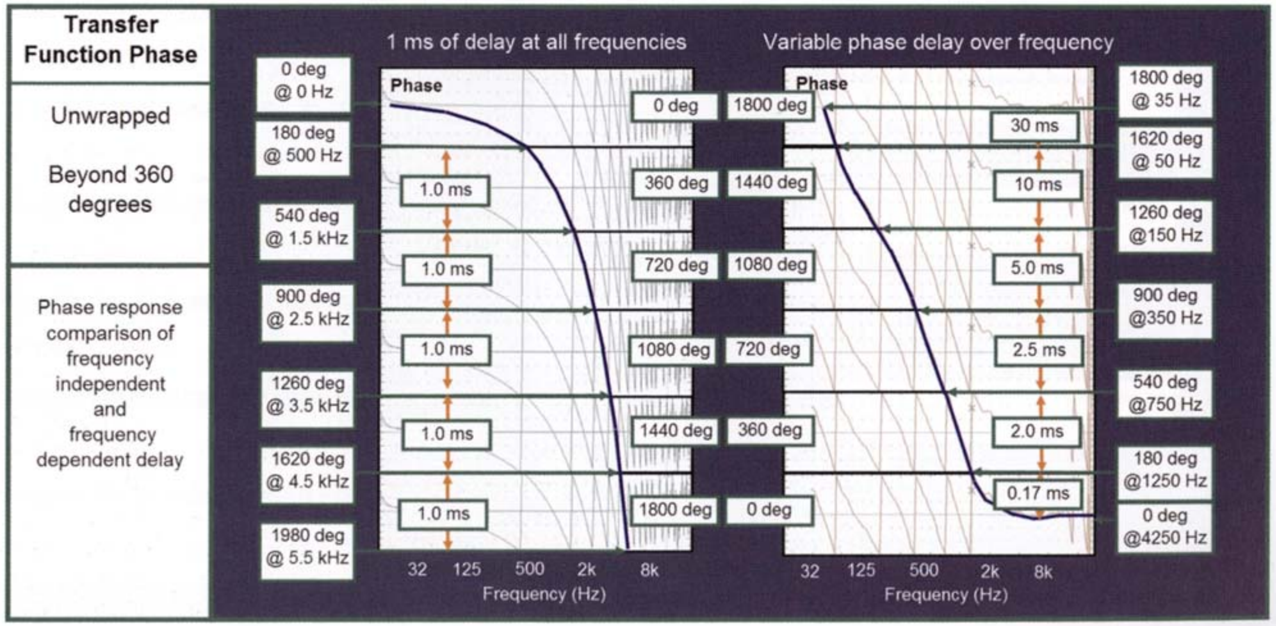


Figure 8.20 Unwrapped phase responses over frequency for the example electronic device (see Fig. 8.15) and example speaker (see Fig. 8.19). The unwrapped representations give a wider context to our view of the phase slope over frequency. The contrast between fixed delay and frequency-dependent delay is clearly seen. All the ghosted phase responses shown in each panel are the same. They are stacked and the trace is unwrapped by connecting the bottom of the first "wrap" to the top of the second

Perspectives I can't tell you the number of times I've heard people say, "Oh they SIM'd or SMAART'd that system and it sounded worse after they left." In response, of course, I explain how any alignment tool is simply a measuring device, and how making this accusation was like saying, "Oh, an electrician came to my house and he volt-ohm'd my breaker panel hut the house burned down." In that case, most would blame the electrician, not the meter, and we should do the same in this situation.

John Huntington

first glance but has important and potentially misleading implications. The reduced phase slope angle could be read as reduced phase delay while it is actually the simultaneous presence of conflicting values. The evidence of the conflict remains on the screen in the form of the comb filtering in the amplitude response, and the ragged edges in the phase response where the phase responses have reached their maximum conflicts. Therefore, in cases where substantial summation is occurring, we will no longer be able to use the phase slope angle as an accurate indicator of phase delay. The combined phase slope does reveal the presence of the component parts, which can be seen superimposed over each other. This is the reason that all references to time in regards to summation phase in this text have always included both of the individual time components (see Figs 2.15 and 2.16).

In the event that the levels are not matched, the compromise value migrates in the direction of the phase of the level dominant party. Once again, size matters. The lower level signal causes slope changes at the same rate over frequency, since this is determined by the time offset. The combined slope angle shows the phase response of the lower level contributor as small disturbances riding above and below the dominant phase trace.

There are some makers of complex analyzers that expand the vertical scale up to thousands of degrees so wraparound can be eliminated, achieving a response similar to the one shown in Fig. 8.15. We should be wary of such results as they are prone to errors when unwrapping phase responses which contain summation from secondary sound sources such as speakers or reflections. The ability to distinguish wraparound from summation phase conflict is a carefully acquired skill.

Coherence

The final component of the frequency response family is coherence. Coherence provides an indication of the stability of the measurement. The coherence computation is a product of the deviation factor in signal averaging discussed previously. If the values for the different samples

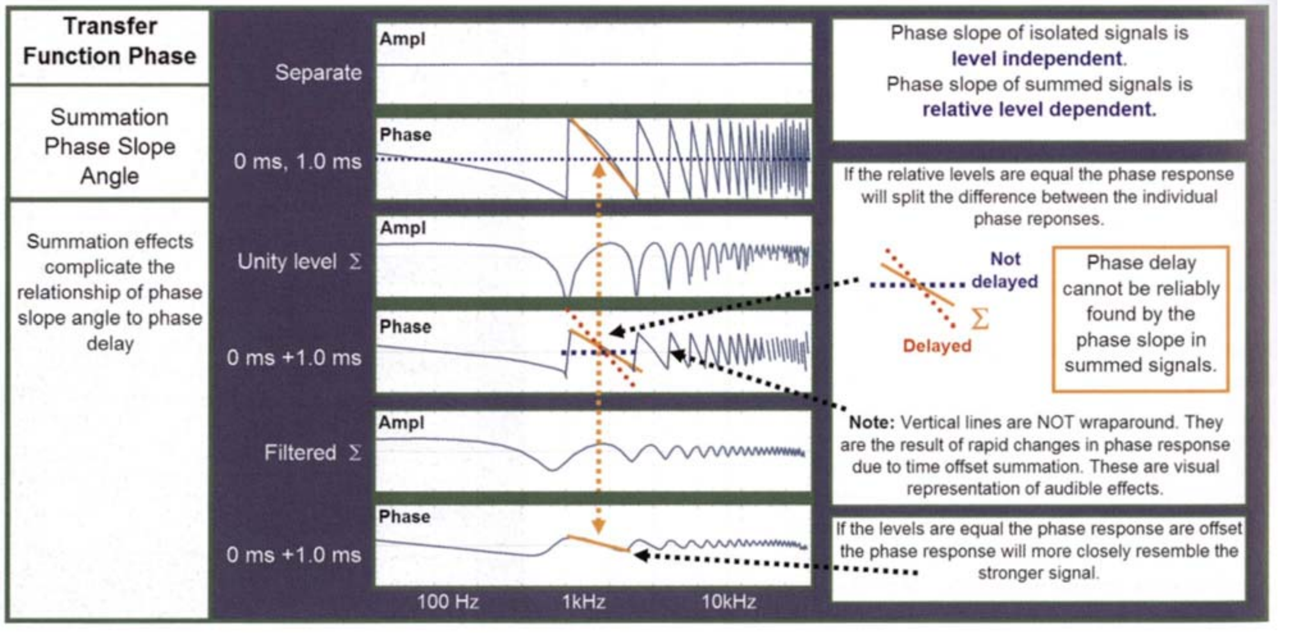


Figure **8.21** Effects of summation on the phase slope angle that comprise an average are consistent, our confidence in the accuracy of the data goes up. Inconsistent answers create doubt, and lower confidence, leaving us with only a gross approximation of the response.

The unstable data that causes low coherence can come from a variety of factors which fall into two categories: errors in the measurement procedures, and instability in the response of the system under test. Naturally we will be very concerned to ensure that our tests are valid before we pronounce that there are defects in the sound system.

A common example of the former, which hounds even the veterans in the field, is forgetting to set the acoustic propagation compensation delay so that the transfer function is executed between signals matched in time. The "Hey dummy! You forgot to set the delay!" screen is shown for reference in Fig. 8.22.

Coherence Defined

The definition of the coherence function as published by a manufacturer of FFT analyzers is the "power at the output relative to the power at the input." (Herlufsen, 1984, p. 28) A number is provided that describes how closely related the output signal is to the input. There are a variety of ways for us to view coherence in our context.

Coherence expressed in analogous terms:

- A data quality index.
- A reliability indicator. It is the analyzer's way of expressing the degree of confidence it has in the data presented.
- The signal-to-noise ratio. Any noise that contaminates our signal degrades the coherence value.
- An indication of the stability of the measurement.

The addition of coherence over frequency to our measurement family provides a whole new dimension. We can now look at a frequency response and assess quality on two levels: the measurement quality and system under test quality. If the coherence is low, either we are measuring it wrong or there is something that is causing inconsistent results. Coherence requires averaging, and looks for discrepancies between each of the individual responses

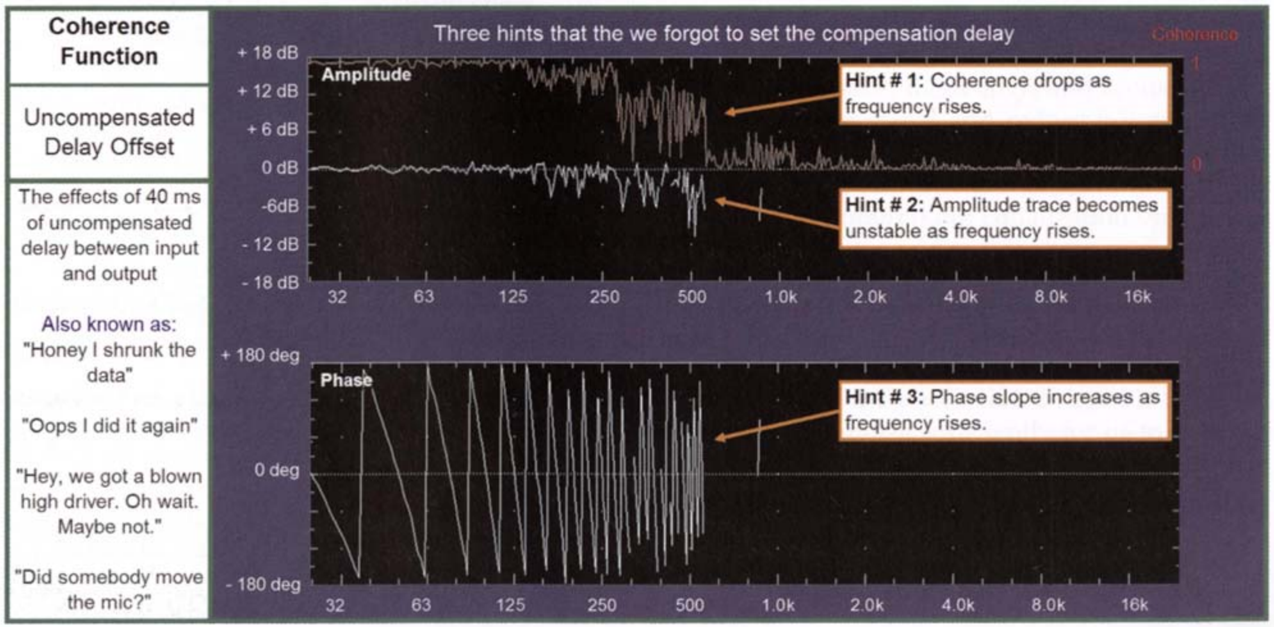


Figure 8.22 Effects of uncompensated delay on the coherence, amplitude and phase

that comprise the average. If every individual response obtains the same values of amplitude and phase we will have perfect coherence. If a new response deviates from the previous values instability is seen by the analyzer, and is reflected as a loss of coherence. Coherence looks at our averaged data and does fact-checking like a detective at a crime scene. If the stories we get recount the same exact details over and over we have high confidence that we are getting a true story. If the details keep changing we lose confidence, and know to take this data with caution or disregard it entirely. Coherence performs our detective work on every frequency bin in our response. In the rough and tumble world of acoustic measurement we will rarely take a measurement where all of the data tells a consistent story.

Coherence is calculated on a scale of 0 to 1 (or 0 per cent to 100 per cent) with 1 being the highest level. Coherence is evaluated for each frequency bin. If there is no difference between input and output (including phase) the coherence is perfect (1). There can be differences, however, which do not degrade coherence, provided that the output signal remains linearly related to the input. Six dB of voltage gain, for example, would linearly alter the output signal relative to the input but leave the coherence unchanged. By contrast, delay between the input and output channels causes the two measurement time records to be drawn from different waveforms. Since the two signals contain unmatched data, their relationship is not stable. This would result in a coherence value less than 1. A third possibility is a completely uncorrelated relationship between the two measured channels which would reduce the coherence value to 0. When the output is linearly linked to the input the relationship is termed **causal,** i.e. the output response was *caused* by the input. When the output signal is unlinked to the input it is termed **non-causal.**

Coherence, above all else, is a measurement of the causal output signal strength relative to the non-causal output noise component. Assuming that we have compensated for any time offset between the measured signals as previously discussed, we can now attribute coherence loss

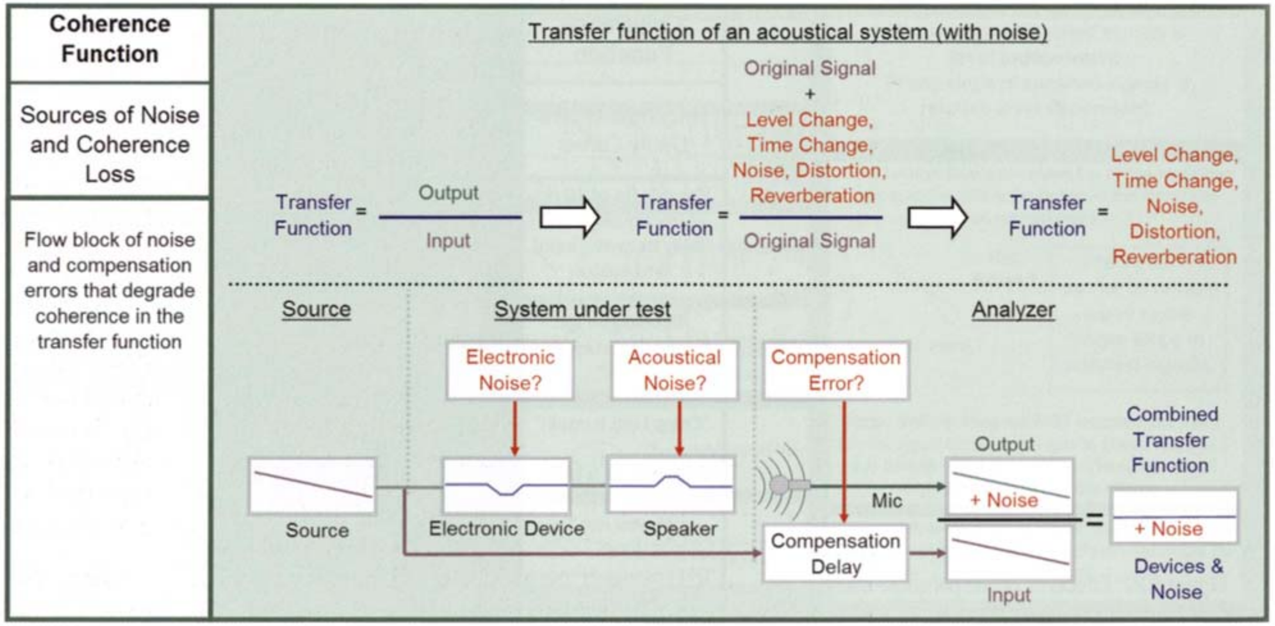


Figure 8.23 Flow block diagram of transfer function with noise sources

to the behavior of the measured device, not an error in measurement. If the output is all signal and no noise, coherence is 1. If signal and noise are equal, coherence is 0.5. If the output is all noise and no signal, the coherence is 0.

We will use some music to create an analogy for the above examples. Consider the case of a musical passage being heard loud and then heard soft. We would recognize the music as different, yet related. Coherence is 1.0. Next let's listen to the opening 10 seconds of a song. Then listen to the song for 10 seconds again, but start five seconds later. The musical passages are half-related. Coherence is 0.5. Finally listen to some good music and then listen to some Kenny G. These two forms of music should be 100 per cent uncorrelated. Coherence would be 0.0.

Factors Affecting Coherence

What are the sources of noise in our measurements, and how does the coherence function detect it? As our analyzer sees it, there are two possible kinds of signals at the output: causal signals that are correlated to the input and those that are not. There are invited guests and party crashers. We can tell who was invited because we have a complete guest list: the input time record. If the waveforms at the output show up on our invitation list we can recognize them. If not, we know they are uninvited. Unfortunately we do not have bouncers to block them at the entrance. The best we can do is to identify them and monitor their effects.

Now let's consider a measurement of a speaker in a room. The input is the original signal.

Causal signals at the output would include:

- the original signal
- copies of the original signal from other speakers
- copies of the original signal from reflections.

The original input signal will be recognized at the output and its level and phase relationship to the input found. Copies of the original would include secondary sources such as reflections or additional speakers that are driven by the same source. It is critical that we are able to differentiate between causal and non-causal signals, because the optimization strategies differ markedly for each. The coherence factor of the causal signal will remain constant over a series of averages, as will the amplitude and phase responses. This is indicative of a stable relationship of the output to the input signal and a stable signal-to-noise ratio. The stable coherence value could be high or low. A strong cancellation from comb filter summation will make a series of high and low coherence values that track the peaks and dips. The coherence at the peaks will be stable and high, while at the dips it will be stable and low. The stability of the coherence removes any doubt as to the fact that the peaks and dips are summation-related and that solutions will be found in management of such interaction. Equalization is an option that is only applicable to stable data; therefore we will be mindful of this parameter.

A non-causal interaction creates instability in the coherence, along with variation in the frequency response data. In such cases the data must have many averages in order to help to steady the instability. Optimization strategies generally are designed to maximize stability by minimizing the amount of non-causal output signal before applying equalization to the stable remainder. An illustrative example would be speaker focus angle and acoustic treatment. We can increase signal and reduce noise with an optimal focus angle and absorption. The coherence value would provide relevant information on optimizing these parameters prior to undertaking equalization on the stable remainder.

The end result is a frequency response that has causal and non-causal elements. The coherence value reflects their mix proportions, with stable values showing strong correlated presence. If the uncorrelated signal is too large the data will never stabilize sufficiently for us to obtain a definitive response of the system. This sends a clear message that some major work needs to be done before equalization is attempted.

There are an infinite number of possibilities for non-causal data sources.

Non-causal sources (include but are not limited to):

- distortion
- hum
- noise
- audience participation
- **HVAC** systems
- forklifts
- moving stage lights
- jackhammers
- late-arriving echoes (from the fixed PPO analyzer point of view)
- late-arriving sound from other speakers (from the fixed PPO analyzer point of view).

The first eight items above are all variations on the same theme. They come and go with no relationship to the original signal. They cannot be stopped, but they can be detected. If they are continuous they can be averaged over time, and their effects seen as deviations above and below the stable frequency response derived from the correlated data. The last two items are a special case, since the signals are late copies of the original. Exactly what is meant by late? To answer this will require a revisiting of the fixed PPO (constant Q) transform.

Direct, Early and Late Arrivals

We will add in one more level of detail: the distinction between causal data that falls within the time record, and that which falls outside. The direct sound is the former. **A** reflection may fall inside the time window, outside it or straddle both. In order to preserve a constant number of points per octave, the time records are long in the **LF** range and short in the **HF** range. If the reflection arrives within the same time record as the direct signal, the reflection will be recognized as causally related. Once the time record acquisition is completed and sent for averaging, the analyzer starts over from scratch, with no knowledge of the previous time record. Reflections that are still bouncing around the room from the previous signal become instant strangers to the analyzer. When that sample is completed, the reflections from the previous sample are treated as noise.

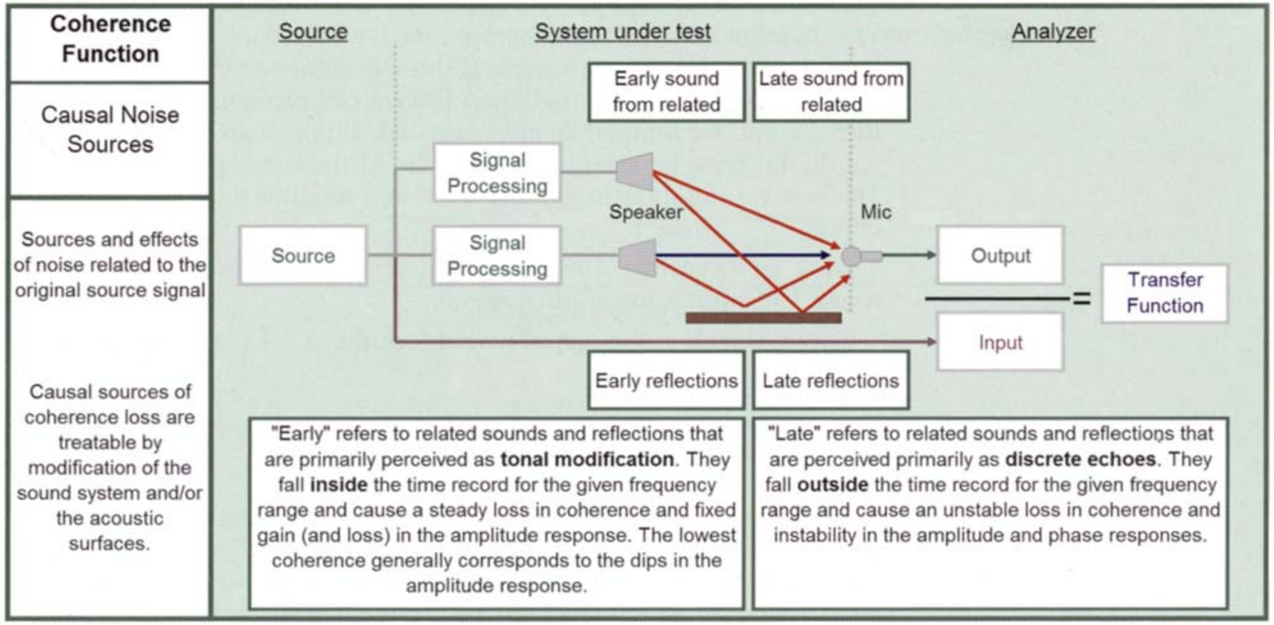


Figure 8.24 The coherence function and causal noise.

How many milliseconds of delay is that threshold? This will depend on the frequency. A 100 ms reflection, for example, falls far outside the short time records in the HF range and well inside the long records of the LF range. The reflection is the child of our input signal, so there is no denying the parental relationship. The distinction in this case is that the analyzer sees the HF portion as non-causal (outside of the time window) and the LF portion as causal (inside the time window).

The composite fixed PPO response is made up of multiple time records that range from a few milliseconds to over half a second. In the resulting response the acoustical properties of the hall are selectively filtered so that a consistent ratio of direct to delayed wavelengths is classified as causal, and later arrivals are classified as noise. Fixed points per octave means, by definition, a fixed number of late wavelengths are allowed in the time window. Recall that the roots of the fixed PPO transform are in our tonal, spatial and echo perception thresholds. The direct sound and the stable summations of early causal signals are the prime candidates for equalization, and are the source of tonal perception. Signals that are perceived as echoes are not shown as stable frequency response deviations, but rather as uncorrelated instability. If the stability is too low, we have strong evidence that equalization solutions will fail, since a stable target curve cannot be found. This will help guide us toward superior solutions for these matters such as acoustic treatment, speaker focus, and phase alignment (delay setting). These considerations are also applicable for signals arriving from other speakers that were driven by the same source. We can look around the room and consider what treatment options remain for various interactive features. For speakers within a coupled array and nearby reflections all options are open. A late arrival from a distant speaker or reflective surface is no longer a candidate for practical equalization above the very low frequency range. The multitime-windowed FFT method (the fixed PPO/constant Q) emphasizes the area of practical equalization by presentation of the high-resolution frequency axis.

Perspectives Learn your tools. No matter what tools you use, learn them. Get comfortable with them. If it becomes clear that it is time for a change, select your new tool with care and then use both the new and old together for a while so you can correlate the two. As you become more at ease with the new tool you can start leaving the other at home. Going from one tool to another cold turkey has always left me with an empty feeling.

Alexander Yuill-Thornton II (Thorny)

A stable response indicates that all system optimization options remain open to us: equalization, relative level, delay setting, speaker focus and acoustic treatment. An unstable response, caused by late reflections and other speakers, is beyond the equalization horizon, but all other options still apply. A completely unrelated non-causal noise source will require solutions outside of the speaker system and electronics. Acoustic isolation from outside sources, or taking away the fork-lift keys, would serve as examples.

Instability in the Causal Signal

There is no guarantee that our causal signal will not change. For example, an equalizer is adjusted while being measured. The output signal would be recognized but its amplitude and phase relationship to the input is changing. This would result in a loss in the coherence reading, even though the device is operating normally and nothing would sound wrong. The coherence is always based on an averaged value. The newest frequency response data is compared to the data already in the averager. If the values are different, the coherence will drop, even if the difference is due to changes in the device, rather than noise in the measurement. Once the changes have stopped and the response has stabilized, the coherence will rise again. Simply put, coherence hates instability, no matter where it comes from.

The behavior of the coherence trace will have three basic trends:

1. temporary drop when the system under test has response changes
2. permanent stable response (high and low) when delayed causal signals are added to the direct signal
3. unstable loss when non-causal noise is added to the output signal.

Impulse Response

The impulse response is another view of the system response. An impulse response contains the same information as a transfer function frequency response as long as they are based on the same time record. The same information is measured about the system but the impulse response provides a vastly different perspective than the

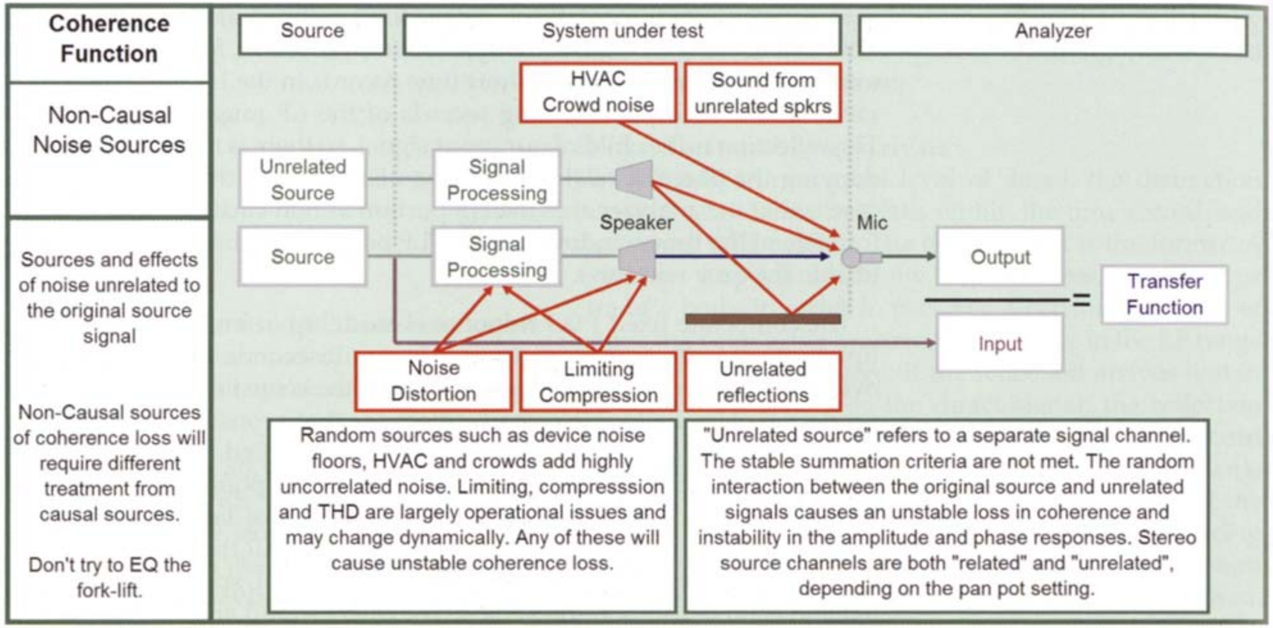


Figure 8.25 The coherence function and non-causal noise.

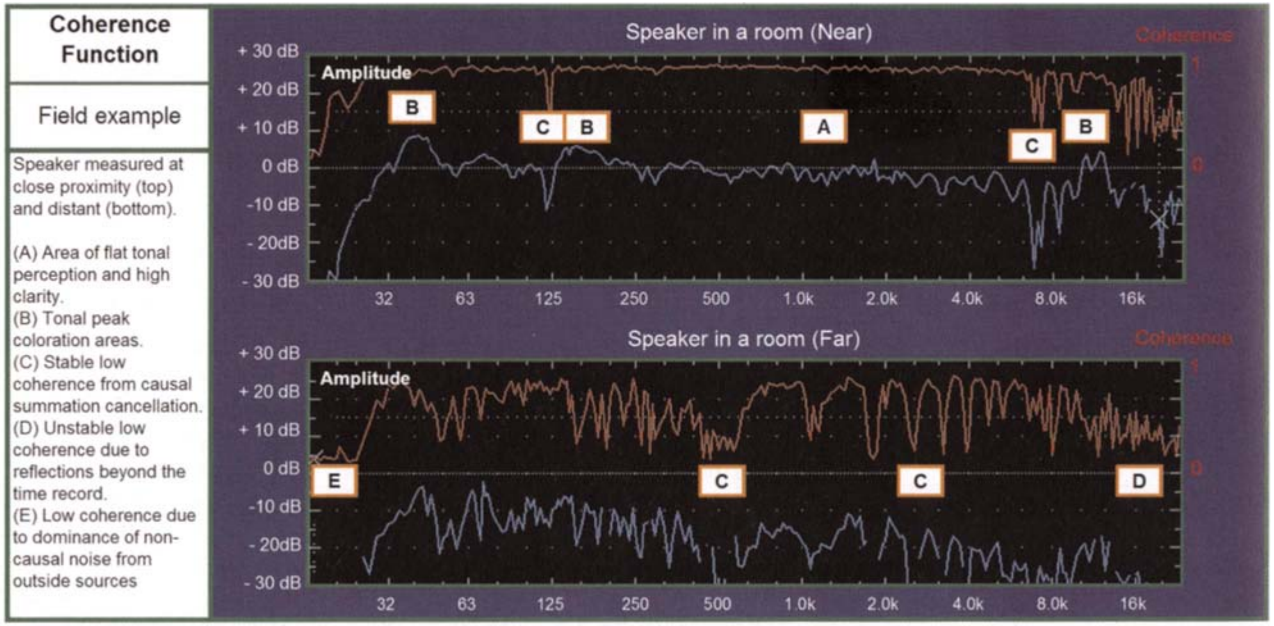


Figure 8.26 Field example of the coherence function.

frequency response. The impulse response gives us a direct view of the time domain of the system under test. The impulse response does not have a frequency axis, but there is frequency information encoded in the trace. Recall that the phase response has no time axis, but as we saw earlier, the timing data can be found.

What is the impulse response? The first matter would be to give it the more accurate term of "calculated" impulse response. The response shown is the calculation of what the system's response over time would be in the event of its excitation by a perfect single impulse. The distinction is that the input signal does not need to be an impulse at all; in fact it can be pink noise or music.

Like its frequency domain counterpart, the impulse response can inform us about the properties of the speaker system, as well as the room. The time-domain representation of the impulse greatly eases the process of separately identifying the speaker's direct sound and the individual room reflections. Whereas the frequency response contains the effects of reflections superimposed on top of the direct sound response, the impulse response displays the arrivals in the order received. These temporal clues allow us to see to identify individual paths by their relative propagation time. The timing sequence, strength and polarity of each arrival can be found. For our speaker system response this will disclose device polarity and its relative level above the reflection pattern. Since the direct sound is our signal we will find a signal-to-noise value in the relative signal strength of our direct and reflected sounds until finally we reach the noise of the system under test (and/or our analyzer). From the room acoustics perspective this timing/level sequence is the decay character of the room.

Before moving forward we need to make a distinction about the equivalence between the transfer function frequency response and the impulse response. Our frequency response measurements are a composite of multiple time records (the fixed point/octave FFT). The frequency domain of each of the individual components of our response would contain the same information as an impulse response of comparable time, but not of simultaneous multiple time windows. Our multiple time records create a quasi-log response in the frequency domain, which selectively shortens our time vista as frequency rises. The impulse response time vista is defined by the length of the measurement. The frequency aspect of the impulse response is linear; since it is based on a single time-windowed FFT.

How does this apply to system optimization? The multi-windowed FFT presents the frequency domain as it perceived in our "tonal and spatial" zones. Long reflection with strong HF content is far outside our tonal perception window, and its effects are beyond our window length in the HF range. We see it in the analyzer only as a loss in coherence or microscopic ripple variance. This reflection is, however, something we can hear as a discrete echo. If we are going to fix it, we will need to identify it. The impulse response will see it, since its window stays open for as long as we ask it. The impulse response opens our view to events beyond the tonal perception horizon and over the line into discrete echo perception. It is as natural to view responses perceived as echoes in the time domain as it was for us to view the tonal response zone through the fixed PPO viewpoint of the frequency domain. The lonely spike that stands out in the crowd of an impulse response is our most likely candidate for discrete echo perception. The impulse response points us toward solutions such as acoustic treatment and speaker positioning that have a high prospect of success.

Acousticians have always favored impulse response measurements and do so with real acoustic impulse sources such as pistols and balloons. Nowadays the preference is towards omnidirectional speaker sources or other devices with known or controlled directionality and impulse response analyzers. The impulse response, whether obtained by old or new methodology, shows the initial arrival of the direct sound and a series of reflections. The level of each reflection can be seen spread along a time line. Huge amounts of information can be learned about the acoustical properties of the room, the rate and character of the decay and more. This information is put to use by acousticians under the premise that a like type of excitation will be used in the hall.

For our amplified sound application, the impulse response will be found only on our analyzer. No pistols or balloons. The foremost roles of this response will be the setting of delay times and the identification of reflections that may benefit from treatment. The fine points of the reverberation time and decay character are largely inapplicable to our application. Why? Recall that in Chapter 4 we discussed the contrasting mechanisms of natural and amplified sound. The amplified sound model isolates the space into a series of subspaces, with minimal emphasis on the room as a homogenized entity. The directional nature of our loudspeakers will create partitioned subspaces in the room. If we want to evaluate decay and reverb, it will need to be done in zones, rather than globally. The next item which reduces the impulse response applicability is the interaction of sources on stage. When signals leak into multiple microphones or arrive from stage monitors, the density of the impulse response for that instrument is effectively multiplied. Every channel on the mixer contains a different amount of leakage, each of which is multiplied by each reflection. The addition of reverberation on the channels at the mix console adds additional density and the saga continues. The reason that this matters is that our transfer function measurements begin at the output of the console. Therefore the measured impulse response shows no hint of the multiplications in arrival density.

This is not to say that these factors do not weigh equally in the frequency domain analysis. They do. But our stated goals in the frequency domain have been very clear: minimum variance, minimum ripple and maximum coherence. We search only to minimize the room interactions. There are no instances where we look at the frequency response with the concern that there is not enough ripple variance. This is not true of the conventional views of impulse response analysis. There is such a thing as not enough reverberation, or decay slopes that are too fast, etc. The evaluation standards for impulse-related parameters, however, are unsurprisingly based on the acoustician's perspective of natural sound transmission with our speaker system substituted as the excitation source. The closest that our amplified sound system will come to this model in practical terms is a monaural center cluster with a direct line input and no outboard reverberation. The omnipresence of stereo speakers, distributed subsystems, open mics and electronic reverberation show the extent to which this model is based on assumptions that are simply non-existent at show time.

This is not to say that acousticians should not evaluate our rooms with impulse responses. It is only to say that the standard "desirable" values are likely to leave us with more reverberation than we need. Those of us doing sound system optimization will find precious few answers in the impulse response that translate directly into action for us. For delay setting, the impulse is the obvious winner, but for this we need little more than an accurate characterization of the direct sound arrival. Data relevant to acoustic treatment considerations on a local level will be shared between the time and frequency domains. Speaker positioning will be done primarily in the frequency domain, but the impulse response may prove useful in the identification of reflections related to position options. Equalization and level setting will be strictly in the frequency domain.

Therefore, we need the impulse response to give clear readings on direct sound and strong reflections. The rest is optional, the utility of which will depend upon how much of our scope of work includes the room acoustics.

There are six main features which can be discerned from the impulse response of a device:

1. relative arrival time
2. relative level
3. polarity
4. phase delay over frequency
5. HF rolloff
6. characteristics 1 to 5 above for any secondary arrivals (echoes, sound from other speakers).

Each of those features has several possible outcomes that can be found in the impulse response.

Relative arrival time has three possibilities:

1. If the input and output signals are synchronized, the impulse is at the horizontal location designated as

synchronized in the analyzer. This point may be on- or off-center, and should have an indication of the amount of internal compensation delay that has been applied to provide the synchronicity.
2. If the output is late the impulse moves to the right by the amount of time offset.
3. If the output is early it moves left by the amount of time offset.

The arrival time of echoes or secondary sources relative to the direct sound is indicated by its horizontal location relative to the direct.

Relative level has three possibilities:

1. If the device gain is unity the impulse amplitude level will be 1 or —1.
2. If the device has a loss in level the impulse amplitude level will be between —1 and 1.
3. If the device has a gain in level the impulse amplitude level will be more than 1 or less than —1.

The complications here are due to the presence of polarity in vertical scale values. Unity gain is 1 if the polarity is normal and —1 if the polarity is inverted. More on this below.

Polarity has two possibilities:

1. If polarity is normal the vertical line goes up.
2. If polarity is reversed the vertical line goes down.

The polarity issue creates a notable side-effect on the amplitude vertical scale. To illustrate this let's consider what happens when gain and polarity both change. It is simple to visualize that 6 db of gain, polarity normal, gives a linear gain value of 2. But what is 6 db of gain, polarity inverted? That would be —2, of course, which can be confused with —6 dB, which it is clearly not. What we have here is a directional component (polarity) independent of level, which is why the vertical scale is linear.

For phase delay:

1. If the phase delay is the same for all frequencies the impulse will be a straight vertical line in one direction.
2. If the phase response is not flat over frequency the impulse will be horizontally stretched. The size and shape of the stretching will vary with the frequency, bandwidth and phase delay time of the affected areas.

HF rolloff has two possibilities:

1. If the device operates without loss over the full frequency range of the analyzer, the impulse will rise as a straight line.
2. If the device has high-frequency rolloff the rise rate of the impulse will be reduced proportionally.

Reflections and secondary arrivals will behave as follows:

1. If there are reflections or other arrivals in the response they will appear as additional impulses.
2. All of the above characteristics also apply to the level, timing, polarity, HF rolloff and phase delay of the echoes and other arrivals.

These are common features to most versions of the impulse response. There are additional computations that can be done to create a variety of manufacturer-specific enhancements to the impulse response. The most pressing limitation of the standard impulse response is the linear vertical scale for the amplitude response, which is not optimally suited for ease of identification and characterization of echoes. Acousticians prefer to enhance the vertical scale through a conversion process termed the Hilbert transform, which creates a log display. Very simplistically this can be described as taking the absolute values of the linear impulse response and converting the vertical scale to log. The negative-going aspects of the impulse are folded upward and join the positive, creating better visibility of the echo peaks above the noise. This computation provides far superior viewing of echoes and displays their relationship to the direct sound in dB. However, we must exercise caution with this display as the process of taking absolute values obscures a parameter that is of little interest to acousticians but critical for audio professionals: the polarity of the signal.

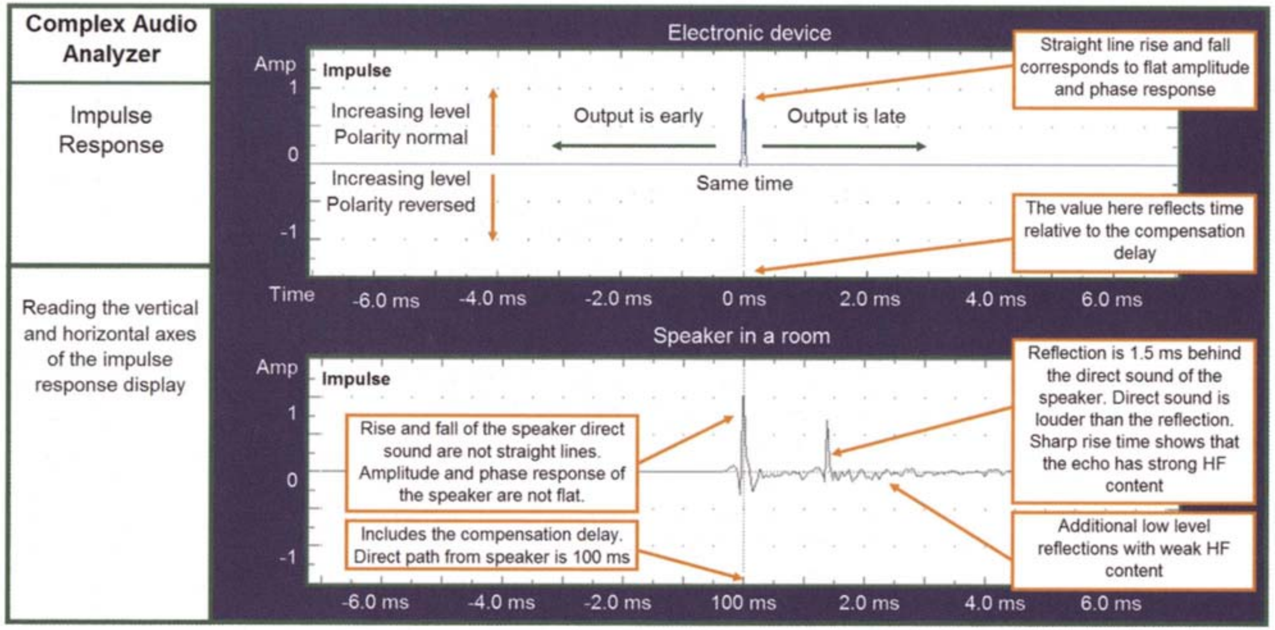


Figure 8.27 Impulse response

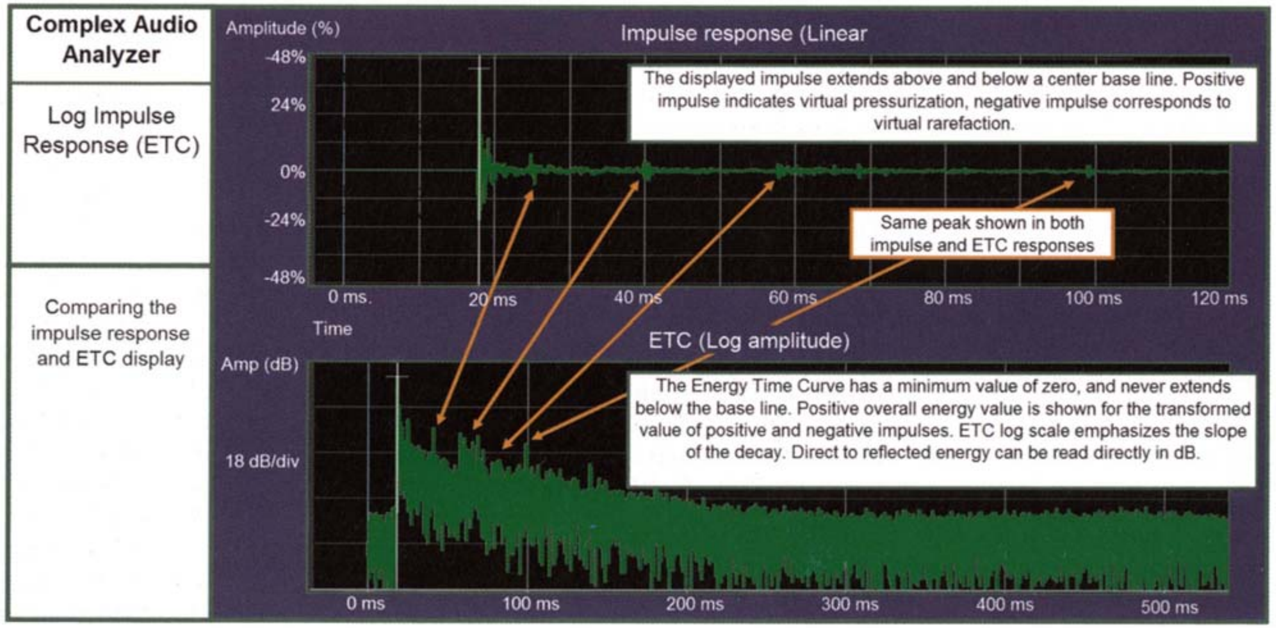


Figure 8.28 Comparison of log and linear impulse response (courtesy of SIA-SMAART)

Transfer Function Applications

Transfer function analysis is the foundation of many of the procedures that comprise the verification and calibration stages of the optimization process. The transfer function capability will allow us to move freely among the system components and subsystems, acquiring the necessary data and verifying the resulting solutions.

Transfer function applications include:

1. architectural acoustical analysis using both the impulse and frequency response
2. speaker positioning using the frequency response
3. level setting using the frequency response
4. delay setting using both the impulse and frequency response
5. equalization using the frequency response
6. polarity verification using the impulse and/or frequency response.

These techniques will be outlined in detail in the following two chapters.

Additional Features

The features presented here barely scratch the surface of the computational power available in the dual-channel FFT, much less the full gamut of modern acoustical analyzers. There are Nyquist plots, cepstrum responses, Wigner distributions, time spectrographs, the modulation transfer function, intensity computations, RASTI, STI II and on and on it goes. Each of these computations contains information about our sound system obtained and/or presented in unique ways. There is no technical reason to exclude these computations from the discussion here. The math behind them is likely to be every bit as sound as the math behind the functions that we have covered in detail. The simple fact of the matter is this: as a practitioner of system optimization for over 20 years, I have not yet found a practical implementation of these transforms that I can translate into direct optimization action, which precludes me from offering advice in such matters. By contrast the basic functions described have provided answers that result in equalizer, delay, level, speaker focus and where to put the fiberglass. Ours is a practical trade, not a research and development foundation. For this reason, the optimization stage will be driven by these basic FFT analyzer functions: single-channel spectrum, transfer function amplitude, phase, coherence and impulse response.

Other Complex Signal Analyzers

The fixed PPO (constant Q) dual-channel FFT analyzer as described above is not the only type of complex signal analyzer. There are many others at the time of this writing and there are certain to be more developed as time goes on. There are many of the standard linear frequency span FFT analyzers, there are specialized stimulus response systems and on and on. We will make no attempt to be comprehensive here, but rather take a moment to discuss the analysis systems that have played a significant role in the field of sound system optimization.

The most significant of the other types is the "calculated semi-anechoic response" family of analyzers. These analyzers strive to capture the response of a system as would occur if the system were free of some or all reflections. This family of devices has far greater "noise immunity" than the fixed points per octave FFT analyzer; i.e. their ability to capture the direct sound alone is unequaled. The extent to which the response is free of the reverberant response depends upon how quickly the measurement time window is closed and is user settable. Therefore, the user may select to view the system with the direct sound only or with some combination of the direct and reverberant sound. Such systems are now based on the FFT and are calculated with linear frequency resolution. The original concept for the system was developed by Richard Heyser as time-delay spectrometry (TDS) and was later manufactured as the TEF system. The basic scheme of the system is as follows: a sine wave sweep is introduced into the system at a constant rate of linear frequency rise over time. A tracking band pass filter sweeps at the same rate as the source. The filter is delayed by the transit time to the measurement mic so that it matches the sine wave frequency at the time of the arrival from the speaker. Both the signal sweep and the filter move upward. By the time the low-frequency reflection has arrived at the mic, the filter has moved to a higher frequency and rejects its contribution to the frequency response. Any reflection slower than the sweep rate will be rejected. The faster the sweep, the fewer room surfaces are seen and the frequency resolution falls proportionally. This response, like the calculated impulse response discussed earlier, is a mathematical construction of the semi-anechoic response rather than an actual one.

Another system developed in the 1980s is the maximum length sequence sound analyzer (MLSSA). This system achieves a similar response by truncating the impulse response. A response is taken of a known periodic source that contains all frequencies in a mathematically calculated non-repeating pattern (the maximum length sequence) and a calculated impulse response is acquired. The impulse response is then modified by setting all of the values past a selected amount of time to zero. This process, termed **truncation,** deletes all of the reflections (or other signals) past the given time from the data. As we discussed earlier, one of the great attributes of the FFT is that it can convert in both directions. In this case the truncated impulse response is used to create a frequency and phase response free of all the ripple variance effect caused by the reflections past the truncation point.

It is possible to obtain responses on the calculated semi-anechoic systems that matches that of the fixed PPO dual-channel FFT platform, but not likely. Which is the "true" response? None of them. They all have a limited length time record, rather than the continuous "real-time" nature of the human hearing mechanism. Which is the best for the task of sound system optimization? This has been the subject of much debate.

If the same techniques and methodologies were applied it is likely that operators of different measurement platforms would concur on three out of the five major categories of sound system optimization: level setting, delay setting, and speaker positioning. These decisions rely primarily on comparisons of different measurements. The process of comparison renders the individual differences in each platform less significant. Two measurement systems may see a given position differently, but they are more likely to see the relationship of that point to others in a similar fashion. Decisions made on a comparative basis are much more immune to the absolute differences in the evaluation of any single point.

The categories of equalization and architectural treatment are where the differences arise. The ripple variance interaction of the speaker with the room is viewed in an absolute sense, and the differences in the data will cause different solutions to be indicated. Because the calculated anechoic systems use a fixed time record for each data reading, the frequency resolution is linear. The cost of using a short time window for all frequencies to remove the room is that it removes details of both the direct and reflected sound in the low frequencies. If a speaker is measured in a room at a significant distance, the calculated anechoic systems will show less low-frequency content than would the fixed PPO FFT. As described previously, there are substantive errors in the data if the frequency resolution is too low. Short time records, to remove early reflections, are fine in the high end, where the wavelengths are short, but will smooth over all the details in the low end. Even the response of something as audible as a 0.1 octave parametric filter set to a 15 dB boost at 40 Hz will be smoothed away with short time records. A typical fixed PPO analyzer has some 24 points per octave resolution from 160 Hz on up. To achieve this we will need a maximum of some 640 ms of time record length, with less and less as we go up in frequency.

Consider the fact that to obtain even the ridiculously low resolution of one point per octave at 50 Hz requires a 20 ms time record (one period). There are likely to be high-frequency reflections in the data with a 20 ms time record (200 wavelengths at 10 kHz). If we shorten the time record to 2 ms, to keep reflections out of the data, we won't have a single data point below 500 Hz. Proponents of those systems advocate taking a series of different speed sweeps or changing time record lengths to get them the data from the different frequency ranges. This is reminiscent of the FFT without the fixed PPO transform as described in the

Preface (the "periscope"). In order to get a complete high-resolution response the analyzer must take measurements one after the other with settings optimized for each range. As a result, like the early versions of the FFT, these systems are so cumbersome to operate that their influence has not been able to spread beyond a small number of devoted practitioners.

In spite of this, the great majority of the techniques described in this book can be applied to the operation of the calculated anechoic measurement systems and no disrespect is intended to their advocates. The setting of delays, the procedures for optimizing speaker angles and many more are functionally equivalent. It is principally in the subject of equalization that the systems are opposed. It has long been the position of the author that there are equalizable portions of the speaker/room summation and that they must be detected to be managed. To remove them from view would deprive us of the option of providing solutions. The fixed PPO FFT analyzer provides superior viewing of the equalizable portions and therefore will be the foundational measurement tool for this text.

Analysis Systems

The perfect analyzer alone does not a sound system optimization tool make. There is more to a car than an engine. The fully assembled analysis system is a tool dedicated to our specific task: tuning the sound system. It may also provide a host of additional capabilities, but the core application is optimization. The major players in the test and measurement industry make research tools. None has come forth with a tool that did not require us to compile an assortment of parts and pieces to make an analysis system. When no other option was available all of the optimization systems consisted of general-purpose analyzer engines with custom audio interface hardware. Since then dedicated systems have been created by, or managed by, audio manufacturers.

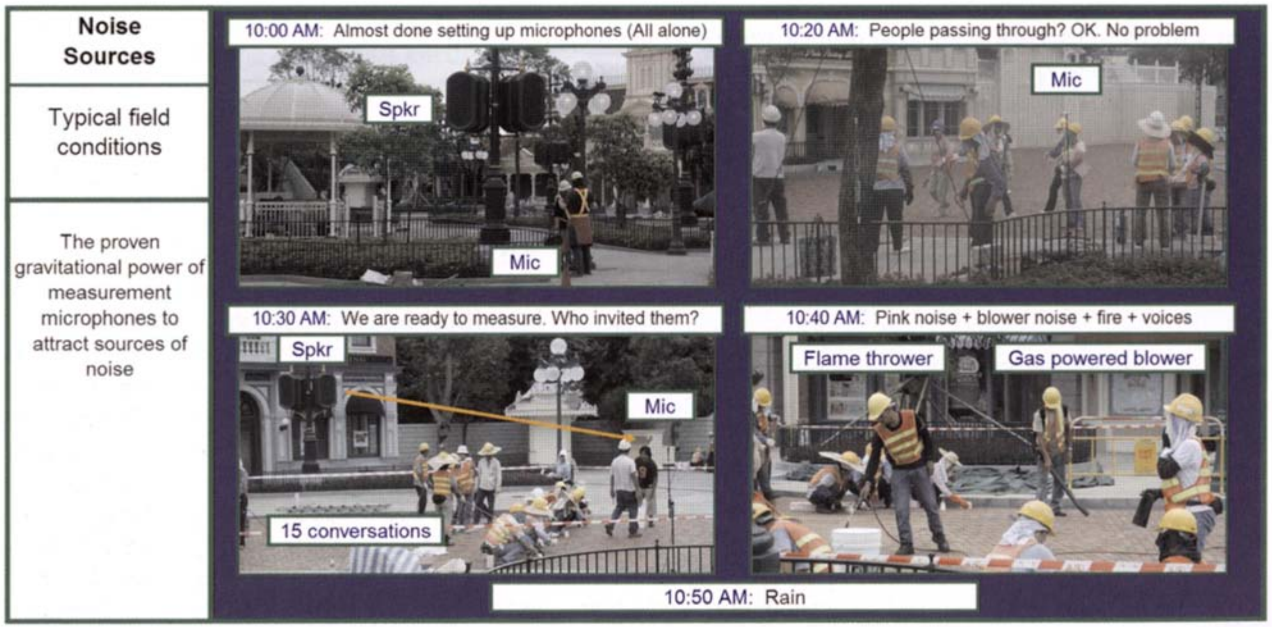


Figure 8.29 When we began our setup there was no one in sight and perfect quiet. By the time we were ready for our first measurements we had dozens of workers sealing the grout, a flame thrower to pre-heat the bricks, a gas-powered blower to clean out the dirt and a brick cutting tool all joining the measurements. This is an illustrative example of "non-causal" coherence degradation

Perspectives To paraphrase Yogi Berra, "Obtaining good data is 90% of the measurement process. The other half is interpretation."

Paul Tucci

A dedicated optimization system contains everything we need to access the sound system and make decisions. The access must be carried out without interrupting the signal flow or injecting hum or noise into the system. We must be able to access both line level signal and microphones for comparison and must have delay compensation for latency and acoustic propagation delays.

The analysis systems that meet our criteria at the time of writing share the common features of the multitime windowed dual-channel, fixed-point-per-octave FFT transform. They are optimized for transfer function frequency response as experienced inside of our tonal and spatial perception zones. They must show relative amplitude, phase and coherence over frequency in resolutions up to at least 24 PPO. We will need an impulse response in order to identify time offsets between speakers and reflections, identify reflections and see discrete echoes that are outside of our multitime windowed frequency response. There are two systems in wide use at the time of writing that fit this model. They are listed here in both chronological and alphabetical order. Each system has unique features/ benefits and hardware approaches. The matter of which platform is the best for a particular application will be left to the reader and the manufacturers. They share the key features just outlined and one additional parameter: the letter S.

- SIM™ (Source Independent Measurement) by Meyer Sound.
- SMAART (Sound Measurement Acoustical Analysis Real-Time Tool) by SIA Software.

Examination is the key to knowledge of our system. We must remember, however, that no disease has ever been cured by diagnosis alone. Treatment will be required when there is a need for a cure. Once the treatment plan is enacted we can verify the result with our examination tools. Only then will we know if our diagnosis was correct. The treatment plans and the proof of their effect for our sound system will comprise the remainder of this book.

Reference

Herlufsen, H. (1984), *Dual Channel FFT Analysis (Part I),* Bruel & Kjaer, Denmark

Verification

verification *v.t.* ***1.*** *a confirmation of truth or authority.* ***2.*** *a formal assertion of validity.* ***3.*** *the process of determining whether or not a set of established requirements are fulfilled.* ***4.*** *establish the truth or correctness of by examination or demonstration. Synonymous with: confirmation, check, or substantiation*

Concise Oxford Dictionary

Introduction

The end product of the optimized design will be speakers in position, acoustic treatment in place and equalizer, delay and level settings programmed into signal processors. These settings and decisions come about as the result of very carefully conducted readings of the system response in a variety of acoustic and electronic locations. One decision leads to the next as a complex system is woven together out of the fabric of the individual elements.

Wouldn't it be depressing and embarrassing to learn we had equalized speakers which had polarity reversals in the wiring, losses due to unbalanced cables, or large amounts of distortion and other assorted unexpected "features." Perhaps an entire series of decisions had been predicated upon the response of a speaker that had cardboard left by the scenery painter in front of it (this really happened). When these types of challenges are found at the beginning they are part of the process of discovery. When they are found near the conclusion they are an embarrassment and huge loss of time. If they are discovered after we have signed off on the job, we may soon find ourselves asking, "would you like fries with that?" Our entire system calibration will come crashing down like a house of cards over false assumptions that should have been discovered before safety cables were tied in to the rigging or settings were input into a processor.

The verification stage ensures that these types of challenges are sorted out in advance of the calibration stage of the operation. This stage consists of individual checkout of each of the system components, and the interconnection wiring. The verification process is a series of routine tests that, for the most part, should turn up results indicating that there is no problem. There is a simple rule of thumb: if we follow all of the verification steps we will find no problems. If we skip verification and assume that things are correct there will be problems which will show up much later, with potentially catastrophic results. Personally, I would sooner leave a system verified and uncalibrated, than calibrated and unverified.

This section outlines the verification procedures, the expected results and how to identify problems.

Test Structure

This is a test. I repeat. This is only a test.

Tests are quests for answers. Our tests are not open-ended, or philosophical. They are very specific questions intended to seek out very specific answers. To obtain reliable answers we must structure the questions very carefully. The structure of the question is the test procedure.

A test procedure has the following features:

- What do we want to know?
- How can we learn it?
- How can we quantify it?
- What is the expected result and/or acceptable measure or quantity?
- What is the outcome, or end result?

First we must define the **subject** of the question. What do we want to learn? The polarity of a processor? The maximum level of a speaker? The subject has two parts: an attribute and an object. Polarity and maximum level are the unknown attributes to be discovered. The processor and speaker are the objects.

The **outcome** is the answer to the question in some form of units or quantities. This could be reverse polarity, or 112 dB SPL if we continue our above examples.

We will then develop a **procedure** to find them. This is a specific set of tests performed for reliable repeatable results. The outcome of the procedure is then viewed in comparison to the expected or acceptable results. Continuing the example above, reverse polarity is an unexpected outcome for our processor. The 112 dB SPL reading at 1 meter is acceptable for a given speaker model, since it is within its manufacturer published range.

The physical object of our inquiry is the **device under test** (DUT). This refers to whatever we are measuring, whether it is a single cable, speaker, electronic component or the entire transmission chain. The test signal is the **source** which, in many of the verification phase tests, is a specific known signal. Some tests require no source at all.

We are done with theoretical constructs. The system is installed and it is time to find out what we have. Let the games begin.

Testing Stages

The verification of an installed system moves through three distinct stages, each of which contains a series of individual test procedures:

- Self-verification: test the analysis system to ensure that it can measure the sound system accurately.
- Pre-verification: check out the system before calibration.
- Post-verification: check out the system after calibration.

The analyzer self-verification is required to ensure that findings about the sound system are actually in the sound system, not the diagnostic tool. Our test signals will need verification. Single channel measurements need a high quality pure sine wave. Transfer measurements need a full range (not necessarily flat) source.

The acoustic side is more complicated since the electrical signal arriving at our analyzer input has gone through a transducer: the measurement microphone. A calibration parameter must be provided to translate the values into dB SPL. A common tool for this is a microphone calibrator which supplies a known acoustic level. The mic sensitivity is derived from the measured electronic level. For users of multiple microphones the relative values of sensitivity and frequency response must be factored in.

The system pre-verification focuses on ensuring that the system has been installed as indicated by the system design specifications. This includes the obvious wiring and electronic verification and a host of other details such as the speaker positions and initial focus angles. The latter stage consists primarily of a verification of symmetry in the calibrated system. For example, a simple stereo system is initially fully pre-verified and then calibrated on one side. The second side is then post-verified to ensure that it is a symmetrical match to the first. Post-verification continues indefinitely. For example a permanently installed system can be continuously verified to ensure that it has maintained its originally calibrated response.

There are many tools that can be put to use for the verification process. A list of the tools discussed in the previous

Examination tool		Verification role
Physical tools	Inclinometer	Determine vertical focus angle for speaker and/or vertical splay angle between speakers.
	Protractor	Determine splay angle between speakers or a speaker and a surface.
	Origami	as above
	Laser pointer	Determine focus angle for the speaker. Especially useful for post calibration verification of symmetry.
	Thermometer	Establish environmental baseline for variable temperature environments.
	Hygrometer	as above
Simple audio tools	VOM	Continuity testing. High voltage testing such as amplifier outputs and line voltage. Essential general purpose tool.
	Polarity Tester	Cable testing
	Listen Box	Continuity testing.Signal routing. Hum and distortion detection.
	Impedance tester	Verify the wiring of speaker lines and the presence of speakers on the line.
	Oscilloscope	Misc. signal tests. Amp outputs and other high voltage testing. Detection of DC and of oscillations beyond the audio band.
	Sound Level Meter	Full or partial bandwidth SPL verification, microphone calibration. General and weighted SPL measurements
	Mic calibrator	SPL calibration of microphone sensitivity
	RTA	Doorstop
Complex tools	Ears	Advance troubleshooting. Signal routing, continuity. THD, noise & frequency response. Essential general purpose tool.
	Eyes	Advance troubleshooting. Path obstruction detection. Symmetry error detection. Noise source detection. Smoke detection.
	Dual Channel FFT analyzer	Electronic and acoustic signal test. Details to follow

Figure 9.1 Verification test reference

Perspectives An important requirement for permanent installations is maintainability. Once the system is tuned it needs to stay that way. Using the pictures from the timing I can verify and correct if necessary any deviations. This is particularly useful when gear has been replaced after a failure.

Bob Maske, Disneyland Resort

chapter is shown in Fig. 9.1. Here we can see the various applications for these tools, most of which will require no further explanation. The remainder of this chapter will be spent on the complex tools found at the bottom of the chart.

In our case the calibration process will include everything from the console outputs to the speaker. Therefore the verification of the components in that part of the path is mandatory. Optionally we may also choose to verify components upstream or outside of the transmission system.

Access Points

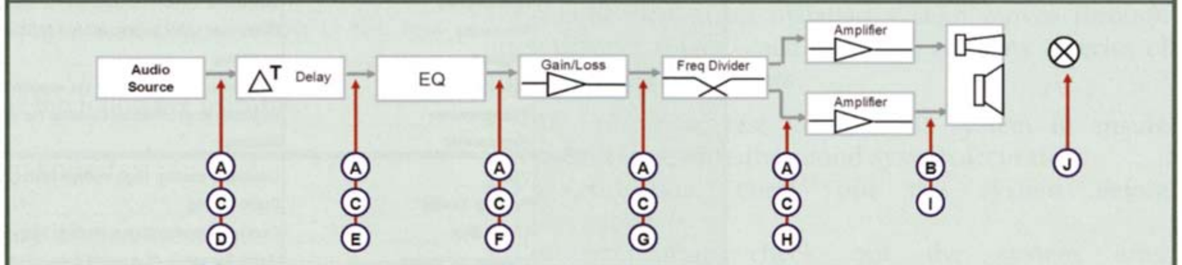
The basic analog signal path flows along a route as shown in Fig. 9.2. Each of the components will need to be verified in various ways, to ensure that the system is ready for calibration. Measurement of a device requires access points at its input and outputs. If these are not provided through patch bays, etc., the device will need to be taken off-line for verification. The required verification parameters for each device are shown in the reference chart. Digital systems will also require verification. Access to the signal at points in the digital network will be required for us to provide verification testing. For our purposes here, we will focus on the analog signal path, which provides easy access. The verification of digital systems can be deduced from this approach at whatever access points it can provide.

Test Setup

There are two general-purpose setups for our verification process: single channel and dual channel. Each configuration is optimized for a particular set of tests. These tests each in turn have electronic and acoustic variations, giving us the total of four basic setups shown in Figs 9.3 and 9.4. Test setup 1 provides single channel measurement of the device (or series of devices) with a known source signal. Test setup 2 is a transfer function measurement scenario that spans any number of devices between the two measurement access points. There are also two additional specialized transfer function setups that provide verification of microphones (Figs 9.5 and 9.6). Refer to these flow block diagrams for the test procedures outlined in this chapter.

Verification can be performed on individual components and on the system as a whole. The component test will inform us about the internal gain structure and other

Verification Access: Analog signal path



Access Pt.	Device Under Test	Verification Parameters
A	Balanced line interconnection	Gain, polarity
B	Unbalanced speaker cable	Polarity
C	Line level device	Gain, polarity, THD, maximum in/output capability, latency, display accuracy, function testing, noise floor, frequency response, phase response
D	Mix console w outboard gear	Gain structure, outboard polarity & latency, meter accuracy
E	Delay line	Anti-aliasing filters, time display accuracy
F	Equalizer	Filter performance, center frequency, level and bandwidth display accuracy
G	Level adjust device	Level display accuracy
H	Active frequency divider	Corner frequency, level and slope display accuracy
I	Power amplifier	Voltage gain, polarity
J	Speaker	Polarity (all ranges), frequency and phase response, THD, maximum SPL

Figure 9.2 Access points for verification and the parameters to be tested

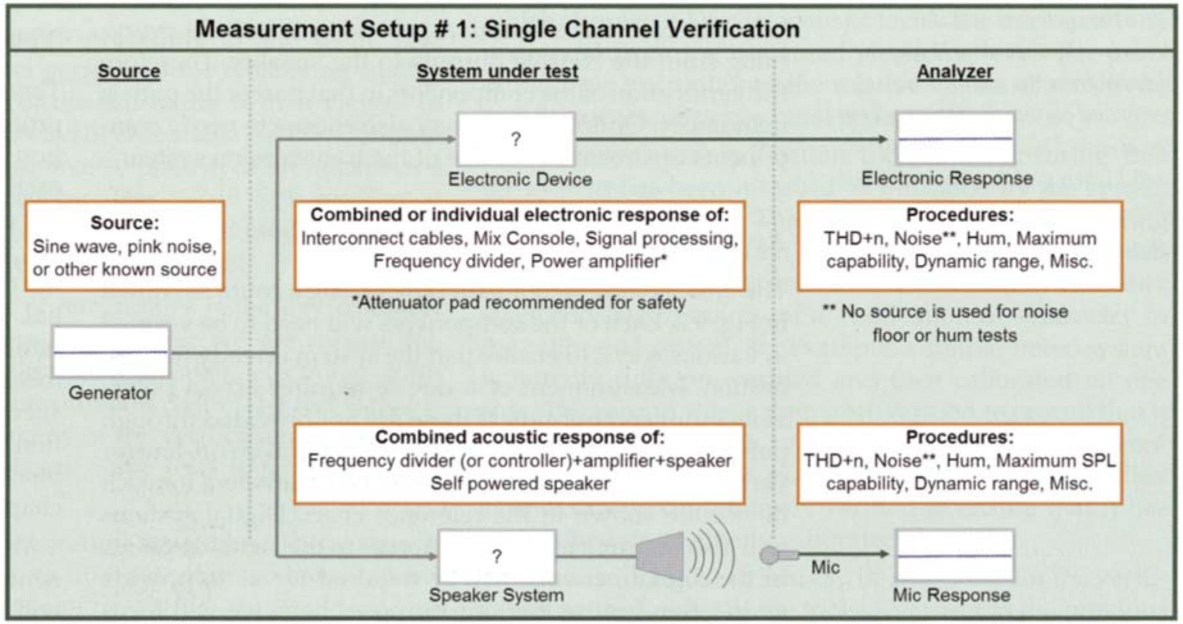
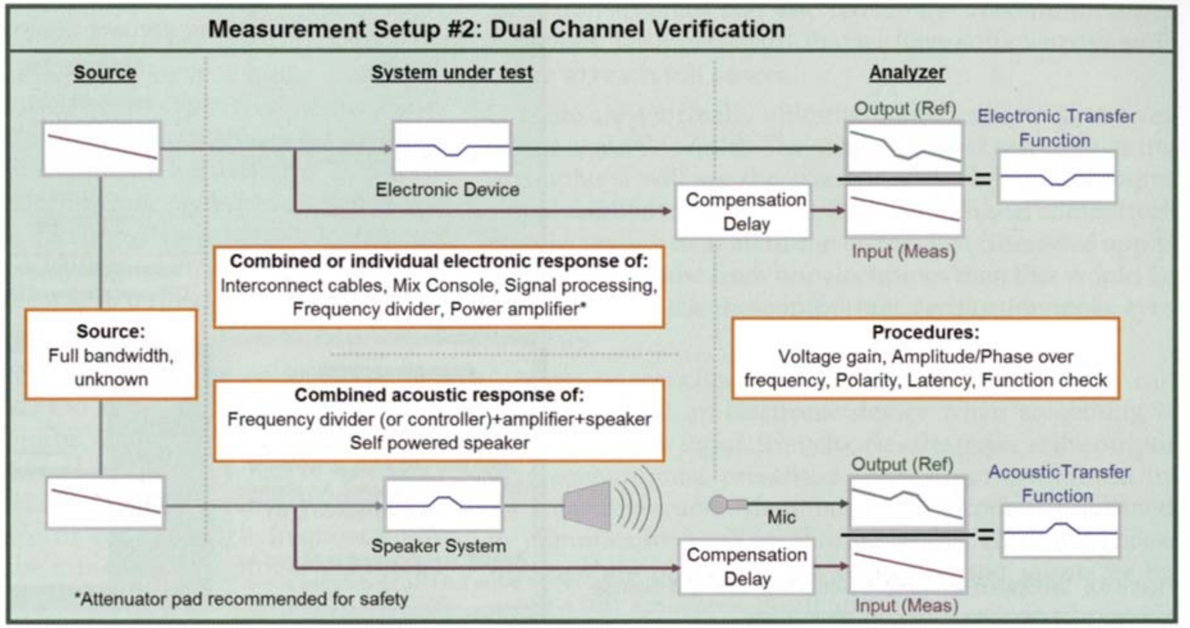


Figure 9.3 Verification test setup 1 flow block of the test setup for single channel verification procedures



Verification test setup 2 flow block of the test setup for dual channel (transfer function) verification procedures

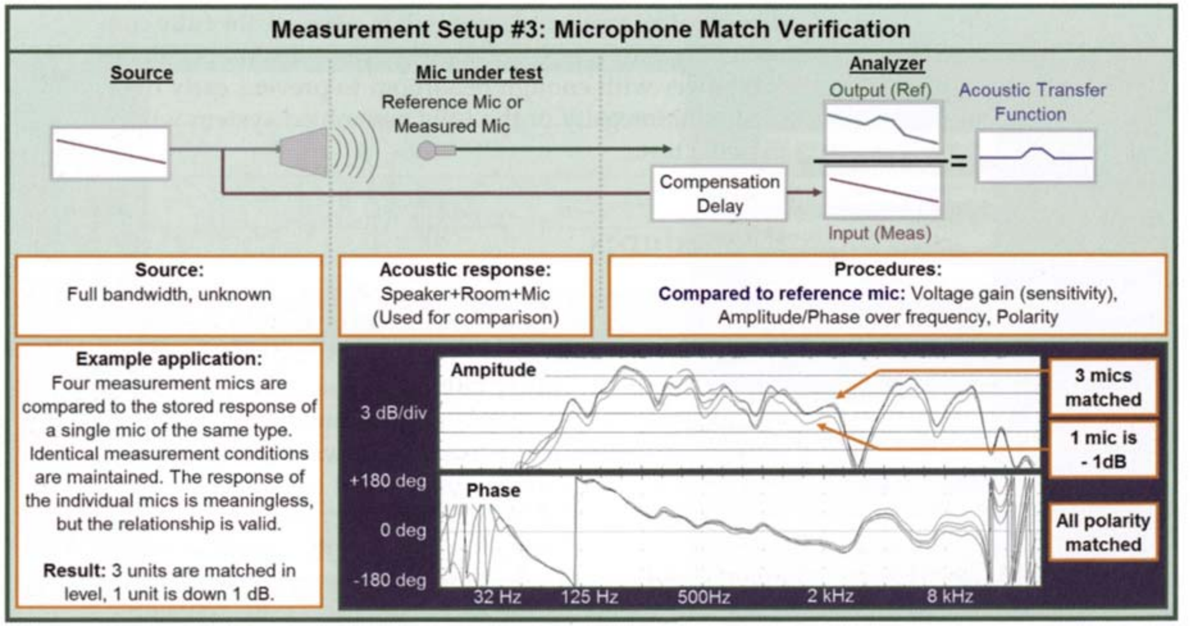


Figure 9.5 Verification test setup 3 setup flow block and example application for microphone level and frequency response matching. Consistent microphone placement is required for accurate results

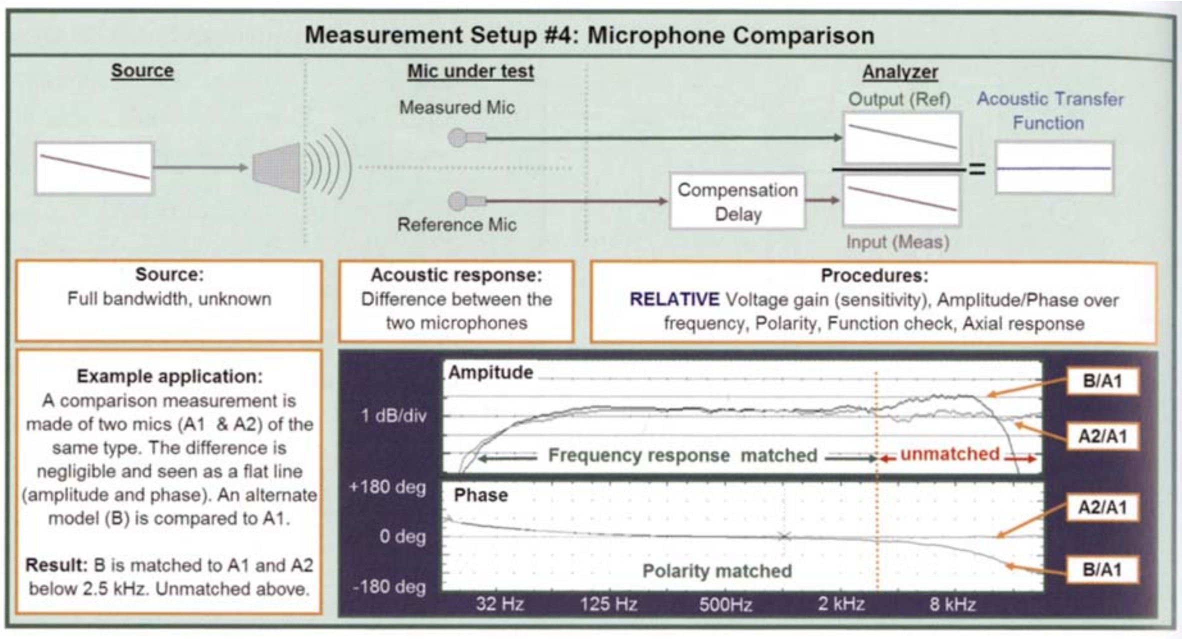


Figure 9.6 Verification test setup **4:** setup flow block and example application for microphone response comparison

Perspectives Even the most highly optimized analyzers on the market are still very complex tools. Start by making very simple measurements in a controlled environment. You need to be able to predict what the response will be when you make a measurement. Then change one parameter of the measurement and predict what the new result will be. Until you can repeatedly get the proper results in a controlled environment you do not have a hope of measuring a complex system in a large space.

Fred Gilpin

features of a particular device. It will not tell us, however, about how it will behave when interconnected to the devices at its input and output. It is, after all, the fully connected system which will need to drive the amplifiers to full power with enough headroom to prevent early overload. The integrity of the fully assembled system will be the final test.

Procedures

There are two basic forms of procedures: direct and comparative. Direct procedures find a result in absolute terms. An example of this would be a device with a maximum output capability of +24dBV. Comparative procedures evaluate the findings of direct procedures and quantify the difference. Carrying our example onward we might find that the *difference* in maximum output capability between two devices was 6dB. Any of the direct procedures can be carried forward to provide a comparative result against a given standard or a previously performed measurement. Therefore we will concentrate on the direct procedures with the understanding that comparative procedures can be structured as required on site.

Noise Over Frequency

No rock concert or sporting event would be complete without someone shouting "Let's make some noise!" Our sound system will need no such encouragement. The upper limit of our dynamic range would be the *noise* the performers are asking for. The noise we are concerned with creates the lower limit of our dynamic range. Hopefully these can be kept far apart. All devices create some noise. Our tests will determine the level over frequency of the noise floor.

The test for noise amounts to "Device under test 'unplugged'," but there can be some additional considerations. One is that the DUT level controls may have an effect on the outcome. If the DUT has both input and output level controls their relative positions are certain to affect this outcome. How the internal gain structure of the device will affect the noise will be detailed in the procedures which follow.

Perspectives *If one just turns on an entire complex system and starts to tweak an EQ to make it sound right, the root cause of the problem is usually missed. What I have found is that if I break the system down into smaller subsystems which can be easily verified, and then combine the subsystems in a controlled manner, I can isolate these kinds of interaction issues, and eliminate the element of magic in the system setup. This can be as simple as checking signal integrity and polarity at each stage of the signal path, for example source outputs, mixer/processing inputs, mixer/processing outputs, amp inputs, amp outputs, speaker inputs. Especially with DSP-based devices where there can be multiple gain stages and processing blocks hidden in the bowels of the software files, it is especially important to be able to trace the integrity of the signal from source to speaker.*

Dave Revel

There are two main types of noise in electronic devices: "hum" and "noise." Hum is a series of steady sine wave tones that are harmonic multiples of the line frequency. Any device has internal hum caused by its own power supply. Best results are obtained with either shorted inputs or using a low resistance terminator but hum can also accumulate in the interconnection process. Wiring schemes will have a strong effect on the strength and harmonic structure of the hum component. Wiring scheme hum is induced by differences in ground current, termed a "ground loop" or by electromagnetic interference (EMI).

The term **noise** refers to the random noise component that is generated inside the active electronic circuits. This is "white noise," the linear version of noise (equal energy per frequency). Audible white noise has half of its energy in the band from 10 kHz to 20 kHz. Therefore we first hear as "hiss" the presence of extreme high frequency random noise. Hiss can be minimized by proper gain structure management. However, we must be wary of hiss reduction schemes that also reduce the maximum output capability. We must ensure that we have sufficient dynamic range to reach full power.

There are a virtually unlimited number of noise sources in our acoustic world. The single channel response at the microphone will see the acoustic response with no input signal applied to the system. If the noise signal comes from our electronics then all of the factors just discussed apply If it does not come from our electronics then this would be a good time to use these important verification tools: eyes and ears.

Note: no conclusions regarding hum and noise can be made about an electronic device when something is plugged into its input. In such cases the noise at the output may contain noise presented at the input or induced by the interconnection. Many noise specifications published by manufacturers call for shorted inputs. The lowest noise readings are obtained with either shorted inputs or by using a low resistance input terminator.

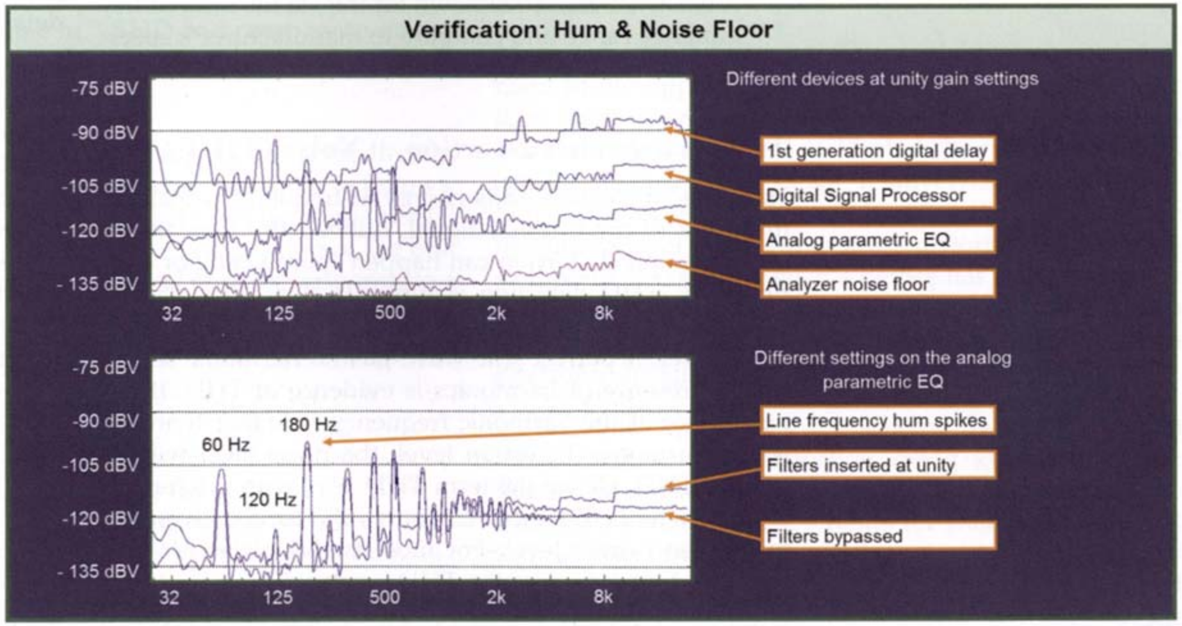


Figure **9.7** Example application of the hum and noise verification procedure

Test: Noise over Frequency

- **Subject:** Determine the level of the hum and noise floor of the DUT
- **Source:** None
- **DUT:** Electronic components, speakers, full signal chain, or entire networks
- **Units:** Volts (dBV) for electronic, dB SPL for acoustic
- **Setup:** #1 (single channel). Disconnect or short the device input. Device output to analyzer
- **Procedure:**
  1. Disconnect source from DUT input to remove the possibility of introducing hum or noise.
  2. Optimize analyzer input gains to the maximum level short of overload.
  3. Measure the output with a single channel FFT high-resolution spectrum. The optimal FFT window should be flat-top (for hum) or Hann (for random noise).
  4. Enable signal averaging, since the random noise component being measured will be variable over time.
  5. Read the outcome in voltage, dBV or dB SPL (the mic sensitivity must be known for this) at the desired frequency range and compare to manufacturer's specification or other threshold of acceptability.

Total Harmonic Distortion + Noise (THD + n)

Harmonic distortion adds energy at frequencies that are multiples of the original signal, dubbed the **fundamental.** Harmonic distortion can happen sporadically or continuously. THD detection is typically done in isolation. A single frequency is sent to the device at a designated drive level. A perfect sine wave has no harmonic series, so the presence of harmonics is evidence of THD. If the noise floor at the harmonic frequency is higher than the actual harmonic distortion level, the noise level will be read as THD. Hence the term THD + n, with n referring to noise. The THD measurement is valid for the given frequency and drive level. For most analog electronic and modern digital devices the THD levels will remain fairly constant over frequency and level. The THD measurement is very much a "source dependent measurement." Once again we return (very briefly) to the subject of resolution. In this case our concern is the purity of the sine wave source signal. It has distortion of its own, which present a lower limit to what we can measure. Before attributing THD to a device under test, we would be wise to test the THD of the generator directly. This would be another facet of our analyzer self-verification process. Most inexpensive sine wave oscillators are not suitable for THD testing because of the high distortion of the sine waves they generate. Many of them are no better than one half percent distortion.

Note: it is possible to receive misleading readings in THD + n measurements using FFT analyzers. Three factors are critical. First, we must optimize the gain structure of the analyzer to ensure that we are not measuring the analyzer noise floor. Second is the purity of the sine wave source signal as discussed above. And finally we must optimize the FFT window for a sine wave input signal to minimize the leakage component in the FFT computation. See the analyzer manufacturer's operating instructions for details.

Test: Total Harmonic Distortion + Noise

- **Subject:** The percentage of THD + n in the DUT
- **Source:** Sine wave
- **DUT:** Electronic components, speakers or full signal chain
- **Units:** Percentage of THD at specified level and frequency
- **Setup:** #1 (single-channel) generator to device input. Device output to analyzer
- **Procedure:**
  1. Drive DUT at desired level and frequency with sine wave source. 1 volt (O dB) @ 1 kHz is typical.
  2. Place the cursor on the fundamental frequency.
  3. Optimize analyzer input gains to minimize noise in the measurement.
  4. Measure the output with a single channel FFT high-resolution spectrum. FFT window should be

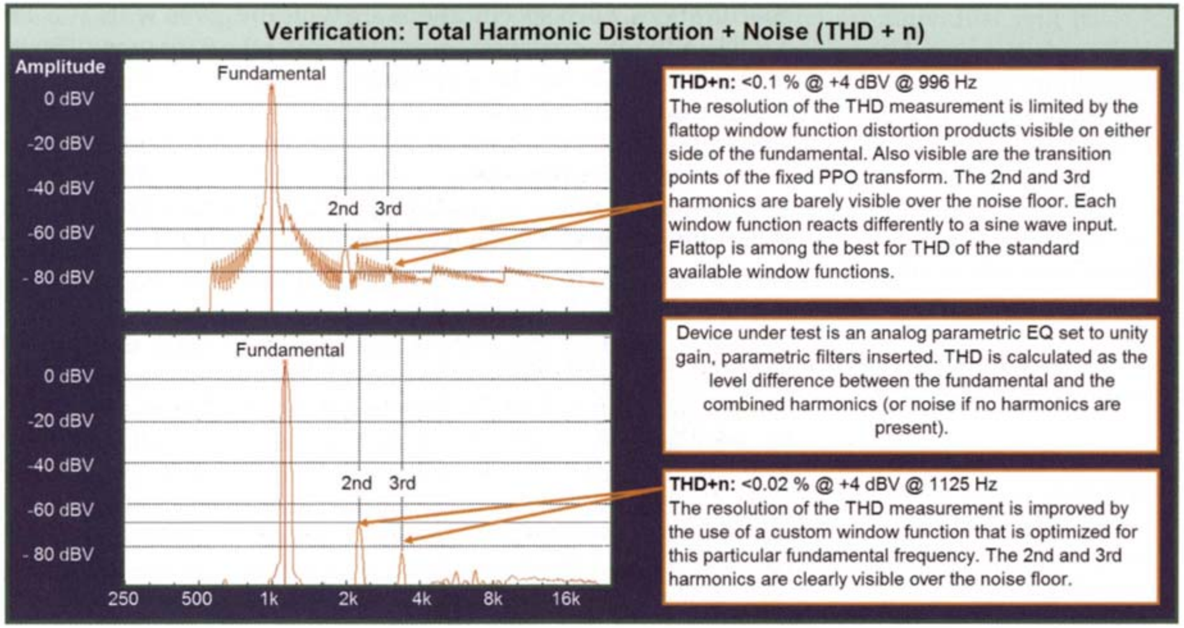


Figure 9 8 Example application of the THD + n verification procedure

flat-top or other window suitable for sine wave analysis.

5. Read the outcome in %THD and compare to manufacturer's specification or other threshold of acceptability.

Note: if the analyzer does not compute the THD directly, the value will have to be determined manually as follows. The distortion is the level offset between the fundamental and each harmonic. The THD is a compilation of the first nine harmonics. Each 20 dB of level offset corresponds to a decimal point change in distortion level for the individual components. As a quick reference the following will get us in the ballpark:

-20dB = 10 per cent

-40db= 1 per cent

-60dB = 0.1 percent

-80dB = 0.01 percent

Maximum Input/Output Capability Over Frequency

Maximum input/output (I/O) capability is a measurement of the upper limit of the dynamic range, the lower limit being the noise floor we measured earlier. A typical electronic device driven at levels below maximum will operate in their linear range, relatively free from distortion and compression. We will know we have reached the maximum level when either gross distortion or compression occurs. Assuming such a device has a flat frequency response, the maximum capability will not change much over frequency. This is a very straightforward test for an electronic device.

Not so for speakers. For speakers, the onset is much more gradual, with distortion rising gradually as the maximum is reached. For a speaker, the maximum SPL we find for a given frequency holds for *only* that frequency. Furthermore, the dB SPL readings for a given frequency will not likely correspond to the manufacturer's stated specifications. Such specifications are typically given with a full bandwidth excitation source and cover the full range of the device. Do not be surprised if a 130 dB SPL rated speaker is capable of only 110 dB SPL when driven with a sine wave. Additional cautions regarding this type of testing with loudspeakers includes the dangers of high SPLs to both listener and loudspeaker alike. The use of ear protection and speaker protection are required.

Now let's return to the electronic device. We have already seen how to identify harmonic distortion in small quantities. Finding the overload point is easy since the THD level abruptly rises as the device reaches clipping. The test amounts to "turn it up until it clips," but there can be some additional considerations. One is that the DUT level controls may have an effect on the outcome. If the DUT has both input and output level controls their relative positions are certain to affect the outcome. Much can be learned about the internal gain structure of a device in this regard.

Pseudo-Unity Gain

A rose is a rose is a rose is a rose.

Gertrude Stein, 1913

Unity gain is not unity gain is not unity gain.

606 McCarthy, 2006

If we measure a device and the output level is equal to the input, we can conclude that it is unity gain. But not all unity gains are created equal. The overall gain of a device is the composite of all of its gain stages. These stages may have internal boosts and cuts that offset each other either by design or by user adjustment of the input and output level controls. It can also happen with or without metering that accurately indicates the internal levels. Does it matter as long as things end up unity gain? Yes. It can change the maximum I/O capability and the level of the noise floor at the output. The internal gain structure of a device will determine how easily we can guide a signal with high dynamic range from input to output without clipping, excessive noise and ready to drive the next stage.

The dynamic range of the system as a whole is limited by its weakest link. If a device with a small dynamic range feeds one with a wide range, the latter device should have no trouble passing the signal through between its upper and lower limits. On the other hand, if a wide dynamic range signal is fed into a restricted range device, something has to give. We can't fit a 120 dB range of signal through a 100 dB device. Barring some form of compression and expansion (like Dolby™) we either lose 20 dB of headroom at the top, gain 20 dB of noise at the bottom or spread the loss between. We cannot fit six liters of water in a five-liter jar. Historically there has been such a dynamic range mismatch between the analog and digital audio domains. The gap between them is closing, but the analog world still enjoys a wider range. One of the places where we must be vigilant about our gain structure is at the transitions between these domains.

Three ways to unity gain from through a device:

1. Unity gain at the input, unity gain at the output: standard for devices with high dynamic range. This will result in the least amount of change in maximum in/out capability, and noise.
2. 20 dB gain at the input, 20 dB loss at the output: sometimes used for devices with low dynamic range. This will result in a 20 dB loss in maximum in/out capability, and minimal addition of noise.
3. 20 dB loss at the input, 20 dB gain at the output: not advisable in most cases. This will result in no loss in maximum in/out capability, and potentially add 20 dB of noise.

The first of these scenarios is unity gain. The second and third are pseudo-unity gain. The boost at the input method became popular in the era of 16- and 18-bit digital devices. The dynamic range difference between analog and digital devices often exceeded 20 dB. Since the digital devices could not span the full range, accommodation was made by removing the top 20 dB of the analog maximum capability. This kept the noise levels fairly constant during ambient conditions. During operation, the issue of the lost headroom could become serious, if the system as a whole is no longer able to drive the amplifiers to full power.

Gain structure management is in many ways an operational issue. We can aid the process by finding the upper and lower range limits of each of the components and their interconnection as a whole. This will provide the highest assurance that we can pass the widest range signal through the system from start to finish.

Internal gain structure that rises and falls is not reserved to any particular form of device. It is present any time there are input AND output controls in a device, whether analog or digital, new or old. A mixing console can achieve unity gain from input to output with more opportunities for internal up and down than a roller coaster. In the end, the goal is to preserve the dynamic capability of the system. This is done by minimizing the amount of rise and fall in internal and interconnection gain structures that inevitably lead us back to the level we started. We can see the implications of all this by testing the dynamic range in different configurations. How much level can we get out? How much noise do we have? Now raise the level here and drop it there. Try again. Is it better or worse?

The value of this test on an individual unit basis has been shown. An additional consideration is the value of this test on the complete signal chain to the amplification system. In the end we will deliver a dynamic signal to the amplifiers. Do we have enough headroom to bring the amplifiers to full power? Is there too much noise?

Note: the purity of the sine tone is not so critical as in the THD shown previously. When the system reaches clipping, the rise in THD is unmistakable.

Test: Maximum I/O Capability **over** Frequency

- Subject: Determine the maximum transmission level through the DUT
- **Source:** Sine wave
- **DUT:** Electronic components, speakers, full signal chain, or entire networks
- **Units:** Volts (dBv) for electronic, dB SPL for acoustic
- **Setup:** (Single channel) generator to device input. Device output to analyzer

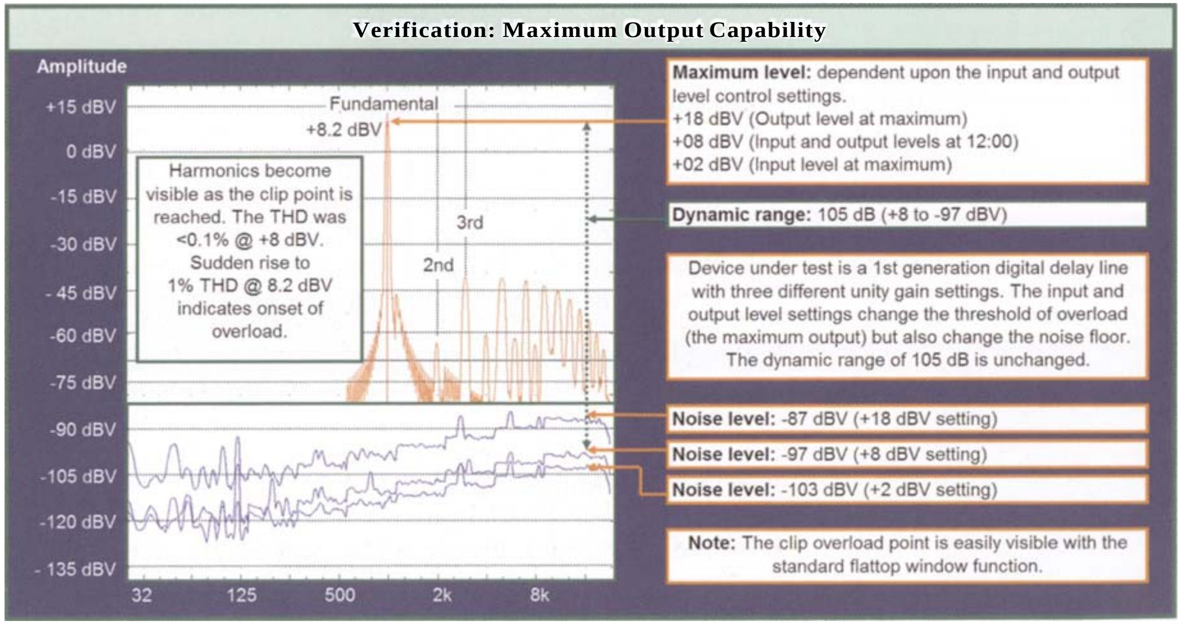


Figure 9.9 Example application of the maximum input/output capability verification procedure

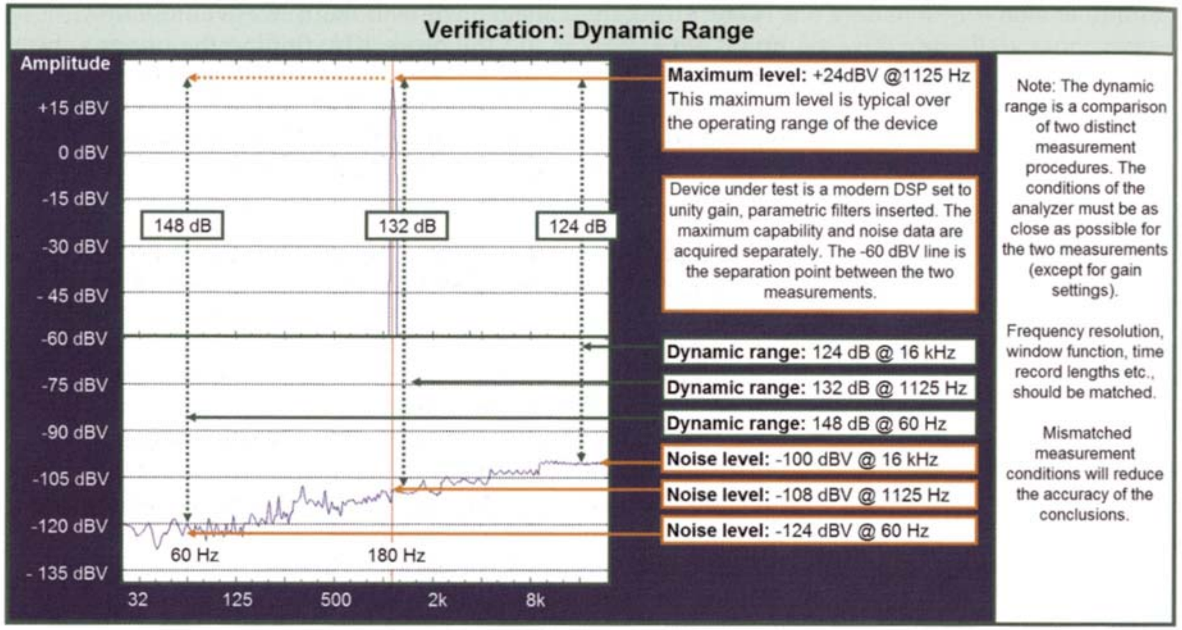


Figure 9.10 Example application of the dynamic range verification procedure

Perspectives Headroom can make or break a system. When you are installing a system from scratch or doing a full system checkout and optimization you will maximize the headroom as part of the process. When you are touring and using installed systems or rental systems from vendors you don't know, you need to make sure there is enough headroom for your event. Without it, the event may sound quite good until the system gets pushed and the system starts clipping the peaks. Not a good place to find oneself. For me 15 dB (average to peak) is the minimum acceptable.

Alexander Yuill-Thornton II (Thorny)

- **Procedure:**
  1. Drive DUT at desired frequency with sine wave source.
  2. Place the cursor on the drive frequency.
  3. Optimize analyzer input gains to prevent overload in the measurement.
  4. Measure the output with a single channel FFT high-resolution spectrum. FFT window should be flat-top or other window suitable for sine wave analysis.
  5. Raise the drive level until distortion reaches unacceptable levels or compression clamps the output level.
  6. Read the outcome in voltage or dB SPL and compare to manufacturer's specification or other threshold of acceptability.

Test: Dynamic Range

- **Subject:** Determine the dynamic range of the DUT
- **Procedure:** Perform the noise floor and maximum in/out capability procedures. The difference between them (in dB) is the dynamic range.

Latency

In Chapter 1 we discussed the issues of unlabeled latency in digital devices. None of them, of course, are really zero, since there is latency delay in the A/D converters and perhaps added internally or in network interfaces. The labeling standard for such devices is to indicate the user selected delay that we are adding to the latency. This is similar to the way our industry sells tickets. A $40 ticket is not really $40.00. It will be $40.00 + service fee + handling charge + tax. We can't buy the ticket without the fees any more than we get our delay without the latency. We need to know the bottom line number for several reasons: to know how long we will have to wait before a signal leaves our speakers, and to anticipate any opportunity where related signals with different latencies might be summed.

There are two basic families of digital audio devices: component and composite. The former is dedicated to a single task such as delay, equalization or spectral crossover setting, while the latter does all of these and more.

These can be run in analog series or parallel, or remain digital via network connections. Component devices tend to be more consistent in their latency, i.e. they will pick one latency and stick to it. Composite devices, especially those with open topology (i.e. customizable DSP units that can be configured in virtually any form), are subject to change without notice.

The following are real examples of what we may find in our system when every digital audio device reads "Oms" delay:

1. Component or composite models from the same company with different latencies.
2. Component or composite models from different companies with different latencies.
3. Composite device with the same latency for each channel.
4. The exact same device with different latency for each channel (the default setting unless we know where to find the override function in the software).
5. Composite device that gives a different latency every time its settings are compiled. Could change for each channel separately or **all** channels together.
6. Composite device that changes latency when certain "zero phase shift" spectral divider filter types are used.
7. Networked composite devices that give different latency for signals that travel the network from those that stay inside the same unit.
8. Signal paths that travel through multiple devices in series (Oms + Oms does *not* equal Oms).

We have to be ready for anything. We must be sure to measure the entire signal path. Relative delay settings that are viewed during the upcoming alignment phase must incorporate all latency accumulations.

Test: Latency

- Subject: The transit time through the DUT
- Source: Independent (noise or music)
- DUT: Electronic components, speakers, full signal chain, or entire networks

Perspectives When it comes down to verification and alignment, the process following console check that gives me the best results is time and amplitude first, equalization last. 1 found that the better the system design, the easier it is to align the system properly.

Todd Meier

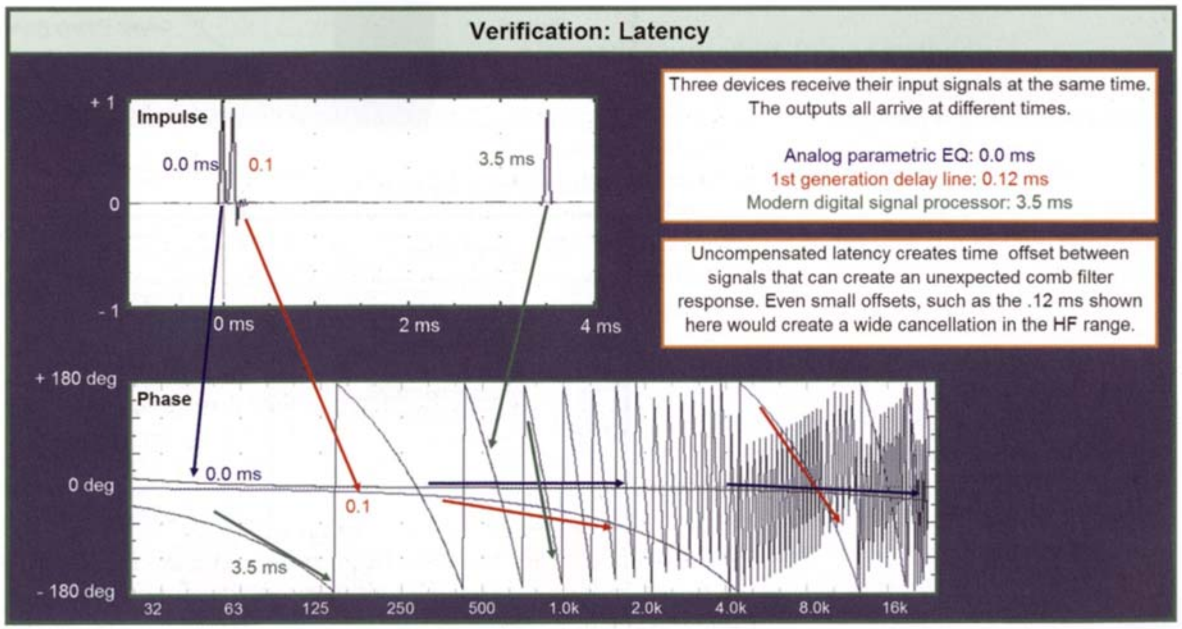


figure 9.11 Example application of the latency verification procedure

- **Units:** ms
- **Acceptable level:** Principal concern is that all devices are matched or at the very least known.
- **Setup:** #2, transfer function. Source to device input and analyzer (as input channel). Device output to analyzer (as output channel).
- **Procedure:**
  1. Drive DUT at any level with the source.
  2. Measure the impulse response.
  3. The latency is the amount of delay in the DUT.
  4. Record or normalize latency values as required to establish a zero time baseline.

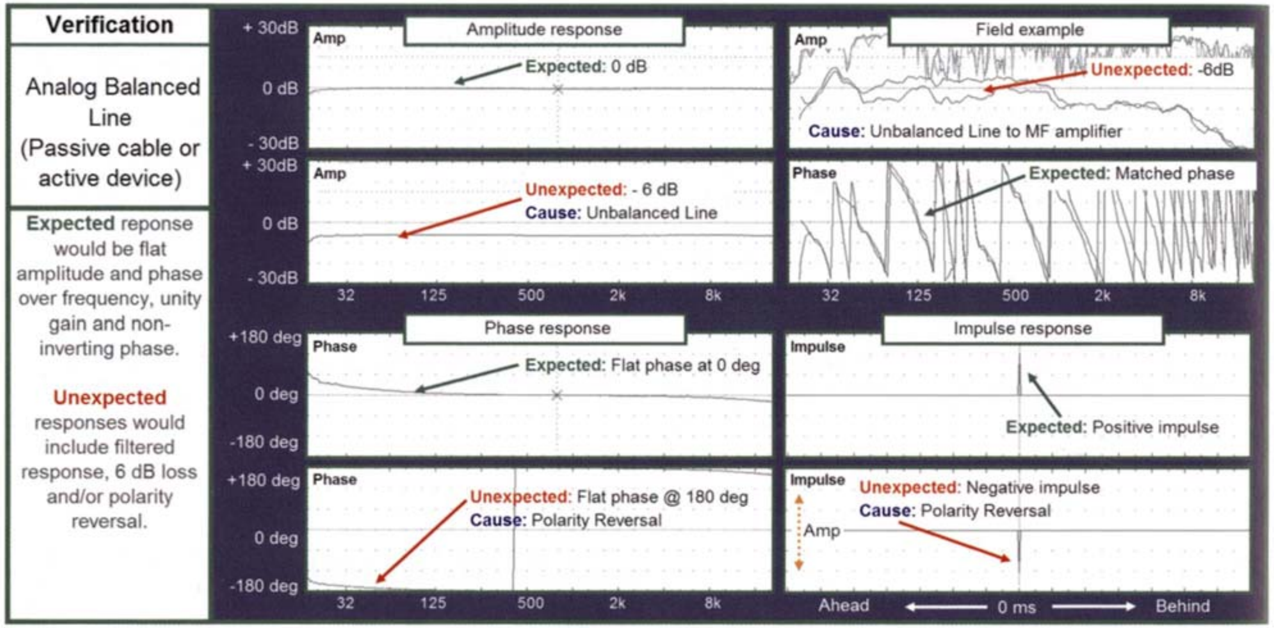
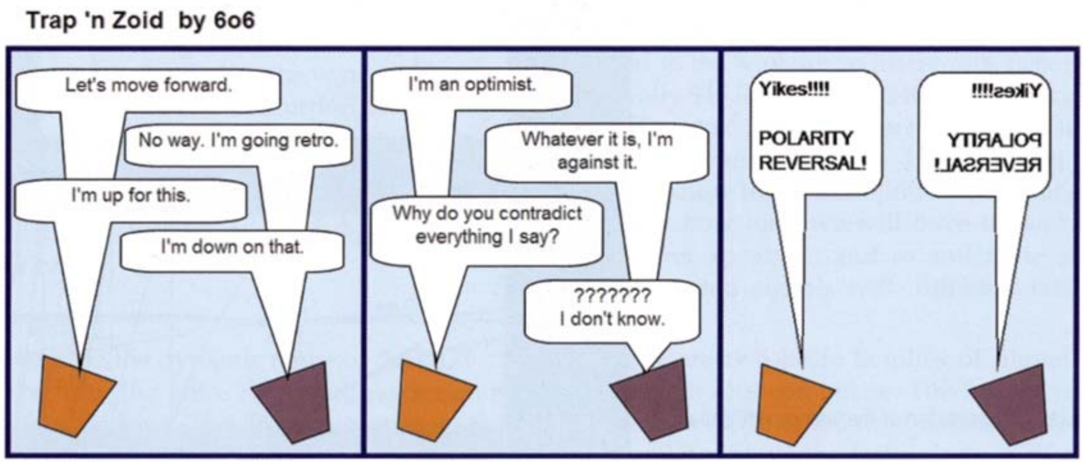


Figure 9.12 Example application of the electronic device polarity and level verification procedures



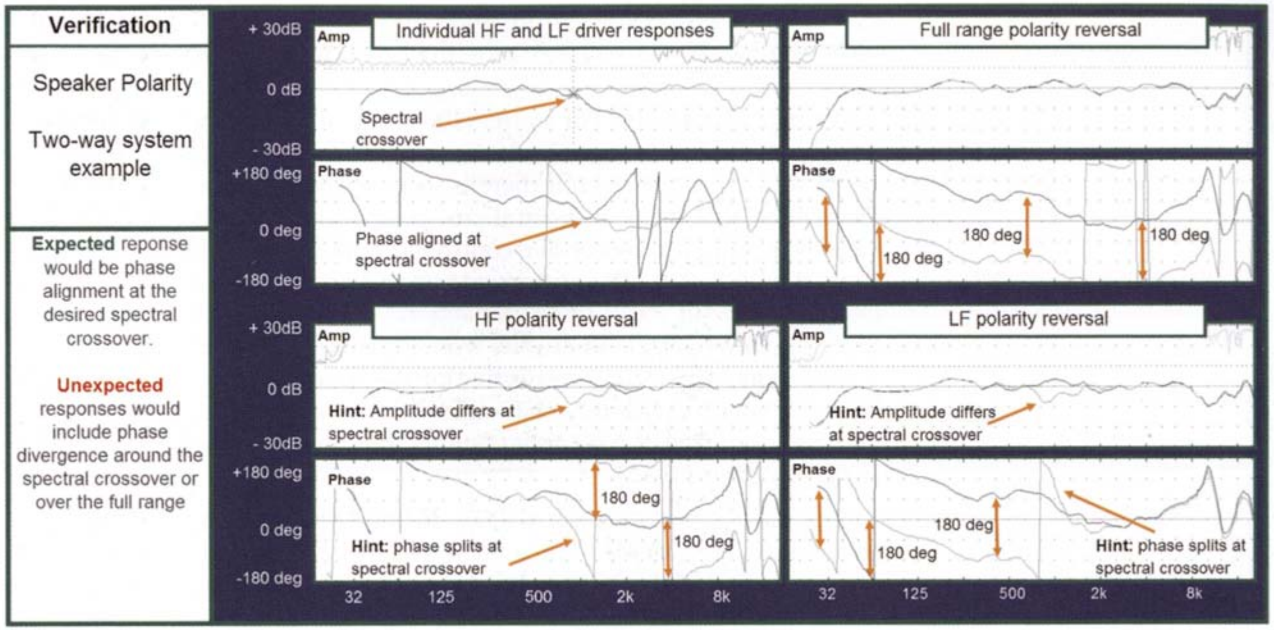


Figure 9.13 Example application of the speaker polarity verification procedure

Perspectives With technologies like SIM and SMAART I'm able to test all aspects of my installations, including the wiring, much quicker and accurately than I used to be able to using a continuity and volt meter alone.

Bob Maske

Polarity

Test: Polarity

- Subject: The polarity of the DUT
- Source: Source-independent (noise or music)
- DUT: Electronic components, speakers or full signal chain
- Units: Normal (non-inverted) or reverse (inverted)
- Acceptable outcome: Non-inverting, unless specified otherwise
- Setup: #2, transfer function. Source to device input and analyzer (as input channel). Device output to analyzer (as output channel).
- Procedure (impulse method):
  1. Drive DUT at any level with the source.
  2. Measure the impulse response (linear type).
  3. If the impulse peak is positive the device is normal (non-inverted). If negative, the device is reverse polarity (inverted).
- Alternative procedure (phase method):
  1. Drive DUT at any level with the source.
  2. Align the internal compensation delay for latency (electronic) or propagation delay (acoustic).
  3. Measure frequency response phase.
  4. Observe the position of the phase trace. If the phase trace is level around 0 degrees the device is normal (non-inverted). If the trace levels around 180 degrees, the device is reverse polarity (inverted).

Frequency Response

The electronic version of this test is very straightforward, so much so that we will immediately move on to the acoustic version. Unless we have an anechoic chamber available to us we will have a difficult time making any conclusive measurement as to the exact nature of the frequency response of a speaker system. There are so many opportunities for ripple variance that making conclusive statements must be done with great caution. Even extremely close measurements will have the floor and more in the data.

What can we verify? For starters the range of the device is determined fairly reliably. We can discern the trends of HF and LF rolloff even if there is local ripple variance. We will be able to see a spectral crossover, though not as clearly as can be done under controlled anechoic conditions. The search for the optimized spectral crossover is a good example of TANSTAAFL in action. As we get closer to the speaker, we gain immunity from the ripple variance. At the same time, our perspective of the spectral crossover becomes increasingly near-sighted and we can make crossover alignment recommendations that will serve us poorly in the far field.

Our most effective work will be in the form of comparison. As long as we reduce the question to one of difference, rather than absolute value, we have leveled the verification playing field.

Comparison verification measurement examples:

- Do speakers of the same model have matched polarity and drive level for each of component drivers?
- Do symmetrically matched speakers have symmetrically matched response?
- Are two speaker models phase-compatible over their shared range?
- Are two speakers phase-compatible at the spectral crossover?

Test: Amplitude **Response over** Frequency

- Subject: The range limits and amount of level variance vs. frequency of the DUT
- Source: Source-independent (noise or music)
- **DUT:** Electronic components, speakers or full signal chain
- **Units:** dB, frequency
- **Setup:** #2, transfer function. Source to device input and analyzer (as input channel). Device output to analyzer (as output channel).

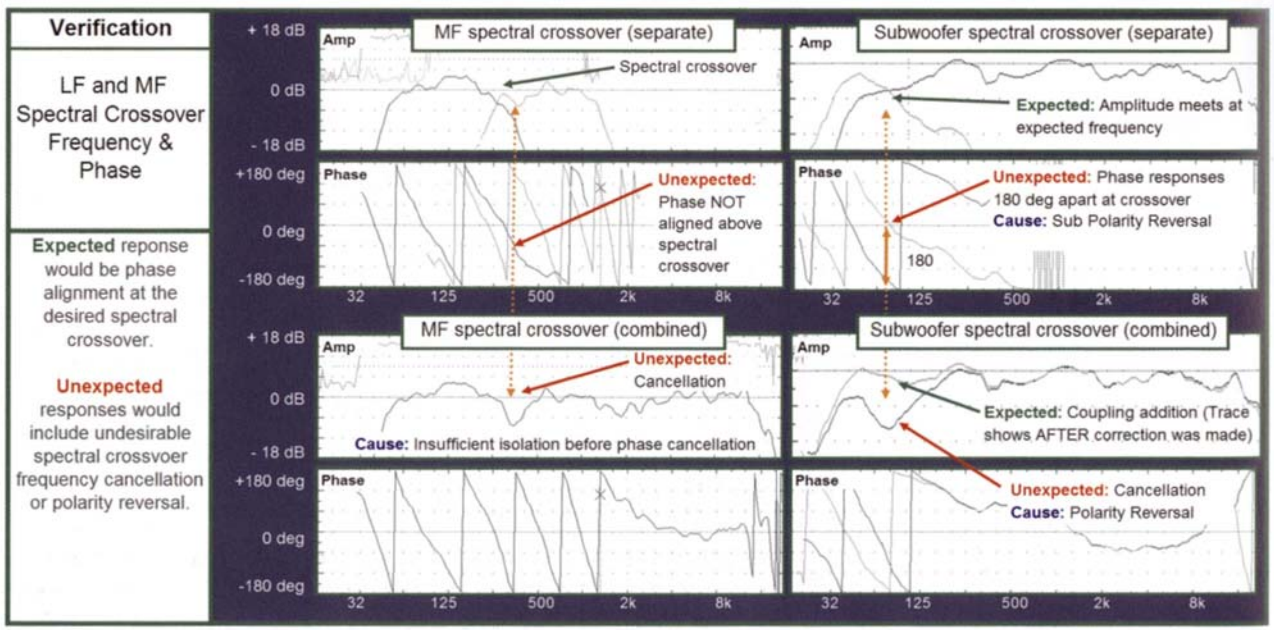


Figure 9.14 Example application of the amplitude and phase response verification procedures to the evaluation of spectral crossovers

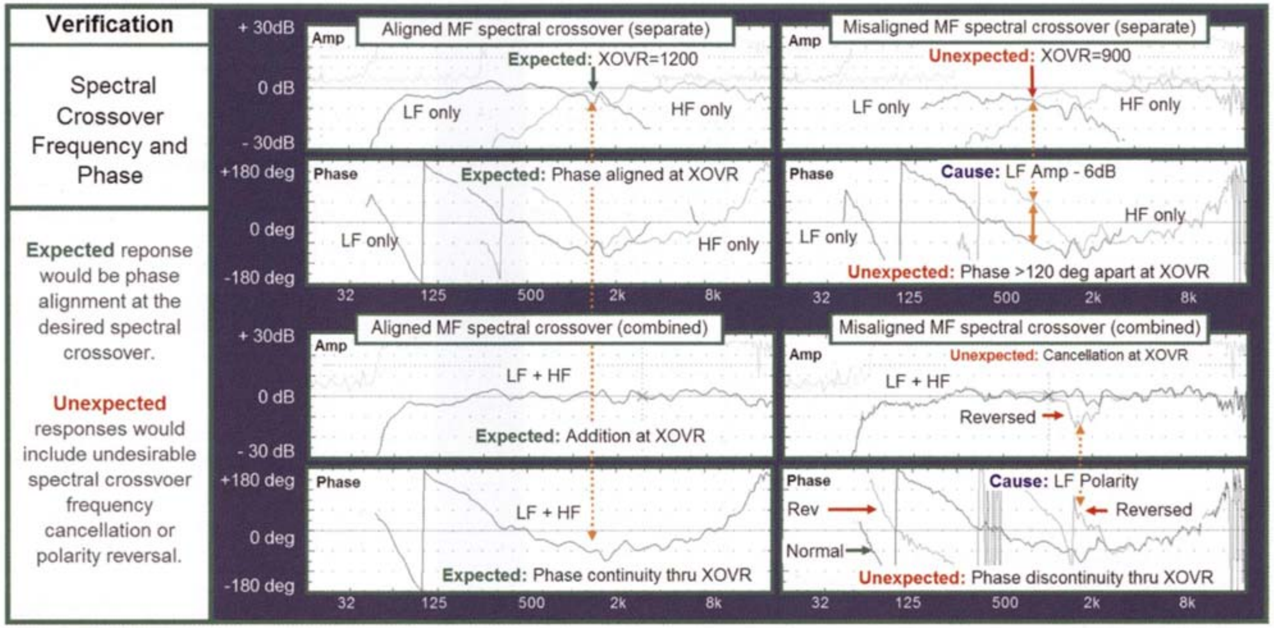


Figure 9.15 Example application of the amplitude and phase response verification procedures to the evaluation of spectral crossovers

- Procedure (range **limits):**
  1. Drive DUT at any level with the source.
  2. Measure transfer function frequency response amplitude.
  3. Position trace to nominal level for reference.
  4. Place cursor at frequency (high and low) where —3 dB (electronic), —6 dB (acoustic) downpoint is reached.
  5. These are the frequency range limits of the DUT.
- Procedure (variance):
  1. Drive DUT at any level with the source.
  2. Measure transfer function frequency response amplitude.
  3. Position trace to nominal level for reference.
  4. Place cursor at the frequencies inside the pass band with the greatest deviation from the nominal level.
  5. This is the level variance (plus/minus) for the DUT.

Note: for speaker measurements such data is inclusive of the ripple variance from the speaker/room and (potentially) speaker/speaker summation. Conclusions about the speaker system alone are limited by this. Conclusions in regard to the performance of the system in the room can be made, although such work is typically considered part of the calibration scope.

Phase Response over Frequency

Once again the cautions regarding acoustical measurements apply. There is an additional fine point in regard to electronic devices: If the latency of the DUT falls somewhere between the compensation delay increments there will be a phase delay remainder in the measurement. This is a time-resolution issue that is a similar situation to the bandwidth resolution (see Fig. 8.4) discussed in the last chapter. For example, if the analyzer compensation delay is limited to 20 microsecond (0.02ms, 20(μs) increments, then what happens when the DUT has a latency that falls at the mid-point between increments, such as 10 μ**S**? The analyzer sees the 10ms of phase delay and charts the response accordingly. This is 72 degrees of phase shift at 20kHz, so

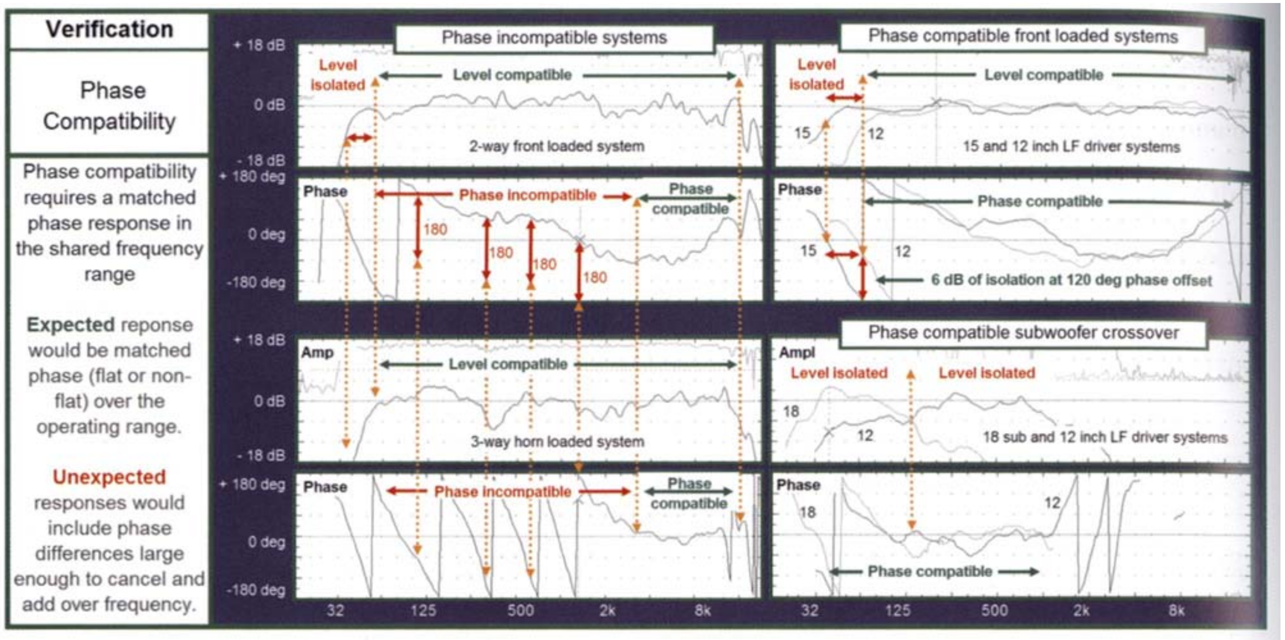


Figure 9 16 Example application of the amplitude and phase response verification procedures to the evaluation of speaker system range and compatibility

it is not likely to go unnoticed. What can we do about it? Some analyzers allow for sub-incremental correction, some don't. Most important, we can be mindful not to conclude that the DUT has phase delay amounts smaller than that which falls within our time resolution window.

Phase shift resolution for a given time compensation increment:

- 20 µ**S**: ±36 degrees @ 10kHz, ±72 degrees @ 20kHz
- 10 µs: ±18 degrees @ 10kHz, ±36 degrees @ 20kHz

Test: Phase Response over Frequency
- Subject: The amount of phase variance over frequency in the DUT
- Source: Source-independent (noise or music)
- **DUT:** Electronic components, speakers or full signal chain
- Units: Degrees
- Setup: #2, transfer function. Source to device input and analyzer (as input channel). Device output to analyzer (as output channel).
- Procedure (variance):
  1. Drive DUT at any level with the source.
  2. Measure transfer function frequency response amplitude.
  3. Position trace to nominal level for reference.
  4. Place cursor at the frequencies inside the passband with the greatest deviation from the nominal level.
  5. This is the phase variance (plus/minus) for the DUT.
  6. A given span can be converted to phase delay using the formulas and techniques shown in the previous chapter.

Note: for speaker measurements such data is inclusive of speaker/room summation. Conclusions about the phase delay in the speaker system alone are severely limited by this.

Compression

Test: Compression (Voltage gain vs. frequency)
- Subject: The threshold of compression in the DUT

Perspectives One of the things I always do when walking into an unknown PA system/rental company is to check every box on the ground. This will save tons of time when you are looking for an odd response from one of the clusters.

Miguel Lourtie

- Source: Source-independent (noise or music)
- DUT: Electronic components, speakers or full signal chain
- Units: Expected (meets specifications or display parameters) or unexpected
- Setup: #2, transfer function. Source to device input and analyzer (as input channel). Device output to analyzer (as output channel).
- Procedure:
  1. Drive DUT at any level with the source.
  2. Optimize analyzer input gains to prevent overload in the measurement.
  3. Measure the transfer function level.
  4. Raise the drive level until compression clamps the output level. This will be seen as a voltage gain change in the transfer function response.
  5. Read the change in voltage gain and compare to the readings on the display of the DUT.

Note: the source signal can alternatively be a sine wave (in this case use test setup 1). If the DUT is a band limiter (broadcast style) or a vocal processor this will be the preferable source to see independent threshold action.

Microphone Verification

In some cases we will have to create special setup configurations to find the answer we seek. One such case is the response of a microphone. In order to characterize a microphone response we will need a known flat acoustic source, which would require a known flat microphone to verify its response. This is a circular argument that we rely on the Bureau of Standards to provide the final say about. For this reason, our purchase of a measurement mic is a choice that we will make very carefully, based upon the credibility of the manufacturers.

For our practical application it is vital that we have microphones which are matched to each other, and that maintain their response over time. These are both verification issues. The verification of constancy will require reproducible conditions. Each month, for example, the mic is placed under the same conditions and measured, with the

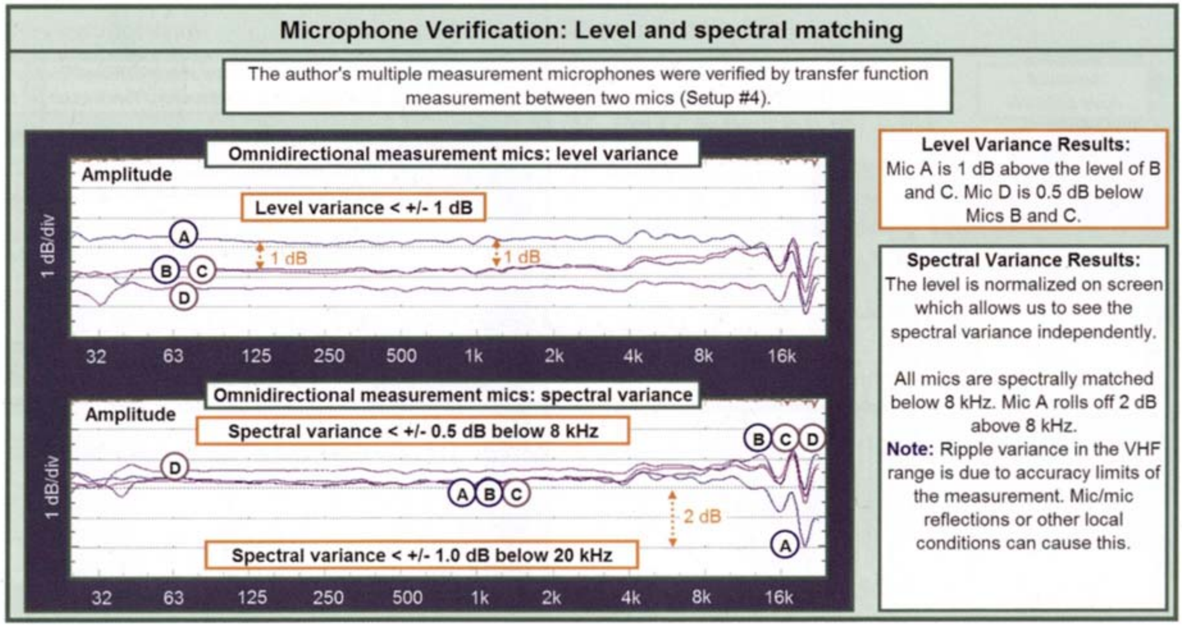


Figure 9.17 Field example of microphone matching transfer function

results compared to the previous. Sets of multiple measurement microphones, such as those we will commonly use for calibration, must be matched in order for us to discern variance in the system response. This verification can be done by a comparison of microphones to each other, a dual acoustic transfer function as shown in Fig. 9.5.

Microphone Match

Test: Microphone match verification

- Subject: The difference in level, polarity and frequency response of the DUT (mic) and another "reference" mic
- Source: Source-independent (noise or music)
- DUT: Microphone
- Units: Expected (meets specifications) or unexpected
- Acceptable outcome: device-specific
- Setup: Test setup #3, transfer function. Source to speaker and to analyzer (as input channel). Measured microphone (DUT) to the analyzer (as output channel). This is a serial verification. The response is stored to become the reference and then the mic is replaced with the alternate mic to be tested.
- Procedure:
  1. Drive DUT at any level with the source.
  2. Optimize analyzer input gains to prevent overload in the measurement.
  3. Measure the transfer function response.
  4. Compensate for any time offset between the arrivals.
  5. Store and recall the data.
  6. Carefully replace the first mic with a second mic (as close to the same position as possible) and obtain a new response. This mic is the new DUT.
  7. The level, polarity and frequency response deviations from the stored traces are the difference between the mics.

Microphone Response

Test: Microphone response verification

- Subject: The difference in level, polarity and frequency response, axial response of the DUT (mic) from a standard microphone reference
- Source: Source-independent (noise or music)
- DUT: Microphone

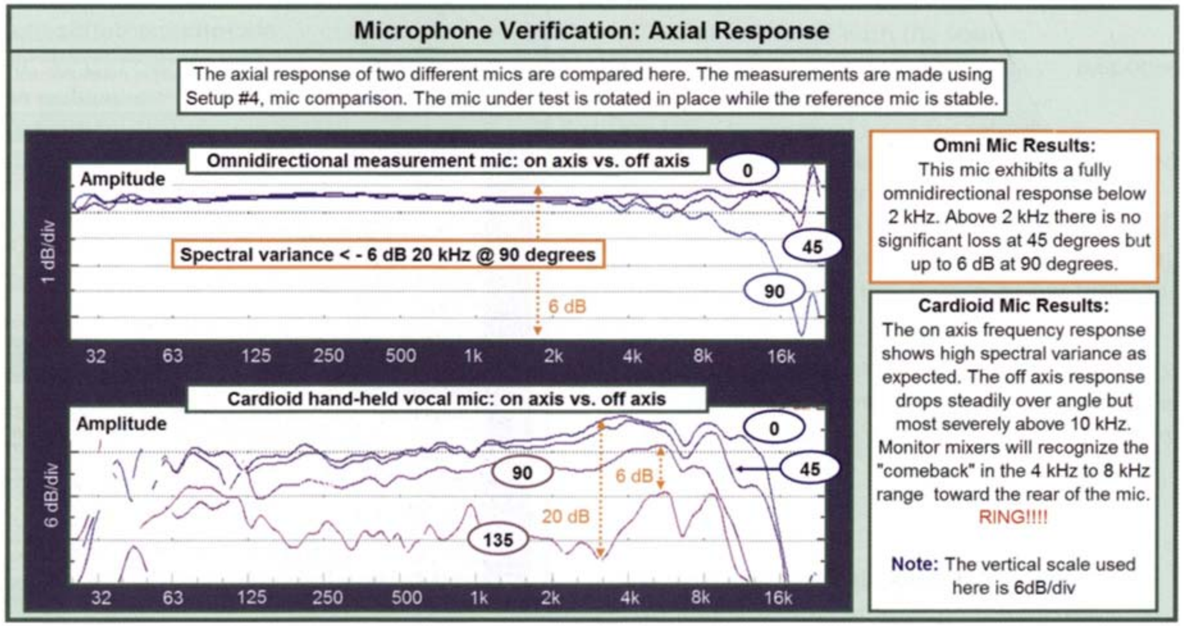


Figure 9.18 Field example of on and off axis response of microphones

Perspectives I was explaining the benefits of employing a sound alignment technician to a potential client and his response was "Why would I want to pay for that? "

d (Mack) mcbryde

- Units: Expected (meets specifications) or unexpected
- Setup: Test setup #4, dual microphone transfer function. Source to speaker. Microphones are placed in matched acoustic orientation to the speaker. Reference microphone to the analyzer (as input channel). Measured microphone (DUT) to the analyzer (as output channel).
- Procedure:
  1. Drive DUT at any level with the source.
  2. Optimize analyzer input gains to prevent overload in the measurement.
  3. Measure the transfer function response.
  4. Compensate for any time offset between the arrivals.
  5. The level, polarity and frequency response deviations are the difference between the mics.

Note: source speakers with a wide coverage pattern (first-order) are preferable, as this reduces the probability that the mics under test do not share the same sound field. Place the mics as close as possible to the speaker to reduce the room/speaker summation effects. If the source speaker is a two-way device the microphones must be placed so that they are not oriented between the drivers. For on-axis characterization, both mics are placed on axis to the source. To characterize the axial response of a mic, the mic under test is rotated in place, keeping its diaphragm in approximately the same plane. This is no substitute for an anechoic chamber research facility, but has real practical value. The precision of this process is limited by reflections between the microphones, and small differences in path lengths of local reflections. Either of these can cause frequency response ripple. The axial trends are, however, clearly evident from these measurements such as the overall spectral tilt in the off axis response of an omnidirectional mic. The axial analysis of a cardioid microphone can be used to find the best placement angle for maximum rejection of a stage monitor.

Post-Calibration Verification

The final verification role comes after the calibration process has been completed. This verification may be part of ongoing maintenance or the checking of "copy and paste" calibration settings, and the symmetric speakers they feed.

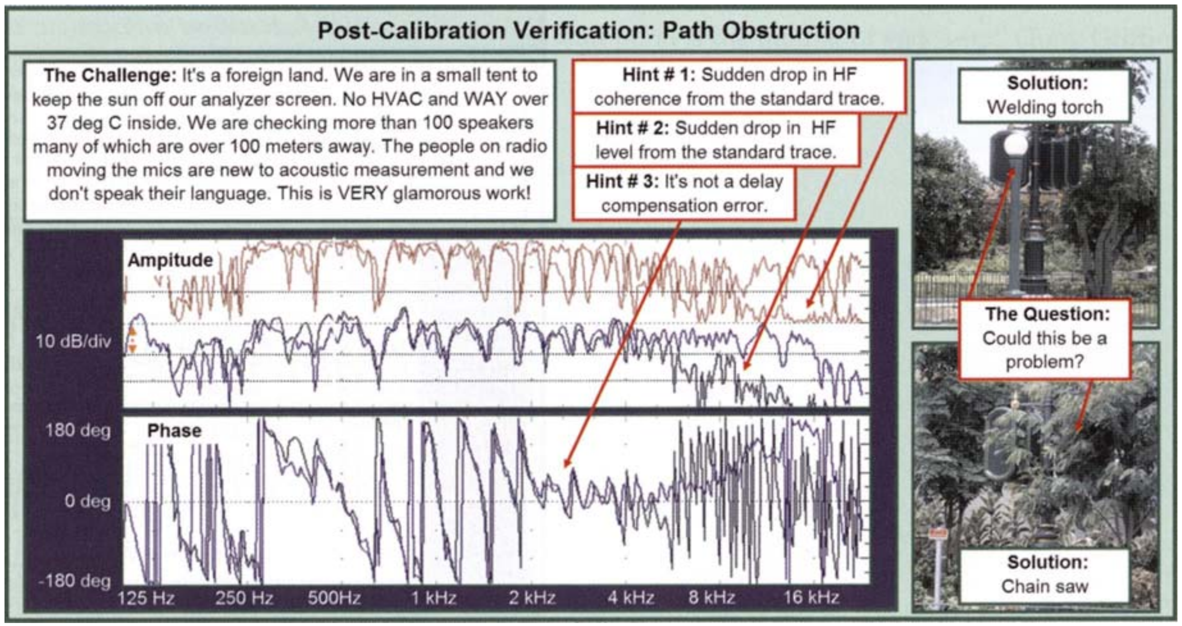


Figure 9 19 Example application of the post-calibration verification

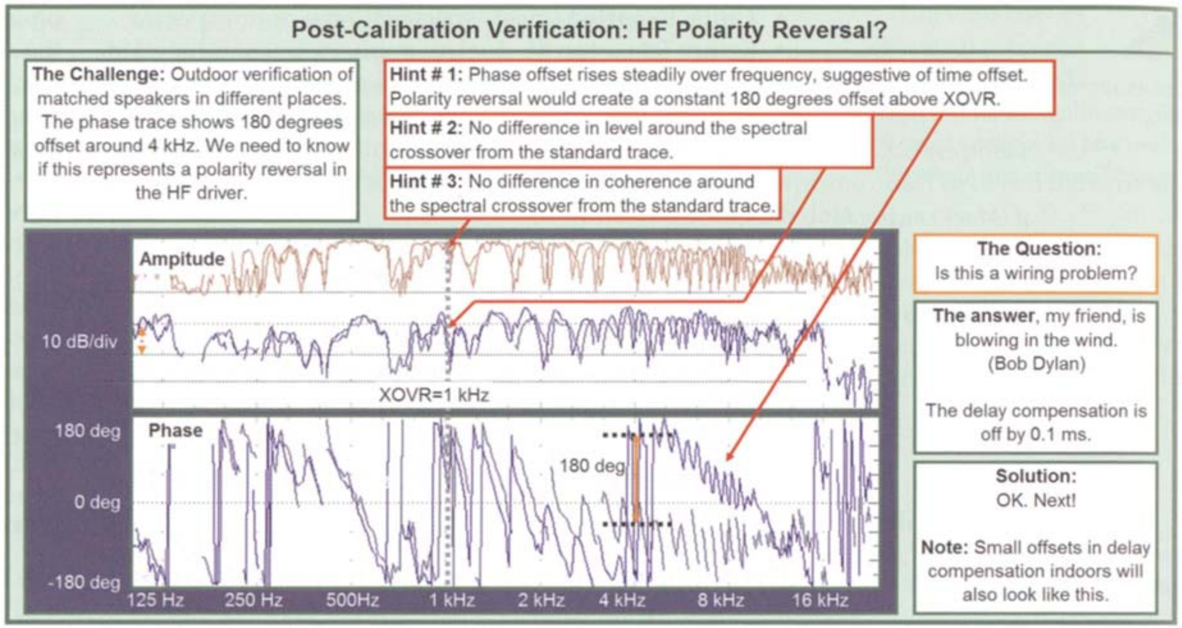


Figure 9.20 Example application of the post-calibration verificatic

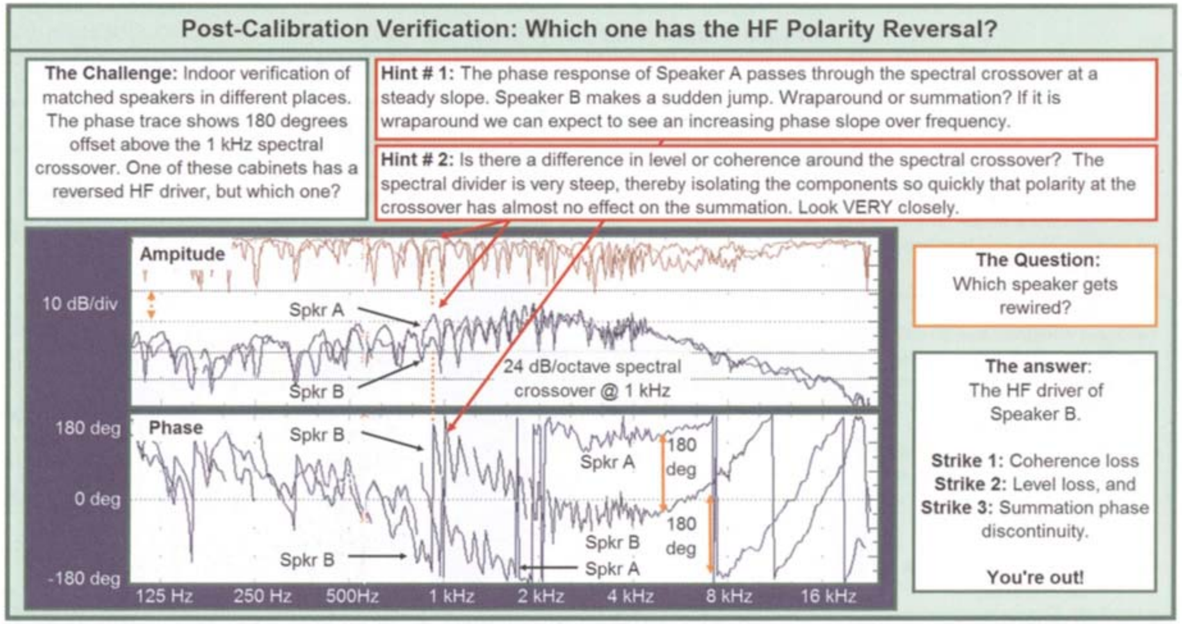


Figure 9.21 Example application of the post-calibration verification

Post-calibration verification measurement examples:

- Do symmetrically matched speakers have symmetrically matched response?
- Were copied processor parameters input correctly and did the processor correctly reproduce them?
- Was this speaker or microphone damaged during the course of the tour?
- Is the response of the system the same as it was 6 months ago?

Most of the work of post-calibration verification consists of "OK, next!" as we move through 16 surround speakers, but the work has its reward when we find the one that falls out of line.

Additional Considerations

There is no limit to the amount of verification that can be performed on the system. The output of any device or combination of devices can be analyzed. The difference between any two devices can be found as well. Get creative!

For ongoing operations such as a touring system, it is not practical or necessary to perform a full set of verifications each night. Standard touring systems are predicated on the reduction of open variables. If the system has only been reconnected, rather than rewired, since the last stop, we can move safely forward. If unexpected results are found in the calibration stage, then we may need to retreat to verification procedures to locate the problem.

A verified system is ready for the final stage of preparation for operation: calibration. A thorough verification stage provides the solid foundation required for the critical decisions that comprise the calibration process. It is tempting to look at the available time and resources and conclude that time would be better spent in the calibration stage rather than verification. Do so at your peril.

In the end, we all come to a level of verification that we are comfortable with. This is application-dependent, with permanent installations requiring the highest verification levels. There is also a relationship aspect. Any work done with a new client must be scrutinized at the highest level. Experienced clients will have done a thorough pre-verification in advance and the process moves along very quickly. Even so, it cannot be skipped.

During my first day of work in professional audio I was told something that I have never forgotten:

"Assumption is the mother of #@$%-up" (Tony Griffin, 1976).

10

Calibration

> ***calibrate*** *v.t. find calibre of; calculate irregularities of before graduating it; graduate with allowance for irregularities; correlate readings (of instrument etc.) with a standard*
>
> Concise Oxford Dictionary

The stage is set to complete the optimization process. The system has been verified such that all of the components are known to be in working order. Speaker positions have been roughed in and we are prepared to input signal processor parameters.

The calibration process proceeds from simple to complex. Each of the verified speaker subsystems is given initial alignment settings. These subsystems are then combined into larger subsystems, with secondary adjustments being added to compensate for the effects of summation. The process is complete when all related subsystems have been grouped into combined systems with minimum response variance throughout the listening space.

Like the verification process just completed, the calibration process is made up of a series of controlled test procedures. Each procedure is designed to give specific answers such as speaker position, delay time, equalization and level setting. These alignment answers are never as cut and dry as those at the verification stage. Attempting to create an exact step-by-step set of procedures that will work for all system designs would be futile. Each sound design has unique combinations of speakers, processing and room acoustics. For the alignment process, the engineer must have a play book ready to adapt on site for the hundreds of contingencies that will arise with the particular job.

This does not mean that the choices of how to implement the alignment settings are arbitrary. Quite the contrary. The approaches shown in this chapter contain more than simply a listing of the author's proven procedures. The rationale behind each procedure is revealed so that the alignment principles can be carried out even when the particulars of the system are not an exact match. These are flexible guidelines to get a result.

Calibration Defined

Calibration is the process of fine-tuning a system for the particular parameters of the situation. Calibration, sometimes called alignment, is concerned with the compensation of parameters that are far beyond the current predictive capability. These must be measured on site and actions taken as required. An illustrative example shows the distinction between calibration and the previous step: the verification process reveals that the equalizer is properly wired. It would not be reasonable to expect, however, that the optimal equalization curve for this venue has been programmed into the equalizer in anticipation of our requirements. That is to be done in the calibration process. The gray line here lies in the placement of speakers. The design and installation processes put the speakers in place.

The verification stage confirms the installation, but it will be the calibration process which will have the final say. The calibration process will determine whether this is indeed the best position and focus angle for the venue. Therefore, some degree of flexibility in focus angle is essential to the calibration process.

Goals

The goals of system calibration have been well established throughout this book. They are neither unique nor original to this text nor relevant only to those using dual-channel FFT analysis. These are fairly universal.

System calibration goals:

- minimum level, spectral and ripple variance over the listening area (sounds the same everywhere)
- maximum coherence (intelligibility, direct/reverberant ratio, clarity)
- maximum power capability (sufficiently loud)
- sonic image control (the sound appears where we want it to appear).

Challenges

The challenges to system calibration are also universal. We must evenly distribute direct sound over the space. We seek to minimize the adverse effects of summation, while taking maximum advantage of the power addition capability found there. Dynamic conditions such as weather changes and the presence of an audience require ongoing monitoring and active calibration of the system in order to maintain consistency over time.

System calibration challenges:

- distribution of direct sound transmission
- speaker/speaker summation
- speaker/room summation
- dynamic conditions
  - humidity
  - temperature
  - audience.

Strategies

Our calibration strategy is built around several factors. The first is the philosophical direction; a guiding principle that aids in the process of making decisions when choices must be made which will not benefit all parties. It is critical for us to be realistic about the fact that we have set forth goals that can never be completely achieved. We must maintain awareness of how to proceed when we have reached the limits of win-win situations.

The next component is access to information. An informed decision will not necessarily be more correct than the random selections of a monkey, but the probability of success rises sharply when access to relevant information is maximized. Our calibration strategy rests upon access to at least three specific points in the transmission chain: the mix console output, the signal processor output and the acoustic response of the speaker system in the space. Anything less than full access to these points for every speaker in the system will force us to make leaps of faith and assumptions.

The third component is the analysis tool kit which was laid out in detail in Chapter 8. The analysis tools provide us with status reports on the current condition of the system and the effects of our actions. This feedback provides us with answers so that we can move forward with the implementation of our goals.

The fourth factor is subdivision. It will not help us to diagnose problems if we cannot act on them to fix them. The key to action is having flexibility in the design to be able to independently set equalization, level and delay parameters.

The fifth factor is methodology: a set of recipes, a playbook, a roadmap. The methods for reaching the goals can be reduced to a series of specific tests for finding the relevant answers.

The final component is context. The data for any given point is neither good nor bad in its own right. Context gives us expectations, an adjustable standard with which to judge the result. Is this the expected response under these circumstances? This is the ever-present question.

Ripple variance of 10 dB may be perfectly acceptable in one context but a sign of major trouble in another. A high degree of pink shift is the desired and expected response in a distant off-axis area. This same response would be unexpected in an on-axis coverage area in the middle of the hall.

Techniques

In the end this all boils down to a series of decisions and signal processing settings. There is absolutely no skill required in this regard. Just point the speakers, treat the walls and turn the knobs. A great deal of skill and discipline is required, however, to achieve minimum variance. Ninety per cent of the work will have been done for us if we have adhered to the minimum variance principles during the design and verification stages. The final ten per cent is the most interesting and educational part of the journey, however, because this is where theory meets cold hard fact.

There are five principal sets of decisions that must be made at the calibration stage.

System calibration techniques:

- optimization of speaker position, focus angle and splay angle
- optimization of the room acoustics
- level setting
- delay setting
- equalization setting.

The determination of these settings will be done by a series of tests that comprise the majority of this chapter.

Approaches to Calibration

It may come as a surprise that we should interrupt this exercise in objectivity and science with a discussion of ethics. This is, however, required in light of the decisions that we will be making with respect to system calibration. The previous process, verification, was a series of tests with clear-cut answers. If we found a speaker with reverse polarity, or 10 per cent THD, we knew that the remedy for this would benefit all members of the audience who were in listening range of that speaker. System calibration contrasts sharply with this model, since adjustments which provide favorable response at one location may degrade the response at others. Sadly, this outcome is virtually assured with equalization, level and delay setting, the cornerstones of the calibration process. The realization that our settings will have mixed effects leads us to the ethical dilemma: how do we decide who gets the benefit, and who gets degradation? It is best if we dispense with the denial here and now. There are no complete calibration solutions that will benefit all. We can, and will, strive toward this but inevitably, all of the win-win options will have been exhausted and we will have to make choices that are win/break even/lose. To find the answer we will interpret our ethical dilemma in terms of well-known socio-political systems and see which model provides us the best decision-making directives.

Anarchy

This political model is based on a total lack of structure and order. This is the absence of governing authority and as a result every individual is essentially self-governing, with no linkage to other systems beyond their body perimeter. This recalls the "wall of sound" described in Chapter 6 where the overlap between speakers is so high that no two seats have the same frequency response. In this case there are no calibration settings that will work beyond a single position, so whoever grabs the controls can calibrate the system for their position. This will be the mix position, of course. Listeners in other areas will not share the benefits of the self-calibration at the mix position. Since there is no continuity of the sound quality to the other locations there is no need for the objectivity of an acoustical analyzer. The system operator can only mix to what they hear. If every seat is different, there is nothing they can do on the mix console to remedy this. It's "every man for himself." This would be comical if not for the extent to which it actually occurs.

Monarchy

In this model, decisions are made by a single party with little or no regard to the effects upon others outside of the royal circle. In our audio model the castle is the mix position and the royal inner circle includes the mix engineer, band manager, producer and court jesters. In this case the mix area is monitored with the finest analyzer available and calibrations are undertaken to ensure that nothing interferes with the maximum concentration of power and perfect transmission in this area. It is announced by regal decree that all of the other seats in the house benefit from this single point calibration.

Capitalism

This model is promulgated on the idea that the quality level should be commensurate with the price of the seats. Under this model we can justify prioritization of resources toward the nearby front seating areas and away from the "cheap" seats. This philosophy is very easy to implement, since the natural course of things favors this. The expensive seats have the advantage of close proximity and therefore enjoy high levels and favorable direct to reverberant ratios. The "cheap" seats are disadvantaged in both categories and yet exceed the quantity of expensive seats by a very large proportion. Unless maximum effort is expended to bridge the quality gap, the vast quantities of distant seats will have far inferior quality to the small minority of highly advantaged seats. It is inevitable, under even the best of circumstances, that the more distant seats will be at a sonic and visual disadvantage. Another variation of the capitalism approach is the Lottery model. In this version we hold to the belief that it is statistically possible to have the winning number: that the possibility exists that there is one perfect tuning that will win it all. Sorry You lose.

Democracy

In the democratic model each seating area is given equal weight and every effort is expended to serve the needs of all parties equally. The mix engineer is the band's artistic representative and the system optimization engineer is charged with equal distribution of the artist's message. When decisions must be made which will benefit one area above the next, the effects are evaluated on a majority basis. There are two principal factors to be weighed in such situations: the number of people positively or negatively affected and the degree of symmetry of effect. If the intensity of the effect is symmetrical; i.e. the plus side is equal to the minus side, it is easy to follow the simple majority. If the effect is asymmetrical we must employ the "triage" method to evaluate the formula. If a small minority are strongly negatively affected, but a large majority are positively affected, the decisions favor the majority. If the effects are strongly asymmetrical but the affected quantities are similar, this weighs against implementation of that particular strategy. This is easily pictured if we take it to the extreme: we can equalize the system to be virtually free of frequency response ripple at a single position. The price for this is increased ripple variance at all other locations. This strategy, if viewed with objectivity, would be difficult to justify ethically as more than an implementation of the monarchy model discussed previously. By contrast a peak at 200 Hz is found over a large majority of the coverage area of the main system. A compensating filter will not benefit all of the seats but its usage can be justified as helpful to the majority.

It seems self-evident that the democratic ideal should be the model for our design and optimization strategies. This strategy requires extraordinary efforts on our behalf to implement. Our desire to implement strategies that benefit the majority will require us to do more than measure in one position and issue proclamations. We must get out there where the people are and find out what is happening in every sector of the population. This is a daunting task. Must we measure 12,000 individual seats in an arena? This is absolutely impractical. Therefore certain seats must be elected to be representative of the district as a whole. This choice is not a random polling sample but rather a very carefully determined placement. The strategy behind this placement is the single most important factor in system optimization. This is where the information on which optimization decisions are implemented and justified. All other choices flow from this strategy.

TANSTAAFL and Triage

The evaluative principles of TANSTAAFL ("There ain't no such thing as a free lunch") and the decision-making structure of acoustic triage were introduced in Chapter 7. These will be equally applicable to the calibration stage.

The TANSTAAFL principle will come into play with each implementation of a calibration setting. There are no effects without side-effects. An improvement at one location will change others. The extent to which the other areas might be affected must not be forgotten in the excitement of discovering a solution for a point in space. The TANSTAAFL principle underscores our need to monitor the effects of a given change at a wide range of areas. One of the most common pitfalls in system calibration is the exaggeration of a single position solution into the belief of global benefit. We must remember: if something seems too good to be true, it probably is.

Very often the solution for one area is degradation for another. We may simply be transporting the problem to a new location. At first glance this sounds like a break-even proposition but that is not necessarily the case. Taking out the garbage does not eliminate it, but it does move it to a far preferable location. If we can find solutions which work in the most populated areas, we can justify the presence of side-effects for the minority. The overlapping areas near a spatial crossover are certain to have high variance. If we can split a spatial crossover down the length of an isle, we will have transported the problem to a position where the paying customers will not be found. TANSTAAFL keeps us on the lookout for side-effects. Acoustic triage helps us decide what to do about them.

Perspectives In the early days of using SIM I had to set the system that was tunable and get permission for placing the system in an ideal location. Most of my energy usually was exhausted at this point. By the time when the main job, SIM tuning, was in gear, I could not help feeling drowsy. Especially, the periodic noise of SIM1 put me to sleep well, br, br, br, bo, bo, bo, da, da, da, bi, hi, hi, beeen, keeen, ss ...

Akira Masn

Access to Information

Our calibration settings should be based on informed decisions. We need access to information in three principal forms: physical, electrical and acoustical. The complexity of the task will be aided by division of the system into separate sections for analysis.

Calibration Subdivision

The signal flows serially through three distinct sections: the source, the signal processing and the speaker system in the room, finally arriving at the listener. Our mission is delivery of the source signal. Any tests performed on the source section are part of the verification process rather than calibration, since the operation of the mix console falls exclusively into the scope of the mix engineer. The transition point out of the artistic sphere occurs at the console outputs. This is the handoff point for the source and we are charged with taking delivery. There are some specific guidelines for acceptance of the source, which will be detailed later. For the moment we will consider the signal to have passed over the "Art/Science" line.

The "Art/Science" line is the scope of work transition point between operation and optimization. It is not the optimization engineer's job to make it sound "good." We don't have such power. Our job is to deliver a system which has the potential to sound "good" to as many audience members as possible. The mix operator's goal is subjective: good sound. Our goal is objective: same sound. My good is not your good, but we can agree on sameness. Generally speaking, a mix engineer will find it easier to achieve their goals when we have achieved ours. An artist prefers to work from a clean canvas.

The source/signal processing transition is a mandatory monitoring point, since we must know exactly what we have received so that we can see how well it holds up to the hazards of transmission and summation. This is the electronic reference point from which our job will be judged. The final monitoring point is our simulated hearing mechanism: the measurement microphone. A transfer function measurement will be made that shows the difference between the reference signal (the source) and the measured signal at the mic. If the transfer data shows an undesirable response, actions can be taken either physically or electronically until the desired result is achieved. An interim measurement point is found at the transition point between the corrective signal processing and the speaker system components, as shown in the flow block diagram in Fig. 10.1. This allows the overall response to be subdivided

428

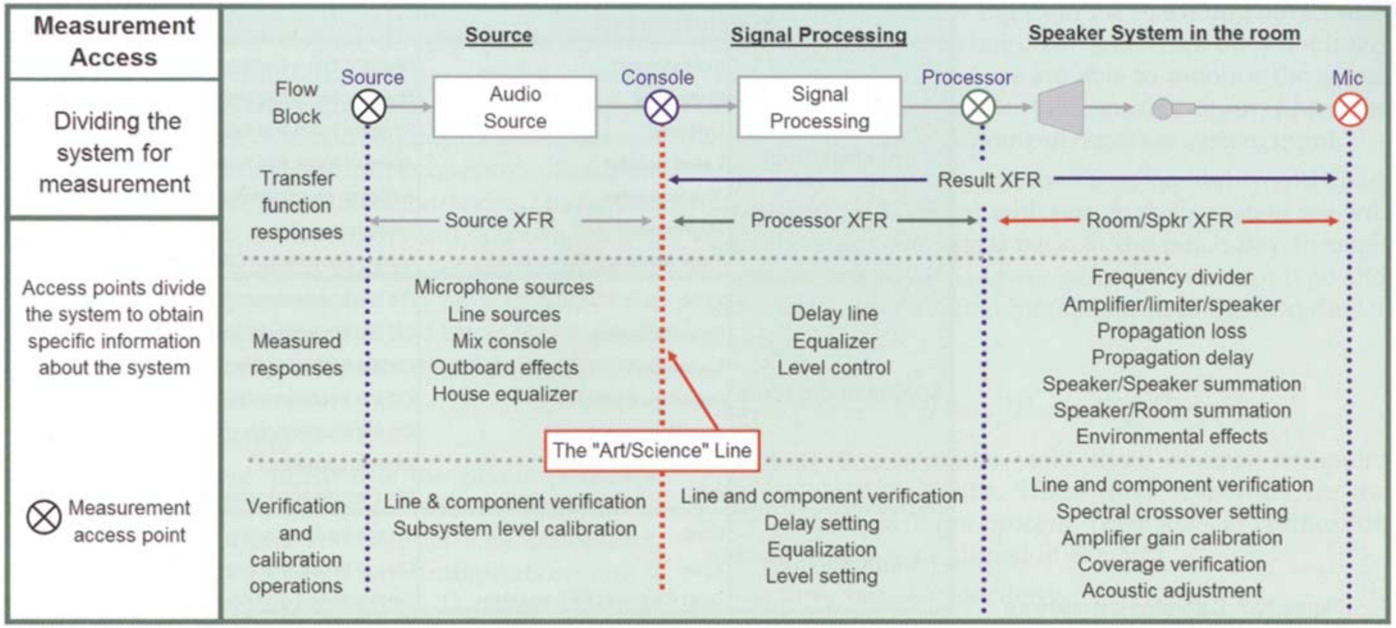


Figure 10.1 Flow block of the electronic and acoustic measurement access points (console, processor and microphone) and the three transfer function responses (room/processor/result)

into two main sections: the electronic response of the signal processing and the speaker system response in the room. This, in turn, allows the electronic correction to be seen in context with the acoustic anomalies for which such correction is attempting to compensate.

Efforts to achieve a minimum variance response will be undertaken in both the signal processing and speaker/room sectors. Physical solutions such as speaker positioning and acoustic treatment are done entirely in the speaker/room sector. The setting of signal processing parameters is done in the processing sector based on data derived from the speaker/room sector. The proof of performance is found in viewing the combination of the processing and speaker/room sectors.

The three monitor points (console output, signal processor output and microphone) yield three distinct transfer function measurements: the processor response, the speaker/room response and the overall result.

The transition between the console source and the signal processing is clear-cut. The second transition, from the signal processing to the speaker system, has some important features which require clarification. The speaker/room system contains more than just the loudspeaker drivers. The line is drawn where the last full-range line level signal is found before entering frequency dividers or dedicated speaker controllers. There are several reasons for this. The first is the fact that a transfer function measurement which uses a post-frequency divider signal as its electronic reference cannot give us an accurate acoustic response. A low passed electrical reference can make a subwoofer appear flat to 10 kHz since the loss is divided out by the transfer function. The second is that equalization and phase alignment parameters which are dedicated to a particular speaker enclosure model are best kept separate from the calibration parameters that will be set for the installed system in the room. If a speaker system has a flat frequency response in free field it then becomes clear that pink shift and ripple variance

Examination tool		Calibration role
Physical tools	Inclinometer	Speaker focus fine tuning
	Protractor	Speaker focus fine tuning
	Origami	Speaker focus fine tuning
	Laser pointer	Speaker focus fine tuning
	Thermometer	Establish environmental baseline for variable temperature environments.
	Hygrometer	as above
	Tape Measure, Meter Stick	Subwoofer spacing. Cluster trim height etc.
Simple audio tools	VOM	Off line troubleshooting of problems found during calibration
	Polarity Tester	Off line troubleshooting of problems found during calibration
	Listen Box	On line troubleshooting of problems found during calibration
	Impedance tester	Off line troubleshooting of problems found during calibration
	Oscilloscope	On line troubleshooting of problems found during calibration
	Sound Level Meter	Minimal application
	RTA	Mix console hood ornament
Complex tools	Ears	On line troubleshooting of problems found during calibration, spatial averaging
	Eyes	On line troubleshooting of problems found during calibration
	Dual Channel FFT analyzer	Electronic and acoustic signal test. Details to follow

Figure 10.2 Calibration test reference

Perspectives Alone on a deserted island with your favorite speakers, what would you choose EQ or time delays? If you chose EQ, you've just been voted off the island... Time delay is the name of the system-tuning game... at least as a first step.

Francois Bergeron

found in the measured response are the effects of speaker/room and speaker/speaker summation. If the speaker is not flat in free field we cannot attribute any expectations to the response we are finding in the room, e.g., a four-element point source array can be expected to have pink shift of up to 12dB from just the speaker/speaker summation. When speaker/room summation is added to this we can expect more. If the speaker is not a flat speaker to begin with, the context of the measured response becomes more challenging to discern. Is the variance we see due to expected mechanisms such as summation and air attenuation or is there something else?

Physical Access

The position, splay angle and focus angles of the speakers can be verified by the standard physical measurement tools. A tape measure provides placement confirmation such as height. The focus angle can be found vertically by an inclinometer. The splay angle between elements can be checked and adjusted by a protractor. The precise location where a speaker is aimed can be found with a laser pointer. In most cases the final focus positions will be determined by acoustic performance as part of the calibration process. The physical tools are extremely useful for symmetry verification after one side has been calibrated.

Electronic Access

Console/Processor/Mic

Access to the three measurement points will need to be done without interrupting the signal flow. This is done by taking a parallel split of the electronic signal, or by sending and returning the feeds. This can be done in the analog or digital domain, just so long as the analyzer can read the signal.

Some analysis systems go beyond simple monitoring and enact control of the speaker muting which will inevitably take place during the alignment process. Such systems require the signal processor output to be routed through the analyzer, rather than the simple parallel connection.

Practical Concerns

In the best-case scenario we have direct physical access to the input and output connections of the signal processing device. There are times when this is not practical, especially in permanent installations or with signal processing devices having non-standardized multipin connectors. Direct access ensures an accurate reflection of the device. The further we get from the actual device inputs and outputs, the less confidence we have that the measured behavior is attributable to the signal processor alone. The most critical factor is the avoidance of insertion errors that cause the response to change when we unpatch our measurement equipment.

Patch Bays and Interrupts

One often-encountered scenario is the phone jack type patch bay. This field of connectors provides insert and interrupt capability to the inputs and outputs of the processing. There are many different patch bay configurations and it is beyond our scope to cover them all. The most common style is the "half normal" configuration. In this scenario the upper jack bay provides a "listen jack" which allows us to monitor the signal flowing from the preceding device as it enters the next in the chain. The listen jack does not interrupt the signal flow, so we are able to monitor the signal at our two desired points: the console output/processor input and the processor output/speaker system input.

If we wish to go beyond the listen capability and enact speaker muting control with our analysis system we will need to return the signal back to the patch bay through the lower jack bay. The lower jacks are interrupt type and allow us to inject a signal into the input of the next device in series.

Room/Processor/Result (Room/EQ/Result)

The three access points yield three distinct two-point transfer function results. These three responses are the heart of the calibration process. Their roles in verification and calibration are outlined in Fig. 10.3.

The three transfer functions:

1. Room/speaker: processor output (speaker system input) vs. mic (speaker system output).

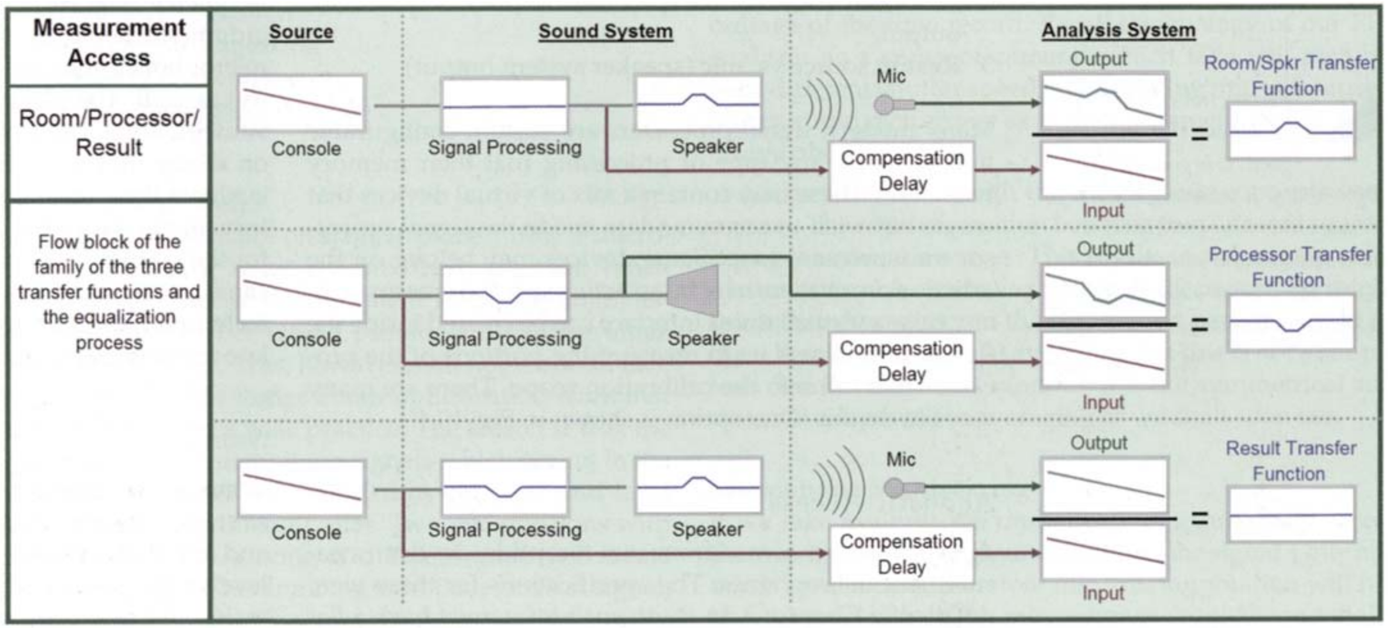


Figure 10.3 Flow block of measurement access points of the three transfer functions and their roles in the equalization process

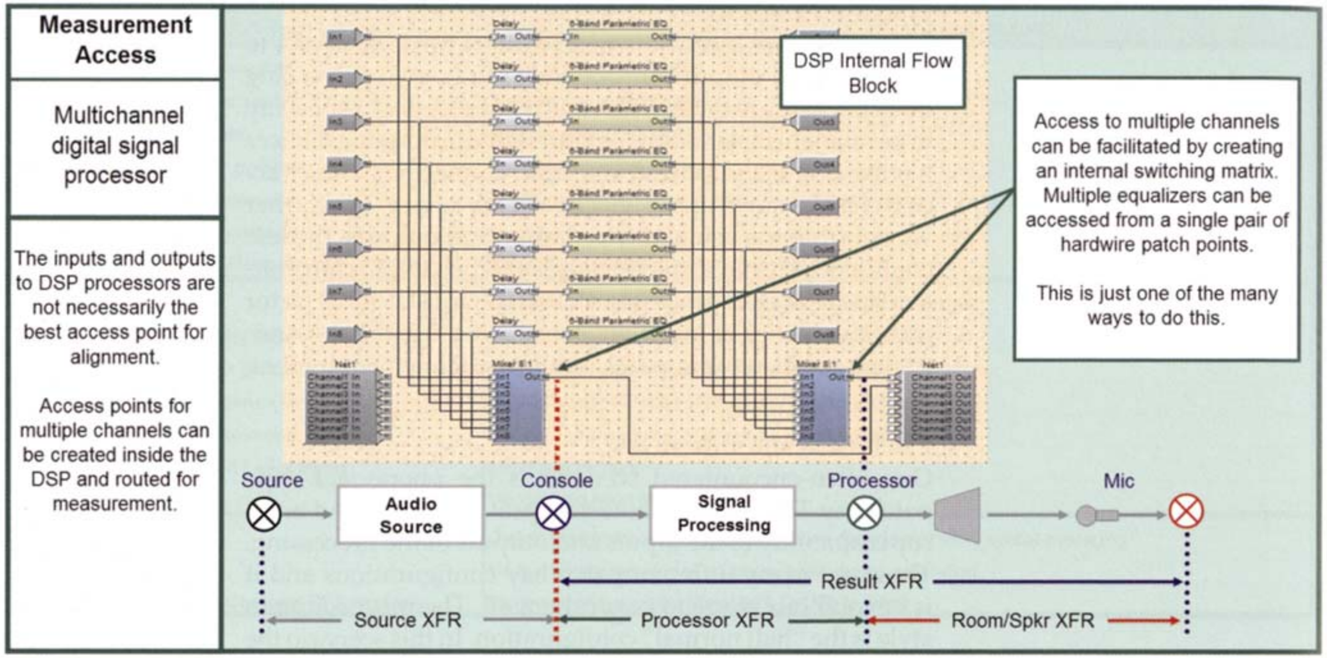


Figure 10.4 Creating measurement access points inside a user-configurable digital signal processor

2. Processor (EQ): source (processor input) vs. processor output.
3. Result: source vs. mic (speaker system output).

Many modern signal processors are custom-configurable into virtually any type of processing that their memory will allow. These may contain a mix of virtual devices that we do not wish to accommodate inside the signal processor measurement loop. Some devices may belong on the artistic side, or there may be an active spectral crossover. In any case, a virtual access interface can be created inside the device that allows us to measure the portions of the processor that fall into the calibration scope. There are many ways to do this. One option is shown in Fig. 10.4.

Perspectives Some people say if it measures good, it is good. Some argue that if it measures bad, it is bad. Others suggest that if it measures bad and sounds good, it is still good. I just suggest another microphone position.

*Paul **Tucci***

Acoustic Access

Acoustic access comes from our reception devices: measurement microphones. The specifications for these were detailed in Chapter 3. In short, such mics must have a flat frequency response if we are going to use them to make judgments about the sound system. If there are multiple microphones, they will need to be level-matched (sensitivity) as well. The placement of the microphones is a major strategic issue. The mix engineer evaluates the sound based on where they are. The same is true for us. We can only evaluate the system where the mics are. There is a long history on the issue of determining a representative response for the speaker system. We know that the sound is not the same everywhere. What can we do then? The principal technique has been to take multiple samples, a practice known as spatial averaging.

Spatial Averaging

Some forms of acoustic measurement seek to find the average response over a specified area. This is common and appropriate when analyzing the distribution of sound level over an area without regard to its degree of variation inside that area. An example of this would be HVAC noise averaged over the seating area, rather than on a seat-by-seat basis. HVAC noise is random and therefore subject to different types of variability than we find in speaker systems. There are not fixed summation relationships between the arrivals from different HVAC ducts as there are from speakers which are transmitting related signals.

The basic principle of spatial averaging is to provide an average response from multiple positions. This averaged response represents the voice of the majority for the area. The intent is that conclusions can be made and actions taken from this representative polling method.

Spatial averaging can take various forms. These include the primitive concept of summing microphones, microphones in motion, microphone multiplexing, the mathematical averaging of individual responses, and the practice of visually examining the trend lines of multiple traces. We will take some time to explore the applicability to our goals of these different forms.

Spatial averaging methods:

1. microphone summation
2. microphones in motion
3. microphone multiplexing
4. mathematical averaging
5. optical averaging.

Summed Microphones

It is an everyday practice on stage to place multiple microphones in different areas and mix them together. When faced with an array of measurement mics, one might assume that this practice would provide a useable average response of the area. This, however, will not work. In fact, this is one of the few things in our optimization work that we can say is never a wise practice. The reason is that the summation of the microphone signals which occurs in the wire (or a mic mixer) is governed by the time and level offset between the different mics. Two mic positions with identical responses arriving at different times will create massive combing. This combing is in the electrical signal, not in any of the acoustic signals in the room. If we made changes to our system based on this we would soon be looking for career opportunities in video.

Microphones in Motion

A single mic can be moved around in the space and its response captured over a period of time. There exist analyzers with "time capture" modes designed for this. An RTA set with an extremely long averaging period can be used to provide a continual capture of the response over a distance covered during the time period. This has advantages over the summation technique, but is equally inapplicable for our purpose. We must remember that for our frequency response data to be useful we must have it in complex form, i.e. amplitudes and phase. An RTA throws out phase so it was eliminated for our purposes earlier. So what about if we move a mic with a dual-channel FFT analyzer? There is the matter of phase again. If the mic is moving, we are destabilizing phase and our coherence will be reduced because the new data does not match the average of the old data. The continual changes conflict with the established average. As frequency rises, the coherence loss increases since the phase shift represents a larger percentage of the time record. Recall the analogy of our FFT analyzer as a series of cameras which take still pictures with different shutter speeds. The moving mic will cause a blurry audio picture just as a moving camera would cause blurry photos.

A secondary issue with the moving mic concerns summation. The relative position between two or more sources changes as the mic moves. This results in a changing summation pattern, as we have previously discussed. We might think that averaging out the summation pattern would be the whole point of spatial averaging but this is not a simple task. Summation is not random, it is not symmetrical and does not adapt well to averaging as we shall soon see.

Microphone Multiplexing

If we take multiple microphones and sequentially access them individually we are multiplexing the signal. Only one mic is on at a time, therefore mic/mic summation will not invalidate the data. The microphones would sequentially feed a common input with a continual stream of the data from those mics. It is analogous to juggling, as we endeavor to keep multiple mics in the air. A response is then computed, as if the data stream were coming from a single mic. This response spatially averages itself by virtue of keeping the data stream filled with representatives of each position. Obviously it is not practical to perform transfer function analysis on a stream of data that comes from different sources at different distances at different levels. This removes phase and coherence from the equation and reduces the outcome to level only As a result, the advocates of this technique use RTAs with a pink noise source. The advocates of the multiplex mic technique do their work in the venue that coincidentally shares its name, the multiplex cinema.

The Star Pattern

The star pattern is a mic position technique initially developed by Roger Gans in the early 1990s. The concept is a form of optical averaging (described later in this section) of five positions within the central coverage area of a speaker. Each of the positions is stored in memory and then viewed, separately or together. The "averaging" is performed with our eyes as we look for trends in the response envelope. Multiple traces laid on top of each other will turn into a forest of narrow dips, but a clear trend is usually evident in the envelopes. The equalization is then employed to match the composite envelope by eye. The difference between the mathematical averaging and our optical method is that our eyes can be trained to delete the deep and narrow dips, focusing instead on the audible envelope.

For the star technique to be successful, its radius of operation must be confined to the isolation zone. The horizontal center/vertical mid-point should be the most isolated of all points in the coverage of a given system. This is where the speaker should be most free from the interaction of other speakers, and hopefully the room. As the mic positions move outward into interactive areas the variance is certain to rise. Equalization affects all areas of the given system equally. Therefore, the most sensible place for equalization decisions is in the most isolated area.

Mathematical Trace Averaging

The most promising avenue in spatial averaging would seem to be taking a series of complex traces and creating a mathematical average of the responses. If three responses showed a peak at 2, 4 and 6 dB respectively, the average would be 4 dB and it would seem to be a defendable strategy to employ a 4 dB cut filter at the given frequency. In this case it would be, but unfortunately not in all, or even close to all, cases. The reason is the inherent asymmetry of the summation mechanism. The peak heights are always smaller than the dip depths, sometimes a little smaller and sometimes *much* smaller. The mathematically averaged response is only aurally representative if all data that makes it up is from the positive side of the summation equation. If data appears from the dark side of summation it will overpower the data from comparably strong positive summation. Accordingly, for any degree of combing in the response we will have a skewed representation of the averaged signal. Our averaged trace will have defendable conclusions for the coupling zone and isolation zone areas, but will be invalid for the combing and combining zones.

How does this occur? We will start with an example. Let's take an average of three traces taken from three different positions relative to a two-speaker array. Each of these traces shows peaks ranging from 5 to 6 dB. This would be an indication of positive summation of two sources within a window of 1 dB and 45 degrees. The average would come out to be 5.5 dB or so (5 + 5.5 + 6/3 = 5.5). Now let's add a fourth trace which is also the result of a 1 dB relative level summation, but is 180 degrees out of phase. This trace comes in at —19 dB. The average is now computed as 5 + 5.5 + 6 — 19.2/4 = —0.4. One cancellation just wiped out three full-power additions! The average reads O dB even though 75 per cent of the positions say there is a large peak. The complete extent of the problem with this approach becomes clear when we consider the fact that the one mic position where we can't hear the sound has the mathematical power to nullify the three positions where we can hear it loud and clear.

Any scheme that uses multiple traces and simply divides them will suffer from this asymmetry. While this

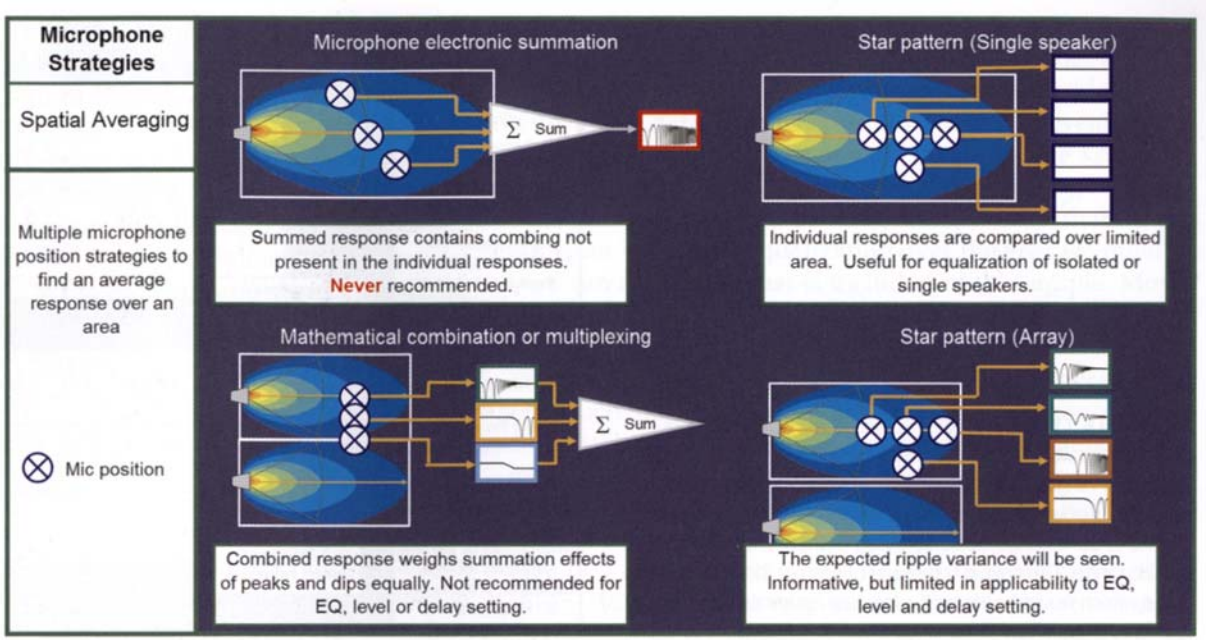


Figure 10.5 Spatial averaging considerations

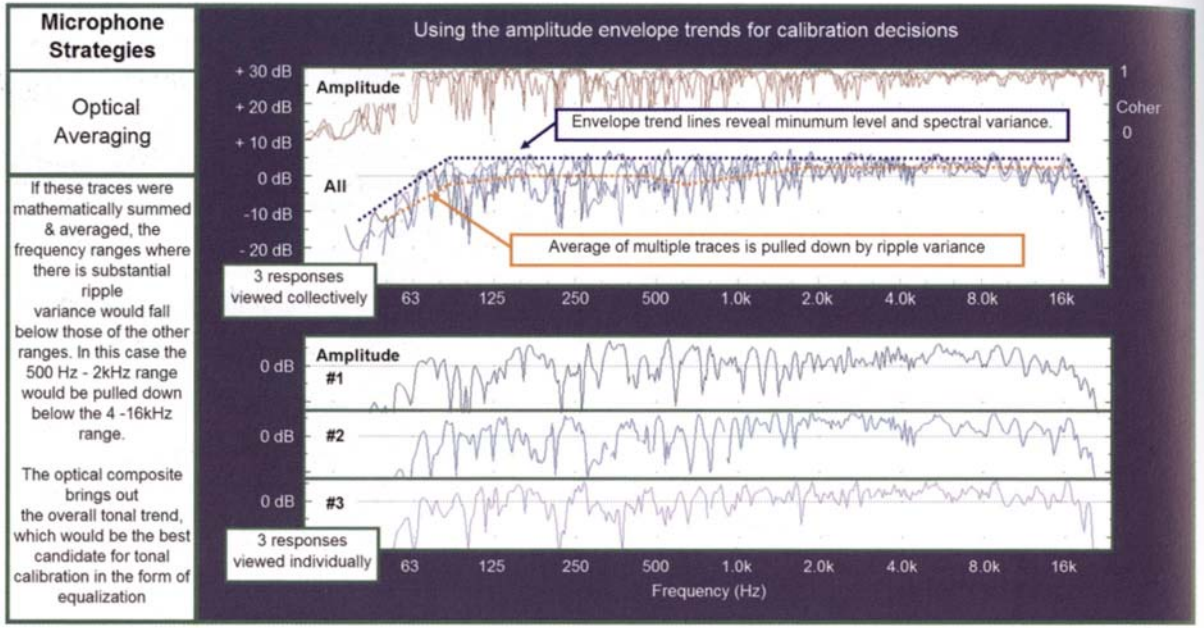
is defendable mathematically, it does not resemble our perception.

Another limitation of the mathematical trace averaging is its lack of contextual clues. The traces are given equal weighting by the math whether or not they should have equal roles or expectations. The average of an on-axis (0 dB) and off-axis (-6 dB) trace is -3 dB. What do we learn from that? Taken individually the data tells us where the coverage edge can be found. Taken together we might conclude we need 3 dB of equalization.

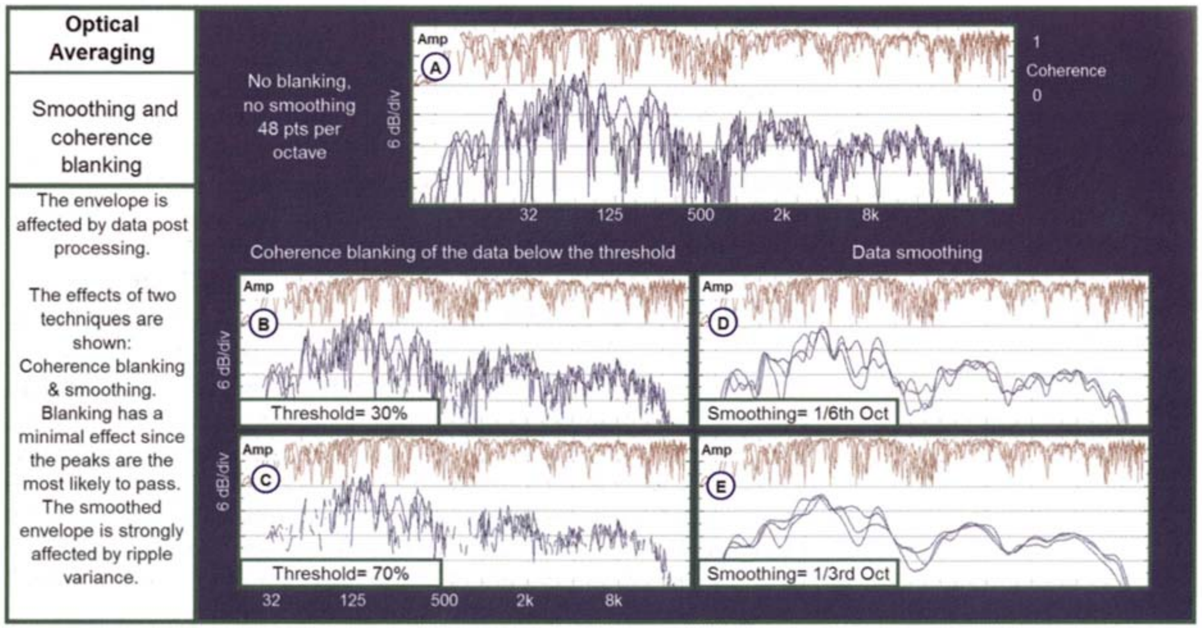
Our hearing is better able to perceive that which we can hear rather than what we can't. This is rather obvious but its implications regarding the asymmetry issue may not be. The importance is our increased sensitivity to the presence of the highest-level values in a frequency range over the lowest. If certain frequencies stand out above the crowd, this is what we hear. There are two ways that frequencies can stand out. There may be a peak in the response, or it could be a survivor in the neighborhood where a lot of cancellation has occurred. In either case our hearing favors the remainder over what has been removed. This differs from the mathematical trace averaging which, by treating peaks and dips the same, allows the deep nulls to be the dominant response. For mathematical trace averaging to be optimized for our application, it will need to treat the upper side of the envelope with preference over the nulls in the underside.

Optical Trace Averaging

The high side of a frequency response curve, the peaks and the survivors, create the audible character of the response, the envelope (described in Chapter 3). We will use the envelope for position, level, and equalization adjustments. Our approach to the peaks and dips will be as asymmetrical as they are. The peaks and survivors will be viewed as the representatives of audible response character, whereas the dips and cancellations will be viewed as evidence of damage to the system response. Our efforts to minimize damage will be focused on position, level, acoustical and



Opticalhaveraginghscreenhdumphofhenvelope.hIf thesehtraceshwerehmathematicallyhsummedhandhaveraged,hthe frequencyhrangeshwherehtherehishsubstantialhripplehvariance wouldhfallhbelowhthosehofhthehotherhranges.hInhthishcasehthe 500hHzhtoh2hkHzhrangehwouldhbehpulledhdownhbelowhtheh4hkHzhto 16hkHzhrange.hThehopticalhcompositehbringshouththehoverallhtonal trend,hwhichhwouldhbehthehbesthcandidatehforhtonalhcalibrationhin thehformhofhequalization



Figureh10.7h Fourhpositionshwithinhthehcoveragehareahofhthe speaker hThehsamehbasichenvelopehtrendsharehrevealedhbyhthe opticalhaveraginghmethod

delay adjustments. There is very little that equalization can do for us in the category of reviving cancellations.

An effective method for finding the envelope is hereby termed "optical averaging". Multiple measurements are taken, such as with the star technique, and the traces are simultaneously overlaid on the same screen. The overall trends of the envelopes can be seen, and common trends found by eye. The more extensive null structures can be selectively ignored for the moment and the focus placed on the most audible characteristic: the spectral outlines of the high points over frequency.

Limitations of Spatial Averaging

All of this begs the question of how much spatial averaging is advisable or appropriate. First we can look at the problem from a statistical point of view. Do we have enough positions to provide a statistically representative sample of the overall response? Each mic position represents an area as large as the microphone diaphragm. From a statistical point of view it is likely that we will have less than 0.1 per cent of the listening area represented. If we did move the mic to a thousand positions then we could average these all together and get something more statistically compelling. Let's say we learned that we have a 6 dB peak at 2 kHz as an average over the space. Would we conclude from this that the splay angle is wrong between the array elements? Not likely. All we can conclude from a pile of spatial averages is *equalization,* which is only a single technique in the overall picture of *optimization.* We need our mic positions to correspond to the specific operations we will perform for optimization: speaker position, acoustic evaluation, relative level, delay setting and equalization. The data that drives these procedures is found in specific locations, not general or averaged ones.

One inherent limitation of spatial averaging is that any scheme which combines multiple measurements into a single value fails to inform us of the degree of variance *between* the sampled positions. It would be possible mathematically to apply standard deviation equations to the data and get such values. What then? Knowing that the averaged value at 2 kHz has a deviation of 10 dB does not help us to decide where to place a filter, or whether we need to reposition a speaker. Only the optical method leaves us with intact data for each position, thereby revealing the individual circumstances at each position. The key question is this: what do we want the spatial averaging to tell us? If we are looking for the "average" response so that we can apply an "average" equalization over the area we must be very careful about what is included in the sample. Movements of even a single seat width will have large effects if they are at positions near the spatial crossover with another speaker or strong reflections. If we average several such positions we will get a meaningless average since each location will comb at different frequency spacings. This is the expected response based upon the principles of summation and would be revealed as such by the individual responses. Viewing the spatial average would give us a curve which would appear to be an appropriate choice. Viewing these responses individually would make it clear that no equalization setting will be appropriate for more than a single point.

The most useful application for spatial averaging is when there are multiple speakers in symmetric array which are controlled by a single equalizer. In this case the response for each array element is viewed on-axis and stored. Since each on-axis area is of comparable size they are equally weighted in terms of audience benefit. The optical averaging method can then be used to view the envelope trends and discern the best equalization fit for the entire area.

In conclusion, any form of response averaging to a composite trace is subject to asymmetry-based weighting in favor of dips. This renders such averaged responses unavailable for equalization. Composite traces also remove the evidence of the normal progressions of variance over the speaker coverage area. This lack of contextual clues renders such traces incapable of aiding the process of speaker positioning. If microphones are strategically placed and individually viewed and compared, the response progression will be revealed. Strategic locations will provide

Perspectives As we walked up the first balcony under the second balcony the general consensus was that the high end was not making it to the back half of the balcony. Looking at the analyzer and comparing the mic at the rear of the orchestra level with the mic at the rear of the first balcony revealed that the level of the horns was the same at each location. It was a build up of low mid energy that was masking the high-frequency energy. The correct solution was to equalize the low mids rather than boost the HF. This demonstrates two important principles the need for separate controllable sections of the system for different parts of the room and the need to make multiple measurements (which is only really practical using multiple microphones) and many comparisons throughout the space.

***Fred** Gilpin*

specific answers in regard to acoustical, position, level, delay and equalizer adjustments.

Microphone Placement in Context

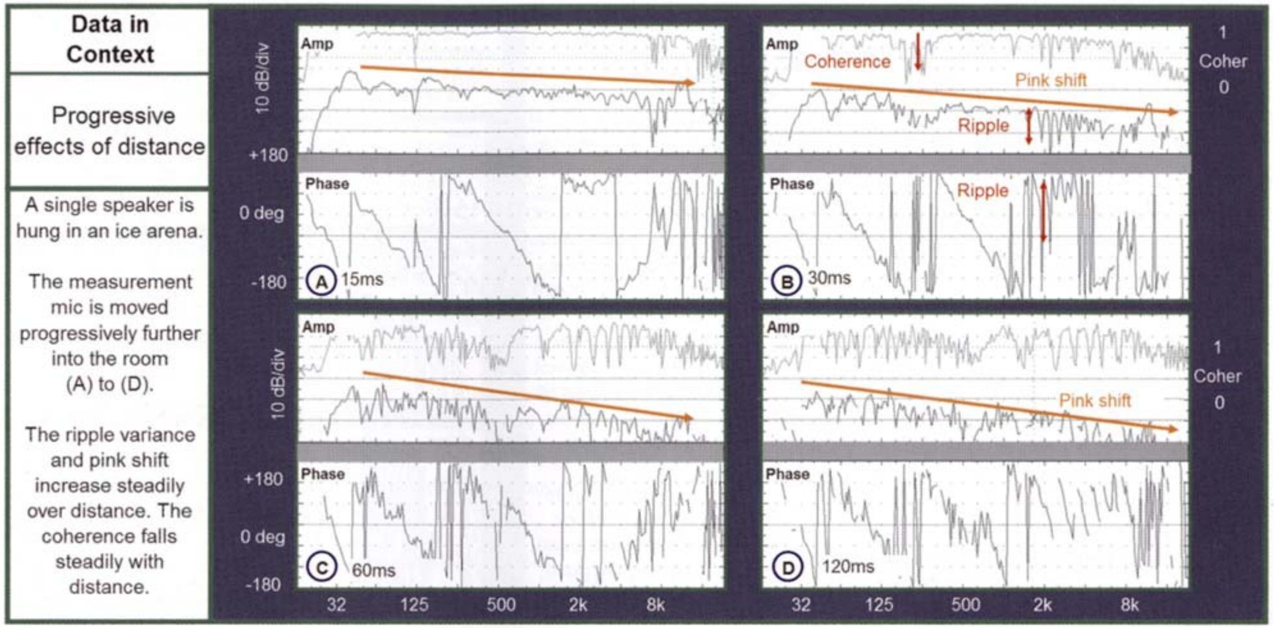
In the previous chapter we discussed how the verification process was structured to provide answers to specific questions. This gave the measured data a context from which we could discern whether this was expected or unexpected. The results led to a plan of action. The calibration stage is no different, except that the expectations are much more difficult to quantify in the hostile circumstances of far-field acoustic measurement. We might have the complex interaction of multiple speakers in a reverberant room arriving at our mic. What the heck is that supposed to look like? One thing is for sure — it won't resemble anything we will have seen in a sales brochure.

Learning to read data in context is an acquired skill. Examples will be presented here to aid the learning process, but there is no substitute for the real thing. To fully grasp this concept will take experience. More than twenty years later, I am still working on it.

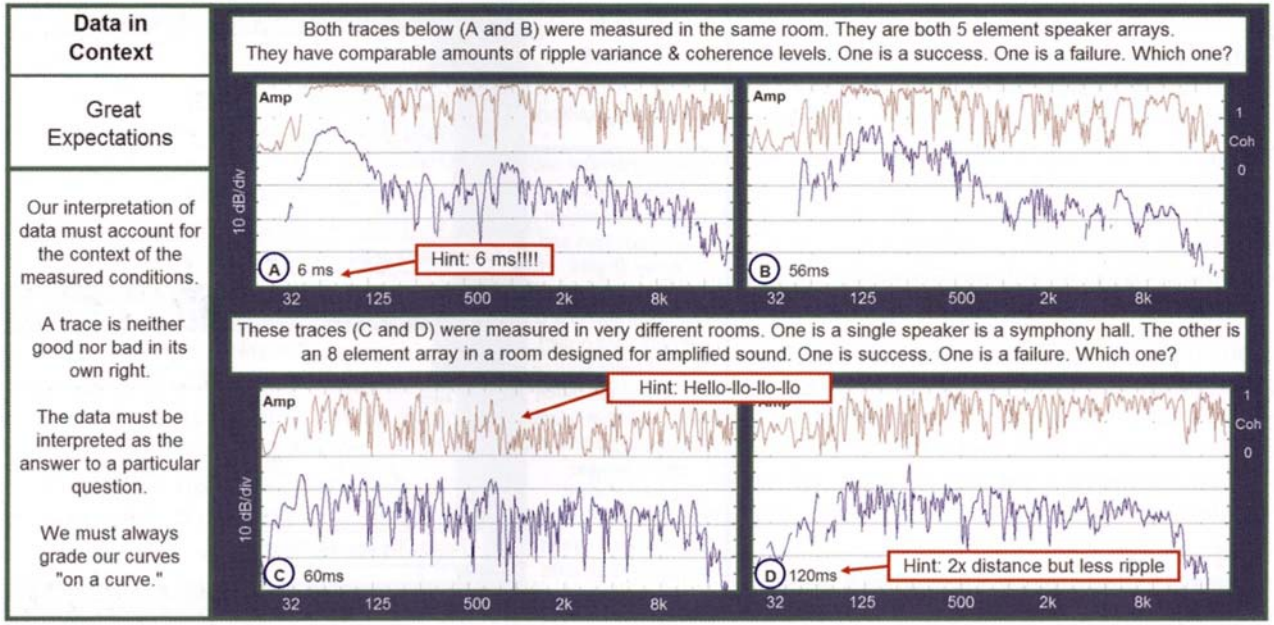
Context gives rise to expectations, which can be met, exceeded or failed. Most of our contextual information comes from the process of comparison. Let's consider some examples.

Great expectations:

1. The on-axis area: our expectations are that this will be our best-case scenario for the speaker. If this area looks worse than other areas we have a contextual clue that perhaps we are not actually on-axis or we are not measuring the speaker we think we are. On-axis should always be the best response. If not, investigate immediately.
2. Distance: how far are we from the source? HF loss at 50 meters is expected. The same response at 10 meters raises the question of whether we are in the on-axis area. A fight for good coherence is expected at 50 meters. We should have things well in hand at 10 meters.
3. The room acoustics: are we in a hall with the acoustics optimized for the symphony? Are we indoors in a glass and cement hockey arena? Are we are in the breezy outdoors? All of these conditions will make us lower our expectations.
4. Local conditions: are we next to something that will give us a strong local reflection? Do we have line of sight to the speaker? The bald guy sitting in the row just ahead of the mic, could that be a reflection source? Yes (true story, Osaka, Japan, 1989). Investigate the local conditions for contextual clues before taking action on the speaker system. Don't equalize the system for the bald guy!
5. The array type: a twelve-element coupled array is expected to create a lot of pink shift. The low-frequency overlap should cause the spectrum to tilt substantially. If we were to see a flat response instead we might think that we have bragging rights for not having to equalize it. The contextual expectation would lead me to begin searching for the LF driver polarity reversals in the array.
6. Inverse square law: It's repeal is unlikely despite marketing claims to the contrary. We measure on-axis at 10 meters. We move back to 20 meters. If the response does not drop 6 dB then we must examine the situation for contextual clues. Did we move off-axis as we moved back? Are we inside the lower levels of a parallel pyramid? These contextual factors lead us to refine our expectations so that the next time we measure we will anticipate their effects in combination with the distance loss of the inverse square law.
7. Off-axis: it is a simple matter to measure a speaker on-axis and declare its placement as optimal. This is like the politician who gives speeches only to their loyalists. Taking the mic out in search of the axial edge takes guts. If the edge is not found where we expect it, we will need to take action. The edge is found by viewing the off-axis response in the context of the on-axis.
8. The spatial crossover area: is an octave-wide 20 dB dip at 5 kHz a problem? It depends on the context. If this is occurring in the center of our most isolated on-axis area we have strong evidence of something serious. If this is found within a few inches of a spatial crossover it is right on schedule.



Figureh10.8h Contextualhevaluationhofhfour measurements.hTheheffectshofhprogressivehdistance onhthehamplitude,hphasehandhcoherence.hAshdistance increaseshthehqualityhofhthehdatahgoeshdown,hashwouldhbe expected



Figureh10.9h Similarhresponsehwithhdifferenthexpectations. Thehexpectedhqualityhofhahshorthrangehmeasurementhis high,hwhilehthathofhahlonghrangehishlow.hHerehthathishnoththe case.hThehshort-rangehmeasurementshrevealhexcessive overlaphinhthehuncoupledhlinehsourceharray

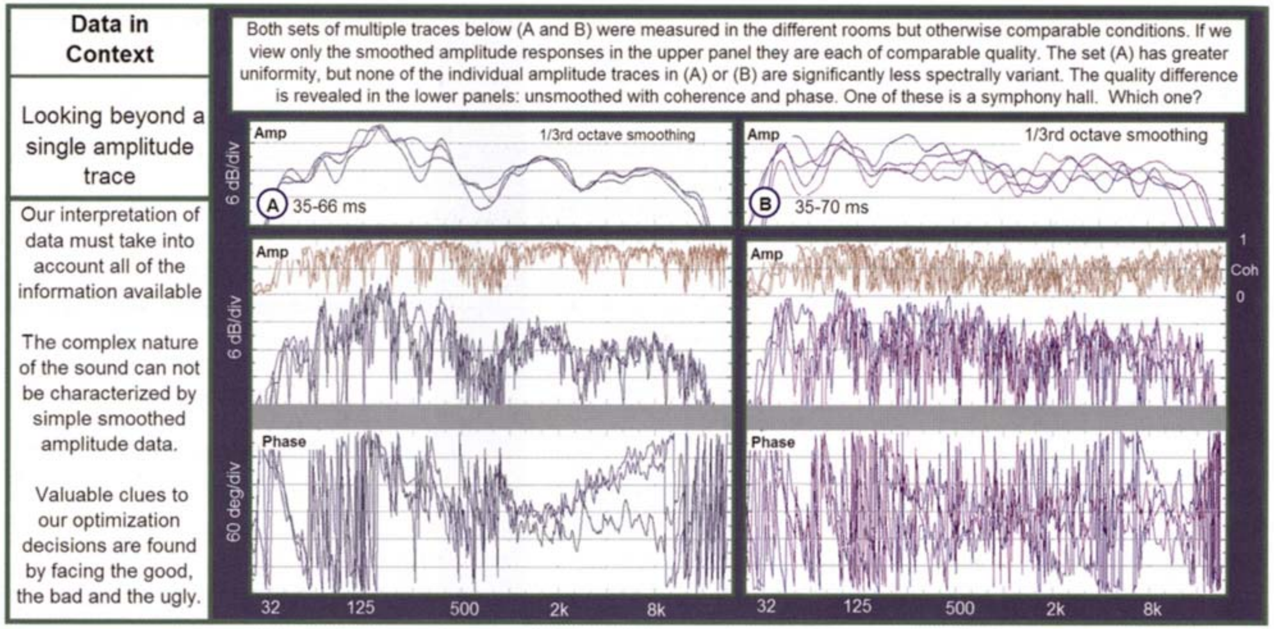


Figure 10.10 An illustration of how looking at amplitude only leaves the data without the clues indicated when other measurements are viewed as well

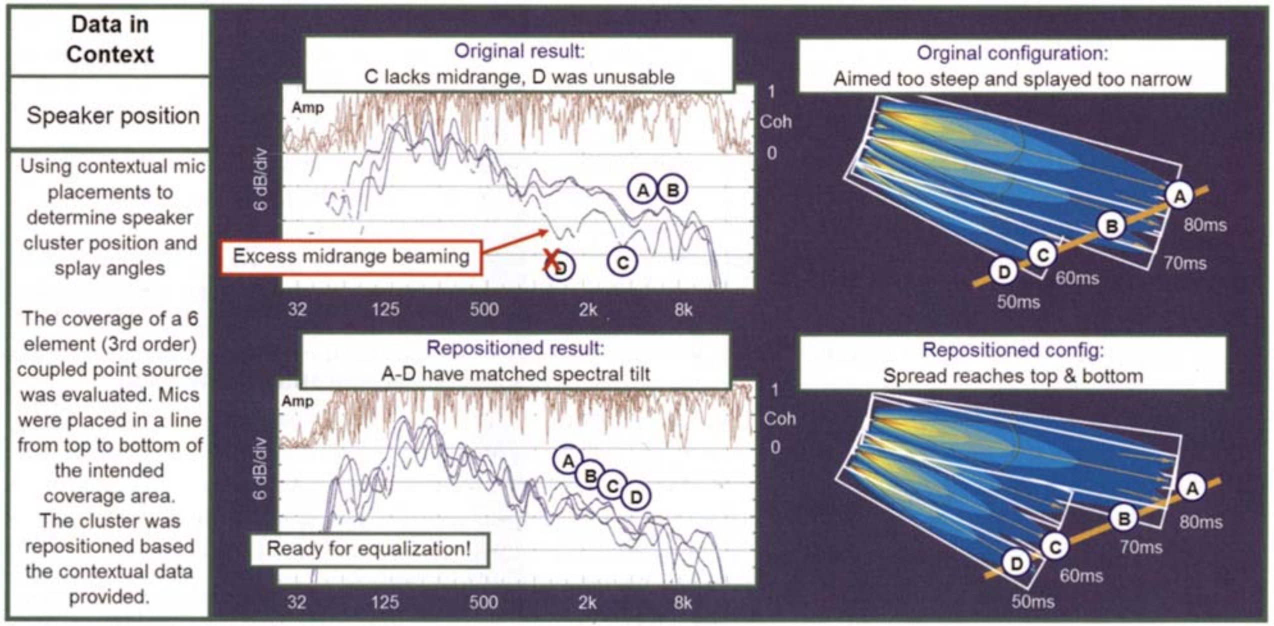


Figure 10.11 Using a contextual mic placement strategy to determine speaker position

Perspectives *This is the most prevalent mistake I see, blind allegiance to a line on the screen, without fully understanding the reasons and factors that make it look the way it does. If something seems like magic, it's only because you don't have all the information you need to understand it.*

Dave Revel

9. Stereo symmetry: we measure the left side. No matter what response was found there we have an expectation: that it will have the same response at the symmetrically opposite position on the right side.
10. Audience presence: we measured the system before the audience came in and stored the response. Now the band is on stage, the temperature has risen and there are changes in the response. Do the changes we see make sense within the framework of the physical changes we know have occurred? This should be carefully considered before reaching for a knob.

As mentioned above, the reading of data in context is a lifelong endeavor. It is also one where we can bring everything we are able to the table: auditory clues, visual clues, past experience with the components, the contractor or the hall, and most of all, our common sense.

Figures 10.8 to 10.11 provide some field examples of contextual data analysis.

Microphone Placement Details

Microphone Placement Classes

Our study of spatial averaging revealed the vulnerability of any method that combines the data from different locations into a single response. This "sonic soup" approach robs the individual responses of their contextual reference, a practice which leaves us without a clear course of action. A different approach is proposed here which relies entirely on individual responses with clearly predefined context. Instead of seeking to find a common average response, we will seek out the differences between the key locations that characterize the expected behavior of the system. We have studied the standard variance and summation progressions and the mic positions will be located at the critical milestones of these progressions. If we know the response at the milestones, we can interpolate the response between them based on the known progressions.

From a statistical point of view there is virtual certainty which will be decisive in our choice of position: the frequency response at every location in the room (except for its perfect symmetric opposite) is unique. Therefore the perfect equalization solution for one location is de facto imperfect for all others. We have studied summation extensively and there is simply no denying this reality. Since no single solution exists, we must move to the next question: is there a location that can be statistically defended as the *best* representative or can we simply make an arbitrary choice? No, there is not a best single location. But yes, there are best *locations.* There is a best location for equalization and a best location for speaker positioning and delay setting, but they are not the same.

Our perfect equalization for that single position is not guaranteed to simply turn sour with a small movement in the space. The applicability of the equalization over the space may fade away gradually (the best-case scenario) or can be become highly inappropriate in a few steps. The best location for equalization will be the one that holds up for the largest area, and yields its applicability most gradually. This is found at the on-axis location.

Why there? The decisive factor is the *rate* of change and can be found in the ripple and spectral variance progressions. The lowest rate of ripple variance change is in the on-axis area of the speaker, because this is the location where it enjoys the highest amount of isolation. This is also the flattest point in the spectral variance progression, with increasing pink shift as we move off-axis. If the equalization is carried out in the on-axis area, two favorable parameters concur: the equalization will hold up for the largest area (due to isolation) and there is the lowest risk of reverse pink shift (the mids and highs being louder than the lows and sounding like a telephone). Any other position will have a lower probability of success in these two most important categories of equalization.

Level setting will also be tied to the on-axis position, since it represents the anchor point of the minimum level variance line. Our hopes of minimum level variance ride first on our ability to set our subsystem relative levels so they can provide the same level in their respective on-axis areas, where they exercise the most dominant control.

Speaker positioning will require at least two mic positions. How can we tell if a speaker is positioned correctly if we have only the on-axis data? It is the relationship of the off-axis to the on-axis data that matters here. Therefore we must view the off-axis response in the relative context of a known on-axis response.

The other key mic position is the spatial crossover to any adjoining system. What can we learn there? The spatial crossover will inform us about speaker positions, since it is the meeting point. We will compare the level at the respective on-axis locations to the level at the spatial crossover and adjust angles to achieve the best line of minimum level variance from the first on-axis point, through the crossover to the second on-axis location.

The certainty of change over the space in the amplitude response is matched by the same certainty of change in phase response. Any two speakers can only stay matched in relative time (phase) over very limited lines of minimum temporal variance. Our delay alignment settings are just as prone to error over the space as were our equalization settings. Is there a best location for delay setting? Yes — the spatial crossover (point of equal level).

Why? We know that the crossover area has the highest rate of change in ripple variance. Even tiny movements here can have more than 20 dB effects on the frequency response. This is the poorest position possible for equalization decisions. It is, however, the wisest location for delay setting, for the exact same reason: the highest rate of change in the ripple variance is here. The key to confining the ripple variance damage to the smallest area is in the phase-aligned spatial crossover. If the phase is aligned at the spatial crossover, the deepest ripple will be confined to the smallest frequency range. The delay offsets will start to creep in as we move away from the spatial crossover, creating ripple variance which moves progressively downward into the lower ranges of the frequency response. The phase-aligned crossover keeps the encroachment of the spatial crossover ripple into the on-axis areas confined to a minimum frequency range before isolation comes to the rescue. The applicability of equalization over the on-axis area is maximized by the optimized delay settings in the spatial crossover area.

There are four classifications of mic positions for the calibration process. These positions have specific locations and roles and provide the data required for all of the calibration procedures.

- ONAX: this refers to an "on-axis" mic position. The ONAX position provides the data for equalization, level setting, architectural modification and speaker position. The ONAX position is found at the point of maximum isolation from neighboring speaker elements. In cases of symmetrical speaker orientation to the space, the location will indeed be "on-axis" to the element in question. In cases where the speaker is asymmetrically oriented, the ONAX mic is located at the mid-point between the coverage edges, rather than the on-axis focus point of the particular speaker. The ONAX position is found on a per element basis. For single speakers this is a simple matter. Multi-element arrays will contain ONAX mic positions for each speaker. Spatial averaging techniques such as the star pattern can be used for multiple positions inside the ONAX area, or for multiple ONAX locations in a symmetrical array. Each mic position must have some degree of isolation between the given speaker and other elements in the array to be considered in the ONAX class.
- OFFAX: the "off-axis" mic position represents the coverage edge. This position is defined by the listening space shape, not the speaker. These may or may not be at the edge of the coverage pattern of the speaker under test. The top row, the bottom row and the last seat before the side wall aisle are representative examples of OFFAX positions. These positions will be analyzed in relation to the ONAX position data. Our goal is that they be as close as possible and not more than 6 dB down (the maximum acceptable variance) from the ONAX response.
- XOVR: this mic position is for the spatial crossover point between two or more speaker elements. The spatial crossover is the point of equal level between the systems, and as a result its exact position must be found, rather than arbitrarily declared. In the case of two elements running at equal levels, the spatial crossover point will be exactly where we expect it: on the geometric center line. When asymmetric levels are applied, the spatial cross-over

point will shift toward the lower level element. The exact position is found by searching for the point of equal level between the elements. The XOVR location will be the spot where we will phase-align the spatial crossover between the systems. The process of "cross-over hunting" will be discussed later in this section.

- SYM: this mic position type refers to a location where there is a symmetrically opposite element found elsewhere in the system. Symmetrical opposites require less detail in their inquiry than originals and as a result we can save precious time. Original positions are one of the three classes above, and there maybe symmetric versions of any of them. Symmetric ONAX positions are required to, at the very least, verify the normal operation of the analogous element. Symmetric OFFAX positions are rarely needed. Assumptions are made only sparingly in this work, but if a matched speaker has matched ONAX response and matched focus, the OFFAX response Will most likely follow. The last type is the XOVR version of the symmetric mic position. This is also rare for the same reasons. If the ONAX response of physically matched components has been verified as matched, the XOVR response will follow.

Microphone Orientation Angle

The exact angle of the microphone is not critical. We have a fudge factor of ±30 degrees within which to work. The mic should simply be directed toward the speaker(s) that are in its scope of work. While the mic may be classified as omnidirectional, it is not (see Chapter 3). Microphones which are placed in acoustic crossover points (spectral or spatial) can be pointed in the direction between the two speakers.

Listening Position Height

We have developed a basic strategy for mic placement. What about the details and practical matters such as exact microphone position and height? The most intuitive assumption is to place the mic at the head height of a seated person since this most closely resembles the position of our listener's ears. Unfortunately this is rarely the position that most closely resembles the response of our listener *during a performance.* The nearby conditions of an empty seating area are quite different from the same area with an audience. The presence of unoccupied seat backs both in front and behind the designated seat create strong local reflections that will be substantially modified, if not eliminated, during performance. When faced with local variables that will change substantively, it is usually best to get some distance from them if possible, to minimize their effects. The seat back reflections are most problematic in steeply raked halls with high hard-backed seats. If the firing angle of the speaker is fairly low there will be strong reflections from the seat back in the row behind the mic. We are working under the assumption that the seats will be occupied during the performance, thereby rendering the reflective response of the seat backs inapplicable to our calibration strategies.

If the microphone is raised to standing head height the nearby seat back reflections are greatly reduced. The listening position is more "free field" than the low position. There is still plenty of room interaction in the response but there are now less of the local effects and more of the large-scale surfaces in the data. The mic has changed its orientation somewhat to the speaker vertical axis as well. In distant locations this is a negligible angular difference but can be substantial in close proximity applications. An example of this is the mic location for a frontfill speaker. In this case there are two factors that work in favor of the low position. First, the rake of the floor is so low that the seat back behind gets very little direct sound. Second, the vertical angle change to the high position is too severe. For our purposes here we will refer to the sitting head height as the "low" position and the standing head height to be the "high" mic position.

Both the high and low positions contain a floor reflection. The floor reflection is the subject of endless controversy as to what steps, if any, should be taken to mitigate its effects on our data. The options include living with it, eliminating it physically, or eliminating it from our data.

Living with it means including it in our data since it will continue to exist when the room is occupied. The extent of the acoustical changes to our floor surface will depend

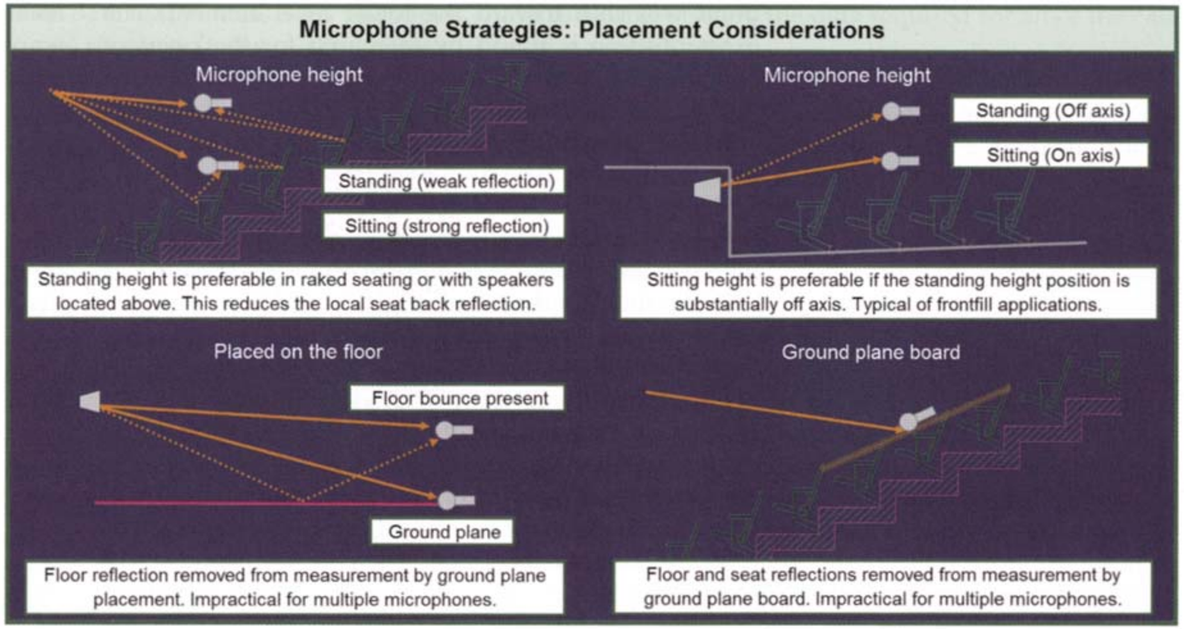


Figure 10.12 Placement considerations for measurement microphones

upon the particulars of our room. At one extreme we have classical music venues which have seats specifically designed to create the same absorption coefficient whether the seat is empty or full. On the other extreme is the hockey arena floor which has not even put its folding chairs down during our setup time so that we have a perfectly flat slab of concrete (or ice covered with plywood, brrr!). We can live with the former, but the latter response will not resemble the response when show time arrives. Most venues are in between. Most have areas that will change and areas that won't.

The ground plane mic placement technique can eliminate the floor reflection from the data. This technique is based on placing the mic directly on the ground, hence the name. Another option is to construct a ground plane, i.e. a flat reflective surface that can be placed anywhere. The intent is to place the mic at the spatial crossover of the speaker and the floor, i.e. at the floor. This gives us a coupling zone summation response with 0 dB and 0 ms offsets. We are not actually eliminating the reflection; we are measuring it at the phase-aligned crossover. Because of this, one should bear in mind that the relative level at this position will rise 6 dB. Level comparisons with other high or low mic positions will need to factor this in. There are a variety of ground plane methods. The quick and dirty method is to lay the mic on the floor and keep a lookout for forklifts. More elegant solutions include plywood sheets with clips to keep the mic in place. The size of the plane will dictate the low-frequency limit of the measurement. The principal limitation of this technique is practical. The calibration process requires either a large number of positions for a mic or a large number of mics. In either case the prospect of hoisting plywood sheets up to the balcony to get a reading is not at all inviting.

The final option is to computationally eliminate the ground reflection from our data. The deed is done by making the time window shorter than the arrival time of the floor reflection. A typical floor bounce from high mic position

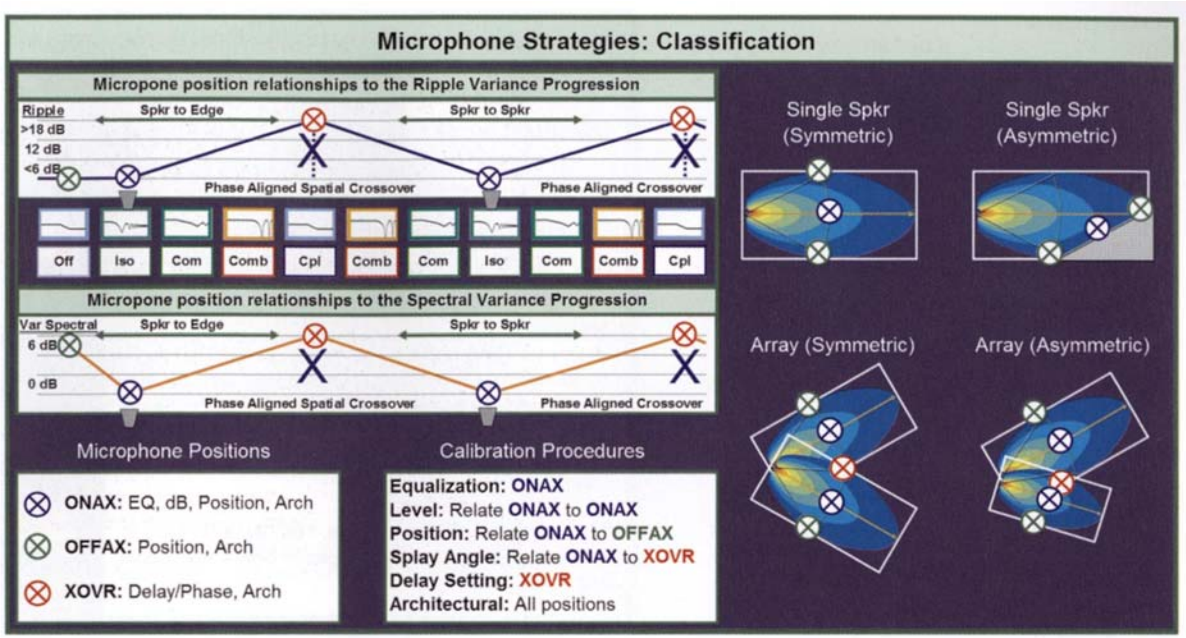


Figure 10.13 Microphone position classification, roles in the calibration

ranges from <2 ms (back of raked arena with hung speakers) to around 8 ms (ceiling speaker firing straight down). To eliminate these will require calculated anechoic responses of less than 2 ms, which means no data below 500 Hz.

Here are some guidelines to help decide whether ground plane measurement is strongly indicated:

1. Evaluate the situation from the viewpoint of the path from the speaker to the mic.
2. If the path will have minor changes between the current conditions and show conditions then no special action need be taken.
3. If the path has a local special condition, such as an aisle running right up to the mic, then move the mic the minimum distance required to reduce the local effect.
4. If the floor will change dramatically then abandon the high or low position and use the ground plane technique.

Hanging Microphones

There are instances where it is necessary to hang microphones, such as during live performance situations. The simplest method is to hang the mics straight down. This is a measure of last resort (pun intended), since at best we will receive data with substantial HF axial loss from the mic directionality. At worst we have a swinging mic that gives us no usable data in the HF range due to poor coherence.

There are, of course, duct tape solutions that can give us an on-axis orientation to the speakers. Such cases may require a tie line to prevent horizontal rotation of the mic.

Microphone Position Strategies

It does not take long to figure out that every listening position is unique. They are all important as well. The adjustments we make for one position will affect some or perhaps even all other locations. It is also clear that we will never have time to measure them all, and even if we could,

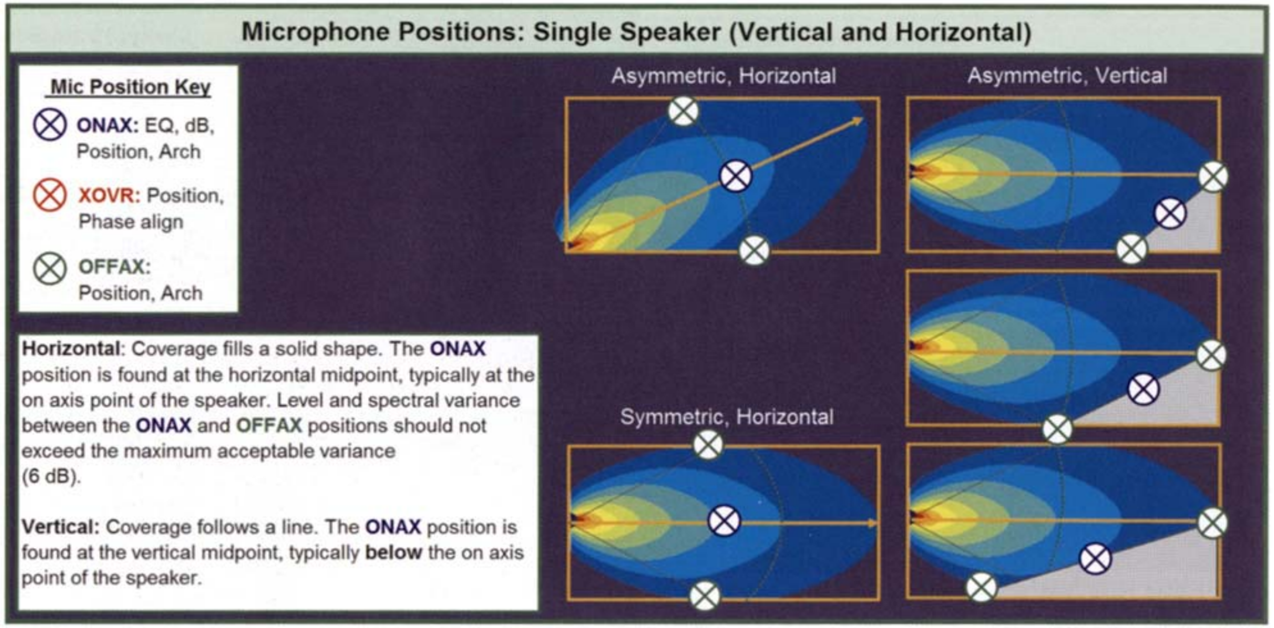


Figure 10.14 Microphone position strategies for a single speaker. Note: the actual length of the coverage pattern (start and stop depth) are not known until both the vertical and horizontal planes have been viewed

what would we do with the conflicting results? How do we decide?

The prioritization of mic positions is based on our goal of minimum variance and our strategies to achieve this. Mics are placed for specific verification and calibration procedures. There are five positions that must be known for every speaker, or speaker subsystem.

The five mic positions to characterize a single speaker:

1. ONAX: the mid-point of coverage (horizontal and vertical)
2. OFFAX: horizontal left edge
3. OFFAX: horizontal right edge
4. OFFAX: vertical top
5. OFFAX: vertical bottom

Single Speaker

Symmetric Orientation

Symmetric orientation of the speaker to the space is typical of horizontal applications as shown in Fig. 10.14. The ONAX position is found in the center of the coverage pattern width and length. The OFFAX positions are found along the coverage shape edges at a comparable distance. It is advisable, within practical reason, to keep the OFFAX mic position equidistant with the ONAX position. With matched distances, the level difference can be attributed only to axial loss. Bear in mind that an equidistant OFFAX position is usually in a different row than the on-axis mic. An equidistant orientation is easily accomplished by comparing the propagation delay times between the two points, which should be roughly equal. The OFFAX response can be anywhere from 0 to 6dB down from the center, depending upon the amount of overlap in the system design. The horizontal edges may well be symmetrical, in which case we will be able to economize positions. However, if there is asymmetry we will need to use both positions.

Asymmetric Orientation

The asymmetric orientation is typical of vertical applications (also shown in Fig. 10.14). The ONAX position is found at the mid-point of the coverage, not at the on-axis

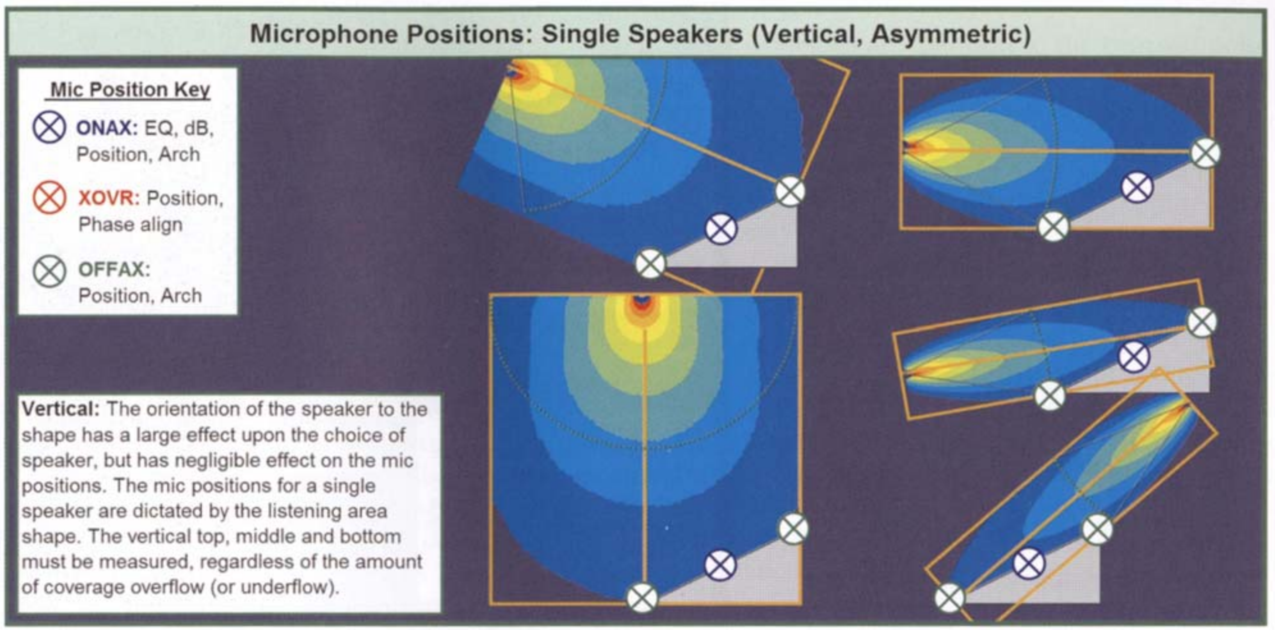


Figure 10.15 Microphone position strategies for a single speaker with an asymmetric orientation to the space

focal point of the speaker. The OFFAX positions bracket the top and bottom of the coverage shape. The upper position is actually "on-axis" to the speaker, but its primary role is as an OFFAX mic. The length of the shape affects the positions of the mic. As the shape stretches, mic positions spread proportionally until the lower limits of the speaker coverage pattern are reached (lower set). The distance between the speaker and the shape will determine the best speaker model, but the mic positions are governed by the listening area shape.

The immunity of the mic positions to the orientation of the speaker is shown in Fig. 10.15. Regardless of angular orientation the mic positions are found in conformance to the listening shape, rather than the speaker pattern.

Coupled Arrays

Line Source

The mic position strategy is limited by the inherent overlap condition of the coupled line source as shown in Fig. 10.16. Only the extreme near field areas contain sufficient isolation to have mic positions that meet the ONAX criteria. These positions are far too close to serve any practical calibration purpose. Once the second level of the parallel pyramid is reached, the overlap dominates. No position contains an isolated response which can be equalized to provide a discernible zone of minimum variance. What is left is a decreasing quantity of spatial crossover positions of increasing complexity. The multi-overlapped spatial crossover cannot be phase-aligned for more than a single point in the space.

Point Source

The microphone position classes are clearly differentiated in the point source array as shown in Fig. 10.17. The ONAX position is found on-axis to each element in the symmetric version. Equalization operations will be performed at this location as will level setting and architectural evaluation. This will also serve as the base point from which to

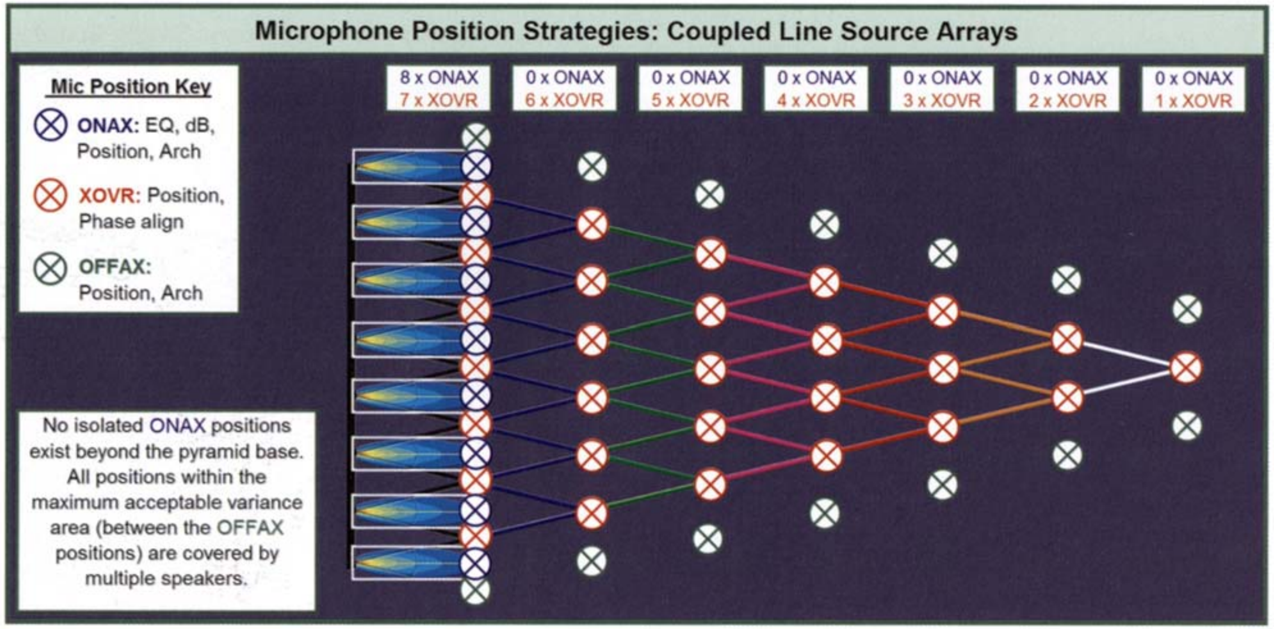
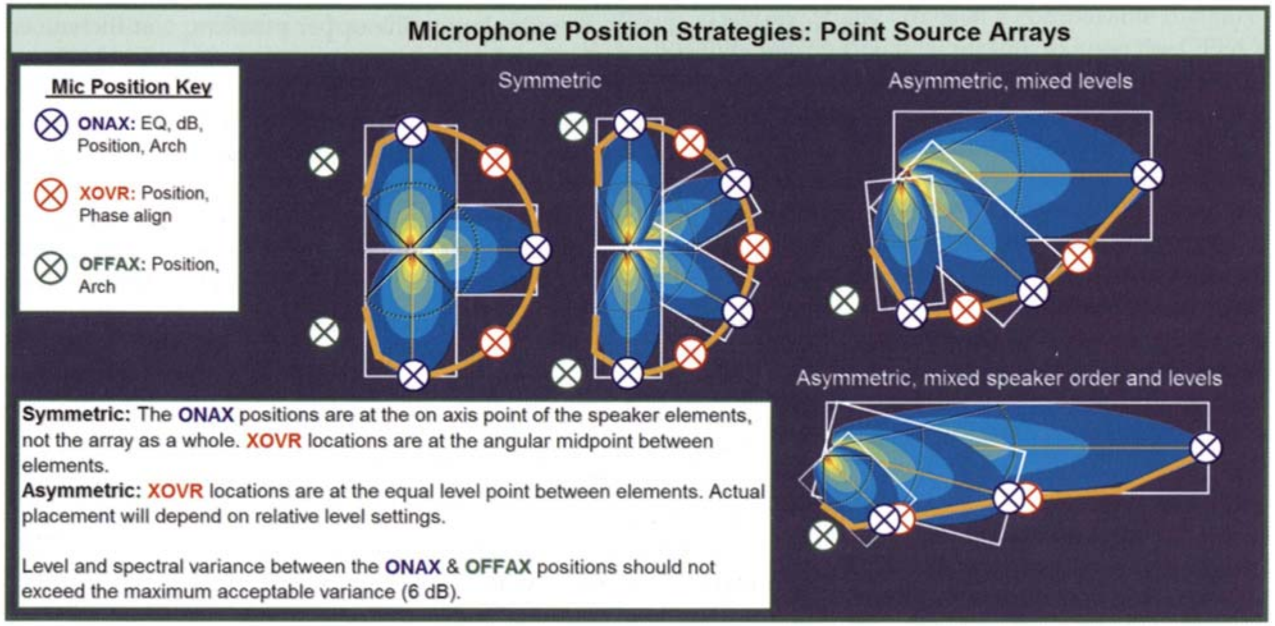


Figure 10.16 Microphone position strategies for the coupled line source



Microphone position strategies for the coupled point source array. Symmetric version is typical of horizontal applications. Asymmetric is typical of vertical applications compare our data from the other mic positions. For arrays with more than two elements, there will be multiple ONAX positions which contain unique data; i.e. not symmetrical opposites. The three- and four-element arrays shown here contain two unique ONAX positions. The multiple ONAX positions can be observed by the optical spatial averaging technique and an equalization curve selected that best serves the overall array.

The XOVR positions will reveal the amount of overlap between the elements. The combined level at XOVR will match the ONAX position if we have implemented the unity class crossover (unity splay angle and matched level). Levels above or below the ONAX reference data indicate overlap or gap crossover classes respectively. The combined data of the ONAX and XOVR microphones will determine the position (in this case the splay angle) of the array elements. The OFFAX position will also determine speaker position by comparison to the ONAX traces. If the OFFAX position is more than 6 dB down from ONAX, the speaker positions will need modification to move us within the maximum acceptable variance. As the OFFAX value approaches unity level to the ONAX we may consider the possibility of repositioning the speakers to obtain less level at OFFAX. This need only be done where nearby surfaces will create strong reflections into the listening areas.

Asymmetric arrays require a different approach in all mic location types, but most notably in the XOVR and OFFAX positions. Spatial averaging of the multiple ONAX positions is no longer applicable when the coverage distances are substantially different. This is not to say that multiple ONAX positions cannot be spatially averaged for a single section, but rather that the sections with substantially different levels must be analyzed and calibrated separately.

Unique ONAX positions are found for each element in asymmetric situations. The level is set at the ONAX location for each element, thereby creating the desired equal level contour. Separate equalization is required for each of the elements as measured at their respective ONAX locations.

Point Destination

Mic position strategies are identical to the coupled point source. This array is so rarely used in its coupled form that our discussion will focus exclusively on the uncoupled version (see below).

Uncoupled Arrays Line Source

Just as the speakers are arrayed in a line, so are the mic positions, as shown in Fig. 10.18. The ONAX position is found at the literal on-axis point of the speaker and at the 50 per cent point in coverage depth. This location will give us the usual data for operations such as equalization, level setting and architectural issues. This location also gives us the reference level for the remaining positions. Each element will require an ONAX microphone. For fully symmetric systems (matched models, level and spacing) the additional ONAX positions are needed only for spatial averaging and verification purposes.

The XOVR location is found at the equal level point between the elements. In level symmetric systems this will be the geometric mid-point. There is no need for further phase alignment at the spatial crossover since the systems should already be synchronous in such systems.

The XOVR positions will reveal the amount of overlap between the elements. The combined level at XOVR will match the ONAX position if we have implemented the unity class crossover (unity aspect ratio spacing and matched level). Combined levels above or below the ONAX references indicate overlap (speakers too close) or gap (speakers too wide) crossover classes respectively. The combined data of the ONAX and XOVR microphones will determine the position (in this case the spacing) of the array elements. The OFFAX position will also determine speaker position by comparison to the ONAX traces. If the OFFAX position is more than 6 dB down from ONAX, the speaker positions will need modification to move us within the maximum acceptable variance.

The calibration steps for asymmetric systems must be carried out in a particular order of operations. The ONAX

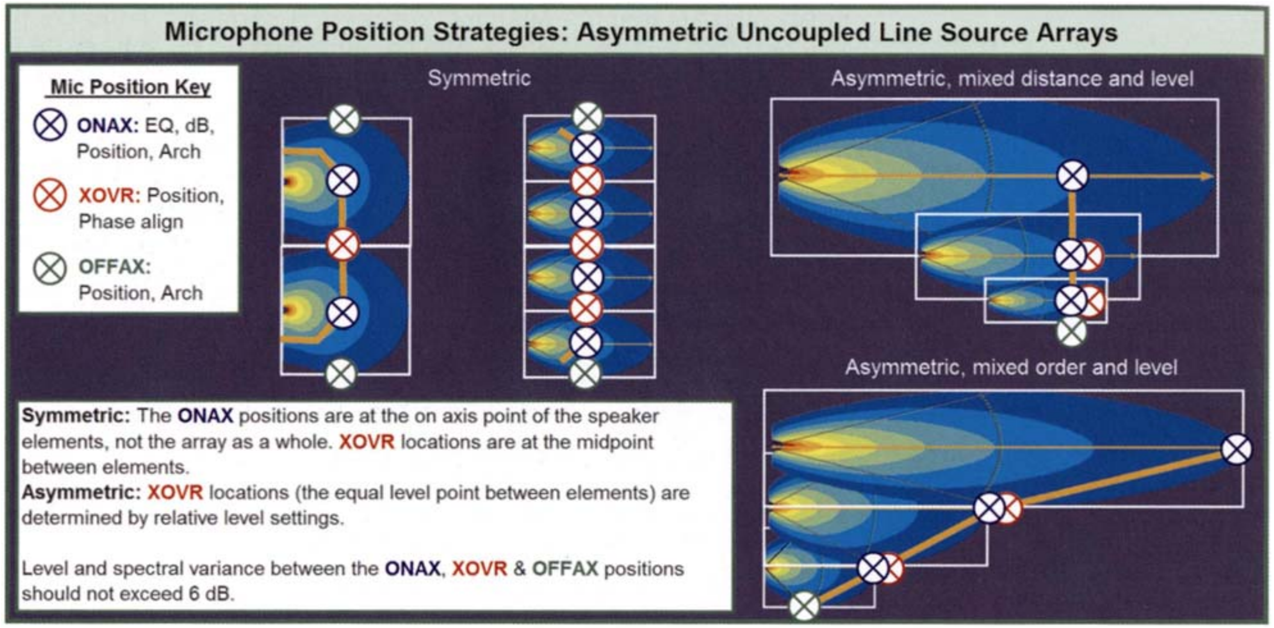


Figure 10.18 Microphone position strategies for the uncoupled line source array. The symmetric version is typical of horizontal applications but can be found in vertical applications in small quantities. The asymmetric version can be found in both vertical and horizontal applications

positions for each system must be measured first. Equalization, level and architectural evaluation must be completed before spatial crossover alignment can commence. In asymmetric systems the XOVR location will not be found at the geometric center between the elements but will have encroached in the direction of the lower level speaker. The spatial crossover will need to be hunted and found, and the time offset compensated to complete the phase alignment. The XOVR mic is moved between the adjacent ONAX positions until the point of equal individual levels is found. Then the XOVR position is used to classify the amount of crossover overlap and to optimize the phase alignment. There is now a linkage from the first position ONAX to the second, through a phase-aligned spatial crossover at the XOVR location.

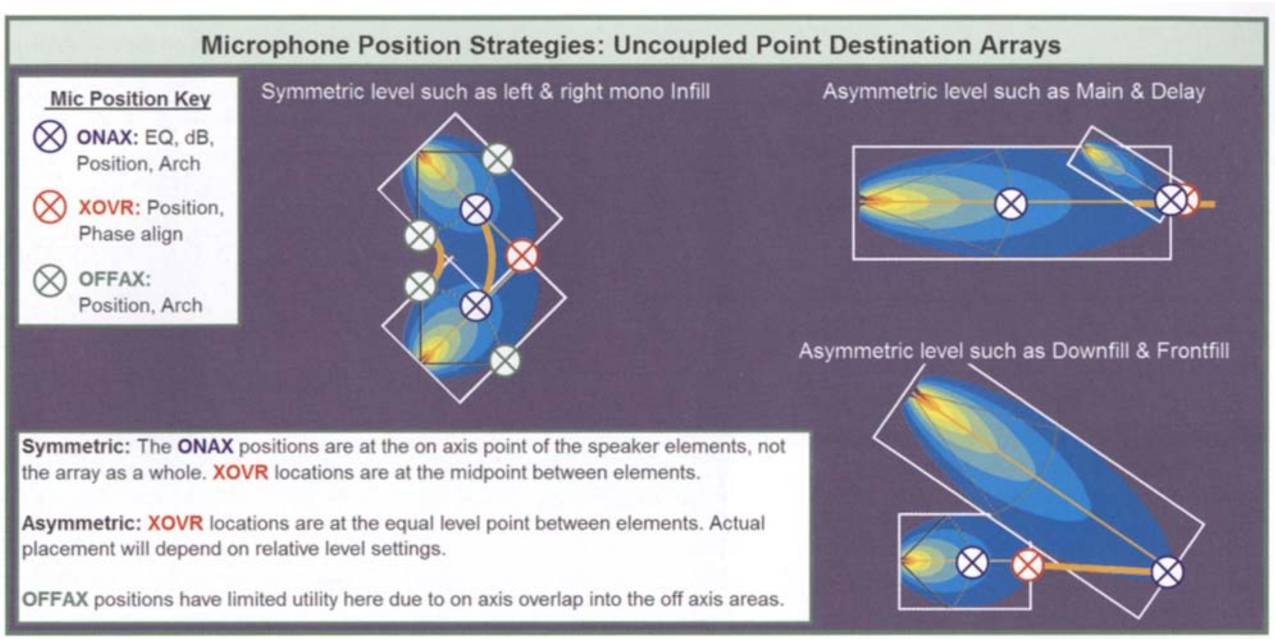
Point Source

The uncoupled point source is the logical hybrid between the coupled point source and the uncoupled line source. Add spacing to the former or angle to the latter and we will arrive at the same place. The mic locations and their roles in the symmetric version have the angular relationship of the coupled point source and the depth of field relationship of the uncouple line source; i.e. 50 per cent of the depth of coverage.

The asymmetric version is also closely related. The mic positions and their roles follow the coupled scenario, with the depth of coverage issue added.

Point Destination

The symmetric version is often found as an infill array to cover central listening areas from side locations. The seemingly obvious mic location would be the point where the speaker patterns meet in the middle. While there is no arguing that this location is literally on-axis to both speakers, it does not classify as an ONAX location due to 0 per cent isolation. That spot is also the XOVR location and is deemed the end of the coverage due to the high ripple



Microphone position strategies for the uncoupled point destination array. The symmetric version is typical of horizontal applications but can be found in vertical application in small quantities. The asymmetric version can be found in both vertical and horizontal applications

variance from this point onward. The usable ONAX points are found along the on-axis line at 50 per cent of the depth of the XOVR location. Equalization, level setting and architectural evaluation are done as usual at the ONAX location. The ONAX level data is used as the reference for the OFFAX positions which are found by radial offset from the ONAX position.

The asymmetric version also ends (for the lower-level speaker) at the XOVR mic position. A representative example is the main/delay system where both systems focus into the same area. The ONAX and OFFAX positions of the level dominant main system are treated like a single speaker with asymmetric orientation. The ONAX position for the delay speaker will be used to set equalization, etc., and to verify the position. In many cases the delay speaker XOVR and ONAX positions are the same. The equalization and phase-aligned spatial crossover are all set at the same location, the mid-point of the delayed system's coverage depth.

Another representative field example is the dominant main downfill system merging with the frontfill. The frontfills are intended to cover only the first few rows so as to prevent excessive ripple variance from the overlap of multiple spatial crossovers. The downfill focus will be asymmetrically oriented to the space and therefore may be beyond the desired frontfill edge. Nonetheless its level dominance allows the downfill to take over the coverage. The XOVR mic position should be the last row of desired frontfill coverage. This is the location for creating a phase-aligned spatial crossover.

Subdivision Strategies

We have just seen a mic placement strategy for single speakers and each of the standard array configurations. There is no limit to the number of elements we might encounter in a single cluster or to the number of subsystem arrays that we will need to weave together into a single fabric. How do we go about ensuring that we have covered all of the positions we will need in order to perform the full set of calibration and verification procedures? Fortunately we will not have to go through the full gamut of array types yet

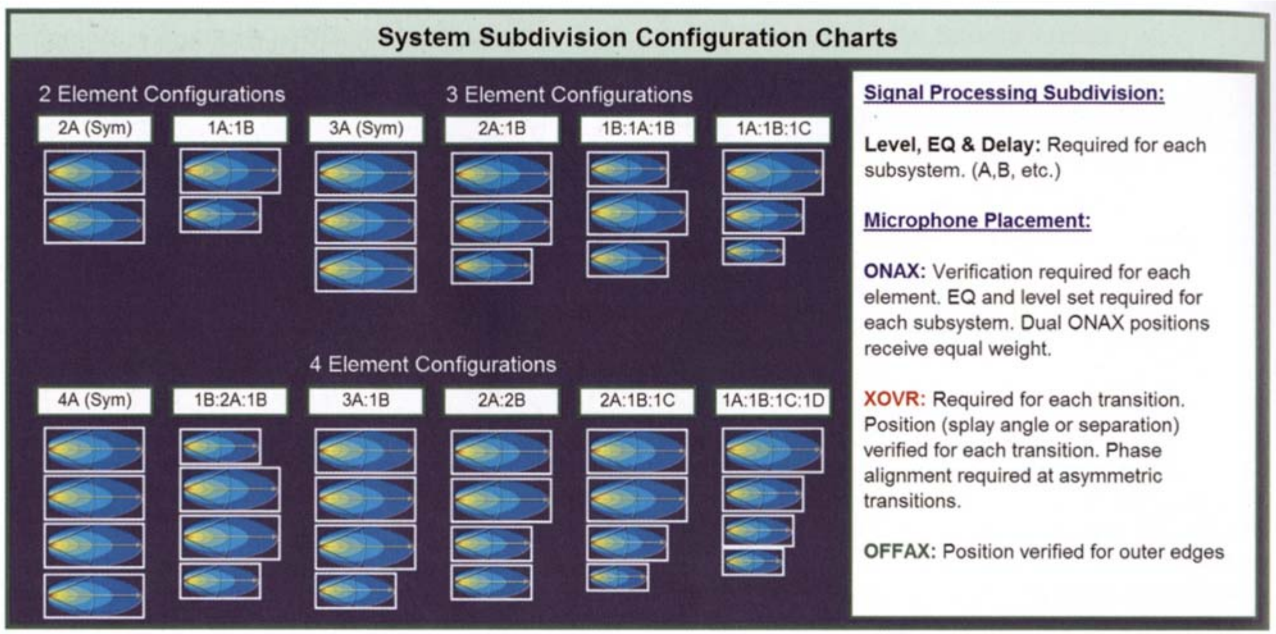


Figure 10.20 Guide to microphone position and system subdivision strategies

again. Instead the answer lies in the quantities of elements and the number of asymmetric transitions. This is the case regardless of the array type, so a single set of logical subdivisions can be employed.

Refer to Fig. 10.20. The criteria for subdivision groupings are shown. The criteria are straightforward: matched elements are grouped together. An unmatched element begins a second group and so on. This illustration shows the matching in terms of level (shown as the size of the aspect ratio icon). Other forms of asymmetry such as speaker order and splay angle would have the same effect on calibration subdivision grouping. It's as simple as this: if you're different, we are going to treat you different. Wherever a unique transition occurs we must be able to provide a unique calibration for optimization.

Procedures

At long last we are ready to perform the calibration procedures. Our exhaustive preparation work will pay us back with the elegant simplicity of these procedures. Our mic positions will guide the process through the order of operations until all of the subsystems are brought together as a single entity. For the next channel, we will do it all again, until all are complete.

Acoustic Evaluation

The room reflections will introduce spectral and ripple variance throughout the space. Our foremost strategy for minimizing these effects is avoidance. This has its practical limits, and therefore we will need to be prepared to identify problems which are caused by these reflections. Once identified, there are a number of possible options such as acoustic treatment, speaker repositioning, relative level adjustment, adding fill speakers or surrender. The treatment option will be the most effective, but as depressing as surrender sounds, it is better than fighting a battle we can't win. Absorptive treatment will almost assuredly reduce variance in all categories and will do so in large portions of the listening area. This is a win-win situation. In such

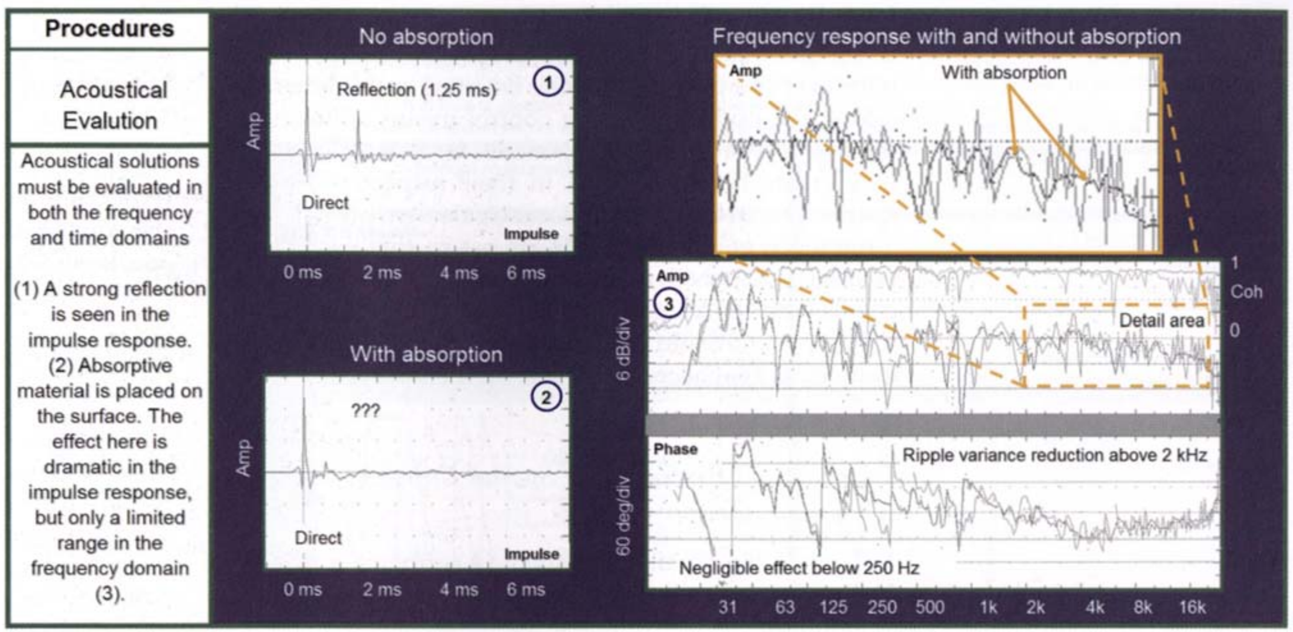


Figure 10.21 Acoustical evaluation considerations

***Perspectives:** When aligning a sound system, before you start EQing make sure the cancellations in the frequency response are not caused by reflections, poor speaker combination or other acoustical phenomena. EQing comb filters simply doesn't work and the only solution to get an even response is to eliminate the source of the reflection/interaction, either by treating the venue or by redirecting/-time aligning the PA system. Looking into the impulse response is a very powerful way to realize what your echo structure looks like and thus assessing your problematic reflective surfaces or*

cases there is little need to be concerned about the details of exactly how much change occurs at each location. If the ripple is reduced, the coherence goes up, and the amount of equalization that would be indicated is reduced. All of this is good news.

The other available options are not so clear-cut. Speaker repositioning in order to reduce reflections will most likely put TANSTAAFL and triage scenarios into play. An example is found in the case of a speaker aimed at the upper seats in the back of the hall. If the optimum speaker focus angle for minimum level variance to the back of the hall causes strong roof reflections, then we are in a quandary. If we reduce the vertical focus angle we can reduce the reflections in the upper area but we are also reducing the direct sound. The chance of a net loss in direct-to-reverberant ratio is as good as the chance of an improvement in the upper area. Meanwhile down on the floor, the angle change has no discernible effect on the direct sound (since such effects are off-axis) but the reduction in reflections is noticeable. Positions on the floor will benefit from less ceiling reflections. In fact, the patrons on the floor would benefit from having the upper speakers turned completely off! There is no simple answer here other than hindsight redesign, which increases the directional control of the upper mains and adds supplemental delays.

Therefore, acoustic evaluations beyond the absorption option will require us to monitor the effects at multiple positions. Those positions would be our ONAX milestones for starters, as those provide us the clearest view of the system. If we are going to attempt to evaluate the acoustics with the maximum degree of isolation, we will want to be at the lowest points in the ripple variance between speaker elements. That point is the ONAX position.

Level Setting

The verification process will have previously established that our gain structure is sufficient to allow us to drive the system to full level. The role of level setting in the calibration process is relative level between subsystems. One of *poor time alignments. Always remember that floor reflections will go away once the venue is packed with people.*

Miguel Lourtie

the key concepts to relative level setting is to take a proactive rather than reactive role. The speakers will be brought into level compliance with the listening area, rather than asking the opposite. Taking control means adjusting the relative levels of isolated elements in the system so that they create the same level in their respective listening spaces. The ONAX position for each element provides such a reference point. The levels are set so that each ONAX position is matched. A line of minimum level variance will connect the ONAX positions. The line will run through the XOVR positions and will match the ONAX positions if the unity class spatial crossover has been employed.

Level Setting Procedure for Spatial Crossover Alignment

1. Turn on the dominant element (A) only.
2. The level at the ONAX A position is the reference standard.
3. Turn on the secondary system (B) only.
4. Set the level for B at ONAX B to match the level reference.
5. Continue with all related subsystems to match the level reference at their ONAX positions.

If the combined levels at the XOVR positions do not match the combined levels at the ONAX positions, refer to the speaker position adjustment procedures.

Level Setting Procedure for Spectral Crossover Alignment

The relative level of the LF and HF systems will be set to meet at a specified spectral crossover frequency. The matter of which element to turn up (or down) and the order of which is measured first is left to the reader's discretion. Once the levels are set we can add delay as required to complete the phase-aligned crossover (see the delay setting procedures in this chapter).

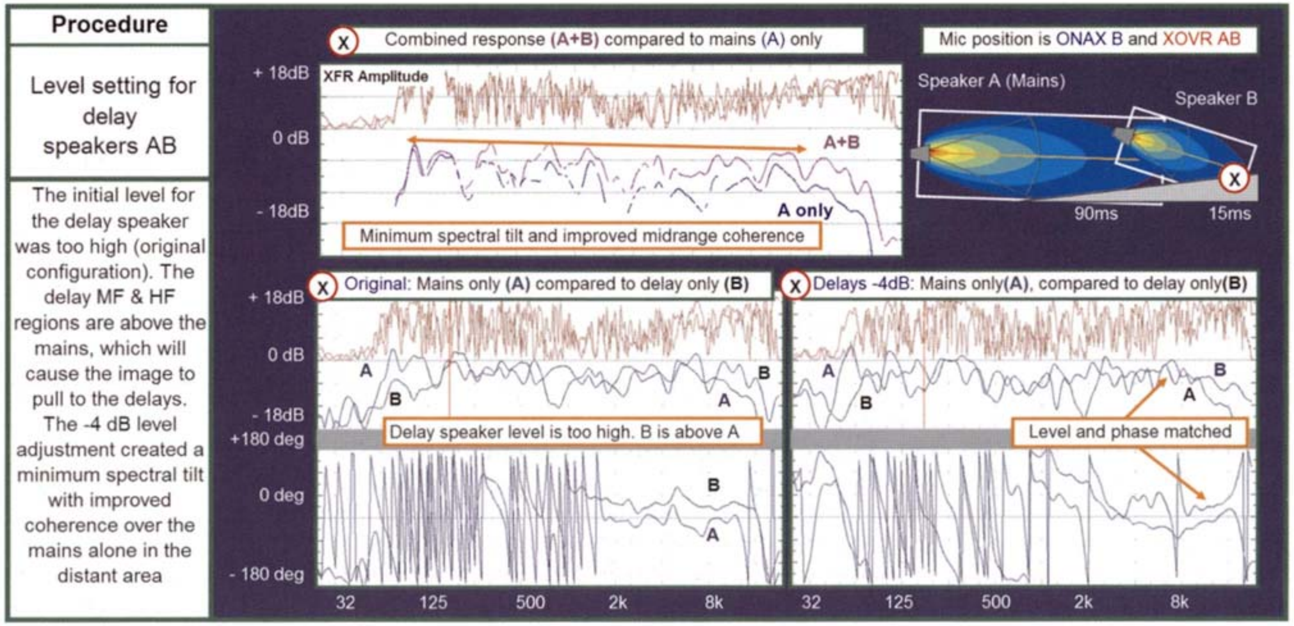


Figure 10.22 An asymmetric combination (A + **B)**

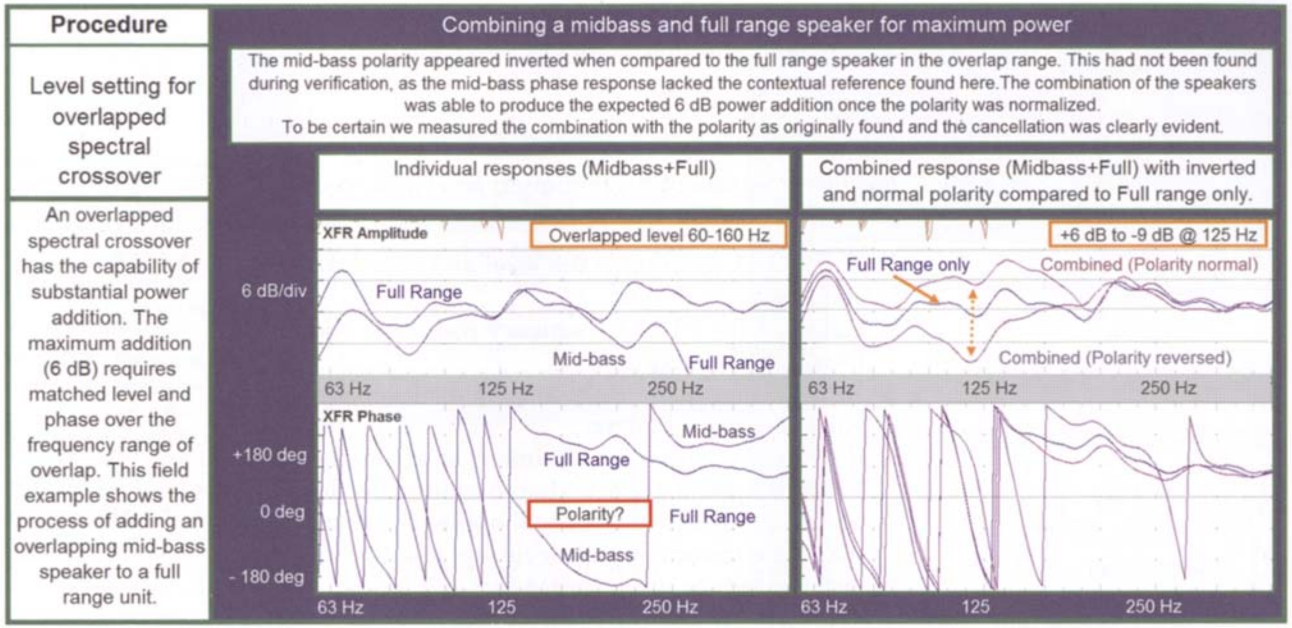


Figure 10.23 A mid-bass spectral crossover

1. Both systems must have the matched drive levels to the measurement reference point.
2. Turn on the HF only. Store the frequency response and recall the trace. Place the cursor on the desired spectral crossover frequency.
3. Turn on the LF only. Without changing the compensation delay in the analyzer, adjust the level control of the LF system until the amplitude responses meet and the desired spectral crossover frequency is found. Store the response and recall the trace.
4. The relative phase responses can be observed. If they do not match around the crossover frequency a phase adjustment will be required (see the delay setting procedures in this chapter).
5. Combine the LF and HF regions. The expected result is combined response addition above either of the individual levels.

Important note: the relative levels of transfer function responses rely on two important known quantities: the matched source levels of the subsystems and matched microphones on the reception end. The verification procedures for these parameters can be found in Chapter 9.

Speaker Position Adjustment

Speaker position adjustment seeks to minimize level, spectral and ripple variance in the key relationships between speakers and the room. The process proceeds in layers of added complexity beginning with the single element and finally encompassing entire arrays and the room. The actions at each layer are discernible to us through the mic position strategies. The pattern emerges by comparison of the ONAX, OFFAX and XOVR positions. The speaker position adjustment will be complete when the relationship between these has reached minimum variance.

Speaker position layers:

1. a single element in the room
2. element to element inside an array
3. array to the room

4. array to array
5. combined multiple arrays to the room

Position Adjustment Procedure A Single Element

Position adjustment for minimum level and spectral variance, single element in the room:

1. The ONAX position is the level and spectral reference standard.
2. Compare the OFFAX frequency responses and level to the ONAX standard.
3. Adjust the speaker position until the level and spectral variance are minimized.
4. If the best position for minimum level variance causes a rise in ripple variance (due to room reflections), then acoustic triage principles will apply. The final position may have to compromise level and spectral variance in order to reduce ripple variance.

Position Adjustment Procedure AA

For symmetric elements within an array (or the symmetric combination of arrays):

1. The ONAX position for one element is the level and spectral reference standard.
2. Compare the XOVR frequency response and level to the ONAX standard. The individual responses should be —6 dB in the isolated frequency range of the elements (a unity class spatial crossover). The combined response at XOVR should match the isolated ONAX standard in the isolated frequency range.
3. The combined response at the ONAX position will show increased level in the non-isolated range. The combined ONAX response can be used as a new standard and compared to the combined XOVR response.
4. Equalization may be applied to reduce the spectral tilt caused by the overlap in the LF range.

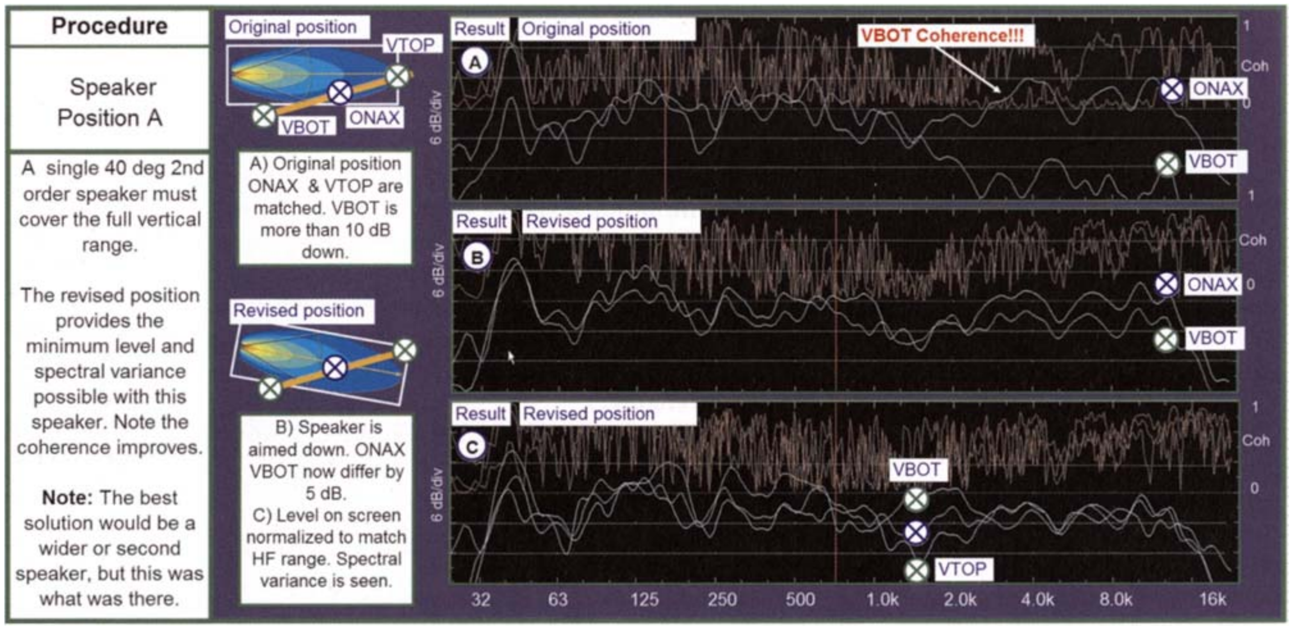


Figure 10.24 Field example of the procedure for a single speaker position

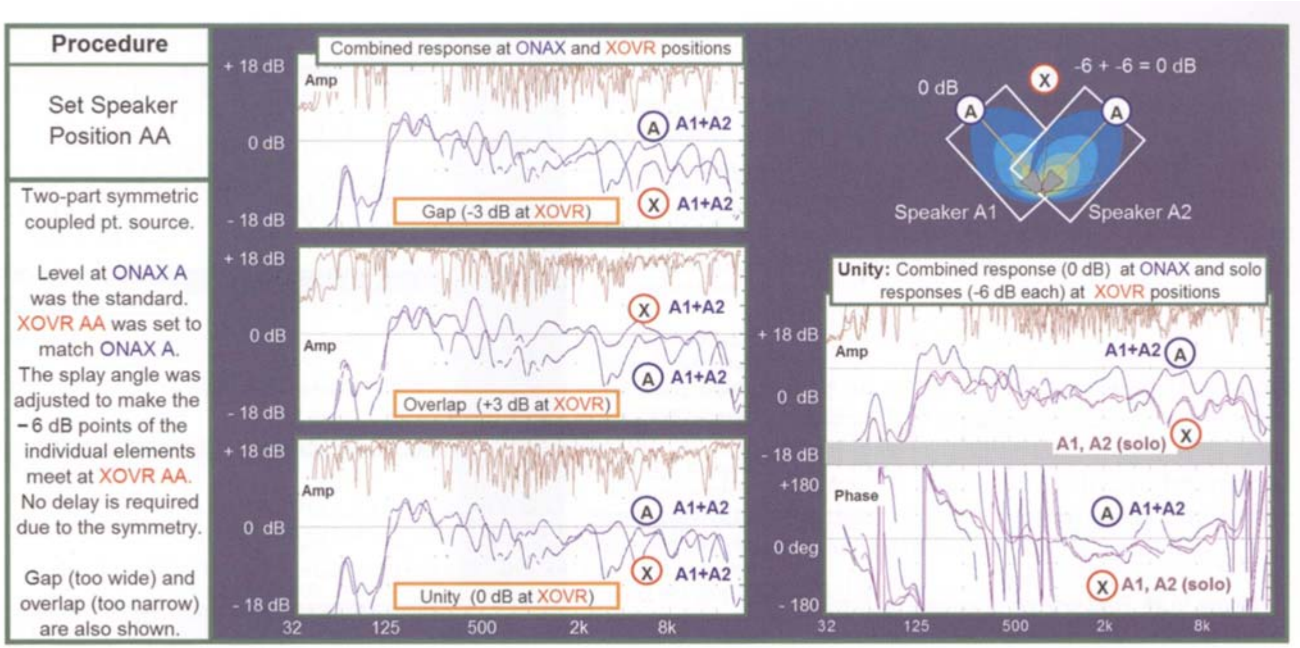


Figure 10,25 Field example of the procedures for a two part symmetric uncoupled point source

Position Adjustment Procedure AB

For asymmetric elements within an array (or the asymmetric combination of arrays):

1. The ONAX position for the dominant element (A) is the level and spectral reference standard.
2. The level and equalization for the secondary system (B) have been set at its ONAX position to match the level and spectral reference standard.
3. Search and find the XOVR position. The individual responses will be matched in the isolated frequency range of the elements, but their levels relative to the ONAX standard will not necessarily be 6 dB down. This is overlap-dependent. The non-isolated frequency range (presumably the LF range) will be stronger from the dominant element. The proportions of isolated and shared response ranges are dependent upon the amount of asymmetry between the subsystems.
4. Combine the systems and store and recall the new response.
5. New responses must now be acquired at the two ONAX positions, since the combination will affect all locations (at least in the LF range).
6. Compare the combined XOVR frequency response and level to the new combined ONAX A standard and the ONAX B response. The combined response at XOVR AB should match the combined responses at the two ONAX positions.
7. Equalization may be sparingly applied to reduce the spectral tilt caused by the overlap in the LF range. Since the level relationship is asymmetric so shall be the spectral tilt. Equalization will be most effective on the dominant system (A) but this system is least in need of additional equalization (due to its level dominance). Cut equalization in the secondary speaker will be ineffective if the element is already level dominated.

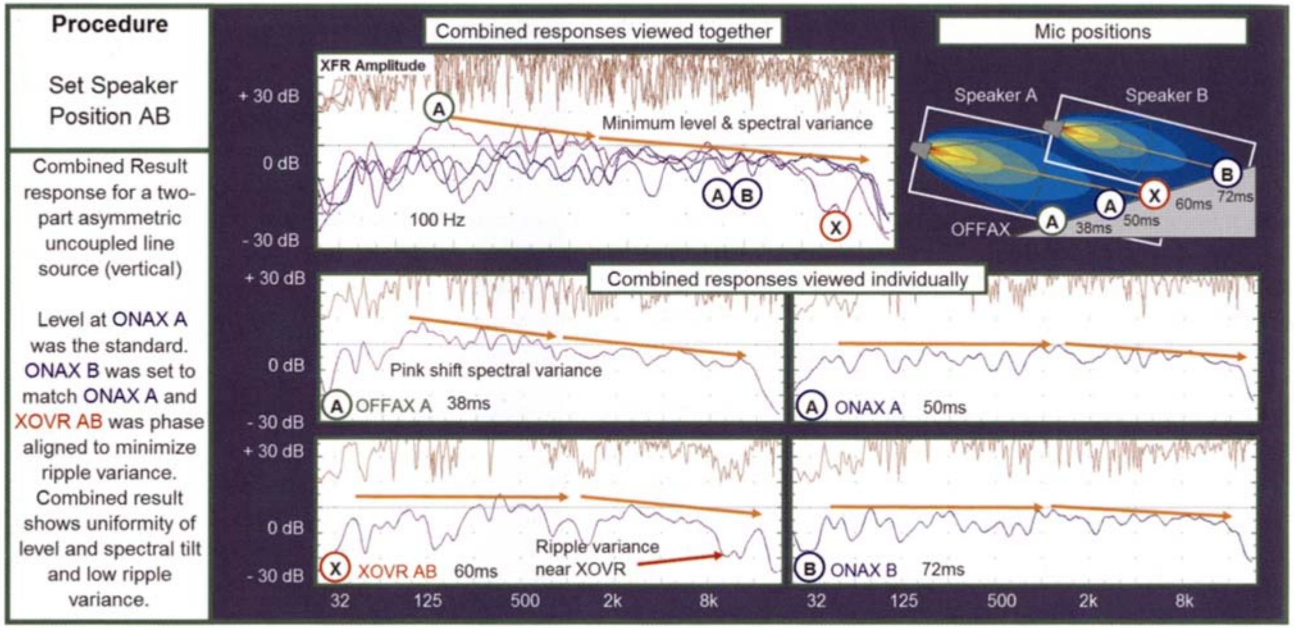


Figure 10.26 Field example of speaker positioning with an asymmetric combination

Example

The remaining levels of speaker position adjustment are simply scaled versions of these three scenarios.

Let's consider an example system with a center cluster and a frontfill system. The cluster is a three-section asymmetric coupled point source in the vertical plane. It is a two-element symmetric point source in the horizontal plane. The first step is to adjust the position of the longest throw section (the top) to the room.

Vertical:

1. Use Procedure A to adjust the top section to the room.
2. Use Procedure AB to adjust the middle section to the top section.
3. Use Procedure AB (again) to adjust the lower section (C) to the combination of the top and middle sections.

Horizontal:

1. Use Procedure AA to adjust the splay angle between the elements.
2. Use Procedure A to check the coverage of the outermost elements with the room.

The frontfills are an eight-element symmetric uncoupled line source.

1. Use Procedure AA to adjust the spacing between the elements.
2. Use Procedure A to check the coverage of the outermost elements to the room.

How do we combine the main array with the frontfills? The main array elements, now combined together, become a single element. This is also the case for the frontfills. Procedure AB would be employed for this two-element asymmetric application.

Equalization

Equalization is a simple process. Just turn the knobs on an equalizer, and the job is done. In fact equalization is sometimes done without even touching the equalizer. All that is required is announcing that the system is equalized.

Equalization has a unique emotional position in the calibration landscape. Because equalization provides a key component to the perceived tonal response of the system it is a subject near and dear to the mix engineer. In actual fact all of the other calibration parameters will play important parts in the tone at the mix position, but none will be scrutinized as closely as the equalization.

Equalization is home of great slogans and truisms:

- "The best EQ is no EQ"
- "He who EQ least, EQ best"

For all the derogatory remarks about equalization you would think we would start seeing technical riders that specify: "No equalizers will be allowed on this system." But we don't. They are always there, and always will be. They are a tool and we have the need. They are often misapplied and this has a lot to do with the bad rap. Equalization cannot make it sound good everywhere in the room. But it can make it sound bad. Our hope is to use this tool for the job for which it is really designed: to make it sound equal.

The Role of Equalization

Here is what we want the equalizer to do.

1. Control of the overall spectral tilt of the system: the overall management of the pink shift caused by summation and air loss effects. This is subject to the artistic discretion of the operator. As spectral tilt increases the listener's sonic perspective is made more distant. This is a global parameter for the entire system.
2. Control of spectral variance: the management of spectral tilt of each subsystem in order to minimize the differences in tilt throughout the listening space. Different equalization curves will be applied to the subsystems in order to bring them into compliance with the artistically desired spectral tilt. This is carried out on a subsystem-by-subsystem basis, in order to create a unified global effect.
3. The reduction of speaker/speaker ripple variance within the subsystem driven by a particular equalizer. This is carried out on a subsystem-by-subsystem basis, in order to create a unified global effect.
4. The reduction of speaker/room ripple variance within the subsystem driven by a particular equalizer. This is carried out on a subsystem-by-subsystem basis, in order to create a unified global effect.

Limitations of Equalization

Equalization affects all areas of the speaker system's coverage in the same way. It is a global solution, and is most effective when facing comparable challenges. It is least effective in the face of widespread local differences. Spectral tilt is the most global of the frequency response modifications, while ripple variance is the most local.

The range of equalization:

1. The range where an equalization filter can remain effective in the reduction of spectral tilt is practically unlimited.
2. An equalization filter cannot reduce spectral variance of a single device. That is governed by the speaker's beamwidth and its position in the space.
3. An equalization filter can reduce the spectral variance between two devices by applying separate filters which match their spectral tilts.
4. The spatial area where an equalization filter can remain effective in the reduction of ripple variance is inversely proportional to frequency.
5. The spatial area where an equalization filter can remain effective in the reduction of ripple variance is inversely proportional to bandwidth.

An example system will help us to consider the implications of these limitations. A system has eight speakers and eight parametric equalizers with five bands each. We can equalize the system as a single feed with all 40 parametric filters available. We will be able to equalize almost all of the ripple variance out of the system with 40 precisely placed filters. The only catch is that such adjustments will only work for one single spot. (Hmm... the mix position?) For that spot we have created the perfect key that decodes the ripple variance. For all other locations, our ripple decoder key does not fit. The key that fixes the mix position increases the ripple variance everywhere else. This is a global "solution" to a local problem. TANSTAAFL: the fix for the mix costs everyone else. Acoustic triage: we used up the entire blood bank for one patient. In the end, all of this effort yields very little. The mix position is artistically self-calibrating, so the system equalization is primarily for the convenience of the mix engineer by reducing the amount of equalization on the individual channels. The issue of primary importance, the difference between the mix area and the other seats, was not changed in the slightest by any of this. That difference would remain if the equalizer were bypassed. That difference is the *real* key.

Now let's rewire the system and separate the equalizers so that each speaker has their own dedicated five-band unit. The same amount of equalization filters can be employed, but in this case they are dedicated to solving the differences between the systems. Those are handled locally, with each equalizer taking care of the most prominent spectral trends from room/speaker summation in their local area. Once each system has been locally neutralized, the amount of ripple variance remaining will be a product of the amount of overlapping speaker summation (which can now be equalized) and the remaining room/speaker effects. If the arrays have been well-designed we are well on our way to minimum variance.

A sound system which has only a single equalizer (or stereo pair) must suppress all spectral and ripple variance by means of speaker position and relative level alone. Systems with a single equalizer and a single level setting, such as the wall of sound and the modern line of sound, must do this entirely by speaker position. If this cannot be accomplished then calibration monarchy must be declared and the variance suppressed by royal decree.

The Equalization Process

There are two basic forms of equalization: single system and combined system. In our example terminology we have equalization A, equalization B and equalization AB. Single system equalization is performed on every subsystem alone before any combination takes place. The focus of the equalization is on the symmetrical speaker/summation within the local subsystem and speaker/room summation in the local area. Spectral tilt due to the HF transmission loss can also be compensated.

Equalization Procedure A

For each single speaker system (A, B, C . . .):

1. Make a transfer function measurement of the room/speaker system at the ONAX position for use as a reference.
2. Measure the equalizer transfer function and adjust the equalizer to create an inverse response.
3. Measure the result transfer function and verify the response.

Combined system equalization is more complex. The speaker/speaker summation between two subsystems is the primary focus. The complication results from the fact that each equalizer affects only one of the two (or more) subsystems that are contributing to the combined response. The equalization is still carried out in the respective ONAX locations. In the low-frequency range the two systems may overlap strongly into each other's ONAX locations. This invasion into the other system's coverage renders them each incapable of independent equalization control. A 6 dB cut in one of the equalizers will have no more than a 3 dB effect on the combined response. An additional 6 dB cut will have almost no effect at all. Why? Because one speaker is now 12 dB down from the other, pushing it into the isolation zone. Equalization for shared response areas will need to be carried out equally in both electrical contributors (the equalizers) to create the expected acoustic result.

As usual we will have two versions of combination: symmetric and asymmetric. For symmetric versions the settings are typically the same for both equalizers.

In asymmetric applications the systems contribute unequally to the combination. In the low-frequency range the louder of the two systems (A) is likely to be dominant

Perspectives: When I started out doing system optimization using parametric equalizers and transfer function measurement techniques I was using a lot of equalization, both in overall cut and in the number of frequencies I was twiddling with. Over the years I have found that simpler is better. I will try to fix something with one filter and only add another when it is clear there is no way to do it with one.

Alexander Yuill-Thornton II (Thorny)

in both ONAX positions. The combined response at ONAX A will be less affected by the overlap than the response at ONAX B. The extent of the differences will depend upon the amount of level asymmetry and isolation between the systems. The level asymmetry renders the secondary system (B) fairly defenseless to the dominant system's contributions. We can cut and cut all day long on a dominated system and have virtually no effect in the reduction of contamination from the dominant system. In fact we are actually worsening the situation by removing our local system from the equation. The best approach in asymmetric situations is to probe and see the effects. If we can see a substantial effect for our equalization changes we are still in the game. If the equalizer knob turns but nothing happens, it would be wise to leave it alone.

Equalization Procedure AA

For each combination of speaker system (symmetric):

1. Compare the combined results at both ONAX positions to the individual results recorded before the combination. In most cases the spectral tilt will rise in the LF range due to overlap.
2. Adjust both equalizers to enact the changes required to restore the spectral tilt to the previous level (if desired).

Note: related symmetric subsystems do not require separate equalizers. The splitting of symmetric systems is a case of over-zealous subdivision. There are cases where systems have dual use and contain a mixture of related and unrelated signals from different matrix outputs. Such subsystems would be aligned as an AA relationship.

Equalization Procedure AB

For each combination of speaker systems (asymmetric):

1. Compare the combined results at both ONAX positions to the individual results recorded before the combination. In most cases the spectral tilt will rise asymmetrically in the LF range due to asymmetric overlap. The dominant system (A) will have only minimal changes, while lower level system will see substantial pink shift.

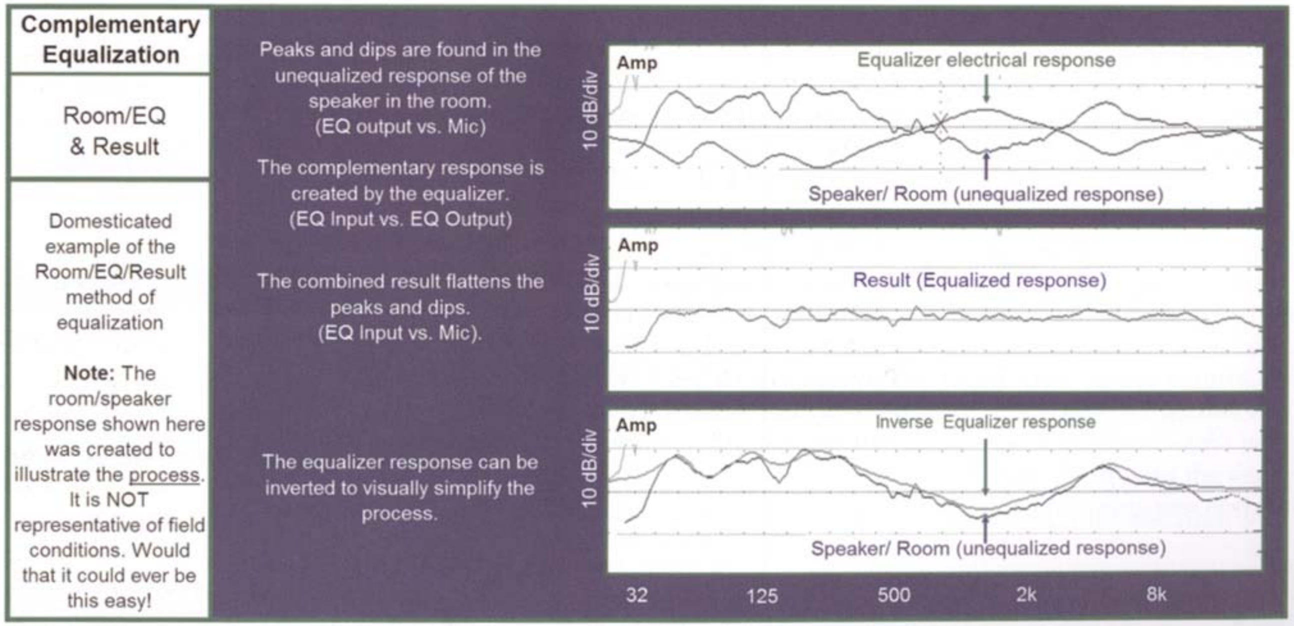


Figure 10 27h Complementaryhequalizationhexample

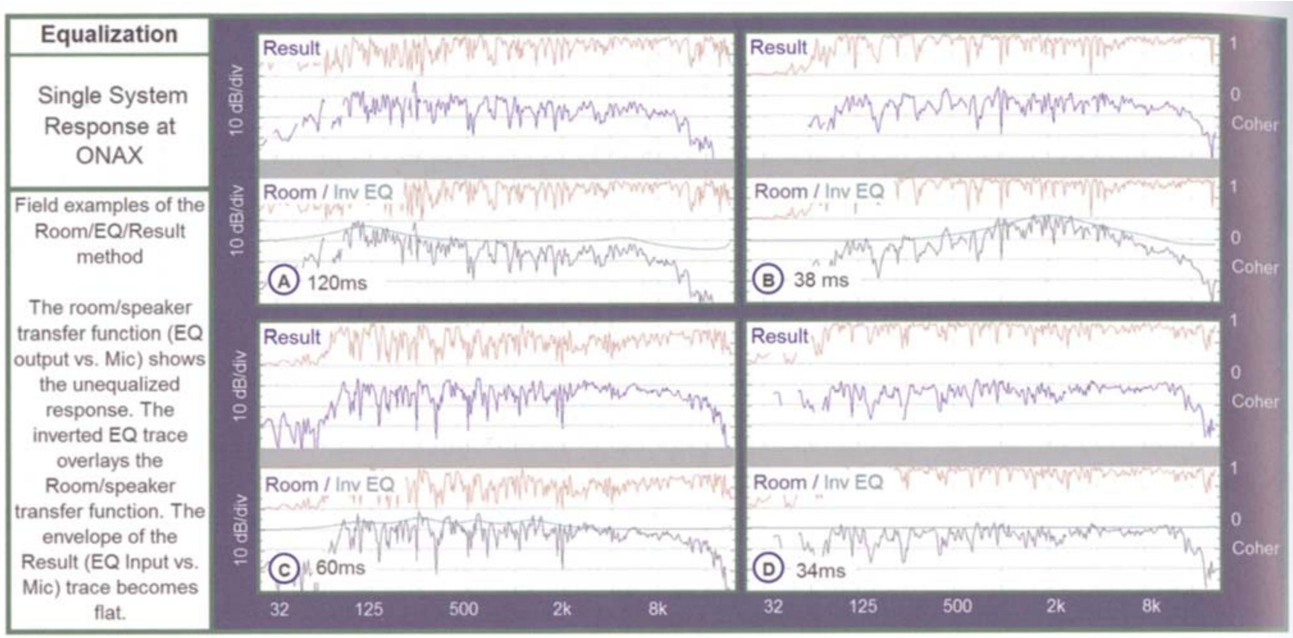


Figure 10.28 Equalization examples from field data

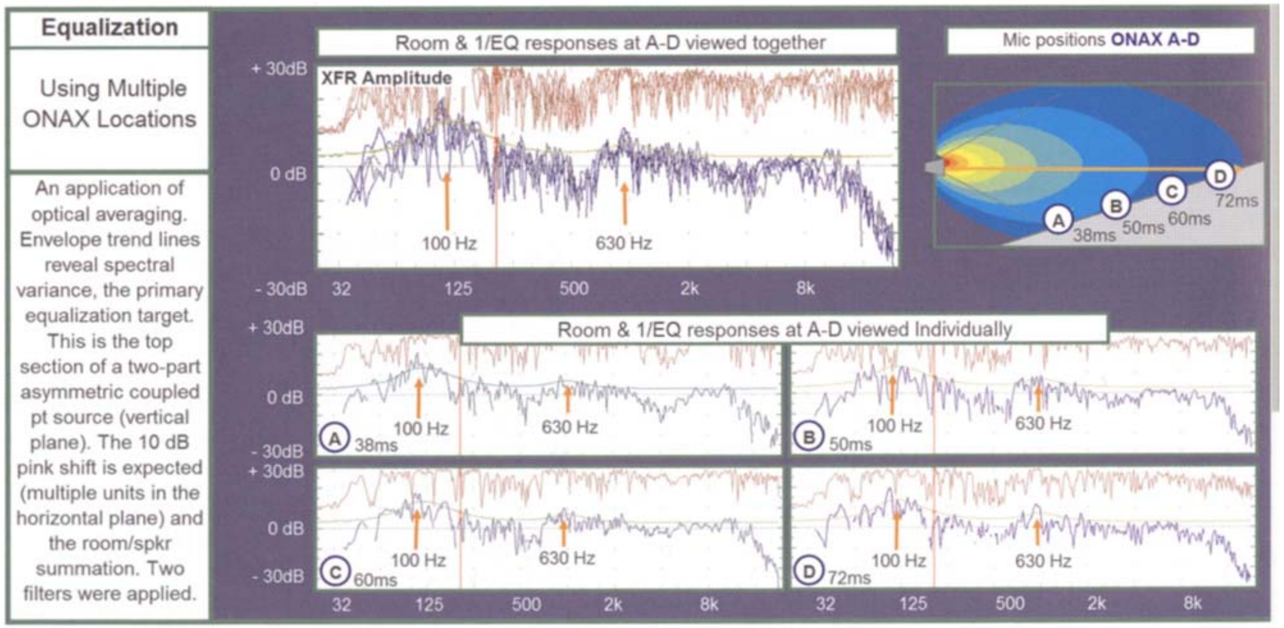


Figure 10.29 Equalization examples from field data

Perspectives: I was trying to explain the result trace, and the goal of making it as straight as possible, and the band's engineer said, "Oh, you're trying to kill the sound."

d. (Mack) mcbryde

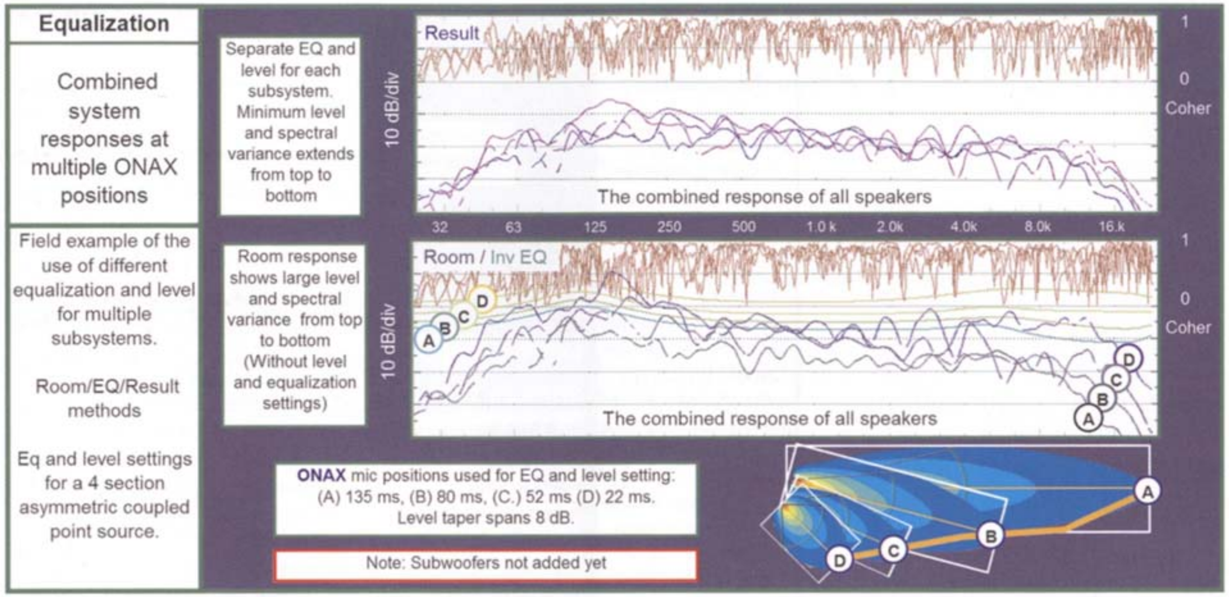


Figure 10.30 Equalization examples from field data

Perspectives: I think about equalization as three distinct areas that need adjustment. The first, the middle frequencies, are for minor corrections the speakers might need or possibly effects caused by arraying. The second area involves the high frequencies, correcting for problems at the upper end of the spectrum and for atmospheric losses. The last is the low frequencies. Here in addition to issues the speakers may have, there is the space the system is in. I think that, in some cases, it is possible to deal with some room issues and improve the overall sound of the system by making careful corrections based on room issues.

Alexander Yuill-Thornton II (Thorny)

The temptation is to remove substantial LF energy from the lower level speaker but this is ineffective due to the level offset (see summation level properties in Chapter 2). Another tendency is to remove LF energy from the main system to accommodate the secondary area. TANSTAAFL applies here. If we remove enough LF energy to flatten the secondary area we risk the telephone transmission syndrome in the mains.

2. Adjust the dominant equalizer only to the extent that it can be changed without compromising the response in its ONAX (A) area. The pink shift left in the secondary system (ONAX (B)) is an application of acoustic triage.

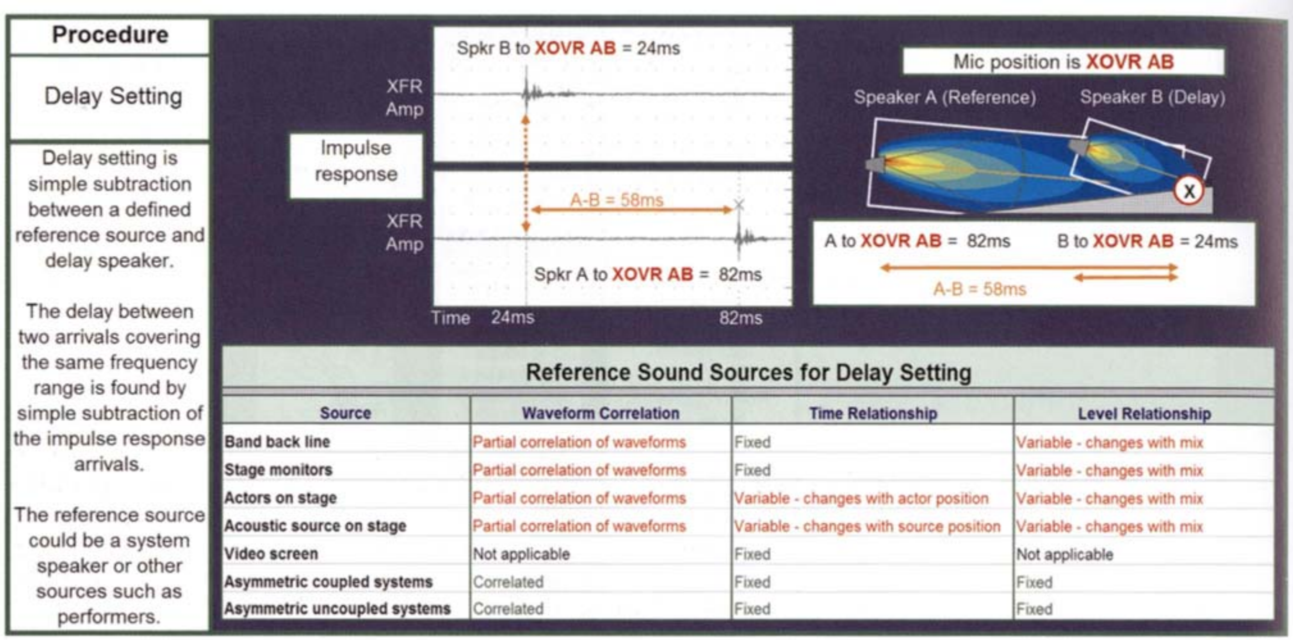
Delay Setting

Delay setting does not carry the emotional weight of equalization but has its own quirks as well. For the most part, delay setting is very straightforward. We measure the time offset between a main and delay speaker and type the number into our delay line. The operation is simple but the strategy behind the setting is more complex.

There are two types of delay setting scenarios. The first and most common is synchronization between two displaced and fixed sound sources. This is the phase-aligned spatial crossover, which is an overriding theme of this book. The settings will be precisely tailored as an integral part of our minimum variance strategy. The procedures are simple and will follow shortly.

The second is the union of fixed and variable sources. This might be the delaying of the speaker system to the back line guitars, the sidefill stage monitors or the actors on stage. This second type is an approximate process, and plays only a limited role in the minimum variance process. The goal here might be enhanced sonic image control. In the case of a band with excessive stage levels, the choice to synchronize to the back line is a form of "if you can't beat 'em, join 'em." It is preferable to ride along with the stage levels (and reduce ripple variance and echo perception) than to lead them in time.

The relationship between these stage sources and our sound system fails the stable summation criteria outlined



Reference Sound Sources for Delay Setting

Source	Waveform Correlation	Time Relationship	Level Relationship
Band back line	Partial correlation of waveforms	Fixed	Variable - changes with mix
Stage monitors	Partial correlation of waveforms	Fixed	Variable - changes with mix
Actors on stage	Partial correlation of waveforms	Variable - changes with actor position	Variable - changes with mix
Acoustic source on stage	Partial correlation of waveforms	Variable - changes with source position	Variable - changes with mix
Video screen	Not applicable	Fixed	Not applicable
Asymmetric coupled systems	Correlated	Fixed	Fixed
Asymmetric uncoupled systems	Correlated	Fixed	Fixed

Figure 10.31 Delay setting and reference speaker considerations

Perspectives Time alignment happens at all scales! A well-tuned system is aligned driver to driver (within a box), box to box, cluster to cluster and in some cases clusters to back-line. When you change the delay in any portion of a system, it is important to know if you are affecting the alignment at any other scale.

Sam Berkow

in Chapter 2. The waveforms are only partially correlated since the sound from the stage is not a match for that which leaves our speakers. The level relationship is variable with changes in the mix and the timing can change as actors move around the stage.

The synchronization to stage sources should be undertaken with caution. Only the absolute minimum number of subsystems (those closest to the stage) should be "aligned" this way. All remaining systems should be synchronized to these speaker systems by the minimum variance methods.

The Precedence Effect

It is a virtual certainty that we will be asked about the precedence effect (also known as the Haas effect) as we go about the business of setting delays. There is a pervasive practice of purposeful de-calibration of delay settings for the supposed benefit of superior sonic imaging attributed to the precedence effect. The strategy is this: find the delay time and then add (5,10,15,20, some number of ms) to the delay so that the main system precedes the delay. The intention is that the precedence effect will aid us to perceive the sound as coming from the main source. Because of its prevalence we will need to meet this strategy head on.

Consider the following:

Recall that the precedence effect is a binaural function. It is applicable to the horizontal plane only.

The offset of x ms is certain to add ripple variance to the summed response of the two systems. If 10ms offset is used, the combing will reach all the way down to 50Hz. Ripple variance decreases coherence and creates the ***experience*** of increased reflections. The job of the delay speaker is to decrease ripple variance, not add to it.

People advocating the setting of delays in this manner have not once mentioned to me the matter of relative level between the mains and delays, seeming only to consider this after the fact. A position is chosen, the delay is set (and then offset) and then the delay speaker level is adjusted to taste. It does not seem to occur to them that they might be turning the level up higher because the intelligibility is down because they misaligned the delay.

The misaligned delay is phase-aligned at some unknown location and has a spatial crossover at some other unknown location. Does this sound like a good plan?

A misaligned delay has lower coherence than a phase-aligned delay. Which one will need to be turned up louder to understand the words under the balcony?

A phase-aligned spatial crossover is based on zero level and time offset. In such cases, the precedence effect would place the image at the center point between the two sources. Is that such a big price to pay for maximum intelligibity, and minimum ripple variance?

When confronted with requests from clients over the years for the precedence effect enhancement, I have proposed a solution as follows: first I ask them to listen to the delays set up as a phase-aligned crossover. Then if they don't like it we can add however much delay they want. In 20 years we have never yet made it to step two.

Delay Setting Procedure AB

Impulse Response Method

The secondary system (B) is to be delayed to synchronize with the dominant system (A).

1. Find the propagation delay from Speaker A to XOVR AB.
2. Find the propagation delay from Speaker B to XOVR AB.
3. The difference between these two readings must be added (or subtracted) to the delay line to synchronize the responses.
4. Add both systems together. The impulse responses should combine to a single impulse.

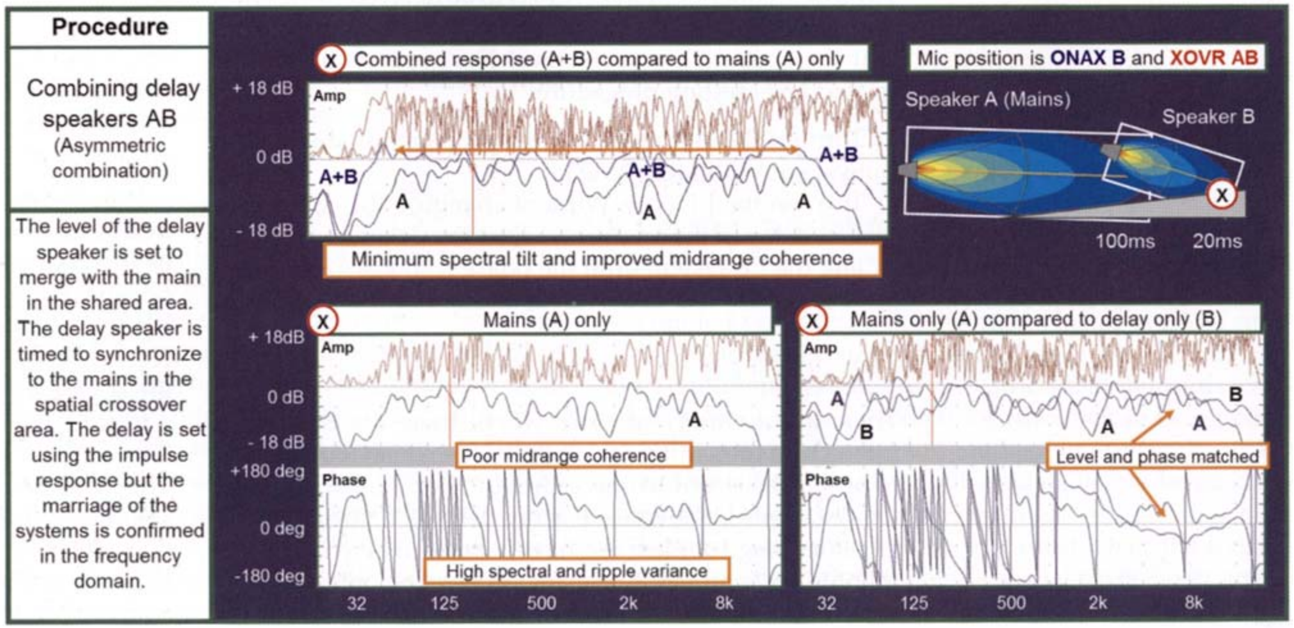
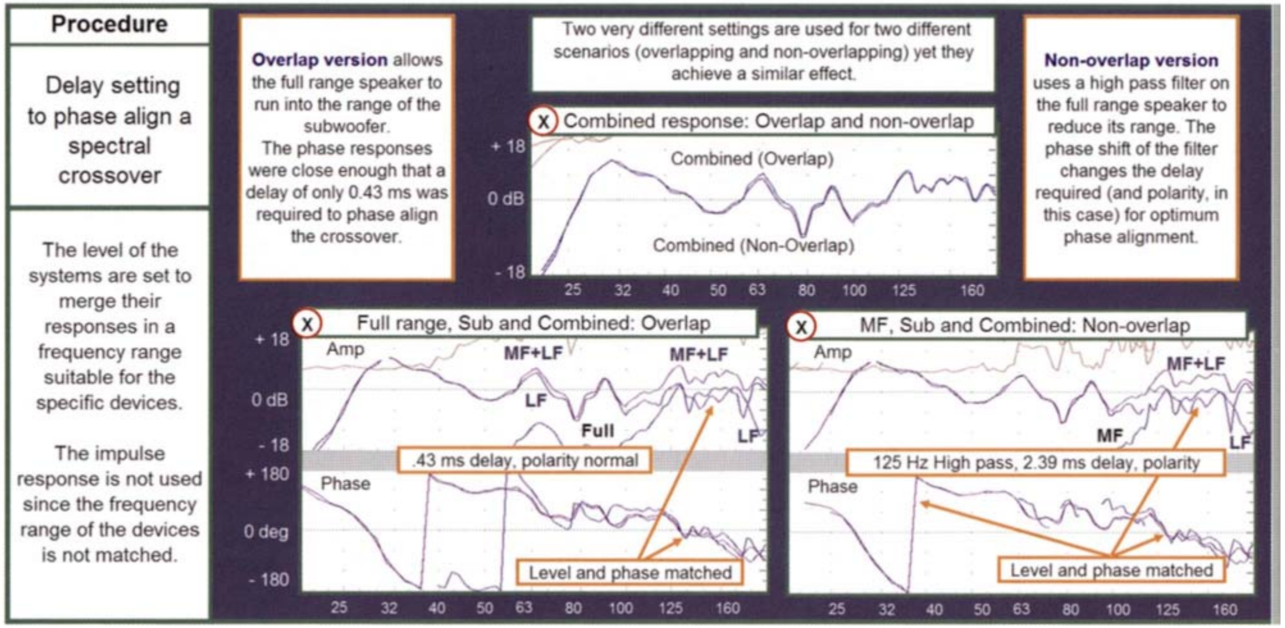


Figure 10.32 Delay setting procedure AB. Delay set—for an asymmetric combination



Delay setting procedure (phase spectral XOVR) delay set—for a phase aligned spectral crossover

Perspectives *I knew my delays were set properly when the audience came and told me the delay speakers were off.*

Don** (Dr Don) **Pearson

Phase Response Method for Spectral Crossover Alignment

The low-frequency system (LF) is to be delayed to synchronize with the high-frequency system (HF). It is assumed that the spectral crossover level has been set as per the related level setting procedure.

1. Turn on the HF only. Find the propagation delay from System A (HF) to ONAX.
2. Turn on the LF only. Without changing the compensation delay in the analyzer, adjust the delay line controlling the LF system until the phase responses overlay.
3. Add the HF and observe the addition.

Delay Dilemmas

There are a variety of cases where there are competing interests in the setting of delays. We can delay our speaker to be synchronized to one source or another, but not both. One example is the frontfill array delayed a virtual source at center stage. Unless we have a perfectly circular stage there is no "one size fits all" solution. The optimal delay time for the inner frontfills (to the source) is shorter than the outers. If we make the inner and outers the same, we will have differing relationships to the source. If we set them differently, they will have different relationships to each other. The seat at the spatial crossover between the inner and outer systems will be affected by a difference in the respective arrivals. This is another case of TANSTAAFL. How do we choose? Triage, as usual.

What is at stake here? What are the variables? If we are delaying to a source on stage we are venturing into areas beyond our control. The levels and timing relationships to our speaker system are "subject to change without notice." We must always remember that sonic imaging is reliant on both level and delay relationships. The difference from the source to the speakers is not just time. It is level also. Are we going to set the inner and outer frontfills to different levels in anticipation of the stage source levels? This may be difficult to assess unless the stage source is stationary. If so the different delay level and delay time is warranted. If not, we are better off ensuring a phase-aligned crossover at the transition between the inners and outers. If we are going to go with just one delay time which should it be? The outers. Why? They are (probably) louder and earlier, of the two systems, compared to the source. Therefore they need the most help. If the timing is set for the outers, the inners will have excess delay, which is the more favorable error for imaging.

Another common delay dilemma is the under balcony array that is delayed to the main speakers flanking the stage. The center delays are far from the sides, while the outers are near. The difference in this case is fixed, and often substantial. If we have an aisle the choice is easy. Split your delay times along that line. If not, it is still advisable to break the delays apart in most cases. The tradeoff is a question of spatial crossover phase alignment. With one delay setting the crossover between the mains and the delay speakers is substantially out of time for entire areas of the delay coverage. In the multiple delay scenario the principal relationship, mains to delays, is optimized and the smaller local horizontal transitions between the delays is compromised. Fewer people are adversely affected. Consider this litmus test: the spatial crossover between the mains and delays must be highly overlapped. The spatial crossovers between the multiple under balcony speaker ***should*** be low overlap. The high overlap crossover takes phase alignment priority.

There are many other types of delay dilemmas. These principles can be applied to aid the decision process.

Perspectives When doing a large, complex system, I identify what looks to be the key subsystem and equalize that. I then listen to it and determine if it is doing what I expected. After 1 have the core where I want it, I start adding other subsystems, listening at each stage to confirm that things are doing what 1 expect. This way I don't end up investing a lot of time and energy going down a path that leads to nowhere.

Alexander Yuill-Thornton II (Thorny)

Order of Operations

We have seen the individual calibration procedures. This is our playbook. Now we must organize them into a game plan. That plan takes shape as an order of operations that begins with each element in each subsystem until all of them are assembled together. The order of operations in system calibration proceeds by the democracy principle: the longest throw system is assumed to cover the largest number of people, and is designated by the letter A. The system with the second highest level is designated B and so on.

Once we have performed our initial calibrations for the individual subsystems we can begin the process of combining these together. It is our hope that the pieces will fit together like a puzzle. If not, we may have to go back and do some rework on the subsystems.

Before the combination process begins, these operations must be completed on all individual systems:

1. Speaker focus as found by the ONAX, OFFAX and XOVR positions.
2. Initial level and equalization setting at the ONAX position(s).
3. Acoustic treatment for the local area as found at any mic position.

The combination process will include:

1. Delay setting at the XOVR position.
2. Level or speaker position adjustment to minimize variance between the ONAX and XOVR positions (if required).
3. Combined equalization adjustment at the ONAX positions (if required).

The strategies for combination follow an order of operations similar to the process involved in mathematics. Let's look at an example equation:

$$(((A1 + A2) + (B1 + B2)) + (C1 + C2))$$

The first operations are the combinations of the 1's and 2's such as (A1 + A2). Next we combine the A's to the B's and then finally add the C's to the running total.

The priorities for calibration order of operations:

1. Coupled subsystems: symmetric subsystem of a coupled array must be joined together before being combined with others, e.g. the horizontal symmetric coupled point source of the main cluster must be combined before the sidefills are added.
2. Coupled subsystems: asymmetric subsystems of a coupled array must be joined together before being combined with others, e.g. the coupled downfill of the main cluster is first joined with the mains before being combined with the uncoupled frontfills.
3. Longest throw: the system that must cover the longest distance takes priority over those that cover closer areas.
4. Closest proximity.

Example:

((((Main upper inner + Main upper outer) + (Main lower inner + main lower outer)) + (Frontfill inner + Outer)))

Now let's look at Fig. 10.34 which will serve as the guide for the reference charts that follow. Each chart contains a number of speaker elements, the class of microphone positions, their functions and the role of signal processing on the calibration. The series of charts shows representative variations of speaker model, desired throw distance and spacing. The spacing may be angular, lateral and both, depending upon the array type. The series of charts is not intended to cover every possible scenario. It does, however, have sufficient quantity for the logic and trends to be clearly seen so that readers can apply these strategies to any array or combination of arrays encountered in the field. Symmetric arrays require the minimum number of mic positions and signal processing channels. They also have the maximum number of SYM mic positions, which are the simplest of the operations. By contrast

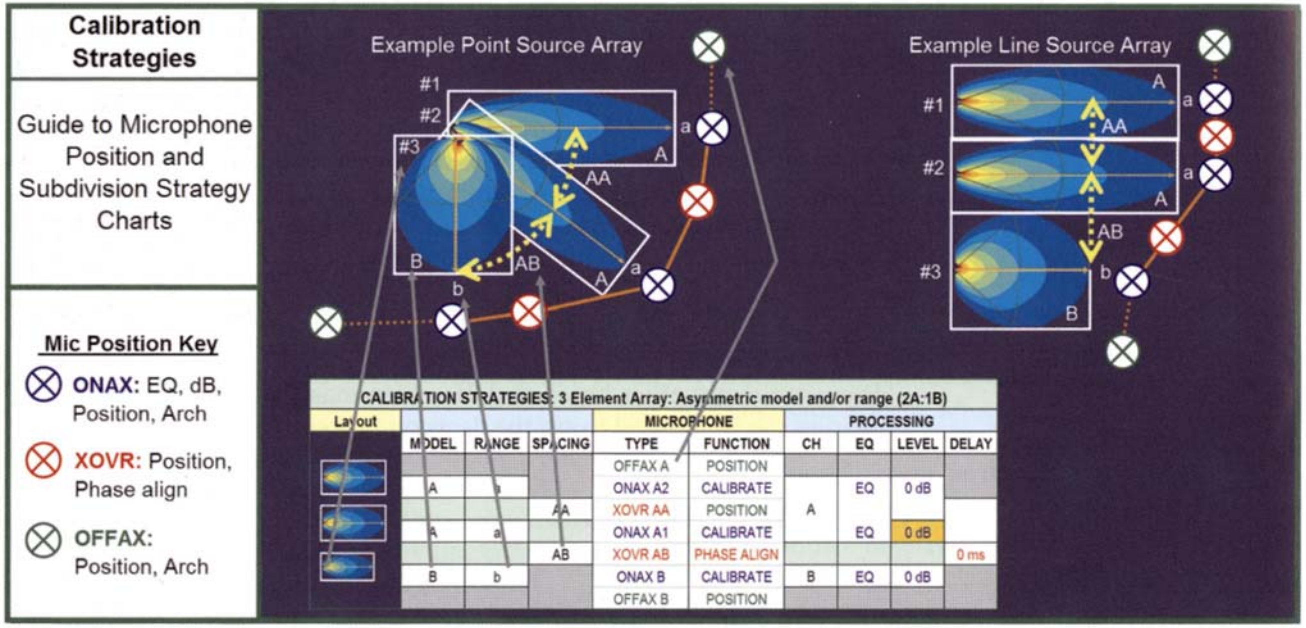


Figure 10.34 Guide to microphone positions and calibration subdivision strategies for the series of charts that follow for differen array quantities and configurations. A representative three-element example array is shown

CALIBRATION STRATEGIES: 1 Element									
Layout				MICROPHONE		PROCESSING			
	MODEL	RANGE	SPACING	TYPE	FUNCTION	CH	EQ	LEVEL	DELAY
				OFFAX A1	POSITION				
	A	a		ONAX A	CALIBRATE	A	EQ	0 dB	
				OFFAX A2	POSITION				

ORDER OF OPERATIONS				
STEP	SPKR	MICROPHONE		PROCEDURES
		TYPE	FUNCTION	
1	A	ONAX A	POSITION	Set Gold Reference Level
2	A	OFFAX A1 & A2	POSITION	Adjust for minimum variance between OFFAX & ONAX
3	A	ONAX A	CALIBRATE	Initial Equalization

Figure 10.35 Calibration strategies for a single element. Asymmetric orientation is shown. Symmetric version requires only a single OFFAX measurement

CALIBRATION STRATEGIES: 2 Element Array: Symmetric (2A)									
Layout				MICROPHONE		PROCESSING			
	MODEL	RANGE	SPACING	TYPE	FUNCTION	CH	EQ	LEVEL	DELAY
				OFFAX A	POSITION				
	A	a		ONAX A1	CALIBRATE		EQ	0 dB	
			AA	XOVR AA	POSITION	A			
	A	a		ONAX A2	SYMMETRIC		EQ	0 dB	
				OFFAX A	SYMMETRIC				

ORDER OF OPERATIONS				
STEP	SPKR	MICROPHONE		PROCEDURES
		TYPE	FUNCTION	
1	A	ONAX A1	POSITION	Set Gold Reference Level
2	A	XOVR AA	POSITION	Set for minimum variance at ONAX A1, XOVR AA, ONAX A2
3	A	OFFAX A	POSITION	Set for minimum variance between OFFAX A & ONAX A1
4	A	ONAX A1 & A2	CALIBRATE	Initial Equalization based on both ONAX positions

CALIBRATION STRATEGIES: 3 Element Array: Symmetric (3A)									
Layout				MICROPHONE		PROCESSING			
	MODEL	RANGE	SPACING	TYPE	FUNCTION	CH	EQ	LEVEL	DELAY
				OFFAX A	POSITION				
	A	a		ONAX A2	CALIBRATE		EQ	0 dB	
			AA	XOVR AA	POSITION				
	A	a		ONAX A1	CALIBRATE	A	EQ	0 dB	
			AA	XOVR AA	SYMMETRIC				
	A	a		ONAX A2	SYMMETRIC		EQ	0 dB	
				OFFAX A	SYMMETRIC				

CALIBRATION STRATEGIES: 4 Element Array: Symmetric (4A)									
Layout				MICROPHONE		PROCESSING			
	MODEL	RANGE	SPACING	TYPE	FUNCTION	CH	EQ	LEVEL	DELAY
				OFFAX A	POSITION				
	A	a		ONAX A2	CALIBRATE		EQ	0 dB	
			AA	XOVR AA	POSITION				
	A	a		ONAX A1	CALIBRATE		EQ	0 dB	
			AA	XOVR AA	SYMMETRIC	A			
	A	a		ONAX A1	SYMMETRIC		EQ	0 dB	
			AA	XOVR AA	SYMMETRIC				
	A	a		ONAX A2	SYMMETRIC		EQ	0 dB	
				OFFAX A	SYMMETRIC				

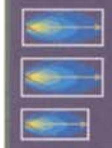
Figure 10.36 Calibration strategies for fully symmetric systems. The approach for higher quantities can be interpolated from the trends shown here

CALIBRATION STRATEGIES: 2 Element Array: Asymmetric model and/or range (1A:1B)									
Layout				MICROPHONE		PROCESSING			
	MODEL	RANGE	SPACING	TYPE	FUNCTION	CH	EQ	LEVEL	DELAY
				OFFAX A	POSITION				
	A	a		ONAX A	CALIBRATE	A	EQ	0 dB	
			AB	XOVR AB	PHASE ALIGN				0 ms
	B	b		ONAX B	CALIBRATE	B	EQ	0 dB	
				OFFAX B	POSITION				

ORDER OF OPERATIONS				
STEP	SPKR	MICROPHONE		PROCEDURES
		TYPE	FUNCTION	
1	A	ONAX A	POSITION	Set Gold Reference Level
2	A	OFFAX A	POSITION	Set for minimum variance between OFFAX A & ONAX A
3	A	ONAX A	CALIBRATE	Initial Equalization Channel A
4	B	ONAX B	CALIBRATE	Set Level B to match Gold Reference A
5	B	ONAX B	CALIBRATE	Initial Equalization Channel B
6	A+B	XOVR AB	POSITION	Set for minimum variance at ONAX A, XOVR AB, ONAX B
7	A+B	XOVR AB	PHASE ALIGN	Set Delay Channel B
8	A+B	ONAX A & ONAX B	CALIBRATE	Combined Equalization A+B
9	A+B	OFFAX B	POSITION	Verify outside coverage edge B

Figure 10.37 Calibration strategies for two element asymmetric systems. This approach can be used for single speakers within an array or for the combination of a previously calibrated set of arrays

CALIBRATION STRATEGIES: 3 Element Array: Asymmetric model and/or range (2A:1B)



Layout	MODEL	RANGE	SPACING	MICROPHONE TYPE	MICROPHONE FUNCTION	PROCESSING CH	EQ	LEVEL	DELAY
				OFFAX A	POSITION				
	A	a		ONAX A2	CALIBRATE	A	EQ	0 dB	
			AA	XOVR AA	POSITION				
	A	a		ONAX A1	CALIBRATE		EQ	0 dB	
			AB	XOVR AB	PHASE ALIGN				0 ms
	B	b		ONAX B	CALIBRATE	B	EQ	0 dB	
				OFFAX B	POSITION				

ORDER OF OPERATIONS

STEP	SPKR	MICROPHONE TYPE	MICROPHONE FUNCTION	PROCEDURES
1	A	ONAX A1	POSITION	Set Gold Reference Level
2	A	XOVR AA	POSITION	Set for minimum variance at ONAX A1, XOVR AA, ONAX A2
3	A	OFFAX A	POSITION	Adjust for minimum variance between OFFAX A & ONAX A1
4	A	ONAX A1	CALIBRATE	Initial Equalization Channel A
5	B	ONAX B	CALIBRATE	Set Level B to match Gold Reference A
6	B	ONAX B	CALIBRATE	Initial Equalization Channel B
7	A+B	XOVR AB	POSITION	Set for minimum variance at ONAX A1, XOVR AB, ONAX B
8	A+B	XOVR AB	PHASE ALIGN	Set Delay Channel B
9	A+B	ONAX A & ONAX B	CALIBRATE	Combined Equalization A+B
10	A+B	OFFAX B	POSITION	Verify outside coverage edge B

Figure 10.38 Calibration subdivision strategies for three elements in the 2A:1B configuration

CALIBRATION STRATEGIES: 3 Element Array: Asymmetric model, range & spacing mix (1A:1B:1C)

Layout	MODEL	RANGE	SPACING	MICROPHONE TYPE	MICROPHONE FUNCTION	PROCESSING CH	EQ	LEVEL	DELAY
				OFFAX A	POSITION				
	A	a		ONAX A	CALIBRATE	A	EQ	0 dB	
			AB	XOVR AB	PHASE ALIGN				0 ms
	A (or B)	b		ONAX B	CALIBRATE	B	EQ	0 dB	
			BC	XOVR BC	PHASE ALIGN				0 ms
	B (or C)	c		ONAX C	CALIBRATE	C	EQ	0 dB	
				OFFAX C	POSITION				

ORDER OF OPERATIONS

STEP	SPKR	MICROPHONE TYPE	MICROPHONE FUNCTION	PROCEDURES
1	A	ONAX A	POSITION	Set Gold Reference Level
2	A	OFFAX A	POSITION	Adjust for minimum variance between OFFAX A & ONAX A
3	A	ONAX A	CALIBRATE	Initial Equalization Channel A
4	B	ONAX B	CALIBRATE	Set Level B to match Gold Reference A
5	B	ONAX B	CALIBRATE	Initial Equalization Channel B
6	A+B	XOVR AB	POSITION	Set for minimum variance at ONAX A, XOVR AB, ONAX B
7	A+B	XOVR AB	PHASE ALIGN	Set delay Channel B
8	A+B	ONAX A & ONAX B	CALIBRATE	Combined Equalization A+B
9	C	ONAX C	CALIBRATE	Set Level C to match Gold Reference A
10	C	ONAX C	CALIBRATE	Initial Equalization Channel C
11	(A+B)+C	XOVR BC	POSITION	Set for minimum variance at ONAX B, XOVR BC, ONAX C
12	(A+B)+C	XOVR BC	PHASE ALIGN	Set delay Channel C
13	(A+B)+C	ONAX A, B, C	CALIBRATE	Combined Equalization (A+B)+C
14	(A+B)+C	OFFAX C	POSITION	Verify outside coverage edge C

Figure 10.39 Calibration subdivision strategies for three elements in the 1A:1B:1C configuration. The sec shown here can be continued indefinitely for additional asymmetric levels

CALIBRATION STRATEGIES: 4 Element Array: Asymmetric model and/or distance (1B:2A:1B)									
Layout				MICROPHONE		PROCESSING			
	MODEL	RANGE	SPACING	TYPE	FUNCTION	CH	EQ	LEVEL	DELAY
				OFFAX B	POSITION				
	B	b		ONAX B1	CALIBRATE	B	EQ	0 dB	
			BA	XOVR AB	PHASE ALIGN				0 ms
	A	a		ONAX A1	CALIBRATE		EQ	0 dB	
			AA	XOVR AA	POSITION	A			
	A	a		ONAX A2	SYMMETRIC		EQ	0 dB	
			AB	XOVR AB	SYMMETRIC				0 ms
	B	b		ONAX B2	SYMMETRIC	B	EQ	0 dB	
				OFFAX B	SYMMETRIC				

ORDER OF OPERATIONS				
STEP	SPKR	MICROPHONE		PROCEDURES
		TYPE	FUNCTION	
1	A	ONAX A1	POSITION	Set Gold Reference Level
2	A	XOVR AA	POSITION	Set for minimum variance at ONAX A1, XOVR AA, ONAX A2
3	A	ONAX A1	CALIBRATE	Initial Equalization Channel A
4	B	ONAX B	CALIBRATE	Set Level B to match Gold Reference A
5	B	ONAX B	CALIBRATE	Initial Equalization Channel B
6	A+B	XOVR AB	POSITION	Set for min. variance at ONAX A1, XOVR AB, ONAX B
7	A+B	XOVR AB	PHASE ALIGN	Set Delay Channel B
8	A+B	ONAX A & ONAX B	CALIBRATE	Combined Equalization A+B
9	A+B	OFFAX B	POSITION	Verify outside coverage edge B

Figureh10.40h Calibrationhstrategieshforhfourhelementshinhtheh1hB:2A:1Bhconfiguration

CALIBRATION STRATEGIES: 4 Element Array: Asymmetric model and/or distance + spacing (3A:1B)									
Layout				MICROPHONE		PROCESSING			
	MODEL	RANGE	SPACING	TYPE	FUNCTION	CH	EQ	LEVEL	DELAY
				OFFAX A	POSITION				
	A	a		ONAX A3	SYMMETRIC		EQ	0 dB	
			AA	XOVR AA	SYMMETRIC				
	A	a		ONAX A1	CALIBRATE	A	EQ	0 dB	
			AA	XOVR AA	CALIBRATE				
	A	a		ONAX A2	CALIBRATE		EQ	0 dB	
			AB	XOVR AB	PHASE ALIGN				0 ms
	B	b		ONAX B	CALIBRATE	B	EQ	0 dB	
				OFFAX B	POSITION				

ORDER OF OPERATIONS				
STEP	SPKR	MICROPHONE		PROCEDURES
		TYPE	FUNCTION	
1	A	ONAX A1	POSITION	Set Gold Reference Level
2	A	XOVR AA	POSITION	Set for minimum variance at ONAX A1, XOVR AA, ONAX A2
3	A	OFFAX A2	POSITION	Set for minimum variance between OFFAX A & ONAX A1
4	A	ONAX A1 & A2	CALIBRATE	Initial Equalization based on both ONAX positions
5	B	ONAX B	CALIBRATE	Set Level B to match Gold Reference A
6	B	XOVR BB	POSITION	Set for minimum variance at ONAX B1, XOVR BB, ONAX B2
7	B	ONAX B	CALIBRATE	Initial Equalization Channel B
8	A+B	XOVR AB	POSITION	Set for min. variance at ONAX A1, XOVR AB, ONAX B
9	A+B	XOVR AB	PHASE ALIGN	Set Delay Channel B
10	A+B	ONAX A & ONAX B	CALIBRATE	Combined Equalization A+B
11	A+B	OFFAX B	POSITION	Verify outside coverage edge B

Figureh10.41h Calibrationhstrategieshforhfourhelementshinhtheh3A:1Bhconfiguration

Perspectives: Some tips about system optimization in small rooms:

1. *I've found that 80 per cent of getting it to sound good in a small room is just patting the speakers and the listener in the right position. This requires some experimentation to balance out the speaker boundary reflections but once you find that proper ratio, minimal room treatments are required, and equalization just becomes icing on the cake. Of course you're sunk if the room's modal distribution is really bad.*
2. *Remember Dustin Hoffman's college graduation party scene in* The Graduate, *where that businessman says to him, "I have just one word for you, son: Plastics." Well, I have just one word for you: Symmetry. If there is just one thing you do right, it should be to set your control room up as symmetrically as possible. What does this mean and why?*

   If your speakers are not placed symmetrically in the room, they will have different frequency responses. This means that your music will sound different in the left and right speakers, your center image will be off center and your music will not properly collapse to mono.
3. *Recording engineers seem to hold on to old habits. When you discuss room tuning, many still believe that a 1/3 octave equalizer is the way to go. In days of yore when the only analyzer was the 2/3 octave variety, that was the only option. Today, with our high-resolution analyzers it only makes sense to use a parametric equalizer since you can select the exact center frequency of the existing problem and adjust the*

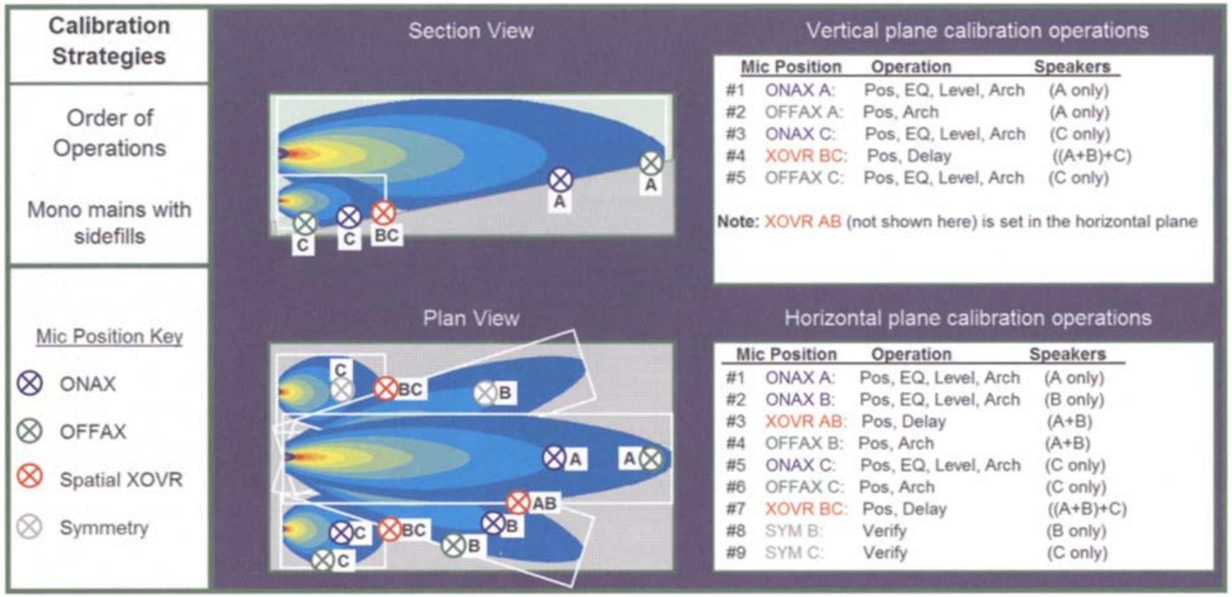


Figure 10.42 Order of operations applied to an example system

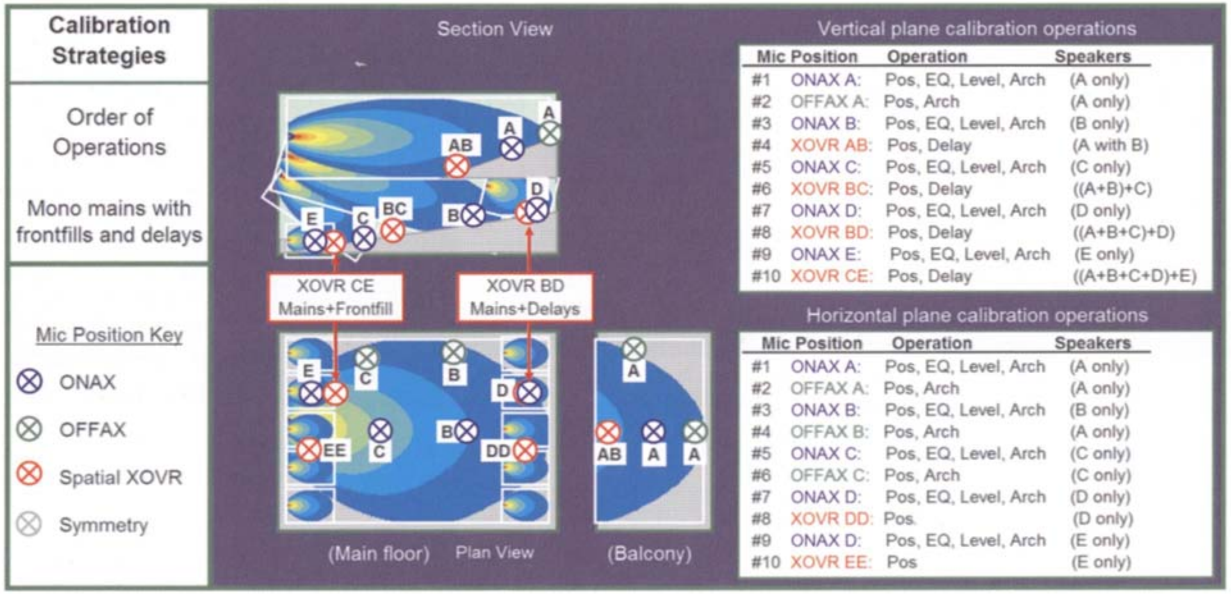


Figure 10.43 Order of operations applied to an example system

bandwidth to fit the existing curve. With a 2/3 octave equalizer you are stuck with fixed Q and frequency centers. That's sort of like doing brain surgery with a butter knife.

4. *A common mistake people make when treating a room such as a vocal booth, small control room or home theater is to throw 1" compressed fiberglass up on all of the walls. This removes flutter echo problems but skews the room response. All of the high end gets absorbed but the low end is left to roll around. Absorptive treatments should be used only where necessary, not globally, and they should be well-balanced for both high and low frequencies.*
5. *Don't believe that old wives' tale that nearfields are not affected by room acoustics. The laws of physics apply to all speakers in a room, so the boundaries will have an effect since most studio monitors are pretty much omnidirectional below 200 Hz. Also, nearfields sitting on the console meter bridge will have severe comb filtering due to the first-order reflection off of the console surface.*

Bob Hodas

Perspectives: One of the things I seem to be fairly good at is getting done what absolutely has to be done for an event to happen. We all want to do everything on our list when installing and optimizing a system, but there are times, especially when touring, when it is clear that there will not be enough time. This is when you have to start

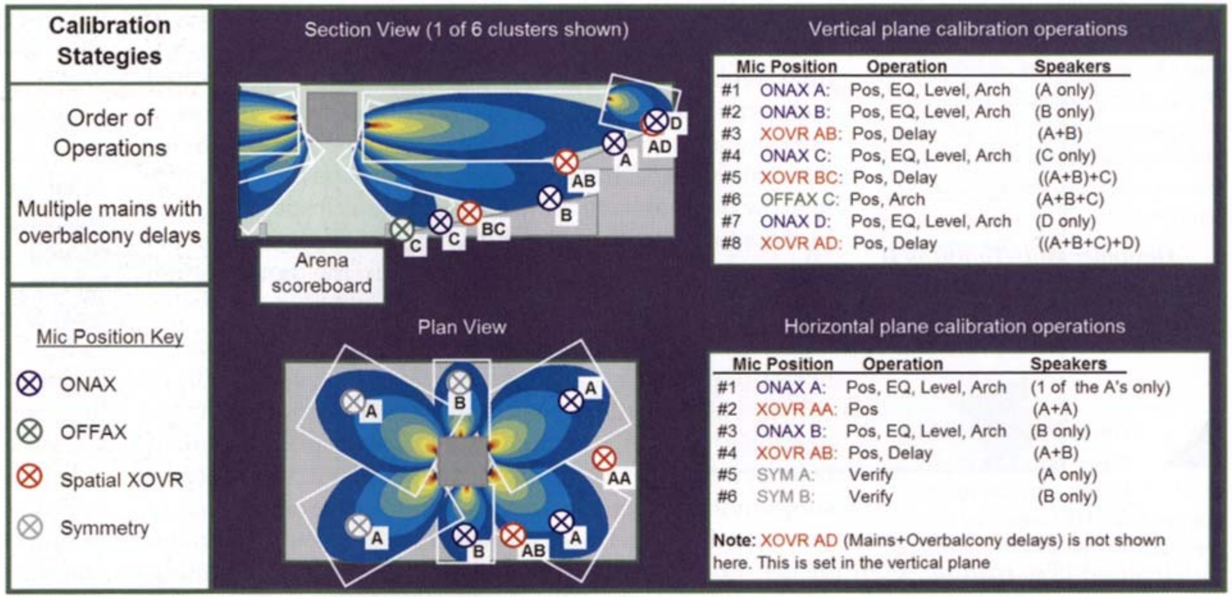


Figure 10.44 Order of operations applied to an example system

all forms of asymmetry (model, distance or spacing) must be fully characterized with dedicated microphone positions and the asymmetry managed by separate signal processing channels. The purpose of these charts is to show a systematic approach to the calibration process.

The calibration of even the most complex sound system design can be broken down into the series of procedures and the order of operations. Step by step, subsystem by subsystem, the pieces are stitched together like quilt pieces into a single fabric. Figures 10.42 to 10.44 show representative examples of the process.

Finishing the Process

The inevitable questions arise: how can we put an end to an iterative process that approaches perfection but can never reach it? We know that we will never be able to make a perfect match for every seat. The fully combined system will be as close as we can come. If this is not acceptable we will need to disassemble the pieces, adjust them, and recombine them again. There are instances where alternative approaches appear to have equal merit. We can first try dividing the space and combining the systems one way, and then try the other. Nothing ends a discussion of alternatives better than "Let's test both theories."

Here is how it ends: we run out of time. Seriously. As long as there is the time there is more that can be learned. Anything we learn today is applicable in advance for tomorrow.

Listening

Our ears are trained professional audio tools. Each ear (and set of ears) has inherent physiological differences. They are not stable physically over time or over our lifespan. If we have just flown into town we are likely to find our ears still *editing your list and resetting your priorities. Sometimes one has to be brutal because if you are running out of time, other departments are probably in the same position. You're editing and prioritizing needs to be done tenth the total event in mind, not just your own department.*

Alexander Yuill-Thornton II (Thorny)

Perspectives: When mixing engineers mix for the first time on a speaker system that has been SIM-tuned, they are apt to feel the lack of low frequencies. Having worked as a SIM engineer for years, I was often questioned on that point. Close observation of each channel on a mixing board shows that high-pass filters are used too much. Such a way of mixing tends to result in a noisy sound with excessive energy concentrated in the high-pitched tones.

Then I would tell the mixing engineer, "Just take my word for it and ..." For example, I would say, "Undo the high-pass filters and let the faders (sound-level) down! If it does not work, all you need to do is put them back. Besides, this is a rehearsal. It can't do any harm. What have we got to lose?"

Any mixing engineers who ventured to take my word were mostly successful in creating excellent sound!

Akira Masu

in the process of adjusting to the local atmospheric pressure for some time. If we have just been exposed to high volume levels on stage our dynamic thresholds will be shifted temporarily. If we have experienced high levels of exposure over the long term we may have permanent shifts in our dynamic and spectral responses. Those of us lucky enough to grow old will have reduced dynamic and spectral range in the normal course of aging. These factors and others make the ear a subjective partner in the optimization process.

The ears are connected to our brains which contain a library of reference data for comparison. Everyone's library carries volumes of accrued aural experience, and everyone's is unique. It is through the library that a measure of objectivity can be introduced which offsets the inherent subjectivity. This enables two people to agree on the quality of a listening experience. How do we know what a violin ***should*** sound like? This knowledge comes from the thousand references to violins in our aural library. Over time we have accrued a map of the expected response of a violin, sets of violins, and details such as how they sound when plucked as opposed to bowed. This is the ear training which we bring to bear in evaluating the final performance of the sound system. We perform moment-to-moment comparisons, internal "transfer functions" as it were, against our memory maps and make conclusions. The question then becomes the degree of matching to the expected response.

The ear/brain system also brings context into the equation. As we listen to a sound we evaluate it with respect to the surroundings. Does this sound normal for the given distance from the source? Does the amount of reverberation seem in scale with the size of the space? Our trained ear/brain system knows that we will have to use evaluative standards that include contextually derived conclusions such as "acceptable for the back rows of a basketball arena."

The ultimate memory map for sound engineers is their personal "reference" program material. We all have love/hate relationships with our reference materials. We have heard them so many times in so many different places on so many different systems that they are permanently burned into our brains. The new system will have to satisfy our ears against this standard.

The process of ear/brain training is a lifelong endeavor. There is no greater aid to this process than our complex audio analyzer. The linkage between what we see on the screen, to our perceptions of a known source such as our reference CD close the loop on the learning process. If we see something really strange on the analyzer we can run to the mic position and listen. If we hear something strange as we walk the building, we can move a mic there and investigate. This helps us to read the traces in context and identify transitional trends in the room. At the end of the optimization process I have often made adjustments to the sound system based on the walk through while listening to my reference material. In most cases the adjustments are minor, but in all cases I try to learn what it was that I missed in the process of interpreting the data. If after-optimization adjustments are needed, I know that the answer can be found in the data. This becomes part of the lesson to be carried forward to the next job.

There is a long-standing question in the industry: which is better, analyzers or ears? This is a red herring. Consider the following question: what would you choose to pound a nail into a piece of wood? A hammer or your hand? The answer is both. The hand alone is slow and painful. The hammer is useless without the hand, but together they make a powerful combination. In the end, neither tool is of any use without a brain to guide the process. So it is with analyzers and ears....

Ongoing Optimization

Using Program Material as the Source

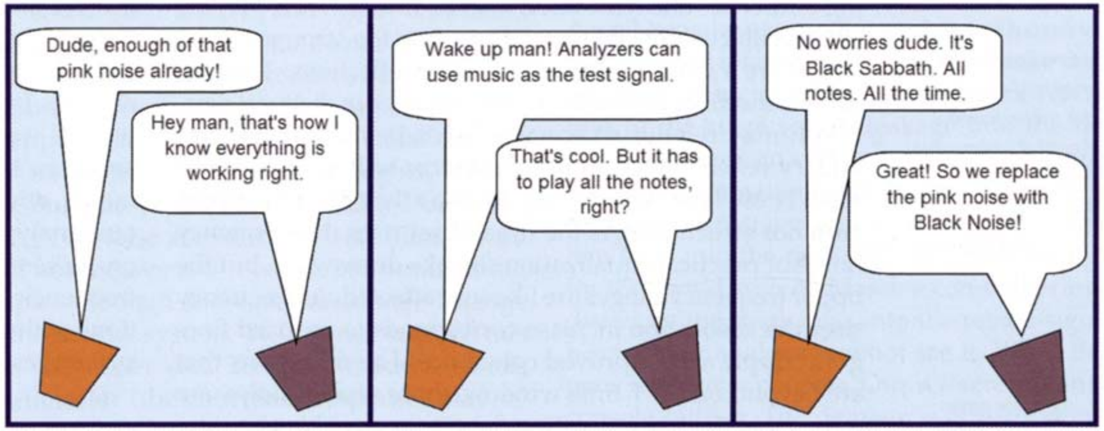
One of the powerful features of dual-channel FFT analysis is the ability to perform accurate transfer function measurements using an unknown source material. We are "source independent" (which is the origin of the acronym SIM™). Source independence means that we can continue to analyze the response of an occupied hall during a performance in progress, using the program material as the reference

Perspectives: On more occasions than I would like to admit I have found myself with an unoptimized system as the rehearsal is starting. Without a measurement system based on transfer functions I would have been up the creek without a paddle. By using transfer function measurements 1 have been able to get the job done without disrupting anyone but the mixing engineer, who has to wait until I have gotten the main system done before they can do any serious work. Using transfer function measurement systems makes setting or checking delays a cinch at any time there is program material.

Alexander Yuill-Thornton II (Thorny)

Perspectives: After tuning the system, the symphony began rehearsals in the afternoon and shortly after, Ray Charles arrived. Through the afternoon, I kept planted on the analyzer. Few filters were used on the main array, however one typical cut was a fairly deep one at about 200 Hz. As rehearsal progressed, I clearly noticed the room response change, needing less cut from the 200Hz filter. I then made the adjustment gently while communicating with the mixing engineer. A little later, it continued to change, and less cut was again applied. I also needed less correction on one of the HF cuts I was using. Now, I started to question myself, why would it be changing? There were still no people in the arena? I then pulled my head out of the analyzer screen and started looking around, and then noticed that the arena staff were quietly installing the floor seating in front of the stage.

Trap `n Zoid by 606



source. The independence has its limits. We will need a source that excites all frequencies, eventually. We just don't care about the order. If the source has low spectral density we will have to wait a long time to get data. If the data is dense we can move almost as quickly as pink noise.

The result is an ongoing optimization process that provides continual unobtrusive monitoring of the system response. Once the sound system is in operation, the control is in the hands of the artistic sector: the mix engineer. To mention that this is a potentially high stress situation would be an understatement. Huge numbers of decisions must be made in real time, and cues must be executed. Ongoing optimization allows objectivity to remain in place to aid the navigational process.

There are a number of ways in which continuing optimization can aid the artistic process. The first is the detection of changes in the response and the sharing of that information. Action can be taken in some cases to restore the response, which allows the mixer to refrain from having to conduct a massive remix effort for a show in progress. In cases where remedial options are limited it can be helpful to the mixer simply to be informed of the differences so that this can be considered in the mix.

There is a lot that can change between the setup time and show time, and even over the course of a performance. Only a portion of those changes are detectible with our analyzer, and a portion of those, in turn, will be treatable with ongoing optimization.

The changes from setup to show time for a single program channel:

1. dynamic room acoustics (changes due to the audience presence)
2. dynamic transmission and summation environmental conditions (changes due to temperature and humidity)
3. leakage from stage emission sources (the band etc.)
4. leakage from stage transmission sources (stage monitors etc.)
5. re-entry summation from the speaker system into the stage mics
6. duplicate entry summation (leakage) between the stage mics.

Audience Presence

The presence of an audience is almost assured to increase the overall absorption in the hall. The effects are not evenly distributed. The floor reflections will undergo the largest change, and the ceiling the least. The extent of the changes will depend largely upon the nature of the seating area prior to occupancy. The most extreme case is the transition

In addition, the arena staff were very quietly bringing the drapes in and closing off the upper level of the arena. This was an amazing realization that day. 1 then fully understood why this tool was so important to live sound, and how much loudspeaker systems can change in a space due to a wide variety of real-world variables.

John Monitto

Perspectives: Some rooms don't change much as the audience fills the space, while others go through major changes. Some of these changes are caused by reflections from the seating area while others are caused by energy that is trapped by architectural features of the space. In both cases the audience can change what is happening. The first can change the. environment on stage for the better, while the other can make some low-frequency adjustments less necessary. The first will creep into measurements but should be generally ignored, while the second can make a big difference in how the low end sounds. If you try to correct for the seat reflections you might have a rather rude awakening after the show starts with the audience in place.

Alexander Yuill-Thornton II
(Thorny)

from an unfurnished hard floor to a densely packed standing audience. The least effect will be found when playing a hall with cushioned seats. In any case the strongest local effects are a modification of the early reflections. This can cause a modification of the ripple variance in the local area, which may require an equalization adjustment. The secondary reflections from other surfaces will be reduced as well. These later reflections will also affect the local ripple variance structures. As the time offset rises the frequency range of practical equalization shrinks downward, but the upper frequency ranges are likewise affected. As frequency rises the absorption increase corresponds to reduced fine-grain ripple and improved coherence. The reflections that are beyond our FFT time window, those seen as noise, are reduced in the room, creating the rise in signal/noise ratio (coherence).

Another factor is the rise in noise level due to audience presence. The screaming of fans, the singing of the congregation and the mandatory coughing during quiet music passages will all degrade the quality of our data. Generally speaking the louder the audience gets, the less it matters what our analyzer says. If they are screaming for the band we are probably OK. If they are screaming at us, then looking at the analyzer might not be the wisest course.

Temperature and Humidity

Temperature and humidity will change both the direct and reflected paths. Increased temperature changes the reflection structure as if the room has shrunk. The direct sound arrives faster to the listener, but so do the reflections (see Fig 2.98). The temperature change may cause the sound speed to change evenly by 1 per cent, but results in each path being changed by a different amount of time. The deck is thereby reshuffled for all of the reflection based ripple variance, which is based on time offset as an absolute number, NOT a percentage. The relationship of mains and delay speakers will also undergo changes as the transmission speed changes. This is well known, and some manufacturers of delay lines have even incorporated temperature sensors to track the changes and automatically compensate. The speaker delay offset over temperature is more easily visualized than the reflection changes, even though they are the same mechanism. A delayed system changes from synchronized to early (or late). A reflection changes from late to not as late (or later). Think of it this way: if we forget to set the delay for our underbalcony speakers they will still change their relationship to the mains when the temperature changes. We can expect our analyzer to see the following response differences on its screens: a redistribution of ripple structure center frequencies, a change in the amplitude range of some portions of the ripple variance, modified coherence and delay system responses that need some fine-tuning.

Humidity change acts like a moving filter in the HF range. Once again the effects are local. In this case, the amount of change scales with the transmission distance. Longer throws have proportionally stronger air absorption effects, and so the changes are more severe over distance. A short throw speaker would see only minimal differences even with large relative humidity variance. By contrast a long throw system will see substantive movement in the HF response. This precludes the option of a global master filter that compensates for the complete system. Any corrective measures will need to take the relative distance into account. Our analyzer can expect to see the following response differences on its screens: a change in the HF response.

Stage Leakage

The presence of live performers on stage presents the opportunity for leakage into the sound system coverage area. This was discussed in some depth in Chapter 4. The amount of leakage from the stage can change on a moment-to-moment basis. Listeners may have a difficult time discerning whether they are hearing the stage leakage or the sound system, and this may be a highly desirable effect. Alternatively, out-of-control band gear and stage monitors are the mix engineer's worst nightmare, as they are unable to maintain control of the mix due to the lack of separation.

Perspectives If during a concert you see a sudden change of frequency response, rather than rush into a correction, first try to understand what is causing it. It will save you the work and embarrassment of correcting something that might be caused by wind or some other weather phenomena which are not correctable by EQ. Move slowly but with confidence.

Miguel Lourtie

Perspectives A cheap trick for setting delay times is to get a pair of walkie-talkies. Using the headphone out on one unit as the. microphone in on the analyzer, send the other one out into the area where the majority of the delay listeners will be and key open the mic. Turn on the main speakers and note the delay time. Now turn off the main speakers and turn on the delay speakers. Insert delay on the delay speakers until it matches the noted delay of the main system. Use delay finder again on both systems to verify correct alignment. This is only useable for delay settings.

Don (Dr Don) Pearson

From our measurement point of view, even the artistically pleasing blend of stage sources and lightly reinforced instruments are crippling to our data acquisition. Leakage is contamination of our data. It reduces our reliability and it reduces our treatment options. If a peak in some frequency range appears in the sound system transfer function frequency response when the show begins, we are called to action. But if the peak is due to leakage from stage sources, this may be a false alarm. An inverse filter placed at the peak will not remove it if it was sent by something that does not pass through the equalizer. Stage leakage short-circuits the sound system. We cannot control this any more than the mixer can. To make matters worse, our polygraph detector (the coherence response) may be fooled by the leakage. The waveform contained in the leaked acoustical signal will be present in our electrical response from the console. How did it get there? Through all of the microphones. Since the waveform is recognized, the coherence can still remain high, and therefore the peak has the appearance of a treatable modification of the response, except that it will not go away, no matter how much equalization we throw at it.

There is no single means of detecting stage leakage. One is to try and equalize it. If it responds, we are fine. If not, it might be leakage. Fishing expeditions like this, however, are not appreciated by the mixer, and should be a measure of last resort. The first consideration is the obvious: what do our ears and eyes tell us? We are in a small club and the guitarist has four Marshall stacks. This is going to be a leakage issue. The next item to consider is plausibility. What mechanism could cause a 10 dB peak to arise between setup and show time? If this was the result of the audience presence, then why do the peaks keep changing shape? Before we grab an equalizer knob or touch a delay we must consider how the change we see could be attributed to the changes in transmission and summation due to the audience presence and environmental conditions. Large-scale changes and song-to-song instability point toward the band. Less dramatic and more stable changes point to the sound system in the room. Those are the ones worth venturing after.

Stage Microphone Summation

A well - tuned system can sound great in every seat when we listen to our reference CD, and then take on a completely different character when the band comes on stage. In addition to the stage source leakage into the house just discussed, there is an even more insidious source of trouble: mic/mic summation (this is also discussed in Chapter 4). Leakage from stage sources or from the main sound system back into the mics becomes part of the source signal for the sound system. Massive comb filtering can be introduced by the stage mics into the mixed signal. We can all hear it. Our analyzer cannot see it. Why? Because it happens in the mix console. Our reference point begins at the console output. The damage is already done. The easiest way to detect this is by reverse logic: the fact that we can hear it, but the analyzer does not see it is the key. We are well versed in the standard progressions of ripple variance in our speaker system. One thing is guaranteed: if we stay put, the ripple stays put. If we move, it moves. The opposite may be true of mic/mic summation. The ripple progression there is governed by relationships on stage. If the stage sources move, the ripple moves. If we are sitting still and the ripple is moving, we have summation progressions changing before they get to the sound system (unless the wind is blowing, which will cause a similar effect). The most effective way of isolating this effect is to keep good records. Store the data at a given location before the performance (or even the sound check). If the transfer function response remains fairly stable in the face of large-scale changes in how it sounds to our ears, we have strong evidence that the solution lies on the other side of the art/science line.

This information can be put to use. Letting the mixer know that we cannot solve this, and that they must deal with it, can provide a huge benefit for the mixer. Combing in one mic channel can make the entire PA sound like it sank underwater. The problem must be solved in the mixer. If we attempt to treat it in the sound system we would be making a huge mistake. First, we would have to use our ears to do this, since it is not seen by the analyzer. That would not be a problem except for the fact that we have now tuned the PA for a stage mic, not the room.

Perspectives Tips from a guy who has made more mistakes optimizing sound systems than you have.

- *Never use arbitrary mic positions for optimization.*
- *The job of the optimization engineer is not tonal quality it is tonal equality.*
- *Always think about sound in its complex form level, frequency, time (phase) and wavelength (Chapter 1).*
- *If you are going to gamble with summation, make sure you know how to win (Chapter 2).*
- *Know when to hold 'em (coupling zone). Know when to fold 'em (isolation zone).*
- *Unlimited range requires a coupled array with some measure of angular isolation (Chapters 2,6).*
- *The range of uncoupled arrays must be limited (Chapters 2, 6).*
- *Where two sound sources meet at equal level, they must be made to meet at equal time (the phase-aligned crossover) (Chapter 2).*
- *Know every acoustic crossover, spectral and spatial, in your system (Chapter 2).*
- *Confine stereo coverage to areas where stereo perception is possible (Chapter 3).*
- *Don't count on the room to fix what is wrong with your sound system (Chapter 4).*
- *Never trust an acoustical prediction program that does not incorporate phase (Chapter 5).*
- *just say* no *to zero degrees in coupled arrays (Chapter 6).*
- *"Line array" speakers have fantastic potential as long they are used as a point source (Chapter 6).*
- *Minimum variance in the coupled point source is achieved by beam*

Every other source into the sound system is now detuned. If we are going to use our ears to make changes that are not indicated by the analyzer we have joined the artistic staff. We will need to inform the mix engineer that we are mixing the show along with them. This might not go over so well. We should always use our ears, but one of the most important ear training skills we must learn is to discern which side of the art/science line we are hearing. Our side of the line is primarily concerned with audible changes over the spatial geometry of the room. If it sounds bad, somewhere, fix it. If it sounds the same bad everywhere, don't fix it. Tell the mixer to fix it.

Feedback

The worst-case scenario of re-entry summation into the microphones is feedback. Feedback is very difficult to detect in a transfer function measurement. Why? Because it is present in both the electrical reference and the acoustic signals. The only hint is that the frequency in question may suddenly have perfect coherence. This is not much help. Feedback detection can be conducted in the FFT analyzer in single-channel mode, where we are looking at the spectrum in absolute terms. The feedback frequency will rise above the crowd and can be identified after it is too late.

Multichannel Program Material

Everything just discussed pertains to a single channel of program. Stereo and other multichannel formats mix the sound in the acoustic space. As if it were not difficult enough already to get clear and stable data, we now add the complication of leakage between sound system channels. Our electrical perspective is always some particular channel. Our acoustical perspective will be a shared perspective that includes related and unrelated material from other channels. Stereo, for example, is a changing mix of related and unrelated signals. At one moment the dominant signals are shared between the systems, at the next moment the sound in our channel of interest is firmly in control, and then later the opposite side is leaking into our

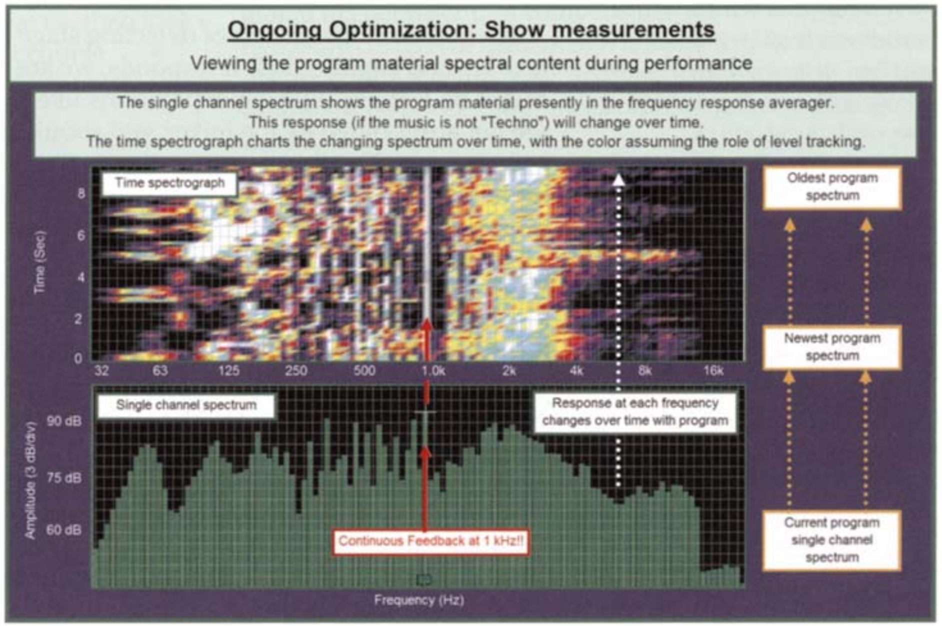


Figure 10.45 Time spectrograph showing feedback (courtesy of SIA-SMAART)

- *spreading in the highs and beam concentration in the lows (Chapter 6).*
- *Remember that the coverage of single speakers isn't fast food — think of aspect ratio, not pizza (Chapter 6).*
- *There are no symmetric solutions for asymmetric problems, and vice versa (Chapter 6).*
- *Asymmetry must be met with an equal and opposite asymmetry (Chapter 6).*
- *Every speaker element, no matter how big or small, plays an individual role, and that solitary identity is never lost (Chapters 6, 7).*
- *Line array is a type of speaker configuration, not a type of speaker (Chapters 2, 6, 7).*
- *Always use the highest-frequency resolution possible for equalization and concentrate on the envelope (Chapter 8).*
- *I would rather leave a system verified and uncalibrated than calibrated and unverified. (Chapter 9).*
- *Every mic position has a defined calibration purpose, and every calibration operation needs a defined mic position (Chapter 10).*
- *The mix position is the only self-calibrating seat in the hall (Chapter 10).*
- *The mix engineer delivers the band to you. You deliver the band to the audience (Chapter 10).*
- *Think globally. Act locally. The best global solutions are a combination of the best local solutions. Optimization solutions are all enacted locally (Chapter 10).*
- *Do not uncalibrate your delays for the precedence effect (Chapter 10).*
- *Don't humiliate the client. Ever. When problems are found in the client's design, it is vitally*

space. Stereo is a wonderfully desirable listening experience. Unfortunately it makes for a very challenging optimization environment. There is no way to win, so let's get that over with right away. The exact center position of an exactly matched stereo system will be the only position free of ripple variance from the overlap of the two systems. It will, however, have the highest level variance over the mix. Signals panned toward the opposite channel will have a small electrical reference level to compare to the acoustic arrival, and will be seen as gain in the transfer function response. If the signal is not fully panned, the console reference signal will recognize the data as valid, which will confuse our coherence response. The coherence will be degraded, however, as the changes in the mix cause changes in the response. The instability will reduce the coherence. The relevant question is: what can we do with this data? It is not stable. Should we be turning equalizer and level controls to stabilize it? Of course not. It is stereo. It is supposed to be changing. So if we cannot get a stable frequency response we have limited use for the mic there. One could argue that we need the mic to make sure that things are not changing at the mix position. If it is stereo, all our mic can do is prove that things are changing at the mix position.

Show Mic Positions

Much emphasis has been placed on the importance of measurement mic locations. Unfortunately, it is rarely possible for these positions to be useable when the hall is occupied. Most often we will be given very limited choices, a fact that greatly reduces the amount and quality of the data available to us during a performance. The role of the mics in a show context differs from previous roles. We will no longer need to establish speaker position splay angles. Therefore OFFAX and XOVR positions will be retired immediately. The principal task ahead is the monitoring of the frequency response changes which result from the dynamic environmental and acoustic properties. If we have an ONAX position on the main system with a high degree of isolation from other subsystems and other channels of sound, we can enjoy a high degree of confidence in our data. In such cases, we can compare the performance response to data captured before the show and make changes to as required to establish continuity. If an isolated ONAX position is not available we are left with fallback positions. The first level would be a less isolated position that has substantial contributions from a related subsystem. An example of this would be a position in the coverage of the main long throw array, but near the area of the spatial crossover into the middle throw system. Adjustments made here will have only a partial effect on the combined response due to the strength of the neighboring system. XOVR positions are the least useable since they are far too volatile to serve an equalization function during the show, just as they were not used for equalization during the setup.

This is a show. We will not be selectively muting parts of the system so that we can get a clear look at the behavior of a particular subsystem. Everything is on all the time. Even the best ONAX position for a subsystem has a very limited range of use. Only the frequency range where that system enjoys dominance will respond to independent adjustment. The further down the hierarchy that our subsystems fall, the less we can do for them in a show context. The low-frequency range is the most likely to be dominated by the main system in the local area. We cannot expect an LF filter change in the subsystem equalizer to exert much control over the combined response.

Delayed systems can be viewed during performances in the hope of maintaining synchronicity under changing environmental conditions. This requires a mic positioned at the spatial XOVR for the delay and the mains. Such mic can be used for time only. This is a very difficult practice in the field, since the delay speakers are set at a low level that blends in with the mains and can be difficult to spot. In any case, if the impulse response is able to reveal the two arrivals, this can be acted on with some confidence. An alternative method for those situations where a mic is not practical in the delay area: make note of the propagation delays before the show at various locations. Compare these to the current propagation delay. Compute the percentage of time change. Modify the delay lines by the percentage change.

The location with the highest probability of show time access, the mix position, is also one of the most challenged.

important to find ways for them to save face.

- *Always disclose your recommendations, actions and the reasons behind them to your client.*
- *Don't let your ego or your desire to be right influence your interpretation of the data or your decisions.*
- *The combination of ears, an analyzer and an experienced open mind is the best analysis* system.

606

In stereo systems the mix position will lack channel isolation and therefore be hard-pressed to establish a stable response. Positions on the outermost horizontal edges of the mix position can find a more isolated response, and provide a clearer picture of "before and after" response. These side positions will have the opposite side arrival but, if the system has limited stereo overlap, some degree of control can be established. None of these are perfect, and often the best we can do is monitor a response solely in relative terms (house empty, house full).

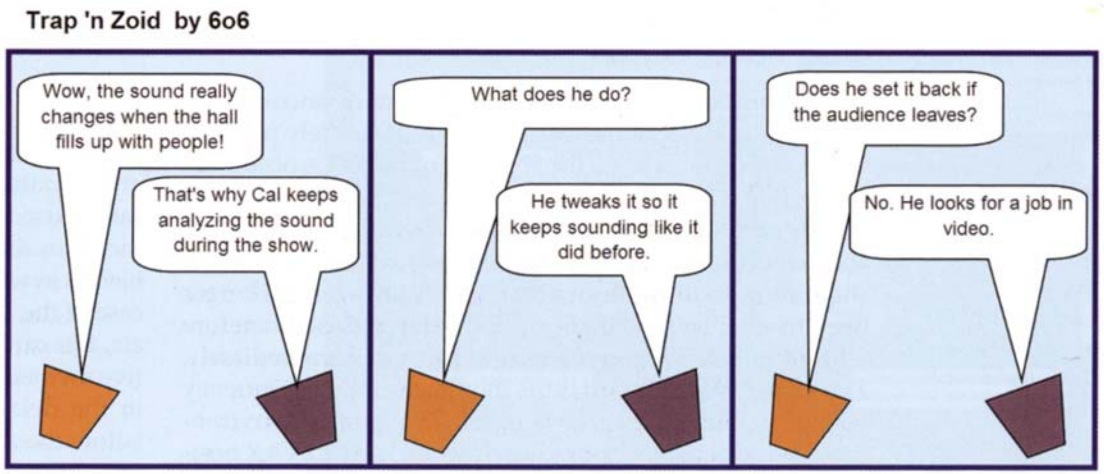
Systems with highly overlapped multichannel interaction are beyond the capability of in-show measurement. Musical theater is routinely done with separate music and vocal systems. There is no way to untangle these in the acoustic space.

Subwoofers as a Program Channel

When subwoofers are driven as an auxiliary channel the low-frequency range is removed from our view. Like the music/voice systems just discussed, the subwoofer range cannot be untangled from the other signals. A transfer function using the mains as a source sees the subs as an unstable, uncontrolled contamination. If the subwoofer send is used as the source the opposite occurs. In either case the electrical reference extends to the full range, so there is no way to separate the acoustic responses.

Final Word

In the end we can see that the occupied hall with the band on stage is not a favorable condition for system optimization. We will be well served to make sure to have performed our optimization before the doors open, so that we can limit our in-show needs as much as possible. The ability to measure with the program material allows us to continue the optimization, but does not mean that we can wait until the band is on stage to begin the process. We still need dedicated time to assemble the individual blocks into a complete system. Only fully assembled systems are ready for the show.



Afterword

This concludes my transmission, for the moment. The cycle of design and optimization has been and will continue to be a learning experience. This is still a young field and the potential for growth is huge. For me personally, I feel that each day has the potential for some new discovery. Rare is the day when that potential is not reached. Over twenty years later "still learning" is the order of the day. I know now why doctors refer to themselves as "practicing" medicine.

The most often-asked question I get about this field goes something like this: "Unlike you, I don't have perfect clients with unlimited budgets, unlimited tools, and unlimited time and support for system tuning. If we can't do it all, what should we do then?"

First of all, I have never had this client, but would love to. Every job happens in the real world, and every job requires us to prioritize and "triage." I have never had the opportunity, on any single job, to perform all of the design and optimization steps shown in this book. I have, however, used all of them over the course of time. They are all in my playbook, ready to be put to use as the situation requires. A coach is not required to use all of the players in their arsenal. They must, however, be prepared to read the situation on the field and have the means at hand to deal with whatever contingencies arise. This book strives to bring into focus the nature of the adversarial forces out there in the practical world of our sound reinforcement environment. This knowledge alone is a powerful ally, even without an analyzer.

In the end we will need to provide our clients with a complete combined system. If we have to streamline the process to get there in the time allowed, we must do so consciously. A skipped step is a leap of faith and a calculated gamble. We need to maintain clear knowledge of where those leaps are, lest they come back to haunt us.

What is the most important? This is a matter of opinion. For me personally, it is like food. We need a variety of the highest-quality ingredients. Good speakers and signal processing. What is the next most important? This is also a matter of opinion. For me it is like real estate. The top priority: location, location, location. Good placement, good angles and good architecture. Level, delay and EQ setting are the finishing processes.

Even more important, however, is maintaining perspective in our role in the big picture. We are members of a multifaceted team. We are there to provide a service to the clients on many levels. The importance of personal relations cannot be overstated. In some cases we are our own clients, stepping out of our lab coats, putting on the artist's beret and mixing the show.

The meeting point between the scientific and artistic sides of our world is the optimized design.

Glossary

Absorption coefficient This indicates how much sound energy will be lost during the transition at a surface. Absorption coefficients range from a maximum of 1.00 (an open window) to 0.00 (100 per cent reflection).

Active balanced interconnection A balanced line connection to or from a powered (active) input or output device.

Active electronic device An audio device that receives power from an external source (or battery) in order to carry out its signal processing functions. An active device is capable of providing amplification of the signal.

Air absorption loss High-frequency attenuation that accrues over transmission distance in air. The humidity, ambient temperature and atmospheric pressure all play a part in the parameters of this filter function.

Amplifier (power) An active electronic transmission device with line level input and speaker level output. The power amplifier has sufficient voltage and current gain to drive a loudspeaker.

Amplitude The level component of the audio waveform, also referred to as **magnitude.** Amplitude can be expressed in absolute or relative terms.

Amplitude threshold An optional feature of transfer analyzers that allows for the analysis to be suspended when insufficient data is presented at the analyzer inputs.

Array A configuration of sound sources defined by their degree of separation and angular orientation.

Aspect ratio A common term in architecture to describe a space as a ratio of length vs. width (or height). This term is also applied for the coverage shape of speakers (interchangeably termed the **forward aspect ratio** here).

Asymmetric Having dissimilar response characteristics in either direction from a defined center line.

Averaging (optical) The finding of a representative response over an area by viewing the individual responses from various locations.

Averaging (signal) A mathematical process of complex audio analyzers that takes multiple data samples and performs complex division to acquire a statistically more accurate calculation of the response.

Averaging (spatial) The finding of a representative response over an area by averaging the individual responses from various locations into a single response.

Balanced The standard two-conductor audio signal transmission configuration chosen for its noise immunity. This is suitable for long distances.

Bandwidth Describes the frequency span of a filter function (in Hz).

Beam concentration The behavior of speaker array elements when they have a high proportion of overlap. Beam concentration is characterized by a narrowing of the coverage area with maximum power addition.

Beam spreading The behavior of speaker array elements when they have a high proportion of isolation. Beam spreading is characterized by a widening of the coverage area with minimal power addition.

Beam steering A technique of asymmetric delay tapering used in subwoofer arrays intended to control the coverage pattern.

Beamwidth A characterization of speaker directional response over frequency. The beamwidth plot shows coverage angle (—6 dB) over frequency.

Binaural localization Horizontal localization mechanism which is driven by the difference in arrivals between the two ears.

Calibration The process of sound system testing and measurement focused upon setting the fine adjust parameters for the system such as equalization, relative level, delay, acoustic evaluation and speaker position adjustments. This process proceeds after the first stage of optimization (verification) has been completed.

Cancellation zone The inverse of the coupling zone. The combination is subtractive only. Phase offset must be between 120 and 180 degrees to prevent addition.

Cardioid (microphones) Unidirectional microphones commonly used on stage. The cardioid action is the result of cancellation zone summation behind the microphone derived from the combination of forward and rear entry of sound at the diaphragm.

Cardioid (subwoofer arrays) A configuration of standard subwoofer elements in a manner which creates a cardioid pattern.

Cardioid (subwoofers) A multidriver, low-frequency system with unidirectional response. The cardioid action is the result of cancellation zone summation behind the speaker derived from the combination of forward and rear firing drivers.

Channel A distinct audio waveform source, such as left and right, surrounds or a special source effect. Each channel must be optimized separately.

Clipping Distortion of the audio waveform that occurs when the signal is driven beyond the linear operating range of a device.

Coherence A measure of the ratio of signal to noise in an FFT transfer function measurement.

Combing zone The summation zone having less than 4 dB of isolation and an unspecified amount of phase offset. Combing zone interaction has the highest ripple variance.

Combining zone The summation zone having between 4 and 10 dB of isolation and an unspecified amount of phase offset. The ripple variance is less than ±6 dB.

Compensated unity splay angle The unity splay angle between array elements with asymmetric relative levels.

Complementary phase equalization The process of creating an inverse response in both amplitude and phase.

Complex audio analyzer A device that performs complex mathematical analysis to provide both the amplitude and phase data relevant to an audio system.

Composite point source The combination of multiple array elements into a virtual single symmetric array element. The components that comprise a composite point source must be matched in level and splay angle.

Compression A slow-acting reduction of the audio signal dynamic range. This is done to prevent clipping or for driver protection.

Constant bandwidth A linear rendering of bandwidth, with each filter (or frequency spacing) having the same bandwidth expressed in Hz. The FFT calculates filters with constant bandwidth.

Constant percentage bandwidth A logarithmic rendering of bandwidth, with each filter (or frequency spacing) having the same percentage bandwidth expressed in octaves, e.g. 1/3rd octave. The RTA filters are constant percentage bandwidth.

Coupled (arrays) Arrays with elements that are within close proximity, i.e. within a single wavelength over a majority of the frequency range of the elements.

Coupling zone The summation zone where the combination of signals is additive only. The phase offset must be between 0 and 120 degrees to prevent subtraction.

Coverage angle The angular spread from on-axis (0 dB) to the —6 dB points on either side.

Coverage bow A method for evaluating speaker coverage which incorporates the equal level contours. The on-axis far (—6 dB) and off-axis near (—6 dB) points are linked by the coverage bow.

Coverage pattern The shape of equal relative level around a sound source for a given propagation plane.

Crest factor The term used to describe the peak to RMS ratio of a waveform.

Critical bandwidth The frequency resolution to which tonal character is audible. 1/6th octave is a typical published value.

Crossover (acoustic) The point where two separate sound sources combine together at equal level.

Crossover (asymmetric) An acoustic crossover where one of the combined elements has different properties than the other. For spectral crossovers this includes level, filter type and speaker parameters; for spatial crossovers this includes level, angle and speaker type.

Crossover (class) A classification based on the combined level at the crossover relative to the isolated areas. This includes unity (0 dB), gap (less than 0 dB) and overlap (more than 0 dB).

Crossover **(order)** The slope rates of the individual elements which combine at a crossover. As the slope rates rise, the crossover order increases. Crossovers may be asymmetric, containing elements with different slope orders.

Crossover (phase-aligned) An acoustic crossover that is matched in both level and phase.

Crossover (spatial) An acoustic crossover in a spatial plane. The spatial crossover point is the location where the two elements combine at equal level.

Crossover (spectral) An acoustic crossover in the frequency domain. The spectral crossover point is the frequency where the high and low drivers operate at equal level.

Crossover hunting The process of finding the asymmetric spatial crossover during the calibration of a sound system.

Cycles **per** second The frequency of an audio signal measured in hertz (Hz).

dB SPL (sound pressure **level)** The quantity of sound level relative to the threshold of human hearing.

dBV A measure of voltage relative to a standard value of 1 volt RMS.

Decibel **(dB)** A unit that describes a ratio between two measures of sound. The decibel is a logarithmic scaling system used to describe ratios with a very large range of values. The decibel is 1/10th of a bel, which is the logarithm of the ratio of two powers.

Dedicated speaker controller An active signal processing device that is manufactured with the optimal settings for a particular model of loudspeaker.

Delay **line** An active electronic transmission device (usually digital) that allows the user to delay the signal for a selected time period.

Diffraction The ability of sound to pass around objects or through openings.

Diffusion The reflection of sound in a manner in which it is scattered in different directions over frequency.

Digital signal processor (DSP) A signal processing device with a wide variety of capabilities, including level, equalization, delay, limiting, compression and frequency division.

Directivity factor (Q) A linear version of the directivity index. The DI value is the 10 log equivalent of the Q factor value. These two values (DI and Q) can be plotted on the same graph with different vertical axis numberings.

Directivity index (DI) The ratio of the amount of energy forward of the speaker to the energy that would be present if the speaker was omnidirectional.

Displacement (source) The physical distance between two sound sources. This fixed displacement will affect the interactions in a given orientation by a fixed amount of time offset.

Displacement (wavelength) The distance between two sound sources expressed as a function of proportional displacement over frequency. The fixed source displacement will affect all interactions in a given orientation by different amounts of wavelength displacement over frequency.

Driver An individual loudspeaker component that covers only a limited frequency range. A speaker enclosure may contain various drivers with dedicated ranges that combine to become a speaker system.

Dynamic range The range between the maximum linear operational level and the noise floor.

Echo perception The listener's subjective experience of the direct sound and late arrivals as being distinct and separate entities.

Element A single sound source within an array.

Emission The origination of sound from a natural source.

End-fire array An array of multiple of subwoofers, placed in a line, one behind the other, with a specific spacing and delay strategy in a timed sequence which creates forward addition and rearward subtraction.

Energy-time curve (ETC) A logarithmic expression (vertical scale) of the impulse response.

Envelope The audible shape of the spectrum, the tonal character. The envelope follows the widest and highest spectral characteristics of the response and does not incorporate narrow dips and nulls.

Equal level contours A radial rendering of a speaker coverage pattern. The on-axis response is normalized to 0 dB and the equal pressure point over angle is plotted radially.

Equal loudness contours (Fletcher-Munson curves) The non-linear nature of the ear's frequency response over level is expressed by this family of curves.

Equalization The process of tonal compensation with an active (or in ancient times, passive) filter set. Equalization is used here primarily to control spectral tilt and thereby minimize spectral variance.

Equalizer An active electronic transmission device with a set of user settable filters.

False perspective Any of the various ways in which the listener is made aware that they are listening to loudspeakers rather than a natural sound source.

FFT The acronym for fast Fourier transform, describing the process of converting the time record data into frequency response data. Also known as the discrete fourier transform (DFT).

Filter (frequency) The action of a system that causes some frequencies to rise (or fall) above (or below) others. Filters in electronic circuits have a wide variety of types, such as shelving, high pass, band pass, and band reject. Examples of filters in acoustic systems include axial attenuation (directional control) over a space, and air absorption loss.

Filter order The gross classification of filter behavior over frequency, designated as first order, second order, third order etc. As filter order rises the rolloff slope becomes steeper. Each filter order connotes an additional 6 dB of rolloff per octave.

Fixed points per octave FFT (constant Q transform) A quasi-log expression of frequency resolution derived from multiple time records of different lengths.

Forward aspect ratio (FAR) A rectangular rendering of the speaker's forward coverage shape expressed as a ratio of length vs. width (or height). The aspect ratio is the fundamental representation of the building block of speaker design: the single speaker element.

Fourier theorem Any complex waveform can be characterized as a combination of individual sine waves, with defined amplitude and phase components.

Frequency The number of cycles per second given in hertz (Hz).

Frequency divider An active (or passive) electronic device that separates the spectrum into frequency bands which are delivered to different speaker drivers for transmission. The waveform will be reconstituted in the acoustical space at the acoustic crossover.

Frequency response The response of a system in various categories over frequency. Here, these include amplitude, relative amplitude, relative phase and coherence.

Full-scale digital The maximum peak voltage before overload for digital audio systems. The actual voltage is model dependant and is often user adjustable.

Graphic equalizer An active (or passive) electronic transmission device with parallel filters having fixed center frequency and bandwidth, and variable level.

Hygrometer An atmospheric instrument that measures humidity.

Impedance The combination of DC resistance and capacitive (or inductive) reactance for a given source or receiver.

Impulse response A rendering of the calculated response of a system as if it were excited by a perfect impulse. This relative amplitude vs. time display is derived from the FFT transfer function measurement.

Inclinometer A device which measures the vertical angle of a device or surface.

Inter-aural **level** difference (ILD) The difference in level at our ears of a horizontally displaced sound source. This is one of the prime factors in horizontal localization.

Inter-aural time difference (ITD) The difference in time arrival at our ears of a horizontally displaced sound source. This is one of the prime factors in horizontal localization.

Inverse square **law** Sound propagation in free field loses 6 dB of SPL for each doubling of distance from the source.

Isobaric contours See Equal level contours.

Isolation zone The summation zone where there is greater than 10 dB of isolation and an unspecified amount of phase offset. The ripple variance is less than ±3 dB.

Latency The transit time through any device, independent of the user-selected settings.

Line level Standard operating audio transmission signal level. The nominal value is 1 volt (0 dBV) with maximum level around +24 dBV.

Line source A speaker array configuration in which the axial (angular) orientation is identical.

Linear frequency axis An equal space per frequency (Hz) rendering of the frequency axis.

Log frequency axis An equal space per octave rendering of the frequency axis.

Maximum acceptable variance A value of 6 dB in level or spectral variance.

Measurement microphone The type of microphone used for acoustic measurement during system optimization. Typical measurement mics are free field omnidirectional type. They are very flat, with low distortion and high dynamic range.

Mic level Low-level audio signal transmission. The nominal values are typically at least 30 dB below line level.

Minimum variance The response of a sound system over a space, which has the minimum differences (less than 6 dB) in level, spectral tilt and ripple.

Mix position The location in the room (usually) where the mix engineer and mix console are located. One of the 15,000 most important seats in an arena.

Noise (causal) A secondary signal which originated from the same parental waveform but has returned to a summation point too late to meet the stable summation criteria. The solutions to causal noise are within the scope of the optimization engineer (and the acoustician).

Noise (non-causal) A secondary signal that is not a related waveform and therefore will meet none of the stable summation criteria. The solutions to non-causal noise will not be found in the sound system.

Noise floor The level of ambient non-causal noise in an electronic device or a complete system.

Nyquist frequency The highest frequency that can be captured at a given sample rate. This frequency is half the sampling rate.

Offset (level) The difference in level (dB) between two sources at a given point. Level offset that is the result of differences in transmission distance is a ratio that will remain constant over scale.

Offset (phase) The difference in arrival time (degrees of phase shift) between two sources at a given frequency and location. Phase offset is frequency-dependent and will double for an octave rise in frequency.

Offset (time) The difference in arrival time (ms) between two sources at a given point. Time offset that is the result of differences in transmission distance, is linearly derived and will not remain constant over scale.

Offset (wavelength) The difference in arrival time (complete cycles) between two sources at a given frequency and location. Wavelength offset is frequency-dependent and will double for an octave rise in frequency.

Optimization The process of sound system test and measurement inclusive of the verification and calibration stages.

Oscilloscope A waveform analysis device with electrical signals displayed as voltage over time.

Panoramic field The horizontal width of the stereo field. The maximum width of the panoramic field at any given location is the angular spread between the stereo sources.

Panoramic perception The experience of apparent sound sources along the horizontal plane between two speakers, also known as stereo perception.

Parallel pyramid The behavior of coupled loudspeakers with matched angular orientation. The spatial crossovers are stacked sequentially in the form of a pyramid.

Parametric equalizer An active electronic transmission device with parallel filters having variable center frequency, bandwidth and level recommended for system optimization.

Pascal An SPL-based standard used in the characterization of sensitivity for microphones. One pascal equals 94 dB SPL. Microphone sensitivity is commonly defined in mv/pascal.

Peak limiting A fast-acting reduction of the audio signal dynamic range. This is done to prevent clipping or for driver protection.

Peak voltage The highest voltage level in a given waveform (V_{pk}).

Peak-to-peak voltage (Vp-p) The voltage between the positive and negative peaks in a given waveform.

Percentage bandwidth The frequency span of a filter function (in octaves).

Perception The subjective experience of the human hearing system.

Period The length of time it takes for a cycle to be completed. A period is the reciprocal of frequency and is given in seconds, although most commonly expressed in milliseconds (ms).

Pink shift See Spectral tilt.

Phase The radial component of the audio waveform expressed in degrees. For a given frequency, the phase value can be converted to time.

Phase delay A delay value (typically in ms) ascribed to a limited span of frequencies, that allows us to characterize frequency-dependent delay.

Pinna (outer ear) The primary mechanism for vertical localization of sound.

Point destination A speaker array configuration in which the axial orientation is inward from the front of the elements, thereby creating a virtual destination which has a common point in front of the elements.

Point source A speaker array configuration in which the axial orientation is outward from the front of the elements, thereby creating a virtual source which has a common point behind the elements.

Polar plot A rendering of a speaker coverage pattern in radial form. The on-axis response is normalized to O dB and the dB loss over angle is plotted radially.

Polarity The measure of waveform orientation above or below the median line. A device that is "normal" polarity has the same orientation from input to output. A device with reverse polarity has opposite orientations from input to output.

Precedence effect (Haas effect) The location of the perceived sonic image can be manipulated by offsetting our binaural localization mechanisms (ILD and ITD). For example, the sonic image may appear in the direction of an early arrival, even if it is not the louder sound. Multichannel speaker systems can use the precedence effect to manipulate sonic image by independent control of relative time and level.

Pressurization In acoustic transmission, the half of the cycle that is higher than the ambient pressure.

Propagation delay The transit time from a source through a medium to a destination. Our primary focus is on acoustic propagation and refers to the transit time between speaker and a listening position.

Protractor method The standard method of determining coverage pattern and aim points of a single speaker which traces an equidistant arc from the on axis point until it reaches the —6 dB point.

Proximity ratio The difference between the closest and farthest seats in the coverage of a given speaker or array

Pseudo-unity gain The voltage gain of the overall device is unity but the unity is created by as much as 20 dB of gain at the input and a tracking amount of loss at the output. This may result in overload at unexpected levels.

Push-pull This is an output stage in an active balanced output with two identical signals with opposite polarities.

Q (filter) The quality factor of a filter circuit. It is a linear scale representation of the filter bandwidth. As the filter narrows, the Q rises.

Q (speaker) See Directivity factor.

Range ratio The relative lengths between the on-axis points of two speaker elements to their respective destinations. The range ratio (in dB) is used to match the relative level between the speakers.

Rarefaction In acoustic transmission, the half of the cycle that is lower than the ambient pressure.

Ray-tracing model The rendering of predicted speaker response as rays of light emitting from a source.

Real-time analyzer (RTA) An acoustic measurement device which uses a bank of log spaced parallel filters to characterize the spectral response.

Refraction The bending of a sound transmission as it passes through layers of media.

Resolution The detail level presented in the measured data which is the basis of the predictive or measured model.

Resolution (angular) The spacing (in degrees) between the measured data points that comprise the speaker data for an acoustic prediction.

Resolution (frequency) The width of each evenly spaced FFT frequency *"line"* or *"bin"* which is calculated by dividing the *sampling rate,* by the number of samples in the *time window.*

Return ratio The difference (in dB) between the distance from a sound source to the front of the balcony and the farthest area angularly adjacent to that balcony front. This ratio is used in the evaluation of solutions for balcony and under balcony coverage.

Ripple variance The range (dB) between the peaks and dips.

Root mean square (RMS) The voltage in an AC circuit that would provide an equivalent to that found in a DC circuit.

Sampling rate The clock frequency of the analog-to-digital conversion.

Sensitivity (microphone) The amount of voltage created by a given mic for a given SPL at the diaphragm. Typically the value is given as a number of mv/pascal.

Sensitivity (speaker) The amount of SPL for a given speaker at a given drive level and distance. Typically the value is given as a number of dB SPL @ 1 watt/1 meter.

Signal processor Any active electronic transmission device charged with the jobs of equalization, level setting or delay.

Sonic image The perceived location of a sound source, regardless of whether a natural source or loudspeaker are present at that location.

Sonic image distortion The extent to which the perceived sound image differs from the intended sound source.

Spatial perception The subjective characterization of the location of a sound source. For summed signals, the listener's subjective experience of the direct sound and a late arrival is spectrally and spatially modified, such that a distinct point of origin is not perceived.

Speaker level The transmission drive level between amplifiers and speakers. High-level audio signal transmissions are typically expressed in watts, rather than voltage, due to the large amounts of power being transmitted.

Speaker order The gross classification of filter behavior over frequency, designated as first order, second order, third order, etc. As speaker order rises the coverage shape contours become steeper. This is analogous to filter slope order with higher order speakers exhibiting increasingly steep response rolloffs. Each speaker order connotes the approximation of an additional 6 dB of high frequency rolloff per quadrant (90 degrees).

Spectral tilt (pink shift) A spectral distribution in favor of more low-frequency energy than high-frequency energy (or the reverse). An example of spectral tilt is off-axis response.

Spectral variance A substantial difference in spectral response between two locations. Locations with matched spectral tilt would be considered minimum spectral variance.

Splay angle The angular orientation between two array elements.

Stable summation criteria The conditions that must be met for summation behavior to maintain sufficient stability to be perceived as an ongoing effect on the system response. These include the origination from a related waveform and sufficient shared duration at the summing junction.

Stereo perception See Panoramic perception.

Subsystems One of a family of systems that together comprise a single channel of the sound system transmission. Subsystems transmit a related waveform to local areas.

Summation duration The length of time that two frequencies share the same location. In related waveforms the summation duration will depend upon the time offset and the transient nature of the waveform.

Summation zones The five categories of summation interaction behavior based upon the relative level and phase of the summed signals. These zones include the coupling, combining, isolation, cancellation and combing zones, which all differ in their distribution of relative level and phase.

Symmetric Having similar response characteristics in either direction from a defined center line.

Systems (fills) The secondary sound sources for a given signal channel. These are the less powerful subsystems and cover the listening areas not handled by the mains. All fill will operate under lower priority status to the mains.

Systems (mains) The primary sound source for a given signal channel. These are the most powerful subsystems and cover the largest percentage of the listening area. All subsystems will operate under lower priority status to the mains.

TANSTAAFL "There ain't no such thing as a free lunch." The concept, attributed to Robert Heinlein, which teaches that no action or solution can occur in isolation without affecting some other action or solution.

Time bandwidth product **(FFT)** The relationship between the length of the time record and the bandwidth. The relationship is reciprocal, therefore the combined value is always one. A short time record creates a wide bandwidth, while a long time record creates a narrow bandwidth.

Time record **(FFT)** (also called the time window) The period of time over which a waveform is sampled, expressed in ms. FFT lines (bins) are the number of samples (division operations) of the time record.

Tonal perception The subjective characterization of the spectral response of the signal. For summed signals, the listener's subjective experience of the direct sound and late arrivals is as a single spectrally modified response.

Total harmonic distortion **(THD)** A measure of the presence of harmonics added to the original signal (the fundamental) by a particular device.

Transfer function A dual-channel system of audio measurement that compares one channel (the reference) to a second channel (measurement). Transfer function measurement illustrates the difference between the two signals.

Transformer balanced A balanced line connection to or from a passive (transformer) input or output device.

Triage The strategy of prioritizing choices in cases where a single solution cannot provide equal benefit in all circumstances.

Unbalanced Single conductor audio transmission configuration suitable only for short distances due to its lack of noise immunity.

Uncoupled **(arrays)** Sound sources that are not within close proximity, i.e. beyond a single wavelength over a majority of the frequency range of the elements.

Uniformity The extent to which we can create a similar experience for all listeners in the hall.

Unity **gain** The condition of an electronic device whose ratio of input to output is 1 (0 dB).

Unity splay angle The splay angle between two speaker elements that will produce minimum level variance between the on-axis points of the two elements and the spatial crossover.

Variable acoustics An architectural design having adjustable acoustical features which can be changed to suit different uses and sonic requirements.

Variation The inverse of uniformity. Principal forms of variation discussed here are level, spectral and ripple variance.

Verification The process of sound system test and measurement focused upon checking that the sound system is correctly installed and fully operational. This process prepares the system for the final stage of optimization: calibration.

Voltage **gain** The ratio of output voltage to input voltage through a device. This can be described linearly ***(2x,*** 4x) or logarithmically (+6 dB, +12 dB, etc.).

Volt/ohm meters (VOMs) An electronic test instrument that provides AC and DC voltage testing, continuity and short-circuit detection.

Wavelength The physical distance required to complete a cycle while propagating through a particular medium. Typically expressed in meters or feet.

Weighting (averages) The extent to which an individual data sample is given preference in the averaged total. An unweighted averaging scheme gives all samples equal statistical value, while weighted schemes gives certain samples a larger proportion.

Weighting (frequency response) The extent to which individual spectral regions are given preference in the averaged total level. An unweighted frequency response scheme gives all frequencies equal treatment in the averaged total.

Window function **(FFT)** A form a waveform shaping of the time record that prevents odd multiples of the time record length distorting the frequency response calculation.

Wraparound (phase) An artifact of the display properties of the FFT analyzer. The cyclical nature of phase requires the display to recycle the phase values such that 0 degrees and 360 degrees occupy the same vertical scale location. The wraparound occurs at the edges of the rectangular display of the circular function.

Bibliography

Beranek, L. L. (1962), *Music, Acoustics & Architecture,* Wiley.

Beranek, L. L. (2004), *Concert Halls and Opera Houses Music, Acoustics, and Architecture,* Springer.

Cantu, L. (1999), Monaural Hearing and Sound Localization, Austin State University, http://hubel.sfasu.edu/courseinfo/SL99/monaural.html

Cavanaugh, W. J. and Wilkes, J. A. (1999), *Architectural Acoustics Principles and Practice,* Wiley.

Duda, R. O. (1998), Sound Localization Research, San Jose State University, http://www-engr.sjsu.edu/~duda/Duda.html

Everest, F. A. (1994), *The Master Handbook of Acoustics,* TAB Books (a division of McGraw-Hill).

Giddings, P. (1990), *Audio System Design and Installation,* Sams.

Herlufsen, H. (1984), *Dual Channel FFT Analysis (Part 1),* Technical Review, advanced techniques in acoustical, electrical and mechanical measurement, No. 1, Bruel & Kjaer.

Martin, K. D. (1994), A Computational Model of Spatial Hearing, Massachusetts Institute of Technology, http://xenia.media.mit.edu/~kdm/proposal/chapter2_1 .html

McCarthy, B. (1998), *Meyer Sound Design Reference for Sound Reinforcement,* Meyer Sound Laboratories.

Tremaine, H. (1979), *Audio Cyclopedia;* Sams.

Index

Italic page numbers refer to illustrations

%ALCONS *see* Percentage of articulation loss of consonants
"Ab" weighting, 52
1 watt/1 meter, 50-1
2-D drawings, 207-12, *208,* 213, 230
3-D drawings, 208, *208-11,* 213, 219, 230

Absorption:
acoustic modeling, 224, 225-6, 229
acoustic treatment, 319, 452-3, *453*
audience presence, 475-6
calibration procedure, 452-3, *453*
coefficient, 225-6, 226
speaker/room summation, 147-9, *148, 149*
specifications for performance space, 203, 204
uncoupled inward angled surface, 147
Absorption coefficient, 225-6, 226
AC *see* Alternating current
Access points, 403, *404,* 428-30, *429, 431*
Accumulated averages, 371, *371*
Acoustic decibels (dB SPL), 14-15
Acoustic evaluation, 183-205, 452-3
Acoustic modeling, 212-30
Acoustic partitions, 193-4
Acoustic power, 44-5
Acoustic pressure *see* dB SPL
Acoustic transmission, 3, *3,* 44-9, *45*
environmental effects, 46-7, 47, *48*
latency, 19
loudspeakers, 47-59
pressurization and rarefaction, 4
unitless decibels, 15
wavelength, 5-8
Acoustic treatment:
analyzer type, 398
calibration procedure, 452-3, *453,* 467
measurement microphone placement, 442
specification, 319
Acoustic triage, 320, 321-2, 427, 428
Acoustical crossovers, 61, 87-112
amplifier voltage gain, 38
classification, 88-9
crossover asymmetry, 98-101, *98-9,* 108-11, *108-11*
crossover audibility, 95, 98, 111-12, 143
crossover classes, 89-93, *89,* 108-11
crossover progressions, 87-8, *88*
definition, 87
delay lines, 22
frequency, 93
frequency dividers, 26-7
level setting, 50
loudspeaker specifications, 48
speaker/room summation, 140-9
Acoustics:
architecture, 191-205
natural vs. amplified transmission, 183-6
specification, 317-18
Active balanced line, 33, *34*
Active line level devices, 20
Active speakers, 31-2
AES *see* Audio Engineering Society
AES/EBU (AES3), 43-4
Ahnert, Dr Wolfgang, 213
Air pressure, acoustic decibels, 14
"All-in-one" box, 233
All-pass filters, 27
Alternating current (AC), 4, 13
Ambient noise floor, 14
Amplified sound, 183-5
detection, 176-8
emission model, 185-6
evaluation of musical-acoustic quality, 193-8, *195-6*
hybrid transmission, 198-9
reception model, 185-6
room specifications, 202-5

Amplified sound *(contd)*
transmission model, 185-6
transmission path, 186-90, *187-8, 190*
Amplified sound detection, 176-8
Amplifier voltage gain, 38, *39*
Amplifiers, 35-40
active speakers, 31-2
level controls, 40
limiters, 28, 29-30, 29-30
polarity, 37
power and impedance, 36
sensitivity, 38, *39*
speaker cables, 40-2
specifications, 36
verification access, *404*
voltage gain, 38, 38-40, *39*
voltage level reference chart, *12*
Amplitude:
decibels, 10-16
definition, 4
Fourier transform, 363
graphic equalizers, 23
parametric filters, 22
shelving filters, 23, 26
summation, 64-6
terminology chart, *8*
transfer function measurement, 376-7
waveforms, *9,10,10*
Amplitude response, 17-18, *18*
active speakers, 31
complementary phase equalization, 24
line level devices, 21
verification procedure, 415-17, *416*
Amplitude thresholding, 372
Analog electrical transmission, 19-42, *19*
Analysis systems, 399-400
Analyzers:
analysis systems, 399-400
calibration access points, 430-2
complex, 362-97
echo perception zone, 169, *169*
listening, 474
mic/mic summation, 477
real-time, 360-2
self-verification, 402
spatial perception zone, 169, *169*
tonal perception zone, 169, *169*
see also FFT analyzers
Anarchy model, 426
Angular resolution, 213, 214, 215-18, *218-19,* 227
Architectural acoustics, 191-205
Array theory, 79
Arrays:
acoustic modeling, *217, 219,* 229-30
acoustical crossover progressions, *88*
calibration strategies reference, *469-71*
coupled, 113, 115-25
design procedures, 331-43
history of, 232-4
loudspeaker sensitivity, 50-1
measurement microphone position strategies, 447-51, *448, 450*
minimum level variance, 253-65
minimum spectral variance, 265, 268-303
minimum variance menu, 310-13, 322
monoaural systems, 322
order of calibration operations, 467
power vs. variance, 310
program material, 317
refraction, 221
ripple variance progression, 240
room shape, 242-7
scaling, 330-1
spatial acoustic crossovers, 102
speaker position adjustment, 456-8
specification, 318
spectral variance, 237
subsystem types, 325-30
summation, 112-39
types, 112-15, 224
uncoupled, 113, 224, 125-39
variation, 232-4
see also Hybrid line/point source arrays; Line arrays; Point destination arrays; Point source arrays
Aspect ratio:
balconies, 348-50
minimum level variance, 242-65
minimum spectral variance, 265-74
Audience:
calibration, 441
listening position height, 443
ongoing optimization, 475-6
Audio analyzer, 355-6
Audio Engineering Society (AES), amplifier polarity, 38
Auditory localization, 153-64
Aural compression, 152
Aural harmonics, 151
Averaging, 370-2
coherence, 387-8, 391
spatial, 432-8, *435*

Balance, 192, 194, 296
Balanced lines:
line level devices, 21, 22, 32-5
power amplifiers, 38
verification access, *404*
Balconies, 346-50, *347, 349-50*
evaluation, 202
multiple main systems, 325, 326
specifications for performance space, 205
see also Over-balcony delays; Under-balcony delays
Balcony bars, 348
Band limiters, 419
Band pass filter, 24
Banded SPL measurements, 15
Bandwidth:
acoustic modeling programs, 215
conversion to Q, *24*
digital audio transmission, 43
equalization filters, 24-6
fast Fourier transform, 363
FFT analyzers, *364,* 365, 367
graphic equalizers, 23
parametric filters, 22-3, 23
shelving filters, 26
specification goals, 315
tonal perception, 166-7
Bathroom plan, 208
Beam concentration, 265-7, 275
coupled line source arrays, 268, 276-9, 281

coupled point source arrays, 271—4, 285-6, 291
hybrid coupled line/point source array, 292
Beam pass-through, 265
asymmetric uncoupled line source, 297
coupled line source array, 276
coupled point source arrays, 273
Beam spreading, 265-7, 275
coupled point source arrays, 271—4, 282-3, *282-3,* 284-5, 287
hybrid coupled line/point source array, 292
uncoupled line source, 297
Beamwidth:
frequency, 56-7, 57
loudspeaker specifications, 48
minimum spectral variance, 265-74, 266
power, 330-1
speaker order, *105,* 309-10
subwoofer arrays, 306
Beat frequencies, 151
Beranek, Leo, 191-2, *195-6*
Berkow, Sam, 201
Binaural localization, 157-64, 169-75
Bit resolution, 43
Blend, 192, 194, *196*
Boost, 23, 24, 26
Boundary interaction, 55
Brilliance, 192, 194, *195*
Budget, specification, 318, 319
BW *see* Bandwidth

"C" weighting, 52
Cable impedance, 32
Cables *see* Interconnection
Calculated semi-anechoic response analyzers, 397-9
Calibration, 424-80
access points, 428-30, *429, 431*
approaches, 426-7
context, 438-41, *439-40*
definition, 424-6
finishing, 473-4
ongoing optimization, 474-80
order of operations, 467-73, *468-73*
oscilloscope, 360
post-calibration verification, 421-3, *421-2*
procedures, 452-67
spatial averaging, 432-8, *435*
strategies reference, *468-71*
subdivision, 428-30
techniques, 426
test reference, *430*
verification, 401, 402-3, 423
Cancellation, 55, 64
polarity, 19
tonal perception, 166-7
Cancellation zones, 71, 72
crossover classes, 92
crossover progressions, 88
filter order, 94
triangulation, 81
Capacitance, 32
Capacitive reactance, 16
Capitalism model, 427
Causal noise, coherence, 388, 389-91, *390*
Ceilings, 203, 205, 207-8, 209
Center clusters, 322, 458
Center frequency, 22, 23, 24
Channels:
multichannel sound, 350-2
program material, 317
signal-processing, 319
specification, 316
stereo systems, 321-2
surround systems, 323
system types, 320-5
Characterizing summation, 226-8
Clarity:
specification goals, 315
see also Definition
Clark, Dave, 201
CLF *see* Common loudspeaker format
Clients, specifications, 314-15
Clipping:
amplified sound detection, 176
amplifier voltage gain, 38
dB SPL continuous, 52
human perception of loudness, 152
limiters, 28, 29
maximum input/output capability, 410
Coherence:
causal noise, 388, 389-91, *390*
definition, 387-9
goals of calibration, 425
non-causal noise, 388, 389-90, 392
stage leakage, 477
transfer function measurement, 386-91, *388*
Comb filtering, 78-9, *78-9*
coherence, 389
coupled line source, 115, *116-18*
crossover audibility, 95
mic/mic summation, 477
tonal, spatial and echo perception, 165, 166, 167-8
uncoupled point destination, 135-8
Combing zones, 71, 72, *74-5, 77*
acoustic modeling programs, 215, *217,* 218
crossover progressions, *88*
filter order, 94
spatial acoustic crossovers, 103
triangulation, 81
Combining zones, 71, 72, 76-7
acoustical crossover progressions, *88*
filter order, 94
spatial acoustic crossovers, 103
triangulation, 81, *82*
unity crossover, 92
Common loudspeaker format (CLF), 218-20
Common (shield), 33, 35, *35*
Common-mode rejection, 33
Compensation delay, 387, *387*
Complementary phase equalization, 24
Composite coupled point source:
balconies, *349-50*
design procedures, 337-41, *339-40*
Compression:
amplified sound detection, 176
human perception of loudness, 152
verification procedure, 418-19
Compromise, 319-20

Computer-aided design (CAD) programs, 207, 208
Conductors, balanced lines, 33
Cone drivers, 55
Constant percentage bandwidth, 363-4
Constant Q transform, 364, 367-8, *369*
Corner frequency, 23, 26-7, 43
Coupled speaker arrays *see* Arrays; Arrays, coupled; Line source arrays; Line source arrays, coupled; Point destination arrays; Point destination arrays, coupled; Point source arrays; Point source arrays, coupled
Coupling zones, 71, *72, 73-7*
coupled line source, *116-18,* 117
crossover asymmetry, 98
crossover classes, 89, *89-92*
crossover progressions, *88*
filter order, 94
spatial acoustic crossovers, 102
triangulation, 81, *82*
wavelength displacement, 84
Coverage angle, 55, *56,* 57-9, *58*
acoustic modeling programs, 218
coupled point source, 120, 257-60, 257, *259,* 283
hybrid coupled line/point source array, 293
line source arrays, 257
maximum acceptable level variance, 249-52
minimum level variance, 242-9, 243, *247,* 283
minimum spectral variance, 266
Coverage bow, *243,* 244, *245, 247,* 278, 279
Coverage patterns, 55-9
acoustic modeling programs, 214
asymmetric, 57-9
loudspeaker specifications, 48
minimum level variance, 242-7, 243, 245
proximity ratio, 246-7
room shape, 242-7
spatial crossovers, 103-11, *106-8*
summation zones geometry, *83*
variation, 234
Coverage shapes:
asymmetric, 243-4, *245,* 334-7, *335-7*
minimum variance, 310
scaling, 331
specification, 317
Crest factor, *8,* 13-14, 151-2
Critical bandwidth, 166-7
Cross section, 207, *208,* 209-12, 209-22
Crossover audibility, 95, 98, 111-12,143
Crossover classes, 89-93, *89,* 113,115
coupled arrays, 115, 226, 120-2, 222
crossover assymetry, 108-11, *108-11*
uncoupled arrays, 125-7, 227, 132, 133, 142
Crossover range, 94, 103-11
Crossover slope, *90,* 142
Crossovers *see* Acoustical crossovers; Spectral dividers
Current, 16-17, 35
Cut, 23, 24, 26
Cycle, 4, *8*

dB *see also* Decibels
dB SPL, 14-15, 45
human perception of loudness, 151-2
log formulas, 11
operating levels, 25
reflections, 45-6
sound level meters, 360
subunits, 14-15
dB SPL continuous, 14-15, 52, 151-2
dB SPL peak (dB SPLpk), 14, 52, 151-2
DB SPL/volt, 51
dB SPL_{pk} *see* dB SPL peak
dBu, 11-12
dBV, 11-13
dBVp-p, 13
dBVpk, 13
dbVRMS, 13
DC *see* Direct current
Decay character, 394
Decibels, 10-16
ratio reference, 22
see also dB
Dedicated speaker controllers, 30-1, 31, 38
Definition (clarity), 192, 194, 295
Delay lines:
active speakers, 31
analog transmission, 22
verification access, *404*
Delay panning, 160-2
Delay setting:
calibration procedure, 463-7, *464-6*
digital audio transmission, 43
measurement microphone placement, 441, 442
Delay systems, 324-5
all-pass filters, 27
balconies, 346-50, *347, 349*
delay setting, 463-7, *464-6*
horizontal localization, 158-64, *158-63*
level setting, *454*
line level devices, 20
ongoing optimization, 479
relative phase cycle, 69-70, *69*
subsystem combinations, 328-30, *329*
temperature and humidity, 476
verification access, *404*
Democracy model, 427, 467
Design:
arrays, 331-43
prediction, 206-31
Deviation, 370
DI *see* Directivity index
Diagonal plane, 343-6, *344-5*
Differential inputs, 21, 33
Diffraction, 223-4, 225
Diffusion, 192, 194, 295
acoustic modeling programs, 223, *224*
specifications for performance space, 203
Digital audio transmission, 8, 19, 42-4
Digital signal processors (DSPs), 30-1, 38
Direct current (DC), 4, 16
Direct sound:
acoustic modeling programs, 220-1
loudness of, 192, 194, 295
Directionality:
dual perspective, 178
microphone reception, 179-80
spatial acoustic crossovers, 102-8

speaker order, 103-8
triangulation, 79
Directivity, 54-9
Directivity factor (Q), 56-7
Directivity index (DI), 56, 58
Displacement effects:
coupled line source, 115, *116*
speaker/room summation, 141
uncoupled line source, 127, 127
uncoupled point destination, *134*
uncoupled point source, *132*
Distance, acoustic transmission loss, 46, *47*
Distance divergence, 209, 209
Distortion:
amplified sound detection, 176, 178
digital audio transmission, 44
linear operating area, 12
Downfill subsystems, 328, 330
Drawings, 207-12
Driver displacement, 95
DSPs *see* Digital signal processors
Dual perspective, 177-8
Dual-channel analyzers *see* FFT analyzers, dual-channel
Duplicate entry summation, 186-90, *190,* 475
Duration, 151
Dynamic range:
attribute of musical-acoustic quality, 192, *196*
digital audio transmission, 43
human perception, 14, 151, 152
limiters, 28-9
variation, 235
verification procedure, 412, *412*
Dynes per square centimeter, 14

Ear, 150, 151-64
Early arrivals, coherence, 390-1
EASE™, 213
Echoes:
coherence, 390-1, *390*
evaluation of musical-acoustic quality, 192, 194, *196*
horizontal localization, 160
impulse response, 393, 395
perception, 164-9, 175
speaker directivity, 54
speaker/room summation, 143
Effects systems, 320
Electromagnetic interference, 33
Electronic crossovers *see* Spectral dividers
Electronic decibels (dBV and dBu), 11-13
Electronic transmission:
analog, 19-42, *19*
digital, 42-4
electronic decibels, 11-13
latency, 19
power, 16-17
summation, 63-4
unitless decibels, 15
voltage, 4
Elevation, 207
Emission, natural vs. amplified sound, 185-6
End-fire subwoofer arrays, 304-5, *304, 309*
Ensemble, 192, 194, *196,* 198
Envelope, 165-7, *166*
Environmental effects:
acoustic modeling programs, 220-1
acoustic transmission loss, 46-7, 47, *48*
speaker/room summation, 147-9, *149*
speaker/speaker summation, 138, *140*
see also Humidity; Temperature
Equal level contours, 56, 56, 58-9, *58-9*
acoustic modeling programs, 214
coverage bow method, 244
forward aspect ratio, 246
minimum level variance, *243-4, 243, 245, 247*
summation zones geometry, 81-4, *83*
Equal loudness contours, 152-3, *153*
Equalization:
active speakers, 31
analyzer type, 398-9
calibration procedure, 458-63, *461-3*
dedicated speaker controllers, 30
digital audio transmission, 43
equal loudness contours, 152-3
evaluation of musical-acoustic quality, 194
measurement microphone placement, 441, 442, 447-9
order of calibration operations, 467
real-time analyzers (RTA), 361
spatial averaging, 434-8
spectral tilt, 238
subsystems, 324-5
tonal, spatial and echo perception, 165, 167-8
Equalizers:
analog transmission, 22-6
line level devices, 20
phase delay, 382, *383*
verification access, *404*
Ethics, calibration, 426-7
Evaluation, 183-205, 452-3, *453*
Examination, 355-400
Excel, log formula reference, *11*
Excursion, current, 35
Exponential averaging, 371

False sonic perspective, 176-7, 239, 323
FAR *see* Forward aspect ratio
Fast Fourier transform (FFT), 363-4
Feedback, 186-7, 199, 478, *478*
FFT analyzers, 363-97
comparison to RTA, 360-1
dual-channel, 374-97
calibration, 424-80
calibration test reference, *430*
ongoing optimization, 474-80
post-calibration, 421-3, 422
test reference, *403*
verification, 403, *405*
verification procedures, 403-21
window functions, 369
feedback, 478
microphones in motion spatial averaging, 433
single-channel, 372-4, *373-4*
THD + n testing, 408-9
FFT lines, 363
FIFO *see* First in, first out

Fill systems:
balconies, 348
design procedure, 242-3
proximity ratio, 246-7
room shape, 202
see also Downfill; Frontfill; Infill; Sidefill
Filters:
crossover asymmetry, 98, *98*
crossover audibility, 95
crossover classes, 92
crossover range, 94
equalizers, 22-6
frequency dividers, 26-7
order, 94
phase response, 17-18
summation zones, 75-7
topology, 94-5
First in, first out (FIFO) averaging, 371, 372
Fixed points per octave (PPO) (constant Q) transform, 364, 367-8, *369,* 390-1
Flanging, 221
Flat phase response, 17-18
Floor:
drawings, 207
measurement microphones placement, 443-5, *444*
specifications for performance space, 204
Forward aspect ratio (FAR), 244-6
coupled line source, 255-7
coupled point source, 258, 332-3, *332*
minimum level variance, 247, 247-9, 248-9
reference chart, 245
uncoupled line source, 260-3, 263-4
Forward extension, 236, 236
Fourier transform, 363
Free field, sound pressure level, 45-6
Freedom from echo, 192,194,*196*
Freedom from noise, 192,194,*196*
Frequency:
acoustic transmission loss, 46-7
amplitude response, 17
banded SPL measurements, 15
beamwidth, 56-7, 57
coverage pattern, 55
definition, 4
echo perception, 165
impedance, 16-17
latency, 19
source matching, 61-2
spectral dividers, 26-7
terminology chart, *8*
time, 4-5, 5
waveform, 8-10, *9*
wavelength, 5-8, 6
wavelength displacement, 84, *85-6*
Frequency dividers *see* Spectral dividers
Frequency effects:
coupled line source, 115-20,*116-18*
coupled point source, *121*
uncoupled line source, 127, 227
uncoupled point destination, 234
Frequency range:
loudspeakers, 52-4, 54
program material, 317
Frequency resolution:
acoustic modeling programs, 213, 214-15, 225-27,219-20, 226
FFT analyzers, 364-5, *364,* 366-7, *367*
Frequency response, 17-19
acoustic modeling programs, 221, *222-5*
coloration, 176
dedicated speaker controllers, 32
equal loundess contours, 152-3
equalizers, 23-4
impulse response, 391-3
loudspeaker specifications, 48
measurement microphone placement, 441
minimum spectral variance, 265-303
ongoing optimization, 479
phase delay, 70
ripple variance, 240-2, 240
speaker order, 105-8, 105, 206-8
spectral variance, 235, 237-40
transfer function measurement, 376-86
verification procedure, 415-17, 426
Front to back ratio, 56
Front/back localization, 157
Frontfill systems, 193-4, 328, 329
delay setting, 466-7
measurement microphone position strategies, 451
speaker position adjustment, 458
specifications for performance space, 204
Full-range speakers, 53, 54, 96-8
Full-scale digital, 43
Fundamental, 408

Gain:
balanced lines, 33
decibels, 10
level-setting devices, 22
limiters, 28, 29
line level devices, 22
line loss, 32
pseudo-unity gain, 410-11
unitless decibels, 15-16
verification procedure, 410-11
Gap area, 127, 227
Gap class crossovers, 88-9, 103, 113, 125, 128-30
Giddings, Phil, 32
"Goldilocks" approach, 330
Graphic equalizers, 23
Grateful Dead, 351
Gray, John, 183
Ground loops, 33, 34
Ground plane mic placement, 444-5, *444*
Ground stacked speakers, speaker/room summation, 143, 144

Haas effect, 464-5
Haas, Helmut, 161-2
Half-space loading, 52, 143, 144
Harmonic distortion, 176
see also Total harmonic distortion
Head-related transfer function (HRTF), 155-7, 256-7, 162-4, 263
Headroom, 14, 40
Hearing damage, 14
Heinlein, Robert, 319
Helmholtz resonators, 225
Heyser, Richard, 397
High-pass filters, 31, 32, 32
Hilbert transform, 395

Hiss, 407, *407*
Horizontal localization, *155,*157-64, *158-63*
HRTF *see* Head-related transfer function
Hum:
coherence, 390
FFT analyzers, 366, 372-3
line level devices, 21
verification procedure, 407, *407*
window functions, 369
Human perception, 150-1
amplified sound detection, 176-8
artificial reverberation, 199
dB SPL continuous, 15
dynamic range, 14, *15*
echo, 167-8
evaluation of concert halls, 191-2, *195-6*
frequency, 17
"line of sound", 202
linear operating range, 14
localization, 153-64
loudness, 151-3
pitch sensing, 62-3
spatial, 168-9
tonal, 164-9
Humidity:
acoustic transmission, 46-7, *48*
measurement, 358
ongoing optimization, 475, 476
Hybrid line/point source arrays ("J" arrays), 233, 234, 256-7, 291-4, 292-5
Hybrid transmission, 198-9
Hygrometer, 358, *403, 430*

ILD *see* Inter-aural level difference
Immediacy of response, 192, 194, *196*
Impedance, 16-17
amplifiers, 36, *37*
analog transmission, 20
digital audio transmission, 44
interconnection, 32
line level devices, 21
speaker cables, 40-2
Impedance economics, 36, 40-2
Impedance testers, 359-60, *403, 430*
Impulse response:
analysis systems, 400
delay setting, 465-6
maximum length sequence sound analyzer, 398
transfer function measurement, 391-5, *396*
"In phase", 67
"In the round", 325-7
Inclinometers, 356, *356, 403,* 430
Inductance, 32
Inductive reactance, 16
Infill subsystems, 327, 328
Intelligibility, 235
Inter-aural level difference (ILD), 158, 170
Inter-aural time difference (ITD), 158, 171
Interconnection, 32-5, 40—4
Intermodulation distortion, 176
Interpolation, 214
Interrupts, 431
Intimacy, 192, 194, *195*, 198
Inverse square law, 45-6, *46*
calibration, 438
loudspeaker sensitivity, 50
maximum SPL, 52
standard level variance progressions, 236, 236
Inverted stack subwoofer arrays, 305-6, *305*
Isobaric contours, 56, *56*
Isolation zones, 71-3, 72, 75-7
crossover progressions, 88, *88*
filter order, 94
spatial acoustic crossovers, 102-3
triangulation, 81, *82*
uncoupled point destination, 136-7
unity crossover, 92
ITD *see* Inter-aural time difference

"J" arrays *see* Hybrid line/point source arrays

Laser pointers, 357-8, *358,403,*430, *430*
Late arrivals:
coherence, 390-1
see also Echoes
Latency, 19
delay lines, 22
digital audio transmission, 43, 44
line level devices, 21
verification procedure, 412-14, *413*
Lateral extension, 236, 236, 237
Layering, minimum level variance, 260
Level:
active speakers, 31
amplifiers, 40
analog transmission, 20
decibels, 10-17
equal loudness contours, 152-3
human perception of loudness, 151-2
multiple input summation, 66
program material, 317
relative phase cycle, 66-7, 67, *68*
ripple variance, 70-86
subsystems, 324-5
summation, 65-6, 65, 66
voltage level reference chart, *12*
see also Relative level
Level effects:
asymmetric coupled line source, 280-1, *281*
asymmetric coupled point source, 281, 284-90, *285, 289-90*
hybrid coupled line/point source array, 292
uncoupled line source, 296-9, *296, 297-300*
Level loss:
coverage patterns, 55-7
polar patterns, 55-6, *56*
Level offset:
asymmetric uncoupled line source, 261-3, 262
crossover asymmetry, 108-11, *109-10*
localization, 154, *154, 155,* 158-64, *158-63*
stereo perception, 170-5, *172-5*
triangulation, 80-2, *80*
vertical localization, 157
Level panning, 160-4

Level setting:
  calibration procedure, 453-5, 454
  crossover asymmetry, 98-101, 99
  dedicated speaker controllers, 30
  digital audio transmission, 43
  measurement microphone placement, 441
  order of calibration operations, 467
Level tapering:
  composite coupled point source, 339
  coupled line source arrays, 280-1, 282
  coupled point source, 284-90, 285-6, *288-9,* 333
  hybrid coupled line/point source array, 294, *295*
  incremental, 339
  power vs. variance, 310
  subwoofer arrays, 307
Level variance, 235, 247-65
  asymmetric coupled point source arrays, 285
  measurement microphone placement, 441
  speaker position adjustment, 455
  standard progressions, 236-7, 236
Level-setting devices, 22, 30
Limiters, 28-30
  active speakers, 31-2
  amplified sound detection, 176
  amplifier voltage gain, 38, 40
  dB SPL continuous, 52
  dedicated speaker controllers, 30
  human perception of loudness, 152
  verification, 419
Line level, 3, 3, 12, 20
Line level devices, 20-35, 20, 22, 404
Line loss, 32, 40-2, 41
Line source arrays:
  coupled, 114, 115-20
    crossover classes, 115, 116
    line length, 275-6
    measurement microphone position, 447, *448*
    minimum level variance, 253-7, *254-5*
    minimum spectral variance, 239, 268, 269, 275-81
    pyramid effect, 119-20, 229-22, 275-9, 280, 281
    subwoofers, 306-8, 307, 309
    summation zones, 115-20, 226
  history, 233-4
  uncoupled, 127-30, 227-32
    design procedure, 341-2, 342
    measurement microphone position, 449-50, 450
    minimum level variance, 256-7, 256, 260-3, 260, *262-4*
    minimum spectral variance, 239, 294-303, *296-300*
    minimum variance menu, 312-13, 322
    multiple main systems, 327
    surround systems, 322, 323-4
  see also Hybrid line/point source arrays
Line testing, 359
Linear frequency axis, 365-6, 365
Linear operating range, 12, 14
Linear voltage gain, 38
Listen box, 359, 359, 430
Listening, 473-4
Liveness, 192, 194, *195*
Localization:
  human perception, 153-64
  stereo, 169-75
Log formulas, decibels, 11, 22
Log frequency axis, 365, 365
Log voltage gain, 38
Longitudinal section, 207, *208, 209*
Longitudinal view, 209-12, 209-22
Loss:
  environmental effects, 46-7, *47, 48*
  sound pressure level, 45-6, 46
  speaker cables, 32, 40-2, 42
  standard level variance progressions, 236, 236
Lottery model, 427
Loudness, 151-3
  evaluation of concert halls, 192, 194, 195
  specification goals, 315
Loudspeakers:
  acoustic modeling, 213-20, 229-30
  acoustic power, 45
  acoustic transmission, 47-59
  active speakers, 31-2
  array types, 318
  cables, 40-2
  calibration, 429-30, 440, 441-2, 453, 455-8
  coverage patterns, 55-7
  current, 35
  dedicated speaker controllers, 30-1
  directivity, 54-9
  drawings, 209-12, 209-22
  equalization, 458-63
  evaluation of musical-acoustic quality, 193
  hidden, 204
  interaction, 55
  interconnection, 40-2
  limiters, 28-9, 29, *30*
  locations, 318
  maximum input/output capability, 51-2, 53, 409-10
  measurement microphone placement, 441-52, *445-8, 450-2*
  models, 318
  muting control, 430, 431
  order:
    acoustic modeling programs, 214-15, 225-26, 218
    beamwidth, 309-10
    classification, 105-8, 205-8
    common loudspeaker format, 219
    coupled line source, 115-17, *116-18,* 268, 269, 275-6, 276
    coupled point source, 222, 122-5, 270-4, 281, 283, *284,* 287-90
    crossover asymmetry, 108, 220
    crossover audibility, 111
    crossover classes, 115
    minimum level variance, 253-4, *254*
    minimum spectral variance, 265-74
    overlapping for power, 331
    spatial acoustic crossovers, 103-8, 105-8
    spectral variance, 237
    uncoupled line source, 127-30, 227, 296-9, 298

uncoupled point destination, *134*, 135-8, *137*
uncoupled point source, 132-3, *132-3*
phase delay, 382, 384, *385*
phase response, 18-19, *18*
placement in respect to microphones, 316-17
poor positions, 315
pressurization and rarefaction, 4
quantities, 318-19
quantity
coupled line source, 268, 269, 275-80, 277-80, *277-8*
coupled point source, 270-4, 270-4, 281
room shape, 242-7
sensitivity, 50-1, *50*
spacing
crossover assymetry, 108, *111*
crossover audibility, 112
point destination, 301, 302
uncoupled line source, 130, *131*, 296-9, *296-300*
uncoupled point source, *300*
speaker/room summation, 140-9
specification, 47-8, 204, 316, 318-19
standard level variance progressions, 236-7, 236
subsystems, 324
triangulation, 79-84
variance, 242-7
verification, *404, 415,* 417, 418
Low-pass filters, 31, *31,32,383,*384

Magnitude *see* Amplitude
Main systems, 320, 325-7, *326*
balconies, 346-8, *347*
design, 331-43
multiple, 325-7, 326
scaling for power, 330
speaker position adjustment, 458
subsystem combinations, 327-9
MAPP Online, 214, 220, 230-1
Mathematical trace averaging, 434-5, *436*
Maximum acceptable level variance, 249-52, 251
Maximum acceptable variance, 243
Maximum input/output (I/O) capability, verification procedure, 409-12, *411*
Maximum length sequence sound analyzer (MLSSA), 398
Measurement:
acoustic modeling programs, 221, 222-5, *228*
averaging, 370-2
monitoring, 44
prediction, 206
tools, 355-97
transfer function, 374-97, *375*
Mechanical overload, 28-30
Media:
acoustic transmission loss, 45-7
definition, 4
digital audio transmission, 42
transmission, 4
waveform, 8
wavelength, 5-6
Microbars, 14
Micropascals, 14
Microphone calibrators, 402, *403*
Microphone level, 12, 20
Microphone multiplexing, 433-4, *435*
Microphones:
amplified transmission, 185
cardioid, 157, 179-80
duplicate entry summation, 185-90, *190,* 477-8
front/back localization, 157
match verification, 419-20, *419*
measurement, 432
calibration access points, 432
calibration strategies, *468*
context, 438-41
listening position height, 443-5, *444*
ongoing optimization, 479-80
orientation angle, 443
placement, 441-52
reception, 179-80
response verification, 420-1, 420
self-verification, 402
spatial averaging, 433-4, *435*
strategies, 445-52, *445-8, 450-2*
verification, *405-6,* 419-21, *419*
omnidirectional, 157, 179-80
re-entry summation (feedback), 186-90, *190*
reception, 178-80
sensitivity, 49-50
speaker placement, 316-17
Microphones in motion, 433, *435*
Mid-bass systems, 53-4, *54*
Minimum phase, 24
Minimum variance:
calibration, 425, 426, 428
level, 242-65, 315
menu, 310-13, *311-12*
ripple, 303-9, 315
specification, 320
spectral, 265-303, 315
Mix consoles:
evaluation, 194
latency, 19
limiters, 28
multichannel sound, 351-2
verification access, *404*
Mix position:
anarchy model, 426
monarchy model, 427
ongoing optimization, 479-80
specifications for performance space, 204
MLSSA *see* Maximum length sequence sound analyzer
Modeling programs, 212-28
Monarchy model, 427
monitors, re-entry summation, 190, *190*
Monoaural systems, system types, 321, 321, 325
Multichannel sound, 350-2, 478-9, 480
Multiplexing, microphones, *433-4, 435*
Multiway arrays, *91*
crossover overlap, 108-11, 112
dual perspective, 178
Music:
acousticians and sound engineers, 201

Music: *(contd)*
crest factor, 151
dB SPL operating levels, *15*
tonal, spatial and echo perception, 164
Musical theater, 176, 198-9, 202

Natural sound, 183-92, *195-6*, 198-9
Negative feedback limiters, 29-30, *30*
Networks, 44
Noise:
ambient noise floor, 14
amplifier level controls, 40
amplitude thresholding, 372
audience presence, 476
balanced lines, 32-3
calculated semi-anechoic response analyzers, 397
coherence, *388*, 389-90
digital audio transmission, 42
evaluation of musical-acoustic quality, 192, 194, *196*
field conditions, *399*
human perception, 151
line level devices, 21
linear operating area, 12
loudspeaker specifications, 47-8
pseudo-unity gain, 33
signal averaging, 370
single-channel analyzers, 372-3
THD + n testing, 408, *409*
transfer function measurement, 375, 376, *388*, 389-90
verification procedure, 406-8
window functions, 369
Noise floor, 14
dB SPL operating levels, *15*
FFT analyzers, 366
human perception, 151
pseudo-unity gain, 33
single-channel analyzers, 372-3
verification procedure, 407, *407*
Noise over frequency, 406-8, *407*
Non-causal noise, coherence, 388, 389-91, *392*
Notch filters, 24
Nyquist frequency, 363

OFFAX microphone position, 442, 446
coupled arrays, *448*, 449
single speakers, 446-7, *446-7*
speaker position adjustment, 455-7
SYM positions, 443
uncoupled arrays, 449-51, *450*
Olson, Harry, 305
ONAX microphone position, 442
acoustic evaluation, 453
coupled arrays, 447-9, *448*
delay setting, 466
equalization, 460-3, *461-3*
level setting, 454
microphone position strategies, 446
ongoing optimization, 479
single speakers, 446-7, *446-7*
speaker position adjustment, 455-7
SYM positions, 443
uncoupled arrays, 449-51, *450*
Ongoing optimization, 474-80
Optical trace averaging, 435-7, 436
Origami angle finder, 356-7, 357, *403*, *430*
Orthogonal views, 208, *208*, *209-11*
Oscilloscope, 360, 408, *430*
"Out of phase", 67
Over-balcony delays, 325, 326, 346-50, 347, 349
Overlap class crossovers, 88-9, *89*
coupled line source, 115
crossover progressions, *88*
multiway arrays, 92, 108-11, *112*
power, 330-1
spatial crossovers, 103, *103*, *104*
spectral crossovers, 95-8
uncoupled arrays, 125, 127, *127*, 128
Overload, 28-30, 47-8, 152, 176

Pain threshold, 14, *15*
Pan pot, 170, *170*
Panning, 160-4, 170-1, *170*, *172-5*
Panoramic effect, 160, 169-75
Panoramic field, 170-1, *170*, *172-5*
Parade routes, 326, 327
Parallel pyramid:
coupled line source, 119-20, *119-21*, 275-9, *280*, 281
coupled point source, 273
hybrid coupled line/point source array, 293
subwoofer coupled line source arrays, 306
uncoupled line source, 128-30
Parallel surface, *143*, 144
Parametric filters, 22-6, 25, 382, *383*
Passive line level devices, 20
Patch bays, 431
Peak limiters, 28, 30, 52, 152
Peak voltage (V_{pk}), 13, *13*
Peak-to-peak voltage (Vp-p), 13, *13*
Peak-to-RMS ratio, 13-14
Percentage of articulation loss of consonants (%ALCONS), 151
Percentage bandwidth, fast Fourier transform, 363
Perception *see* Human perception
Period *see* Time period
Perspective *see* Sonic perspective
Phase:
all-pass filters, 27
dedicated speaker controllers, 30
definition, 4
delay lines, 22
Fourier transform, 363
polarity tester, 359
real-time analyzers, 361
spectral dividers, 27
summation, 64-5, 65, 66-70
terminology chart, *8*
transfer function measurement, 377-86
unity crossover, 92
waveform, 8-10, *9*
wraparound, 378-80, *379*
see also Relative phase
Phase cycle, 66-7, *67-8*, 87-8
Phase delay:
crossover classes, 92-3
impulse response, 395
relative phase cycle, 69-70

transfer function measurement, 380-6, 382
Phase popper *see* Polarity tester
Phase response, 17-19, *18*
acoustic modeling programs, 213, 218, 227, *228,* 230
acoustical crossovers, 87
active speakers, 31
complementary phase equalization, 24
crossover level asymmetry, 100
dedicated speaker controllers, *31*
line level devices, 21
single-channel analyzers, 373
verification procedure, *416,* 417-18, *418*
Phase shift, 43, 78-9
Phase slope, 69, 378, 380-6, *381, 383*
Phase-aligned crossover, 102
Phons, 152
Pink noise, 151, 239, 375
Pink shift, 238-9, 239-40, 246-7, 278-9, 323
Pistonphone, 49-50
Pitch, 5, 62-3, 150
Plan views, 207, *208-11,* 209-12
Point destination arrays:
coupled, 113, *114,* 115, 479
infill subsystems, 327
uncoupled, 133-8
design procedure, 342-3
measurement microphone placement, *451*
minimum level variance, 263-5, 264
minimum spectral variance, 301-3, 302-3
summation zones, *134-9*
Point source arrays:
coupled, *114,* 120-5
asymmetric, 258-60, *259*
asymmetric-composite, 337-41, *339-40*
design procedure, 331-2, 332, 333-7, *334*
diagonal plane, 344-5
history, 233
line length, 288, 290-1, 291
main systems, 331-2
measurement microphone placement, 447-9, *448*
minimum level variance, 257-60, 257, 258, 259
minimum spectral variance, 237, 239, 270-4, *270-4,* 281-91, *291*
minimum variance menu, 310, 311-12, *311*
power vs. variance, 310
sidefill subsystems, 327
subwoofers, 308-9, *308*
summation zones, 126
symmetric, 258, *259*
history, 233, 234
triangulation, 81
uncoupled, 130-3
design procedure, 342, *343*
history, 233
measurement microphone placement, 450
minimum level variance, 263, 264
minimum spectral variance, 299-301, *300*
minimum variance menu, 310, 322, 313
multiple main systems, 327
speaker position adjustment, 457
see also Hybrid line/point source arrays
Points per octave (PPO), 367-8, *369*
Polar patterns, 55-6, *56,* 58, *58*
Polarity, 19
active speakers, 31
balanced lines, 33
electrical summation, 64
impulse response, 395
line level devices, 21
power amplifiers, 37, 38
relative phase, 378
unbalanced lines, 33, 35
verification procedure, *414,* 415, *415*
Polarity testers, 359, *403, 430*
Post-calibration verification, 421-3, *421-2*
Post-verification, 402
Power, 16-17
acoustic transmission, 44-5
amplifiers, 36
goals of calibration, 425
log formulas, 11, 21
program material, 317
scaling, 330
specification goals, 315
vs. variance, 232-4, 310
PPO *see* Points per octave
Pre-verification, 359, 402
Precedence effect, 161-2, 263, 464
Prediction, 206-31
Predictive limiters, 29, 29
Pressurization, 4
Program material:
channels, 317
loudness, 151
ongoing optimization, 474-5
reference material, 474
speaker array types, 318
specification, 317
Propagation plane, 209-12, *209-11*
Protractors, 55, *56,* 57-9, *58,* 212, 356, *357*
calibration of speaker positions, 430
calibration test reference, *430*
coverage patterns, 242-3, *243,* 244
minimum level variance, 245
test reference, *403*
Proximity effect, 180
Proximity ratio, 246-7, 318
balconies, 346
composite coupled point source, 339
surround systems, 322-3, 324
Pseudo-unity gain, 33, 410-11
Pyramid effect *see* Parallel pyramid
Pythagorean theorem, 80, *80,* 171

Q, 24-6, *24,* 56-7, 58
Quasi-phase response, 227

Racetracks, 326, 327
Radial coverage, 257-60, 257
Radial extension, 236-7, 236
Radial method, 55, *56,* 57-9, *58*

Rake, 212
Range finder, 358
Rarefaction, 4
Ray-tracing models, 221, 222-5, 225
Re-entry summation, 186-90, *190*, 199, 475, 477-8
Reactance, 16-17
Real-time analyzer (RTA), 360-2, *403*, *430*, 433, 434
Reception, 150-80
  microphones, 178-80
  natural vs. amplified sound, 185-6
  specification goals, 315
Reference materials, 474
Reflections:
  acoustic modeling, 222, 223, 224, 229-30
  amplified sound transmission, *187*, 189-90, *190*
  artificial reverberation, 199-201
  audience presence, 475-6
  calibration procedure, 452-3, *453*
  coherence, 390-1, *390*
  diffusion, 223, 224
  empty seats, 443, 444, *444*
  floor, 443-5, 444
  impulse response, 393-4, 395
  localization, 154-7, *156*, 158
  natural sound transmission path, *188*, 189, *189*
  natural vs. amplified sound, 198
  sonic perspective, 177
  speaker directivity, 54
  speaker/room summation, 140-9
  specifications for performance space, 203-5
  spectral tilt, 238
  SPL, 45-6
Refraction, 221, 223
Reinforcement, 198-9
Relative amplitude, 64-6, 376-7
Relative arrival time, 394-5
Relative level:
  acoustic crossover, 87
  crossover asymmetry, 100
  horizontal localization, 158-64, *158-63*
  impulse response, 395
  speaker/room summation, 141
  summation zones, *68*, 71-8
Relative phase:
  acoustic crossover, 87
  phase delay, 380-6, *381*
  polarity, 378
  summation, 64-5, 65, 66-70
  summation zones, 71-8
  transfer function measurement, 377-86
  wavelength displacement, 84
  wraparound, 378-80, 379
Relative phase cycle, 66-7, 67, *68*
Relative time, horizontal localization, 158-64, *158-63*
Resistance, 16-17, 32
Resolution, 213-14, 230, 363
Resonance, 225
Response ripple, 70-86, *70*
Return ratio (RR), 346-8, 347
Reverberation:
  artificial, 199-201
  crossover audibility, 95
  evaluation of musical-acoustic quality, 193
  impulse response, 394
  loudness of, 192, 194, *195*
  natural vs. amplified sound, 198
  specification goals, 315
  transmission path, 189-90
  variable acoustics, 198
Ripple variance, 70-86, *70*, 235
  acoustic modeling programs, 218
  beamwidth, 265, 267, 273
  coupled line source, 276
  coupled point source, 283, 286-7, 288, 290-1, 333
  equalization, 459-60
  FFT analyzers, 366
  line source array, 279
  measurement microphone placement, 441, 442
  mic/mic summation, 477
  minimum, 303-9, 315
  point destination, 301-2
  precedence effect, 464-5
  spatial reference, *241*
  speaker position adjustment, 455
  standard progression, 240-2, 240
  uncoupled line source, 296, 297, 299
RMS *see* Root-mean-square
Room:
  acoustic modeling, 229-30
  acoustic properties, 191-205, 317-18
  aspect ratio, 244-6
  calibration, 438, 443-4, *444*, 452-3, 453
  diffraction, 224
  diffusion, 223
  drawings, 207-12
  impulse response, 391-5
  minimum level variance, 248-65
  natural vs. amplified sound, 183-6, 187, *188-90*, 189-90
  ongoing optimization, 475
  resonance, 225
  shape, 202
  speaker/room summation, 140-9
  specification, 202-5, 316, 317-18
  specification limitations, 316
  variance, 242-7
Root-mean-square (RMS) averaging, 370
Root-mean-square (RMS) limiters, 13-14, 28, 30, 52
Root-mean-square (RMS) voltage, 13, *13*
RTA *see* Real-time analyzer

Sampling frequency:
  AES/EBU standards, 43-4
  fast Fourier transform, 363
Scalable design, 330-1
Scale rulers, 212
Seating boxes, 203
Seats, 443, 444, 444
Self-powered speakers, 51
Self-verification, 402
Sensitivity, 38, *39*, 49-51, 50
Shelving filters, 22-3, 25, 26
Side seating, 202-3
Sidefill subsystems, 327, 327, 328
Signal averaging, 370-2
SIM™ (Source Independent Measurement), 362, 400
Sine waves, 4-5, 5, *9*, 10, 13-14, *13*

Single speakers:
calibration strategies reference, *468*
measurement microphone placement, 446-7, *446-7*
minimum variance, 248-52, 248-52, 265-68, 310
speaker position adjustment, 456, *456*
subsystem types, 325
Single-channel analyzers, 372-4, *373-4*
SMAART (Sound Measurement Acoustical Analysis Real-Time Tool), 400
Soffit mounting, 143
Sonic image:
calibration, 425
channels, 316
delay setting, 463, 464-5, 466
distortion, 154, 154, 178
localization, 153-64
precedence effect, 464-5
specifications for performance space, 202, 204
Sonic perspective, 176-8, 239, 323, 459
Sound level meters, 360, *403, 430*
Sound pressure level (SPL), 14-15
acoustic transmission, 44-6, 46
coverage patterns, 55-7, 56
crossover asymmetry, 98-100, *98*
human perception of loudness, 151-2
level tapering, 339
maximum, 51-2, 53
program material, 317
scaling, 330
sensitivity, 49-51
sound level meters, 360
subwoofer arrays, 308-9
vs. minimum variance, 242, *243*
weighting, 52
see also dB SPL
Sound pressure level (SPL) meters, 51
Source:
calibration subdivision, 428-9, *429*
test procedure, 402
transfer function measurement, 375, 375
Source matching, 61-2, 62
Source/room summation, 186-90
Source/speaker summation, 186-90
"Spaciousness", 158
Spatial averaging, 432-8, 441
Spatial crossovers, 101-12
arrays, 112-39
calibration, 438
classes, 103, *103*
comparison with spectral crossovers, 101-3
definition, 87
delay setting, 463
FFT analyzers, 366
level setting, 454
measurement microphone placement, 442, 442-3
ripple variance, 240-2
speaker/room summation, 140-9, 141
Spatial dividers, 101-12
acoustical crossover progressions, 88, *88*
comparison with spectral dividers, 101-3
definition, 87
speaker/room summation, 140, 141
Spatial enhancement, 200
Spatial perception, 164-9
Spatial separation, 315
Speaker level, 3, 3, 22, 20, 40-2
Speaker level devices, 20, 35-42
Speaker load, 36, 37
Speaker/room summation, 186-90
absorption effects, 147-9, 248, 249
acoustic modeling, 229
calibration, 430
coupled point source, 286
pink shift, 240
Speaker/speaker summation, 186-90, 222, 229, 240
Speakers *see* Loudspeakers
Specification, 314-52
Spectral crossovers, 89-101
alignment example, 97
comparison with spatial crossovers, 101-3
component acoustic responses, 97
definition, 87
delay setting, 466, 466
level setting, 454-5, *455*
summation acoustic responses, 97
summation zones, 93, 101
verification procedure, 416, *416-17*
Spectral dividers (frequency dividers), 26-7, 89-101
acoustical crossover progressions, 88, 88
active speakers, 31, 31-2
amplifier voltage gain, 38
comparison with spatial dividers, 101-3
crossover classes, 89
curve family, 27
dedicated speaker controllers, 30
definition, 87
limiters, 28, *29, 30*
line level devices, 20
manufacturer's recommendations, 100-1
phase delay, 384
verification access, 404
Spectral tilt, 237-9, 459
Spectral variance, 235, 265-303
equalization, 459-60
minimum, 265-303, 315
proximity ratio, 246-7
speaker position adjustment, 455
standard progressions, 237—40, 238
Speech, 15, 151
Speed of sound, 6-8, 7
SPL *see* Sound pressure level
Splay angles:
calibration, 430
coupled outward angles surface, 144
coupled point source, 120-5, *121-6*, 281-4, 287, 288-90, 308-9, 332-7
hybrid coupled line/point source array, *292-5*
measurement, 356-8
minimum level variance, 257-60, 257, 259, 264
minimum spectral variance, *265-7*, 270-4, 270-4
speaker/room summation, 141

Splay angles: *(contd)*
subwoofers, 308-9
uncoupled outward angled surface, 146
uncoupled point source, 130-3, 132
Sports stadia, 326, 327
Stage, specifications, 204-5
Stage leakage, 186-90, 463-4, 475, 476-8
Star pattern, 434, 435, 442
Stereo:
calibration, 441
channels, 316
localization, 160
main systems, 325
ongoing optimization, 478-9
perception, 169-75
program material, 317
room shape, 202
scaling, 171-5
side-effects, 175
summation, 62
system types, 321-2, 322
Subsystems, 324-30
amplifiers, 36
calibration, 424
definition, 320-1
equalization, 460-3
level setting, 453-5
measurement microphone placement, 451-2, *452*
monoaural systems, 321
ongoing optimization, 479
order of calibration operations, 467-73, *468-73*
stereo systems, 321-2
types, 325-30
Subwoofers:
acoustical crossover progressions, 88, *88*
coupled line source, 306-8, *307*
coupled point source, 308-9, *308*
crossover audibility, 95
dual perspective, 178
end-fire arrays, 304-5, 304, *309*
frequency range, 53, 54
ongoing optimization, 480
overlap class spectral crossover, 96-8
phase delay, 384
program material, 317
ripple variance, 303-9
speaker/room summation, 144
two-element in-line technique, 305, 305
Summation, 60-149
acoustic modeling programs, 215, 217, 218-21, 219, 222, 226-9
acoustical crossovers, 87-112
arrays, 112-39
artificial reverberation, 199-201
calibration, 424
characterization, 226-8
definition, 61
delay lines, 22
duplicate entry, 186-7, 189, 190, 290, 477-8
duration, 62-3, 63
environmental effects, 138, 240
equalization, 23
equalization filters, 24
evaluation, 195-6
flow block, 62
frequency dividers, 26-7
frequency range, 54
geometry, 79-86
horizontal localization, 158
icons, 73, 77
measurement microphone placement, 441
microphones in motion spatial averaging, 433
multiple input level reference, 66
ongoing optimization, 475
phase slope, 384-6, *386*
quantity, 63, 63
re-entry (feedback), 186-7, 189, 190, 290
response ripple, 70-86
ripple variance, 235, 240-2
sonic perspective, 177
source direction, 64, 64
source matching, 61-2, 62
spatial perception, 164
speaker interaction, 55
speaker/room, 140-9
specification goals, 315
spectral tilt, 237-9, *238*
stereo side-effects, 175
summed microphones spatial averaging, 433
tonal perception, 165, 166-7
tonal, spatial and echo perception, 167-8
transfer function measurement, 384-6, *386*
transmission path, 186-90, 187-90
variation, 232-4
wavelength, 5
zones, 71-8, 72
Summation zones, 71-8, 72-7
acoustic modeling programs, 221, 222-5
acoustical crossover progressions, 87-8, *88*
arrays, 113
coupled line source, 115-20, 226
coupled point source, 122-5, 222-6
crossover asymmetry, 108-11, *109-10*
crossover classes, 89-93, 89-92
crossover range, 94
drawing series guide, 224, 115
ripple variance progression, 240-2
spatial acoustic crossovers, 102-3
speaker/room summation, 141-2, 143-7, 244-7
spectral acoustic crossover, 93, 202
triangulation, 79-80, 81, 82, 83
uncoupled line source, 127-30, 228-32
uncoupled point destination, 134-9
uncoupled point source, 132-3, 232-3
uniformity of experience, 194
Summed microphones, 433, 435
Surface area, 45-6
Surface detail omissions, 226
Surfaces, 140-9, 319
Surround systems, 320, 322-4, 323
SYM mic position, 443
Symbols, terminology chart, 8
Symmetric unity splay equation, 333
Synchronization, 178
Systems:
calibration access points, 431-2, 432
definition, 320-1
equalization, 23
subdivision, 324-30, 428-30, 429

types, 320-5
see also Main systems; Subsystems

TANSTAAFL ("There ain't no such thing as a free lunch."), 319-24, 428
Tape measure, 430
TDS *see* Time-delay spectrometry
Technical ground, 35, 35
Temperature:
acoustic modeling programs, 220-1
acoustic transmission, 46-7, *48*
measurement, 358
ongoing optimization, 475, 476
speaker/room summation, 147-9, 249
speaker/speaker summation, 138, 240
wavelength, 6-8, 7
Terminology reference, *8*
Testing, structure, 401-19
Texture, 192,194, 296
THD *see* Total harmonic distortion
Thermal overload, 28-30
Thermometers, 358, *403, 430*
Thresholding, 372
Tilt, 236
see also Spectral tilt
Time bandwidth product, 363
Time offset:
comb filtering, 78-9, 79
localization, 154,155
relative phase cycle, 69-70, 69
stereo perception, 171-5, *172-5*
summation zones, 73, 73-7
tonal, spatial and echo perception, 165-9, 265
triangulation, 80, 80-2, *80*
vertical localization, 157
Time period:
definition, 4
frequency, 4-5, 5
temperature effects on speed of sound, 7
terminology chart, *8*
wavelength chart, 6
Time record, 362-3, 364-5, 364
Time-delay spectrometry (TDS), 397
Tonal balance, 203
Tonal envelope, 165-7,166
Tonal fusion zone, 165
Tonal perception, 164-9,175
Tonal quality, 192, 196,198
Tonal variance, 198
Total harmonic distortion (THD), 21, 48, 369, 372-3,408-9
Touring systems, 233-4, 423
Transducers, 32, 49, 52-4
Transduction, 4
Transfer function measurement, 374-97
analysis systems, 400
applications, 397
calibration access points, 430-2
calibration subdivision, 428-9
impulse response, 391-5, 396
verification test set-up, 403, *405*
Transformers, 33, 34
Transient peak, 13-14
Transmission, 3-59
acoustic, 44-9, 45
acoustic modeling programs, 220-5
analog electrical, 19-42, *19*
digital, 42-4
interconnection, 32-5
loudspeakers, 47-59
natural vs. amplified sound, 185-6, 186-90, 287
natural/amplified hybrids, 198-9
speaker level devices, 35-42, 36
specification goals, 315
Trapezoid, 233
Triangulation, 79-84
binaural localization, 157
diagonal plane, 346
drawings, 210-12, 222
ripple variance, 241-2
uncoupled point destination, 133,135, 137-8
Truncation, 398
Two-element in-line technique, 305, *305*

Unbalanced lines, 33-5, 35, *404*
Uncoupled speaker arrays *see* Arrays; Line source arrays; Point destination arrays; Point source arrays
Under-balcony delays, 202, 346-50, 347, 349
acoustic effects, 318
delay setting, 467
dual perspective, 177
evaluation, 193
extensions, 237
multiple main systems, 325, 326
scaling for power, 330
sonic image, 316
specifications for performance space, 204
stereo systems, 322
variation, 235
Uniform frequency response, 197
Uniformity, 192,194,*196,*198
Unitless decibels, 15-16
Units, *8*
Unity class crossovers, 88-9, *89,* 96-7, 103, 103-4,113
addition formula, 92
coupled point source, 120-2,122
crossover assymetry, 108, 110
speaker order, 105-8
uncoupled arrays, 125,127-32, 227, 133,135
Unity gain, 10, 410-11
Unity splay angle, 332,333-7, *334-7*
Unity splay equation, 333—4
Unweighted averaging, 371

Value engineering, 36
Variable acoustics, 198
Variable electro-acoustics, 199-201
Variation, 232-313
definition, 235
goals of calibration, 425, 426
minimum variance principles, 235-42
standard progressions of, 235-42
stereo systems, 322
vs. power, 232-4, 242, 243, 310
Vector averaging, 370
Venue *see* Room
Verification, 401-23
access points, 403, 404
calibration, 424-5

Verification *(contd)*
  microphones, 419-21, *419*
  oscilloscope, 360
  post-calibration, 421-3, 422-2
  procedures, 406
  test reference, *403*
  test set-up, 403-6
  testing stages, 402
Vertical localization, 154-7, *154, 156,* 162-4, *163*
Video screens, 178
Vocal processors, 419
Volt/ohm meters (VOMs), 358-9, *403,*
Voltage:
  AC signals, 13
  electronic decibels, 11-13
  electronic transmission, 4
  gain, 10
  log formulas, 11, *11*
  power, 16-17
  sensitivity, 49-51
  speaker level devices, 35
  voltage level reference chart, *12*
  waveforms, *10*
Voltage gain, amplifiers, 38-40
Voltage source, 21, 32-5
VOM *see* Volt/ohm meters
Vp-p *see* Peak-to-peak voltage
V_{pk} *see* Peak voltage

"Wall of sound", 233-4, 426
Walls, 203-4, 207, 319
Warmth, 192, 194, *195,* 203
Waveforms, 8-10, *8, 9, 10*
  crest factor, 13-14, *13*
  oscilloscope, 360
  transmission, 3
Wavelength:
  definition, 4
  displacement, 84, *85-6*
  handy references, 6, 7, *8*
  relationship to period and frequency, 5-8
  speaker directionality, 55
  speaker/room summation, 143
  temperature effects, 6-8, *7*
Weighted averaging, 370-1, *371*
Weighting, 52
White noise, 407
Window functions, 368-70
Wraparound, transfer function measurement, 378-80, *379,* 384-6

XOVR microphone position, 442-3
  coupled arrays, *448,* 449
  delay setting, 465, *466*
  level setting, 454
  single speakers, 446, *446-7*
  speaker position adjustment, 455-7
  SYM positions, 443
  uncoupled arrays, 449-51, *450*

AUDIO

- The first book devoted exclusively to the emerging field of sound system optimization using modern tools and practices
- Offers you expert guidance on the popular tools of the trade, including dual channel FFT analyzers, acoustic prediction programs, modern speaker arrays and digital signal processors
- Illustrated throughout with clear color images and diagrams

"Bob's deep understanding of, and dedication to, the scientific principles of audio is greatly evidenced in this book. The experience and knowledge imparted—both rigorous and accessible—are certain to benefit the industry as a whole for years to come." John Meyer, Meyer Sound

"Pro Audio needs more tools like this, well thought out, well researched and well written, a valuable tool!" Sam Berkow, SIA Acoustics

"Unlike the many other books that offer fine discourses on theory, this book also provides you with the intellectual tools you need to make the tough calls in the field, when time is tight and things haven't gone exactly according to plan. Setting priorities is the hardest part and Bob McCarthy lays it out—loud and clear." Dave Clark, Engineering Harmonics Inc.

"Writing a meaningful book on live sound loudspeaker system measurement and optimization is perhaps one of the most daunting tasks one could ever dream up. Bob McCarthy has accomplished this and his new book provides very detailed guidelines for those of us who are involved in the design and commissioning of live sound loudspeaker systems." Tom Young, Electroacoustic Design Services

Bob McCarthy has been a leader in sound system optimization with FFT analyzers since its beginnings in 1984. He has developed techniques applicable to the alignment of sound systems of all types, scales and levels of complexity. Bob has conducted training seminars all over the world on the subject of sound system optimization and design for more than 20 years.

ELSEVIER

Focal Press

Focal Press
An imprint of Elsevier
www.focalpress.com

ISBN 0-2405-2020-3

9 780240 520209